# EXHIBIT A

## CONFIDENTIAL
## OUTSIDE COUNSELS' EYES ONLY

# EXHIBIT A

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

**Infringement Claim Chart for U.S. Patent No. 11,091,794**

Products and systems ("Accused Products and/or Systems") made, used, sold, offered for sale, and/or imported by Pacific Biosciences of California, Inc. ("PacBio") practice all elements of claims 1–3, 7–8, 10–16, 18–19 ("Asserted Claims") of U.S. Patent No. 11,091,794 ("'794 Patent"). The Accused Products and/or Systems include at least those products and systems identified in Plaintiffs' Disclosures Pursuant to Paragraph 7(a) of the Scheduling Order (April 12, 2023), and any amendments thereto. PacBio directly infringes the Asserted Claims at least through its making, use, sale, offer for sale, and/or importation of the Accused Products and/or Systems. PacBio indirectly infringes the Asserted Claims through its making, use, sale, offer for sale, and/or importation of the Accused Products and/or Systems, as well as any acts taken to induce direct infringement by others. Plaintiffs reserve their rights to amend or supplement this infringement claim chart as their investigation and discovery proceed in this action.

References:

- Methylation Detection with HiFi Sequencing Video, available at https://www.pacb.com/videos/methylation-detection-with-pacbio-hifi-sequencing/ ("HiFi Sequencing Video").
- Measuring DNA Methylation with 5-Base HIFI Sequencing Application Note, available at https://www.pacb.com/wp-content/uploads/application-brief-measuring-dna-methylation-with-5-base-hifi-sequencing.pdf ("PacBio Measuring DNA Methylation Application Brief"), produced at PBIO_00000010–11.
- Epigenetics, available at https://www.pacb.com/products-and-services/applications/epigenetics/, attached as Exhibit F to the Complaint ("PacBio Epigenetics Webpage").
- PacBioFileFormats Documentation, Release 11.0.0 (Feb. 2, 2022), available at https://pacbiofileformats.readthedocs.io/_/downloads/en/11.0/pdf/, attached as Exhibit G to the Complaint ("PacBioFileFormats Documentation").
- HiFi Sequencing, available at https://www.pacb.com/technology/hifi-sequencing/, attached as Exhibit H to the Complaint ("PacBio HiFi Sequencing Webpage").
- PacBio SMRT® Link User Guide Sequel® II and IIe Systems (v11.0) (April 2022), available at https://www.pacb.com/wp-content/uploads/SMRT_Link_User_Guide_v11.0.pdf, attached as Exhibit I to the Complaint ("SMRT® Link User Guide, v11.0"), produced at PBIO_00000560–701.
- PacBio SMRT® Link User Guide Sequel® II and IIe Systems (v11.1) (November 2022), available at https://www.pacb.com/wp-content/uploads/SMRT_Link_User_Guide_v11.1.pdf, attached as Exhibit M to the Complaint ("SMRT® Link User Guide, v11.1"), produced at PBIO_00000702–854.

## EXHIBIT A

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

- PacBio SMRT® Link User Guide Revio™, Sequel® II and Sequel® IIe Systems (v12.0) (April 2023), available at https://www.pacb.com/wp-content/uploads/SMRT_Link_User_Guide_v12.0.pdf ("SMRT® Link User Guide, v12.0"), produced at PBIO_00000855–1003.
- PacBio SMRT® Tools Reference Guide (v11.0) (April 2022), available at https://www.pacb.com/wp-content/uploads/SMRT_Tools_Reference_Guide_v11.0.pdf, attached as Exhibit J to the Complaint ("SMRT® Tools Reference Guide, v11.0"), produced at PBIO_00001004–1120.
- PacBio SMRT® Tools Reference Guide (v11.1) (November 2022), available at https://www.pacb.com/wp-content/uploads/SMRT_Tools_Reference_Guide_v11.1.pdf, attached as Exhibit L to the Complaint ("SMRT® Tools Reference Guide, v11.1"), produced at PBIO_00001121–1254.
- PacBio SMRT® Tools Reference Guide (v12.0) (March 2023), available at https://www.pacb.com/wp-content/uploads/SMRT_Tools_Reference_Guide_v12.0.pdf ("SMRT® Tools Reference Guide, v12.0"), produced at PBIO_00001255–1389.
- Technical Overview: HiFi Library Preparation Using SMRTbell® Express Template Prep Kit 2.0 for De Novo Assembly and Variant Detection Applications (April 2021), available at https://www.pacb.com/wp-content/uploads/HiFi-Library-Preparation-Using-SMRTbell-Express-TPK-2.0-Customer-Training.pdf ("HiFi Library Preparation Customer Training").
- Preparing Whole Genome and Metagenome Libraries Using SMRTbell® Prep Kit 3.0 (March 2023), available at https://www.pacb.com/wp-content/uploads/Procedure-checklist-Preparing-whole-genome-and-metagenome-libraries-using-SMRTbell-prep-kit-3.0.pdf ("Preparing Whole Genome and Metagenome Libraries Using SMRTbell® Prep Kit 3.0").
- PacBio and Children's Mercy Kansas City Expand Collaboration Taking a Multi-Omics Approach to Characterize Rare Disease (April 21, 2022), available at https://www.pacb.com/press_releases/pacbio-and-childrens-mercy-kansas-city-expand-collaboration-taking-a-multi-omics-approach-to-characterize-rare-disease/, attached as Exhibit E to the Complaint ("PacBio and Children's Mercy Kansas City Press Release").
- Rhoads, A., & Au, K. F. (2015). PacBio Sequencing and Its Applications. GENOMICS, PROTEOMICS & BIOINFORMATICS, 13(5), 278–289. https://doi.org/10.1016/j.gpb.2015.08.002, available at https://www.sciencedirect.com/science/article/pii/S1672022915001345 ("Rhoads & Au").
- Sequencing 101: from DNA to discovery - the steps of SMRT® sequencing (July 7, 2020), available at https://www.pacb.com/blog/steps-of-smrt-sequencing/ ("Sequencing 101")

## EXHIBIT A

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

- Preparing DNA for PacBio HiFi Sequencing — Extraction and Quality Control (Technical note) (available at https://www.pacb.com/wp-content/uploads/Technical-Note-Preparing-DNA-for-PacBio-HiFi-Sequencing-Extraction-and-Quality-Control.pdf) ("Preparing DNA for PacBio HiFi Sequencing")
- HiFi WGS at Scale on the Sequel® IIe System (Brochure) (available at https://www.pacb.com/wp-content/uploads/HiFi-whole-genome-sequencing-at-scale-on-the-Sequel-IIe-system.pdf) ("HiFi Sequel® II System Brochure")
- Alternative size selection methods for SMRTbell® prep kit 3.0 (Technical note) (available at https://www.pacb.com/wp-content/uploads/Technical-note-Alternative-size-selection-methods-for-SMRTbell-prep-kit-3.0.pdf) ("Alternative size selection methods for SMRTbell® prep kit 3.0").
- "No-Amp Targeted SMRT® Sequencing Using a CRISPR-Cas9 Enrichment Method" (Poster Presentation) (available at https://www.pacb.com/wp-content/uploads/McLaughlin-ASHG-2018-No-Amp-Targeted-SMRT-Sequencing-Using-a-CRISPR-Cas9-Enrichment-Method.pdf) ("No-Amp Targeted SMRT® Sequencing Using a CRISPR-Cas9 Enrichment Method").
- Long, Processive enzymatic DNA synthesis using 100% dye-labeled terminal phosphate-linked nucleotides (July 7, 2019) (available at https://www.pacb.com/publications/long-processive-enzymatic-dna-synthesis-using-100-dye-labeled-terminal-phosphate-linked-nucleotides/) ("Long, Processive enzymatic DNA synthesis using 100% dye-labeled terminal phosphate-linked nucleotides").
- "Introduction to SMRT® Sequencing" ("Introduction to SMRT® Sequencing Video"), produced at PBIO_00000001.
- "PacBio Sequencing - How It Works" ("PacBio Sequencing - How It Works Video"), produced at PBIO_00000002.
- "Customer Overview of System and Software Release v11.0" ("Customer Overview of System and Software Release v11.0"), produced at PBIO_00000012–46.
- "DNA Methylation and 5-base HiFi Sequencing Slides" ("DNA Methylation and 5-base HiFi Sequencing Slides"), produced at PBIO_00000047–79.
- "SMRTbell library preparation overview - Revio system v12.0.pdf" ("SMRTbell library preparation overview - Revio system"), produced at PBIO_00001390–1427.
- "SMRTbell library preparation overview - Sequel® IIe system v11.0.pdf" ("SMRTbell library preparation overview - Sequel® IIe system"), produced at PBIO_00001428–1472.
- "Release v11.0 Sales Training Part 2_v1.0" ("Release v11.0 Sales Training Part 2_v1.0"), produced at PBIO_00001473–1543.

## EXHIBIT A

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

- "Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0" ("Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0"), produced at PBIO_00001544–1610.
- "Sequel® IIe System v11.0 Release Overview Training - NA EMEA FAS and Tech Support - Part 1 040722_v1.0" ("Sequel® IIe System v11.0 Release Overview Training - NA EMEA FAS and Tech Support - Part 1 040722_v1.0"), produced at PBIO_00001611–1704.
- "Sequel® IIe System v11.0 Release Overview Training - NA EMEA FAS and Tech Support - Part 2 041422_v1.0" ("Sequel® IIe System v11.0 Release Overview Training - NA EMEA FAS and Tech Support - Part 2 041422_v1.0"), produced at PBIO_00001705–1824.
- "Sequel® IIe System v11.0 Release Overview Training BFX FAS Training 041422_v1.0" ("Sequel® IIe System v11.0 Release Overview Training BFX FAS Training 041422_v1.0"), produced at PBIO_00001825–1871.

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

**A. [Claim 1]**

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| [1.pre] A method for detecting a modification of a nucleotide in a nucleic acid molecule, the method comprising: | All of the Accused Products and/or Systems perform a method for detecting a modification of a nucleotide in a nucleic acid molecule using a Single Molecule, Real-Time (SMRT) sequencing technology known as HiFi Sequencing.<br><br>**Sequencing 101: from DNA to discovery — the steps of SMRT sequencing**<br><br>Starting a sequencing project can be daunting. First of all, there are several types of sequencing technologies, each based on unique processes. At PacBio, we use a technology called Single Molecule, Real-Time (SMRT) sequencing.<br><br>*See* Sequencing 101 (annotated)<br><br>For example, the epigenome workflows with PacBio HiFi reads using Accused Products include a DNA extraction step[1] followed by: (1) a library prep step, which includes the use of SMRTbell® prep kits[2], and/or barcoding kits[3], and/or AMPure® PB bead size selection kits[4]; (2) a sequencing step using a |

[1] The DNA extraction step can be performed using one or more of the following PacBio kits: Nanobind® kits (e.g., PN: 102-301-900, PN: 102-302-000, PN: 102-302-100, PN: 102-762-800, PN: 102-762-700) and/or Short Read Eliminator (SRE) kits (e.g., PN: 102-208-200, PN: 102-208-300, PN: 102-208-400).

[2] SMRTbell® prep kits include (a) SMRTbell® prep kit 3.0 (PN: 102-182-700), which is compatible with the Sequel® II, Sequel® IIe, and Revio™ Systems; and (b) HiFi express template prep kit 2.0 (PN: 102-088-900), which is compatible with the Sequel® II and Sequel® IIe Systems.

[3] Barcoding kits include (a) SMRTbell® barcoded adapter plate 3.0 (PN:102-009-200), which is compatible with the Sequel® II, Sequel® IIe, and Revio™ Systems; and (b) HiFi express template prep kit 2.0 (PN: 102-088-900), which is compatible with the Sequel® II and Sequel® IIe Systems.

[4] AMPure® PB bead size selection kits include AMPure PB beads size selection kit (PN: 102-182-500), which is compatible with use with the SMRTbell® prep kit 3.0 for the Sequel® II, Sequel® IIe, and Revio™ Systems.

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | PacBio long read sequencing system (e.g., Sequel® II, Sequel® IIe, and Revio™ Systems)[5]; and (3) a data analysis step using PacBio's SMRT® Link software[6].<br><br>That workflow is indicated below for the Sequel® II (e.g., Sequel® II and Sequel® IIe) and Revio™ systems:<br><br>Scale your whole genome sequencing (WGS) and epigenome workflows with PacBio® HiFi reads employing a single technician working one day a week. Sequence 25 human genomes at ≥ 30-fold coverage each month using four Sequel® II systems.<br><br>1 technician + 1 day per week =<br>5 SMRT® Cells per Sequel IIe system for up to 4 Sequel IIe systems<br><br>LIBRARY PREP<br>SMRTbell® prep kit 3.0 + AMPure® PB bead size selection<br>• 1 µg of DNA per SMRT Cell 8M<br>• Automation-friendly<br>• Up to 24 WGS samples prepared in 6.5 hours<br>• Single workflow for HiFi reads + 5mC base calling<br><br>SMRT® SEQUENCING<br>Sequel IIe system<br>• 6.5 days of walk-away automated runtime<br>• On-instrument base and 5mC calling allowing low compute costs<br>• Data transferred to server automatically after each SMRT Cell 8M is processed<br><br>DATA ANALYSIS<br>SMRT® Link<br>• HiFi reads + 5mC calling allow complete view of the genome and epigenome from a single data type<br>• Low data storage costs with 35 GB file sizes per SMRT Cell (175 GB per week per instrument)<br><br>*See, e.g.,* HiFi Sequel® II System Brochure |

[5] Sequel® II and Sequel® IIe Systems are used with a SMRT® cell (e.g., SMRT® Cell 8M tray (PN: 101-389-001)), which is loaded with a SMRTbell® library bound to polymerase using a Polymerase binding kit (e.g., Sequel® II binding kit 2.0 (PN: 101-789-500), and/or Sequel® II binding kit 2.1 (PN:101-820-500), and/or Binding Kit 3.1 and cleanup beads (PN:102-333-400), and/or Binding kit 3.2 and cleanup beads (PN: 102-333-300), and/or Sequencing kit (Sequel® II sequencing kit 2.0 (PN:101-820-200)). The Revio™ system is used with a Revio™ SMRT® cell (Revio™ SMRT® Cell tray (PN: 102-202-200)), which is loaded with a SMRTbell® library bound to polymerase using a Polymerase binding kit (Revio™ polymerase kit (PN: 102-817-600)), and/or Sequencing kit (Revio™ sequencing plate (PN: 102-587-400)).

[6] PacBio's SMRT® Link software is compatible with Revio™ (v12.0), Sequel® II (v12.0, v11.1, v11.0), and Sequel® IIe (v12.0, v11.1, v11.0) systems (*See* Software Downloads, available at https://www.pacb.com/support/software-downloads/).

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | **Sequencing specifications**<br>The Revio system utilizes nanofabricated Revio SMRT® Cells, each of which has 25 million zero-mode waveguide wells. Revio has four independent stages, allowing sequencing of multiple SMRT Cells in parallel. The onboard compute provides accurate basecalling with Google DeepConsensus, plus methylation calling in every run.<br><br>Library: 15–20 kb<br>Run time[1]: 24 hours<br>HiFi yield[2,3]: 1 Revio SMRT Cell: 90 Gb; 4 Revio SMRT Cells: 360 Gb<br>Base quality[2]: 90% Q30+<br>Methylation: 5mC at CpG sites<br><br>1. Run time specification is for the sequencing reaction.<br>2. HiFi yield and base quality is based on HG002/GM24385 human DNA extracted with Nanobind CBB kit and prepared with SMRTbell prep kit 3.0.<br>3. HiFi yield is dependent on library fragment size. Yield is typically lower for shorter libraries.<br><br>*See, e.g.,* Revio™ Specification Sheet<br><br>The methods and products that are used in this workflow are used to generate HiFi reads and allows 5mC base calling. i.e., detecting methylation, of nucleic acid molecules<br><br>For example, all of the Accused Products and/or Systems can detect 5mC methylation (a modification of a nucleotide) in a DNA molecule using HiFi Sequencing as shown below:<br><br>**HUMAN AND EUKARYOTE EPIGENETICS**<br>Use 5-base HIFI sequencing to study epigenetics of humans and other eukaryotes. Measure 5mC methylation at CpG sites directly from the sequencing instrument without the need for bisulfite treatment or other special library preparation methods.<br>Download example data<br><br>*See, e.g.,* PacBio Epigenetics Webpage (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

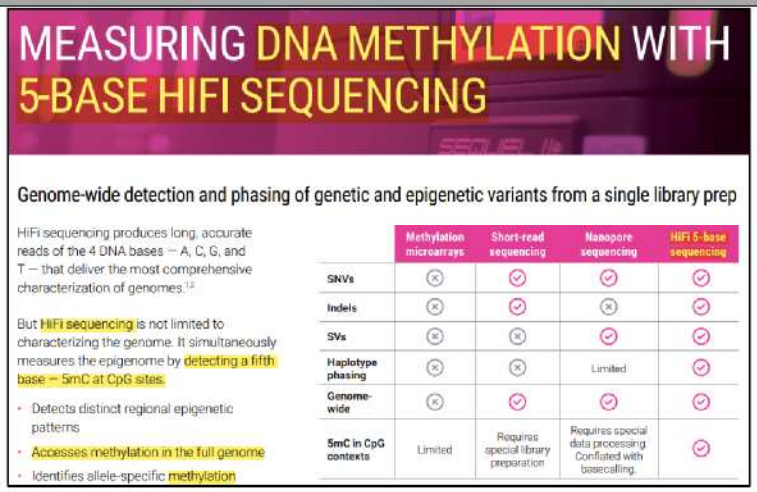
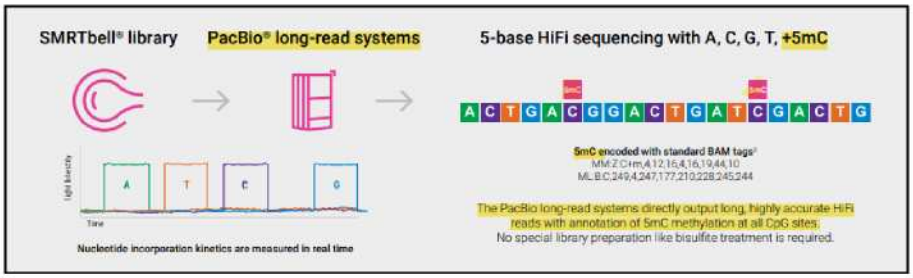
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>MEASURING DNA METHYLATION WITH 5-BASE HIFI SEQUENCING<br>Genome-wide detection and phasing of genetic and epigenetic variants from a single library prep<br>HiFi sequencing produces long, accurate reads of the 4 DNA bases — A, C, G, and T — that deliver the most comprehensive characterization of genomes.[1,2]<br>But HiFi sequencing is not limited to characterizing the genome. It simultaneously measures the epigenome by detecting a fifth base — 5mC at CpG sites.<br>• Detects distinct regional epigenetic patterns<br>• Accesses methylation in the full genome<br>• Identifies allele-specific methylation<br>Methylation microarrays / Short-read sequencing / Nanopore sequencing / HiFi 5-base sequencing<br>SNVs: ✗ / ✓ / ✓ / ✓<br>Indels: ✗ / ✓ / ✗ / ✓<br>SVs: ✗ / ✗ / ✓ / ✓<br>Haplotype phasing: ✗ / ✗ / Limited / ✓<br>Genome-wide: ✗ / ✓ / ✓ / ✓<br>5mC in CpG contexts: Limited / Requires special library preparation / Requires special data processing. Conflated with basecalling. / ✓<br>*See, e.g.,* PacBio Measuring DNA Methylation Application Brief (annotated) at PBIO_00000010.<br>  <br>SMRTbell® library<br>PacBio® long-read systems<br>5-base HiFi sequencing with A, C, G, T, +5mC<br>A C T G A C G G A C T G A T C G A C T G<br>5mC encoded with standard BAM tags[3]<br>MM:Z:C+m,4,12,16,4,16,19,44,10<br>ML:B:C,249,4,247,177,210,228,245,244<br>The PacBio long-read systems directly output long, highly accurate HiFi reads with annotation of 5mC methylation at all CpG sites.<br>No special library preparation like bisulfite treatment is required.<br>Light Intensity<br>Time<br>Nucleotide incorporation kinetics are measured in real time<br>*See, e.g.,* PacBio Measuring DNA Methylation Application Brief (annotated) at PBIO_00000010. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

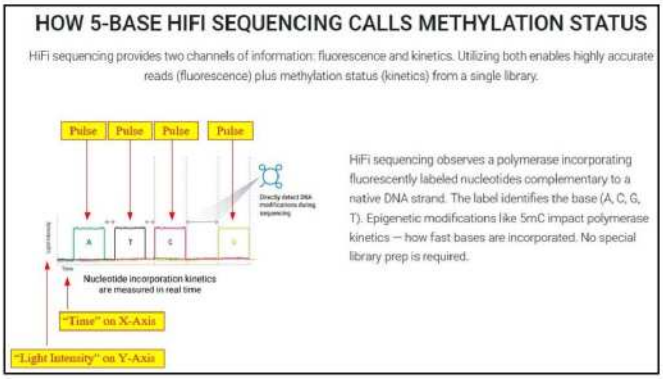
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | **EPIGENETIC ANALYSIS WITH 5-BASE HIFI SEQUENCING**<br>PacBio HiFi sequencing simultaneously calls the four DNA bases and 5mC from untreated genomic DNA. Achieve genome-wide detection and phasing of genetic and epigenetic variants from a single, standard HiFi library prep with long and accurate reads.<br>With the power of long-read sequencing, you can achieve:<br>Epigenetics in every run — no bisulfite treatment required<br>Unlike methods that require chemical conversion of DNA, HiFi sequencing detects modifications in native DNA through impacts on the kinetics of base incorporation.<br>High accuracy of sequence and methylation<br>Methylation detection with HiFi sequencing is highly concordant to bisulfite sequencing.<br>Access the full genome<br>Access difficult regions of the genome like repeats and centromeres that are beyond the reach of short-read sequencing.<br>Phasing<br>Identify allele-specific methylation, whether due to parental imprinting, genetic variation, or repeat expansions.<br>*See, e.g.,* PacBio Epigenetics Webpage (annotated).<br>*See also* Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002; Customer Overview of System and Software Release v11.0 at PBIO_00000013–33; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000048–78; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001479, 1489, 1492–1516; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001574–1589; Sequel® IIe System v11.0 Release Overview Training - NA EMEA FAS and Tech Support - Part 1 040722_v1.0 at PBIO_00001677–1679; Sequel® IIe System v11.0 Release Overview Training - NA EMEA FAS and |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

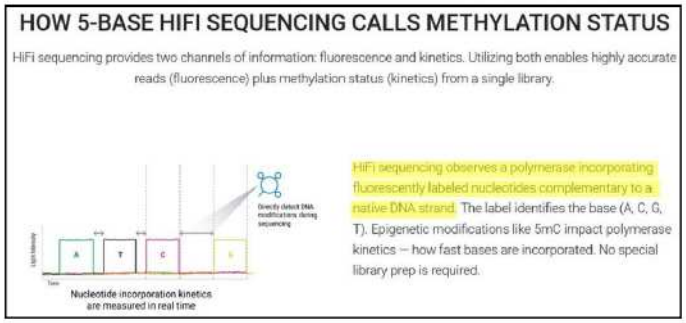
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | Tech Support - Part 2 041422_v1.0 at PBIO_00001776–1778; Sequel® IIe System v11.0 Release Overview Training BFX FAS Training 041422_v1.0 at PBIO_00001835–1850. |
| [1.a] (a) receiving data acquired by measuring pulses in an optical signal corresponding to nucleotides sequenced in a sample nucleic acid molecule and obtaining, from the data, values for the following properties: | The Accused Products and/or Systems (e.g., PacBio's Sequel® II, Sequel® IIe, and Revio™ long-read sequencing systems) receive data acquired by measuring pulses in an optical signal corresponding to nucleotides sequenced in a sample nucleic acid molecule and obtaining, from the data, values for each nucleotide.<br><br>For example, the Accused Products and/or Systems measure an optical signal—light—that is emitted during HiFi sequencing. Those optical signals are then recorded (i.e., received) as data "in a format that can then be analyzed" to obtain values of the properties discussed further below.<br><br>Step 3: sequencing<br><br>Once your library is prepared, the PacBio sequencing system, such as the Sequel II system, takes over. At the heart of SMRT sequencing is the SMRT Cell, which contains millions of tiny wells called zero-mode waveguides (ZMWs). Single molecules of DNA are immobilized in these wells, and as the polymerase incorporates each nucleotide, light is emitted, and nucleotide incorporation is measured in real time. The reactions are recorded in a format that can then be analyzed using on-instrument tools, as well as additional software.<br><br>*See, e.g.,* Sequencing 101 (annotated)<br><br>The light emitted and the data acquired during HiFi sequencing comes from fluorescently labeled nucleotides from a DNA strand.<br><br>For example, PacBio explains that "the fundamental characteristics of PacBio sequencing" includes "observing in real time a DNA polymerase synthesizing a DNA strand incorporating fluorescently labeled nucleotides" and that the fluorescence "labels the color of the ***pulses*** that are observed" which "indicate the identity of the base A, C, G, or T" nucleotides. HiFi Sequencing Video at 2:35–2:54. Thus, PacBio describes the fluorescent labeling performed by the Accused Products and/or Systems as using pulses of light. |

Those "pulses" of fluorescent color are shown on PacBio's website:




*See, e.g.*, PacBio Epigenetics Webpage (annotated).

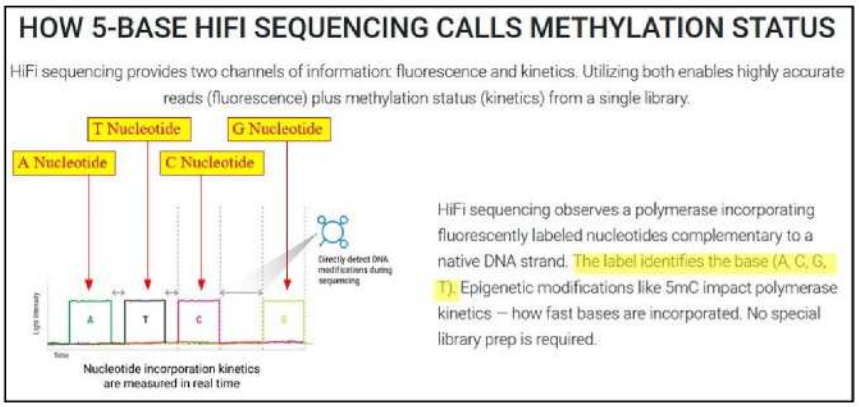
In addition, that same webpage explains that "HiFi sequencing observes a polymerase incorporating fluorescently labeled nucleotides complementary to a native DNA strand" as shown below:




*See, e.g.*, PacBio Epigenetics Webpage (annotated).

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | *See also* Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002; Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000056–59; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576. |
| [1.a.1] for each nucleotide: an identity of the nucleotide, | As to claim element [1.a.1], the data acquired by measuring pulses in the optical signal include values, for each nucleotide, for the identity of the nucleotide.   *See, e.g.,* PacBio Epigenetics Webpage (annotated). |

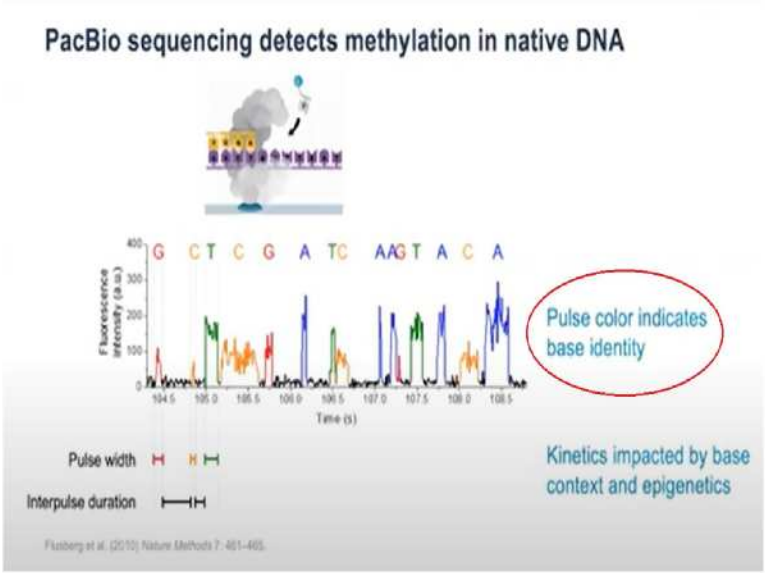
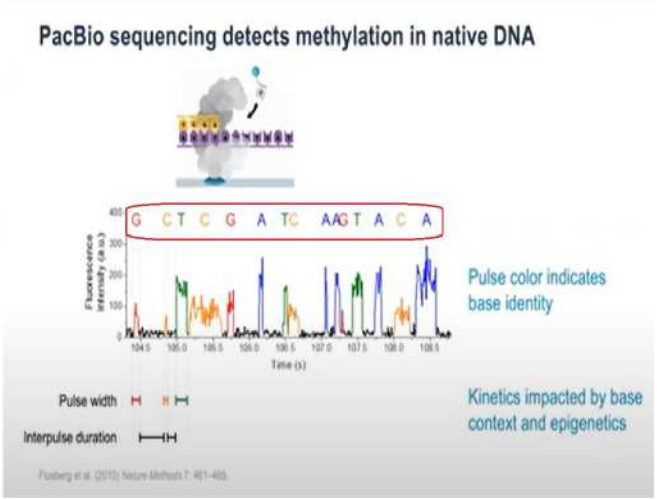
| | |
|---|---|
| | PacBio explains the fluorescent labeling processing of the Accused Products and/or Systems detecting "the color of the pulses that are observed" which "indicate the identity of the base A, C, G, or T" nucleotides. HiFi Sequencing Video at 2:45–2:54.<br>  <br><br>*See, e.g.,* HiFi Sequencing Video at 2:45–2:54 (annotated).<br>*See also* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492.<br><br>*See also* Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002; Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000056–59; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| [1.a.2] [for each nucleotide:]<br>a position of the nucleotide within the sample nucleic acid molecule, | As to claim element [1.a.2], the data acquired by measuring pulses in the optical signal include values, for each nucleotide, for the position of the nucleotide within the sample nucleic acid molecule.<br>PacBio shows that it captures the positions of the bases over the time of the sequencing read:<br>  <br><br>*See, e.g.,* HiFi Sequencing Video at 2:45–2:54 (annotated).<br>*See also* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492. |

For example, the Accused Products and/or Systems receive data corresponding to the subsequence or subread of a sample nucleic acid—referred to as a "query"—which is aligned with a target or reference sequence, thereby obtaining the position of each nucleotide within the sample nucleic acid.

### 2.9 Use of read tags for per-read information

| Tag | Type | Description |
|---|---|---|
| qs | i | 0-based start of query in the ZMW read (absent in CCS) |
| qe | i | 0-based end of query in the ZMW read (absent in CCS) |
| ws | i | Start of first base of the query ('qs') in approximate raw frame count since start of movie. For a CCS read, the start of the first base of the first incorporated subread. |
| we | i | Start of last base of the query ('qe - 1') in approximate raw frame count since start of movie. For a CCS read, the start of the last base of the last incorporated subread. |
| zm | i | ZMW hole number |
| np | i | NumPasses (1 for subreads, variable for CCS—encodes number of *complete* passes of the insert) |
| ec | f | Effective coverage for CCS reads, the average subread coverage across all windows (only present in CCS reads) |
| rq | f | Float in [0, 1] encoding expected accuracy |
| sn | B,f | 4 floats for the average signal-to-noise ratio of A, C, G, and T (in that order) over the HQRegion |

*See, e.g.,* PacBioFileFormats Documentation, p. 10.

PacBio shows that the position is included in the "feature vector" it generates from the reads, which Dr. Wenger refers to as part of "kinetics": "The basic idea is for every CpG site in a read, so every consecutive CG in a read to produce a feature vector that consists of the kinetics, the pulse widths, and inter-pulse durations in a 16 base window around that CpG site on both strands of the sequence." HiFi Sequencing Video at 7:21 - 7:47.

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

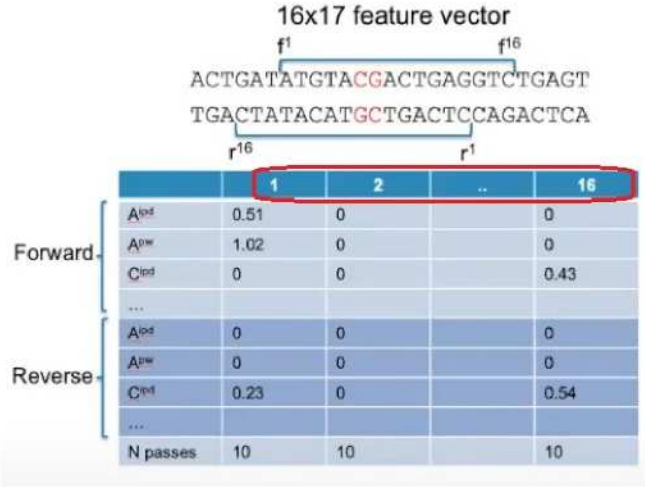
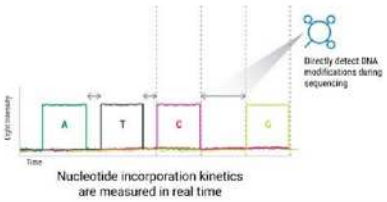
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Sequencing Video at 7:21 - 7:47.<br><br>*See also* Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002; Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000056–59; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576. |
| [1.a.3] [for each nucleotide:]<br>a width of the pulse corresponding to the nucleotide, and | As to claim element [1.a.3], the data acquired by measuring pulses in the optical signal include values, for each nucleotide, for the width of the pulse corresponding to the nucleotide. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | For example, the Accused Products and/or Systems receive data corresponding to the time it takes to incorporate a base, which indicates the width of the pulse corresponding to the nucleotide.<br><br>**HOW 5-BASE HIFI SEQUENCING CALLS METHYLATION STATUS**<br>HiFi sequencing provides two channels of information: fluorescence and kinetics. Utilizing both enables highly accurate reads (fluorescence) plus methylation status (kinetics) from a single library.<br>  <br><br>HiFi sequencing observes a polymerase incorporating fluorescently labeled nucleotides complementary to a native DNA strand. The label identifies the base (A, C, G, T). Epigenetic modifications like 5mC impact polymerase kinetics — how fast bases are incorporated. No special library prep is required.<br><br>*See, e.g.,* PacBio Epigenetics Webpage (annotated). |

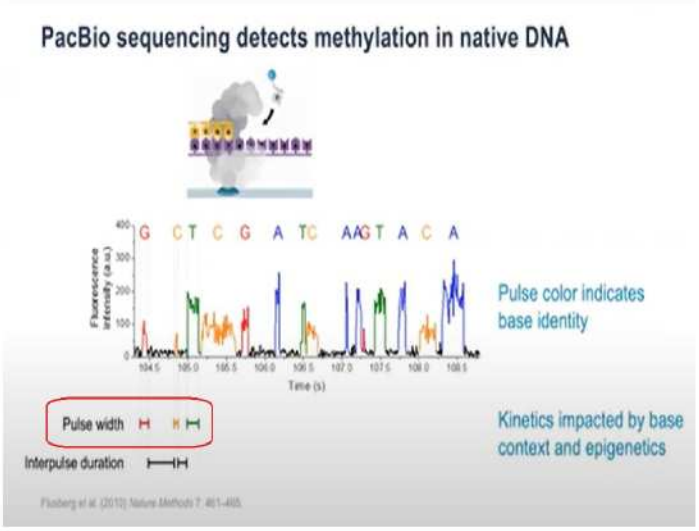
| | |
|---|---|
| | PacBio shows that it captures the pulse widths during sequencing reads:<br>  <br><br>*See, e.g.,* HiFi Sequencing Video at 2:45–2:54 (annotated).<br>*See, e.g.,* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492.<br><br>PacBio also shows that the pulse width is included in the "feature vector" it generates from the reads: "The basic idea is for every CpG site in a read, so every consecutive CG in a read to produce a feature vector that consists of the kinetics, the ***pulse widths***, and inter-pulse durations in a 16 base window around that CpG site on both strands of the sequence." HiFi Sequencing Video at 7:21 - 7:47 (emphasis added). |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

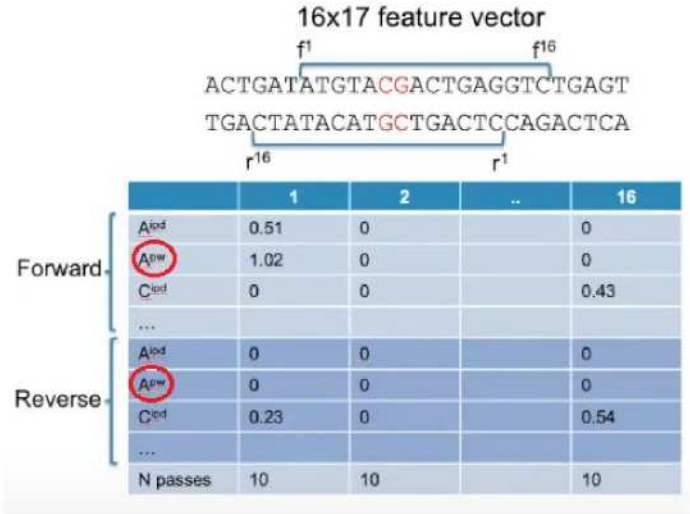
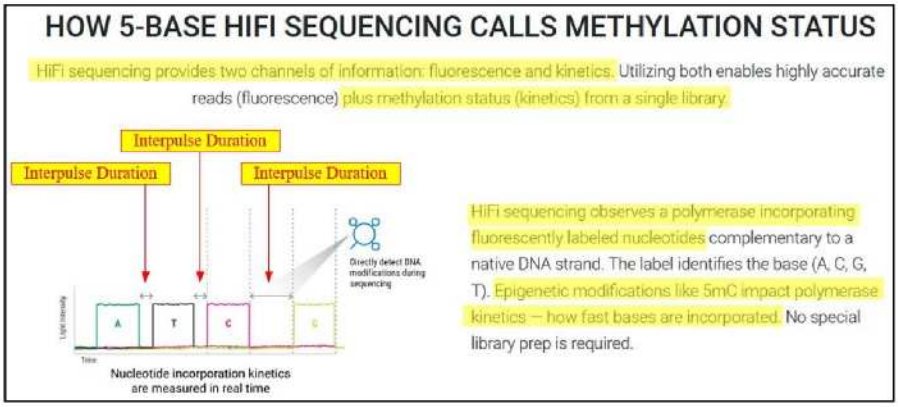
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Sequencing Video at 7:21 - 7:47 (annotated).<br><br>*See also* Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000056–59; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576. |
| [1.a.4] [for each nucleotide:] an interpulse duration representing a time between the pulse corresponding to | As to claim element [1.a.4], the data acquired by measuring pulses in the optical signal include values, for each nucleotide, for a time between the pulse corresponding to the nucleotide and a pulse corresponding to a neighboring nucleotide. |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| the nucleotide and a pulse corresponding to a neighboring nucleotide; | For example, the Accused Products and/or Systems receive data corresponding to the time between adjacent incorporations, which indicates the interpulse duration.<br>  <br><br>*See, e.g.,* PacBio Epigenetics Webpage (annotated). |

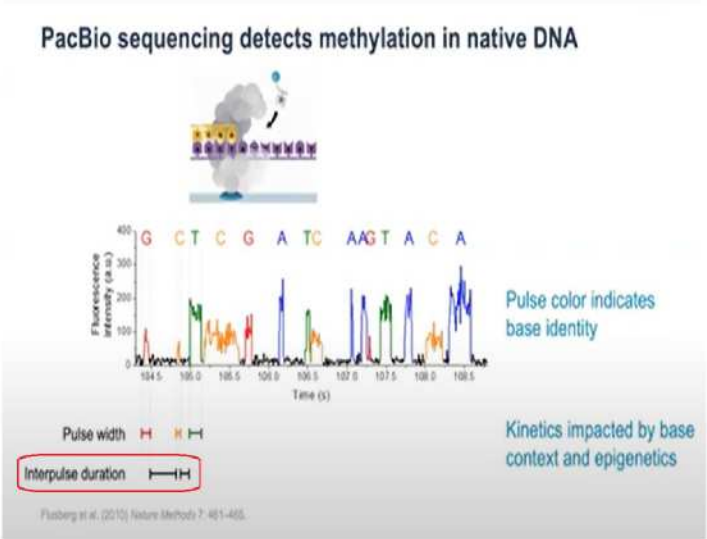
| | |
|---|---|
| | PacBio shows that it captures interpulse duration during sequencing reads:<br>  <br><br>*See, e.g.,* HiFi Sequencing Video at 2:45–2:54 (annotated).<br>*See also* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492.<br><br>PacBio also shows that the pulse width is included in the "feature vector" it generates from the reads: "The basic idea is for every CpG site in a read, so every consecutive CG in a read to produce a feature vector that consists of the kinetics, the pulse widths, and ***inter-pulse durations*** in a 16 base window around that CpG site on both strands of the sequence." HiFi Sequencing Video at 7:21 - 7:47 (emphasis added). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

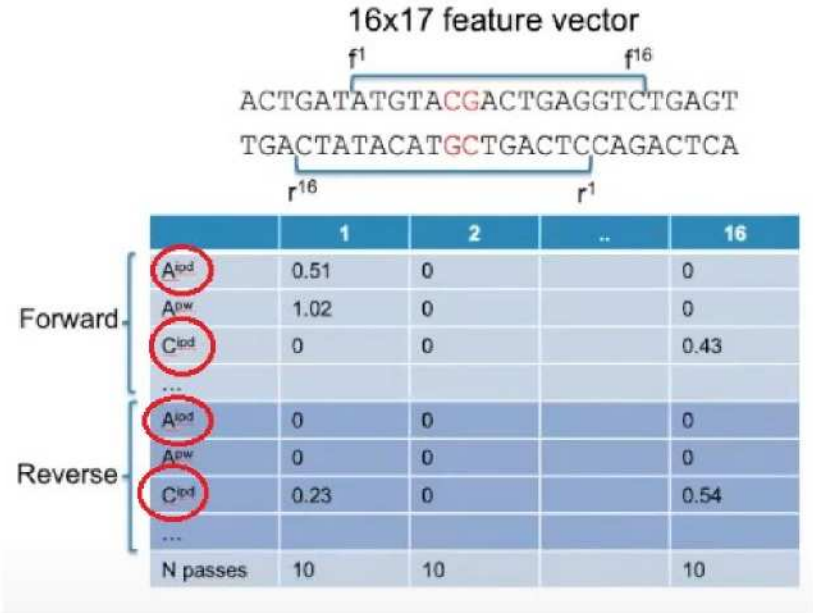
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Sequencing Video at 7:21 - 7:47 (annotated).<br><br>*See also* Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000056–59; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576. |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| [1.b] (b) creating an input data structure, the input data structure comprising a window of the nucleotides sequenced in the sample nucleic acid molecule, wherein the input data structure includes, for each nucleotide within the window, the properties: | The Accused Products and/or Systems create an input data structure that comprises a window of the nucleotides sequenced in the sample nucleic acid molecule.<br><br>For example, the PacBio products first receive a file or set of files in BAM format that contain a plurality of HiFi reads obtained from the sequencing data ("the BAM file").<br><br>As of the 3.0 release of SMRTanalysis, PacBio is embracing the industry standard BAM format for (both aligned and unaligned) basecall data files. We have also formulated a BAM companion file format (*bam.pbi*) enabling fast access to a richer set of per-read information as well as compatibility for software built around the legacy *cmp.h5* format.<br><br>*See, e.g.,* PacBioFileFormats Documentation at. 5. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | **primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>**Algorithm**<br>The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M. SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.<br><br>**Usage**<br>`primrose [options] <HiFi INPUT> <HiFi OUTPUT>`<br><br>Required: `HiFi INPUT` — Description: Input BAM file or ConsensusReadSet XML file.<br>Required: `HiFi OUPUT` — Description: Output BAM file or ConsensusReadSet XML file.<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*Also see* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. |

| | |
|---|---|
| | The BAM file contains sequence alignment data with specifications that are standardized so that the information provided in the fields are generally applicable to raw or aligned sequence reads across different sequencing platforms.<br><br>**1.1 BAM everywhere**<br><br>*Unaligned* BAM files representing the *subreads* will be produced natively by the PacBio instrument. The subreads BAM will be the starting point for secondary analysis. In addition, the artifacts resulting from removing adapter and LQ region sequences will be retained in a `scraps.bam` file.<br><br>The circular consensus tool/workflow (CCS) will take as input an unaligned subreads BAM file and produce an output BAM file containing unaligned *consensus* reads.<br><br>Alignment (mapping) programs take these unaligned BAM files as input and will produce *aligned* BAM files, faithfully retaining all tags and headers.<br><br>*See, e.g.,* PacBioFileFormats Documentation at 8.<br><br>The BAM format is a binary, compressed, record-oriented container format for raw or aligned sequence reads. The associated SAM format is a text representation of the same data. The specifications for BAM/SAM are maintained by the SAM/BAM Format Specification Working Group.<br><br>PacBio-produced BAM files are fully compatible with the BAM specification. In this document we describe the way we make use of the extensibility mechanisms of the BAM specification to encode PacBio-specific information, as well as conventions we adhere to.<br><br>An example file adhering to this specification will be maintained in the *pbcore* Python library.<br><br>*See, e.g.,* PacBioFileFormats Documentation at 9.<br><br>The HiFi reads in the BAM file are then used to create the data structure of claim element [1.b], i.e., an input data structure, for claim element [1.c].<br><br>To create the HiFi reads in BAM format for the BAM file(s), the PacBio Products perform an analysis called circular consensus sequencing (CCS). |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | **WHAT ARE HIFI READS?**<br>HiFi reads are produced using circular consensus sequencing (CCS) mode on PacBio long-read systems. HiFi reads provide base-level resolution with 99.9% single-molecule read accuracy.<br>HiFi reads can be used across a wide range of SMRT sequencing applications, from whole genome sequencing for *de novo* assembly, comprehensive variant detection, epigenetic characterization, RNA sequencing and more.<br>*See, e.g.,* PacBio HiFi Sequencing Webpage (annotated).<br>*See also* SMRT® Link User Guide, v11.0 at PBIO_00000651–652, PBIO_00000665–667; SMRT® Link User Guide, v11.1 at PBIO_00000793–803, PBIO_00000818–820; SMRT® Link User Guide, v12.0 at PBIO_00000959–962, PBIO_00000979–981; SMRT® Tools Reference Guide, v11.0 at PBIO_00001012–1020; SMRT® Tools Reference Guide, v11.1 at PBIO_00001130–1138.<br><br>As noted above, PacBio used a 16-base-pair window of nucleotides for what it refers to as a "feature vector."<br>For example, Dr. Wenger during the development of this windowing process, every CpG site in a HiFi read is used to produce "a feature vector that consists of the kinetics, the pulse widths, and interpulse durations in a 16-base window around that CpG site on both strands of the sequence." *See* HiFi Sequencing Video, 7:25-7:47.<br>Each feature vector, which is a smaller data structure built from the BAM file, serves as an input to a convolutional neural network (CNN). *See* HiFi Sequencing Video, 7:08. |

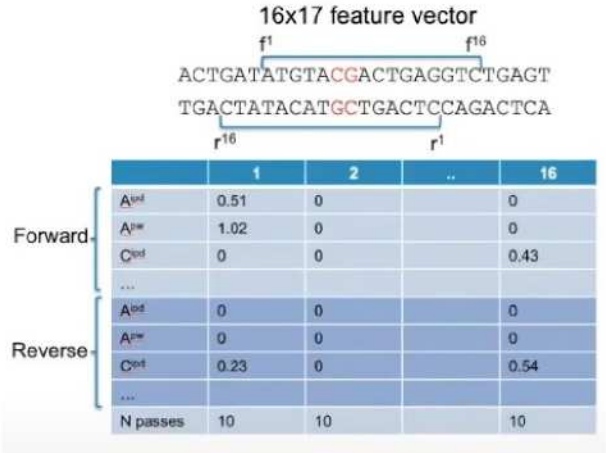
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Sequencing Video, 7:25-7:47.<br>Thus, the HiFi reads and feature vectors created from the BAM file, i.e., the "input data structure," comprise "a window of the nucleotides sequenced in the sample nucleic acid molecule."<br>To detect 5mC methylation, the CCS process analyzes kinetic data, resulting in the creation of a BAM file and feature vectors that include both conventional and PacBio-specific information. *See* PacBioFileFormats Documentation at 5 ("In this document we describe the way we make use of the extensibility mechanisms of the BAM specification to encode PacBio-specific information, as well as conventions we adhere to.").<br>*See also* Customer Overview of System and Software Release v11.0 at PBIO_00000013–33; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000056–59; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001580. |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| [1.b.1] the identity of the nucleotide, | As to claim element [1.b.1], the input data structure includes, for each nucleotide within the window, the identity of the nucleotide.<br><br>For example, an input BAM file, including read tags and feature vectors for HiFi per-read-base kinetic information, includes the identity of the nucleotide as indicated in the SEQ field. *See* PacBioFileFormats Documentation at 15.<br><br>**2.11 Use of read tags for HiFi per-read-base kinetic information**<br><br>The following read tags contain averaged kinetic information (IPD/PulseWidth) from subreads when applying CCS to generate HiFi reads. These are computed and stored independently for both orientations of the insert. Forward is defined & stored with respect to the orientation represented in SEQ and is considered to be the native orientation. Reverse tags are stored in the opposite direction, e.g. from the last base to the first. As with other PacBio-specific tags, aligners will not re-orient these fields.<br><br>Tag \| Type \| Description<br>fi \| B,C \| Forward IPD (codec V1)<br>ri \| B,C \| Reverse IPD (codec V1)<br>fp \| B,C \| Forward PulseWidth (codec V1)<br>rp \| B,C \| Reverse PulseWidth (codec V1)<br>fn \| i \| Forward number of complete passes (zero or more)<br>rn \| i \| Reverse number of complete passes (zero or more)<br><br>The following clipping example illustrates the coordinate system for these tags, shown as stored in the BAM file:<br><br>--------<br>Original<br>--------<br><br>SEQ: A A C C G T T A G C<br>fi/fp: f0, f1, f2, f3, f4, f5, f6, f7, f8, f9<br>ri/rp: r9, r8, r7, r6, r5, r4, r3, r2, r1, r0<br><br>----------<br>Clipped to [1, 4)<br>----------<br><br>SEQ: A C C<br>fi/fp: f1, f2, f3<br>ri/rp: r3, r2, r1<br><br>*See, e.g.,* PacBioFileFormats Documentation, p. 15 (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

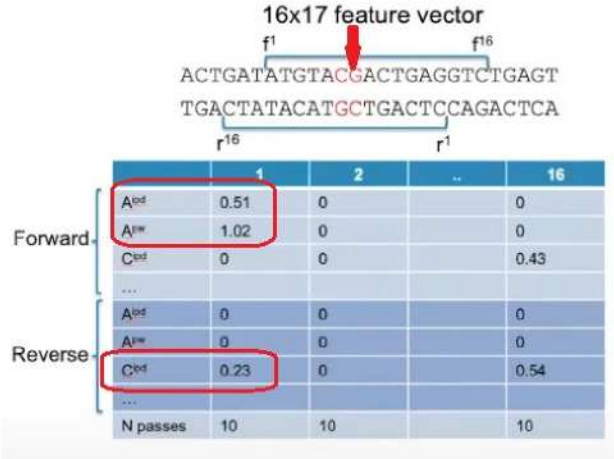
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | During the development of this windowing process, every CpG site in a HiFi read is used to produce "a feature vector that consists of the kinetics, the pulse widths, and interpulse durations in a 16-base window around that CpG site on both strands of the sequence." *See* HiFi Sequencing Video, 7:25-7:47.<br><br>Each feature vector, which is a smaller data structure built from the BAM file, serves as an input to a convolutional neural network (CNN). *See* HiFi Sequencing Video, 7:08.<br><br>The identity of the nucleotides, for example each CpG site and the 16-base window around that CpG site (arrow), is a property of the feature vector:<br><br>  <br><br><br>*See, e.g.,* HiFi Sequencing Video, 7:25-7:47 (annotated).<br><br>*See also* Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002; Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576, 1580. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

<table>
<tr><th>U.S. Patent No. 11,091,794<br>CLAIM 1</th><th>Pacific Biosciences of California, Inc.</th></tr>
<tr><td>[1.b.2] a position of the nucleotide with respect to a target position within the window,</td><td>

As to claim element [1.b.2], the input data structure includes, for each nucleotide within the window, a position of the nucleotide with respect to a target position within the window.

For example, each feature vector includes a position of the nucleotide with respect to a target position within the window by aligning the start and end positions of a sequence read (qStart, qEnd) to that of a target sequence (tStart, tEnd). *See* PacBioFileFormats Documentation at 10.

## 2.3 *Query* versus *aligned query* terminology

A sequence read presented to an aligner is termed a *query*; typically this query will be a subsequence of an entire PacBio ZMW read—most commonly, it will be a *subread*, which is basecalls from a single pass of the insert DNA molecule.

Upon alignment, generally only a subsequence of the query will align to the reference genome, and that subsequence is referred to as the *aligned query*. Under *soft-clipping*, the entirety of the query is stored in the aligned BAM, but the CIGAR field indicates that some bases at either end are excluded from the alignment.

Abstractly, we denote the extent of the *query* in ZMW read as *[qStart, qEnd)* and the extent of the aligned subinterval as *[aStart, aEnd)* The following graphic illustrates these intervals:

```
          qStart                      qEnd
0         |  aStart               aEnd  |
[--...----+--+---------------------+-----+-----...------)  < "ZMW read" coord. system
             ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~               <  query; "-" = aligning
subseq.
[--...--------+---------...----------+-----------...------)  < "ref." / "target" coord.
system
0             tStart          tEnd
```

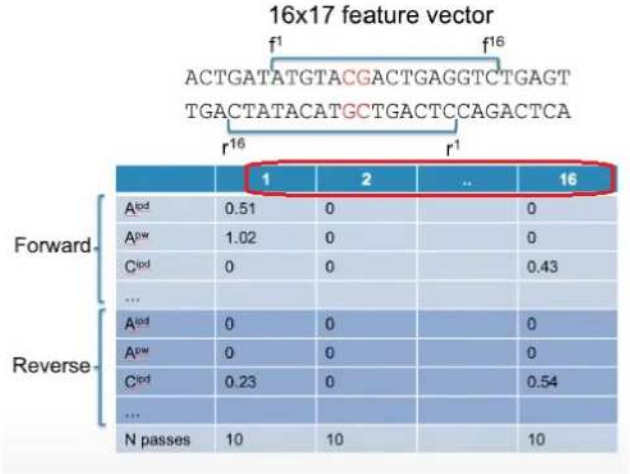
In our BAM files, the qStart, qEnd are contained in the qs and qe tags, (and reflected in the QNAME); the bounds of the *aligned query* in the ZMW read can be determined by adjusting qs and qe by the number of soft-clipped bases at the ends of the alignment (as found in the CIGAR).

*See, e.g.,* PacBioFileFormats Documentation at 10 (annotated).

During the development of this windowing process, every CpG site in a HiFi read is used to produce "a feature vector that consists of the kinetics, the pulse widths, and interpulse durations in a 16-base window around that CpG site on both strands of the sequence." *See* HiFi Sequencing Video, 7:25-7:47.

</td></tr>
</table>

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | Each feature vector, which is a smaller data structure built from the BAM file, serves as an input to a convolutional neural network (CNN). *See* HiFi Sequencing Video, 7:08.<br>The position of the nucleotides is a component of the feature vector (positions circled below):<br>  <br><br>*See, e.g.,* HiFi Sequencing Video, 7:25-7:47 (annotated).<br>*See also* Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002; Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576, 1580. |
| [1.b.3] the width of the pulse corresponding to the nucleotide, and | As to claim element [1.b.3], the input data structure includes, for each nucleotide within the window, the width of the pulse corresponding to the nucleotide. |

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | For example, the BAM file, including read tags and feature vectors for HiFi per-read-base kinetic information, includes the width of the pulse corresponding to the nucleotide as indicated in the fp and rp tags. *See* PacBioFileFormats Documentation at 15. |

**2.11 Use of read tags for HiFi per-read-base kinetic information**

The following read tags contain averaged kinetic information (IPD/PulseWidth) from subreads when applying CCS to generate HiFi reads. These are computed and stored independently for both orientations of the insert. Forward is defined & stored with respect to the orientation represented in `SEQ` and is considered to be the native orientation. Reverse tags are stored in the opposite direction, e.g. from the last base to the first. As with other PacBio-specific tags, aligners will not re-orient these fields.

| Tag | Type | Description |
|---|---|---|
| fi | B,C | Forward IPD (codec V1) |
| ri | B,C | Reverse IPD (codec V1) |
| fp | B,C | Forward PulseWidth (codec V1) |
| rp | B,C | Reverse PulseWidth (codec V1) |
| fn | i | Forward number of complete passes (zero or more) |
| rn | i | Reverse number of complete passes (zero or more) |

The following clipping example illustrates the coordinate system for these tags, shown as stored in the BAM file:

```
--------
Original
--------

    SEQ:  A   A   C   C   G   T   T   A   G   C
  fi/fp: f0, f1, f2, f3, f4, f5, f6, f7, f8, f9
  ri/rp: r9, r8, r7, r6, r5, r4, r3, r2, r1, r0

-----------------
Clipped to [1, 4)
-----------------

    SEQ:  A   C   C
  fi/fp: f1, f2, f3
  ri/rp: r3, r2, r1
```

*See, e.g.,* PacBioFileFormats Documentation at 15 (annotated).

During the development of this windowing process, every CpG site in a HiFi read is used to produce "a feature vector that consists of the kinetics, the ***pulse widths***, and interpulse durations in a 16-base

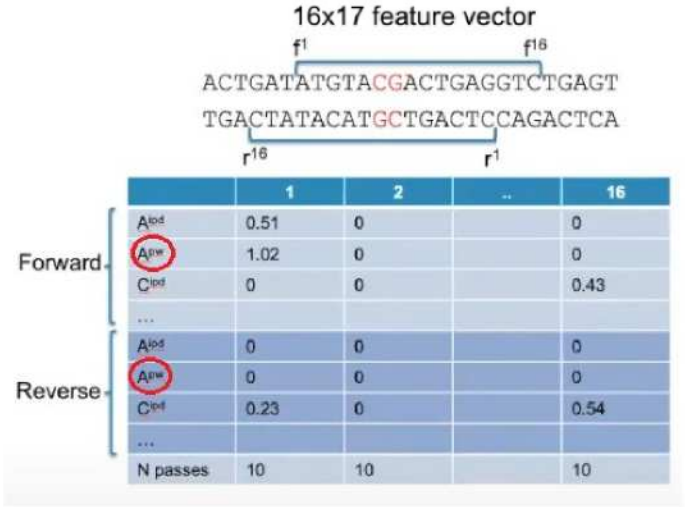
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | window around that CpG site on both strands of the sequence." *See* HiFi Sequencing Video, 7:25-7:47 (emphasis added).<br><br>Each feature vector, which is a smaller data structure built from the BAM file, serves as an input to a convolutional neural network (CNN). *See* HiFi Sequencing Video, 7:08.<br><br>The pulse width is a component of the feature vector (pulse width (pw) circled below):<br><br>  <br><br>*See, e.g.,* HiFi Sequencing Video, 7:25-7:47 (annotated).<br><br>*See also* Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576, 1580. |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| [1.b.4] the interpulse duration; | As to claim element [1.b.4], the input data structure includes, for each nucleotide within the window, the interpulse duration.<br><br>For example, the BAM file, including read tags and feature vectors for HiFi per-read-base kinetic information, includes the interpulse duration or "IPD" as indicated in the fi and ri tags. *See* PacBioFileFormats Documentation at 15.<br><br>**2.11 Use of read tags for HiFi per-read-base kinetic information**<br><br>The following read tags contain averaged kinetic information (IPD/PulseWidth) from subreads when applying CCS to generate HiFi reads. These are computed and stored independently for both orientations of the insert. Forward is defined & stored with respect to the orientation represented in SEQ and is considered to be the native orientation. Reverse tags are stored in the opposite direction, e.g. from the last base to the first. As with other PacBio-specific tags, aligners will not re-orient these fields.<br><br>Tag \| Type \| Description<br>fi \| B,C \| Forward IPD (codec V1)<br>ri \| B,C \| Reverse IPD (codec V1)<br>fp \| B,C \| Forward PulseWidth (codec V1)<br>rp \| B,C \| Reverse PulseWidth (codec V1)<br>fn \| i \| Forward number of complete passes (zero or more)<br>rn \| i \| Reverse number of complete passes (zero or more)<br><br>The following clipping example illustrates the coordinate system for these tags, shown as stored in the BAM file:<br><br>--------<br>Original<br>--------<br><br>SEQ: A A C C G T T A G C<br>fi/fp: f0, f1, f2, f3, f4, f5, f6, f7, f8, f9<br>ri/rp: r9, r8, r7, r6, r5, r4, r3, r2, r1, r0<br><br>----------<br>Clipped to [1, 4)<br>----------<br><br>SEQ: A C C<br>fi/fp: f1, f2, f3<br>ri/rp: r3, r2, r1<br><br>*See, e.g.,* PacBioFileFormats Documentation at 15 (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | During the development of this windowing process, every CpG site in a HiFi read is used to produce "a feature vector that consists of the kinetics, the pulse widths, and ***interpulse durations*** in a 16-base window around that CpG site on both strands of the sequence." *See* HiFi Sequencing Video, 7:25-7:47 (emphasis added).<br><br>Each feature vector, which is a smaller data structure built from the BAM file, serves as an input to a convolutional neural network (CNN). *See* HiFi Sequencing Video, 7:08.<br><br>The interpulse duration is a component of the feature vector (interpulse duration (ipd) circled below):<br><br>16x17 feature vector<br>$f^1$ $f^{16}$<br>ACTGATATGTACGACTGAGGTCTGAGT<br>TGACTATACATGCTGACTCCAGACTCA<br>$r^{16}$ $r^1$<br><br>Forward: $A^{ipd}$ 0.51 0 0; $A^{pw}$ 1.02 0 0; $C^{ipd}$ 0 0 0.43; ...<br>Reverse: $A^{ipd}$ 0 0 0; $A^{pw}$ 0 0 0; $C^{ipd}$ 0.23 0 0.54; ...<br>N passes 10 10 10 |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | *See, e.g.,* HiFi Sequencing Video, 7:25-7:47 (annotated).<br><br>*See also* Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001575–1576, 1580. |
| [1.c] (c) inputting the input data structure into a model, the model trained by: | The Accused Products and/or Systems input the input data structure into a model.<br><br>For example, each feature vector, which is a smaller data structure built from the BAM file, serves as an input to a convolutional neural network (CNN) as shown below:<br><br>• A convolutional neural network model processes polymerase kinetics to determine the methylation status of each CpG site in a HiFi read.<br>• The model runs directly on the Sequel IIe system and is also available in SMRT Link.<br>• Methylation status is output using the BAM standard MM and ML tags.<br><br>*See, e.g.,* PacBio Epigenetics Webpage. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

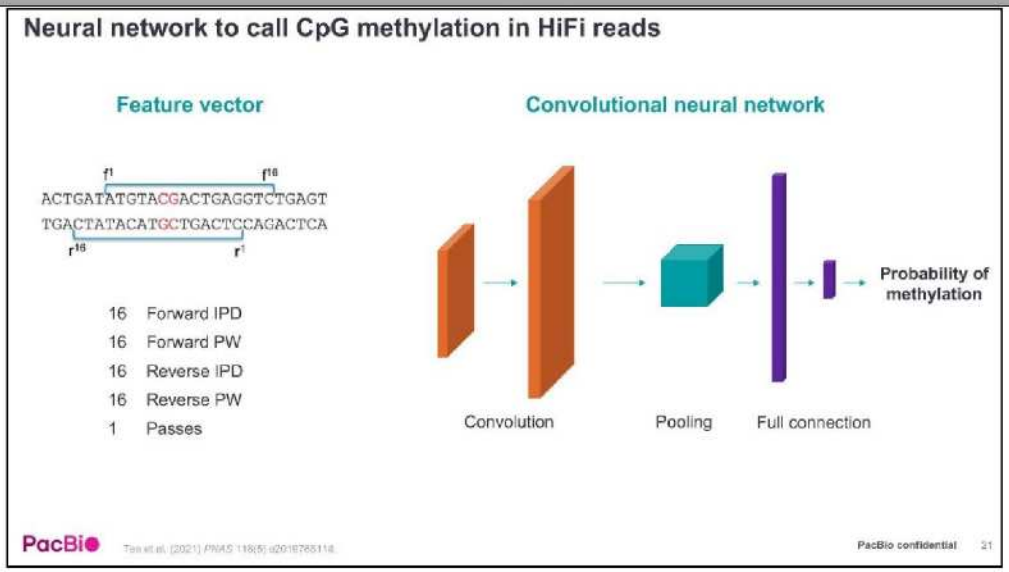
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493.<br><br>The Accused Products and/or Systems use the "primrose" tool. The primrose tool uses a convolutional neural network (CNN) to predict methylation state (5mC) of each CpG in a HiFi read. |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | **primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. *See also* HiFi Sequencing Video, 7:08.<br><br>The input for the primrose tool is a plurality of feature vectors created or built from a BAM file (i.e., the "input data structure" above). The inputs include the kinetic data (including nucleotide identity), pulse width, and inter-pulse distance for each window. |

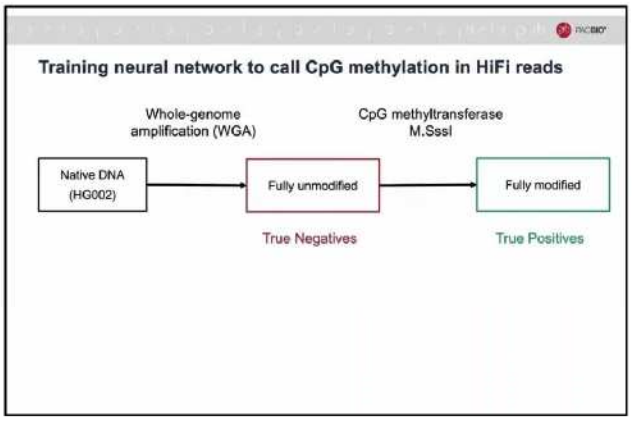
**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | **primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>**Algorithm**<br>The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M. SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.<br><br>**Usage**<br>`primrose [options] <HiFi INPUT> <HiFi OUTPUT>`<br><br>Required: `HiFi INPUT` — Description: Input BAM file or ConsensusReadSet XML file.<br>Required: `HiFi OUPUT` — Description: Output BAM file or ConsensusReadSet XML file.<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371; Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000056–59; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001492; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001578, 1580. |
| [1.c.1] receiving a first plurality of first data structures, each first data | As to claim element [1.c.1], the model is trained by receiving a first plurality of first data structures, each first data structure of the first plurality of data structures corresponding to a respective window of nucleotides sequenced in a respective nucleic acid molecule of a plurality of first nucleic acid molecules, |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| structure of the first plurality of data structures corresponding to a respective window of nucleotides sequenced in a respective nucleic acid molecule of a plurality of first nucleic acid molecules, wherein each of the first nucleic acid molecules is sequenced by measuring pulses in the optical signal corresponding to the nucleotides, wherein the modification has a known first state in a nucleotide at a target position in each window of each first nucleic acid molecule, each first data structure comprising values for the same properties as the input data structure, | wherein each of the first nucleic acid molecules is sequenced by measuring pulses in the optical signal corresponding to the nucleotides, wherein the modification has a known first state in a nucleotide at a target position in each window of each first nucleic acid molecule, each first data structure comprising values for the same properties as the input data structure. For example, the Primrose tool receives a first plurality of first data structures as claimed as shown below:<br><br>**primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>**Algorithm**<br>The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M. SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.<br><br>**Usage**<br>`primrose [options] <HiFi INPUT> <HiFi OUTPUT>`<br><br>Required: `HiFi INPUT` — Description: Input BAM file or ConsensusReadSet XML file.<br>Required: `HiFi OUPUT` — Description: Output BAM file or ConsensusReadSet XML file.<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | "In addition to developing the features for this model and the model structure, the other key attribute to producing a neural network model is training data and for that you want examples of methylation where you have sequences that are either fully methylated or fully unmethylated, and in order to train this model we produced data in a manner where we start with native human DNA for the sample HG002, and in this native DNA of course some of the sites are methylated and some are not. We then performed whole genome amplification using a PCR based approach, which produces then fully unmodified DNA. This library or a library was prepared from that DNA in sequenced, and that serves as a set, a large set, genome-wide true negatives of DNA strands that are not modified. Then to produce fully modified sites, we treated this whole genome amplified HG002 DNA with a CpG methyltransferase enzyme, which efficiently adds methylation to any CpG site in the DNA, and this then produces fully modified sequences, which again were sequences to produce a large set of true positives genome wide. These two sets were passed into the neural network to train the model." (HiFi Sequencing Video, 8:04–9:23.)<br>  <br><br>*See, e.g.,* HiFi Sequencing Video, 8:30.<br>*See also* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001578, 1580. |

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| [1.c.2] storing a plurality of first training samples, each including one of the first plurality of first data structures and a first label indicating the first state of the nucleotide at the target position, and | As to claim element [1.c.2], the model is trained by storing a plurality of first training samples, each including one of the first plurality of first data structures and a first label indicating the first state of the nucleotide at the target position. For example, the Primrose tool stores a plurality of first training samples as claimed as shown below: |

**primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.

`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)

**Algorithm**

The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M. SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.

**Usage**

```
primrose [options] <HiFi INPUT> <HiFi OUTPUT>
```

| Required | Description |
|---|---|
| `HiFi INPUT` | Input BAM file or ConsensusReadSet XML file. |
| `HiFi OUPUT` | Output BAM file or ConsensusReadSet XML file. |

*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).

*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371.

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

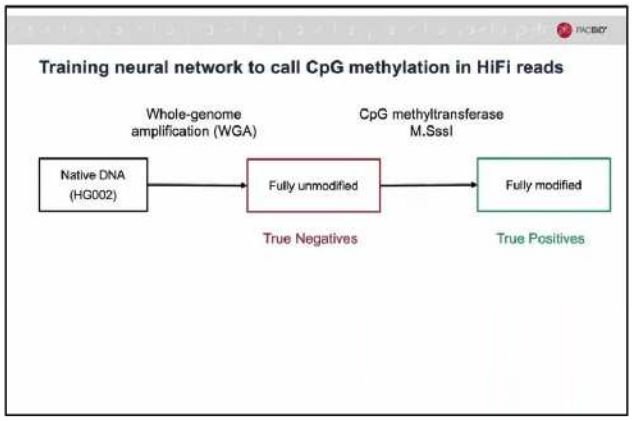
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | "In addition to developing the features for this model and the model structure, the other key attribute to producing a neural network model is training data and for that you want examples of methylation where you have sequences that are either fully methylated or fully unmethylated, and in order to train this model we produced data in a manner where we start with native human DNA for the sample HG002, and in this native DNA of course some of the sites are methylated and some are not. We then performed whole genome amplification using a PCR based approach, which produces then fully unmodified DNA. This library or a library was prepared from that DNA in sequenced, and that serves as a set, a large set, genome-wide true negatives of DNA strands that are not modified. Then to produce fully modified sites, we treated this whole genome amplified HG002 DNA with a CpG methyltransferase enzyme, which efficiently adds methylation to any CpG site in the DNA, and this then produces fully modified sequences, which again were sequences to produce a large set of true positives genome wide. These two sets were passed into the neural network to train the model." (HiFi Sequencing Video, 8:04–9:23.)<br>  <br><br>*See, e.g.,* HiFi Sequencing Video, 8:30.<br>*See also* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001578, 1580. |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| [1.c.3] optimizing, using the plurality of first training samples, parameters of the model based on outputs of the model matching or not matching corresponding labels of the first labels when the first plurality of first data structures is input to the model, wherein an output of the model specifies whether the nucleotide at the target position in the respective window has the modification, | As to claim element [1.c.3], the model is trained by optimizing, using the plurality of first training samples, parameters of the model based on outputs of the model matching or not matching corresponding labels of the first labels when the first plurality of first data structures is input to the model, wherein an output of the model specifies whether the nucleotide at the target position in the respective window has the modification. For example, the Primrose tool optimizing, using the plurality of first training samples, parameters of the model based on outputs of the model as claimed, as shown below:<br><br>**primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>**Algorithm**<br>The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M.SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.<br><br>**Usage**<br>`primrose [options] <HiFi INPUT> <HiFi OUTPUT>`<br><br>Required: `HiFi INPUT` — Input BAM file or ConsensusReadSet XML file.<br>Required: `HiFi OUPUT` — Output BAM file or ConsensusReadSet XML file.<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

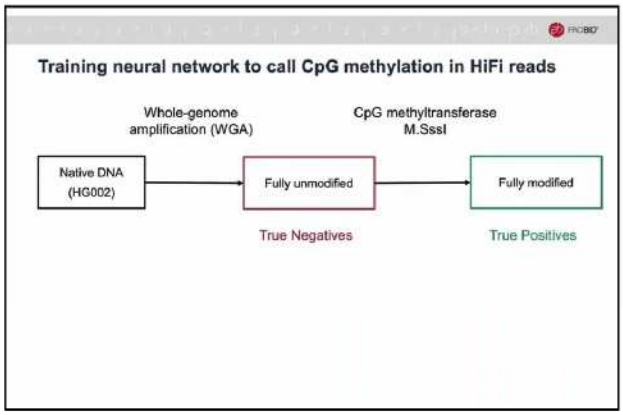
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | "In addition to developing the features for this model and the model structure, the other key attribute to producing a neural network model is training data and for that you want examples of methylation where you have sequences that are either fully methylated or fully unmethylated, and in order to train this model we produced data in a manner where we start with native human DNA for the sample HG002, and in this native DNA of course some of the sites are methylated and some are not. We then performed whole genome amplification using a PCR based approach, which produces then fully unmodified DNA. This library or a library was prepared from that DNA in sequenced, and that serves as a set, a large set, genome-wide true negatives of DNA strands that are not modified. Then to produce fully modified sites, we treated this whole genome amplified HG002 DNA with a CpG methyltransferase enzyme, which efficiently adds methylation to any CpG site in the DNA, and this then produces fully modified sequences, which again were sequences to produce a large set of true positives genome wide. These two sets were passed into the neural network to train the model." (HiFi Sequencing Video, 8:04–9:23.)<br>  <br><br>*See, e.g.,* HiFi Sequencing Video, 8:30.<br>*See also* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001578, 1580. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| [1.d] (d) determining, using the model, whether the modification is present in a nucleotide at the target position within the window in the input data structure. | The Accused Products and/or Systems determine, using the model, whether the modification (e.g., 5mC methylation) is present in a nucleotide at the target position within the window in the input data structure (e.g., each CpG site within a HiFi read).<br><br>For example, the Accused Products and/or Systems can detect 5mC methylation in a DNA strand.<br><br>**EPIGENETIC ANALYSIS WITH 5-BASE HIFI SEQUENCING**<br><br>PacBio HiFi sequencing simultaneously calls the four DNA bases and 5mC from untreated genomic DNA. Achieve genome-wide detection and phasing of genetic and epigenetic variants from a single, standard HiFi library prep with long and accurate reads.<br><br>With the power of long-read sequencing, you can achieve:<br><br>Epigenetics in every run — no bisulfite treatment required<br><br>Unlike methods that require chemical conversion of DNA, HiFi sequencing detects modifications in native DNA through impacts on the kinetics of base incorporation.<br><br>High accuracy of sequence and methylation<br><br>Methylation detection with HiFi sequencing is highly concordant to bisulfite sequencing.<br><br>Access the full genome<br><br>Access difficult regions of the genome like repeats and centromeres that are beyond the reach of short-read sequencing.<br><br>Phasing<br><br>Identify allele-specific methylation, whether due to parental imprinting, genetic variation, or repeat expansions.<br><br>*See, e.g.,* PacBio Epigenetics Webpage (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

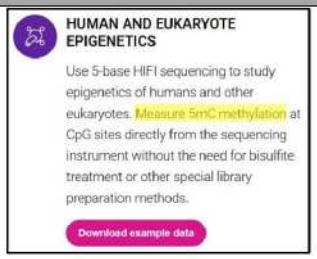
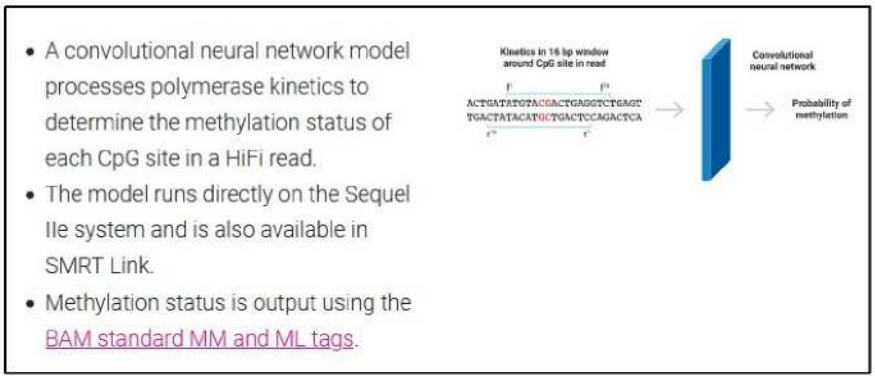
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* PacBio Epigenetics Webpage (annotated).<br><br>The Accused Products and/or Systems can detect 5mC methylation in a DNA strand by inputting each feature vector, which is a smaller data structure built from the BAM file, into to a convolutional neural network (CNN) as demonstrated below:<br>  <br>*See, e.g.,* PacBio Epigenetics Webpage. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

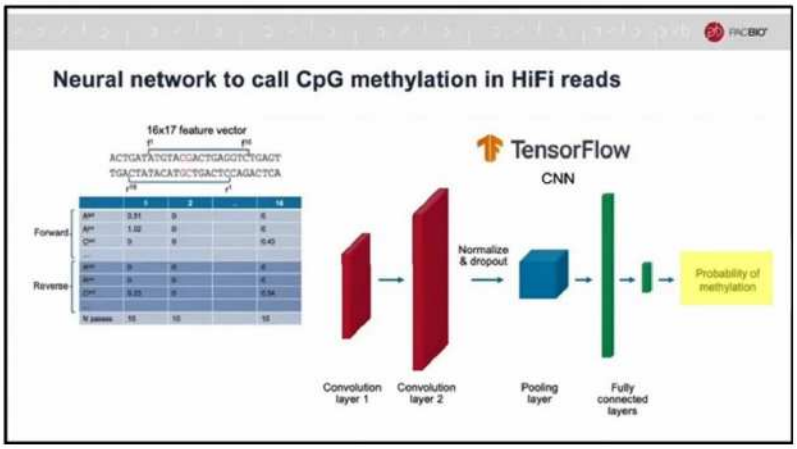
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Sequencing Video, 7:08 (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

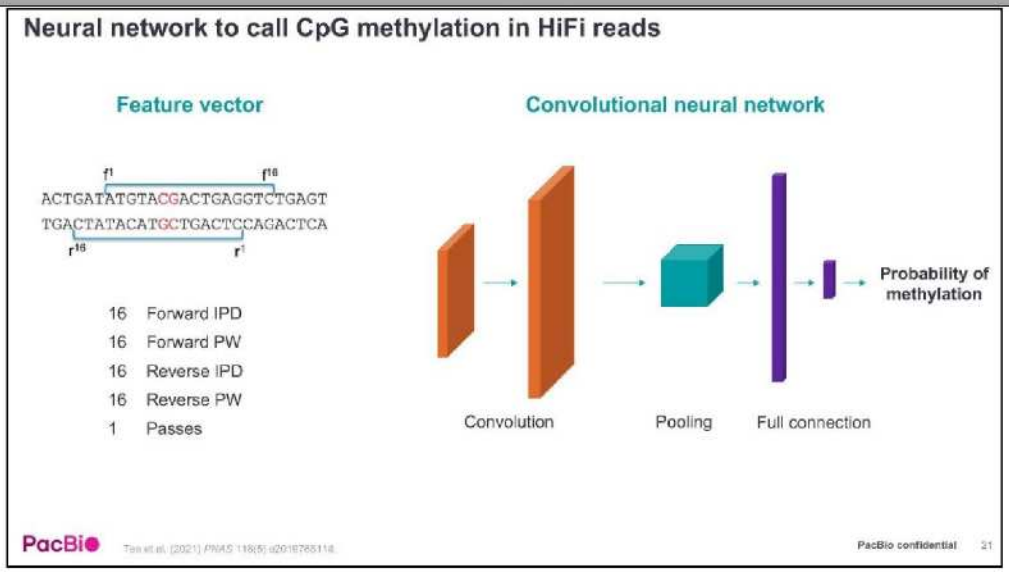
| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 1 | Pacific Biosciences of California, Inc. |
|---|---|
| | **primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>**Algorithm**<br>The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M. SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.<br><br>**Usage**<br>`primrose [options] <HiFi INPUT> <HiFi OUTPUT>`<br><br>Required: `HiFi INPUT` — Description: Input BAM file or ConsensusReadSet XML file.<br>Required: `HiFi OUPUT` — Description: Output BAM file or ConsensusReadSet XML file.<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001578, 1580. |

**B. [Claim 2]**

| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| [2.pre] The method of claim 1, wherein: | [*See* Claim 1 above] |
| [2.a] the input data structure is one input data structure of a plurality of input data structures,<br><br>[2.b] the sample nucleic acid molecule is one sample nucleic acid molecule of a plurality of sample nucleic acid molecules,<br><br>[2.c] the plurality of sample nucleic acid molecules are obtained from a biological sample of a subject, and<br><br>[2.d] each input data structure corresponds to a respective window of nucleotides sequenced in a respective sample nucleic acid molecule of the plurality of sample nucleic acid molecules, and | The Accused Products and/or Systems are capable of processing multiple sample nucleic acid molecules (i.e., a plurality) obtained from a biological sample of a subject as recited in [2.c]. When the plurality of sample nucleic acid molecules are processed, each sample nucleic acid molecule includes at least one window of nucleotides that are sequenced and corresponds to a data structure as recited in [2.b] and [2.d]. Consequently, there are a plurality of input data structures as recited in [2.a], each of which correspond to a respective window of nucleotides sequenced in a respective sample nucleic acid molecule of the plurality of sample nucleic acid molecules as recited in [2.d].<br><br>For example, in accordance with [2.c], PacBio teaches its end users that Single Molecule Real-Time (SMRT) sequencing technology involves sequencing long fragments of native DNA (i.e., a plurality of sample nucleic acid molecules) that is extracted from a biological sample to produce HiFi reads:<br><br>**Introduction**<br><br>Single Molecule, Real-Time (SMRT®) sequencing uses the natural process of DNA replication to sequence long fragments of native DNA in order to produce highly accurate long reads, or HiFi reads. As such, starting with high-quality, high molecular weight (HMW) genomic DNA (gDNA) will result in longer libraries and better performance during sequencing. This technical note is intended to give recommendations, tips and tricks for the extraction of DNA, as well as assessing and preserving the quality and size of your DNA sample to be used for HiFi sequencing.<br><br>*See, e.g.,* Preparing DNA for PacBio HiFi Sequencing at 1. |

| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | The extracted DNA thus includes a plurality of individual fragments (i.e., a plurality of sample nucleic acid molecules).<br><br>PacBio further instructs the use of its Nanobind® DNA extraction kits[7] as "an excellent way to produce HiFi reads" by obtaining the plurality of sample nucleic acid molecules from a variety of biological samples—including blood, cells, and tissues—that are used in combination with the other Accused Products (SMRTbell® prep kits, Sequel® II System, Sequel® IIe System, Revio™ System) in accordance with the methods recited in the Asserted Claims:<br><br>**DNA extraction**<br><br>**Nanobind®** DNA extraction kits are an excellent way to produce HiFi reads on the PacBio systems. The following is a list of kits that our customers have used in combination with the SMRTbell® prep kit 3.0 protocol for high yields and consistent sequencing results. DNA yields will vary by sample quality, type, and input amount. See our **technical note on human blood sequencing** for detailed information.<br><br>**Kit** / **Sample amount** / **Typical yields**<br>**Nanobind CBB kit** PN 102-301-900 / • Blood: 200 µL • Cells: 1 x $10^6$ • Bacteria: 5 x $10^8$–5 x $10^9$ / • 5–34 µg depending on sample type and input amount<br>**Nanobind tissue kit** PN 102-302-100 / • Depends on tissue and preservation method; typically around 25 mg / • 5–100 µg<br>**Nanobind plant nuclei kit** PN 102-302-000 / • Up to 10 g / • 5–20 µg<br><br>*See, e.g.,* Preparing DNA for PacBio HiFi Sequencing at 2. |

[7] PacBio Nanobind® DNA extraction kits that can used with the asserted method claims include PN: 102-301-900, PN: 102-302-000, PN: 102-302-100, PN: 102-762-800, PN: 102-762-700.

After preparing a library (or plurality) of sample nucleic acid molecules from the extracted DNA (i.e., a SMRTbell® library), each sample nucleic acid molecule that is part of the plurality of sample nucleic acid molecules is immobilized in a tiny well known as a "zero-mode waveguide" or "ZMW" contained in a PacBio "SMRT® Cell." Each SMRT® Cell contains millions of ZMWs such that millions of individual sample nucleic acid molecules can be sequenced and analyzed in accordance with at least [2.b]-[2.d].

Step 3: sequencing

Once your library is prepared, the PacBio sequencing system, such as the Sequel II system, takes over. At the heart of SMRT sequencing is the SMRT Cell, which contains millions of tiny wells called zero-mode waveguides (ZMWs). Single molecules of DNA are immobilized in these wells, and as the polymerase incorporates each nucleotide, light is emitted, and nucleotide incorporation is measured in real time. The reactions are recorded in a format that can then be analyzed using on-instrument tools, as well as additional software.

*See, e.g.,* Sequencing 101.

Thus, sequencing is performed on a plurality of DNA molecules in the SMRT® sequencing process, at least because the process is performed in each of a large number of zero-mode waveguides (ZMW), each zero-mode waveguide having a DNA molecule (as an insert within a SMRTbell® template molecule) from the subject.

PacBio SMRT sequencing operates within a silicon chip (a **SMRTcell**) fabricated to contain a large number of microscopic holes (**ZMWs**, or **zero-mode waveguides**), each assigned a **hole number**.

Within a ZMW, PacBio SMRT sequencing is performed on a circularized molecule called a **SMRTbell**. The SMRTbell, depicted below, consists of:

- the customer's double-stranded DNA **insert** (with sequence $I$, read following the arrow)
- (optional) double-stranded DNA **barcodes** (sequences $B_L, B_R$) used for multiplexing DNA samples. While the barcodes are optional, they must be present at both ends if present at all. Barcodes may or may not be *symmetric*, where symmetric means $B_L = B_R^{RC}$.
- SMRTbell **adapters** (sequences $A_L, A_R$), each consisting of a double stranded stem and a single-stranded hairpin loop. Adapters may or may not be *symmetric*, where symmetric means $A_L = A_R$.

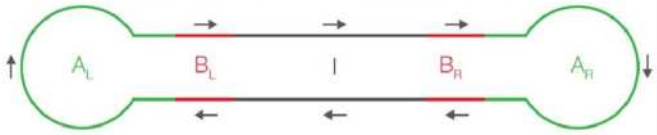



Fig. 1: A schematic drawing of a SMRTbell

*See, e.g.,* PacBioFileFormats Documentation, p. 7 (annotated).

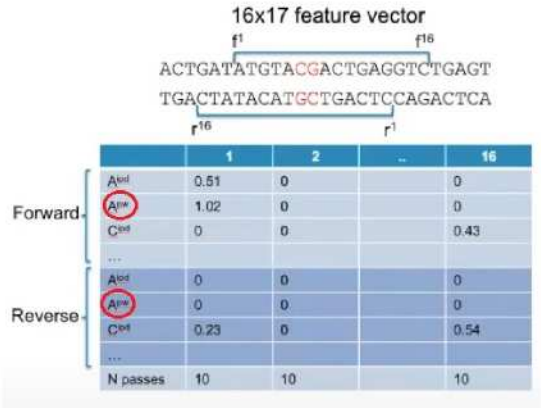
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | The SMRTbell® template produces HiFi reads which include a feature vector, i.e., a window, of nucleotides comprising data related to the window of nucleotides on both forward and reverse strands of the double-stranded DNA insert that is used as the input data structure:<br>  <br><br>*See, e.g.,* HiFi Sequencing Video, 7:25-7:47 (annotated).<br>As such, each respective input data structure corresponds to a respective window of nucleotides sequenced in a respective sample nucleic acid molecule (i.e., a HiFi read) of the plurality of sample nucleic acid molecules sequenced in accordance with PacBio's SMRT® sequencing technology.<br><br>*See also* Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002. |
| [2.e] the method further comprising: | |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

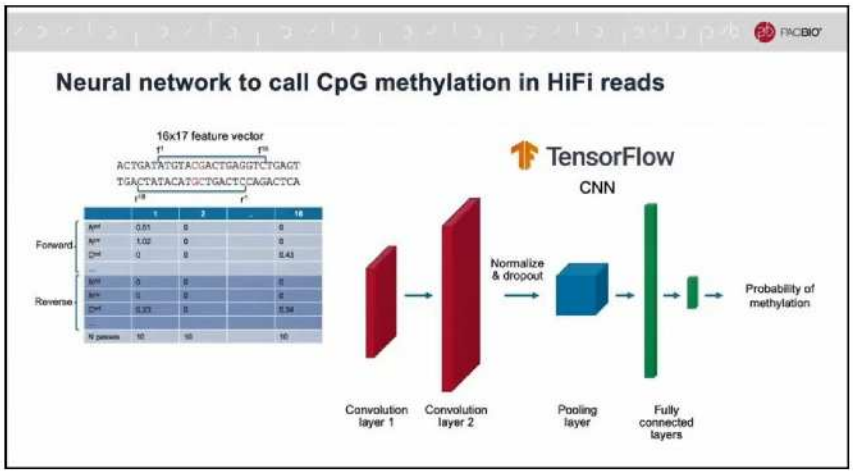
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| [2.e.1] receiving the plurality of input data structures, | The Accused Products and/or Systems receive the plurality of input data structures.<br><br>For example, the Accused Products and/or Systems receive the HiFi sequencing data from the input BAM file. *See* claim element [1.c].<br><br>  <br><br>*See, e.g.,* HiFi Sequencing Video, 7:08. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

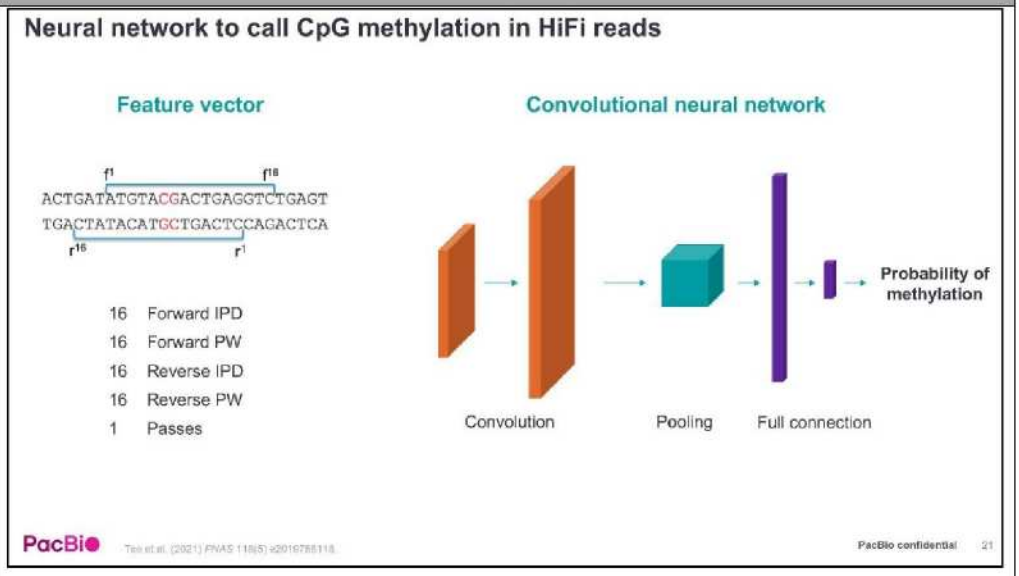
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | **primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>**Algorithm**<br>The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M. SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.<br><br>**Usage**<br>`primrose [options] <HiFi INPUT> <HiFi OUTPUT>`<br><br>Required: `HiFi INPUT` — Description: Input BAM file or ConsensusReadSet XML file.<br>Required: `HiFi OUPUT` — Description: Output BAM file or ConsensusReadSet XML file.<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

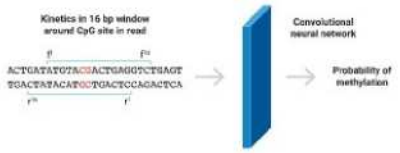
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | • A convolutional neural network model processes polymerase kinetics to determine the methylation status of each CpG site in a HiFi read.<br>• The model runs directly on the Sequel IIe system and is also available in SMRT Link.<br>• Methylation status is output using the BAM standard MM and ML tags.<br>  <br><br>*See, e.g.,* PacBio Epigenetics Webpage. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

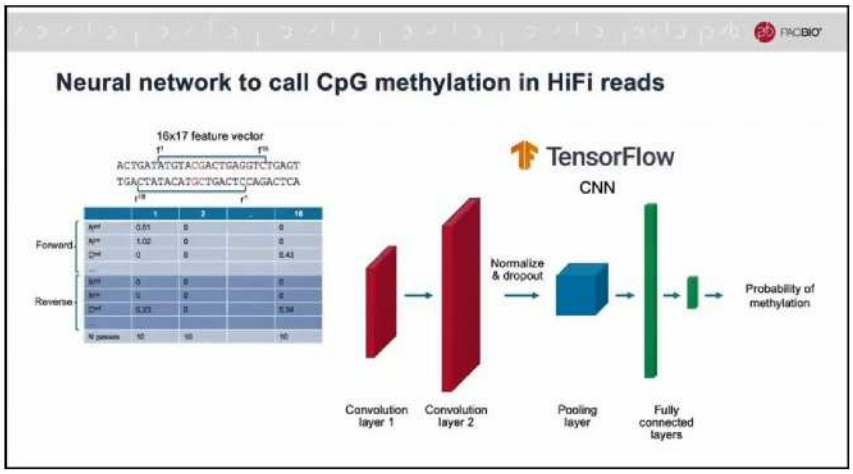
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| [2.e.2] inputting the plurality of input data structures into the model, and | The Accused Products and/or Systems input the plurality of input data structures into the model.<br><br>For example, the Accused Products and/or Systems receive the HiFi sequencing data from the input BAM file and input it to the convolutional neural network (CNN). *See* claim element [1.c].<br><br>  <br><br><br>*See, e.g.,* HiFi Sequencing Video, 7:08. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

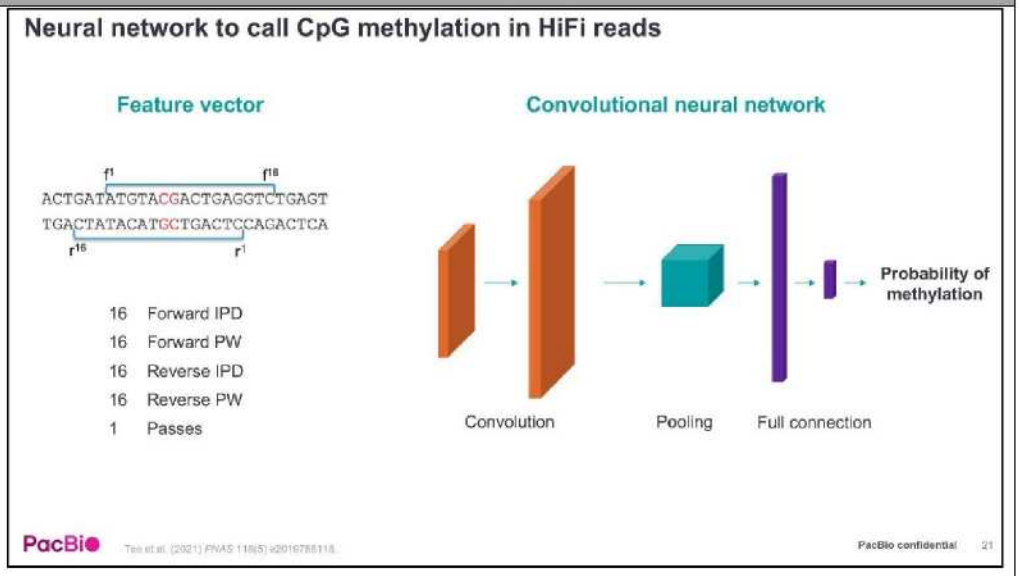
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | **primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>**Algorithm**<br>The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M. SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.<br><br>**Usage**<br>`primrose [options] <HiFi INPUT> <HiFi OUTPUT>`<br><br>Required: `HiFi INPUT` — Description: Input BAM file or ConsensusReadSet XML file.<br>Required: `HiFi OUPUT` — Description: Output BAM file or ConsensusReadSet XML file.<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

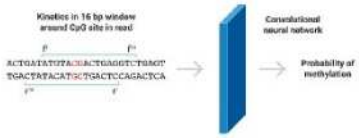
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | • A convolutional neural network model processes polymerase kinetics to determine the methylation status of each CpG site in a HiFi read.<br>• The model runs directly on the Sequel IIe system and is also available in SMRT Link.<br>• Methylation status is output using the BAM standard MM and ML tags.<br>  <br><br>*See* PacBio Epigenetics Webpage. |
| [2.e.3] determining, using the model, whether a modification is present in a nucleotide at a target location in the respective window of each input data structure. | The Accused Products and/or Systems determine, using the model, whether a modification is present in a nucleotide at a target location in the respective window of each input data structure.<br>For example, the Accused Products and/or Systems can detect 5mC methylation in a DNA strand, including at CpG sites in the windows corresponding to the sequenced DNA molecules. *See* claim element [1.d].<br>**HUMAN AND EUKARYOTE EPIGENETICS**<br>Use 5-base HiFi sequencing to study epigenetics of humans and other eukaryotes. Measure 5mC methylation at CpG sites directly from the sequencing instrument without the need for bisulfite treatment or other special library preparation methods.<br>Download example data<br>*See, e.g.,* PacBio Epigenetics Webpage (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | **EPIGENETIC ANALYSIS WITH 5-BASE HIFI SEQUENCING**<br>PacBio HiFi sequencing simultaneously calls the four DNA bases and 5mC from untreated genomic DNA. Achieve genome-wide detection and phasing of genetic and epigenetic variants from a single, standard HiFi library prep with long and accurate reads.<br>With the power of long-read sequencing, you can achieve:<br>Epigenetics in every run — no bisulfite treatment required<br>Unlike methods that require chemical conversion of DNA, HiFi sequencing detects modifications in native DNA through impacts on the kinetics of base incorporation.<br>High accuracy of sequence and methylation<br>Methylation detection with HiFi sequencing is highly concordant to bisulfite sequencing.<br>Access the full genome<br>Access difficult regions of the genome like repeats and centromeres that are beyond the reach of short-read sequencing.<br>Phasing<br>Identify allele-specific methylation, whether due to parental imprinting, genetic variation, or repeat expansions.<br>*See, e.g.,* PacBio Epigenetics Webpage (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

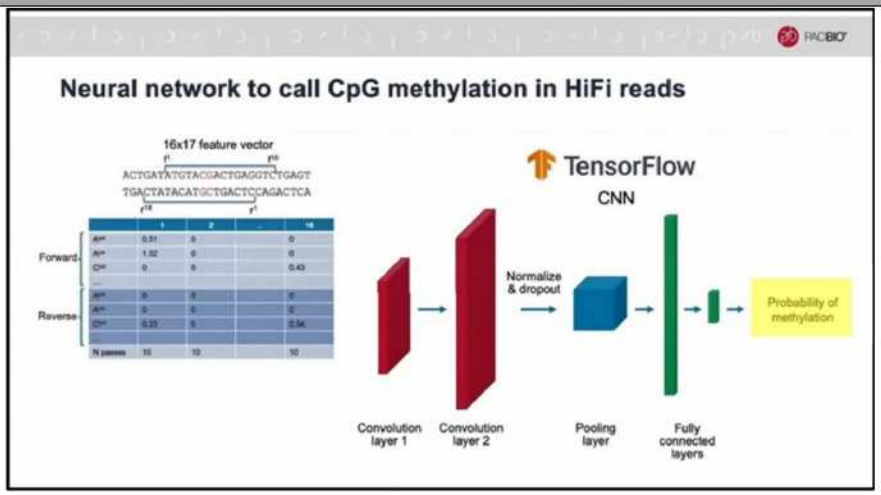
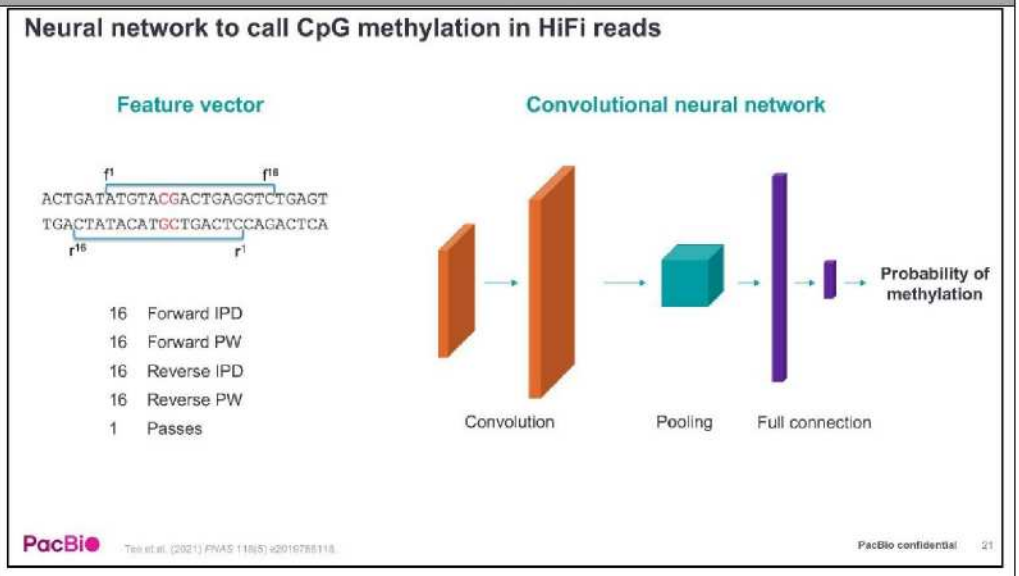
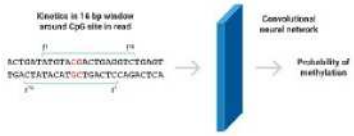
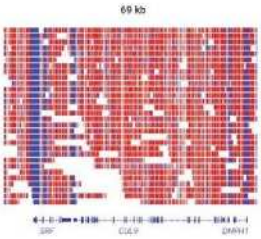
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Sequencing Video, 7:08 (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

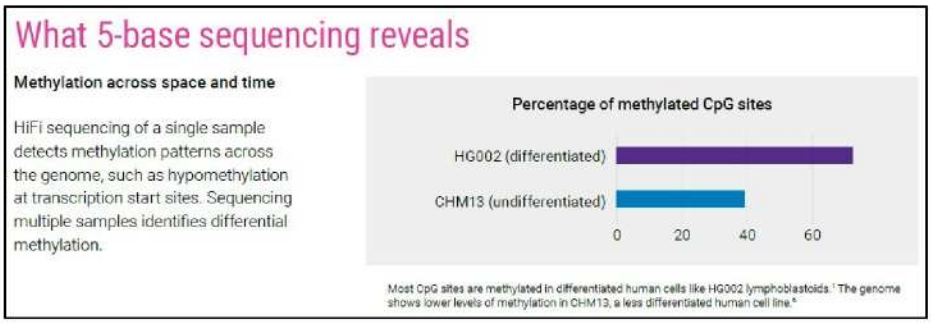
| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493. |

| U.S. Patent No. 11,091,794 CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | **primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>**Algorithm**<br>The CNN is trained using invitro modified controls of methylated and unmethylated human DNA. The unmethylated control comprises a human shotgun library that has undergone whole genome amplification (WGA), a process that includes PCR, and therefore removes all modifications. The methylated control is generated by subjecting a human WGS library to invitro methylation using a CpG methyltransferase enzyme (`M. SssI`). Kinetic data, pulse width and inter-pulse distance, over a 16 base pair window for both the forward and reverse strand, is used as input to the CNN. The output of the CNN is a probability scale measure of whether the CpG is symmetrically 5mC-modified.<br><br>**Usage**<br>`primrose [options] <HiFi INPUT> <HiFi OUTPUT>`<br><br>Required: `HiFi INPUT` — Description: Input BAM file or ConsensusReadSet XML file.<br>Required: `HiFi OUPUT` — Description: Output BAM file or ConsensusReadSet XML file.<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. |

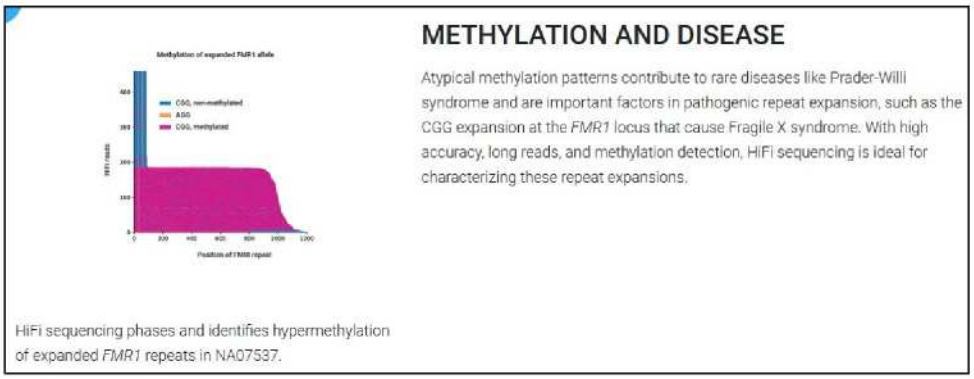
**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794<br>CLAIM 2 | Pacific Biosciences of California, Inc. |
|---|---|
| | • A convolutional neural network model processes polymerase kinetics to determine the methylation status of each CpG site in a HiFi read.<br>• The model runs directly on the Sequel IIe system and is also available in SMRT Link.<br>• Methylation status is output using the BAM standard MM and ML tags.<br>  <br><br>*See, e.g.,* PacBio Epigenetics Webpage.<br>The determination of whether a modification is present in a nucleotide at a target location in the respective window of each input data structure allows determination of regional methylation patterns across a genomic region:<br>  <br><br>**REGIONAL METHYLATION PATTERNS**<br>Methylation levels vary across the genome in many species. In vertebrates like human most CpG sites are methylated. Active gene transcription start sites are often hypomethylated.<br>In this example genomic region, 5-base HiFi sequencing of the human HG002 sample shows overall hypermethylation (red) with hypomethylation (blue) specifically at transcription start sites.<br>*See, e.g.,* PacBio Epigenetics Webpage.<br>*See also* Customer Overview of System and Software Release v11.0 at PBIO_00000028–33; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000067–77. |

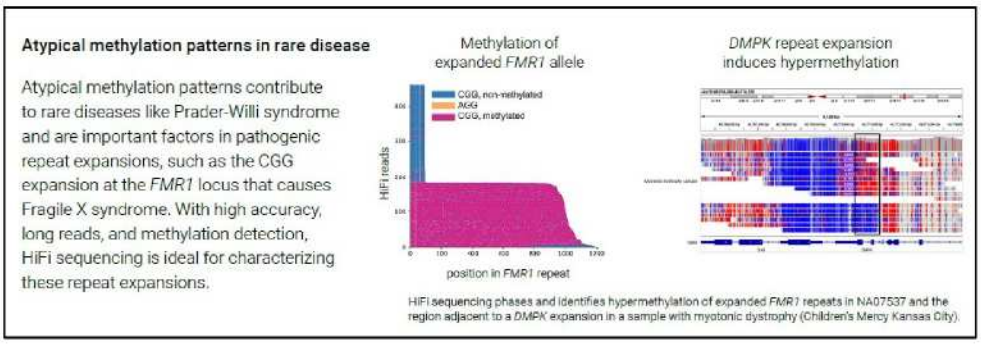
**C. [Claim 3]**

| U.S. Patent No. 11,091,794 CLAIM 3 | Pacific Biosciences of California, Inc. |
|---|---|
| [3.pre] The method of claim 2, further comprising: | [*See* Claim 3 above] |
| [3.a] determining the modification is present at one or more nucleotides, and | The Accused Products and/or Systems determine whether the modification is present at one or more nucleotides. *See* claim element [1.d], [2.e.3]. |
| [3.b] determining a classification of a disorder using the presence of the modification at one or more nucleotides. | The Accused Products and/or Systems determine a classification of a disorder using the presence of the modification at one or more nucleotides.<br>For example, methylation patterns like hypomethylation and/or hypermethylation"<br>  <br><br>*See, e.g.,* PacBio Measuring DNA Methylation Application Brief at PBIO_00000011. |

PacBio uses results of such methylation patterns to classify/characterize diseases such as Prader-Willi syndrome and Fragile X syndrome as shown below:




*See, e.g.,* PacBio Epigenetics Webpage.




*See, e.g.,* PacBio Measuring DNA Methylation Application Brief at PBIO_00000011.

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 3 | Pacific Biosciences of California, Inc. |
|---|---|
| | *See also* PacBio and Children's Mercy Kansas City Press Release. ("PacBio (NASDAQ: PACB), a leading provider of high-quality, highly accurate sequencing solutions, today announced an expanded research collaboration with Children's Mercy Kansas City, one of the nation's top pediatric medical systems, to use the multi-omics capabilities of PacBio's Sequel® IIe system in the study of genetic disease. The research will apply direct methylation detection and Iso-Seq™ full-length RNA sequencing in their study. A better understanding of the genetic underpinnings of rare disease may ultimately help drive diagnostic yield and potentially enhance care for patients. Children's Mercy is a leader in using PacBio's genomic sequencing technology, approaching 1,000 samples sequenced since 2020."). <br><br> *See also* Customer Overview of System and Software Release v11.0 at PBIO_00000028–33; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000067–77. |

**D. [Claim 7]**

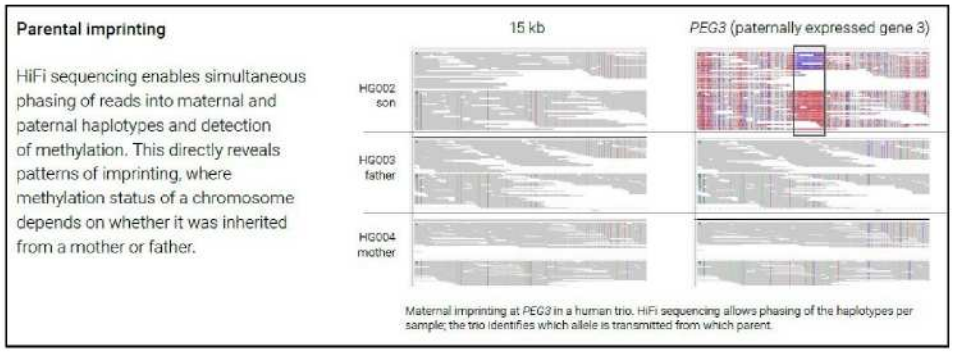
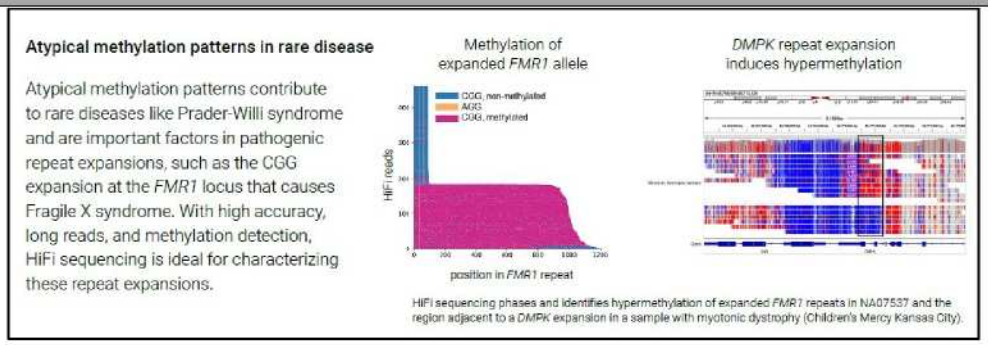
| U.S. Patent No. 11,091,794 CLAIM 7 | Pacific Biosciences of California, Inc. |
|---|---|
| [7.pre] The method of claim 2, | [*See* Claim 2 above] |
| [7.a] wherein the modification is a methylation, the method further comprising: | *See* claim element [1.pre], [1.d]. |
| [7.b] determining the modification is present at one or more nucleotides, and | *See* claim element [3.a]. |
| [7.c] determining a clinically-relevant DNA fraction, a fetal methylation profile, a maternal methylation profile, a presence of an imprinting gene region, or a tissue of origin using the presence of the modification at one or more nucleotides. | The Accused Products and/or Systems determine a clinically-relevant DNA fraction, a fetal methylation profile, a maternal methylation profile, a presence of an imprinting gene region, or a tissue of origin using the presence of the modification at one or more nucleotides as shown below:   *See, e.g.*, PacBio Measuring DNA Methylation Application Brief at PBIO_00000011. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 7 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.*, PacBio Measuring DNA Methylation Application Brief at PBIO_00000011.<br><br>*See also* Customer Overview of System and Software Release v11.0 at PBIO_00000028–33; DNA Methylation and 5-base HiFi Sequencing Slides at PBIO_00000067–77. |

**E. [Claim 8]**

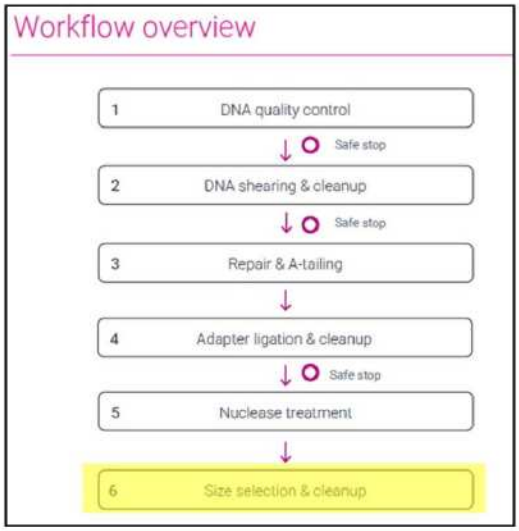
| U.S. Patent No. 11,091,794 CLAIM 8 | Pacific Biosciences of California, Inc. |
|---|---|
| [8.pre] The method of claim 2, | [*See* Claim 2 above] |
| [8.a] wherein each sample nucleic acid molecule of the plurality of sample nucleic acid molecules has a size greater than a cutoff size. | In the Accused Products and/or Systems, each sample nucleic acid molecule of the plurality of sample nucleic acid molecules has a size greater than a cutoff size.<br><br>For example, PacBio gives its users a workflow overview for producing a SMRTbell® library that includes a step of size selection:<br>  <br><br>*See, e.g.,* Preparing Whole Genome and Metagenome Libraries Using SMRTbell® Prep Kit 3.0 at 2 (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 8 | Pacific Biosciences of California, Inc. |
|---|---|
| | Further, Pac Bio instructs that a size cutoff be used in preparing SMRTbell® libraries using SMRTbell® prep kit 3.0 with PacBio's size-selection beads (i.e., AMPure® PB Beads) or other alternative size selection methods as shown below in PacBio's Technical note titles "Alternative size selection methods for SMRTbell® prep kit 3.0" (available at https://www.pacb.com/wp-content/uploads/Technical-note-Alternative-size-selection-methods-for-SMRTbell-prep-kit-3.0.pdf):<br><br>In *Procedure & checklist — Preparing whole genome and metagenome sequencing libraries using SMRTbell prep kit 3.0*, size-selection beads are used to remove short DNA fragments and enrich for the long fragments that are most valuable for downstream analysis. Although size-selection beads have many advantages, alternative methods offer more flexibility in defining a size cutoff. This technical note provides guidance for size selection of WGS libraries prepared with the SMRTbell prep kit 3.0.<br><br>**Size-selection beads (Standard)** / **Gel cassettes** / **Short read eliminator**<br>**Size cutoff**: Fixed (>5 kb) / ✓ Adjustable / Fixed (> 10 kb)<br>**Hands-on time**: ✓ Low / ✓ Low / ✓ Low<br>**Run time**: ✓ Minutes / Hours / ✓ Minutes<br>**Automation**: ✓ Standard liquid handler / Separate instrument / n/a<br>**Cost**: ✓ Low / High / ✓ Low<br>**DNA recovery**: ✓ High / Low / Low<br>**DNA input to library prep per SMRT® Cell**: ✓ 1 µg / 1.5 µg / 1.5 µg<br><br>*See also* SMRTbell library preparation overview - Revio system at PBIO_000001411–12; SMRTbell library preparation overview - Sequel® IIe system at 1447–48; |

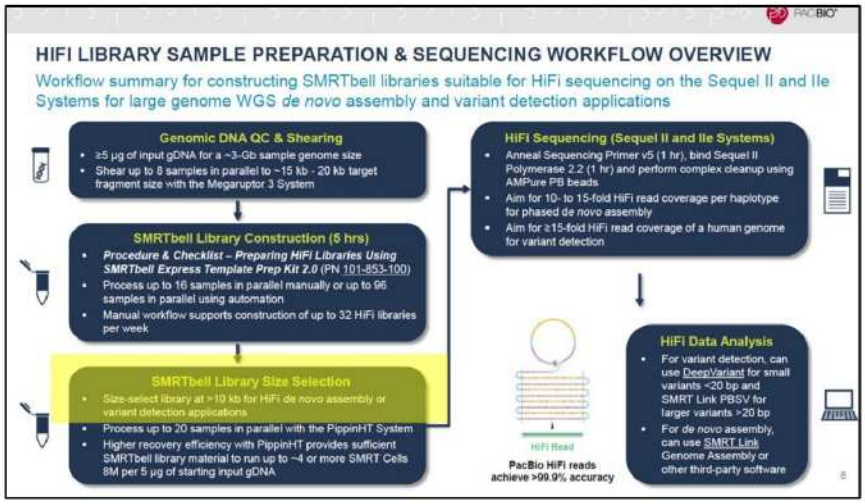
| | |
|---|---|
| | In another example, PacBio instructs in its Customer Training that a cutoff size (e.g., > 10 kb) be used to perform size selection of DNA molecules for constructing SMRTbell® libraries suitable for HiFi sequencing as follows:   *See, e.g.,* HiFi Library Preparation Customer Training at 8 (annotated). |

**EXHIBIT A**

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

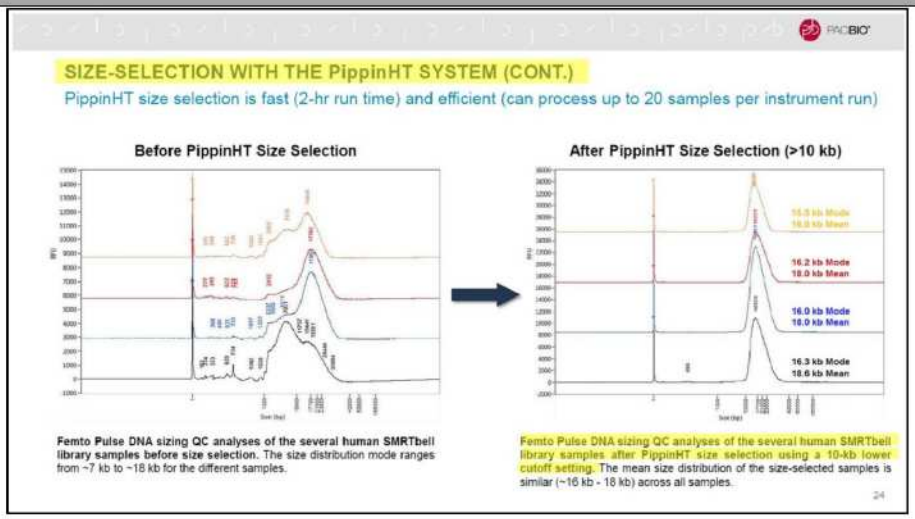
| U.S. Patent No. 11,091,794<br>CLAIM 8 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Library Preparation Customer Training at 24 (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

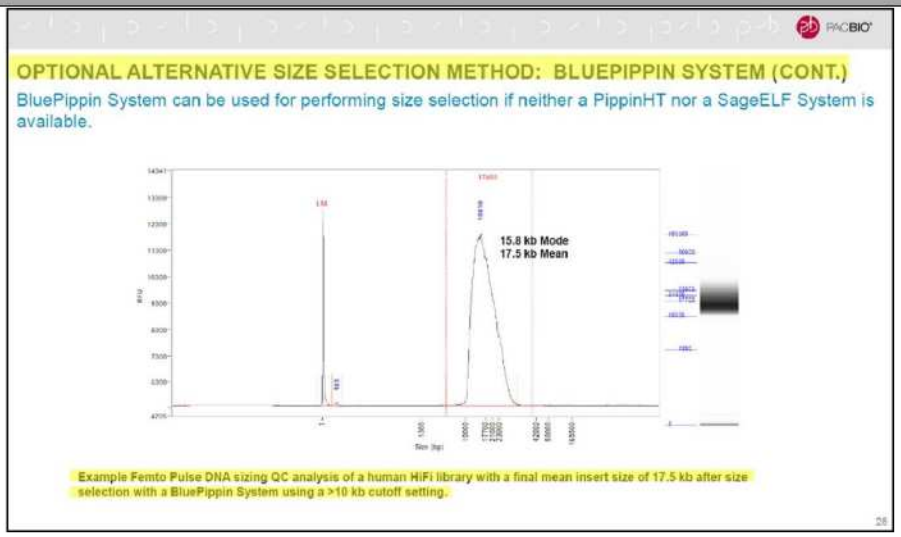
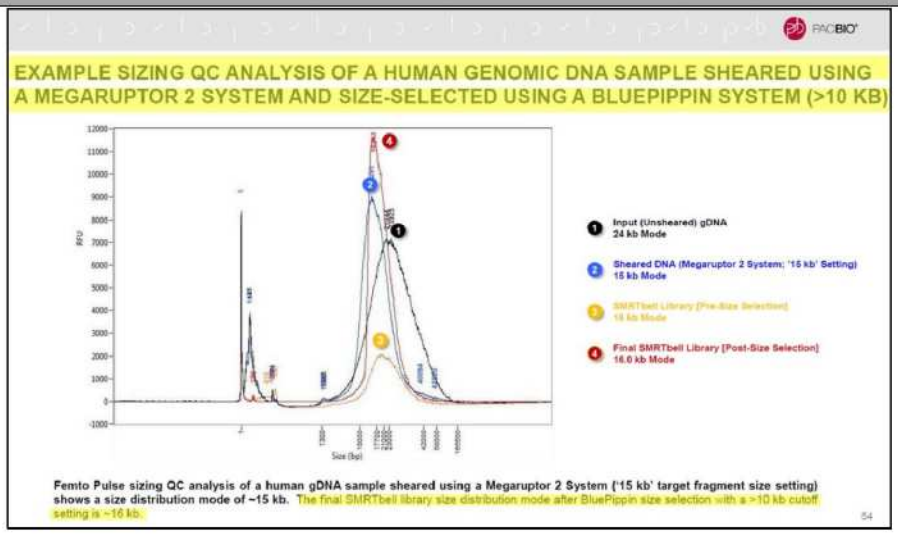
| U.S. Patent No. 11,091,794 CLAIM 8 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Library Preparation Customer Training at 28 (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 8 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Library Preparation Customer Training at 54 (annotated).<br>*See also* HiFi Library Preparation Customer Training at 57.<br>In yet another example, PacBio provides instructions to deplete short DNA fragments below a particular size (e.g., <5 kb, <10kb, <40 kb) prior to shearing and library preparation, leaving sample nucleic molecules with a size greater than the cutoff size using its short read eliminator kits.[8] See, e.g., Short Read Eliminator (SRE) XS and SL kits, Guide & overview (available at https://www.pacb.com/wp-content/uploads/Guide-overview-Short-Read-Eliminator-SRE-XS-and-XL-kits.pdf). |

[8] PacBio Short Read Eliminator Kits that can used with the asserted method claims include PN: 102-208-200, PN: 102-208-300, PN: 102-208-300.

**F. [Claim 10]**

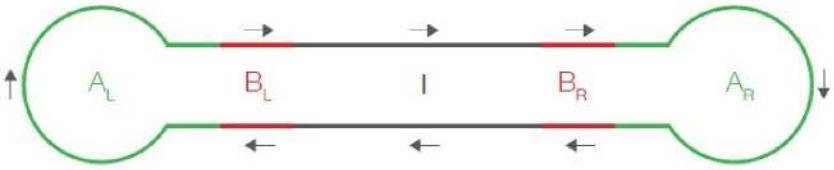
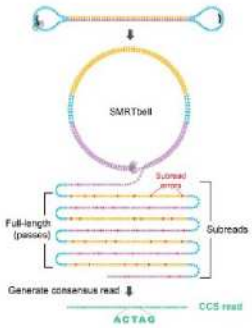
| U.S. Patent No. 11,091,794 CLAIM 10 | Pacific Biosciences of California, Inc. |
|---|---|
| [10.pre] The method of claim 1, | [*See* Claim 1 above] |
| [10.a] further comprising sequencing the sample nucleic acid molecule. | The Accused Products and/or Systems sequence the sample nucleic acid molecule.<br><br>For example, as shown in PacBio's PacBioFileFormats Documentation, "PacBio SMRT® sequencing is performed on a circularized molecule called a SMRTbell" that is generated using PacBio SMRTbell® prep kits. The SMRTbell® includes "the customer's double-stranded DNA insert" (sequence *I*), i.e., the sample nucleic acid molecule:<br><br>PacBio SMRT sequencing operates within a silicon chip (a **SMRTcell**) fabricated to contain a large number of microscopic holes (**ZMWs**, or **zero-mode waveguides**), each assigned a **hole number**.<br><br>Within a ZMW, PacBio SMRT sequencing is performed on a circularized molecule called a **SMRTbell**. The SMRTbell, depicted below, consists of:<br><br>• the customer's double-stranded DNA **insert** (with sequence $I$, read following the arrow)<br>• (optional) double-stranded DNA **barcodes** (sequences $B_L, B_R$) used for multiplexing DNA samples. While the barcodes are optional, they must be present at both ends if present at all. Barcodes may or may not be *symmetric*, where symmetric means $B_L = B_R^{RC}$.<br>• SMRTbell **adapters** (sequences $A_L, A_R$), each consisting of a double stranded stem and a single-stranded hairpin loop. Adapters may or may not be *symmetric*, where symmetric means $A_L = A_R$.<br><br>  <br><br>Fig. 1: A schematic drawing of a SMRTbell<br><br>*See, e.g.*, PacBioFileFormats Documentation at 7. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 10 | Pacific Biosciences of California, Inc. |
|---|---|
| | The Accused Products and/or Systems then perform HiFi sequencing on the SMRTbell® template using circulate consensus sequencing (CCS).<br><br>**ccs** Circular Consensus Sequencing (CCS) Analysis computes consensus sequences from multiple "passes" around a circularized single DNA molecule (SMRTbell® template). CCS analysis uses the Arrow framework to achieve optimal consensus results given the number of passes available.<br><br>  <br><br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001130.<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001012; SMRT® Tools Reference Guide, v12.0 at PBIO_00001264; Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002. |

**G. [Claim 11]**

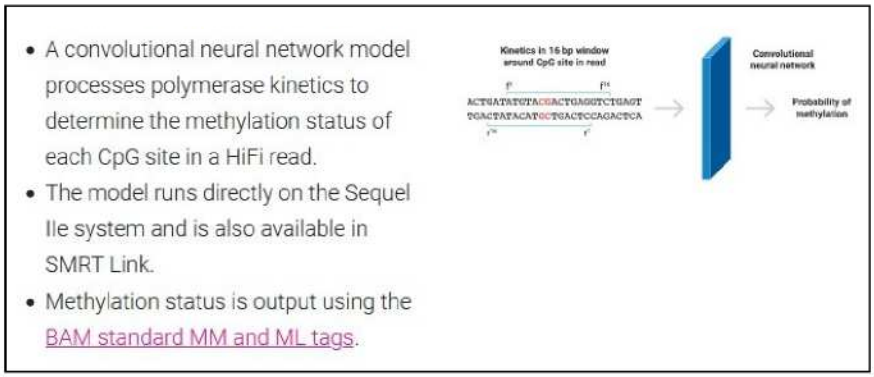
| U.S. Patent No. 11,091,794 CLAIM 11 | Pacific Biosciences of California, Inc. |
|---|---|
| [11.pre] The method of claim 1, | [*See* Claim 1 above] |
| [11.a] wherein the model includes a machine learning model, a principal component analysis, a convolutional neural network, or a logistic regression. | The Accused Products and/or Systems use a convolutional neural network for the model:<br>  <br><br>*See, e.g.,* PacBio Epigenetics Webpage. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

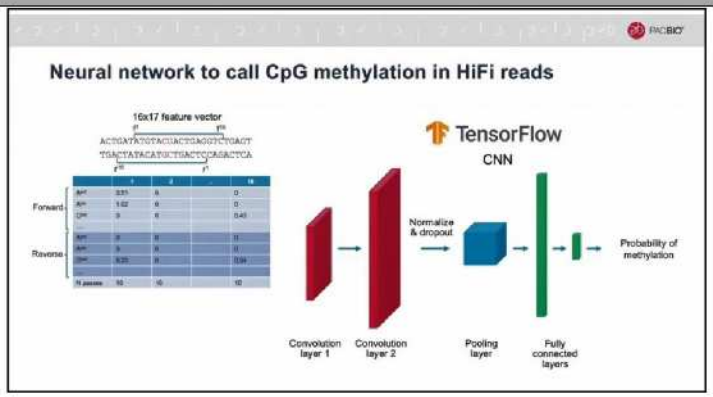
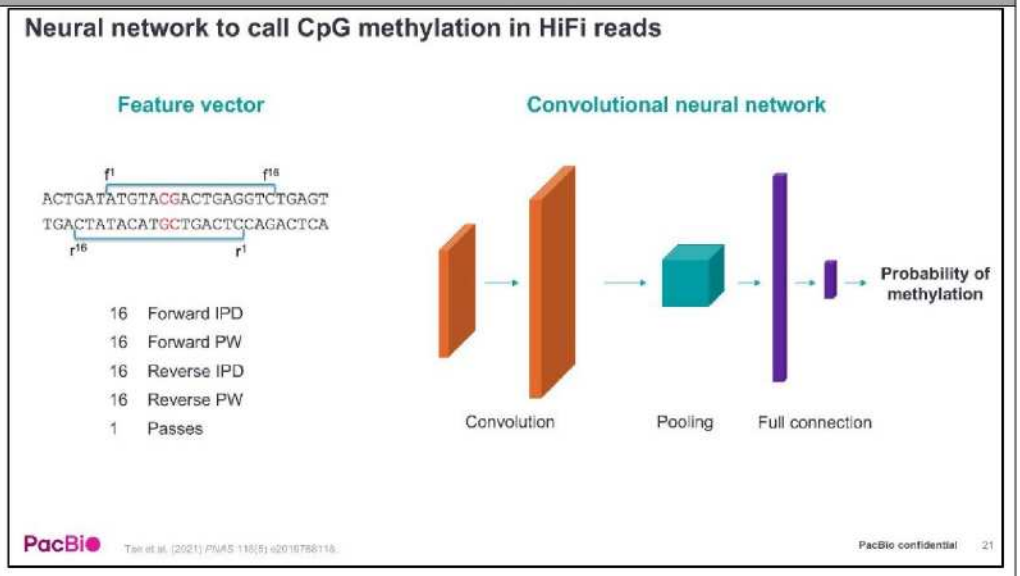
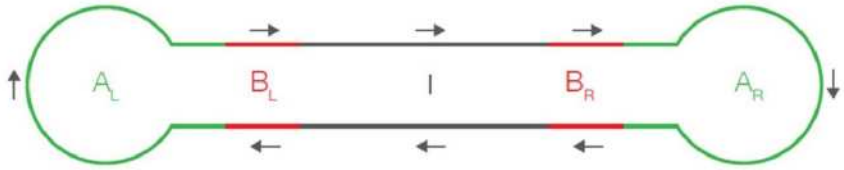
| U.S. Patent No. 11,091,794 CLAIM 11 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* HiFi Sequencing Video, 7:08.<br><br>**primrose** The `primrose` tool analyzes the kinetic signatures of cytosine bases in CpG motifs to identify the presence of 5mC.<br><br>`primrose` uses a convolution neural network (CNN) to predict the methylation state (5mC) of each CpG in a HiFi read. Methylation is assumed to be symmetric between strands with output reported in the forward direction with respect to the HiFi read sequence. The output uses the `Mm` and `Ml` tags as defined in the SAM Format Optional Fields Specification. (See here for details.)<br><br>*See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236 (annotated).<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 11 | Pacific Biosciences of California, Inc. |
|---|---|
| |  <br>*See, e.g.,* Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493.<br><br>*See also* SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371; Customer Overview of System and Software Release v11.0 at PBIO_00000021–22; Release v11.0 Sales Training Part 2_v1.0 at PBIO_00001493; Sales Training - DNA Methylation and 5-base HiFI Sequencing_v1.0 at PBIO_00001578, 1580. |

**H. [Claim 12]**

| U.S. Patent No. 11,091,794 CLAIM 12 | Pacific Biosciences of California, Inc. |
|---|---|
| [12.pre] The method of claim 1, wherein: | [*See* Claim 1 above] |
| [12.a] the window of nucleotides corresponding to the input data structure comprises nucleotides on a first strand of the sample nucleic acid molecule and nucleotides on a second strand of the sample nucleic acid molecule, and | In the Accused Products and/or Systems, the window of nucleotides corresponding to the input data structure comprises nucleotides on a first strand of the sample nucleic acid molecule and nucleotides on a second strand of the sample nucleic acid molecule.<br><br>For example, the Accused Products and/or Systems perform sequencing on double-stranded DNA inserts within a SMRTbell® template:<br><br>PacBio SMRT sequencing operates within a silicon chip (a **SMRTcell**) fabricated to contain a large number of microscopic holes (**ZMWs**, or **zero-mode waveguides**), each assigned a **hole number**.<br><br>Within a ZMW, PacBio SMRT sequencing is performed on a circularized molecule called a **SMRTbell**. The SMRTbell, depicted below, consists of:<br><br>• the customer's double-stranded DNA insert (with sequence $I$, read following the arrow)<br>• (optional) double-stranded DNA **barcodes** (sequences $B_L, B_R$) used for multiplexing DNA samples. While the barcodes are optional, they must be present at both ends if present at all. Barcodes may or may not be *symmetric*, where symmetric means $B_L = B_R^{RC}$.<br>• SMRTbell **adapters** (sequences $A_L, A_R$), each consisting of a double stranded stem and a single-stranded hairpin loop. Adapters may or may not be *symmetric*, where symmetric means $A_L = A_R$.<br><br>  <br><br>Fig. 1: A schematic drawing of a SMRTbell<br><br>*See, e.g.,* PacBioFileFormats Documentation at 7 (annotated). |

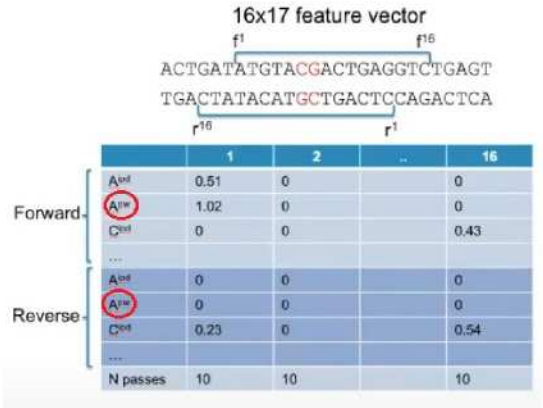
| U.S. Patent No. 11,091,794 CLAIM 12 | Pacific Biosciences of California, Inc. |
|---|---|
| | The SMRTbell® template produces HiFi reads which include a feature vector, i.e., a window, of nucleotides comprising data related to the window of nucleotides on both forward and reverse strands of the double-stranded DNA insert that is used as the input:<br>  <br><br>*See, e.g.,* HiFi Sequencing Video, 7:25-7:47 (annotated).<br><br>*See also* Introduction to SMRT® Sequencing Video at PBIO_00000001; PacBio Sequencing - How It Works Video at PBIO_00000002. |
| [12.b] the input data structure further comprises for each nucleotide within the window a value of a strand property, the strand property indicating the nucleotide being present on | In the Accused Products and/or Systems, the input data structure includes for each nucleotide within the window a value of a strand property, the strand property indicating the nucleotide being present on either the first strand or the second strand.<br><br>For example, as shown above, the feature vector indicates a value (e.g., pulse width, $A^{pw}$, as shown circled in [12.a] above) indicating the nucleotide is present on the forward strand(s). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 12 | Pacific Biosciences of California, Inc. |
|---|---|
| either the first strand or the second strand. | |

# EXHIBIT A

CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

**I. [Claim 13]**

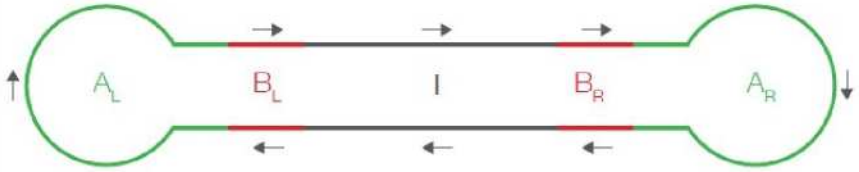
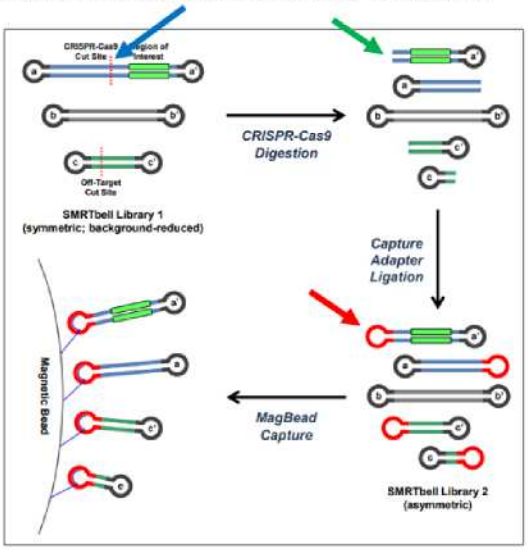
| U.S. Patent No. 11,091,794 CLAIM 13 | Pacific Biosciences of California, Inc. |
|---|---|
| [13.pre] The method of claim 12, | [*See* Claim 12 above] |
| [13.a] wherein the sample nucleic acid molecule is a circular DNA molecule | The sample nucleic acid is a circular DNA molecule. For example, as shown in PacBio's PacBioFileFormats Documentation, "PacBio SMRT® sequencing is performed on a *circularized* molecule called a SMRTbell" (emphasis added) that is generated using PacBio SMRTbell® prep kits. The SMRTbell® includes "the customer's double-stranded DNA insert" (sequence *I*), i.e., the sample nucleic acid molecule:<br><br>PacBio SMRT sequencing operates within a silicon chip (a **SMRTcell**) fabricated to contain a large number of microscopic holes (**ZMWs**, or **zero-mode waveguides**), each assigned a **hole number**.<br><br>Within a ZMW, PacBio SMRT sequencing is performed on a circularized molecule called a **SMRTbell**. The SMRTbell, depicted below, consists of:<br><br>• the customer's double-stranded DNA **insert** (with sequence $I$, read following the arrow)<br>• (optional) double-stranded DNA **barcodes** (sequences $B_L, B_R$) used for multiplexing DNA samples. While the barcodes are optional, they must be present at both ends if present at all. Barcodes may or may not be *symmetric*, where symmetric means $B_L = B_R^{RC}$.<br>• SMRTbell **adapters** (sequences $A_L, A_R$), each consisting of a double stranded stem and a single-stranded hairpin loop. Adapters may or may not be *symmetric*, where symmetric means $A_L = A_R$.<br><br>  <br>Fig. 1: A schematic drawing of a SMRTbell<br><br>*See, e.g.*, PacBioFileFormats Documentation at 7. |

**EXHIBIT A**

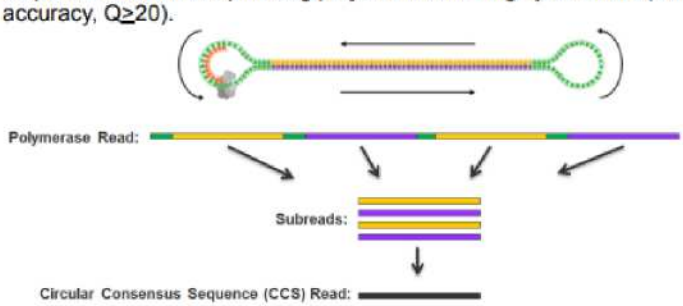
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 13 | Pacific Biosciences of California, Inc. |
|---|---|
| [13.b] formed by: cutting a double-stranded DNA molecule using a Cas9 complex to form a cut double-stranded DNA molecule, and ligating a hairpin adaptor onto an end of the cut double-stranded DNA molecule. | The circular DNA molecule is formed by: cutting a double-stranded DNA molecule using a Cas9 complex to form a cut double-stranded DNA molecule, and ligating a hairpin adaptor onto an end of the cut double-stranded DNA molecule. For example, PacBio describes a method "for targeted SMRT® Sequencing using the CRISPR-Cas9 system that requires no PCR amplification" and "preserves epigenetic modifications." *See* PacBio poster presentation titled "No-Amp Targeted SMRT® Sequencing Using a CRISPR-Cas9 Enrichment Method" (available at https://www.pacb.com/wp-content/uploads/McLaughlin-ASHG-2018-No-Amp-Targeted-SMRT-Sequencing-Using-a-CRISPR-Cas9-Enrichment-Method.pdf) ("PacBio Cas9 Poster"). The method includes cutting a double-stranded DNA molecule using a Cas9 complex (blue arrow) to form a cut double-stranded DNA molecule (green arrow) and ligating a hairpin adaptor onto an end of the cut double-stranded DNA molecule (red arrow).   *See, e.g.,* PacBio Cas9 Poster (Fig. 3) (annotated) |

**J. [Claim 14]**

| U.S. Patent No. 11,091,794 CLAIM 14 | Pacific Biosciences of California, Inc. |
|---|---|
| [14.pre] The method of claim 1, | [*See* Claim 1 above] |
| [14.a] wherein the nucleotides within the window are determined using a circular consensus sequence and without alignment of the sequenced nucleotides to a reference genome. | The Accused Products and/or Systems use *circular* consensus sequence sequencing (CCS) to determine the nucleotides within the window.<br><br>**WHAT ARE HIFI READS?**<br>HiFi reads are produced using circular consensus sequencing (CCS) mode on PacBio long-read systems. HiFi reads provide base-level resolution with 99.9% single-molecule read accuracy.<br>HiFi reads can be used across a wide range of SMRT sequencing applications, from whole genome sequencing for *de novo* assembly, comprehensive variant detection, epigenetic characterization, RNA sequencing and more.<br><br>*See, e.g.,* PacBio HiFi Sequencing Webpage (annotated).<br><br>*See also* SMRT® Link User Guide, v11.0 at PBIO_00000651–652, PBIO_00000665–667; SMRT® Link User Guide, v11.1 at PBIO_00000793–803, PBIO_00000818–820; SMRT® Link User Guide, v12.0 at PBIO_00000959–962, PBIO_00000979–981; SMRT® Tools Reference Guide, v11.0 at PBIO_00001012–1020; SMRT® Tools Reference Guide, v11.1 at PBIO_00001130–1138; SMRT® Tools Reference Guide, v12.0 at PBIO_00001264–1272.<br><br>In addition, as shown in PacBio's PacBioFileFormats Documentation, "PacBio SMRT® sequencing is performed on a *circularized* molecule called a SMRTbell" that is generated using PacBio SMRTbell® prep kits. *See, e.g.,* PacBioFileFormats Documentation at 7:<br><br>PacBio SMRT sequencing operates within a silicon chip (a **SMRTcell**) fabricated to contain a large number of microscopic holes (**ZMWs**, or **zero-mode waveguides**), each assigned a **hole number**.<br>Within a ZMW, PacBio SMRT sequencing is performed on a circularized molecule called a **SMRTbell**. The SMRTbell, depicted below, consists of: |

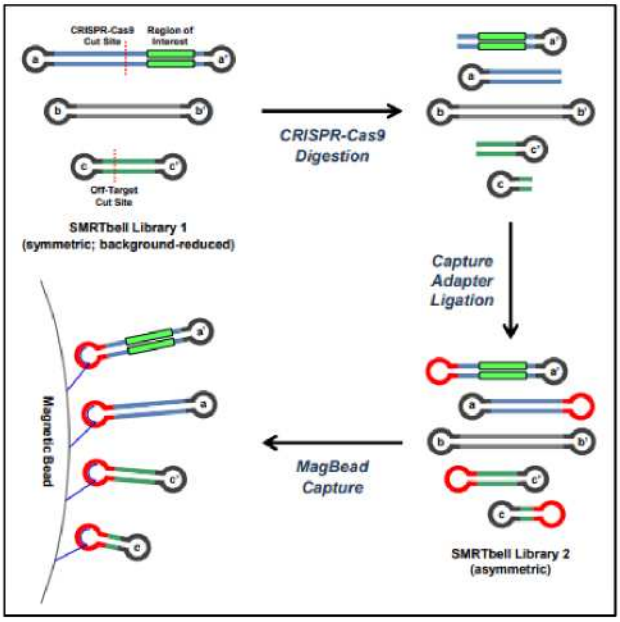
**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 14 | Pacific Biosciences of California, Inc. |
|---|---|
| | The Accused Products and/or Systems use the circular consensus sequence without alignment of the sequenced nucleotides to a reference genome. For example, in PacBio's "Glossary of Terms" provided on its website, "[g]enerating a CCS read *does not include or require alignment* against a reference sequence but does require at least two full-pass subreads from the insert, as shown below:<br><br>• **circular consensus sequencing (CCS) read**: The consensus sequence resulting from alignment between subreads taken from a single ZMW. Generating a CCS read does **not** include or require alignment against a reference sequence but **does** require at least two full-pass subreads from the insert. CCS reads are advantageous for amplicon and RNA sequencing projects and are highly accurate (>99% accuracy, Q≥20).<br>  <br><br>*See, e.g.,* Pacific Biosciences Glossary of Terms (emphasis added above).<br><br>*See also* PacBio Sequencing - How It Works Video at PBIO_00000002. |

**K. [Claim 15]**

| U.S. Patent No. 11,091,794 CLAIM 15 | Pacific Biosciences of California, Inc. |
|---|---|
| [15.pre] The method of claim 1, | [*See* Claim 1 above] |
| [15.a] wherein each nucleotide within the window is enriched or filtered. | In the Accused Products and/or Systems, each nucleotide within the window is enriched or filtered.<br><br>For example, PacBio instructs each nucleotide within the window to be enriched or filtered by depleting short fragments in a SMRTbell® library—thereby enriching for long fragments—using AMPure PB beads size selection kits or with SMRTbell® cleanup beads provided in SMRTbell® Prep Kits:<br><br>**6. AMPure PB bead size selection or cleanup with SMRTbell cleanup beads**<br>The AMPure PB bead size selection step will clean the library and deplete DNA fragments shorter than 5 kb. Please see the Technical note – HiFi WGS Performance with AMPure PB bead size selection for more information on performance of this method.<br>To remove SMRTbell molecules with inserts greater than 5 kb from the SMRTbell library, please see the Technical Note - Alternative size selection methods for SMRTbell prep kit 3.0. If performing an alternative method, or skipping size selection altogether, proceed to the bead binding, wash, and elution steps by adding a 1X concentration of the SMRTbell cleanup beads to each library (50 µL).<br><br>*See, e.g.*, Preparing whole genome and metagenome libraries using SMRTbell® prep kit 3.0 at 12.<br><br>**Introduction**<br>In *Procedure & checklist — Preparing whole genome and metagenome sequencing libraries using SMRTbell prep kit 3.0*, size-selection beads are used to remove short DNA fragments and enrich for the long fragments that are most valuable for downstream analysis. Although size-selection beads have many advantages, alternative methods offer more flexibility in defining a size cutoff. This technical note provides guidance for size selection of WGS libraries prepared with the SMRTbell prep kit 3.0.<br><br>*See, e.g.*, Alternative size selection methods for SMRTbell® prep kit 3.0.<br><br>*See also*, claim 8 (size cutoff) |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 15 | Pacific Biosciences of California, Inc. |
|---|---|
| | As another example, PacBio demonstrates that a window within a particular genomic region of interest can be enriched using a CRISPR-Cas9 enrichment methodology, as shown below:   *See, e.g.,* No-Amp Targeted SMRT® Sequencing Using a CRISPR-Cas9 Enrichment Method, Figure 2. |

**L. [Claim 16]**

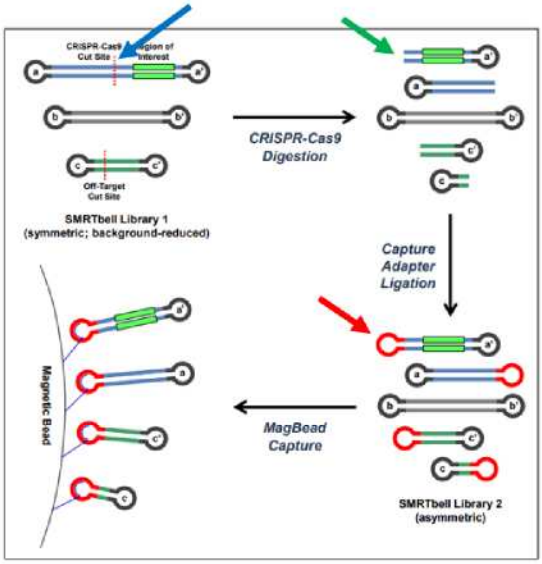
| U.S. Patent No. 11,091,794 CLAIM 16 | Pacific Biosciences of California, Inc. |
|---|---|
| [16.pre] The method of claim 15, | [*See* Claim 15 above] |
| [16.a] wherein each nucleotide within the window is enriched by: cutting a double-stranded DNA molecule using a Cas9 complex to form a cut double-stranded DNA molecule, and ligating a hairpin adaptor onto an end of the cut double-stranded DNA molecule, or | In the Accused Products and/or Systems, each nucleotide within the window is enriched by cutting a double-stranded DNA molecule using a Cas9 complex to form a cut double-stranded DNA molecule, and ligating a hairpin adaptor onto an end of the cut double-stranded DNA molecule<br><br>For example, the circular DNA molecule is formed by: cutting a double-stranded DNA molecule using a Cas9 complex to form a cut double-stranded DNA molecule, and ligating a hairpin adaptor onto an end of the cut double-stranded DNA molecule. For example, PacBio describes a method "for targeted SMRT® Sequencing using the CRISPR-Cas9 system that requires no PCR amplification" and "preserves epigenetic modifications." *See, e.g.,* No-Amp Targeted SMRT® Sequencing Using a CRISPR-Cas9 Enrichment Method. The method includes cutting a double-stranded DNA molecule using a Cas9 complex (blue arrow) to form a cut double-stranded DNA molecule (green arrow) and ligating a hairpin adaptor onto an end of the cut double-stranded DNA molecule (red arrow):<br><br>  <br><br>No-Amp Targeted SMRT® Sequencing Using a CRISPR-Cas9 Enrichment Method (Fig. 2) (annotated) |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

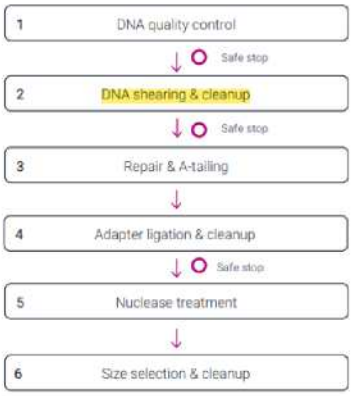
| U.S. Patent No. 11,091,794 CLAIM 16 | Pacific Biosciences of California, Inc. |
|---|---|
| [16.b] filtered by: selecting double-stranded DNA molecules having a size with a size range. | In the Accused Products and/or Systems, each nucleotide within the window is filtered by: selecting double-stranded DNA molecules having a size with a size range<br><br>For example, PacBio gives its users a workflow overview for producing a SMRTbell® library that includes a step of DNA shearing:<br><br>  <br><br><br>*See, e.g.,* Procedure & checklist for SMRTbell® prep kit 3.0 at 2.<br><br>The step of DNA Shearing results in filtering the DNA by selecting double-stranded DNA molecules having a size within a target size range as follows:<br><br>DNA Shearing \| Human, plant, and animal samples \| Microbial and Metagenomic samples<br>DNA shearing (Megaruptor 3 system) \| Speed 31 \| Speed 40<br>Target fragment lengths \| 15–18 kb \| 7–12 kb<br><br>*See, e.g.,* Procedure & checklist for SMRTbell® prep kit 3.0 at 1. |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

**M. [Claim 18]**

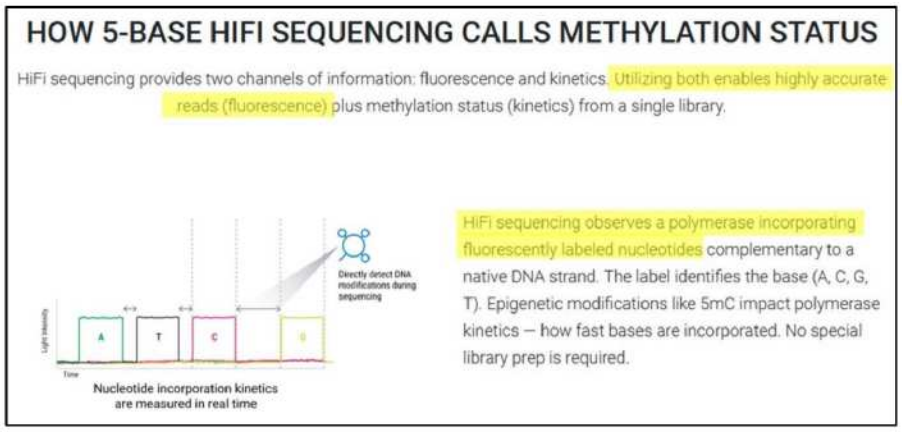
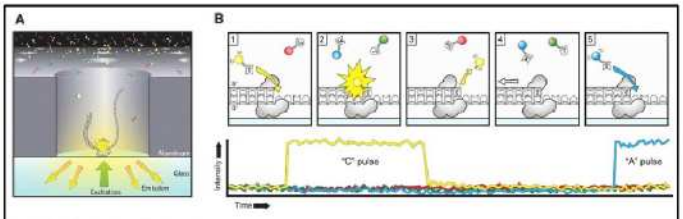
| U.S. Patent No. 11,091,794 CLAIM 18 | Pacific Biosciences of California, Inc. |
|---|---|
| [18.pre] The method of claim 1, | [*See* Claim 1 above] |
| [18.a] wherein the optical signal is a fluorescence signal from a dye-labeled nucleotide. | In the Accused Products and/or Systems, the optical signal is a fluorescence signal from a dye-labeled nucleotide.<br>For example, the methods discussed herein use a fluorescence signal as the optical signal:<br>  <br><br>*See, e.g.,* PacBio Epigenetics Webpage (annotated). |

**EXHIBIT A**
CONFIDENTIAL - OUTSIDE COUNSELS' EYES ONLY

| U.S. Patent No. 11,091,794 CLAIM 18 | Pacific Biosciences of California, Inc. |
|---|---|
| | PacBio indicates that fluorescent signals can be detected, for example, using a "fluorescent dye suitable for single molecule fluorescence detection." *See* Long, Processive enzymatic DNA synthesis using 100% dye-labeled terminal phosphate-linked nucleotides (right).<br><br>JULY 7, 2019 \| **Long, processive enzymatic DNA synthesis using 100% dye-labeled terminal phosphate-linked nucleotides.**<br><br>**Authors**: Korlach, Jonas and Bibillo, Arek and Wegener, Jeffrey and Peluso, Paul and Pham, Thang T and Park, Insil and Clark, Sonya and Otto, Geoff A and Turner, Stephen W<br><br>We demonstrate the efficient synthesis of DNA with complete replacement of the four deoxyribonucleoside triphosphate (dNTP) substrates with nucleotides carrying fluorescent labels. A different, spectrally separable fluorescent dye suitable for single molecule fluorescence detection was conjugated to each of the four dNTPs via linkage to the terminal phosphate. Using these modified nucleotides, DNA synthesis by phi 29 DNA polymerase was observed to be processive for products thousands of bases in length, with labeled nucleotide affinities and DNA polymerization rates approaching unmodified dNTP levels. Results presented here show the compatibility of these nucleotides for single-molecule, real-time DNA sequencing applications.<br><br>The use of fluorescent dye-labeled nucleotides to detect a fluorescence signal is also disclosed in Rhoads & Au at Fig. 3 (right). ("Each of the four nucleotides is labeled with a different fluorescent dye (indicated in red, yellow, green, and blue, respectively for G, C, T, and A) so that they have distinct emission spectrums.").<br><br>  <br><br>**Figure 3 Sequencing via light pulses**<br>**A.** A SMRTbell (gray) diffuses into a ZMW, and the adaptor binds to a polymerase immobilized at the bottom. **B.** Each of the four nucleotides is labeled with a different fluorescent dye (indicated in red, yellow, green, and blue, respectively for G, C, T, and A) so that they have distinct emission spectrums. As a nucleotide is held in the detection volume by the polymerase, a light pulse is produced that identifies the base. (1) A fluorescently-labeled nucleotide associates with the template in the active site of the polymerase. (2) The fluorescence output of the color corresponding to the incorporated base (yellow for base C as an example here) is elevated. (3) The dye-linker-pyrophosphate product is cleaved from the nucleotide and diffuses out of the ZMW, ending the fluorescence pulse. (4) The polymerase translocates to the next position. (5) The next nucleotide associates with the template in the active site of the polymerase, initiating the next fluorescence pulse, which corresponds to base A here. The figure is adapted from [4] with permission from The American Association for the Advancement of Science.<br><br>*See also* Introduction to SMRT® Sequencing Video at PBIO_00000001. |

**N. [Claim 19]**

| U.S. Patent No. 11,091,794 CLAIM 19 | Pacific Biosciences of California, Inc. |
|---|---|
| [19.pre] The method of claim 1, | [*See* Claim 1 above] |
| [19.a] wherein each window associated with the first plurality of data structures comprises 4 consecutive nucleotides on a first strand of each first nucleic acid molecule. | Each window associated with the first plurality of data structures of the Accused Products comprises 4 consecutive nucleotides on a first strand of each first nucleic acid molecule. For example, each window used as in input to the convolutional neural network (CNN) employed by the Primrose tool includes "a 16 base pair window for both the forward and reverse strand"—which comprises 4 consecutive nucleotides on a first strand of a HiFi read. *See, e.g.,* SMRT® Tools Reference Guide, v11.1 at PBIO_00001236; SMRT® Tools Reference Guide, v11.0 at PBIO_00001108; SMRT® Tools Reference Guide, v12.0 at PBIO_00001371. |