A diploid human genome (HG002) was sequenced with HiFi reads to around 30-fold coverage (SRA accession: SRX5527202). Reads were aligned to the human reference genome using the SMRT Link mapping analysis. Variants were detected using DeepVariant® and the Structural Variant Calling analysis in SMRT Link®. Variant calls were measured against the Genome in a Bottle (GIAB) small variant benchmark (v3.3.2) for SNVs and indels, and v0.6 benchmark for SVs. Phasing was performed using WhatsHap®. Reads were randomly subsampled to evaluate the impact of coverage on variant detection.
**DATA ANALYSIS SOLUTIONS**

- Call variants with HiFi reads using SMRT Analysis\(^2\), and GATK\(^8\) or Google DeepVariant\(^9\).
- Google DeepVariant performance is better than GATK, particularly for indels.
- Detect all variant types including SNVs, indels, SVs, and CNVs\(^4\).
- Comprehensively profile all variants in a human genome with the highest precision and recall\(^4,5,6\).
- Expand variant calling into previously inaccessible regions of the genome, including repetitive regions and medically relevant genes that are difficult to map\(^4,5\).
- Output data in standard file formats – BAM and VCF – for seamless integration with downstream analysis tools.
- Phase small variants into phase blocks using WhatsHap\(^7\).
- Confirm variant calls visually with IGV\(^10\) and GenomeRibbon\(^11\).

### HiFi READS INCREASE VARIANT DISCOVERY POWER

<table>
<thead>
<tr>
<th>SNVs</th>
<th>Small Indels</th>
<th>SVs</th>
</tr>
</thead>
<tbody>
<tr>
<td>99.3</td>
<td>98.6</td>
<td>95.1</td>
</tr>
<tr>
<td>95.1</td>
<td>93.0</td>
<td>95.4</td>
</tr>
<tr>
<td>96.8</td>
<td>96.8</td>
<td>99.3</td>
</tr>
</tbody>
</table>

**Coverage of the Cytochrome P450 2D6 (CYP2D6) and CYP2D7 genes with HiFi reads\(^6\) and NGS reads visualized in IGV\(^10\). CYP2D6 is responsible for the metabolism and elimination of approximately 25% of clinically used drugs\(^12\).**

### EXPAND VARIANT DETECTION IN MEDICALLY RELEVANT GENES

- Call variants with HiFi reads using SMRT Analysis\(^2\), and GATK\(^8\) or Google DeepVariant\(^9\).
- Google DeepVariant performance is better than GATK, particularly for indels.
- Detect all variant types including SNVs, indels, SVs, and CNVs\(^4\).
- Comprehensively profile all variants in a human genome with the highest precision and recall\(^4,5,6\).
- Expand variant calling into previously inaccessible regions of the genome, including repetitive regions and medically relevant genes that are difficult to map\(^4,5\).
- Output data in standard file formats – BAM and VCF – for seamless integration with downstream analysis tools.
- Phase small variants into phase blocks using WhatsHap\(^7\).
- Confirm variant calls visually with IGV\(^10\) and GenomeRibbon\(^11\).

For Research Use Only. Not for use in diagnostic procedures. © Copyright 2019, Pacific Biosciences of California, Inc. All rights reserved. Information in this document is subject to change without notice. Pacific Biosciences assumes no responsibility for any errors or omissions in this document. Certain notices, terms, conditions and/or use restrictions may pertain to your use of Pacific Biosciences products and/or third party products. Please refer to the applicable Pacific Biosciences Terms and Conditions of Sale and to the applicable license terms at http://www.pacbio.com/legal-and-terms/terms-and-conditions-of-sale/.

Pacific Biosciences, the Pacific Biosciences logo, PacBio, SMRT, SMRTbell, Iso-Seq, and Sequel are trademarks of Pacific Biosciences. BluePippin and SageELF are trademarks of Sage Science. NGS-go and NGSengine are trademarks of GenDx. Femto Pulse and Fragment Analyzer are trademarks of Agilent Technologies Inc. All other trademarks are the sole property of their respective owners.

**KEY REFERENCES**

2. SMRT Link user guide (v7.0.0). PacBio Documentation.
3. Sequel II System Data Release: HG002 SV and SNVs (HiFi Reads powered by CCS).
5. Wagner, J. et al. (2019) Expanding the Genome in a Bottle benchmark callsets with high-confidence small variant calls from long and linked read sequencing technologies. Advances in Genome Biology and Technology. Marco Island, FL.
12. NCBI Gene: CYP2D6 cytochrome P450 family 2 subfamily D member 6, Homo sapiens (human).