SMRT® Link Release Notes (v10.1)

SMRT Link Server Installation

- SMRT Link server software is supported on:
  - English-language CentOS 7.x, supported until end-of-life 6/30/2024.
  - English-language CentOS 8.x, supported until end-of-life 12/31/2021.
  - Ubuntu 18.04 and 20.04 64-bit Linux® distributions.
  - These supported versions **also** apply to SMRT Link compute nodes.
- **Note**: PacBio advises **against** new installations of CentOS for use with SMRT Link.
- SMRT Link is **not** guaranteed to work on Linux versions that are no longer supported by the Operating Systems' vendors.
- SMRT Link server software **cannot** be installed on systems running other versions of UNIX, macOS® or Windows®.
- Several SMRT Link v10.1 features are computationally-intensive and require adherence to the computational and storage requirements listed in the document **SMRT Link Software Installation (v10.1)**.
- To install **only** command-line SMRT Tools, use the `--smrttools-only` option with the installation command, whether for a new installation or an upgrade. Examples:
```
smrtlink-*.run --rootdir smrtlink --smrttools-only
smrtlink-*.run --rootdir smrtlink --smrttools-only --upgrade
```

Supported Chemistry

- SMRT Link v10.1 supports all chemistry versions for **Sequel® II Systems** and chemistry v2.1 and later for the **Sequel System**.

Included with the SMRT Link Installation

Barcode Sequences:

- gDNA_Sample_Amplification_Adapter
- IsoSeq_Primers_12_Barcodes_v1
- IsoSeqPrimers_v2 (Includes the content of IsoSeqPrimers as well as support for NEB and Clontech primers.)
- Sequel_16_barcodes_v1
- Sequel_16_barcodes_v3
- Sequel_96_barcodes_v1
- Sequel_384_barcodes_v1
- Sequel_64_M13barcodes_v1
- HiFiViral_SARS-CoV-2_M13barcodes

Reference Sequences:

- Sequel DNA Internal Control
- HIV_HXB2
- HLA_11locus_clustering_guide
- lambdaNEB

- SARS-CoV-2 Genome NC_045512.2
- SARS-CoV-2 Guide for PBAA

SMRT Link Database Note

- SMRT Link v10.1 no longer includes weekly automatic database backups. A database backup is still automatically performed once, during installation or upgrade. Failure to back up the SMRT Link database on a regular schedule risks losing **all** records in SMRT Link (including users, Data Sets, analyses, barcodes, and references) if a file system or reconfiguration error occurs. The underlying sequencing or analysis files, such as BAM files, are **not** affected. We **strongly** recommend asking your local Linux System Administrator to schedule regular weekly backups of the SMRT Link database using standard Linux utilities. For additional details, please contact PacBio Technical Support.

Iso-Seq® Method Notes

- The recommended analysis workflow for Iso-Seq experiments is to generate the CCS Reads and then analyze them using the Iso-Seq application. For details on suggested analysis parameters, see the **SMRT Link User Guide**, Iso-Seq Analysis Application section.
- **Note**: Continuous Long Read input for Iso-Seq analysis will be **deprecated** in the next SMRT Link release.

New Features (v10.1)

SMRT Link - SMRT® Analysis

- New SARS-CoV-2 Analysis Application - Uses HiFi Reads as input.
- **Note**: The Resequencing application will be **discontinued** in future releases.
- In Mapping applications, alignment concordance is reported as BLAST-style alignment identity (`matches/alignment columns`). Alignments are filtered by gap-compressed identity (`matches/ matches + mismatches + insertion events + deletion events.`)
- All Mapping applications for HiFi data include a CCS On-target report.
- Iso-Seq Demultiplexing improvements: Iso-Seq Analysis supports multiplexed samples by identifying isoforms per barcode. Iso-Seq Analysis also supports both joint and separate clustering of barcoded samples.

SMRT Link - Sample Setup

- New Calculation CSV import/export feature.

New Features (v10.0)

SMRT Link

- Support for new Sequel IIe System.
- SMRT Link end-to-end workflow enabled for Amazon AWS.
- Enhanced visual contrast throughout SMRT Link; optimized fonts, color schema and magnification.

SMRT Link - Sample Setup

- Now application-based for easier use.

SMRT Link - Run Design

- (**Sequel IIe only**) Added support for on-instrument HiFi Reads generation.
- Now application-based for easier use.

SMRT Link - SMRT® Analysis

- New applications:
  - Genome Assembly - Uses HiFi Reads as input.
  - Export Reads
  - CCS with Demultiplexing
- Mapping Application
  - Works with HiFi Reads and Continuous Long Reads input.
- Demultiplex Barcodes Application:
  - New GUI for easier editing of barcode/Bio Sample name pairs.

Fixed Issues (v10.0, v10.1)

- SV calling: The `pbsv call` command now correctly reports insertions present in the `svsig.gz` file.
- SV calling: Joint SV calling on demultiplexed Data Sets from the same cell/collection now works correctly. Demultiplexed Data Sets are now analyzed separately, and the Bio Sample name is used correctly.
- Sample Setup: Columns in the edit/print view can now be drag-and-dropped.
- Copying to the clipboard now works as expected.
- Exporting large analysis directories now works correctly and does not fail.
- The outputs analysis directory now includes symbolic links to the BAM files.
- BAM files consolidation for microbial assemblies now works correctly.
- Absolute file paths are now included in the `subreadset.xml` file.
- Login for local SMRT Link WSO2 users is now enabled.
- Run Design now does **not** allow setting up invalid barcoded Iso-Seq Run Designs.
- Instrument type selection, as specified in the global SMRT Link settings, now persists correctly.
- If the header line in the CSV file containing the Bio Sample names is accidentally omitted, the demultiplexing analysis still proceeds for **all** barcodes as intended. We recommend using the template CSV file that can be downloaded from SMRT Link so that the Bio Sample names can be added next to the corresponding barcodes.
- Using the hyphen character "-" in barcode and Bio Sample names no longer causes the Demultiplex Barcodes application to fail.

- Fixed an issue when using the interactive barcoded sample editor in Run Design or SMRT Analysis with asymmetric barcodes. Dragging large numbers of samples from the **Available Barcodes** column to the **Included Barcodes** column now works correctly.
- Fixed an issue where specifying an input FASTA file for `gcpp` without a FASTA index file (`.fai`) with the same name and path caused the software to fail.
- HGAP4 analysis no longer fails if the Genome Size is set to more than 2.0 GB.
- After changing a value in an applications' **Advanced Parameters** dialog, the **Cancel** button now displays correctly.
- After clicking **Save** In **Data Management > Create Data Set**, the **Save** button now works correctly.

Known Issues (v10.0, v10.1)

- Run Designs created in v10.0 or v10.1 will **not** work correctly with ICS v8.0 on a Sequel System. The Run Designs are **not** backwards-compatible and lose application information.
- When opening a saved Run Design, you will sometimes be asked to save changes when no changes were made.
- Run Design: The **Import from Sample Setup** feature does not distinguish between Sample Setup designs created for Sequel II and for Sequel IIe, instead showing **both** if either instrument is currently selected in Run Design. As of this release, the imported information is equivalent for both instrument types, so the user may select any of the displayed entries and the feature should function as expected.
- When creating a new Run Design and an invalid value is entered in the Pre-Extension Time field, the **Save** button is correctly disabled, but the allowed value range is no longer displayed.
- **Some** Sequel runs generated using SMRT Link versions prior to v6.0.0 cannot be viewed, and display the message `getDefaultInstrumentTypeForChipType: Unexpected chip name`.
- A cached URL containing the string `/welcome` at the end of the SMRT Link URL (Example: `https://URL/sl/welcome`) in the browser's history causes an error when accessing SMRT Link. Delete `welcome` from the URL to access SMRT Link.
- Motif detection is designed for microbial genomes and has **not** been tested on non-microbial genomes; it may run out of memory on large genomes.
- In the Structural Variant Calling application's **Advanced Parameters** dialog, when entering a floating-point number (such as `1.3`) for **Minimum % of Reads that Support Variant (any one sample)**, only `1` is passed to the application. **Workaround**: Enter an integer number only.
- When copying an analysis using the Demultiplex Barcodes application (using **Copy** from an Analysis Results page or **Copy From** on the New Analysis page), the input of sample names for each barcode is **not** preserved from the copied analysis. In the second step of the New Analysis wizard, users must **either** re-enter the sample names using

the Interactive Barcode Selector and Sample Name Editor, or re-upload a Barcoded Sample File. The **Start** button is **not** enabled until users do so. Note that when using the Barcode Selector and Sample Name Editor there is a checkbox-option to Download as a file for later use. Using this feature makes it easier to re-enter the barcode sample names in the event of copying a Demultiplex Barcodes analysis.

- When creating a user, ensure that the new user profile has the `Username` attribute populated with the account/login name. This is **required** for the user search in the configuration and project pages to find local users. (See **SMRT Link Software Installation (v10.1)** for details.)

- When using Bio Sample Names with a PacBio analysis application, you can enter names that include spaces. Please **avoid** using spaces in Bio Sample Names as spaces may lead to third-party compatibility issues.

- Occasionally, Run QC table metrics do **not** print out correctly. As a workaround, click **Export** to export the metrics in CSV format, or take a screen shot of the metrics table.

- Using the **Advanced Search** functionality in the Data Management and SMRT Analysis modules and entering a "date =" query using a precise time stamp of date/hours/minutes/seconds causes an error. Searching for the date **without** hours/minutes/seconds succeeds.

- Several CCS Analysis plots display an increase in "Other Reads" at 5 passes. This is a cosmetic issue caused by a reporting artifact in CCS Analysis, as the number of passes for a given ZMW can be erroneously inflated when filtered at certain stages. This issue has **no** impact on polished CCS Reads or HiFi yield. The reporting discrepancy will be addressed in the next release.

- Iso-Seq summary.csv file: For separately clustered samples (or single samples), the per-sample transcript counts currently specify the primer pair instead of the sample name. Jointly clustered samples have separate demultiplexed transcript count files with the correct sample.