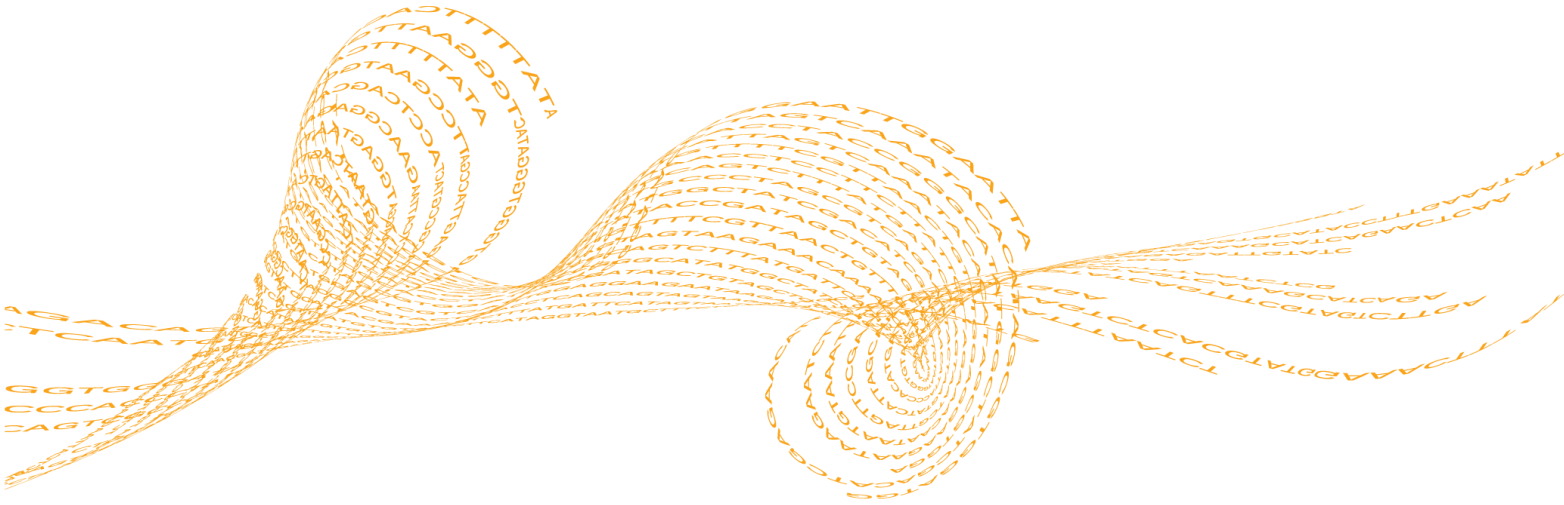
BaseSpace® Onsite v2.1 HT

System Guide

For Research Use Only. Not for use in diagnostic procedures.



ILLUMINA PROPRIETARY

Document # 15049148 v01
October 2015

illumina®

Read Before Using this Product

This Product, and its use and disposition, is subject to the following terms and conditions. If Purchaser does not agree to these terms and conditions then Purchaser is not authorized by Illumina to use this Product and Purchaser must not use this Product.

1. **Definitions**. "**Application Specific IP**" means Illumina owned or controlled intellectual property rights that pertain to this Product (and use thereof) only with regard to specific field(s) or specific application(s). Application Specific IP excludes all Illumina owned or controlled intellectual property that cover aspects or features of this Product (or use thereof) that are common to this Product in all possible applications and all possible fields of use (the "**Core IP**"). Application Specific IP and Core IP are separate, non-overlapping, subsets of all Illumina owned or controlled intellectual property. By way of non-limiting example, Illumina intellectual property rights for specific diagnostic methods, for specific forensic methods, or for specific nucleic acid biomarkers, sequences, or combinations of biomarkers or sequences are examples of Application Specific IP. "**Consumable(s)**" means Illumina branded reagents and consumable items that are intended by Illumina for use with, and are to be consumed through the use of, Hardware. "**Documentation**" means Illumina's user manual for this Product, including without limitation, package inserts, and any other documentation that accompany this Product or that are referenced by the Product or in the packaging for the Product in effect on the date of shipment from Illumina. Documentation includes this document. "**Hardware**" means Illumina branded instruments, accessories or peripherals. "**Illumina**" means Illumina, Inc. or an Illumina affiliate, as applicable. "**Product**" means the product that this document accompanies (e.g., Hardware, Consumables, or Software). "**Purchaser**" is the person or entity that rightfully and legally acquires this Product from Illumina or an Illumina authorized dealer. "**Software**" means Illumina branded software (e.g., Hardware operating software, data analysis software). All Software is licensed and not sold and may be subject to additional terms found in the Software's end user license agreement. "**Specifications**" means Illumina's written specifications for this Product in effect on the date that the Product ships from Illumina.
2. **Research Use Only Rights**. Subject to these terms and conditions and unless otherwise agreed upon in writing by an officer of Illumina, Purchaser is granted only a non-exclusive, non-transferable, personal, non-sublicensable right under Illumina's Core IP, in existence on the date that this Product ships from Illumina, solely to use this Product in Purchaser's facility for Purchaser's internal research purposes (which includes research services provided to third parties) and solely in accordance with this Product's Documentation, **but specifically excluding any use that** (a) would require rights or a license from Illumina to Application Specific IP, (b) is a re-use of a previously used Consumable, (c) is the disassembling, reverse-engineering, reverse-compiling, or reverse-assembling of this

Product, (d) is the separation, extraction, or isolation of components of this Product or other unauthorized analysis of this Product, (e) gains access to or determines the methods of operation of this Product, (f) is the use of non-Illumina reagent/consumables with Illumina's Hardware (does not apply if the Specifications or Documentation state otherwise), or (g) is the transfer to a third-party of, or sub-licensing of, Software or any third-party software. All Software, whether provided separately, installed on, or embedded in a Product, is licensed to Purchaser and not sold. Except as expressly stated in this Section, no right or license under any of Illumina's intellectual property rights is or are granted expressly, by implication, or by estoppel.

Purchaser is solely responsible for determining whether Purchaser has all intellectual property rights that are necessary for Purchaser's intended uses of this Product, including without limitation, any rights from third parties or rights to Application Specific IP. Illumina makes no guarantee or warranty that purchaser's specific intended uses will not infringe the intellectual property rights of a third party or Application Specific IP.

3 **Regulatory**. This Product has not been approved, cleared, or licensed by the United States Food and Drug Administration or any other regulatory entity whether foreign or domestic for any specific intended use, whether research, commercial, diagnostic, or otherwise. This Product is labeled For Research Use Only. Purchaser must ensure it has any regulatory approvals that are necessary for Purchaser's intended uses of this Product.

4 **Unauthorized Uses**. Purchaser agrees: (a) to use each Consumable only one time, and (b) to use only Illumina consumables/reagents with Illumina Hardware. The limitations in (a)-(b) do not apply if the Documentation or Specifications for this Product state otherwise. Purchaser agrees not to, nor authorize any third party to, engage in any of the following activities: (i) disassemble, reverse-engineer, reverse-compile, or reverse-assemble the Product, (ii) separate, extract, or isolate components of this Product or subject this Product or components thereof to any analysis not expressly authorized in this Product's Documentation, (iii) gain access to or attempt to determine the methods of operation of this Product, or (iv) transfer to a third-party, or grant a sublicense, to any Software or any third-party software. Purchaser further agrees that the contents of and methods of operation of this Product are proprietary to Illumina and this Product contains or embodies trade secrets of Illumina. The conditions and restrictions found in these terms and conditions are bargained for conditions of sale and therefore control the sale of and use of this Product by Purchaser.

5 **Limited Liability. TO THE EXTENT PERMITTED BY LAW, IN NO EVENT SHALL ILLUMINA OR ITS SUPPLIERS BE LIABLE TO PURCHASER OR ANY THIRD PARTY FOR COSTS OF PROCUREMENT OF SUBSTITUTE PRODUCTS OR SERVICES, LOST PROFITS, DATA OR BUSINESS, OR FOR ANY INDIRECT, SPECIAL, INCIDENTAL, EXEMPLARY, CONSEQUENTIAL, OR PUNITIVE DAMAGES OF ANY KIND ARISING OUT OF OR IN CONNECTION WITH, WITHOUT LIMITATION, THE SALE OF THIS PRODUCT, ITS USE, ILLUMINA'S PERFORMANCE HEREUNDER OR ANY OF THESE TERMS AND CONDITIONS, HOWEVER ARISING OR CAUSED AND ON ANY THEORY OF LIABILITY (WHETHER IN CONTRACT, TORT (INCLUDING NEGLIGENCE), STRICT LIABILITY OR OTHERWISE).**

6 **ILLUMINA'S TOTAL AND CUMULATIVE LIABILITY TO PURCHASER OR ANY THIRD PARTY ARISING OUT OF OR IN CONNECTION WITH THESE TERMS AND CONDITIONS, INCLUDING WITHOUT LIMITATION, THIS PRODUCT (INCLUDING USE THEREOF) AND ILLUMINA'S PERFORMANCE HEREUNDER, WHETHER IN CONTRACT, TORT (INCLUDING NEGLIGENCE), STRICT LIABILITY OR OTHERWISE, SHALL IN NO EVENT EXCEED THE AMOUNT PAID TO ILLUMINA FOR THIS PRODUCT.**

7 **Limitations on Illumina Provided Warranties. TO THE EXTENT PERMITTED BY LAW AND SUBJECT TO THE EXPRESS PRODUCT WARRANTY MADE HEREIN ILLUMINA MAKES NO (AND EXPRESSLY DISCLAIMS ALL) WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THIS PRODUCT, INCLUDING WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, NONINFRINGEMENT, OR ARISING FROM COURSE OF PERFORMANCE, DEALING, USAGE OR TRADE. WITHOUT LIMITING THE GENERALITY OF THE FOREGOING, ILLUMINA MAKES NO CLAIM, REPRESENTATION, OR WARRANTY OF ANY KIND AS TO THE UTILITY OF THIS PRODUCT FOR PURCHASER'S INTENDED USES.**

8 **Product Warranty**. All warranties are personal to the Purchaser and may not be transferred or assigned to a third-party, including an affiliate of Purchaser. All warranties are facility specific and do not transfer if the Product is moved to another facility of Purchaser, unless Illumina conducts such move.

a **Warranty for Consumables**. Illumina warrants that Consumables, other than custom Consumables, will conform to their Specifications until the later of (i) 3 months from the date of shipment from Illumina, and

(ii) any expiration date or the end of the shelf-life pre-printed on such Consumable by Illumina, but in no event later than 12 months from the date of shipment. With respect to custom Consumables (i.e., Consumables made to specifications or designs made by Purchaser or provided to Illumina by, or on behalf of, Purchaser), Illumina only warrants that the custom Consumables will be made and tested in accordance with Illumina's standard manufacturing and quality control processes. Illumina makes no warranty that custom Consumables will work as intended by Purchaser or for Purchaser's intended uses.

b **Warranty for Hardware**. Illumina warrants that Hardware, other than Upgraded Components, will conform to its Specifications for a period of 12 months after its shipment date from Illumina unless the Hardware includes Illumina provided installation in which case the warranty period begins on the date of installation or 30 days after the date it was delivered, whichever occurs first ("Base Hardware Warranty"). "Upgraded Components" means Illumina provided components, modifications, or enhancements to Hardware that was previously acquired by Purchaser. Illumina warrants that Upgraded Components will conform to their Specifications for a period of 90 days from the date the Upgraded Components are installed. Upgraded Components do not extend the warranty for the Hardware unless the upgrade was conducted by Illumina at Illumina's facilities in which case the upgraded Hardware shipped to Purchaser comes with a Base Hardware Warranty.

c **Exclusions from Warranty Coverage**. The foregoing warranties do not apply to the extent a non-conformance is due to (i) abuse, misuse, neglect, negligence, accident, improper storage, or use contrary to the Documentation or Specifications, (ii) improper handling, installation, maintenance, or repair (other than if performed by Illumina's personnel), (iii) unauthorized alterations, (iv) Force Majeure events, or (v) use with a third party's good not provided by Illumina (unless the Product's Documentation or Specifications expressly state such third party's good is for use with the Product).

d **Procedure for Warranty Coverage**. In order to be eligible for repair or replacement under this warranty Purchaser must (i) promptly contact Illumina's support department to report the non-conformance, (ii) cooperate with Illumina in confirming or diagnosing the non-conformance, and (iii) return this Product, transportation charges prepaid to Illumina following Illumina's instructions or, if agreed by Illumina and Purchaser, grant Illumina's authorized repair personnel access to this Product in order to confirm the non-conformance and make repairs.

e **Sole Remedy under Warranty**. Illumina will, at its option, repair or replace non-conforming Product that it confirms is covered by this warranty. Repaired or replaced Consumables come with a 30-day warranty. Hardware may be repaired or replaced with functionally equivalent, reconditioned, or new Hardware or components (if only a component of Hardware is non-conforming). If the Hardware is replaced in its entirety, the warranty period for the replacement is 90 days from the date of shipment or the remaining period on the original Hardware warranty, whichever is shorter. If only a component is being repaired or replaced, the warranty period for such component is 90 days from the date of shipment or the remaining period on the original Hardware warranty, whichever ends later. The preceding states Purchaser's sole remedy and Illumina's sole obligations under the warranty provided hereunder.

f **Third-Party Goods and Warranty**. Illumina has no warranty obligations with respect to any goods originating from a third party and supplied to Purchaser hereunder. Third-party goods are those that are labeled or branded with a third-party's name. The warranty for third-party goods, if any, is provided by the original manufacturer. Upon written request Illumina will attempt to pass through any such warranty to Purchaser.

9 **Indemnification**.

a **Infringement Indemnification by Illumina**. Subject to these terms and conditions, including without limitation, the Exclusions to Illumina's Indemnification Obligations (Section 9(b) below), the Conditions to Indemnification Obligations (Section 9(d) below), Illumina shall (i) defend, indemnify and hold harmless Purchaser against any third-party claim or action alleging that this Product when used for research use purposes, in accordance with these terms and conditions, and in accordance with this Product's Documentation and Specifications infringes the valid and enforceable intellectual property rights of a third party, and (ii) pay all settlements entered into, and all final judgments and costs (including reasonable attorneys' fees) awarded against Purchaser in connection with such infringement claim. If this Product or any part thereof, becomes, or in Illumina's opinion may become, the subject of an infringement claim, Illumina shall have the right, at its option, to (A) procure for Purchaser the right to continue using this Product, (B) modify or replace this Product with a substantially equivalent non-infringing substitute, or (C) require the return of this Product and terminate the rights, license, and any other permissions provided to Purchaser with respect this Product and refund to Purchaser the depreciated value (as shown in

Purchaser's official records) of the returned Product at the time of such return; provided that, no refund will be given for used-up or expired Consumables. This Section states the entire liability of Illumina for any infringement of third party intellectual property rights.

b **Exclusions to Illumina Indemnification Obligations**. Illumina has no obligation to defend, indemnify or hold harmless Purchaser for any Illumina Infringement Claim to the extent such infringement arises from: (i) the use of this Product in any manner or for any purpose outside the scope of research use purposes, (ii) the use of this Product in any manner not in accordance with its Specifications, its Documentation, the rights expressly granted to Purchaser hereunder, or any breach by Purchaser of these terms and conditions, (iii) the use of this Product in combination with any other products, materials, or services not supplied by Illumina, (iv) the use of this Product to perform any assay or other process not supplied by Illumina, or (v) Illumina's compliance with specifications or instructions for this Product furnished by, or on behalf of, Purchaser (each of (i) – (v), is referred to as an "Excluded Claim").

c **Indemnification by Purchaser**. Purchaser shall defend, indemnify and hold harmless Illumina, its affiliates, their non-affiliate collaborators and development partners that contributed to the development of this Product, and their respective officers, directors, representatives and employees against any claims, liabilities, damages, fines, penalties, causes of action, and losses of any and every kind, including without limitation, personal injury or death claims, and infringement of a third party's intellectual property rights, resulting from, relating to, or arising out of (i) Purchaser's breach of any of these terms and conditions, (ii) Purchaser's use of this Product outside of the scope of research use purposes, (iii) any use of this Product not in accordance with this Product's Specifications or Documentation, or (iv) any Excluded Claim.

d **Conditions to Indemnification Obligations**. The parties' indemnification obligations are conditioned upon the party seeking indemnification (i) promptly notifying the other party in writing of such claim or action, (ii) giving the other party exclusive control and authority over the defense and settlement of such claim or action, (iii) not admitting infringement of any intellectual property right without prior written consent of the other party, (iv) not entering into any settlement or compromise of any such claim or action without the other party's prior written consent, and (v) providing reasonable assistance to the other party in the defense of the claim or action; provided that, the party reimburses the indemnified party for its reasonable out-of-pocket expenses incurred in providing such assistance.

e **Third-Party Goods and Indemnification**. Illumina has no indemnification obligations with respect to any goods originating from a third party and supplied to Purchaser. Third-party goods are those that are labeled or branded with a third-party's name. Purchaser's indemnification rights, if any, with respect to third party goods shall be pursuant to the original manufacturer's or licensor's indemnity. Upon written request Illumina will attempt to pass through such indemnity, if any, to Purchaser.

Revision History

Part #	Date	Description of Change
Document # 15049148 v01	October 2015	Updated document for BaseSpace Onsite HT System v2.1: • Updated supported apps list • Added information about annotation options • Updated supported browsers • Added information about new instruments supported: NextSeq 550, MiSeq, NeoPrep, and HiSeq 3000/4000 • Updated User Interface descriptions to match current release • Added NeoPrep setup workflow for Prep tab • Updated Download Run File Package with information about BaseSpace Onsite Downloader • Updated Admin Tasks-Updates section about capability to upload additional genomes files
Part # 15049148 Rev. E	May 2015	Minor changes for final launch.
Part # 15049148 Rev. D	April 2015	Supporting BaseSpace Onsite v2.0 using BaseSpace version 3.14.
Part # 15049148 Rev. C	August 2014	Added log package descriptions.
Part # 15049148 Rev. B	June 2014	Added descriptions for the following features: • Prep tab library import • Upload VCF file • Custom library • VariantStudio app
Part # 15049148 Rev. A	March 2014	Initial Release

Introduction

The BaseSpace Onsite LT System is a genomics analysis platform that is a directly integrated end-to-end solution for the MiSeq System, using MiSeq Control Software (MCS) v2.6 or later.

The BaseSpace Onsite HT System is a genomics analysis platform that is a directly integrated end-to-end solution for the following systems:

- **NextSeq 500 System**
- **NextSeq 550 System**—For sequencing analysis only, with NextSeq Control Software (NCS) v2.0.0.10 or later
- **NeoPrep System**—With NeoPrep Control Software v1.1 or later
- **HiSeq 1000, HiSeq 1500, HiSeq 2000, HiSeq 2500 Systems**—With HiSeq Control Software (HCS) v2.2.68 or later
- **HiSeq 3000, HiSeq 4000 Systems**—With HCS v3.3.52 or later
- **MiSeq System**—With MiSeq Control Software (MCS) v2.6 or later

BaseSpace Onsite has the following features:

- You can prepare a NextSeq and NeoPrep run on the BaseSpace Onsite Prep Tab, and then start that run from your instrument.
- The instrument seamlessly pushes the base call (*.bcl) files and associated files to BaseSpace Onsite for automatic analysis and storage. There is no need for a manual and time-consuming data-transfer step.
- BaseSpace Onsite provides a mechanism to share data with others and easily scale storage and computing needs.
- BaseSpace Onsite runs locally; no need to connect to the cloud.

Workflow Model

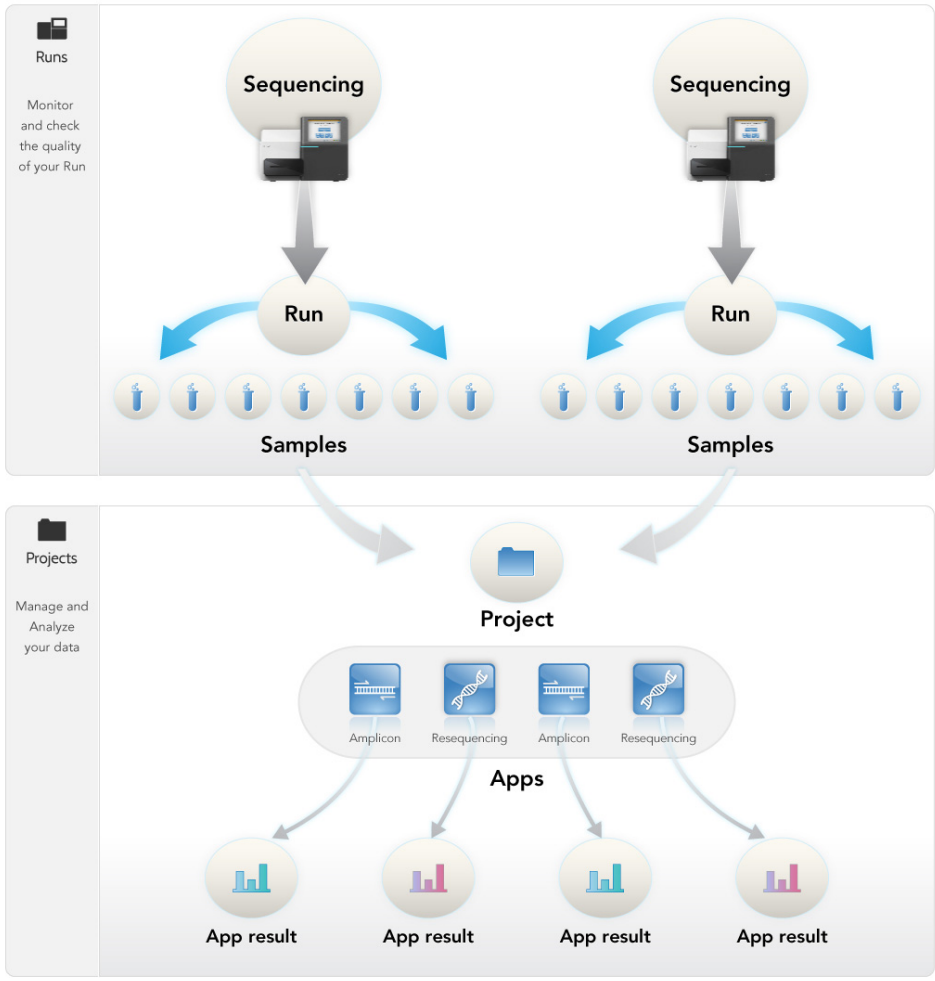
Processing a flow cell on a sequencing instrument produces various files, collectively referred to as a run. A run contains log files, instrument health data, run metrics, sample sheet, and base call information (*.bcl files). The base call information is demultiplexed in BaseSpace Onsite to create the samples used in secondary analysis.

Samples are analyzed by launching BaseSpace Onsite apps. BaseSpace Onsite apps are processing software and routines that interact with BaseSpace Onsite data through the API.

The result files from an app session are stored in an analysis. Analyses are created to record every time an app is launched. For example, when a resequencing app executes alignment and variant calling, an analysis is created that contains the app results for each sample. App results generally contain BAM and VCF files, but they can also contain other file types. App results can also be used as inputs to apps.

Finally, projects are simple containers that store samples and analyses.

Figure 1 BaseSpace Onsite Data Model



BaseSpace Onsite Apps

BaseSpace Onsite supports the following apps.

App	Description
Illumina Core Apps	The following Illumina Core Apps are available: • Isaac Whole-Genome Sequencing v4 • Isaac Enrichment v1.0 • Isaac Enrichment v2.1 • BWA Whole-Genome Sequencing v1 • BWA Enrichment v1.0 • BWA Enrichment v2.1 • Tumor Normal v1.0 • TopHat Alignment v1 (multilaunch feature not available) • Cufflinks Assembly and Differential Expression v1 • TruSeq Amplicon v1.1 • TruSeq Long Reads v1.1 • TruSeq Phasing v1.1 • Amplicon DS v1.1 • 16S Metagenomics v1.0.1 • MethylSeq v1.0 • Small RNA v1.0 • TruSeq Targed RNA v1.0 For documentation, see the Illumina support page for BaseSpace.
VariantStudio	Illumina VariantStudio enables you to view annotation information for variants and transcripts and filter to the variants of interest in one easy to use application. See *Run the VariantStudio App* on page 43 for instructions on launching this app, and the *BaseSpace VariantStudio User Guide* for instructions on using the software application.
Integrative Genomics Viewer	Illumina has generated an app with the Integrative Genomics Viewer (IGV) of the Broad Institute for visualizing your sequence data in great detail. See *Launch the IGV App* on page 42 for instructions on using this app.
BaseSpace Labs Apps	The following apps are developed using an accelerated development process. Illumina provides limited support for BaseSpace Labs Apps and are provided without a warranty. • FastQC v1.0 • FASTQ Toolkit v1.0 • Kraken Metagenomics v1.0 • Variant Calling Assessment Tool (VCAT) v.2.0
Third-Party Apps	Currently, BaseSpace Onsite provides SPAdes Genome Assembler 3.5. For more information, see the SPAdes App page on the Apps tab. **Note**—The NovoAlign app is no longer available. However, legacy analysis results are still viewable.

App	Description
BaseMount	BaseMount is a tool that allows command line read access to your BaseSpace Onsite data. The BaseMount app is a stand-alone installable application that you can install on your own computer. Then you can connect BaseMount to the data on your BaseSpace Onsite server or servers. For more information, see the following links: • BaseMount Web Page • BaseMount Help • BaseSpace Blog **Note**—BaseSpace is the default API endpoint for BaseMount. Make sure to launch BaseMount using the correct server configuration option for BaseSpace Onsite: `basemount --config {config_file_prefix} --api-server=http://{BSO_windows_IP}:8080 {mount-point_folder_path}` • The `--config` parameter is optional. Use this parameter to mount against multiple user accounts. • The `BSO_windows_IP` is the IP address that you enter in your browser to access your BaseSpace Onsite account. The API url is at port 8080. • The `mount-point_folder_path` is the path to a user-created directory on the system where BaseMount is running.

Annotation Options

Depending on the application used, the following annotation options are available.

Option	Description
Illumina Annotation Service RefSeq or Ensembl	This option utilizes the Illumina Annotation Service (hosted at annotation.basespace.illumina.com) to provide annotation of variants (SNPs, insertions, and deletions) in the sample VCF files. This service requires an external network connection and also requires you to whitelist the IP address of your BaseSpace Onsite system.
Basic Annotation with local genome files from the UCSC hg19 genome	This option utilizes a basic set of annotations, including SNPs and indels, using the genome files hosted locally on the BaseSpace Onsite system. This option does ***not*** require an external network connection.
On-node analysis RefSeq or Ensembl	This option is a localized version of Illumina Annotation Service provided with the BaseSpace Onsite v2.1 software update. It provides annotations for known variants as provided by Illumina Annotation Service, and can also predict novel variants. This option does ***not*** require an external network connection.
None	No annotation is performed with this option. The reports generated for analyses using this option shows either a 0 or N/A for the various annotations.

The following apps provide either the Illumina Annotation Service or None options:

- Amplicon DS v1.0
- BWA Enrichment v2.1
- Isaac Enrichment v2.1
- TruSeq Amplicon v1.1
- Tumor Normal v1.0

The following apps provide the Illumina Annotation Service, Basic, or None options:

- BWA Enrichment v1.0
- BWA Whole Genome Sequencing v1.0
- Isaac Enrichment v1.0
- Isaac Whole Genome Sequencing v2

The Isaac Whole Genome Sequencing v4.0 application can perform annotation on-node, without the need for an external network connection to the Illumina Annotation Service.

All other applications on BaseSpace Onsite do not provide annotation options. To perform annotation on the results of these analyses, you can use the VariantStudio application, which also performs annotation using the Illumina Annotation Service.

Getting Started

Make sure to have a compatible desktop computer. BaseSpace Onsite v2.1 is compatible with Internet Explorer (10 or higher), Chrome (43 or higher), Firefox (38 or higher), and Safari (8 or higher). Hardware requirements for these browsers can be found on the developer website. A PC running Windows 7 or greater is required to access the BaseSpace Onsite.

See the following topics to get started with BaseSpace Onsite.

- *Using BaseSpace Onsite with NextSeq* on page 12
- *Using BaseSpace Onsite with NeoPrep* on page 13
- *Using BaseSpace Onsite with HiSeq* on page 13
- *Using BaseSpace Onsite with MiSeq* on page 15
- *Getting Shared Data* on page 16

NOTE

When connecting your instrument to BaseSpace Onsite, you are asked to enter the full path to your BaseSpace Onsite server. Your BaseSpace Onsite IP address can be found in the broker configuration file of the sequencing system that you want to connect BaseSpace Onsite to. The BaseSpace Onsite IP address is included as part of the API URL value.

Using BaseSpace Onsite with NextSeq

BaseSpace Onsite is the Illumina analysis local environment. BaseSpace Onsite facilitates your experiments on the NextSeq system in 2 different ways:

- BaseSpace Onsite helps to organize your samples and experiments, and preps runs for NextSeq.
- BaseSpace Onsite stores and analyzes your sequencing run data, providing the following benefits:
  - Enables browser-based data management and analysis
  - Provides tools for institute-wide collaboration and sharing
  - Eliminates the need to manage your own storage and computing solutions

BaseSpace Onsite has limited storage capacity, and checks the free space available before uploading a run. If there is not enough available space, you see an error message; see *Storage Check* on page 68 for more information.

NOTE

Only use BaseSpace Onsite with NextSeq 550 when performing sequencing runs. To analyze array scanning data from NextSeq 550, use BlueFuse Multi Software.

NOTE

When you bookmark the BaseSpace Onsite location in your browser, make sure to bookmark the Dashboard page after logging on. Do not bookmark the Login page.

NextSeq Connection

Perform the following steps to connect NextSeq to BaseSpace Onsite.

1. If you are using NextSeq 550, make sure that you have NCS v2.0.0.10 or later. If you do not, upgrade your control software.
2. Make sure that you have a stable connection of at least 10 Mbps upload speed from the NextSeq.
3. From the Manage Instrument screen, click **System Configuration**.
4. Click **BaseSpace Configuration**.

5 Select BaseSpace Onsite.

6 In the **Server Name** field, enter the full path to your BaseSpace Onsite server.

7 [Optional] Select **Browse** and navigate to a secondary network **Output Folder** location to save a copy of data in addition to the BaseSpace Onsite server.

8 Click **Save**.

9 Log in to BaseSpace Onsite when setting up the run on the NextSeq system.

Using BaseSpace Onsite with NeoPrep

The NeoPrep library prep workflow is integrated with BaseSpace Onsite. For instruments configured for BaseSpace Onsite, library information and run parameters are specified on the BaseSpace Onsite Prep tab before setting up the run on the instrument. Runs that were set up in BaseSpace Onsite appear on the instrument interface during run setup. As the run progresses, output files are streamed in real time to BaseSpace Onsite or BaseSpace Onsite.

NeoPrep Connection

Perform the following steps to connect NeoPrep to BaseSpace Onsite.

1 Make sure that you have NeoPrep Control Software version v1.1 or later. If you do not, upgrade your control software.

2 From the Manage Instrument screen on the NeoPrep, select **System Configuration**.

3 Select **BaseSpace Configuration**.

4 Select **BaseSpace Onsite**.

5 In the **Api Url** field, enter the full path to your BaseSpace Onsite server.

6 [Optional] Select **...**, and navigate to a secondary network **Output Folder** location to save a copy of data in addition to the BaseSpace Onsite server.

CAUTION
Do not specify a location on the instrument computer for the output folder. Illumina recommends using a network location only. The capacity of the instrument hard drive is sufficient for multiple runs, however a full drive can prevent subsequent runs.

7 Select **Next**.

8 Set the BaseSpace Onsite parameters as follows.
- Enter a BaseSpace Onsite **User Name** and **Password** to register the instrument with BaseSpace Onsite.
- Select **Use default login and bypass the BaseSpace login screen** to set the registered user name and password as the default login. This setting bypasses the BaseSpace screen during NeoPrep run setup.

9 Select **Finish**.

Using BaseSpace Onsite with HiSeq

The HiSeq sequencing systems feature an option to send instrument health and sequencing data to BaseSpace Onsite in real time to streamline both instrument quality control and analysis. Real-time monitoring of runs enables fast troubleshooting. Free alignment and variant calling provide many easy to use workflows that tailor analysis for diverse biological applications.

When you set up the run on the HiSeq, select the option to log in to BaseSpace Onsite. If you have a problem with the data upload between HiSeq and BaseSpace Onsite, see *HiSeq Connection* on page 14.

NOTE
Raw data from the run is also stored on the instrument, or in the location of the output folder that you specified in the Storage screen.

BaseSpace Onsite has limited storage capacity, and checks the free space available before uploading a run. If there is not enough available space, you see an error message; see *Storage Check* on page 68 for more information.

BaseSpace Onsite automatically disconnects from the HiSeq at the end of the run or when all base calling files have finished uploading. If the internet connection is interrupted, analysis files continue uploading after the connection is restored from the point when the interruption occurred.

When the last base call file is uploaded to BaseSpace Onsite, secondary analysis of your data begins. For information about how to run tasks, see *How To Use BaseSpace Onsite* on page 27.

HiSeq Connection

Perform the following steps to connect HiSeq to BaseSpace Onsite.

1. Make sure that you have the correct HCS version:
   - HiSeq 1000, HiSeq, 1500, HiSeq 2000, HiSeq 2500—HCS v2.2.68 or later
   - HiSeq 3000, HiSeq 4000—HCS v3.3.52 or later

   If you do not, upgrade your HCS.
2. Make sure that you have a stable internet connection of at least 10 Mbps upload speed from the HiSeq.
3. During run configuration, the Integration screen on the HiSeq enables you to define where your run data are output and stored. Perform the following steps.
   a. Select **BaseSpace Onsite**.
   b. Select **Storage and Analysis**—Sends run data to BaseSpace for remote monitoring and secondary analysis. A sample sheet is required with this option.
   c. Enter your **BaseSpace Onsite** account information.
   d. Go to the Storage screen and specify the storage options.
4. If BaseSpace Onsite is not available, open Windows Services and start or restart Illumina BaseSpace Onsite Broker.
   a. Click the Windows **Start** button.
   b. Right-click **Computer**, and then select **Manage**.
   c. On the left side, under Services and Applications, choose **Services**.
   d. Scroll down the list to find Illumina BaseSpace Broker.
   e. Right-click Illumina BaseSpace Broker. If available, click **Start**. If the Start option is grayed out, click **Restart**. The service starts, or closes then restarts.
   f. Close the Computer management window.

NOTE
To use BaseSpace Onsite, load a sample sheet at the start of your run.

For more information, see the system guide for your HiSeq instrument.

When you begin your sequencing run on the HiSeq, the BaseSpace Onsite icon changes to indicate that the HiSeq is connected to BaseSpace Onsite and data files are being transferred.

Using BaseSpace Onsite with MiSeq

The MiSeq System features an option to send instrument health and sequencing data to BaseSpace Onsite in real time to streamline both instrument quality control and analysis. Real-time monitoring of runs enables fast troubleshooting. Free alignment and variant calling provide many easy to use workflows that tailor analysis for diverse biological applications.

When you set up the run on the MiSeq, select the option to log in to BaseSpace Onsite. For more information, see *MiSeq Connection* on page 15.

NOTE
Raw data from the run is also stored on the instrument, or in the location of the output folder that you specified in the Storage screen.

BaseSpace Onsite has limited storage capacity, and checks the free space available before uploading a run. If there is not enough available space, you see an error message; see *Storage Check* on page 68 for more information.

BaseSpace Onsite automatically disconnects from the MiSeq at the end of the run or when all RTA analysis files have finished uploading. If the internet connection is interrupted, analysis files continue uploading after the connection is restored from the point when the interruption occurred.

When the last base call file is uploaded to BaseSpace Onsite, secondary analysis of your data begins. For information about how to run tasks, see *How To Use BaseSpace Onsite* on page 27.

NOTE
Unlike BaseSpace, BaseSpace Onsite does not have a MiSeq Reporter app. BaseSpace Onsite directs all workflow types uploaded from MiSeq through the GenerateFASTQ app. To perform additional secondary analysis, manually launch the relevant app or apps on the resulting FASTQ samples.

MiSeq Connection

Perform the following steps to connect MiSeq to BaseSpace Onsite.

1. Make sure that you have MCS version 2.6 or later. If you do not, upgrade your MCS.
2. Make sure that you have a stable internet connection of at least 10 Mbps upload speed from the MiSeq.
3. During run configuration, enter the full path to your BaseSpace Onsite server.
4. Select or clear **When using BaseSpace or BaseSpace Onsite, replicate analysis locally on MiSeq**.
   The Replicate Analysis Locally setting specifies analysis processing locations when using BaseSpace Onsite. The setting provides the option to perform analysis both locally on the instrument and in BaseSpace Onsite. Consider the following items when selecting or deselecting this option:
   - If you select this option, MiSeq Reporter launches automatically after the run and performs analysis locally.
   - If you do not select this option, MiSeq Reporter does not launch automatically after the run and analysis is performed in BaseSpace Onsite only.
   - If performing the VeriSeq PGS workflow with BlueFuse Multi, select this option.

5 Click **Save and Return**.

6 When starting a sequencing run, select **Use BaseSpace Onsite for storage and analysis** on the BaseSpace Options screen.

NOTE
To use BaseSpace Onsite, load a sample sheet at the start of your run.

For more information, see the *MiSeq System Guide (document # 15027617)*.

When you begin your sequencing run on the MiSeq, the BaseSpace Onsite icon changes to indicate that the MiSeq is connected to BaseSpace Onsite and data files are being transferred.

Getting Shared Data

If you receive a link to shared data in BaseSpace Onsite, click the link.

Use your BaseSpace Onsite account to log on; the first time you visit you are asked to accept the BaseSpace Onsite agreement. After that, you are ready to run BaseSpace Onsite. Ask your local administrator to set up a new account.

If someone shares data with you, you receive a notification. The shared data show up in your project list. Now you can use the BaseSpace Onsite tools to look at and download the data. For information about how to run tasks, see *How To Use BaseSpace Onsite* on page 27.

NOTE
The owner of the data can disable the sharing feature at any time.

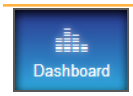
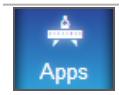
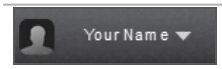
NOTE
When you bookmark the BaseSpace Onsite location in your browser, make sure to bookmark the Dashboard page after logging on. Do not bookmark the Login page.

BaseSpace Onsite User Interface

This section describes the various aspects of the BaseSpace Onsite UI.

Toolbar

The BaseSpace Onsite toolbar elements are listed in the following table.

Icon	Element	Description
	Dashboard Tab	See *Dashboard Tab* on page 18.
	Prep Tab	See *Prep Tab* on page 20.
	Runs Tab	See *Runs Tab* on page 22.
	Projects Tab	See *Projects Tab* on page 24.
	Apps Tab	See *Apps Tab* on page 26.
	Support Page	The BaseSpace Onsite Support page provides access to the BaseSpace Onsite Knowledge Base, User Guide, and Illumina Technical Support. Access to this information requires that your computer is connected to the internet.
	Search	The Search box allows you to find runs, projects, samples, files, or apps. For more information, see *Search for Runs, Projects, Samples, Files, and Apps* on page 65.
	Account	The Account drop-down list provides access to: • MyAccount—See *MyAccount* on page 18 • Trash—See *Delete Items* on page 61 • Terms—Leads to the User Agreement • Admin Panel—Only available if you have admin privileges; see *Admin Panel* on page 18 • Blog—Leads to the blog, which includes latest news, developments, and updates • Sign out

MyAccount

MyAccount provides access to the Settings, Transfer History, Library Prep Kits, and Storage pages.

Settings

On the Settings page you can edit your notifications settings, edit your profile, or update your profile picture.

Transfer History

The Transfer History page allows you to review projects or runs that have been transferred. For more information, see *Transfer Owner* on page 61.

Library Prep Kits

The Library Prep Kits page shows all the available library prep kits in Sample Prep. To enter a custom kit, click **New**. Custom kits allow you to define nonstandard index combinations and to define your own default layout for the indexes on a plate.

Storage

The Storage page shows the total amount of storage used (owned by you and shared) and the amount of storage used in Runs, Projects, and Trash.

Admin Panel

The Admin Panel allows you to manage analysis, notifications, storage, users, system health, planned runs, software updates, and alarms. See *Admin Tasks* on page 66 for a description.

NOTE
You need administrator privileges to see and work in the admin panel.

Dashboard Tab

After login, the first tab you see is the dashboard. The dashboard provides access to storage information, the developers page, the newsfeed, notifications, latest runs, and latest analyses. The dashboard is always accessible in BaseSpace Onsite from the top ribbon selector.

NOTE
If a run or project is not showing on BaseSpace Onsite, it is possible your data has not been sent to BaseSpace Onsite. Set the BaseSpace Onsite option on your sequencing instrument. For more information, see the system guide for your instrument.

Storage

The storage pane shows the total amount of storage used (used by you and shared). Click the **Details** button to open the Storage page under My Account.

Developers

Click the **Details** button on the Developers pane to go to the BaseSpace developer portal. Access to the BaseSpace developer portal requires an external internet connection outside

of your local network.

Newsfeed

The Newsfeed pane shows the most recent posts from the BaseSpace blog. Click an article name to open the article, or click the **Read more** link at the bottom of the pane to open the BaseSpace blog.

The Newsfeed pane requires an external internet connection outside of your local network to show content. Access to the BaseSpace blog also requires an external internet connection.

Notifications

This pane shows invitations to shared projects, along with other notifications and alerts. The most recent notifications listed first. There are multiple types of notifications:

- Runs
  - Run in progress
  - Run completed
  - Run error
- Collaborators
  - Collaborator joined a project/run of which you are a member
  - Collaborator invited you to a project/run
  - (optionally) collaborator has included a personal message
  - Collaborator recommended an App
  - Collaborator accepted your offer to transfer ownership
  - Collaborator offered to transfer ownership to you.
- Analyses by you
  - Analysis in progress
  - Analysis completed
  - Analysis error
- Analyses by collaborators
  - Analysis in progress
  - Analysis completed
  - Analysis error
- Uploads, additions, or deletions to/from a project of which you are a member
  - By you
  - By a collaborator

NOTE
If you did not configure SMTP during install, all notifications show up on the dashboard. No notifications are sent through email.

Latest Runs

The Latest Runs pane shows the 3 most recent runs with their dates and status, and is updated automatically.

Click **All Runs** to open the Runs tab. Clicking a run opens the Runs tab with the run loaded. For more information, see *Runs Tab* on page 22.

Latest Analyses

The Latest Analyses pane shows the 3 most recent analysis with their dates and status, and is updated automatically.

Click **All Analyses** to open the Analyses page, which lists all analysis results, along with the application used, project name, date updated, size of the result, and status.

Clicking an analysis on the dashboard pane opens the specific analysis report. For more information, see *Analyses Page* on page 25.

Prep Tab

The Prep tab enables you to set up a run on the NeoPrep or NextSeq system.

This tab is only available for NeoPrep and NextSeq. Other sequencing instruments use a sample sheet to provide sample information to BaseSpace.

To use the Prep functionality for a NeoPrep run, click the NeoPrep option. For more information, see *NeoPrep* on page 20.

To use the Prep functionality for a NextSeq run, click the 4 steps under Manual Prep.

1. **Biological Samples**—Contains information about the samples that are going to be sequenced. See *Biological Samples* on page 20.
2. **Libraries**—Consists of biological samples that are prepped and contain adapters. Each library usually derives from a single biological sample, though biological samples can be used in multiple libraries. See *Libraries* on page 21.
3. **Pools**—Consists of groups of libraries that share analysis parameters. Pools can consist of 1 or multiple libraries. See *Pools* on page 21.
4. **Planned Runs**—Contains pools that run with the same analysis parameters, on the same machine, at the same time. Planned runs can consist of 1 or multiple pools. See *Planned Runs* on page 22.

NeoPrep

The NeoPrep Runs Prep page allows you to create library cards and manage NeoPrep runs. The top of the page shows the number of runs that are ready, running, or complete.

Using the Show Runs functionality, you can filter the list of runs to include or exclude runs according to its status. Click 1 or more of the following statuses: **Ready**, **Running, Completed**, **Failed**, and **Planning**. Statuses with orange toggle switch icons are included in the list. Statuses with gray toggle switch icons are excluded. Using the filters for Ready, Running, and Complete statuses affects the run status count at the top of the page.

To create a run, click **Create New Run** at the upper right corner of the page. For more information, see *Prepare a NeoPrep Run* on page 27.

Manual Prep

Biological Samples

The Biological Samples list shows all available samples that you have created on your account.

If you want information about the samples, you can perform the following:

- Sort the list by clicking the column headers.

- Click a sample to go to the sample page.

This page provides the following actions to prepare your analysis:

- Create a sample.
- Import new samples.
- Select a sample and edit its properties.
- Select 1 or more samples and continue with Prep Libraries.

NOTE
You can select multiple samples by using the following methods:

- Select multiple checkboxes.
- Click anywhere on a sample row while holding **Ctrl** button to add to a selection.
- Click anywhere on a sample row while holding **Shift** button to select all samples in between.
- Click the checkbox next to Plate/Tube ID to select all samples on the current page.

The box next to the Biological Samples header tracks the total number of samples, and how many are selected. Click **X** next to the selection count to clear the current selection.

For more information about these actions, see *Create New Biological Samples* on page 29, *Import Biological Samples* on page 29, and *Use Existing Biological Samples* on page 30.

Libraries

The Libraries list shows all available plates or tubes with libraries that you have created on your account. You can sort the list by clicking the column headers, or click a plate to see its properties and associated libraries.

This page provides the following actions to prepare your analysis:

- Click a plate, then click **Edit** to edit its properties or libraries.
- Select 1 or more plates or tubes and move to Pool Libraries.
- Import libraries and associate them to new biological samples at the same time. When going through the Import workflow, you can also download a template import CSV file. For more information, see *Import Samples and Libraries* on page 32.

NOTE
You can select multiple libraries by using the following methods:

- Select multiple checkboxes.
- Click anywhere on a library row while holding **Ctrl** button to add to a selection.
- Click anywhere on a library row while holding **Shift** button to select all libraries in between.
- Click the checkbox next to Plate/Tube ID to select all samples on the current page.
- Use the import function.

The box next to the Libraries header tracks the total number of libraries, and how many are selected. Click **X** next to the selection count to clear the current selection.

For more information about these actions, see *Prep Libraries* on page 30. When prepping a library, you can also create a custom library kit; see *Set Up Custom Library Prep Kit* on page 31.

Pools

The Pools list shows all available pools of libraries that you have created on your account. You can sort the list by clicking the column headers, or click a pool to see its properties and associated libraries.

This page provides the following actions to prepare your analysis:

- Click a pool, and then click the **Edit** button to edit the notes.
- Select a pool and move to Plan Run.

- Merge pools from multiple plates.

The box next to the Pools header tracks the total number of pools, and how many are selected.

For more information about these actions, see *Pool Libraries* on page 33.

Planned Runs

The Planned Runs tab you see the Planned Runs list shows all planned runs that you have created on your account.

You can sort the list by clicking the column headers, or click a run to see or edit its properties. For more information about these actions, see *Plan Runs* on page 34.

The runs can have the following states:

- **Ready to Sequence**—The run can be started from the NextSeq System.
- **Planning**—The run does not show up on the NextSeq System, because it is still in the planning stage.

NOTE

You can select multiple samples by using the following methods:

- Select multiple checkboxes.
- Click anywhere on a planned run row while holding **Ctrl** button to add to a selection.
- Click anywhere on a planned run row while holding **Shift** button to select all runs in between.
- Click the checkbox next Experiment Name to select all planned runs on the current page.

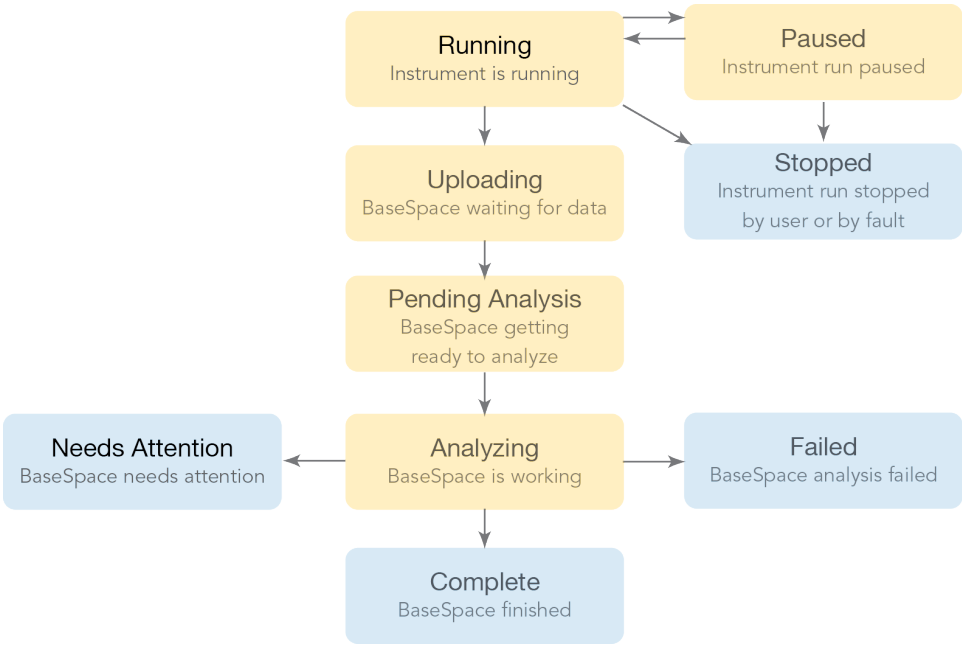
The box next to the Planned Runs header tracks the total number of runs, and how many are selected. Click **X** next to the selection count to clear the current selection.

When sequencing on a run starts, the run is removed automatically from the Planned Runs list.

Runs Tab

The Runs button leads to the runs list, which allows you to sort your runs based on experiment name, state, workflow, created date, machine, and owner.

The following run states are possible. The blue boxes indicate final states.



If you want to look at a run in detail, click the name to view metrics in more detail. For more information, see *Run Overview Page* on page 23.

You manage the runs by using the buttons above the list:

- **Share**—Manage sharing a run with a particular collaborator. See *Share a Run Using the Email Option* on page 58.
- **Get Link**—Forward the sharing link to any number of collaborators. See *Share a Run with Get Link* on page 57.
- **Download Run**—Download files from this run. See *Download Run File Package* on page 45.
- **Transfer Owner**—Hand control of data over to a collaborator or customer. This button is visible if a run is selected. See *Transfer Owner* on page 61.
- **Move to Trash**—Delete a project. This button is visible if a run is selected. See *Delete Project* on page 62.
- **View Trash**—View the deleted items in the trash, so you can restore the items or empty the trash. See *Delete Items* on page 61.

NOTE
Runs and projects have separate permissions. If you share a project, you do not share the runs contained within the project.

Run Overview Page

The Run Overview page provides 5 panes:

- **Run Details**—Provides a summary of the run with links to view files and download and share options. For more information, see *Share Data* on page 56, *View Files and Results* on page 35, or *Download Files* on page 44.
- **Samples** —Provides a list of all the app results in the run, the associated projects, and the number of samples in that analysis.
- **Charts**—Shows an intensity by cycle chart. Clicking the header takes you to the Charts page, which contains 5 charts with run metrics.
- **Run Summary**—Shows tables with basic data quality metrics. Clicking the header takes you to the Run Summary page.

- **Indexing QC**—Lists count information for indexes used in the run. Clicking the header takes you to the Indexing QC page.

In addition, the Side Navigation ribbon provides easy navigation in the Run Details area.

It contains links to the Overview, Run Samples List, Charts, Run Summary, Indexing QC, Run Settings, and Files pages.

Run Samples List

The samples list allows you to sort the samples in your run based on sample ID, app, date created, and project. If you want to look at a sample, app result, or project in detail, click the links to get to the following pages:

- *Sample Overview Page* on page 25.
- *Analyses Page* on page 25.
- *Project Overview Page* on page 24.

In addition, the Side Navigation ribbon provides easy navigation in the Run Details area.

It contains links to the Overview, Run Samples List, Charts, Run Summary, Indexing QC, Run Settings, and Files pages.

Projects Tab

The Projects button opens a list of your projects. You can sort the list by name, last update, or owner. Clicking a project provides access to the app results and samples within that project.

You manage the projects by using the buttons above the list:

- **New Project**—Generate a new project. See *Set Up a New Project* on page 59.
- **Edit project**—Edit the name and description of the project. See *Edit Project Details* on page 58.
- **Share Project**—Manage sharing a project with a particular collaborator. See *Share a Project Using the Email Option* on page 57.
- **Get Link**—Forward the sharing link to any number of collaborators. See *Share a Project using the Get Link option* on page 56.
- **Transfer Owner**—Hand control of data over to a collaborator or customer. See *Transfer Owner* on page 61.
- **Move to Trash**—Delete a project. This button is visible if a project is selected. See *Delete Project* on page 62.
- **View Trash**—View the deleted items in the trash, so you can restore them or empty the trash. See *Delete Items* on page 61.

NOTE

Runs and projects have separate permissions. If you share a project, you do not share the runs contained within the project.

Project Overview Page

The Project Overview page provides access to 3 panes with information about the project:

- **About**—Provides summary information about the project, including the project owner, shared status, date created, and collaborators.
- **Analyses**—Provides a list of all the App Sessions in the project. This tab can be sorted based on analysis name, last modified date created, status, or application

used to generate the analysis. Clicking the analysis links to the app results for that sample. For more information, see *Analyses Page* on page 25.

- **Samples**—Provides a list of all the samples in the project. Clicking a sample links to the page for that sample. For more information, see *Sample Overview Page* on page 25. Selecting the samples allows you to launch it in an app, copy to a different project, or combine with another result.

NOTE
You can access these panes through the left navigation bar.

Project Toolbar

The Project toolbar provides the following actions:

- **Launch app**—Run apps on your sample. Clicking the app name leads to a page with more information about launching that app, including access permissions. See *Analyze Samples Further* on page 42.
- **Download Project**—Download all files in a project. See *Download Project or Analysis Package* on page 1.
- **Import**—Upload files to a project. See *Upload Files* on page 59.
- **Share project**—Manage sharing a project with a particular collaborator. See *Share a Project Using the Email Option* on page 57.
- **Get link**—Forward the sharing link to any number of collaborators. See *Share a Project using the Get Link option* on page 56.
- **Edit project**—Edit the name and description of the project. See *Edit Project Details* on page 58.
- **Transfer Owner**—Hand control of data over to a collaborator or customer. See *Transfer Owner* on page 61.
- **Move to Trash**—Move the project to the trash. Deleting a project deletes all associated data, including samples and analyses. See *Delete Project* on page 62.
- **View Trash**—View all deleted runs, projects, analyses, and samples. See *Delete Items* on page 61.

Options that are not available for the particular analysis or sample are grayed out.

If you have selected samples in the **Samples** pane, you can perform additional actions:

- **Copy to...**—Copy samples from this project to another. See *Copy Samples* on page 59.
- **Combine**—Combine samples. See *Combine Samples* on page 59.

The app session states are defined as follows.

State	Description
Running	The app is processing or uploading data.
Complete	Processing and file upload has finished and the data are now available to use.
Aborted	This AppResult or Sample has been aborted and cannot be resumed.
Needs Attention	Cannot continue without user intervention.

Analyses Page

The Analyses page provides access to the results for that app session. There is a general information pane to the left, and several graphs, depending on the app run.

Sample Overview Page

The Sample Overview page provides 2 panes:

- **Sample Details**—Provides a summary of the run with a links to launch a custom BaseSpace Onsite app on your sample. Clicking the app name leads to a page with more information about that app, including access permissions.
  Running apps can incur a charge.
- **Files**—Provides a list of files associated with that sample. You can either look at all FASTQ files, or look at files specific for an app session.

For more information, see *View Files and Results* on page 35. You can also download selected files; see *Download Multiple FASTQ Files* on page 45.

Apps Tab

The Apps button leads to the Apps page, which provides an overview of the custom BaseSpace Onsite apps that you can run.

- Clicking the app name leads to a page with more information about that app, including the version, a link to the developer, and their app support contact details.
- Clicking the Launch button leads you through the launch pages, which allow you to set up the app session. Specify parameters like the project, sample, or output folder used by the app, depending on the app, and accept access permissions.
- You can search for apps using Search.

How To Use BaseSpace Onsite

The following topics describe how to run different functions in BaseSpace Onsite.

Prepare a NeoPrep Run

You can prepare NeoPrep runs through the BaseSpace Onsite Prep tab. Use this option if you want to perform library preparation on a NeoPrep instrument, and have the data stream seamlessly to BaseSpace Onsite.

To create a NeoPrep run, perform the following steps.

1. Log in to BaseSpace Onsite. If it is your first time logging in, accept the user agreement.
2. Click the **Prep** icon.
3. Click the **NeoPrep** option.
4. Click **Create New Run** and enter the required information in the following screens:
   - Configure Run
   - Assign Samples to Wells
   - Review Run Details

Configure Run

1. Select **TruSeq Nano DNA** or **TruSeq Stranded mRNA** in the Protocol dropdown.
2. Select a version of the protocol.
3. [Optional] Enter notes about the run.
4. Enter a run name.
5. Click **Select Project(s)** to select a default project.
   Previously unassigned and any new biological samples created for the NeoPrep run are assigned to the default project. This selection does not change the project assignments of any previously assigned biological samples that are used in the NeoPrep run.
6. Select a project, and then click **Confirm**.
7. Click **Next**.
   The Configure Run screen expands to allow you to enter more information.
8. If performing quantification or normalization in addition to library prep, select the **Quantitate** or **Normalize** checkboxes.
   Library Prep is selected by default and cannot be deselected.
9. Select the number of samples in the Sample Count dropdown.
10. If available, select the insert size.

11 Select the number of PCR cycles for the run. The supported number of cycles is indicated.
If you select a number that is not recommended, you are warned that run success and library quality are only guaranteed with the supported number. Click **OK** to acknowledge the message.

12 Click **Next**.

Assign Samples to Wells

1 For each well, enter the following information:

- **Sample ID**—Select a sample from the dialog, and then click **Select**. Alternatively, click **New** to create a biological sample, and then enter the sample ID, sample name, species, and nucleic acid type.

  NOTE
  The sample ID and sample name can only contain alphanumeric characters, dashes, or underscores. The sample ID must be unique and short. The sample name can be more descriptive to provide a human-readable identifier.

  NOTE
  You can use the same sample in multiple wells, as long as different library IDs are used.

- **Library ID**—Enter a library ID. You cannot use the same library ID for multiple wells.
- **Index**—Select an index. You cannot use the same index for multiple wells.
- **Insert Size**—Select the insert size.

2 Click **Next**.

Review Run Details

1 Review the run details.
If any information is incorrect, click **Back** to return to the previous screens.

2 Click **Finish** to create the run.

Prepare a NextSeq Run

You can prepare NextSeq runs through the BaseSpace Onsite Prep tab, which organizes samples, libraries, pools, and run in a single environment. This option allows NextSeq sequencing data stream seamlessly to BaseSpace Onsite.

Do not use it to prepare sequencing runs for other sequencing instruments. If you do have a NextSeq sequencing system but do not want to use BaseSpace Onsite, you can also start a run straight on the instrument.

To prepare a NextSeq run, perform the following steps.

1 Log in to BaseSpace Onsite. If it is your first time logging in, accept the user agreement.

2 Click the **Prep** icon.

3 Set up a NextSeq run on the Prep Tab in 4 consecutive steps.

a **Biological Samples**—Contains information about the samples that are going to be sequenced. You can create samples, import samples, or use existing samples.

For instructions, see *Create New Biological Samples* on page 29, *Import Biological Samples* on page 29, or *Use Existing Biological Samples* on page 30.

b **Libraries**—Consists of biological samples that are prepped and contain adapters. Each library usually derives from a single biological sample, though biological samples can be used in multiple libraries; see *Libraries* on page 21. You can also import biological samples and libraries in one step; see *Import Samples and Libraries* on page 32.

c **Pools**—Consists of groups of libraries that share analysis parameters. Pools can consist of 1 or multiple libraries. See *Pools* on page 21.

d **Planned Runs**—Contains pools that run with the same analysis parameters, on the same machine, at the same time. Planned runs can consist of 1 or multiple pools. See *Planned Runs* on page 22.

Create New Biological Samples

If you want to create a new biological sample, perform the following steps.

NOTE
Use the import function to create several new samples. For more information, see *Import Biological Samples* on page 29.

1 Click the **Prep** icon.

2 Click **Biological Samples**.

3 Click the **+ Create** button.

4 Fill out the required fields Sample ID, Name, and Nucleic Acid type.

NOTE
The sample ID and sample name can only contain alphanumeric characters, dashes, or underscores. The sample ID must be unique and short. The sample name can be more descriptive to provide a human-readable identifier.

5 [Optional] Fill out the Organism (species) field.

6 [Optional] Fill out the Project fields. You can also generate a new project. A project is optional here, but required later, because the output data gets stored to the project.

7 When finished, perform 1 of the following:

- If you only want to select the newly created samples, click **Next: Prep Libraries**. Continue with *Prep Libraries* on page 30.
- If you want to select multiple samples, click **Save & Continue Later**. This selection takes you back to the Biological Samples list, with the recently created sample at the top of the list. Continue with *Use Existing Biological Samples* on page 30.

Import Biological Samples

If you want to import new biological samples, perform the following steps.

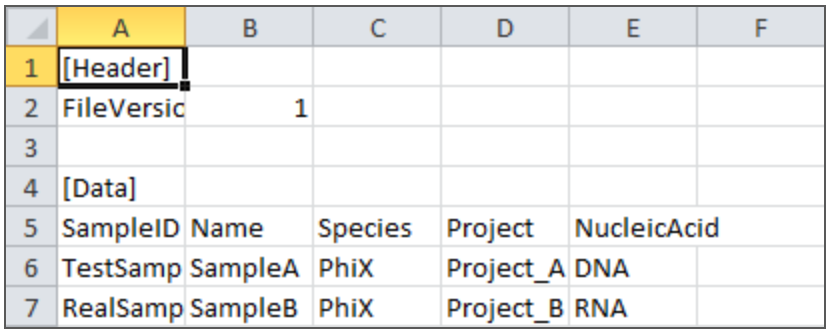
1 Click the **Prep** icon.

2 Click **Biological Samples**.

3 Click the **Import** button.

4 If you have not generated an import file yet, click the template link and fill out the samples. Note the following items when filling out the template:
   - The sample ID and sample name can only contain alphanumeric characters, dashes, or underscores. The sample ID must be unique and short. The sample name can be more descriptive to provide a human-readable identifier.
   - The Organism (species) field is optional.
   - The Project field is optional, but a project is required later because the output are stored to the project.
   - Fill out the Nucleic Acid column with DNA or RNA.

Figure 2 Import Sample Template



	A	B	C	D	E	F
1	[Header]					
2	FileVersic	1				
3						
4	[Data]					
5	SampleID	Name	Species	Project	NucleicAcid	
6	TestSamp	SampleA	PhiX	Project_A	DNA	
7	RealSamp	SampleB	PhiX	Project_B	RNA	

5 Click the **Choose File** button.

6 Browse to the import file, and then click **Open**.

7 Click **Import**.

8 When finished, perform 1 of the following:
   - If you only want to select the newly created samples, click **Next: Prep Libraries**. Continue with *Prep Libraries* on page 30.
   - If you want to select multiple samples, click **Save & Continue Later**. This selection takes you back to the Biological Samples list, with the recently created sample at the top of the list. Continue with *Use Existing Biological Samples* on page 30.

Use Existing Biological Samples

The Biological Samples list shows all available samples that you have created on your account.

1 To select existing samples, perform 1 of the following steps in the Biological Samples list:
   - Select the checkboxes.
   - Click the sample. If you want to select multiple samples, hold the **Ctrl** button.
   - Select all samples by selecting the checkbox next to the SampleID header.

2 Click the **Prep Libraries** button in the top navigation bar.

Prep Libraries

On the Prep Libraries page, you assign indexes to biological samples, based on the indexes available in the library preparation chosen. Every used well or tube contains a separate library. The best practice is to set up the libraries in BaseSpace Onsite first. Then export a file of your library settings, and use that to pipette the biological samples into the proper wells or tubes.

NOTE
If you do not want to use indexed sequencing, you still assign your biological sample to an index. When you set up your sequencing run, you can specify that you do not want to sequence the index.

1 Select the library prep type. BaseSpace Onsite now automatically assigns indexes to wells or tubes, depending on the format of the library prep type.

2 Enter the plate ID. The ID must be unique.

3 Click the **Auto Prep** button to fill the plate or tubes automatically with all samples listed.
Alternatively, manually drag the samples to the wells or tubes.

   a Select 1 or more samples. Hold the Shift key to select multiple samples. To select multiple samples on Firefox or Internet Explorer, click the well twice.
   b Drag the selected samples to a position.
   c Make sure that the indexes have been assigned to the proper samples. Hovering over a position reveals the sample that is assigned to that position. You can drag samples from position to position.

4 Save a file of your library settings by clicking the **Download CSV** button. Use this file in the lab to indicate which biological samples get pipetted into specific wells.

5 When finished, perform 1 of the following.
   - If you want to select the new plate or tubes, click the **Pool Libraries** button. Continue with *Pool Libraries* on page 33.
   - If you want to select multiple library preps or plates, perform the following steps.

      a Click **Save & Continue Later**. This selection takes you to the Libraries list, with the recently created set-up at the top of the list.
      b Select the checkboxes in the Libraries list.
      c Click the **Pool Libraries** button in the top navigation bar.

NOTE
If any samples are not assigned to a project, you cannot continue. Select the sample, click the **Set Project** button, and assign it to a project. You can also generate a new project.

Set Up Custom Library Prep Kit

When you have selected your biological samples, perform the following steps to set up a custom library prep kit.

1 When prepping a library, select **+ Custom Library Prep Kit** in the Library Prep Kit dropdown menu.
The Custom Library Prep Kit Definition page opens.

2 Enter a name for the custom prep that meets the following requirements:
   - Unique for your account
   - Contains only alphanumeric, hyphen, underscores, and spaces
   - Contains ≤ 50 characters

3 Select at least 1 of the supported read types.

4 Select at least 1 of the indexing strategies. You cannot only select **None**.

5 Fill out the default number of cycles.

6 Click **template** to download the index definition file template.

7 Fill out the **Settings** section:
   - For single read only—No adapter (blank), or 1 adapter sequence for Read 1.
   - For paired-end—No adapter (blank), or 2 adapter sequences, 1 for Read 1 and 1 for Read 2.
   - Make sure that each adapter sequence meets the following criteria:
     - Sequence of A, T, C, or G character
     - Length from 1 to 20 characters

8 Fill out the **Index1Sequences** and **Index2Sequences** sections:
   - For Single Index, with or without None:
     - 1 to 100 Index 1 names
     - For each Index 1 name an associated Index 1 sequence
   - For Dual Index, with or without None and Single Index:
     - 1 to 100 Index 1 names
     - For each Index 1 name an associated Index 1 sequence
     - 1 to 100 Index 2 names
     - For each Index 2 name an associated Index 2 sequence
   - Each index name meets the following criteria:
     - Unique within the file
     - Length from 1 to 8 characters alphanumeric, hyphen, or underscore characters
   - Each index sequence meets the following criteria:
     - Sequence of A, T, C, or G characters
     - Length from 1 to 20 characters
     - All index sequence lengths (Read 1 and Read 2) are equal
     - Index 1 sequences are unique within the file set of Index 1 sequences
     - Index 2 sequences are unique within the file set of Index 2 sequences

9 If the supported indexing strategy specifies Single Index, you can set up **Default Layout By Well**:
   - Each well unique from A01 to H12
   - For each well, an associated index name exists in the specified Index1Sequences section

10 If the supported indexing strategy specifies Single Index or Dual Index, you can set up **Default Layout By Column**:
   - Each column number unique from 1 to 12
   - For each column, an associated index name exists in the specified Index1Sequences section

11 If the supported indexing strategy specifies Dual Index, you can set up **Default Layout By Row**:
   - Each row letter unique from A to H
   - For each row, an associated index name exists in the specified Index2Sequences section

12 Click the **Choose.csv File** button to select and upload your custom index file.

13 Click **Create New Kit** to complete the process.
Your custom library prep is added to the library kit dropdown.

Import Samples and Libraries

To import libraries and associate them to new biological samples at the same time, perform the following steps.

1 Click the **Prep** icon.

2 Click **Libraries**.

3 Click the **Import** button.

4 If you have not generated an import file yet, click the template link and fill out the samples. Note the following items when filling out the template:
- The sample ID and sample name can only contain alphanumeric characters, dashes, or underscores. The sample ID must be unique and short. The sample name can be more descriptive to provide a human-readable identifier.
- The Species field is optional.
- The Project field is optional, but a project is required later because the output data are stored to the project.
- Fill out the Nucleic Acid column with DNA or RNA.

Figure 3 Import Sample Template

[Header]									
FileVersic	1								
LibraryPre	Nextera XT								
Container	Plate96								
Container	TestPlate4385								
Notes	This is a test plate for importing libraries								
[Data]									
SampleID	Name	Species	Project	NucleicAc	Well	Index1Na	Index1Sec	Index2Na	Index2Sequence
TestSamp	SampleNa	PhiX	TestImpor	DNA	A01	N701	TAAGGCG	S501	TAGATCGC
TestSamp	SampleNa	PhiX	TestImpor	DNA	B01	N701	TAAGGCG	S502	CTCTCTAT

5 Click the **Choose File** button.

6 Browse to the import file, and then click **Open**.

7 Click **Import**.

8 When finished, perform 1 of the following:
- If you only want to select the newly created samples, click the **Pool Libraries** button. Continue with *Pool Libraries* on page 33.
- If you want to select other libraries, click **Save & Continue Later**. This selection takes you back to the Libraries list, with the recently created sample at the top of the list.

Pool Libraries

The Pool Libraries page allows you to pool samples and sequence them in the same run, using the same analysis parameters.

1 Fill out the first pool ID. Pool ID must be unique.

2 If needed, you can create additional pools on the right by clicking the **+ Add Pool** button and filling out the pool IDs.
- Colors of the wells correspond to the colors of the pools.
- You can hover over the wells to see the library IDs.

3 Drag and drop individual samples from their well on the plate to a pool.
You can select multiple samples by holding **Shift**. To select multiple samples on Firefox or Internet Explorer, click the well twice.

4 If you want to pool libraries from multiple plates, use the Plate drop-down menu to specify the plate.
Alternatively, you can merge pools using the following steps.

   a Click **Save & Continue Later**. This selection takes you to the Pools list, with the recently created plate at the top of the list.
   b Select the checkboxes in the Pools list.
   c Click the **Merge Pools** button in the top navigation bar.

5 Click the **Plan Run** button.

Plan Runs

On the Planned Runs page, you can set up the parameters for the sequencing run on your NextSeq instrument.

1 Enter a name for your planned run.

2 [Optional] Enter the reagent barcode you plan to use, which links a reagent kit to this run.

3 Make sure that the **rehyb** checkbox is not selected. Currently, there is no supported rehyb workflow available for NextSeq.

4 Fill out the Enter Cycles section:
- Single- vs. paired-end
- Number of cycles per read

5 Verify the Review Indexes section for the indexing strategy. For indexing, it is set according to the index/library prep type chosen previously. If you choose to override this default indexing scheme, you are required to select the Index type (Single, Dual, or No Index). Make sure that you enter the number of index cycles accordingly. If you have selected multiple libraries, you cannot specify No Index.
BaseSpace Onsite automatically checks if the indexes chosen all start with 2 Gs; if so, it warns you to change your index strategy.

6 Verify the pool that is included in the planned run.

7 When your settings are complete, choose 1 of these options to continue:
- Click the **Sequence** button, which opens the Planned Runs list, and sets the state of the recently planned run to *Ready to Sequence*.
- Click the **Save & Continue Later** button, which opens the Planned Runs list, and sets the state of the recently planned run to *Planning*.

NOTE
A planned run must be in the *Ready to Sequence* state in order for it to show up in the Planned Runs list in the control software on the instrument.

8 If you want to change a planned run to the *Ready to Sequence* state, select the planned run from the list. Click the **Sequence** arrow link in the top navigation bar on the Planned Runs list page.

Your run now shows up in the Planned Runs list in the control software on your NextSeq sequencing system. Complete the run from your sequencing instrument. A sample sheet is not required. BaseSpace Onsite automatically generates FASTQ files when the sequencing run is complete.

NOTE
You can connect as many instruments as you have BaseSpace Onsite nodes installed, up to a maximum of 6.

View Files and Results

The following topics describe how to view files and results in BaseSpace Onsite.

View Files from a Run

BaseSpace Onsite gives you an option to view your run files or download them individually.

To view files, perform the following steps.

1 Click the **Runs** icon.

2 Click the desired run.

3 From the Run Overview Page, select the **Files** icon from the left navigation menu.

4 Select the desired file to view.

See *BaseSpace Onsite Files* on page 46 for a description of the available files.

View Indexing QC Page

The Indexing QC page lists count information for indexes used in the run. The Indexing QC is only available if the run is an index run.

By viewing the indexing QC results, you can see unexpected results for a sample with a particular index and use the information for troubleshooting purposes. You can also use this feature to confirm that all indexed samples were represented properly.

To view indexing QC results, perform the following steps.

1 Click the **Runs** icon.

2 Click the desired run.

3 There are 2 methods to go to the Indexing QC page:
- From the Run Overview page, click the Indexing QC link.
- From the Run Overview page, click the **Indexing QC** icon from the left navigation menu.

You can select the displayed lane through the drop-down list.

The first table provides an overall summary of the indexing performance for that lane, including the following information.

Total Reads	The total number of reads for this lane.
PF Reads	The total number of passing filter reads for this lane.
% Reads Identified (PF)	The total fraction of passing filter reads assigned to an index.
CV	The coefficient of variation for the number of counts across all indexes.
Min	The lowest representation for any index.
Max	The highest representation for any index.

Further information is provided regarding the frequency of individual indexes in both table and graph form. The table includes the following columns.

Index Number	A unique number assigned to each index by BaseSpace Onsite for display purposes.
Sample ID	The sample ID assigned to an index in the sample sheet.
Project	The project assigned to an index in the sample sheet.
Index 1 (I7)	The sequence for the first Index Read.
Index 2 (I5)	The sequence for the second Index Read.
% Reads Identified (PF)	The number of reads (only includes Passing Filter reads) mapped to this index.

This information is also displayed in graphical form. In the graphical display, indexes are ordered according to the unique Index Number assigned by BaseSpace Onsite.

View Run Charts

The Charts page shows charts with run metrics.

1 Click the **Runs** icon.

2 Click the desired run.

3 There are 2 methods to go to the Charts page:
- From the Run Overview page, click the Charts link.
- From the Run Overview page, click the **Charts** icon from the left navigation menu.

Flow Cell Chart

The Flow Cell chart shows color-coded graphical quality metrics per tile for the entire flow cell.

Use the Flow Cell chart to judge local differences per cycle, per lane, or per read in sequencing metrics on a flow cell. It is also an easy way to see the %Q30 metric, which is an excellent single metric to judge a run. Do not use the Flow Cell chart to look at downstream analysis metrics.

The Flow Cell chart has the following features:
- You can select the displayed metric, surface, cycle, and base through the drop-down lists.
- The color bar to the right of the chart indicates the values that the colors represent.
- The chart is displayed with tailored scaling by default.
- Tiles that have not been measured or are not monitored are gray.

You can monitor the following quality metrics using the Flow Cell chart.

Intensity	This chart shows the intensity by color and cycle of the 90% percentile of the data for each tile.
FWHM	The average full width of clusters at half maximum (in pixels). Used to display focus quality.

% Base	The percentage of clusters for which the selected base (A, C, T, or G) has been called.
%Q > 20, %Q > 30	The percentage of bases with a quality score of > 20 or > 30, respectively. These charts are generated after the 25th cycle, and the values represent the current scored cycle.
Median Q-Score	The median Q-Score for each tile over all bases for the current cycle. These charts are generated after the 25th cycle. This plot is best used to examine the Q-scores of your run as it progresses. Bear in mind that the %Q30 plot can give an over simplified view due to its reliance on a single threshold.
Density	The density of clusters for each tile (in thousands per mm^2).
Density PF	The density of clusters passing filter for each tile (in thousands per mm^2).
Clusters	The number of clusters for each tile (in millions).
Clusters PF	The number of clusters passing filter for each tile (in millions).
Error Rate	The calculated error rate, as determined by a spiked in PhiX control sample. If no PhiX control sample is run in the lane, this chart is not available.
% Phasing, % Prephasing.	The estimated percentage of molecules in a cluster for which sequencing falls behind (phasing) or jumps ahead (prephasing) the current cycle within a read.
% Aligned	The percentage of reads from clusters in each tile that aligned to the PhiX genome.
Perfect Reads	The percentage of reads that align perfectly, as determined by a spiked in PhiX control sample. If no PhiX control sample is run in the lane, this chart is all gray.
Corrected Intensity	The intensity corrected for cross talk between the color channels by the matrix estimation and phasing and prephasing.
Called Intensity	The intensity for the called base.
Signal to Noise	The signal to noise ratio is calculated as mean called intensity divided by standard deviation of noncalled intensities.

Note the variable scales used on these different parameters.

Data By Cycle Plot

The Data by Cycle plot shows the progression of quality metrics during a run as a line graph.

Use the Data By Cycle plot to judge the progression of quality metrics during a run on a cycle by cycle basis. Do not use the Data By Cycle plot to look at downstream analysis metrics, or aggregate analysis for a whole lane regardless of cycle.

The Data by Cycle plots have the following features:

- You can select the displayed metric and base through the drop-down lists.
- The symbol in the top right-hand corner toggles the plot between pane view and full screen view.

You can monitor the following quality metrics with this plot.

Intensity	This chart shows the intensity by color and cycle of the 90% percentile of the data for each tile.
FWHM	The average full width of clusters at half maximum (in pixels). Used to display focus quality.
% Base	The percentage of clusters for which the selected base (A, C, T, or G) has been called.
%Q > 20, %Q > 30	The percentage of bases with a quality score of > 20 or > 30, respectively. These charts are generated after the 25th cycle, and the values represent the current scored cycle.
Median Q-Score	The median Q-Score for each tile over all bases for the current cycle. These charts are generated after the 25th cycle. This plot is best used to examine the Q-scores of your run as it progresses. Bear in mind that the %Q30 plot can give an over simplified view due to its reliance on a single threshold.
Error Rate	The calculated error rate, as determined by a spiked in PhiX control sample. If no PhiX control sample is run in the lane, this chart is not available.
Perfect Reads	The percentage of reads that align perfectly, as determined by a spiked in PhiX control sample. If no PhiX control sample is run in the lane, this chart is all gray.
Corrected Intensity	The intensity corrected for cross talk between the color channels by the matrix estimation and phasing and prephasing.
Called Intensity	The intensity for the called base.
Signal to Noise	The signal to noise ratio is calculated as mean called intensity divided by standard deviation of noncalled intensities.

You can expand a chart by clicking the expand button  .

Q-score Distribution

The Q-score Distribution plot shows a bar graph that allows you to view the number of bases by quality score. The quality score is cumulative for current cycle and previous cycles, and only bases from reads that pass the quality filter are included.

Use it to judge the Q-score distribution for a run, which is an excellent indicator for run performance. Do not use the Q-score Distribution plot to look at metrics other than quality scores.

The Q-score Distribution pane shows plots that allow you to view the number of reads by quality score. The quality score is cumulative for current cycle and previous cycles, and only reads that pass the quality filter are included.

These plots have the following features:

- You can select the displayed read, and cycle through the drop-down lists.
- The symbol in the top right-hand corner toggles the plot between pane view and full screen view.

The Q-score is based on the Phred scale. The following list shows Q-scores and the corresponding chance that the base call is wrong:

- Q10—10% chance of wrong base call
- Q20—1% chance of wrong base call
- Q30—0.1% chance of wrong base call

- Q40—0.01% chance of wrong base call

You can slide the threshold (set at ≥ Q30 by default) to examine the proportion of bases at or above any particular Q-score. When using Q-score binning, this plot reflects the subset of Q-scores used.

Data by Lane Plot

The Data by Lane plots allow you to view quality metrics per lane.

Use the Data By Lane plot to judge the difference in quality metrics between lanes. Do not use the Data By Lane plot to look at alignment or variant calling analysis metrics.

The Data by Lane plots have the following features:

- You can select the displayed metric through the drop-down lists.
- The symbol in the top right-hand corner toggles the plot between pane view and full screen view.

The plots share several characteristics:

- The plots show the distribution of mean values for a given parameter across all tiles in a given lane.
- The red line indicates the median tile value for the parameter displayed.
- Blue boxes are for raw clusters, green boxes for clusters passing filter.
- The box outlines the interquartile range (the middle 50% of the data) for the tiles analyzed for the data point.
- The error bars delineate the minimum and maximum without outliers.
- The outliers are the values that are more than 1.5 times the interquartile range below the 25th percentile, or more than 1.5 times the interquartile range above the 75th percentile. Outliers are indicated as dots.

You can monitor the following quality metrics with this plot:

- The density of clusters for each tile (in thousands per mm^2).
- The number of clusters for each tile (in millions).
- The estimated percentage of molecules in a cluster for which sequencing falls behind (phasing) or jumps ahead (prephasing) the current cycle within a read.
- The percentage of reads from clusters in each tile that aligned to the PhiX genome.

You can expand a chart by clicking the expand button .

Q-score Heat Map

The Q-score heat map shows plots that allow you to view the Q-score by cycle.

View the Q-score heat map for a quick overview of the Q-scores over the cycles. Do not use the Q-score heat map to look at metrics other than quality scores.

These plots have the following features:

- The color bars to the right of each chart indicate the values that the colors represent. The charts are displayed with tailored scaling; the scale is always 0% to 100% of maximum value.
- The symbol in the top right-hand corner toggles the plot between pane view and full screen view.

You can expand a chart by clicking the expand button .

View Run Samples List

The Run Samples List contains a list of all the samples in the run.

Use this option when you want to see a list of all the samples in the run or to navigate to details regarding a specific sample.

To view the Run Samples List, perform the following steps.

1 Click the **Runs** icon.

2 Click the desired run.

3 There are 2 methods to go to the Run Samples List:
- From the Runs Overview page, click the Samples link.
- From the Runs Overview page, click the **Samples** icon from the left navigation menu.

You can now click a sample to see the sample overview; for more information, see *Sample Overview Page* on page 25.

View Run Summary

The Run Summary page has the overall statistics about the run.

Use this option when you want to view information about the run such as percent alignment, cycles, and densities.

To view the Run Summary page, perform the following steps.

1 Click the **Runs** icon.

2 Click the desired run.

3 There are 2 methods to go to the Run Summary:
- From the Run Overview Page, select Run Summary button.
- From the Run Overview Page, select the **Run Summary** icon from the left navigation menu.

The following metrics are displayed in the top table, split out by read and total.

Level	The sequencing read level.
Yield Total	The number of bases sequenced, which is updated as the run progresses.
Projected Total Yield	The projected number of bases expected to be sequenced at the end of the run.
Aligned	The percentage of the sample that aligned to the PhiX genome, which is determined for each level or read independently.
Error Rate	The calculated error rate of the reads that aligned to PhiX.
Intensity Cycle 1	The average of the A channel intensity measured at the first cycle averaged over filtered clusters. For the NextSeq 500 System, the red channel is used.
%Q ≥ 30	The percentage of bases with a quality score of 30 or higher, respectively. This chart is generated after the 25th cycle, and the values represent the current cycle.

The following metrics are available in the Read tables, split out by lane.

Tiles	The number of tiles per lane.

Density	The density of clusters (in thousands per mm^2) detected by image analysis, +/- 1 standard deviation.
Clusters PF	The percentage of clusters passing filtering, +/- 1 standard deviation.
Phas./Prephas.	The value used by RTA for the percentage of molecules in a cluster for which sequencing falls behind (phasing) or jumps ahead (prephasing) the current cycle within a read. For MiSeq and NextSeq, RTA generates phasing and prephasing estimates empirically for every cycle. The value displayed here is therefore not used in the actual phasing/prephasing calculations, but is an aggregate value determined from the first 25 cycles. For most applications, the value reported is very close to the value that is applied. For low diversity samples or samples with unbalanced base composition, the reported value can diverge from the values being applied because the value changes from cycle to cycle.
Reads	The number of clusters (in millions).
Reads PF	The number of clusters (in millions) passing filtering.
%Q ≥ 30	The percentage of bases with a quality score of 30 or higher, respectively. This chart is generated after the 25th cycle, and the values represent the current cycle.
Yield	The number of bases sequenced which passed filter.
Cycles Err Rated	The number of cycles that have been error-rated using PhiX, starting at cycle 1.
Aligned	The percentage that aligned to the PhiX genome.
Error Rate	The calculated error rate, as determined by the PhiX alignment. Subsequent columns display the error rate for cycles 1–35, 1–75, and 1–100.
Intensity Cycle 1	The average of the A channel intensity measured at the first cycle averaged over filtered clusters.
%Intensity Cycle 20	The corresponding intensity statistic at cycle 20 as a percentage of that value at the first cycle. 100%x(Intensity at cycle 20)/(Intensity at cycle 1).

View the Project Sample List

The Project Sample List contains the list of samples in a project. Use this option when you want to see a list of all the samples in the project or navigate to details regarding a specific sample.

To view the Project Sample List, perform the following steps.

1 Click the **Projects** icon.

2 Click the desired project.

3 Click the **Samples** link from the left navigation menu.

You can now click a sample to see the sample overview; for more information, see *Sample Overview Page* on page 25.

View the Analyses List

The Analyses List contains a list of app sessions in a project. Use this option when you want to navigate to details regarding a specific app session.

To view the Analyses List, perform the following steps.

1 Click the **Projects** icon.

2 Click the desired project.

You can now click an Analysis to see the results; for more information, see *Analyses Page* on page 25.

Analyze Samples Further

The following topics describe how to analyze samples further in BaseSpace Onsite.

Launch Apps

You can launch apps that perform additional analysis, visualization, or annotation of data.

To launch an app, perform the following steps.

1 There are 2 ways to start an app:
- Navigate to the project, sample, or analysis that you want to run the app on, click the **Launch Apps** button, and select the desired app from the drop-down list.
- Go to the Apps button, select the desired app from the list and click **Launch**. If you are running BWA Enrichment v2.1 or Isaac Enrichment v2.1, read *Enrichment v2.1 Apps Adapter Trimming* on page 44.

2 Read the End-User License Agreement and permissions, and then click **Accept**.

The app now guides you through the start-up process. BaseSpace Onsite has limited storage capacity, and checks the free space available before starting an app. If there is not enough available space, BaseSpace Onsite displays an error message. For more information, see *Storage Check* on page 68.

Launch the IGV App

The Integrative Genomics Viewer (IGV) of the Broad Institute is a fully featured genome browser that allows you to visualize your sequence data in great detail. Illumina has modified IGV to display alignment and variant data from BaseSpace Onsite (BAM and VCF files).

IGV enables you to perform variant analysis after launching Resequencing or Amplicon workflows in BaseSpace Onsite. IGV is run on a project, which is the highest level directory and contains one or more AppResults. IGV retains all its native functions, including loading data from your local computer.

NOTE
Make sure that the Java Runtime Environment is installed on the computer in order for IGV to work properly. Download Java here: java.com/en/.

Perform the following steps to launch the IGV app.

1 Click the **Projects** icon.

2 Click desired project.

3 Click the **Launch Apps** button and select the IGV application from the drop-down list.

4 Select the **Accept** button.

5 Depending on your browser, it asks you to open or save the *.jnlp file.
- For Internet Explorer, click the **Open** button.
- For Chrome, click the **Keep** button and then click file to open.
- For Firefox, select the **Open with Java(TM) Web Start Launcher (default)** option.

The IGV App opens on your desktop with the requested project loaded.

BaseSpace Onsite Data in IGV

The BaseSpace Onsite file browser shows data in BaseSpace Onsite that is available for viewing in IGV. The directory structure shown is according to how data are organized in BaseSpace Onsite.

A project is the highest level directory and it contains one or more AppResults. If an AppResult was the result of analyzing a single sample, then the sample name is appended to the AppResult name. Each AppResult contains zero or more files.

The file browser shows alignment (BAM) and variant (VCF) files, and, if produced, BED, GTF, and BEDGRAPH files. Double-click a file to load it as an IGV track. First load VCF files before BAM files because read tracks can take up an entire IGV screen, which requires scrolling to see variants.

Additional Reference Genomes

IGV contains several installed reference genomes:
- *Homo sapiens*: Human hg19
- *Mus musculus*: Mouse mm9
- *Saccharomyces cerevisiae*: *S. cerevisiae* (sacCer2)
- *Arabidopsis thaliana*: *A. thaliana* (TAIR10)

Run the VariantStudio App

The Illumina VariantStudio data analysis software application enables researchers to identify and classify disease-relevant variants quickly, and then communicate significant findings in concise and actionable reports.

This application provides an intuitive framework for nonexpert users and offers the following features:
- Flexible filtering options
- Streamlined variant classification
- Rapid and rich annotations
- Customizable reporting

Use VariantStudio if you want to perform the following:
- Explore and isolate key variants
- Categorize variants and determine biological impact

To launch the VariantStudio app, perform the following steps.

1 Click the **Apps** button.

2 Select the VariantStudio app from the list and click **Launch**.

3 Select the project you want to run the app on. You can only select projects you own.

4 Click **Continue**.

5 If you use the app for the first time, install VariantStudio:

   a Click the **Install VariantStudio** button.
   b Run the setup.exe file. Your web browser can ask you to save the file first. After the download has completed, double-click the setup file.
   c If you are prompted with a security warning, click **Install**.

6 Click the **Launch VariantStudio**.
   The VariantStudio application opens on your desktop with the requested project loaded. For instructions on how to run VariantStudio, see the *VariantStudio User Guide*.

Enrichment v2.1 Apps Adapter Trimming

BWA Enrichment v2.1 and Isaac Enrichment v2.1 Apps do not support adapter trimming in this BaseSpace Onsite release. Here are some notes on the use of these apps:

- **NextSeq**—You do not need to use adapter trimming for NextSeq data in any of the secondary analysis apps. Adapter trimming is performed automatically in FASTQ generation based on the library prep kit type selected during run setup in the Prep Tab workflow.
- **HiSeq**—You can go back to the uploaded run and requeue FASTQ generation by fixing the sample sheet. To fix the sample sheet, add the adapter sequences to the Settings subsection, and then requeue the analysis. For more information, see *Fix Sample Sheet / Rerun Workflow* on page 63.
- **MiSeq**—You can only upload MiSeq FASTQ files through BaseSpace Onsite Import. Trim the adapters from the FASTQ files before import.

Download Files

The following topics describe how to download files in BaseSpace Onsite.

Download Individual Files

BaseSpace Onsite allows you to download data as a package, individually, or as a group of FASTQ files. Use this option when you want to download individual files, and do not need all files for a run, sample, project, or analysis.

To download multiple FASTQ files, perform the following steps.

1 Click the **Runs** icon or **Projects** icon.

2 Navigate to the file you want to download

3 Click the file.

4 Click the **Download** button.
   BaseSpace Onsite now downloads the files to the desired location.

Download Multiple FASTQ Files

BaseSpace Onsite allows you to download data as a package, individually, or as a group of FASTQ files. If you want to download FASTQ files per sample, use this option to save time.

To download multiple FASTQ files, perform the following steps.

1 Click the **Runs** icon or **Projects** icon.

2 Click the desired run or project.

3 Click the desired sample in the Samples pane.

4 In the Files pane, select the checkboxes for the desired FASTQ files.

5 Click the **Download Selected** button.
The BaseSpace Onsite Downloader is not signed, which can lead to the following warnings.
- A Windows popup appears when trying to install the setup.exe. Choose **Run anyway** to install the BaseSpace Onsite Downloader.
- Antivirus software can flag the file. Make sure that you allow this file to be installed.

The BaseSpace Onsite Downloader guides you through the download process, and starts the download of the files to the desired location.

Download Run File Package

BaseSpace Onsite allows you to download data as a package, individually, or as a group of FASTQ files. This topic describes how to download a package of files in a run.

The packages available depend on your workflow; packages that are grayed out are not available for download. There are 4 types of data packages:
- Variant Data, containing VCF files with variant calls.
- Aligned Data, containing BAM files with aligned reads.
- Unaligned Data, containing FASTQ files with unaligned reads.
- SAV Data, containing files describing the set-up of the run and InterOp files.

For more information about file types, see *BaseSpace Onsite Files* on page 46.

To download a run file package, perform the following steps.

1 Click the **Runs** icon.

2 Click the desired run.

3 Click **Download Run**.

4 If not yet installed, click **Install the Downloader** to install BaseSpace Onsite Downloader.
BaseSpace Onsite Downloader is a desktop client that allows for fast, reliable downloads transferred securely over SSL.

5 Select the files to be downloaded.

6 Click **Download your files**.
BaseSpace Onsite Downloader is launched.

7 In the Confirm Download dialog, browse to a download folder, and then click **Start Download**.
BaseSpace Onsite Downloader tracks the progress of the download.

BaseSpace Onsite Files

BaseSpace Onsite uses and produces various files. See the topics in this section for details.

BAM Files

What is it?

The Sequence Alignment/Map (SAM) format is a generic alignment format for storing read alignments against reference sequences, supporting short and long reads (up to 128 Mb) produced by different sequencing platforms. SAM is a text format file that is human-readable. The Binary Alignment/Map (BAM) keeps the same information as SAM, but in a compressed, binary format that is only machine readable.

When to use it

Allows you to see alignments. Use it for direct interpretation or as a starting point for further analysis with downstream analysis tools that are compatible with BAM. BAM files are suitable for viewing with an external viewer such as IGV or the UCSC Genome Browser.

When not to use it

Do not use it with tools that are not compatible with the BAM format. Do not use with applications that cannot handle large files, as BAM files can get large, depending on the application and data.

How to use it

If you use an app in BaseSpace that uses BAM files as input, the app locates the file when launched. If using BAM files in other tools, download the file to use it in the external tool.

Detailed Description

Go to samtools.sourceforge.net/SAM1.pdf to see the exact SAM specification.

gVCF Files

What is it?

This application also produces the Genome Variant Call Format file (gVCF). gVCF was developed to store sequencing information for both variant and nonvariant positions, which is required for human clinical applications. gVCF is a set of conventions applied to the standard variant call format (VCF) 4.1 as documented by the 1000 Genomes Project. These conventions allow representation of genotype, annotation, and other information across all sites in the genome in a compact format. Typical human whole-genome sequencing results expressed in gVCF with annotation are less than 1 Gbyte, or about 1/100 the size of the BAM file used for variant calling. If you are performing

targeted sequencing, gVCF is also an appropriate choice to represent and compress the results.

gVCF is a text file format, stored as a gzip compressed file (*.genome.vcf.gz). Compression is further achieved by joining contiguous nonvariant regions with similar properties into single 'block' VCF records. To maximize the utility of gVCF, especially for high stringency applications, the properties of the compressed blocks are conservative. Block properties like depth and genotype quality reflect the minimum of any site in the block. The gVCF file can be indexed (creating a *.tbi file) and used with existing VCF tools such as tabix and IGV, making it convenient both for direct interpretation and as a starting point for further analysis.

For more information, see sites.google.com/site/gvcftools/home/about-gvcf.

When to use it

Use it for direct interpretation or as a starting point for further analysis with downstream analysis that is compatible with gVCF, such as tabix and IGV.

When not to use it

Do not use it with tools that are not compatible with the gVCF format.

How to use it

Apps that use gVCF files find it when kicked off and directed to the sample. If using gVCF files in other tools, download the file to use it in the outside tool.

Detailed Description

The following conventions are used in the variant caller gVCF files.

Samples per File

There is only one sample per gVCF file.

Nonvariant Blocks Using END Key

Contiguous nonvariant segments of the genome can be represented as single records in gVCF. These records use the standard 'END' INFO key to indicate the extent of the record. Even though the record can span multiple bases, only the first base is provided in the REF field to reduce file size.

The following is a simplified segment of a gVCF file, describing a segment of nonvariant calls (starting with an A) on chromosome 1 from position 51845 to 51862.

```
##INFO=<ID=END,Number=1,Type=Integer,Description="End position
   of the variant described in this record">#CHROM POS ID REF
   ALT QUAL FILTER INFO FORMAT NA19238chr1 51845 . A . . PASS
   END=51862
```

Any field provided for a block of sites, such as read depth (using the DP key), shows the minimum value that is observed among all sites encompassed by the block. Each sample value shown for the block, such as the depth (DP), is restricted to a range where the maximum value is within 30% or 3 of the minimum. For example, for sample value range [x,y], y <= x+max(3,x*0.3). This range restriction applies to each of the sample values printed in the final block record.

Indel Regions

Sites that are "filled in" inside deletions have additional changes:

All deletions:

- Sites inside any deletion are marked with the deletion filters, in addition to any filters that have already been applied to the site.
- Sites inside deletions cannot have a genotype or alternate allele quality score higher than the corresponding value from the enclosing indel.

Heterozygous deletions:

- Sites inside heterozygous deletions are altered to have haploid genotype entries (eg, "0" instead of "0/0", "1" instead of "1/1").
- Heterozygous SNV calls inside heterozygous deletions are marked with the "SiteConflict" filter and their genotype is unchanged.

Homozygous deletions:

- Homozygous reference and no-call sites inside homozygous deletions have genotype "."
- Sites inside homozygous deletions that have a nonreference genotype are marked with a “SiteConflict” filter, and their genotype is unchanged.
- Site and genotype quality are set to "."

The described modifications reflect the notion that the site confidence is bound within the enclosing indel confidence.

On occasion, the variant caller produces multiple overlapping indel calls that cannot be resolved into 2 haplotypes. If this case, all indels and sites in the region of the overlap are marked with the *IndelConflict* filter.

Genotype Quality for Variant and Nonvariant Sites

The gVCF file uses an adapted version of genotype quality for variant and nonvariant site filtration. This value is associated with the key GQX. The GQX value is intended to represent the minimum of {Phred genotype quality assuming the site is variant, Phred genotype quality assuming the site is nonvariant}. The reason for using this value is to allow a single value to be used as the primary quality filter for both variant and nonvariant sites. Filtering on this value corresponds to a conservative assumption appropriate for applications where reference genotype calls must be determined at the same stringency as variant genotypes, ie:

- An assertion that a site is homozygous reference at GQX >= 30 is made assuming the site is variant.
- An assertion that a site is a nonreference genotype at GQX >= 30 is made assuming the site is nonvariant.

Section Descriptions

The gVCF file contains the following sections:

- Metainformation lines start with ## and contain metadata, config information, and define the values that the INFO, FILTER, and FORMAT fields can have.
- The header line starts with # and names the fields that the data lines use. These fields are #CHROM, POS, ID, REF, ALT, QUAL, FILTER, INFO, FORMAT, followed by one or more sample columns.
- Data lines that contain information about one or more positions in the genome.

If you extract the variant lines from a gVCF file, you produce a conventional variant VCF file.

Field Descriptions

The fixed fields #CHROM, POS, ID, REF, ALT, QUAL are defined in the VCF 4.1 standard provided by the 1000 Genomes Project. The fields ID, INFO, FORMAT, and sample are described in the metainformation.

- **CHROM**: Chromosome: an identifier from the reference genome or an angle-bracketed ID String ("<ID>") pointing to a contig.
- **POS**: Position: The reference position, with the first base having position 1. Positions are sorted numerically, in increasing order, within each reference sequence CHROM. There can be multiple records with the same POS. Telomeres are indicated by using positions 0 or N+1, where N is the length of the corresponding chromosome or contig.
- **ID**: Semicolon separated list of unique identifiers where available. If this ID is a dbSNP variant, it is encouraged to use the rs number. No identifier is present in more than 1 data record. If there is no identifier available, then the missing value is used.
- **REF**: Reference bases: A,C,G,T,N; there can be multiple bases. The value in the POS field refers to the position of the first base in the string. For simple insertions and deletions in which either the REF or 1 of the ALT alleles would otherwise be null/empty, the REF and ALT strings include the base before the event. This modification is reflected in the POS field. The exception is when the event occurs at position 1 on the contig, in which case they include the base after the event. If any of the ALT alleles is a symbolic allele (an angle-bracketed ID String "<ID>"), the padding base is required. In that case, POS denotes the coordinate of the base preceding the polymorphism.
- **ALT**: Comma-separated list of alternate nonreference alleles called on at least 1 of the samples. Options are:
  - Base strings made up of the bases A,C,G,T,N
  - Angle-bracketed ID String (”<ID>”)
  - Break-end replacement string as described in the section on break-ends.
- If there are no alternative alleles, then the missing value is used.
- **QUAL**: Phred-scaled quality score for the assertion made in ALT. ie -10log_10 probability (call in ALT is wrong). If ALT is ”.” (no variant), this score is -10log_10 p (variant). If ALT is not ”.”, this score is -10log_10 p(no variant). High QUAL scores indicate high confidence calls. Although traditionally people use integer Phred scores, this field is permitted to be a floating point to enable higher resolution for low confidence calls if desired. If unknown, the missing value is specified. (Numeric)
- **FILTER**: PASS if this position has passed all filters, ie a call is made at this position. Otherwise, if the site has not passed all filters, a semicolon-separated list of codes for filters that fail. gVCF files use the following values:
  - *PASS*: position has passed all filters.
  - *IndelConflict*: Locus is in region with conflicting indel calls.
  - *SiteConflict*: Site genotype conflicts with proximal indel call, which is typically a heterozygous SNV call made inside a heterozygous deletion.
  - *LowGQX*: Locus GQX (minimum of {Genotype quality assuming variant position,Genotype quality assuming nonvariant position}) is less than 30 or not present.
  - *HighDPFRatio*: The fraction of base calls filtered out at a site is greater than 0.3.
  - *HighSNVSB*: SNV strand bias value (SNVSB) exceeds 10. High strand bias indicates a potential high false-positive rate for SNVs.
  - *HighSNVHPOL*: SNV contextual homopolymer length (SNVHPOL) exceeds 6.

- *HighREFREP*: Indel contains an allele that occurs in a homopolymer or dinucleotide track with a reference repeat greater than 8.
- *HighDepth*: Locus depth is greater than 3x the mean chromosome depth.

- **INFO**: Additional information. INFO fields are encoded as a semicolon-separated series of short keys with optional values in the format: <key>=<data>[,data]. gVCF files use the following values:
  - *END*: End position of the region described in this record.
  - *BLOCKAVG_min30p3a*: nonvariant site block. All sites in a block are constrained to be nonvariant, have the same filter value, and have all sample values in range [x,y], y <= max(x+3,(x*1.3)). All printed site block sample values are the minimum observed in the region spanned by the block.
  - *SNVSB*: SNV site strand bias.
  - *SNVHPOL*: SNV contextual homopolymer length.
  - *CIGAR*: CIGAR alignment for each alternate indel allele.
  - *RU*: Smallest repeating sequence unit extended or contracted in the indel allele relative to the reference. If longer than 20 bases, RUs are not reported.
  - *REFREP*: Number of times RU is repeated in reference.
  - *IDREP*: Number of times RU is repeated in indel allele.
- **FORMAT**: Format of the sample field. FORMAT specifies the data types and order of the subfields. gVCF files use the following values:
  - *GT*: Genotype.
  - *GQ*: Genotype Quality.
  - *GQX*: Minimum of {Genotype quality assuming variant position, Genotype quality assuming nonvariant position}.
  - *DP*: Filtered base call depth used for site genotyping.
  - *DPF*: Base calls filtered from input before site genotyping.
  - *AD*: Allelic depths for the ref and alt alleles in the order listed. For indels, this value only includes reads that confidently support each allele (posterior probability 0.999 or higher that read contains indicated allele vs all other intersecting indel alleles).
  - *DPI*: Read depth associated with indel, taken from the site preceding the indel.
- **SAMPLE**: Sample fields as defined by the header.

VCF Files

What is it?

VCF is a text file format that contains information about variants found at specific positions in a reference genome. The file format consists of metainformation lines, a header line, and then data lines. Each data line contains information about a single variant.

When to use it

Use it for direct interpretation or as a starting point for further analysis with downstream analysis that is compatible with VCF, such as IGV or the UCSC Genome Browser.

When not to use it

Do not use it with tools that are not compatible with the VCF format.

NOTE
Windows recognizes VCF files as an Outlook contact file. Do not open VCF files in Outlook.

How to use it

If you use an app in BaseSpace that uses VCF files as input, the app locates the file when launched. If using VCF files in other tools, download the file to use it in the external tool.

Detailed Description

The file naming convention for VCF files is as follows: SampleName_S#.vcf (where # is the sample number determined by ordering in the sample sheet).

The header of the VCF file describes the tags used in the remainder of the file and has the column header:

```
##fileformat=VCFv4.1
##fileDate=20120317
##source=SequenceAnalysisReport.vshost.exe
##reference=
##phasing=none
##INFO=<ID=DP,Number=1,Type=Integer,Description="Total Depth">
##INFO=<ID=TI,Number=.,Type=String,Description="Transcript ID">
##INFO=<ID=GI,Number=.,Type=String,Description="Gene ID">
##INFO=<ID=CD,Number=0,Type=Flag,Description="Coding Region">
##FILTER=<ID=q20,Description="Quality below 20">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
##FORMAT=<ID=GQ,Number=1,Type=Integer,Description="Genotype
   Quality">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT SAMPLE
```

A sample line of the VCF file, with the data that is used to populate each column described:

```
chr22 16285888 rs76548004 T C 17 d15;q20 DP=11;TI=NM_
   001136213;GI=POTEH;CD GT:GQ 1/0:17
```

Setting	Description
ALT	The alleles that differ from the reference read. For example, an insertion of a single T could show reference A and alternate AT.
CHROM	The chromosome of the reference genome. Chromosomes appear in the same order as the reference FASTA file (generally karyotype order)
FILTER	If all filters are passed, the' PASS' is written. The possible filters are as follows: • q20 – The variant score is less than 20. (Configurable using the VariantFilterQualityCutoff setting in the config file) • r8 – For an Indel, the number of repeats in the reference (of a 1- or 2-base repeat) is greater than 8. (Configurable using the IndelRepeatFilterCutoff setting in the config file)

Setting	Description
FORMAT	The format column lists fields (separated by colons), for example, "GT:GQ". The list of fields provided depends on the variant caller used. The available fields are as follows: AD – Entry of the form X,Y where X is the number of reference calls, Y the number of alternate calls GQ – Genotype quality GT – Genotype. 0 corresponds to the reference base, 1 corresponds to the first entry in the ALT column, 2 corresponds to the second entry in the ALT column, and so on. The '/' indicates that there is no phasing information. NL – Noise level; an estimate of base calling noise at this position SB – Strand bias at this position. Larger negative values indicate more bias; values near 0 indicate little strand bias. VF – Variant frequency. The percentage of reads supporting the alternate allele.
ID	The rs number for the SNP obtained from dbSNP. If there are multiple rs numbers at this location, the list is semicolon delimited. If no dbSNP entry exists at this position, the missing value ('.') is used.
INFO	The possible entries in the INFO column: • AD – Entry of the form X,Y where X is the number of reference calls, Y the number of alternate calls. • CD – A flag indicating that the SNP occurs within the coding region of at least 1 RefGene entry • DP – The depth (number of base calls aligned to this position) • GI – A comma-separated list of gene IDs read from RefGene • NL – Noise level; an estimate of base calling noise at this position. • TI – A comma-separated list of transcript IDs read from RefGene • SB – Strand bias at this position. • VF – Variant frequency. The number of reads supporting the alternate allele.
POS	The 1-based position of this variant in the reference chromosome. The convention for VCF files is that, for SNPs, this base is the reference base with the variant. For indels or deletions, this base is the reference base **immediately before** the variant. Variants are in order of position.
QUAL	A Phred-scaled quality score assigned by the variant caller. Higher scores indicate higher confidence in the variant (and lower probability of errors). For a quality score of Q, the estimated probability of an error is 10-(Q/10). For example, the set of Q30 calls has a 0.1% error rate. Many variant callers assign quality scores (based on their statistical models) which are high relative to the error rate observed in practice.
REF	The reference genotype. For example, a deletion of a single T can read reference TT and alternate T.
SAMPLE	The sample column gives the values specified in the FORMAT column. One MAXGT sample column is provided for the normal genotyping (assuming the reference). For reference, a second column is provided for genotyping assuming the site is polymorphic. See the Starling documentation for more details.

FASTQ Files

What is it?

BaseSpace Onsite converts *.bcl files into FASTQ files, which contain base call and quality information for all reads passing filtering.

When to use it.

FASTQ files can be used as sequence input for alignment and other secondary analysis software.

When not to use it.

Do not use it with tools that are not compatible with the FASTQ format.

How to use it

BaseSpace Onsite automatically generates FASTQ files in sample sheet-driven workflow apps. Other apps that perform alignment and variant calling also automatically use FASTQ files.

Naming

FASTQ files are named with the sample name and the sample number, which is a numeric assignment based on the order that the sample is listed in the sample sheet. Example:

Data\Intensities\BaseCalls\SampleName_S1_L001_R1_001.fastq.gz

- **SampleName**—The sample name provided in the sample sheet. If a sample name is not provided, the file name includes the sample ID, which is a required field in the sample sheet and must be unique.
- **S1**—The sample number based on the order that samples are listed in the sample sheet starting with 1. In this example, S1 indicates that this sample is the first sample listed in the sample sheet.

  NOTE
  Reads that cannot be assigned to any sample are written to a FASTQ file for sample number 0, and excluded from downstream analysis.

- **L001**—The lane number.
- **R1**—The read. In this example, R1 means Read 1. For a paired-end run, there is at least one file with R2 in the file name for Read 2. When generated, index reads are **I1** or **I2**.
- **001**—The last segment is always 001.

Compression

FASTQ files are saved compressed in the GNU zip format (an open source file compression program), indicated by the .gz file extension.

Format

Each entry in a FASTQ file consists of 4 lines:

- Sequence identifier
- Sequence
- Quality score identifier line (consisting only of a +)
- Quality score

```
@<instrument>:<run number>:<flowcell ID>:<lane>:<tile>:<x-
   pos>:<y-pos>:<UMI> <read>:<is filtered>:<control
   number>:<index>
```

The following table describes the elements.

Element	Requirements	Description
@	@	Each sequence identifier line starts with @.
<instrument>	Characters allowed: a–z, A–Z, 0–9 and underscore	Instrument ID.
<run number>	Numerical	Run number on instrument.
<flowcell ID>	Characters allowed: a–z, A–Z, 0–9	
<lane>	Numerical	Lane number.
<tile>	Numerical	Tile number.
<x_pos>	Numerical	X coordinate of cluster.
<y_pos>	Numerical	Y coordinate of cluster.
<UMI>	Restricted characters: A/T/G/C/N	Optional, appears when UMI is specified in sample sheet. UMI sequences for Read 1 and Read 2, seperated by a plus [+].
<read>	Numerical	Read number. 1 can be single read or Read 2 of paired-end.
<is filtered>	Y or N	Y if the read is filtered (did not pass), N otherwise.
<control number>	Numerical	0 when none of the control bits are on, otherwise it is an even number.
<index>	Restricted characters: A/T/G/C/N	Index of the read.

An example of a valid entry is as follows; note the space preceding the read number element:

```
@SIM:1:FCX:1:15:6329:1045:GATTACT+GTCTTAAC 1:N:0:ATCCGA
TCGCACTCAACGCCCTGCATATGACAAGACAGAATC
+
<>;##=><9=AAAAAAAAA9#:<#<;<<<????#=
```

Quality Scores

A quality score (or Q-score) expresses an error probability. In particular, it serves as a convenient and compact way to communicate small error probabilities.

Given an assertion, A, the quality score, Q(A), expresses the probability that A is not true, P(~A), according to the relationship:

$$Q(A) = -10 \log_{10}(P(\sim A))$$

where *P(~A)* is the estimated probability of an assertion A being wrong.

The relationship between the quality score and error probability is demonstrated with the following table.

Quality score, Q(A)	Error probability, P(~A)
10	0.1
20	0.01
30	0.001

NOTE
On the systems we currently support, Q-scores are automatically binned. The specific binning applied depends on the current Q-table. See the white paper *Reducing Whole Genome Data Storage Footprint* for more information, available from www.illumina.com.

Quality Scores Encoding

In FASTQ files, quality scores are encoded into a compact form, which uses only 1 byte per quality value. In this encoding, the quality score is represented as the character with an ASCII code equal to its value + 33. The following table demonstrates the relationship between the encoding character, its ASCII code, and the quality score represented.

NOTE
When Q-score binning is in use, the subset of Q-scores applied by the bins is displayed.

Table 1 ASCII Characters Encoding Q-scores 0–40

Symbol	ASCII Code	Q-Score	Symbol	ASCII Code	Q-Score
!	33	0	6	54	21
"	34	1	7	55	22
#	35	2	8	56	23
$	36	3	9	57	24
%	37	4	:	58	25
&	38	5	;	59	26
'	39	6	<	60	27
(	40	7	=	61	28
)	41	8	>	62	29
*	42	9	?	63	30
+	43	10	@	64	31
,	44	11	A	65	32
-	45	12	B	66	33
.	46	13	C	67	34
/	47	14	D	68	35
0	48	15	E	69	36
1	49	16	F	70	37
2	50	17	G	71	38
3	51	18	H	72	39
4	52	19	I	73	40
5	53	20			

Health Runs

A user can choose whether to send anonymous system health information to Illumina. Health runs help Illumina diagnose issues and improve our products. The information consists of InterOp files and log files, and is not tied to any user account. This option is on by default.

Share Data

Data in BaseSpace Onsite can be shared with collaborators in a couple of different ways. You can either share data at a run or project level, via an email invitation or through a hyperlink. With the email invitation option, only the accounts with the specified email can view shared data. Sharing via a hyperlink option allows anyone with access to the hyperlink to be able to view the shared data, as long as the hyperlink is still active.

You can only share data to collaborators who have an account on the same BaseSpace Onsite.

Sharing is for read-only access. If you want a collaborator to have write access, see *Transfer Owner* on page 61.

NOTE

Runs and projects have separate permissions. If you share a run, the project associated with that run is not shared automatically. Therefore, samples and app results are not accessible to collaborators of the run.

The following topics describe how to share data.

Share a Project using the Get Link option

Sharing using the Get Link option allows you to share a project or a run with any collaborator who is in your company, and has an account on the BaseSpace Onsite system. The hyperlink can be turned on or off by setting the activate or deactivate option. Anyone can access the project or run when the link is activated. Furthermore, anyone who previously accepted the link still has access to the run even when the link is deactivated.

Use the Get Link option when you do not want to assign a project to a specific person. However, do not use this option if you want to confine the list of who has access to the project.

NOTE

If you want more control, use the email share option to specify who can view the project. See*Share a Project Using the Email Option* on page 57.

To use the Get Link option, perform the following steps.

1. Click the **Projects** icon.
2. Click the desired project.
3. Click the **Get Link** button.
4. Click the **Activate** button.
5. Copy the URL to share with collaborators.

The link is active until the Deactivate option is selected. To deactivate a link, perform the following steps.

6 Navigate to the shared item.

7 Click the **Get Link** button.

8 Click the **Deactivate** button.

Share a Project Using the Email Option

Use the Share Project option to share a Project using an email link. The specified collaborators receive an email with a link to the Project and only that person can view the corresponding data.

The email option allows greater control over who can view your data. Sharing using the Get Link options gives anyone access to your data, as long as the link is left activated.

To email a link, perform the following steps.

1 Click the **Projects** icon.

2 Click the desired project.

3 Click **Share Project**.

4 In the Share Settings dialog box, enter the collaborators email address, and then click **Invite**.

NOTE
The invitation email address must match your BaseSpace Onsite login email address. Otherwise, your collaborator is not able to view the project.

5 Click **Save Settings**.

Share a Run with Get Link

Sharing using the Get Link option allows you to share a run with any collaborator who has access to the link. The hyperlink can be turned on or off by setting the activate or deactivate option. Anyone can access the project or run when the link is activated. Furthermore, anyone who previously accepted the link still has access to the run even though the link is deactivated.

Sharing runs with the Get Link option is similar to sharing projects with the Get Link option.

NOTE
If you want more control, use the email share option to specify who can view the project. See *Share a Run Using the Email Option* on page 58.

To share a run using the Get Link option, perform the following steps.

1 Click the **Runs** icon.

2 Click the desired run.

3 Click the **More** button and select the Get Link option.

4 Click the **Activate** button.

5 Copy the URL to share with collaborators.

The link is active until the Deactivate option is selected. To deactivate a link, perform the following steps.

6 Navigate to the run

7 Click the **Get Link** button.

8 Click the **Deactivate** button.

NOTE
Runs and projects have separate permissions. If you share a run, the project associated with that run is not shared automatically. Therefore, samples and app results are not accessible to collaborators of the run.

Share a Run Using the Email Option

Use the Share Project option to share a Run with a specified collaborator via an email link. The specified collaborators receive an email with a link to the Run and only that person can view the corresponding data.

The email option allows greater control over who can view your data. Sharing using the Get Link options gives anyone access to your data, as long as the link is left activated.

To email a link, perform the following steps.

1 Click the **Runs** icon.

2 Click the desired run.

3 Click **Share**.

4 In the Share Settings dialog box, enter the collaborators email address, and then click **Invite**.

NOTE
The invitation email address must match your BaseSpace Onsite login email address. Otherwise, your collaborator is not able to view the project.

5 Click **Save Settings**.

NOTE
Runs and projects have separate permissions. If you share a run, the project associated with that run is not shared automatically. Therefore, samples and app results are not accessible to collaborators of the run.

Project and Sample Management

The following topics describe how to manage projects and samples in BaseSpace Onsite.

- *Combine Samples* on page 59
- *Transfer Owner* on page 61

Edit Project Details

Use the Edit Project option to change details regarding the project, such as the description or project name.

To edit a project, perform the following steps.

1 Click the **Projects** icon.

2 Click the desired project.

3 Click **Edit Project**.

4 Change project details in the Edit Project dialog box.

5 Click **Save**.

Set Up a New Project

The following list includes reasons to set up a new project.

- To analyze a sample in the context of 2 different projects
- To transfer ownership of samples to a collaborator, but still keep a copy yourself
- To split a project into multiple projects

To set up a new project, perform the following steps.

1. Click the **Projects** icon.
2. Click **New Project** in the top left corner.
3. Enter a new name and description.
4. Click **Create**.

To copy samples into the new project, see *Copy Samples* on page 59.

Combine Samples

Use the following method to combine the data from 2 or more different sequencing runs on the same sample. You can only combine samples that have the same read lengths.

To combine samples, perform the following steps.

1. Click the **Projects** icon.
2. Click the desired project.
3. Click the **Samples** link from the left navigation menu.
4. Select the checkboxes of the samples you want to combine.
5. Click **Combine**.
6. Click **Combine** in the pop-up screen.

Copy Samples

You can copy samples from a project to another project for the following reasons:

- To analyze a sample in the context of 2 different projects
- To transfer ownership of samples to a collaborator, but still keep a copy yourself
- To assign a sample to the correct project

To copy samples, perform the following steps.

1. Click the **Projects** icon.
2. Click the desired project.
3. Click the **Samples** link from the left navigation menu.
4. Select the checkboxes of the samples you want to copy.
5. Click **Copy**.
6. Select the new project in the drop-down list.
7. Click the **Copy** button.

Upload Files

Some apps need additional files generated outside of BaseSpace Onsite. Some appls also provide downstream analysis for results generated outside of BaseSpace Onsite.

To upload files, perform the following steps.

1 Click the **Projects** icon.

2 Go to the project you want to add the file to.

NOTE
You can only upload files to projects you own.

3 Click **Import**.

4 Select the file to upload using 1 of the following methods:
- Drag and drop the file in the **Drag and Drop** box.
- Browse to the file through the **select file** link and click **Open**.

5 BaseSpace Onsite now uploads the file to your project.

6 Enter the analysis name in the **Name of analysis** box.

7 [Optional] Associate the file you are importing with the samples used as inputs. By making this association, the analysis is listed on the sample detail page. You can also locate these uploaded files later by navigating to 1 of the samples.

8 Click **Complete Import**.

FASTQ File Upload Requirements

The FASTQ file is a text format file used to represent sequences. Each record has 4 lines. The 4 lines of data are an identifier (read descriptor), the sequence, +, and the quality scores. For a detailed description of the FASTQ format, see *FASTQ Files* on page 52.

Make sure that the FASTQ file adheres to the following upload requirements:

- The uploader only supports gzipped FASTQ files generated on Illumina instruments
- The name of the FASTQ files conforms to the following convention:
  `SampleName_SampleNumber_Lane_Read_FlowCellIndex.fastq.gz`
  For example,
  `SampleName_S1_L001_R1_001.fastq.gz`
  `SampleName_S1_L001_R2_001.fastq.gz)`
- The read descriptor in the FASTQ files conforms to the following convention:
  `@Instrument:RunID:FlowCellID:Lane:Tile:X:Y`
  `ReadNum:FilterFlag:0:SampleNumber:`
  For example, a Read 1 descriptor could look like this line:
  `@M00900:62:000000000-A2CYG:1:1101:18016:2491 1:N:0:13`
  And the corresponding Read 2 would have a 2 in the ReadNum field:
  `@M00900:62:000000000-A2CYG:1:1101:18016:2491 2:N:0:13`

Quality considerations:

- The number of base calls for each read equals the number of quality scores.
- The number of entries for Read 1 equals the number of entries for Read 2.
- The uploader determines whether files are paired-end based on matching file names in which the only difference is the ReadNum.
- For paired-end reads, the descriptor must match for every entry for both reads 1 and 2.
- Each read has passed filter.

The uploader rejects any upload session containing files that do not meet the requirements. If a file is rejected, a message is provided. This message states why the upload failed and what to do to remedy the problem.

Transfer Owner

You can use the Transfer Owner option to hand control of data over to a collaborator or customer. You can use this option for the following reasons:

- To give control of your data to a collaborator. You can only transfer data to collaborators who have an account on BaseSpace Onsite.

NOTE
If items from a project are in the trash, you cannot transfer ownership of the project.

To use this option, perform the following steps.

1. Select the project or run you want to transfer.
   - To transfer a project, perform the following steps.
     a. Click the **Projects** icon.
     b. Click the desired project.
     c. Click the **Transfer Owner** button.
   - To transfer a run, perform the following steps.
     a. Click the **Runs** icon.
     b. Click the desired run.
     c. Click the **More** button, and then select **Transfer Ownership**.
2. Enter new owner email and an optional message in the Transfer Ownership dialog box.
3. Click **Continue**.

Delete Items

You can delete items in a 2-step process:

1. Move the item to trash.
2. Clean up the trash.

For more information, see the following topics.

- *Delete Run* on page 61
- *Delete Project* on page 62
- *Delete Analysis* on page 62
- *Delete Sample* on page 63
- *Empty Trash* on page 63
- *Restore Trash* on page 63

Delete Run

To move a project out of your Runs list into the trash, perform the following steps.

1. Click the **Runs** icon.
2. Click the desired run.
3. Click **Move to Trash**.
4. If there are analyses or samples associated with this run, BaseSpace Onsite asks if you also want to delete those analyses and samples. Select the checkbox if you want to delete the associated items.
5. Click **Confirm**.

If you are the owner of the run, the run now appears in the trash.
You can permanently delete the run from trash by emptying all items in the trash; see *Empty Trash* on page 63 You can also restore the run to its original location; see *Restore Trash* on page 63.

Note the following items about deleting a run:

- If you received the run as a share, the deleted run does not appear in the trash. If you want to restore a run shared with you, click the previously sent share link or contact the owner.
- Runs cannot be removed if they are in a nonterminal state (eg, running, uploading, or analyzing).
- Do not perform this action if you want to archive a run.

Delete Project

To move a project out of your Projects list into the trash, perform the following steps.

1. Click the **Projects** icon.
2. Click the desired project.
3. Click **Move to Trash**.
4. Click **Confirm**.
   If you are the owner of the project, the project now appears in the trash.
   You can permanently delete it from trash by emptying all items in the trash; see *Empty Trash* on page 63 You can also restore the project to its original location; see *Restore Trash* on page 63.

Note the following items about deleting a project:

- If you received the project as a share, the deleted project does not appear in the trash. If you want to restore a project shared with you, click the previously sent share link or contact the owner.
- Do not perform this action if you want to archive a project.

Delete Analysis

To move an analysis out of your Analysis Results list into the trash, perform the following steps.

1. Click the **Projects** icon.
2. Click the desired project.
3. Select the analysis results to be deleted.
4. Click **Move to Trash**.
5. Click **Confirm**.
   The analysis is now in the trash.
   You can permanently delete the analysis from BaseSpace Onsite by emptying all items in the trash; see *Empty Trash* on page 63. You can also restore the analysis to its original location; see *Restore Trash* on page 63.

Note the following items about deleting an analysis result:

- You cannot delete an analysis that is in a nonterminal state (eg, pending execution or running).
- Deleting an analysis from a project that is in ownership transfer can cause a delay until the transfer has completed.
- Do not perform this action if you want to archive an analysis result.

Delete Sample

To move a sample out of your Samples list into the trash, perform the following steps.

1 Click the **Projects** icon.

2 Click the desired project.

3 Click the **Samples** tab on the left navigation bar.

4 Select the samples to be deleted.

5 Click **Move to Trash**.

6 Click **Confirm**.
The sample is now in the trash.
You can permanently delete the sample from BaseSpace Onsite by emptying all items in the trash; see *Empty Trash* on page 63. You can also restore the analysis to its original location; see *Restore Trash* on page 63.

Note the following items about deleting a sample:

- If you received the sample as a share, the deleted sample does not appear in the trash. If you want to restore a sample shared with you, click the previously sent share link or contact the owner.
- Do not perform this action if you want to archive a sample.

Empty Trash

After you have deleted an item, it is present in the Trash if you are the owner of that item. With this method, you can delete all items in the trash from BaseSpace Onsite permanently.

To delete an item from BaseSpace Onsite permanently, perform the following steps.

1 Click **View Trash**.

2 Click **Empty Trash**.

3 Click **Confirm**.
The items are now ready to be purged from BaseSpace Onsite by the admin; see *Purge Deleted Items* on page 68.

Restore Trash

After you have deleted an item, it is present in the Trash if you are the owner of that item. To restore an item from the trash to its original location, perform the following steps.

1 Click **View Trash**.

2 Select the items you want to restore.

3 Click **Restore Trash**.
The item is now restored to its original location. Restored items keep all their original attributes except for the share recipients.

Fix Sample Sheet / Rerun Workflow

The Fix Sample Sheet page lets you correct errors in your sample sheet for HiSeq runs, or set up a new analysis to requeue. Use this feature for the following reasons:

- To fix errors in the sample sheet—Errors in the sample sheet can prevent BaseSpace Onsite from processing a run. This option allows BaseSpace Onsite to finish the analysis.
- To change analysis parameters—If first analysis is suboptimal, you can resubmit the sample sheet and requeue the run with new analysis parameters one time.
- To change indexing details—If index settings for samples were wrong, you can correct the settings.

NOTE
You can only submit a corrected sample sheet and requeue the run 1 time.

To fix a sample sheet and rerun a workflow, perform the following steps.

1. Open the Fix Sample Sheet page using 1 of the following methods:
   - If a run has a Needs Attention state, open the run and click **Fix Sample Sheet**.
   - Go to a run, click the **More** drop-down list, and then select **Fix Sample Sheet**.

   The Fix Sample Sheet page opens. If BaseSpace Onsite has detected an error, it shows the issue above the black sample sheet editor.
2. Make changes depending on the complexity of the change:
   - **Easy fix**—Edit the sample sheet in the sample sheet editor. BaseSpace Onsite keeps validating the sample sheet as you edit; any remaining issues are displayed above the sample sheet editor.
   - **Complex Fix**—Use Illumina Experiment Manager (IEM) to create a sample sheet.

     a. If necessary, install IEM and open the program.
     b. Import the original sample sheet from your system in IEM and edit it, or generate a new sample sheet. For more information, see the *Illumina Experiment Manager User Guide (part # 15031335)*.
     c. Copy and paste the sample sheet into the Sample Sheet Editor in BaseSpace Onsite.
     BaseSpace Onsite validates the sample sheet; any issues are displayed above the sample sheet editor.
3. When you are done editing and the sample sheet is valid, click the **Queue Analysis** button.
   BaseSpace Onsite starts analyzing the run using the new sample sheet.
   You can only resubmit a sample sheet and requeue the run 1 time.

NOTE
If your edits result in an invalid sample sheet, the **Queue Analysis** button is not available. You can return to the original using the **Load Original** button.

Common Sample Sheet Fixes

If a sample sheet is invalid, it could be because the genome path is not set up correctly. This situation is indicated through the *Genome Path Unknown Genome* warning (as in the example). The paths of the standard BaseSpace Onsite genomes have to conform to the following relative paths:

```
Arabidopsis_thaliana\NCBI\build9.1\Sequence\WholeGenomeFASTA
Bos_taurus\Ensembl\UMD3.1\Sequence\WholeGenomeFASTA
Escherichia_coli_K_12_DH10B\NCBI\2008-03-
   17\Sequence\WholeGenomeFASTA
Homo_sapiens\UCSC\hg19\Sequence\WholeGenomeFASTA
Mus_musculus\UCSC\mm9\Sequence\WholeGenomeFASTA
PhiX\Illumina\RTA\Sequence\WholeGenomeFASTA
Rattus_norvegicus\UCSC\rn4\Sequence\WholeGenomeFASTA
```

```
Saccharomyces_cerevisiae\UCSC\sacCer2\Sequence\WholeGenomeFASTA
Staphylococcus_aureus_NCTC_8325\NCBI\2006-02-
   13\Sequence\WholeGenomeFASTA
```

Fix Indexes in the Prep Tab

You can correct errors in your indexes through the Prep tab and regenerate the FASTQ files. You can use this feature to change indexes and regenerate FASTQ files for samples have already been sequenced.

You ***cannot*** use this feature for the following situations:

- For runs that are not set up with the Prep tab.
- For runs where the wrong library prep kit was selected in the Prep tab.

To fix indexes in the Prep Tab, perform the following steps.

1. Go to the run with the wrong indexes.
2. Click the Run Settings button in the navigation task pane.
3. Go to the bottom of the page, and click the pool.
4. Click the Plate ID of the plate.
5. Click **Edit**.
6. Correct the index in the dropdown menu.
7. Go to the affected run.
8. In the **More** dropdown menu, select **Generate FASTQ Files**.
   BaseSpace Onsite now starts regenerating the FASTQ files with the corrected indexes. The new FASTQ files get added to the sample list and you can identify the new files by date.

NOTE
If you do not want to identify the samples by date, you can also rename the sample_ID, or assign the new FASTQ files to a new project in the Prep tab.

Search for Runs, Projects, Samples, Files, and Apps

The Search function allows you to find runs, projects, samples, files, and apps.

1. Click the search icon.
2. Type in the run, project, or sample name in the search field and hit enter.
3. Select the desired run, project, sample, file, or app in the search results. You can also filter the search results by these categories using the drop-down list at the right of the results page.

Admin Tasks

The admin of BaseSpace Onsite has the following tasks:

- Manage analysis, notifications, storage, users, system health, planned runs, software updates, and alarms using the Admin Panel. For more information, see *Manage BaseSpace Onsite* on page 66.
- Replace hard drive or power supply. For more information, see *Replacement Procedures* on page 69.

In addition, BaseSpace Onsite sends an email alert if there is an issue with your system. *Error Codes* on page 71 provides a list of possible codes and descriptions to help you troubleshoot.

Manage BaseSpace Onsite

If you have admin privileges, the Account drop-down list provides access to the Admin Panel. The Admin Panel allows you to manage analysis, notifications, storage, users, system health, planned runs, software updates, and alarms.

Manage Analysis

On the Analysis page, you see the analyses that are currently running, and the analyses that are queued. You can sort the analyses by the column headers.

You can perform the following actions:

- Select an active analysis and click **Stop** to stop the analysis.
- Select a queued analysis and click **Remove** to remove the analysis from the queue.
- Click **Settings** to set the administrator notifications you get through email.

Manage Storage

The Storage page contains 3 tabs:

- **Active Storage**—Provides an overview of the total amount of active storage used and free, and the storage by user. Active samples, runs, and analyses are stored on the BaseSpace Onsite system, and can be used for analysis. At the bottom, you see the list of active samples, runs, and analyses stored; you can sort this list by clicking the column headers. Any changes that add or remove data usually take 3–5 minutes to be reflected in the pie chart. Active samples, runs and analyses are stored on the BaseSpace Onsite system, and can be used for analysis.

  NOTE
  If there are many merged samples and copied samples on the system, the usage per-user values are overestimated in the Storage Use By Owners pie chart. The Total Active Storage is not affected.

- **Archive Storage**—Provides an overview of the archived analyses. Archived samples, runs, and analyses are stored on the archive system, and cannot be used for analysis on BaseSpace Onsite without restoring. At the bottom, you see the list of archived samples, runs, and analyses; you can sort this list by clicking the column headers.
- **Deleted Items**—Provides an overview of the items that have been deleted. At the bottom, you see the list of deleted samples, runs, and analyses; you can sort this list by clicking the column headers. You can restore these items to active storage, or purge them from BaseSpace Onsite, which deletes them permanently.

NOTE
Illumina highly recommends that you set up an archive location.

Archive or Backup Analysis

To archive or backup an analysis, perform the following steps.

1 On the Storage page, go to the **Archive Storage** tab.

2 Select the analysis.

3 Click **Archive**.

4 A dialog box appears, asking you if you want to keep the data in active storage.
- If you keep the data in active storage, you can keep working with the analysis, while making a backup in the archive.
- If you do not keep the analysis in active storage, the data are archived, and you cannot work with the analysis in BaseSpace Onsite. You can always move it back to active storage as described in *Restore Analysis From Archive* on page 67.

The maximum speed for archiving is 25 Mb/s, so it does not interfere with other BaseSpace Onsite tasks.

Restore Analysis From Archive

To restore an analysis from archive to active status, perform the following steps.

1 On the Storage page, go to the **Archive Storage** tab.

2 Select the analysis.

3 Click **Unarchive**.
If there is sufficient free space, BaseSpace Onsite restores the analysis. If there is not enough space, BaseSpace Onsite displays an error message.

NOTE
You cannot restore an archived item that has been deleted and purged from BaseSpace Onsite.

Set Up Archive Location

To set up an archive location, perform the following steps.

1 On the Storage page, go to the **Archive Storage** tab.

2 Click **Mount**

3 Fill out the form.
You can set up 1 archive location per BaseSpace Onsite system.

Restore Deleted Item

To restore a deleted item to active status, perform the following steps.

1 On the Storage page, go to the **Deleted Items** tab.

2 Select the analysis.

3 Click **Unarchive**.
If there is sufficient free space, BaseSpace Onsite restores the item. If there is not enough space, BaseSpace Onsite displays an error message.

Purge Deleted Items

To purge deleted items from BaseSpace Onsite permanently, perform the following:

1 On the Storage page, go to the **Deleted Items** tab.

2 Click **Purge**.
This action removes all items in the trash from BaseSpace Onsite permanently.

Storage Check

BaseSpace Onsite has limited storage capacity, and checks the free space available before uploading a run or starting an app. The necessary space is then reserved until the process completes.

If there is not enough space, you see an error message and the run or app does not start. If the space check fails before starting FASTQ generation, the run gets into ***Needs Attention*** state. To manage available space, see *Archive or Backup Analysis* on page 67.

Manage Users

On the Users page, you see a list of current BaseSpace Onsite accounts, and their roles. You can sort the users by the column headers.

To set up a new user, click **Add User** and fill out the form.

The new user gets an email with a link to set up a new password. If the user does not get an email, the user can go to the BaseSpace Onsite login page and click **Forgot Password** to resend the email.

If you did not configure SMTP during install, be aware of the following items:

- You enter the password for the new user.
- BaseSpace Onsite does not enforce usernames to be in the form of an email, which can cause a problem with HiSeq integration. Make sure to enforce that all users create a username in the format of an email.
- If you enable the SMTP support after using BaseSpace Onsite without SMTP support, all users that are logged in must log out and then log back in. Otherwise, they do not receive notifications by email or on the dashboard.

Monitor System Health

On the System Health page, you see the BaseSpace Onsite system health alerts. You can sort the alerts by the column headers.

Many sensors are monitored for health in the BaseSpace Onsite server. If a sensor indicates a failure, BaseSpace Onsite sends an alert to the administrator. When you receive an alert, contact Illumina Support to diagnose the error and, if necessary, arrange a site visit to correct the problem.

Some of these errors, such as a failure of a disk drive or 1 of the power supplies, you can resolve. See *Replacement Procedures* on page 69 for instructions. Illumina Support can also guide you through the process of replacing the faulty component.

If you want to remove an alert, select the alert and click **Dismiss**.

Unlock Planned Runs

On the Planned Runs page, you see the runs that are currently planned. You can sort the runs by the column headers.

The sequencing system locks planned runs when they are selected. In rare instances, the sequencing system leaves the run in locked state without starting the run. These runs must be manually unlocked to access again. Unlocking runs allows the users to edit the run in the Prep tab, and makes the run available for selection on other sequencing systems. Unlocking also frees up the reserved run storage space on the BaseSpace Onsite system.

To unlock a run, select the planned run and click **Unlock**.

Updates

To update the server software and upload additional genome content, perform the following steps.

1 Plug the external USB drive with the software update into the BaseSpace Onsite head node server.

2 Go to the Updates page.

3 Click **Detect Drive**.
The Updates wizard leads you through the update.

System Logs

The System Logs page provides a download of the log files. You can download all log files, or packages that contain the system state log files, active analysis log files, or delete log files.

If an issue arises with your BaseSpace Onsite, Illumina support uses these files to troubleshoot your system. To speed up troubleshooting, download the appropriate package before calling Illumina support.

About Page / Licenses

The About page provides a download containing the licenses for third-party software components.

Manage Alarms on the Settings Page

If there are problems with the hard drive, BaseSpace Onsite sounds a loud alarm. The **Disable Alarm** button on the Settings page allows you to turn off a current alarm.

See *Replace Hard Drive* on page 70 to continue.

CAUTION
Do not disable the alarm and ignore the warning. Failure to address the warning could lead to irretrievable data loss.

Replacement Procedures

You have received a spare hard drive and power supply module with the instrument. If needed, you can perform the replacement without calling Illumina, using the instructions in this section.

To order more, use the following material numbers.

Part	Nomenclature	Material Number
Power Supply	PSU 750W 1U CRPS 80PLUS PLATINUM	10535S
Hard drive	Hard Drive, 2 TB SATA, 6GB/S	15049450S
	Hard Drive, 4 TB SATA, 6GB/S	20000882

Replace Power Supply Module

To replace the power supply module, perform the following steps.

1 Remove the power cord from the power supply.

2 Press the green and black tab and slide the power supply out.

Figure 4 Slide Power Supply Out

3 Push the new power supply into the slot and make sure that it is seated properly.

Figure 5 Reseat Power Supply

4 Attach the power cord to the power supply.

Replace Hard Drive

To replace the hard drive, perform the following steps.

1 Lift the green tab on hard drive tray, and then pull out the drive.

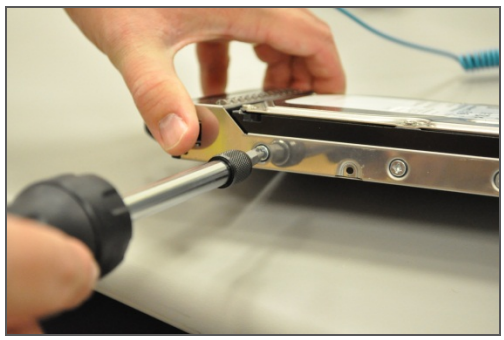
Figure 6 Lift Tab

Figure 7 Pull Drive Out

2 Look at the label on the drive and check whether it is a 2 TB or 4 TB drive.

3 Remove the 3 screws on each side of the carrier, then remove the old drive from the carrier.

Figure 8 Remove Drive From Carrier



4 Attach the new drive to the carrier.

5 Insert the new drive into the bay and make sure the tray latches closed. The new drive is automatically rebuilt.

6 Go to the Settings page in the Admin Panel, and click **Enable Alarm**.

Data Recovery

To recover data, contact Illumina Technical Support.

Error Codes

If there is an issue with your system, BaseSpace Onsite sends an email alert. The possible codes and descriptions are listed in this topic to help you troubleshoot.

Error Code/ Item ID	Item Name	Message	Status
Pwr_ Unit_ Status	Power Unit Status	The power unit has detected a shutdown	Error
Pwr_ Unit_ Status	Power Unit Status	The power unit has detected that the system has been turned on	OK

Error Code/ Item ID	Item Name	Message	Status
Pwr_ Unit_ Status	Power Unit Status	The power unit has detected a loss of AC power	Error
Pwr_ Unit_ Status	Power Unit Status	The power unit has detected that the AC power has been restored	OK
Pwr_ Unit_ Status	Power Unit Status	The power unit has detected a soft-power failure	Error
Pwr_ Unit_ Status	Power Unit Status	The power unit has recovered from a soft-power failure	OK
Pwr_ Unit_ Status	Power Unit Status	The power unit has detected an unexpected failure	Error
Pwr_ Unit_ Status	Power Unit Status	The power unit has recovered from an unexpected failure	OK
PU_$C	Power supply unit number $C	The power supply is no longer redundant due to power unit number $C failing	Error
PU_$C	Power supply unit number $C	The power supply is no longer redundant due to power unit number $C failing	Error
PU_$C	Power supply unit number $C	The power supply is no longer redundant due to power unit number $C failing	Error
PU_$C	Power supply unit number $C	The power supply is no longer redundant due to power unit number $C failing	Error
PU_$C	Power supply unit number $C	The power supply is no longer redundant due to power unit number $C failing	Error
PU_$C	Power supply unit number $C	The power supply is no longer redundant due to power unit number $C failing	Error
PU_$C	Power supply unit number $C	The power supply is no longer redundant due to power unit number $C failing	Error
PU_$C	Power supply unit number $C	The power supply is redundant again	OK
IPMI	IPMI Watchdog	The hardware monitor has detected an expired timer	Error
IPMI	IPMI Watchdog	The hardware monitor has detected hard reset of the system	Error
IPMI	IPMI Watchdog	The hardware monitor has detected that the system is shutting down	Error

Error Code/ Item ID	Item Name	Message	Status
IPMI	IPMI Watchdog	The hardware monitor has detected that the system is restarting	Error
IPMI	IPMI Watchdog	The hardware monitor has detected a timer interrupt	Error
Phy_Sec	Physical Security	The system cover has been opened	Error
Phy_Sec	Physical Security	The system cover has been closed	OK
Phy_Sec	Physical Security	The system has been unplugged from the network	Error
Phy_Sec	Physical Security	The system network connection has been restored	OK
FPB	Front Panel Board	The front panel board has detected a critical interrupt error	Error
SMI	System Board Timeout	The system board has detected a timeout	Error
SMI	System Board Timeout	The system board has recovered from a timeout	OK
SE	System Event	There was a PEF Action detected	Error
SB	System Board	The system board has detected that the power button has been pressed	OK
SB	System Board	The system board has detected that the reset button has been pressed	OK
VR	Voltage Watchdog	The voltage sensor has detected that the voltage is not within normal range	Error
VR	Voltage Watchdog	The voltage sensor has detected that the voltage is back within normal range	OK
F_$C	Fan number $C	The fans are no longer redundant due to fan number $C failing	Error
F_$C	Fan number $C	The fans are no longer redundant due to fan number $C failing	Error
F_$C	Fan number $C	The fans are no longer redundant due to fan number $C failing	Error
F_$C	Fan number $C	The fans are no longer redundant due to fan number $C failing	Error
F_$C	Fan number $C	The fans are no longer redundant due to fan number $C failing	Error
F_$C	Fan number $C	The fans are no longer redundant due to fan number $C failing	Error

Error Code/ Item ID	Item Name	Message	Status
F_$C	Fan number $C	The fans are no longer redundant due to fan number $C failing	Error
F_$C	Fan number $C	The fans are redundant again	OK
TEMP	System Board Temperature	The system board temperature has exceeded the normal range	Error
TEMP	System Board Temperature	The system board temperature is back within normal range	OK
BMC_ FW	BMC Firmware Health	The BMC board has detected a sensor failure	Error
F1	Fan 1	Fan number 1 is not functioning	Error
F1	Fan 1	Fan number 1 is functioning	OK
F2	Fan 2	Fan number 2 is not functioning	Error
F2	Fan 2	Fan number 2 is functioning	OK
F3	Fan 3	Fan number 3 is not functioning	Error
F3	Fan 3	Fan number 3 is functioning	OK
F4	Fan 4	Fan number 4 is not functioning	Error
F4	Fan 4	Fan number 4 is functioning	OK
F5	Fan 5	Fan number 5 is not functioning	Error
F5	Fan 5	Fan number 5 is functioning	OK
PS_$C	Power supply number $C	The number $C AC power supply is no longer available	Error
PS_$C	Power supply number $C	The number $C AC power supply has been connected	OK
PS_$C	Power supply number $C	The number $C AC power supply has recovered from failure	OK
PS_$C	Power supply number $C	The number $C AC power supply has detected a failure	Error
PS_$C	Power supply number $C	The number $C AC power supply has recovered from a predictive failure	OK
PS_$C	Power supply number $C	The number $C AC power supply has detected a predictive failure	Error
PS_$C	Power supply number $C	The number $C AC power supply has been lost	Error
PS_$C	Power supply number $C	The number $C AC power supply has been restored	OK

Error Code/ Item ID	Item Name	Message	Status
PS_$C	Power supply number $C	The number $C AC power supply has resolved the configuration error for a vendor mismatch	OK
PS_$C	Power supply number $C	The number $C AC power supply has detected a configuration error for a vendor mismatch	Error
PS_$C	Power supply number $C	The number $C AC power supply has resolved the configuration error for a revision mismatch	OK
PS_$C	Power supply number $C	The number $C AC power supply has detected a configuration error for a revision mismatch	Error
PS_$C	Power supply number $C	The number $C AC power supply has resolved the configuration error for a missing processor	OK
PS_$C	Power supply number $C	The number $C AC power supply has detected a configuration error for a missing processor	Error
PS_$C	Power supply number $C	The number $C AC power supply has resolved an unexpected configuration error	OK
PS_$C	Power supply number $C	The number $C AC power supply has detected an unexpected configuration error	Error
CBPB	Chassis Back Panel Board	The panel board in the back of the chassis has changed to being offline	Error
CBPB	Chassis Back Panel Board	The panel board in the back of the chassis is back online	OK
P_$C	Processor number $C	The processor number $C has exceeded the normal range	Error
P_$C	Processor number $C	The processor number $C is back within the normal range	OK
P_$C	Processor number $C	The processor number $C has been detected	OK
P_$C	Processor number $C	The processor number $C is not being detected	OK
P_ERR_ $C	Processor number $C digital state error	The processor number $C has detected a digital state error	Error
P_ERR_ $C	Processor number $C digital state error	The processor number $C has recovered from a digital state error	OK

Error Code/ Item ID	Item Name	Message	Status
P_ TEMP_ $C	System temperature	The system temperature has exceeded its threshold	Error
P_ TEMP_ $C	System temperature	The system temperature is back within normal operating range	OK
PS_ FAN_ $C	Power Supply Fan $C	The power supply fan number $C has failed	Error
PS_ FAN_ $C	Power Supply Fan $C	The power supply fan number $C has been restored	OK
M_$C	Memory Sensor number $C	The thermal sensor at memory location $C has exceeded its threshold	Error
M_$C	Memory Sensor number $C	The thermal sensor at memory location $C is back within normal range	OK
HD_$C	Disk in slot $C	Disk rebuild in progress at slot $C	Error
HD_$C	Disk in slot $C	Disk rebuild at slot $C has completed	OK
HD_$C	Disk in slot $C	Disk at slot $C is no longer online	Error
HD_$C	Disk in slot $C	Disk at slot $C is back online	OK
HD_$C	Disk in slot $C	Disk at slot $C is no longer working	
HD_$C	Disk in slot $C	Disk at slot $C is no functioning correctly again	
LV_$B	Logical volume $B with capacity $C $D	The logical volume at $B size $C $D is not currently active	Error

Technical Assistance

For technical assistance, contact Illumina Technical Support.

Table 2 Illumina General Contact Information

Website	www.illumina.com
Email	techsupport@illumina.com

Table 3 Illumina Customer Support Telephone Numbers

Region	Contact Number	Region	Contact Number
North America	1.800.809.4566	Italy	800.874909
Australia	1.800.775.688	Netherlands	0800.0223859
Austria	0800.296575	New Zealand	0800.451.650
Belgium	0800.81102	Norway	800.16836
Denmark	80882346	Spain	900.812168
Finland	0800.918363	Sweden	020790181
France	0800.911850	Switzerland	0800.563118
Germany	0800.180.8994	United Kingdom	0800.917.0041
Ireland	1.800.812949	Other countries	+44.1799.534000

Safety data sheets (SDSs)—Available on the Illumina website at support.illumina.com/sds.html.

Product documentation—Available for download in PDF from the Illumina website. Go to support.illumina.com, select a product, then select **Documentation & Literature**.

Illumina
San Diego, California 92122 U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com