

BaseSpace™ Sequence Hub

Data management and analysis that is simple enough for labs getting started, or powerful enough for rapidly scaling up next-generation sequencing (NGS) operations.

Highlights

- Touchpoint-free automation**
 Configure analysis workflows and automated quality control steps to remove touchpoints and streamline data processing
- Real-time data upload and run monitoring**
 View run progress as data uploads to the data repository and begin analysis immediately after run completes
- One-click analysis with over 80 bioinformatics tools**
 Access and launch a growing collection of bioinformatics tools easily with BaseSpace Apps
- Global collaboration and data sharing**
 Configure options to disseminate data to peers, create working groups, or engage the scientific community

Introduction

Next-generation sequencing (NGS) has revolutionized the way and rate at which biomedical research is conducted. As the cost of sequencing decreases, the volume of NGS-generated data increases, creating new bottlenecks. The challenges of secure data storage and management, complex data analysis, and sharing results with collaborators can result in nonuniform methods within institutions and labs, conflicting results, and increased operational overhead. BaseSpace Sequence Hub is a genomics cloud-computing platform designed to bring simplified data management and analytical sequencing tools directly to investigators in a user-friendly format (Figure 1). BaseSpace Sequence Hub provides flexibility and convenience with an array of tools, significantly expanding the possibilities of yielding meaningful results from NGS data.

Figure 1: Intuitive BaseSpace Sequence Hub Dashboard—The Notifications panel has widgets that highlight the latest sharing, ownership transfer activities, occasional alerts on new features, bugs, etc. The Latest Runs panel has widgets that provide real-time status of the sequencing run. The Latest Analyses panel has widgets that show the status of user app sessions.

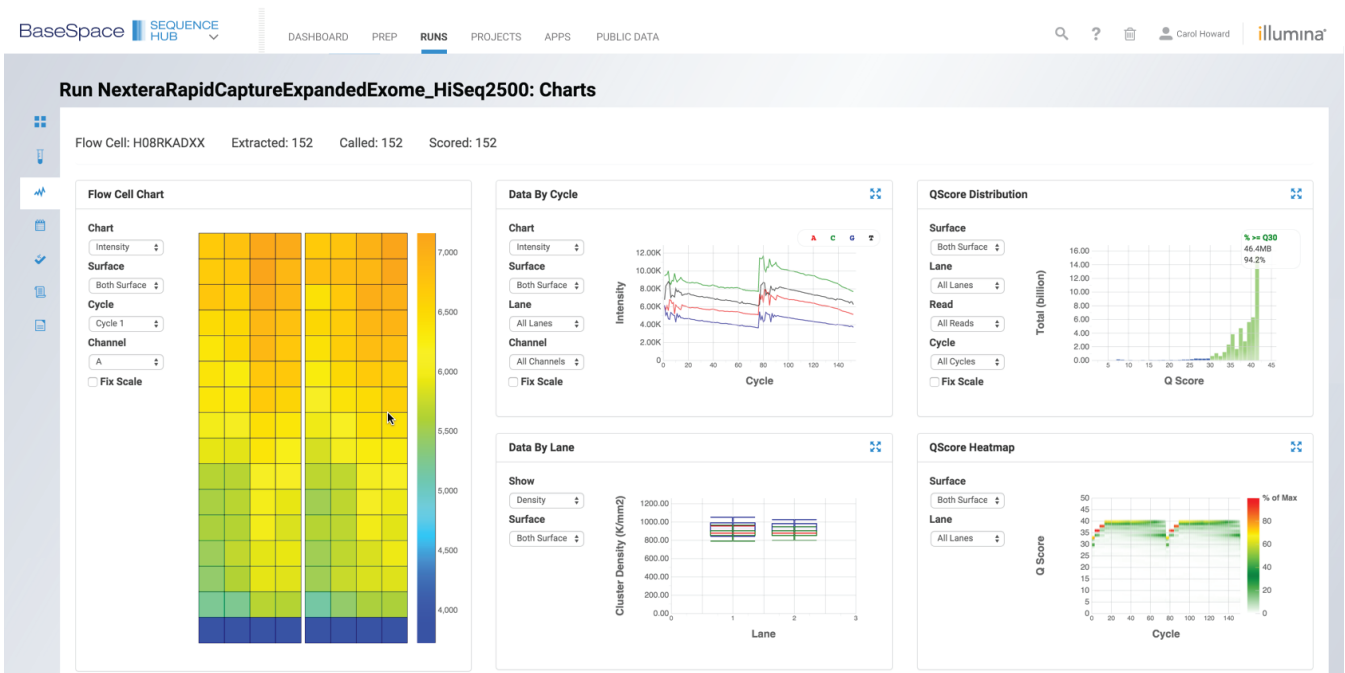


Figure 2: Monitor run data in real time—Sequence Analysis Viewer (SAV) capabilities are built into BaseSpace Sequence Hub user interface, allowing real-time, cycle-by-cycle monitoring. The Charts view shows data by lane and by cycle, with Q-Score distribution and heat map features. Each graph can be expanded to full size.

Scalable bioinformatics infrastructure

Labs pursuing NGS traditionally required the services of highly trained bioinformaticians and a dedicated infrastructure to perform data management, analysis, and storage. BaseSpace Sequence Hub relieves many of these burdens by automating bioinformatic analysis using cloud-based software applications. These push-button apps were designed with the biologist in mind, enabling users to produce biologically relevant results from raw data that can be used in downstream analysis tools. All BaseSpace Sequence Hub accounts come with 1 TB of free storage, but this storage is scalable according to changing laboratory needs.

Run setup and management

BaseSpace Sequence Hub makes biosample and run management easy using the Prep Tab feature, an intuitive, graphical environment for one-stop library and run preparation.

- Prepare and manage biological samples, libraries, pools, and planned sequencing runs directly in BaseSpace Sequence Hub
- Import biological samples or library information in batch mode for large experiments

The features available in Prep Tab allow for easy integration of BaseSpace Sequence Hub with library preparation and sequencing platforms. Using the Prep Tab, the entire workflow can be planned from biosample creation and library prep to pooling and sequencing. When a run has been planned and is ready to start, no additional setup is required at the instrument.

- Prep Tab supports all Illumina library prep kits
- Prep Tab can also be used for custom library kits
- MiSeq™, NovaSeq™, and HiSeq™ instruments (including the HiSeq X™ System) can be set up in BaseSpace Sequence Hub using sample sheets.

Real-time monitoring

BaseSpace Sequence Hub is the only cloud platform directly integrated with Illumina sequencing systems. With the Runs Dashboard, users can monitor data, by lane or by cycle, as data are generated in real time on the sequencer. Additionally, users can view quality performance metrics from their browsers (Figure 2) or on mobile devices using the BaseSpace Mobile App (Figure 3).¹ Data are seamlessly pushed to BaseSpace Sequence Hub for automatic analysis and storage upon run completion, with the option of retaining data for local hosting and analysis on the instrument.

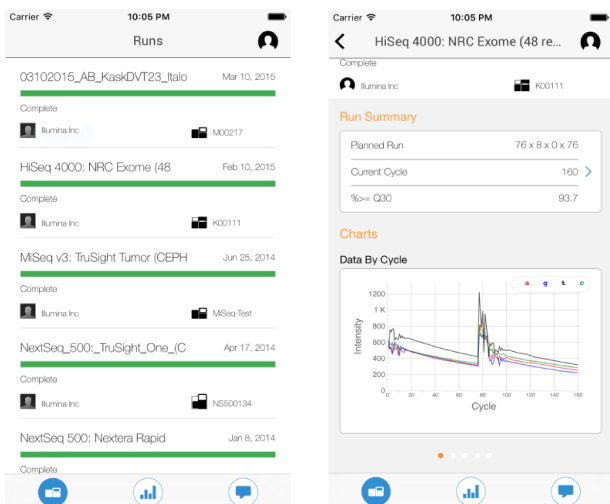


Figure 3: View runs on the BaseSpace Mobile App—Runs can be viewed on the BaseSpace Mobile App on any iOS-compatible device. The Mobile App also provides status updates on analysis, including push-button notifications when runs and analyses are completed. Download the Mobile App for free at the iTunes app store.¹

BaseSpace apps, powerful yet simple

Analysis of complex sequencing data sets is a challenge at any scale. BaseSpace Sequence Hub provides a continuously growing list of powerful apps (analysis workflows and tools), allowing researchers to set up and perform complex data analyses. A simple interface links data sets directly to bioinformatics pipelines based on open source and commercial tools (Figure 4). BaseSpace Apps meet the diverse needs of any researcher, regardless of informatics experience, in an expansive analysis ecosystem.²

In addition to Illumina developed apps, BaseSpace Sequence Hub hosts third-party premium apps developed by commercial partners, including Edico Genome and Sentieon. Note: app availability varies by regional deployment of BaseSpace Sequence Hub.

Develop custom apps to analyze data

Working with customized pipelines and tools within BaseSpace Sequence Hub simplifies bioinformatics processes by allowing the user to bring their analysis methods to the data in a flexible platform. BaseSpace Sequence Hub supports third-party software development by providing a robust app development platform. The BaseSpace Native App Engine and extensive application program interfaces support development of apps to perform analyses and create custom reports.³ Customized apps can be kept private, shared between collaborators, or made publicly available to all BaseSpace Sequence Hub users.



Figure 4: Launch analytical tools on demand—Browse and explore a growing list of apps from the bioinformatics community in the BaseSpace Apps Store, and launch selected apps with a single click, directly from the data set. For more information, go to www.illumina.com/BaseSpaceApps.

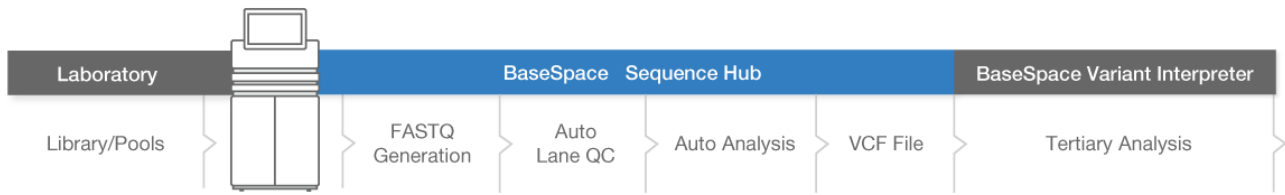


Figure 5: Automated workflow in BaseSpace Sequence Hub—BaseSpace Sequence Hub enables users to configure customized analysis workflows which include automatic FASTQ generation, automatic sample aggregation (when applicable) automatic Lane QC, automatic launching of analyses with BaseSpace Apps, and automated secondary analysis QC.

Streamline analysis with automated workflows

A typical bioinformatics analysis workflow includes many steps (Figure 5). From reviewing lane metrics and post-run demultiplexing to merging data from multiple runs, setting up secondary analyses, and reviewing results, the process is time-consuming and susceptible to human error. BaseSpace Sequence Hub enables users to configure customized workflows to completely automate the process from sequencing run completion to secondary analysis and results review prior to data delivery or tertiary analyses.

These features also enable high-volume laboratories to maximize efficiencies by tracking biosample status throughout the data analysis workflow, so users can review the progress of individual biosamples as they proceed down individual analysis paths.

Collaborate better with workgroups

The ability to form a team through the Workgroup feature is available with an upgrade to a BaseSpace Professional subscription or BaseSpace Enterprise subscription. Each BaseSpace Professional subscription is provided with a single workgroup, while BaseSpace Enterprise tier customers can create any number of workgroups for better management of access to data. This feature enables simplified collaboration on a global scale (Figure 6):

- The team administrator (subscriber) can invite other users to the Workgroup
- All team members will have access with individual login
- Team members can switch between individual and Workgroup spaces
- In Workgroup context, team members can access all runs, analyses, and storage usage that belongs to the Workgroup

In large labs with multiple users, accounts and passwords are often shared between technicians, bioinformaticians, lab managers, etc. The Workgroup feature enables each invited individual to log in with individual passwords. This also mitigates issues that may arise when any user leaves the lab.

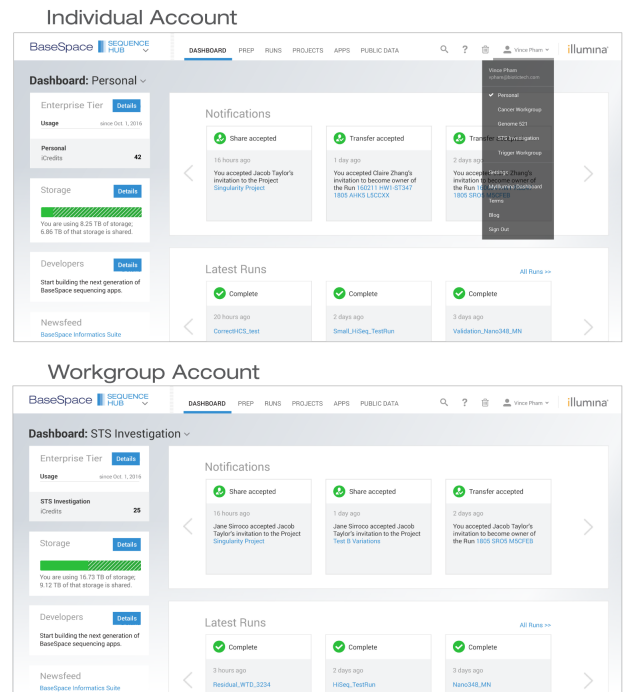


Figure 6: Workgroup feature—With Workgroup, users can log in with personal credentials, then switch context between individual accounts and Workgroup accounts. In Workgroup context, users can view all runs and analyses common to the Workgroup, as well as use the storage and computation hours purchased by the Workgroup, as shown in the different dashboards.

Collaboration on a global scale

Researchers frequently need to collaborate and share access to sequencing data and results. BaseSpace Sequence Hub enables users to share raw sequencing data and analysis results, easily and securely, with collaborators across the globe. Shareable links can be easily created and emailed to partners, allowing instantaneous access to shared data and results. Also, data delivery is simplified with the ability to transfer runs and projects effortlessly to collaborators or customers. BaseSpace Sequence Hub makes big data portable and accessible to the people who need it most (Figure 7).

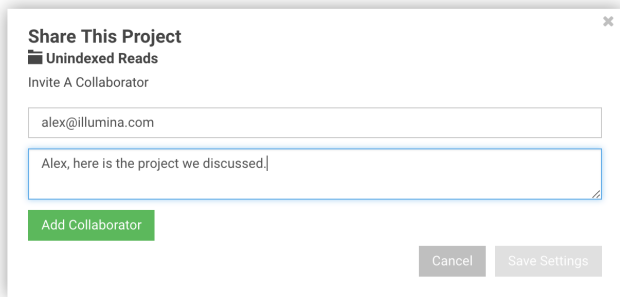


Figure 7: Collaboration tool—Flexible collaboration tools simplify data sharing, enable expansion of collaboration circles, and easily keep track of who shares data.

Bioinformatics Professional Services

To address the challenges associated with managing and analyzing the vast amounts of data generated by NGS, Illumina offers a range of professional services for both bioinformatics analysis and consulting.

Driven by a need for quality data and analysis, Illumina has developed an extensive list of bioinformatics offerings. These cover standard analysis workflows, standard output files, and visualization tools to more application-specific workflows and personalized consultation on topics such as data quality, data analysis, data management, and pipeline analysis.

Illumina Bioinformatics Professional Services^{4,5} was developed and delivered by professionals with deep domain experience, and can be of value to any lab.

Enhanced security

Security is of paramount importance when making the decision to move genomic data to cloud-based analysis and storage. In BaseSpace Sequence Hub, data are protected through various physical, electronic, and administrative measures. Data for upload are encrypted using the AES256 standard and protected by secure sockets later (SSL). Data within BaseSpace Sequence Hub are hosted on Amazon Web Services (AWS), which is compliant with a wide variety of industry-accepted security standards.⁶ Enterprise subscriptions offer an additional level of security. Enterprise customers are provided their own domain and the ability to use their own SAML 2.0 supported authentication service to manage users and passwords. BaseSpace Sequence Hub also supports Enterprise

customers in a Health Insurance Portability and Accountability Act (HIPAA)-regulated environment with a signed Business Associate Agreement (BAA). For more information about security features, read the [BaseSpace Sequence Hub Security and Privacy white paper](#).⁷

Flexible billing plans

BaseSpace Sequence Hub offers customers the option to use only what they need, providing both a monthly billing feature for the use of paid apps and storage, and an unlimited plan which includes storage and consumption for each connected instrument during the first year of a subscription. All BaseSpace Sequence Hub accounts come with 1 TB of free storage and access to a limited number of free apps. Log in to BaseSpace Sequence Hub and visit the app page for more information about apps and pricing. Customers can purchase subscriptions for specified storage amounts and use the pay-as-you-go feature for any additional storage and paid apps. Billing features vary with the type of subscription (Table 1).

Table 1: Billing features for BaseSpace Sequence Hub subscriptions

Features	Basic	Professional	Enterprise
Included storage	1 TB	1 TB	1 TB
Compute	30-Day Free Trial with 250 complimentary iCredits	250 complimentary iCredits ^a	250 complimentary iCredits ^a
Payment plans	n/a	Monthly billing or unlimited storage + compute ^a	Monthly billing or unlimited storage + compute ^a
Run setup	✓	✓	✓
Run monitoring	✓	✓	✓
Data sharing	✓	✓	✓
Apps	Free apps only (after trial period expires)	All	All
API access	✓	✓	✓
Command-line access	✓	✓	✓
Comprehensive security framework	✓	✓	✓
Number of users	One	Unlimited	Unlimited
Workgroups		One	Unlimited
Bioinformatics Professional Services		8 hours ^b	24 hours ^b
Multiple workgroups			✓
Private Domain			✓
Single sign-on			✓
Supports customers in a HIPAA-regulated environment ^c			✓
Access control			✓
Audit trail			✓
Price	Free	Contact a sales rep	Contact a sales rep

a. For new customers only. Unlimited plans are only available for the first year. Subsequent years are Monthly Billing.

b. With the purchase of Unlimited Storage and Compute.

c. Applicable in the US only.

The screenshot shows two app pricing cards. The first card is for 'Whole Genome Sequencing' by Illumina, Inc., with a version of 7.0.1 and a compute cost of 3.00 iCredits per node hour. The second card is for 'Sentieon DNaseq FASTQ to VCF' by Sentieon Inc., with a version of 1.0.0 and a compute cost of 3.00 iCredits per node hour. Both cards include a 'Launch Application' button and a 'READ MORE' link.

Figure 8: Paid app pricing—Individual apps can be purchased with iCredits on a price-per-node hour basis, single-use fee, or annual subscription fee.

The monthly billing feature in BaseSpace Sequence Hub uses iCredits to track the use of paid apps and storage. For paid apps, the iCredits system tracks central processing unit (CPU) cycles used, and app license fees for a limited number of third-party apps (Figure 8). Each paid app using CPU cycles is priced in iCredits per node hour, and is tracked in per-minute billing increments. Third-party app license fees are listed in iCredits and are tracked upon app launch. Any storage use above baseline is priced at .03 iCredits/GB per month, and is tracked in per-day billing increments. A bill is sent each month for usage of paid apps and additional storage.

Learn more

For more details on BaseSpace Sequence Hub, or to sign up for a free BaseSpace account, visit www.illumina.com/basespace

Ordering information

Products	Catalog no.
MiSeq Unlimited Plan	20020622
NextSeq™ Unlimited Plan	20020623
HiSeq 2500 Unlimited Plan	20020624
HiSeq 3000 Unlimited Plan	20020625
HiSeq 4000 Unlimited Plan	20020626
HiSeq X Unlimited Plan	20020627
NovaSeq 6000 Unlimited Plan	20020628
BaseSpace Sequence Hub Enterprise Annual Subscription	SW-411-1003
BaseSpace Sequence Hub Live Storage 1 TB	SW-410-1002
BaseSpace Sequence Hub Live Storage 5 TB	SW-410-1003
BaseSpace Sequence Hub Live Storage 10 TB	SW-411-1001
BaseSpace Sequence Hub Live Storage 50 TB	SW-411-1002

References

1. iTunes BaseSpace App. itunes.apple.com/us/app/basespace/id942794217?mt=8. Accessed June 12, 2017.
2. BaseSpace Genomics Computing. www.illumina.com/BaseSpaceApps. Accessed June 12, 2017.
3. BaseSpace Developers. developer.basespace.illumina.com. Accessed June 12, 2017.
4. Illumina (2017) Illumina Bioinformatics Professional Services. (<https://www.illumina.com/content/dam/illumina-marketing/documents/services/illumina-bioinformatics-professional-services-data-sheet-070-2016-006.pdf>).
5. Illumina Consulting. www.illumina.com/services/instrument-services-training/consulting.html. Accessed June 12, 2017.
6. AWS Cloud Security. aws.amazon.com/security/. Accessed June 12, 2017.
7. Illumina (2016) BaseSpace Sequence Hub Security and Privacy. (www.illumina.com/content/dam/illumina-marketing/documents/products/whitepapers/basespace-sequence-hub-security-and-privacy-white-paper-970-2016-020.pdf).