BaseSpace™ Sequence Hub Security and Privacy

Numerous security features help protect genomic data.

Introduction

The BaseSpace Sequence Hub is a software-as-a-service (SaaS) offering for genomics data management and analysis, placing high priority on protecting the confidentiality, integrity, and availability of customer data. BaseSpace Sequence Hub has a set of comprehensive policies, processes, and controls for data security and privacy.

BaseSpace Sequence Hub is implemented as a single instance, multitenant deployment. Twice monthly updates are transparent, requiring no user action. All major upgrades are communicated via release notes at least two weeks in advance.

Based on business needs and customer input, the software development life cycle (SDLC) of BaseSpace Sequence Hub determines prioritization of features, content, functionality, and bug remediation. Software is engineered and unit-tested using contemporary agile development methodologies. Service updates are relatively small and frequent, reducing the risk of customer impact. Validation of changes in BaseSpace Sequence Hub occurs in a test environment that is segregated from a staging/production environment, involving automated regressions and manual testing.

To address platform security, the well-tested approach of Amazon Web Services (AWS)[1] is combined with internal testing procedures. Together, these methods provide a cloud genomics solution that meets the security provided by many institutional IT infrastructures. This white paper describes the underlying security features of BaseSpace Sequence Hub, which supports customer compliance with various regulations and standards, such as the Health Insurance Portability and Accountability Act (HIPAA) in the US, and the General Data Protection Regulation (GDPR) in the European Union (EU).

Security practices in BaseSpace Sequence Hub

BaseSpace Sequence Hub imports data directly from the sequencing instrument during the run, enabling customers to begin data analysis as soon as the run completes. Several security measures protect data in transit while communication occurs between the sequencing instruments and the data analysis and storage servers.

Users can store data both locally and in BaseSpace Sequence Hub (MiniSeq™ System excluded). Illumina sequencing platforms (MiniSeq, MiSeq™, NextSeq™, HiSeq™, and NovaSeq™ Systems) support encryption in transit and verification of data generated during a run. The brokering software interacts with the application program interface (API), allowing network interruptions, latencies, and incomplete transmissions to be caught and requeued automatically.

During run setup, the user initiates the decision to send data to BaseSpace Sequence Hub. If this option is chosen, the run is authenticated against, and tracked to, a user account. BaseSpace Sequence Hub Enterprise also supports single sign-on, which is supported by third-party authentication using the standards of security assertion markup language (SAML) 2.0.

Users are required to remove all direct identifiers of individuals (eg, names, dates of birth) from genomic information when they upload it to BaseSpace Sequence Hub.[2] This is accomplished by providing unique codes, such as bar codes or random sample identifiers, to each sample prior to uploading. It is recommended that BaseSpace Sequence Hub customers store any direct identifier information in a separate encrypted system outside of BaseSpace Sequence Hub.

Security in transit

BaseSpace Sequence Hub communicates with instruments through a web-based API. All traffic between the sequencing instrument and the BaseSpace Sequence Hub uses Transport Layer Security, an internet standard that encrypts sensitive communications as they pass over the internet. All service methods require API key signatures, and service is refused to all others. Requests are monitored for abuse.

Access to BaseSpace Sequence Hub

To access BaseSpace Sequence Hub, users log in via a web portal. Users can be identified with BaseSpace Sequence Hub authentication or using single sign-on for enterprise customers. BaseSpace Sequence Hub Enterprise customers can dictate policies for password length and complexity, and configure account lockout and lockout duration to protect against password brute forcing.

Invalid login attempts and logoffs are recorded by the system. If enterprise customers use a single sign-on, login activity may also be monitored from the customer systems. Changes, reads, updates, deletions, and shares of customer data are also logged for BaseSpace Sequence Hub Enterprise customers. Logs can be monitored for suspicious user activity and are available as CSV files or via API. All computation instances run within Virtual Private Clouds, providing a logically isolated section of the AWS cloud, where AWS resources reside in a virtual network defined by Illumina.

Data integrity

Through AWS, BaseSpace Sequence Hub stores customer data synchronously across multiple availability zones, performs regular data integrity checks, and self-heals to protect against data loss. However, BaseSpace Sequence Hub is not an unlimited backup system. There is no mechanism to retrieve deleted data.

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

Encryption at rest

Customer data in BaseSpace Sequence Hub is encrypted at rest using the AES-256 standard.

Preventing network and application vulnerabilities

Boundary controls monitor and regulate communications at the external boundary of the network and at key internal boundaries. These boundary controls employ rule sets, access control lists, and configurations to enforce the flow of information to specific information system services. Access control lists, or traffic flow policies, are established on each managed interface to regulate the flow of traffic.

Additional controls include:

- Periodic penetration tests by a third-party security firm
- Periodic network scanning
- Policy against use of email for data delivery, mitigating risk from attachments that could contain malware
- Prioritized response for critical security issues (eg, padding oracle on downgraded legacy encryption [POODLE])
- System hosts (virtual instances) deployed as known fixed images

Data sharing by users

BaseSpace Sequence Hub is designed as a collaborative system. Users are responsible for following internal organizational policies for regulating who can share or transfer data. Users share data and grant access rights within the application by sending a request to share to another registered user.

Users can temporarily share data access with technical support when legally permitted. On occasion, troubleshooting or user training can be done without sharing data via screen-sharing or remote computing tools such as GoToAssist or WebEx.

Technical support and quality control (QC)

BaseSpace Sequence Hub is supported by an expert technical support team, accessible by phone or email. All customer contacts, customer events, and responses are tracked.

BaseSpace Sequence Hub provides QC tools for assessing run quality. These tools are also available as standalone software. During analyses, each software application provides QC logs corresponding to the analysis performed

Data center security

BaseSpace Sequence Hub is built on preexisting cloud infrastructure provided by AWS, and therefore shares several AWS standards and accreditations (Table 1). More information on AWS security features is available on the Amazon website.[1]

Illumina employee security practices

Background checks are performed on all Illumina employment candidates in the United States. The background check includes education, university degrees, previous employment, and criminal records. Documented policies and procedures are in place to guide personnel in preventing, detecting, containing, and correlating security violations.

A security awareness and training program communicates Illumina security policies to employees that support BaseSpace Sequence Hub. An automated compliance monitoring system tracks employee compliance with training requirements. All Illumina employees supporting BaseSpace Sequence Hub are aware of disciplinary action for failure to comply with Illumina security policies.

All Illumina personnel that support BaseSpace Sequence Hub are trained annually on appropriate handling of customer data. Download of customer data is restricted. Illumina personnel are granted access to the BaseSpace Sequence Hub systems on an as-needed basis. Access to the system is logged and documented in an automated ticketing system.

When personnel leave Illumina, access to the production environment, Illumina applications, and IT systems is revoked. All equipment and badges owned by Illumina are also returned.

Table 1: Amazon Web Services standards and accreditations

Feature	Description
Service Organization Controls 1/SSAE 16/ISAE 3402	An audit verifying that AWS controls to protect customer data are properly designed and that the individual controls are operating effectively.
Federal Information Security Management Act (FISMA) Moderate	An accreditation granted by the US Government to strengthen federal information system security. For reference, the NIH data centers are rated FISMA moderate.
Payment Card Industry Data Security Standard Level 1	A standard setup to increase electronic payment security. AWS is rated at the highest level.
ISO 27001	A widely recognized international security standard that specifies security management best practices and comprehensive security controls.
Federal Information Processing Standard Publication 140-2	A US government computer security standard that specifies the requirements for cryptography modules.

HIPAA security and privacy

BaseSpace Sequence Hub Enterprise was designed and implemented to be HIPAA-compliant. The United States Congress enacted HIPAA in 1996,[3] and, thereafter, the Department of Health and Human Services (HHS) implemented multiple regulations to carry out the law in practice. Among other things, HIPAA established national standards for the security and privacy of protected health information (PHI). Major provisions for HIPAA include the Security Rule, Privacy Rule, and Breach Notification Rule. Visit the HHS Health Information Privacy website[3] for more information on HIPAA, its history, and links to the specific HIPAA regulations.

HIPAA compliance

HIPAA Security Rule requirements include administrative, physical, and technical safeguards to ensure the confidentiality, integrity, and availability of electronic protected health information (ePHI). BaseSpace Sequence Hub offers several features and controls to meet the requirements of the HIPAA Security Rule (Table 2).

Customer responsibilities under HIPAA

Customers required to comply with HIPAA are ultimately responsible for ensuring that they have a HIPAA-compliant program in place and that they use BaseSpace Sequence Hub in a manner to ensure their compliance.

Illumina will execute a Business Associate Agreement with BaseSpace Sequence Hub Enterprise customers upon request. Illumina also has a Business Associate Agreement with AWS for BaseSpace Sequence Hub. Establishing a Business Associate Agreement with Illumina helps support customer HIPAA compliance.

GDPR privacy and compliance

BaseSpace Sequence Hub is designed and built to meet the standards laid out in the GDPR by the EU. This extends to all global instances of BaseSpace Sequence Hub.

GDPR Compliance

GDPR extends the scope of EU data protection to all companies processing data of EU residents. As a data processor, BaseSpace Sequence Hub is designed to comply with GDPR requirements (Table 2).

Customer Responsibilities under GDPR

European customers are ultimately responsible for ensuring that they have a GDPR-compliant practice in place. As data controllers, customers should use BaseSpace Sequence Hub in a manner that ensures their compliance.

Table 2: Security controls in BaseSpace Sequence Hub

Security Control	Description
Administrative Controls	Policies and procedures to prevent, detect, contain, and correct security violations
	Security official responsible for developing and implementing security policies and controls
	Procedures to make sure that workforce member access to customer data is appropriate and approved
	Processes to authorize access to customer data
	Workforce members trained for security policies
	Processes for incident reporting
	Periodic evaluation of environmental and operational changes that impact the security of the data
	Privacy Impact Assessments (PIAs) performed for all new features that handle user data
Physical Controls	Implemented facility access controls
	BaseSpace Sequence Hub hosted in secure data centers
	Policies regarding workstation security
	Policies and procedures for mobile devices
	Maintained inventory of devices supporting BaseSpace Sequence Hub
Technical Controls	Unique user ID for each user
	User authentication by BaseSpace Sequence Hub or the identity management system of the customer organization
	Protection of integrity of data in transit
	Transport Layer Security–based encryption in transit
	User-initiated data deletion capability

Customer-implemented security controls

The use of BaseSpace Sequence Hub puts several responsibilities in the hands of the customer, which aligns with the AWS model of shared responsibility. Customers should perform risk assessments to account for the use of SaaS solutions, and outcomes of the risk assessment should be reflected in a review of privacy and security controls for each customer.

Customer policies should be reviewed to reflect the use of SaaS solutions. For example, password policies should prohibit the sharing of BaseSpace Sequence Hub accounts and passwords. Institutions should establish processes and procedures for the approval of access and implement regular reviews of access that has been granted to all users. Additionally, customers should review and establish best practices encompassing the content of the data submitted to BaseSpace Sequence Hub. For example, naming policies should prohibit the introduction of identifying subject information.

Workstations used to access BaseSpace Sequence Hub should have proper protections installed, such as antivirus software, host-based firewalls, centralized logging, etc. Business continuity and disaster recovery plans should be updated to account for the use of BaseSpace Sequence Hub.

Breach notification

BaseSpace Sequence Hub customers are responsible for notifying individuals whose data may have been compromised as part of a breach. An audit trail API will be made available to Enterprise administrators, containing information about every instance of attempts to access user data (all file types that are potentially accessible). This includes invalid logon attempts, logoffs, downloads, views, and shares. The log includes date, time, user, and a description of each action. The description of data modification comprises the name of the tool, or the API call, used to modify the data. An API enables users to administer the audit log in an external system.

CLIA and CAP

Many Illumina customers perform sequencing on human samples. Such laboratories are under the authority of the Centers for Medicare and Medicaid Services (CMS),[4] as described by the Clinical Laboratory Improvement Amendments of 1988 (CLIA Regulations).[5] The CLIA regulations establish quality standards for laboratory testing performed on human specimens for diagnosis, prevention, treatment of disease, or assessment of health.

CLIA regulations are designed to ensure the accuracy, reliability, and timeliness of test results. Regulations include quality standards for proficiency testing, test management, quality control, personnel qualifications, and quality assurance.

Clinical labs can choose to be evaluated under more rigorous standards set by the College of American Pathologists (CAP).[6] From a regulatory perspective, CAP standards have been recognized as above and beyond what is required by CLIA regulations. Therefore accreditation by CAP is formally deemed by CMS to certify compliance with CLIA regulations as well.

BaseSpace Sequence Hub support for CLIA and/or CAP

CLIA and/or CAP labs can use BaseSpace Sequence Hub to store, manage, and analyze sequencing data. Use of BaseSpace Sequence Hub does not require CLIA and/or CAP validation because it does not interpret data received from health care providers. BaseSpace Sequence Hub provides several key features that enable labs to ensure data integrity, accuracy, and reliability. The ability to demonstrate reproducibility and to track the origin of analysis results enables customer adherence to CLIA and/or CAP standards:

- A checksum is performed on all data uploaded directly from the sequencing instrument to ensure integrity with the source data
- All data, including genomic data, in S3 is immutable
- BaseSpace Sequence Hub apps are version-controlled; procedures are in place to prevent modification to published apps
- Functions that can alter the interpretation of a result are versioned; users can continue to use the previous version until a new round of validation is complete
- Detailed logs describe every analysis performed

Learn more

To learn more about BaseSpace Sequence Hub, or to sign up for a free BaseSpace account, visit www.illumina.com/informatics.html.

References

1. Amazon Web Services. aws.amazon.com. Accessed July 19, 2018.
2. BaseSpace Sequence Hub User Terms of Use. basespace.illumina.com/agreements/current/details?category=USER. Accessed July 19, 2018.
3. HHS Health Information Privacy. www.hhs.gov/hipaa/. Accessed July 19, 2018.
4. Centers for Medicate and Medicaid Services. www.cms.gov. Accessed July 19, 2018.
5. Clinical Laboratory Improvement Amendments (CLIA). www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/index.html. Accessed July 19, 2018.
6. CAP Guidelines. www.cap.org/web/home/protocols-and-guidelines/cap-guidelines/current-cap-guidelines. Accessed July 19, 2018.

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

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