



SLAM DUNK

SLAMseq Data Analysis Pipeline
SLAMdunk on BlueBee® Genomics Platform

User Guide

Catalog Numbers:
061, 062 (SLAMseq Kinetics Kits)
015 (QuantSeq 3' mRNA-Seq Library Prep Kits)

FOR RESEARCH USE ONLY. NOT INTENDED FOR DIAGNOSTIC OR THERAPEUTIC USE.

INFORMATION IN THIS DOCUMENT IS SUBJECT TO CHANGE WITHOUT NOTICE.

Lexogen does not assume any responsibility for errors that may appear in this document.

PATENTS AND TRADEMARKS

The SLAMseq and QuantSeq 3' mRNA-Seq Library Prep Kits are covered by issued and/or pending patents. QuantSeq is a trademark of Lexogen. Lexogen is a registered trademark (EU, CH, USA). SLAMseq is a trademark of the Institute of Molecular Biotechnology GmbH (IMBA).

All other brands and names contained in this user guide are the property of their respective owners.

Lexogen does not assume responsibility for patent infringements or violations that may occur with the use of its products.

LIABILITY AND LIMITED USE LABEL LICENSE: FOR RESEARCH USE ONLY

This document is proprietary to Lexogen. The SLAMseq and QuantSeq kits as well as SLAMdunk data analysis pipeline are intended for use in research and development only. They need to be handled by qualified and experienced personnel to ensure safety and proper use. Lexogen does not assume liability for any damage caused by the improper use or the failure to read and explicitly follow this user guide. Furthermore, Lexogen does not assume warranty for merchantability or suitability of the product for a particular purpose.

The purchase of the product does not convey the rights to resell, distribute, further sublicense, repackage, or modify the product or any of its components. This document and its content shall not be used or distributed for any other purpose and/or otherwise communicated, disclosed, or reproduced in any way without the prior written consent of Lexogen.

For information on purchasing additional rights or a license for use other than research, please contact Lexogen.

WARRANTY

Lexogen is committed to providing excellent products. Lexogen warrants that the product performs to the standards described in this user guide up to the expiration date. Should this product fail to meet these standards due to any reason other than misuse, improper handling, or storage, Lexogen will replace the product free of charge or issue a credit for the purchase price. Lexogen does not provide any warranty if product components are replaced with substitutes.

Under no circumstances shall the liability of this warranty exceed the purchase price of this product.

We reserve the right to change, alter, or modify any product without notice to enhance its performance.

LITERATURE CITATION

SLAMdunk was developed in a collaboration between the Zuber Group at the Research Institute of Molecular Pathology (IMP), the von Haesler Group at the Max F. Perutz Laboratories (MFPL) and the Ameres Group at the Institute of Molecular Biotechnology (IMBA) at the Vienna Biocenter. An opens source version of SLAMdunk was made available on Github <https://github.com/t-neumann/slamdunk/>.

When describing a procedure for publication using this Pipeline product, please refer to it as Lexogen's SLAMdunk integrated data analysis pipeline on Bluebee platform.

For citing a peer reviewed scientific publication of the whole workflow involving the SLAMseq method, QuantSeq kit and SLAMdunk data analysis please refer to: Herzog V. et. al., (2017) Thiol-linked alkylation to assess expression dynamics. *Nature Methods*, doi:10.1038/nmeth.4435.

CONTACT INFORMATION

Lexogen GmbH

Campus Vienna Biocenter 5
1030 Vienna, Austria
www.lexogen.com
E-mail: info@lexogen.com

Support

E-mail: support@lexogen.com
Tel. +43 (0) 1 3451212-41
Fax. +43 (0) 1 3451212-99

Table of Contents

1. Overview	4
2. Registration	5
3. Get Connected	6
3.1 BlueBee® Service Connector	6
3.2 Cloud Connectors	7
3.2.1 BaseSpace	7
3.2.2 Google	8
3.2.3 Amazon	8
4. Run Your Pipeline.	10
4.1 Sample Selection and Pipeline Configuration	11
4.2 Workflow.	11
5. Explore Your Results	13
6. Support.	16
7. Data Security	17
8. Revision History	17

1. Overview

BlueBee® Genomics provides an easily accessible and secure platform for performing automated analysis of SLAMseq sequencing data. The SLAMseq Data Analysis Pipeline, SLAMdunk, is available for human and mouse samples, and are adapted specifically for the analysis of SLAMseq sequencing data. Access to the Data Analysis Pipeline is possible with the activation codes purchased at our [Lexogen webstore](#).

The SLAMdunk software on the BlueBee® platform allows easy analysis of sequencing data from a SLAMseq experiment!

Register and Activate Your Code

Activation codes for SLAMdunk can be purchased at our [Lexogen webstore](#). Enter this code when registering on the BlueBee® Genomics Platform (see **2. Registration**).

Connect and Upload Your Data

Zipped individual FASTQ files are the required input for SLAMdunk. Uploading and downloading data is handled by the BlueBee® Service Connector, a Cloud Connector (BaseSpace, Google, Amazon), and/or the Drag & Drop browser option (see **3. Get Connected**).

View and Retrieve Your Data

Results of the analysis pipeline can be viewed and downloaded via the web interface as well as through the BlueBee® Service and Cloud Connectors (BaseSpace, Google, Amazon) (see **3. Get Connected**). Your data and results are not stored indefinitely and should be downloaded promptly after the pipeline run is completed (see **5. Explore Your Results**).

Define Data Access

Access to data is controlled through the concept of projects, in which user access privileges can be defined. Data is always owned by a specific account. The administrators of this account can share projects, setup specific sub-projects, link the appropriate data available for designated users, and define user permissions (e.g., viewing only, or with upload/download and data processing permissions). For data security information, see section **7. Data Security**.


2. Registration

For your registration, please keep your kit activation code handy as you will need it to verify your access. The code includes a combination of letters and numbers in the following format:

7YTM-D2IP-QAMX-HAPT

To register yourself as a first time user of the BlueBee® Analysis Platform, go to: <https://lexogen.bluebee.com/slamdunk>.

First, click **Register** to begin the registration process.



Research Use Only
powered by **BlueBee**

Figure 1. Login page.

Enter your email address and activation code in the required field and click **NEXT**, and fill out all other data to complete the registration.

NOTE: A run describes the data analysis run for a certain number of samples. The allocated size of the purchased activation code in GB reflects the size per run with a maximum data input (e.g., if you have an activation code for 12 GB, you can run as many samples as you need but not exceeding a total of 12 GB per run). This can be 24 samples of 500 MB each or 48 samples of 250 MB each.

IMPORTANT: Please note your data allocation requirements before purchasing an analysis package. Unused GB cannot be reused/recycled.

3. Get Connected

There are three options to upload your FASTQ files: BlueBee® Service Connector, Cloud Connector, and Drag & Drop.

3.1 BlueBee® Service Connector

Once registered, you can install and configure the BlueBee® Connector. Go to **Settings** in the top right corner and scroll down to **Local Connector Settings**.

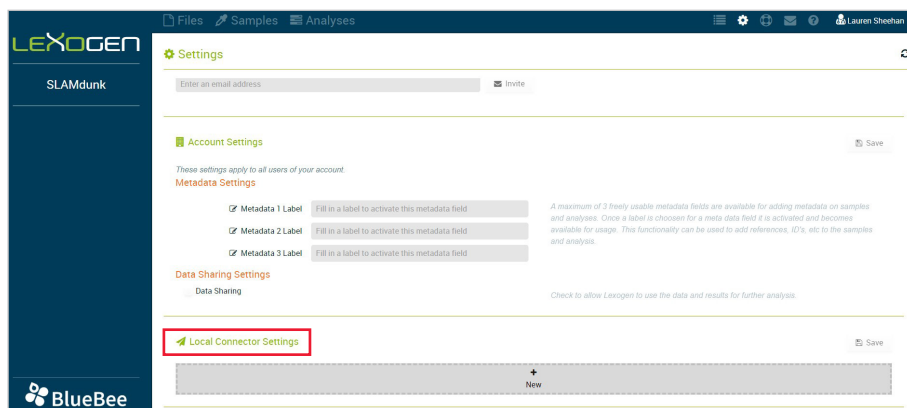


Figure 2. Local Connector settings.

Click **New** to add information on the BlueBee® Service Connector. The BlueBee® Service Connector is a lightweight piece of software enabling several key functionalities and security features for use of the BlueBee® service. The connector uniquely identifies the client and facilitates end-to-end process auditing. It collects data coming straight from the specified server, computer, or sequencer and initiates the transfer of raw data in an automated way (no manual intervention required). Additionally, the software encrypts, compresses the raw sequence data and drastically accelerates the data transfer to the appropriate computing centre. The software not only encrypts but also compresses the data, thereby drastically accelerating the data transfer to the appropriate computing centre.

After clicking **New**, configure what folder the connector pulls FASTQ files from (**upload location**), and what folder the connector drops output files you scheduled for download (**download location**). Please note, if you would like the connector to only upload or download data, leave the respective location empty.

After filling in the upload and download details, press **Save** at the top right of the screen. The connector software for the operating system you choose will be downloaded automatically.

You will see the progress of your uploads on the left of your homepage under **Local Connector Uploads**. Once files are finished uploading, they will be located under the **Files** tab.

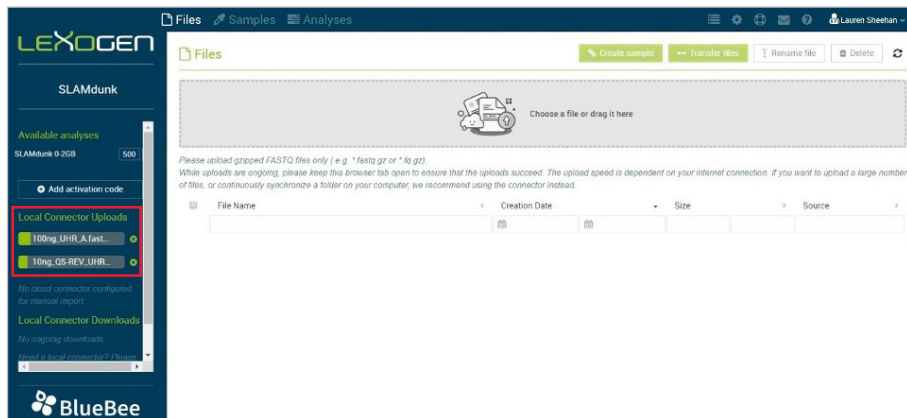


Figure 3. Uploads in progress with the BlueBee® Service Connector.

3.2 Cloud Connectors

You can import data to BlueBee® from the following cloud providers: BaseSpace, Google, and Amazon. To begin, click on the + button on the respective connector you wish to import data from.

PLEASE NOTE: It is possible to have multiple Cloud Connectors. For each Cloud Connector added, provide the connector with a unique name.

To make a connection between the bucket or project located on the cloud provider, specific information is required. Instructions on how to find this information is listed below under each respective provider. After entering this specific information, continue with creating an import rule.

3.2.1 BaseSpace

In the instructions below, we are assuming the user already has a project containing data. If you still need to create a project and upload data, please consult the [BaseSpace help pages](#). To connect the BlueBee® Cloud Connector with BaseSpace you will need:

- **URL:** The user can keep the predefined URL
- **Access Token:** BaseSpace provides [online information](#) on how to request your token (scroll down to 'Steps for developing your first App') or you can follow the next steps:
 1. Go to the *My Apps* page within the BaseSpace developers environment or follow [this link](#).
 2. If you haven't created an application yet, follow these additional steps first:
 - a. Click on 'Steps for developing your first App'

- b. Click on 'Register'
 - c. Accept the Terms and Conditions
3. Click the 'Create a new Application' button
 - a. Name: has to be unique, e.g. company name + 'BlueBee connector'
 - b. Provide your company or organization name
 - c. As App Type select 'Web/Desktop = Other'
 - d. Short Description: e.g. 'BlueBee connector'
 - e. Click Create Application
 - f. Fill in the Release Notes and Long Description as desired
 - g. Select a Category as desired
 - h. Select 'read' for the Application Permission Settings
 - i. 'Save'
4. Scroll up to find the *Credentials* tab, copy the Access Token and paste it into the corresponding field in the Cloud Connector on the BlueBee® platform.

3.2.2 Google

To connect to Google, you will need a JSON file with your credentials.

1. On the [Google Cloud Platform](#), navigate to the '[Service accounts](#)' page under *AIM & Admin* section
2. Follow the steps for creating a new service account
 - The service account permissions should at least contain the role 'Storage Object Viewer' for Storage
3. Create a key for the service account of type JSON
4. Download the JSON file and upload this in the Cloud Connector for Google on the BlueBee® platform

3.2.3 Amazon

Amazon provides information on how to get your access token in their [documentation online](#). Follow the instructions for creating an IAM user (Console) and use the characteristics below while setting up the user:

- Type of access = Programmatic access
- Minimal permission for user group = ListBucket & GetObject

In the last step of the user creation, you can either copy the Access Key ID and the secret Access key or download the CSV file containing this information necessary for configuring the cloud connector on the platform.

Import Rule for Cloud Connectors

Depending on which cloud provider you want to import from, other fields might be shown in the import rule.

- **Automatic import:** if ticked, files uploaded to this cloud project/bucket will immediately imported to BlueBee®. To import files manually, untick this box and use the 'List' button.
- **Unique Name:** name you want to give to this import rule.
- **Bucket (Amazon, Google):** Name of the bucket containing data you want to import.
- **BaseSpace project (BaseSpace):** Name of the project containing data you want to import from BaseSpace.

Save the import rule after configuring it. It can be edited afterwards.

Importing Data from a Cloud Connector

There are two way of importing data: Automatically or Manually

Automatic Import: This can be enabled with a tickbox in the import rule. When you upload files into your cloud project/bucket, these files will immediately be imported to the BlueBee® platform as soon as the upload is complete.

Manual Import: Manual import can be used for several reasons such as:

- You do not want all of the files in your cloud bucket/project to be imported to BlueBee® . In this case, make sure automatic import is disabled.
- You want to import data that has been imported before.

For both actions you can use the 'List' button present in the import rule of the cloud connector. By default, all files will be listed that match your import rule and that haven't been imported yet. To view files that have been imported before, matching the import rule, enable the tickboxes on top.

4. Run Your Pipeline

Once data is uploaded, check off the small box to the left of the file you wish to analyze (Figure 4). If you wish to merge several files into one sample, please check off those boxes and click **Create Sample** and select Multi-file or Single-file sample (Figure 5). We recommend merging files when a sample was sequenced in two or more lanes, or if the files are technical replicates. Enter information related to your file(s) (i.e., species, Lexogen kit, description of files). Once finished, click the **Create** button on the bottom left. Your samples can now be found under the **Samples** tab.

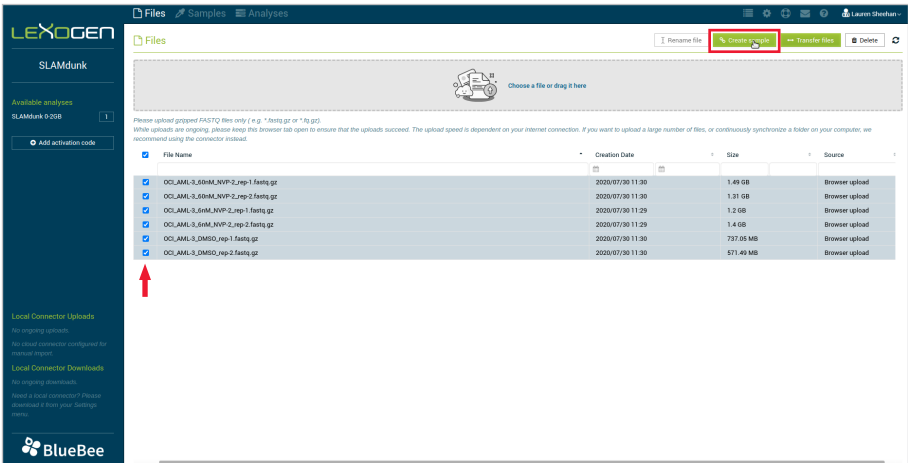


Figure 4. Selection of fastq files.

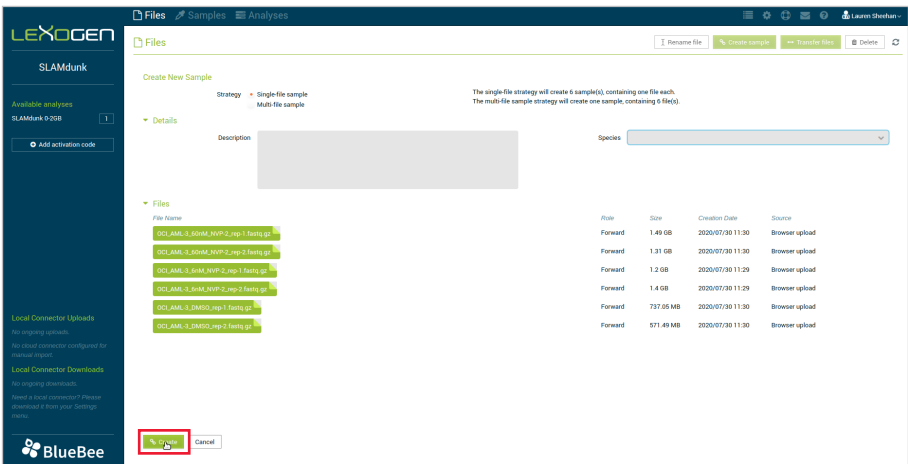


Figure 5. Creating a new sample.

4.1 Sample Selection and Pipeline Configuration

Prior to executing SLAMdunk, the user must select the files which are to be processed. Ideally, the user will select all samples belonging to an experiment/timecourse. This guarantees that summaries and comparative overviews are derived for all conditions of interest. To do this, under the **Samples** tab, check off the small box to the left of the samples you wish to analyze together. Additionally, select the minimum fraction to call a variant. The default for this is 0.8. Finally, you will need to input an Analysis ID and reference genome. Once complete, click the **Analyses** button to start your run.

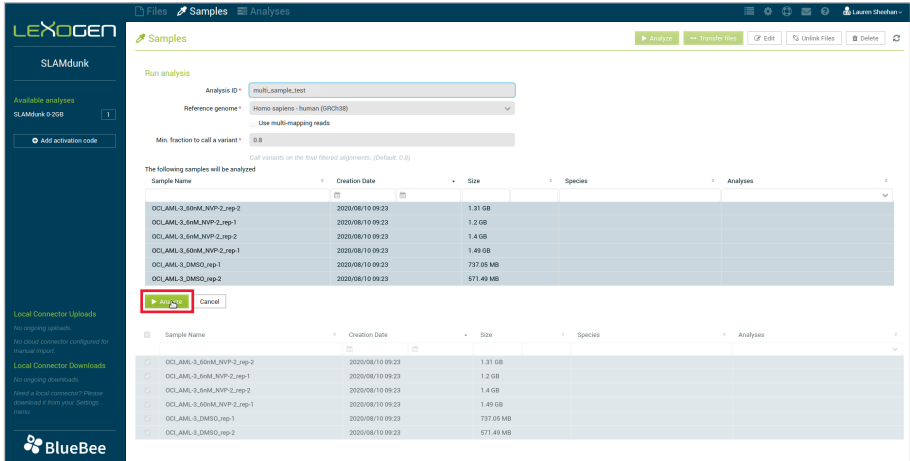


Figure 6. Analyzing sample(s).

4.2 Workflow

Once a new run has been started, the SLAMdunk pipeline executes 5 major steps:

1. The *map* step is used to map reads to a given genome using NextGenMap's SLAMSeq alignment settings. Output files from this step will be individual bam files for each of the samples in the input.
2. The *filter* step is used to filter the raw alignments from the *map* step using multiple quality criteria to obtain the final alignments for all subsequent analyses. Results of this step can be found in directory "filter".
3. The *snp* step is used to call variants on the final filtered alignments of the *filter* step. Any called T->C SNPs from this step will be excluded in the subsequent analyses to reduce the false-positive number. Output files from this step will be vcf files for each of the samples in the input. Results of this step can be found in directory "snp".
4. The *count* step calculates all relevant numbers on statistics of SLAMSeq reads for each given 3'UTR. Central output will be *tcount* files for the samples in the input. Results of this step can

be found in directory “count”.

- The alleyoop step will produce summary statistics, a PCA for all samples, and several conversion rate plots (read based, UTR based, read position based). Results of this step can be found in file summary_PCA.pdf and directories “multiqc” and “stats”.

Under the **Analyses** tab, all runs will appear and progress can be monitored under the **Status** column. Once complete, **Succeeded** will appear under the Status column and you can click the sample name under the **Analysis ID** column for details of that run.

NOTE: Please refresh your browser to obtain real-time status on the progress of uploads under the Analyses tab.

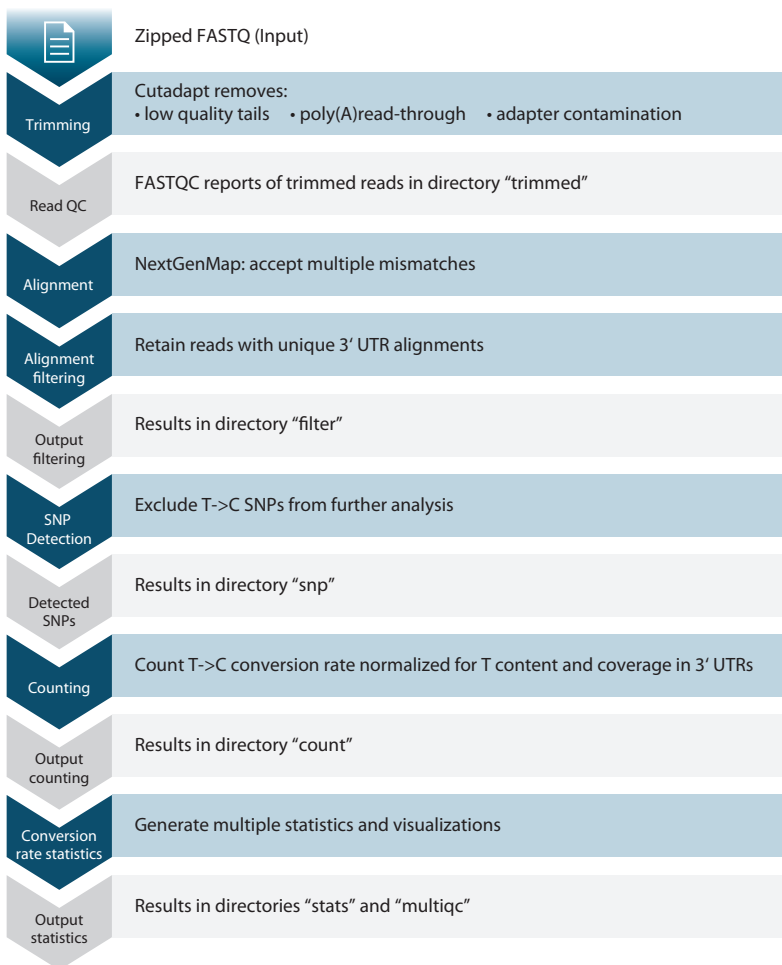


Figure 7. The SLAMdunk Analysis Pipeline Workflow.

5. Explore Your Results

Under the **Analyses** tab, you can click a sample to view the results. Navigate this section by clicking one of the options on the left side: Analysis Details, MultiQC, and Output Files.

Analysis Details: multi_sample_test - SLAMdunk 2.0.4 - 2020/08/04 11:13

Analysis Detail

Analysis ID: multi_sample_test Start Date: Aug 4 2020
 Pipeline: SLAMdunk 2.0.4 Runtime: 17h 11m
 Status: Analysis succeeded

Input

Min. fraction to call a variant: 0.8
 Call variants on the final filtered alignments. (Default: 0.8)

Use multi-mapping reads

Sample Name	Creation Date	Size	Species	Analyses
OCL_AML-3_DM50_rep1	2020/08/04 11:11	737.65 MB		1/2
OCL_AML-3_60MM_NVP-2_rep1	2020/08/04 11:11	1.49 GB		1/2
OCL_AML-3_60MM_NVP-2_rep1	2020/08/04 11:11	1.2 GB		1/2
OCL_AML-3_60MM_NVP-2_rep2	2020/08/04 11:11	1.31 GB		1/2
OCL_AML-3_60MM_NVP-2_rep2	2020/08/04 11:11	571.49 MB		1/2
OCL_AML-3_60MM_NVP-2_rep2	2020/08/04 11:11	1.4 GB		1/2

Metadata

Species:
 Reference genome: Homo sapiens - human (GRCh38)

Figure 8. SLAMdunk Analyses data output.

MultiQC
 sample analysis

Report generated on 2020-08-04, 20:08 based on data in: /yyer/11b/csd/r1/g1/278/0368-3882-6492-8ee2-5263174999ccr/SLAMdunk_Selcher_080303

General Statistics

Sample Name	Counted	Retained	Mapped	Sequenced	% Trimmed	% Dups	% GC	M Segs
OCL_AML-3_60MM_NVP-2_rep-1	12.73 M	21.44 M	25.83 M	27.13 M	0.7%	76.0%	46%	27.1
OCL_AML-3_60MM_NVP-2_rep-2	11.11 M	18.35 M	22.17 M	23.21 M	0.5%	68.7%	46%	23.2
OCL_AML-3_60M_NVP-2_rep-1	9.43 M	15.63 M	20.09 M	20.81 M	0.5%	59.6%	45%	20.8
OCL_AML-3_60M_NVP-2_rep-2	11.38 M	19.67 M	23.59 M	24.55 M	0.5%	58.3%	45%	24.5
OCL_AML-3_DM50_rep-1	4.88 M	10.23 M	12.38 M	12.95 M	1.2%	64.3%	46%	13.0
OCL_AML-3_DM50_rep-2	4.08 M	8.07 M	9.61 M	10.02 M	0.5%	68.4%	46%	10.0

Slamdunk
 Slamdunk is a tool to analyze SLAMseq data.

Filter statistics
 This table shows the number of reads filtered with each filter criterion during filtering phase of slamdunk.

Figure 9. Example of MultiQC output in environment.

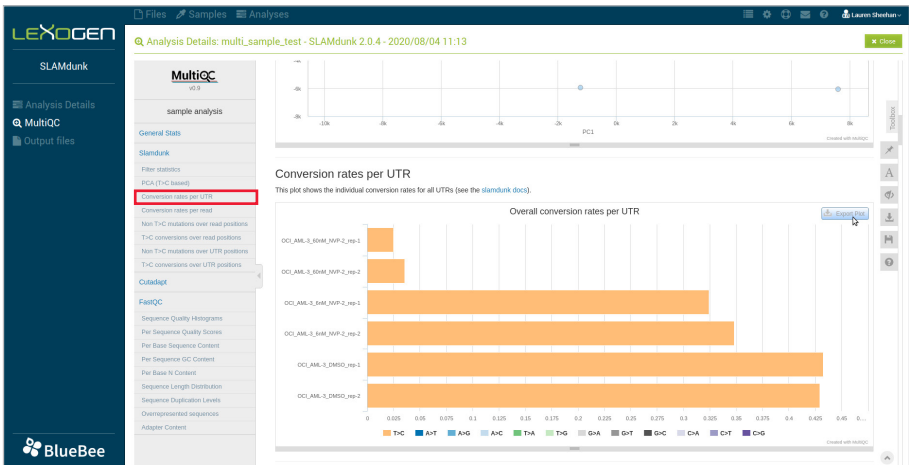


Figure 10. Example of downloadable conversion rate plot.

You can directly download your files using the **Output Files** option. You can either download files with your browser or with your connector.



Figure 11. Downloading your output files.

Alternatively, you can select multiple files and click the **Download** button to perform a batch download.

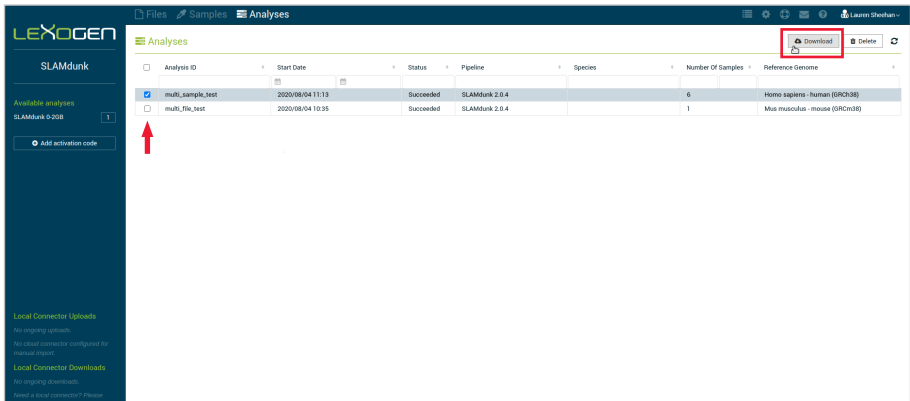


Figure 12. Output data batch downloads.

Data Storage

BlueBee® provides secure cloud-based storage of the results files generated by the pipeline runs. However, results are not stored permanently. We recommend downloading all results files as soon as the pipeline run has finished.

File type	Duration kept in BlueBee
Uploaded FASTQ files	If not used in a pipeline, 6 days after upload
Output .bam files	1 month (and then permanently deleted)
Output files (except .bam files)	1 year

6. Support

If you have any questions, please fill out a support ticket under the **Support** tab or contact us at support@lexogen.com.

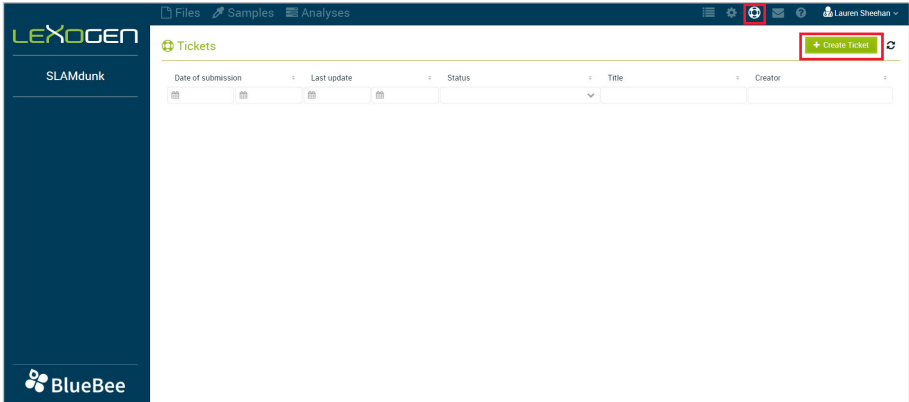


Figure 13. Create a ticket.

Additionally, you can find our frequently asked questions by clicking on the **Help** tab.

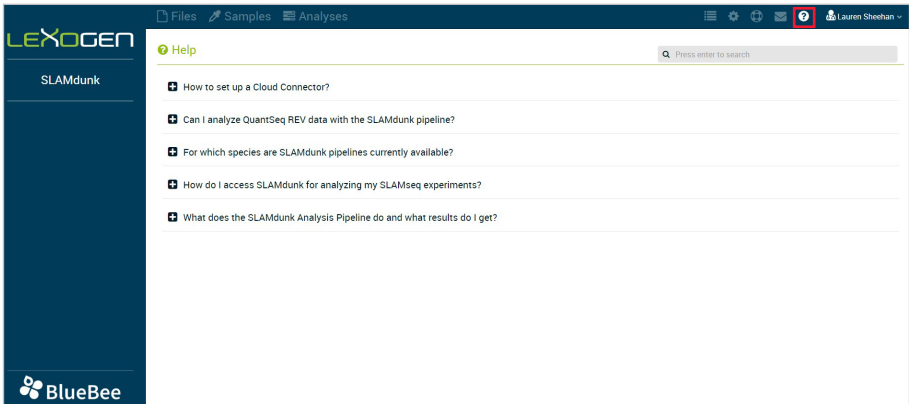


Figure 14. Frequently asked questions.

7. Data Security

BlueBee® can ensure compliance with data protection and security requirements, also for clinical sequencing data. For more information the BlueBee® Genomics Platform Data Security and Compliance Whitepaper can be found in the downloads section on www.bluebee.com. For further information on data security and data protection please contact BlueBee® Technical Support at support@bluebee.com.

8. Revision History

Publication No.	Change	Page
063UG147V0200 Aug. 11, 2020	Updated User Guide to fit the new BlueBee Web Solutions interface - adjustments of all text and images.	1-18
063UG147V0100 Nov. 9, 2017	Initial Release.	

The background of the slide is decorated with a pattern of semi-transparent blue spheres of various sizes, connected by thin, light blue lines that create a network-like structure. The spheres have a glossy, 3D effect with highlights and shadows.

SLAMseq Data Analysis Pipeline SLAMdunk on BlueBee® Genomics Platform

Lexogen GmbH
Campus Vienna Biocenter 5
1030 Vienna, Austria
Telephone: +43 (0) 1 345 1212-41
Fax: +43 (0) 1 345 1212-99
E-mail: info@lexogen.com
© Lexogen, 2020