

Importing 10x Genomics Matrix Files

- [Importing single cell data](#)
 - [Importing matrices into Partek Flow \(this Market Exchange Format is popular for public repositories\)](#)
 - [Importing matrices in h5 format \(this Hierarchical Data Format is recommended for multiple samples\)](#)
- [Importing spatial data](#)
 - [Importing Xenium Output Bundle](#)

Importing single cell data

Partek Flow supports the import of [filtered gene-barcode matrices](#) generated by 10x Genomics' [Cell Ranger pipeline](#).

Below is a video summarizing the import of these files:

Your browser does not support the HTML5 video element

Importing matrices into Partek Flow (this Market Exchange Format is popular for public repositories)

To import the matrices into Partek Flow, create a new project and click **Add data** then select **Import scRNA count feature-barcode-mtx** under **Single cell > scRNA-Seq**.

The screenshot shows the 'Select the format' dialog in the Partek Flow interface. At the top, there are tabs for 'Single cell', 'Bulk', and 'Other'. Under 'Single cell', there are sub-tabs for 'scRNA-Seq', 'Spatial transcriptomics', 'scATAC-Seq', 'V(D)J', and 'Flow/Mass Cytometry'. The 'scRNA-Seq' tab is selected. Below the tabs, the dialog displays five options for data formats, each with a radio button and a description:

- Import scRNA count feature-barcode-mtx** (selected): This sparse matrix output is common for 10x Genomics, Fluent Biosciences and Parse Biosciences. Each sample has 3 files (two .csv with one .mtx or two .tsv with one .mtx for each sample).
- 10x Genomics Cell Ranger counts h5**: This compressed binary format is preferred for 10x Genomics Cell Ranger output. There is 1 filtered .h5 file per sample and multiple files can be selected.
- Full count matrix**: This rectangular cell-by-feature count matrix is common for BD Rhapsody. There is one file for one or more samples (txt, csv, tsv, txt.gz, csv.gz, tsv.gz).
- Seurat Rds**: This R object is for data processed by Seurat (.rds).
- h5ad**: This AnnData object in the h5ad file format is for data processed by Scanpy.
- fastq**: The fastq format is used for unaligned reads. Acceptable file types are fastq, fastq.gz, fastq.bz2, fq, fq.gz, fq.bz2.

Figure 10. Importing single cell data

Samples can be added using the **Add sample** button. Each sample should be given a name and three files should be uploaded per sample using the **Browse** button.

Import three feature-barcode matrix files for each sample.

Samples and files

Import three feature-barcode matrix files (features.tsv, barcodes.tsv and matrix.mtx, or all_genes.csv, cell_metadata.csv and DGE.mtx) for each sample. If necessary, when transferring the files to the server, please remember to follow a helpful naming convention.

+ Add sample

Sample name	Cells	Features	Files	Action
Sample 1	0	0		
Sample 2	0	0		

Feature annotation

☐ Use annotation file

Annotation file can be selected for detecting the fraction of mitochondrial counts

Deduplication method

Condenses features with multiple expression values for a given cell. Mean will calculate the mean expression value for that feature. Maximum will retain the maximum value for each feature and Sum will add up all available values.

☒ Mean
☐ Maximum
☐ Sum

Count value format

☒ Raw count
☐ Normalized count

None

Report

☒ All features
☐ Features with non-zero values across all samples

Cells with total read count at least

A low total read count threshold will result in a large number of cells which might take a long time to import

☒ 400

Back
Finish

Figure 11. Import three feature-barcode matrix files for each sample

If you have not already, transfer the files to the server to be accessed when you click **Browse**. Follow the directions [here](#) to add files to the server. Make sure the files are **decompressed** before they are uploaded to the server.

By default, the Cell Ranger pipeline output will have a folder called `filtered_gene_bc_matrices` (Figure 3). It is helpful to rename and organize the files prior to transfer using the File browser.

There are folders nested within the matrix folder, typically representing the reference genome it was aligned to. Navigate to the lowest subfolder, this should contain three files:

- barcodes.tsv
- genes.tsv
- matrix.mtx

Select all 3 files for import into Partek Flow

Copyright © 2018 by Partek Incorporated. All Rights Reserved. Reproduction of this material without express written consent from Partek Incorporated is strictly prohibited.

Name	Date modified	Type	Size
analysis	3/9/2018 9:23 AM	File folder	
filtered_gene_bc_matrices	3/9/2018 9:23 AM	File folder	
raw_gene_bc_matrices	3/9/2018 9:23 AM	File folder	
cloupe.cloupe	3/9/2018 9:23 AM	CLOUPE File	36,529 KB
filtered_gene_bc_matrices_h5.h5	3/9/2018 9:23 AM	H5 File	9,736 KB
metrics_summary.csv	3/9/2018 9:23 AM	Microsoft Excel ...	1 KB
molecule_info.h5	3/9/2018 9:23 AM	H5 File	191,301 KB
possorted_genome_bam.bam	3/9/2018 9:33 AM	BAM File	32,421,469 ...
possorted_genome_bam.bam.bai	3/9/2018 9:23 AM	BAI File	8,508 KB
raw_gene_bc_matrices_h5.h5	3/9/2018 9:23 AM	H5 File	17,052 KB
web_summary.html	3/9/2018 9:23 AM	Chrome HTML ...	3,710 KB

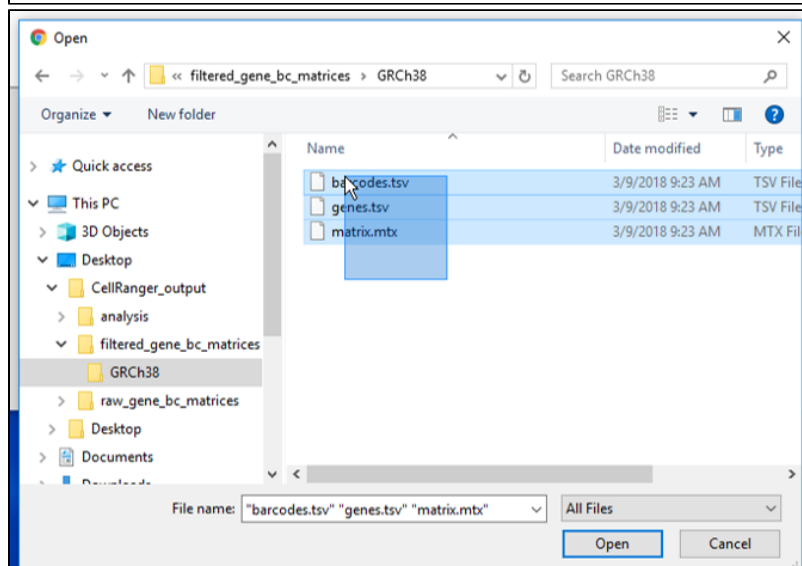


Figure 12. Filtered matrix folder from Cell Ranger pipeline

Specify the annotation file used when running the pipeline for additional information such as mitochondrial counts (Figure 4). Other information can also be specified, such as the count value format. All features can be reported or features with non-zero values across all samples can be reported and the read count threshold can be modified to make the import more efficient.

Feature annotation

☒ Use annotation file

Annotation file can be selected for detecting the fraction of mitochondrial counts

Assembly

Mus musculus (mouse) - mm39

Annotation model

RefSeq Transcripts 98 - 2021-05-05 (Administrator)

Deduplication method

Condenses features with multiple expression values for a given cell. Mean will calculate the mean expression value for that feature. Maximum will retain the maximum value for each feature and Sum will add up all available values.

☒ Mean
 ☐ Maximum
 ☐ Sum

Count value format

☒ Raw count
 ☐ Normalized count

None

Report

☒ All features
 ☐ Features with non-zero values across all samples

Cells with total read count at least

A low total read count threshold will result in a large number of cells which might take a long time to import

☒ 400

Back

Finish

Figure 13. Configuring matrix metadata

Click **Finish** when you have completed configuration. This will queue the import task.

Importing matrices in h5 format (this Hierarchical Data Format is recommended for multiple samples)

The Cell Ranger pipeline can also generate the same filtered gene barcode matrix in h5 format. This gives you the ability to select just one file per matrix and select multiple matrices to import in batch. To import an h5 matrix, select the **Import scRNA full count matrix or h5** option (Figure 1). **Browse** for the files and modify any configuration options. Remember the files need to be transferred to the server.

Select file

Select single cell files

☒ Partek Flow Server
 ☐ URL

No files selected

Browse

To move files from your local computer to the Partek server, please Transfer files first.

Back

Next

Select file

Current directory

/home/flow/FlowData/project_output/user4/uploads

Goto

▼

project_output

▼

GarmentClassifierDownloadResult-5747514633003329

▼

Project_10X_PBMC_CITE-seq_H5_6

▼

Project_10X_mouse_brain_coronal_41

▼

Project_10X_Spatial_40

▼

Project_2_2

▼

Project_Canine_acanthomatous_ameloblasts_11

▼

Project_COVID-19_Severity_large-scale_10

▼

Project_GSE43245_17

▼

Project_GSE93395_16

▼

Project_Haemopoiesis_Trajectory_39

▼

Project_Liver-Muscle_33

▼

Project_MIS-C_study_(GSE183716)_15

▼

Project_Mouse_VDJ_5pGEx_14

▼

Project_NSCLC_(lg_B_cells)_VDJ_34

▼

Project_PBMC_5K_19

▼

Project_PBMC_5K_1_23

▼

Project_SC2108_18

▼

Project_scATAC-Seq_Demonstration_36

▼

Project_scRNA-seq of 23 Korean colorec_3

▼

Project_scRNA-seq Analysis Training Ba_25

▼

Project_Survival Analysis Example_13

▼

Project_Test_45

▼

Project_VDJ_Test_31

▼

Project_Visium_FFPE_Prostate_Cancer_26

▼

Project_Visium_Human_Prostate_FFPE_7

▼

Project_Visium_Spatial_for_FFPE_Demons_1

▼

user4

▼

upl69d6

▼

Coronal_1

▼

Coronal_2

4 files selected

☐	Name	Size
☐	atac_pbmc_1k_nextgem_S1_L001_1_001.fastq.gz	211.50 MB
☐	atac_pbmc_1k_nextgem_S1_L001_R1_001.fastq.gz	797.23 MB
☐	atac_pbmc_1k_nextgem_S1_L001_R2_001.fastq.gz	441.55 MB
☐	atac_pbmc_1k_nextgem_S1_L001_R3_001.fastq.gz	789.02 MB
☐	atac_pbmc_1k_nextgem_S1_L002_1_001.fastq.gz	211.80 MB
☐	atac_pbmc_1k_nextgem_S1_L002_R1_001.fastq.gz	797.17 MB
☐	atac_pbmc_1k_nextgem_S1_L002_R2_001.fastq.gz	440.73 MB
☐	atac_pbmc_1k_nextgem_S1_L002_R3_001.fastq.gz	786.82 MB
☒	filtered_feature_bc_matrix_WTL1.h5	12.33 MB
☒	filtered_feature_bc_matrix_WTL2.h5	12.97 MB
☒	filtered_feature_bc_matrix_WTS1.h5	9.30 MB
☒	filtered_feature_bc_matrix_WTS2.h5	11.14 MB
☐	GSE132465_GEO_processed_CRC_10X_cell_annotation.txt.gz	423.85 KB
☐	GSE132465_GEO_processed_CRC_10X_raw_UMI_count_matrix...	122.97 MB
☐	GSM5712427_SISC_Lymphatics_barcodes.tsv.gz	44.28 KB
☐	GSM5712427_SISC_Lymphatics_features.tsv.gz	284.11 KB
☐	GSM5712427_SISC_Lymphatics_matrix.mtx.gz	48.04 MB
☐	GSM5712428_Colon_Lymphatics_barcodes.tsv.gz	43.25 KB
☐	GSM5712428_Colon_Lymphatics_features.tsv.gz	284.11 KB

Valid files are: txt, csv, tsv, h5, loom, fcs or gz

Continue

Figure 14. Importing matrix in h5 format

This feature is also useful for importing multiple samples in batch. Simply put all h5 files from your experiment on a single folder, navigate to the folder and select all the matrices you would like to import.

Configure all the relevant sample metadata, including sample name and the annotation that was used to generate the matrices, and click **Finish** when completed. Note that all matrices *must have been generated using the same reference genome and annotation* to be imported into the same project.

Importing spatial data

Importing Xenium Output Bundle

Raw output data generated by the 10x Genomics' Xenium Onboard Analysis pipeline consists of decoded transcript counts and morphology images. The raw output and other standard output files derived from them are compiled into a zipped file called Xenium Output Bundle.

To import the **Xenium Output Bundle** into Partek Flow, create a new project and click **Add data**, then select **Import 10x Genomics Xenium** under **Single cell > Spatial**, click **Next**.

Single cellBulkMicroarrayOther

scRNA-SeqSpatialscATAC-SeqV(D)JFlow/Mass Cytometry

Select the format

☐ 10x Genomics Visium Space Ranger output
10x Genomics Space Ranger output can be count matrix data as 1 filtered .h5 file per sample or sparse matrix files for each sample as 3 files (two .csv with one .mtx or two .tsv with one .mtx for each sample). The spatial output files should be in compressed format (.zip). The high resolution image can be uploaded and is optional.

☒ 10x Genomics Xenium
10x Genomics Xenium data should include the unzipped Xenium Output Bundle with the preferred input image file (TIFF) for each sample.

☐ NanoString CosMx
NanoString CosMx data should include 5 files (exprMat_file.csv, metadata_file.csv, polygons.csv, tx_file.csv, fov_positions_file.csv) and an image folder (CellComposite) per sample

☐ 10x Genomics Visium fastq
Unaligned fastq reads (fastq, fastq.gz, fastq.bz2, fq, fq.gz, fq.bz2) can be processed using the 10x Genomics Space Ranger task. Please follow a naming convention only containing letters, digits, underscores and dashes.

BackNext

Figure 15. Importing Xenium spatial data

Samples can be added using the **Add sample** button. Each sample should be given a name and a folder containing the required 6 files: cell_feature_matrix.h5, cells.csv.gz, cell_boundaries.csv.gz, nucleus_boundaries.csv.gz, transcripts.csv.gz, morphology_focus.ome.tif should be uploaded per sample using the **Browse** button. The required 6 files should be all included in the Xenium Output Bundle folder.

Samples and files

Specify a folder containing the required 6 files: cell_feature_matrix.h5, cells.csv.gz, cell_boundaries.csv.gz, nucleus_boundaries.csv.gz, transcripts.csv.gz, morphology_focus.ome.tif

[+ Add sample](#)

Sample name	Cells	Features	Folder	Action
Sample 1	0	0		
Sample 2	0	0		

Feature annotation

☐ **Use annotation file**
Select the file that has been used to generate the feature counts (e.g. gene or protein information).

Deduplication method
If the feature ID is not unique, the feature will be summarized by the selected method.

☒ Mean
 ☐ Maximum
 ☐ Sum

Data format

Count value format

☒ Raw counts
 ☐ Normalized counts with log base

Filtering

Features to report

☒ All features
 ☐ Features with non-zero values across all samples

☒ **Report cells with total read count at least**
 Disabling this option or choosing a low count threshold might result in a long import time.

[Back](#)
[Finish](#)

Figure 16. Import Xenium Output Bundle folder for each sample

If you have not already, transfer the files to the server to be accessed when you click **Browse**. Follow the directions [here](#) to add files to the server. You will need to **decompress** the Xenium Output Bundle zip file before they are uploaded to the server. After decompression, you can **drag and drop** the entire folder into the Transfer files dialog, all individual files in the folder will be listed in the Transfer files dialog after drag & drop, with no folder structure. The folder structure will be restored after upload is completed.

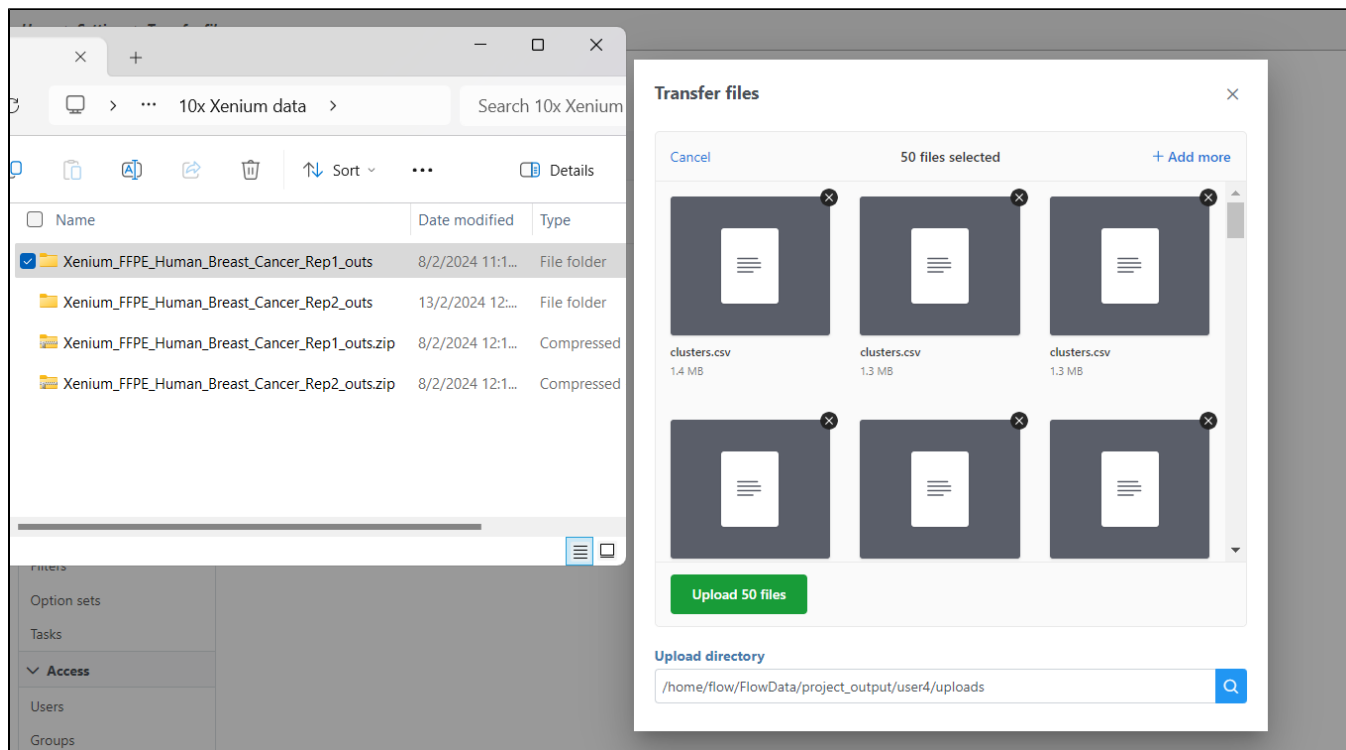


Figure 17. Drag & drop unzipped Xenium Output Bundle folder into Transfer files dialog

Once you have uploaded the folder into the server, you can continue to select the folder for each sample from **Browse**. Once the folder is selected, the *Cell* and *Features* values will auto-populate. You can choose an annotation file that matches what was used to generate the feature count. Then, click **Finish** to start importing the data into your project.

Samples and files
Specify a folder containing the required 6 files: cell_feature_matrix.h5, cells.csv.gz, cell_boundaries.csv.gz, nucleus_boundaries.csv.gz, transcripts.csv.gz, morphology_focus.ome.tif

+ Add sample

Sample name	Cells	Features	Folder	Action
Sample 1	167780	541	outs	
Sample 2	118752	541	outs	

Feature annotation
☒ Use annotation file
Select the file that has been used to generate the feature counts (e.g. gene or protein information).

Assembly
Homo sapiens (human) - hg38

Annotation model
Ensembl Transcripts release 105 (Administrator)

Primary feature identifier
☒ Feature name (Values: ABCC11, ACTA2, ACTG2, ADAM9, ADGRES, ADH1B, ADIPOQ,...)
☐ Feature ID (Values: ENSG00000121270, ENSG00000107796, ENSG00000163017,...)

Deduplication method
If the feature ID is not unique, the feature will be summarized by the selected method.
☒ Mean ☐ Maximum ☐ Sum

Data format
Count value format
☒ Raw counts ☐ Normalized counts with log base

None

Filtering

Features to report
☒ All features
☐ Features with non-zero values across all samples

☒ Report cells with total read count at least
Disabling this option or choosing a low count threshold might result in a long import time.

1

Back Finish

Figure 18. Add Xenium Output Bundle and select annotation

Additional Assistance

If you need additional assistance, please visit [our support page](#) to submit a help ticket or find phone numbers for regional support.



Your Rating:  Results:  26 rates