## Import a GEO / ENA project

- How to import a study from GEO / ENA
  Common Issues
- - · Error Message The project did not yield any data. Double-check the project ID, or try importing the data manually
  - The project was imported, but the Analyses tab is empty and there are no FASTQ files
  - Something is missing or the import failed
- FAQ
  - What are GEO and ENA?
    - How do I know if a GEO project is also in ENA?

## How to import a study from GEO / ENA

If a project is publicly available in the Gene Expression Omnibus (GEO) and European Nucleotide Archive (ENA) databases, you can import associated FASTQ files and sample attributes automatically into Partek Flow.

• On the Homepage click New Project to create a project and give the project a name



Click Add data

Analyses	Metadata	Log	Project settings	Notebook	Data viewer	Attachments
Add data	Welcome to You'll need s Click the blu For more he	your pro samples Je button	iject. before you can run a on the left then choo e watch this video (	n analysis. se your assay a l) or see our doo	and file type. cumentation.	

• Select fastq as the file type after choosing between Single cell or Bulk as the assay types

Sin	gle cell	Bulk

- Click Next .
- Choose GEO / ENA
- Enter the BioProject ID of the data set you would like to download. The format of a BioProject ID is PRJNA followed by one to six numbers (e.g. PRJNA381606)

File select		
O Partek Flow Serve	er 🔿 URL 🔘 GEO / ENA	
BioProject ID 🕖	PRJNA381606	
Upload directory	/home/flow/FlowData/project_output	

A GEO ID can also be used in the format GSE followed by one to five numbers (e.g. GSE71578).

• Click Finish

It may take a while for the download to complete depending on the size of the data. FASTQ files are downloaded from the ENA BioProject page.

• FASTQ files will be added as an Unaligned reads data node in the Analyses tab

Analyses	Metadata	Log	Project settings	Notebook	Data viewer	Attachments		
Import sample files	Unaligned reads							

### **Common Issues**

# Error Message - The project did not yield any data. Double-check the project ID, or try importing the data manually

If the study is not publicly available in both GEO and ENA, project import will not succeed.

#### The project was imported, but the Analyses tab is empty and there are no FASTQ files

If there is an ENA project, but the FASTQ files are not available through ENA, the project will be created, but data will not be imported.

#### Something is missing or the import failed

A variety of other issues and irregularities can cause imports to not succeed or partially succeed, including, but not limited to, a BioProject having multiple associated GSE IDs, incomplete information on the GEO or ENA page, and either the GEO or ENA project not being publicly available.

#### FAQ

#### What are GEO and ENA?

The Gene Expression Omnibus (GEO) and the European Nucleotide Archive (ENA) are web-accessible public repositories for genomic data and experiments. Access and learn more about their resources at their respective websites:

GEO - https://www.ncbi.nlm.nih.gov/geo/

ENA - https://www.ebi.ac.uk/ena

#### How do I know if a GEO project is also in ENA?

• You can search ENA using the GEO ID (e.g., GSE71578) to check if there is a matching ENA project.

EMBL-EBI			Services	Research	Training	About us
European Nucleotide Archive		GSE97388 Examples: <u>BN000065</u> , <u>histone</u>	Advanc Sequer			ed ice
Home Search & Browse Submit & Update Software	e About ENA Support					
Search results for GSE97388			Show mo	re data fro	m EMBL-	BI
Study Study (1)	Study (1 results found)					
Submission Submission (Read/Analysis) (1)	SRP103018 In vitro differentiation of human embryonic stem cells i View all 1 results	nto ovarian follicle-like cells				
	Submission (Read/Analysis) (1 results found)					
	SRA551479 Submitted by Gene Expression Omnibus on 05-JUL-203 View all 1 results	17				
Powered by <u>EBI Search</u>						

• Open the Study result to view the BioProject ID (e.g., PRJNA381606) and a table with information about the samples and files included in the project

EMBL-EBI																
European Nucleotide Archive								Exa	mples: <u>BN000065</u> ,	histone			S	dvanced equence		
Home Searc	h & Browse S	ubmit & Updat	e Software	About ENA	Sup	port										
Study: P	RJNA3816	606													<u>Contac</u>	<u>t Helpdesk</u> 🖂
In vitro differentiation of human embryonic stem cells into ovarian follicle-like cells																
View: Project	XML Study XI	4L											Dow	nload: P	roject XML	Study XML
Name Homo sapiens				Submit Basic M	tting C edical	entre Sciences, T	singhua Univ	ersity			<b>Organism</b> Homo sapien	5				
Secondary ac SRP103018	cession(s)															
Understandin derived from differentiate i meiosis. By e granulosa cel cells, and ova	Description Understanding the unique mechanisms of human oogenesis necessitates the development of an in vitro system of stem cell differentiation into oocytes. Specialized cell types and organoids have been derived from human pluripotent stem cells in vitro, but generating a human ovarian follicle remains a challenge. Here we report that human embryonic stem cells (hESCs) can be induced to differentiate into ovarian follicle-like cells in vitro. First, we find that two RNA-binding proteins specifically expressed in germ cells, DAZL and BOULE, regulate the exit from pluripotent, and entry into meiosis. By expressing DAZL and BOULE with recombinant human GDF9 and BMP15, these meiotic germ cells are further induced to form ovarian follicle-like cells (FLCs), including oocytes and granulosa cells. This robust in vitro differentiation system will allow the study of the unique molecular mechanisms underlying human pluripotent stem cell differentiation into late PGCs, meiotic germ cells, and ovarian follicles. Overall design: Including 6 samples, 4 controls: ES_1, ES_2, SDE_1, SDE_2; 2 samples: FLC_1 (HSF6), FLC_2 (H9)															
Lineage Eukaryota, Mel	azoa, Chordata,	Craniata, Vertel	brata, Euteleos	itomi, Mamma	alia, Eu	theria, Eua	rchontoglire	, Primate	s, Haplori	hini, Catar	rhini, Homini	dae, <u>Homo</u>				
Navigation	Read F	iles	Portal		Attri	butes	Pu	lications		Parent P	rojects					
O Rulla Davrala																
U Buik Downic	ad Files 🔼 (If t	ne downloader a	ipp doesn't op	en, please try	using	Pirefox to la	aunen Ie.)									
Download:	-	6 o	f 6 results in	TEXT												
Select columns																
Showing resu	lts 1 - 6 of 6 re	sults														
Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrumen model	t Library layout	FASTQ files (FTP)	FASTQ files (Galaxy)	Submitted files (FTP)	Submitted files (Galaxy)	NCBI SRA file (FTP)	NCBI SRA file (Galaxy)	CRAM Index files (FTP)	CRAM Index files (Galaxy)
PRJNA381606	SAMN06681557	SRS2098866	SRX2705208	SRR5413258	9606	Homo sapiens	Illumina HiSeq 2500	PAIRED	File 1 File 2	File 1 File 2			File 1	File 1		
PRJNA381606	SAMN06681556	SRS2098867	SRX2705209	SRR5413259	9606	Homo sapiens	Illumina HiSeq 2500	PAIRED	File 1 File 2	File 1 File 2			File 1	File 1		

## Additional Assistance

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

