

Setup

We're aware some of your tasks have failed and have determined that the issue was the number of duplicate files in the main project which were copied over

This can be resolved by re-running the task with fewer files and we have made instructions on how to do this in your projects

Delete Old Files

1. Navigate to your project and click on the **Files** tab beneath the blue navigation bar
2. Select all of the folders and click on the **three dots** next to the download button

The screenshot illustrates the steps to delete old files in the CGC Platform. It shows the navigation bar with the 'Files' tab highlighted (1). Below, the 'Files' section shows a table of folders with checkboxes for selection (2). A context menu is open over the 'Download' button, showing options like 'Delete' (3).

Navigation Bar: Dashboard, **Files**, Apps, Tasks

Files Section: Edit Metadata, Copy, Move, Tags, Download, ...

<input checked="" type="checkbox"/>	Name	Experimental strategy
<input checked="" type="checkbox"/>	workflow_outputs	-
<input checked="" type="checkbox"/>	sc-RNAseq	-
<input checked="" type="checkbox"/>	reference_files	-
<input checked="" type="checkbox"/>	bulk-mnaseq-samples	-

Context Menu: Archive, Restore, **Delete**, Export metadata manifest (CSV), Export metadata

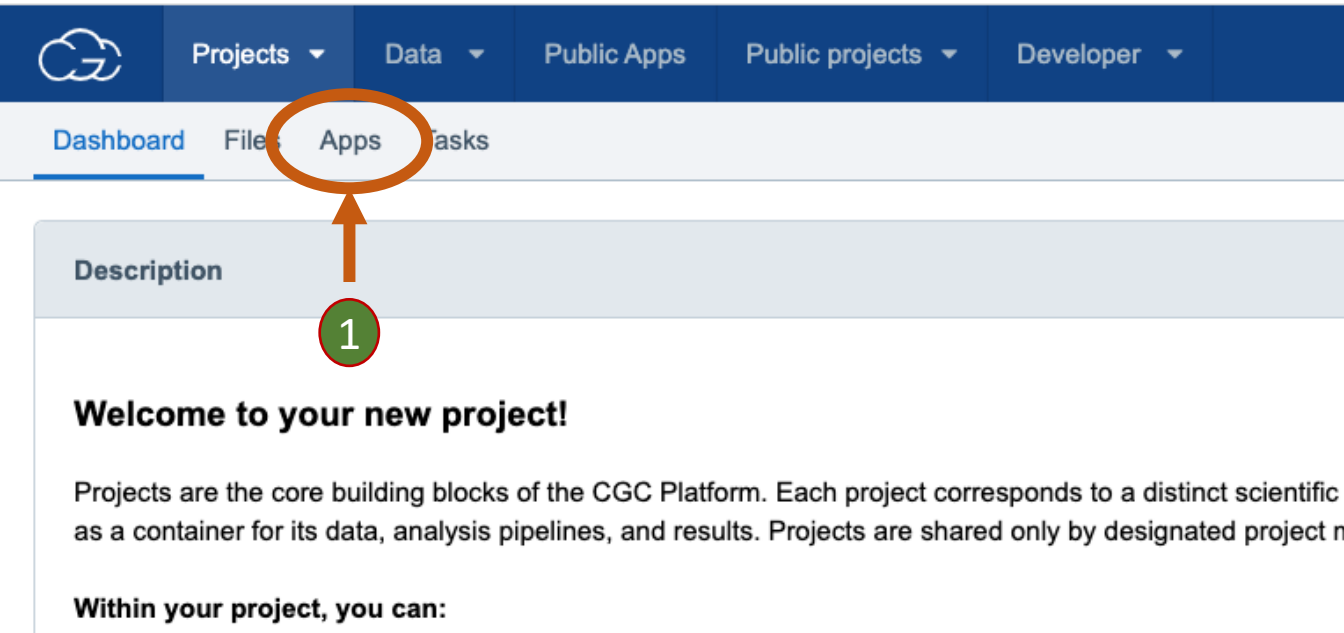
Copy Files Over To Your Project From Purdue Lecture Series

1. Navigate to the project **Purdue Lecture Series**
2. Sort by **Name**
3. Select the **sc-RNAseq**, **reference_files**, and **bulk-rnaseq-samples** folders
4. Click **copy** then select your project

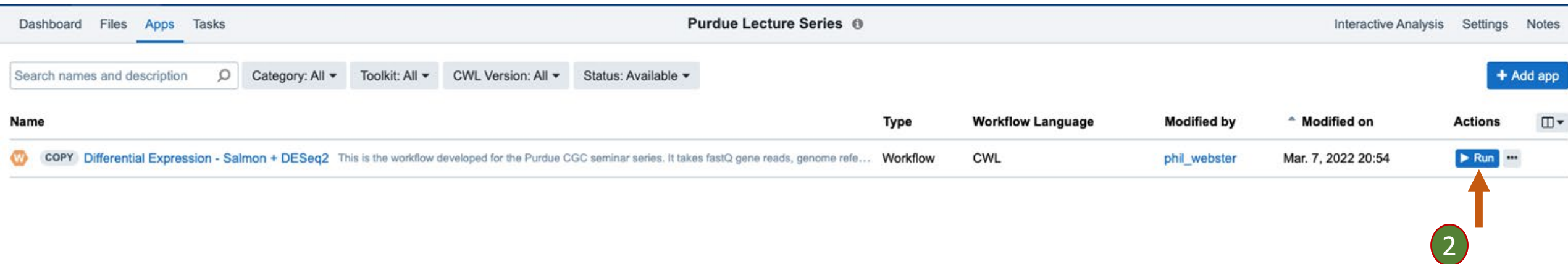
The screenshot shows the Genomics Workbench interface for the 'Purdue Lecture Series' project. The interface includes a navigation bar with 'Projects', 'Data', 'Public Apps', 'Public projects', and 'Developer'. The 'Files' tab is selected, and the 'Copy' button is highlighted. The 'Name' dropdown menu is open, showing a list of projects. The table below shows the following data:

Experimental strategy	Paired-end	Extension	Size	Sample ID	induced_hbv_replication
-	-	-	-	-	-
-	-	-	-	-	-
-	-	CSV	0.2 KiB	-	-
-	-	TSV	21.8 MiB	-	-
-	-	TSV	3.8 MiB	-	-
-	-	-	-	-	-

Navigating to the Workflow




1. Navigate to **your project**. (Not **Purdue Lecture Series** project).
2. Click on the **Apps** tab under the blue navigation bar. The workflow should already be there from class.
3. Then click the **Run** button next to the workflow we copied over to our projects



Setup The Workflow: FastQ Files


1. We will fill out our three inputs, begin by clicking **Select File(s)** next to **FASTQ read files**




DRAFT Differential Expression - Salmon + DESeq2 run - 03-09-22 03:35:18 

Last update by phil_webster on Mar. 8, 2022 22:35
App: Differential Expression - Salmon + DESeq2 - Revision: 1


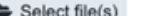
[Task Inputs](#) Execution Settings

Inputs



Batching  Off

FASTQ read files *    1



No files selected
This field is required and cannot be empty.

GTF annotation *  

No files selected
This field is required and cannot be empty.



Genome FASTA  

No files selected


Transcript FASTA or Salmon Index *  

No files selected
This field is required and cannot be empty.

App Settings


 

DESeq2 (#deseq2_1_26_0)


Covariate of interest * 

No value

This field is required and cannot be empty.













Factor level - reference 

No value

Factor level - test 

No value

Output Settings

DESeq2 analysis results. 	No value
Expression matrix genes 	No value
Expression matrix transcripts 	No value
Gene-level quantification 	No value
HTML report 	No value
HTML reports 	No value
Normalized counts 	No value
RData file 	No value
Report zip 	No value
Salmon Quant archive 	No value
Salmon quant log 	No value
Transcript-level quantification 	No value
pheno_out	No value

Selecting FastQ Files

1. When selecting the FastQ files first click on **Tags** at the end of the search bar next to the **+** symbol and select **BULK**
2. The 154 files should filter down to 30
3. Select all files and click **Save selection**

The screenshot shows a web interface for selecting files. At the top, a header reads "Select files for 'FASTQ read files'". Below this are navigation tabs: "Current Project", "Projects", and "Public Files". A search bar contains the text "FASTQ read files". To the right of the search bar are filter buttons: "Extension: 4 selected", "Sample ID: All", "Task ID: All", and "Tags: BULK". A red circle with the number "1" points to the "Tags: BULK" button. A dropdown menu is open, titled "Filter by: Tags", with a search input field and a "Clear selected" option. Under "Showing 6", the "BULK (30)" option is selected and highlighted. Other options include "GRCh38 (5)", "SB-GENERATED (1)", "SUGGESTED (3)", "Single-Cell (126)", "mseq (156)", and "No value (98)". A red circle with the number "2" points to the search bar. Below the filters is a table of files with columns: "Name", "Experimental strategy", "Path", "Extension", "Size", "Sample ID", and "induced_hbv_replication". The first five rows are visible, each with a checked checkbox and a "30" indicator. A red circle with the number "3" points to a "Save selection" button in the top right corner.

Name	Experimental strategy	Path	Extension	Size	Sample ID	induced_hbv_replication
<input checked="" type="checkbox"/> SRR9059002_2.fastq BULK RNASEQ	-	Files / bulk-maseq-	FASTQ	3.4 GiB	SRR9059002	HBV_pos
<input checked="" type="checkbox"/> SRR9059002_1.fastq BULK RNASEQ	-	Files / bulk-maseq-	FASTQ	3.4 GiB	SRR9059002	HBV_pos
<input checked="" type="checkbox"/> SRR9059001_2.fastq BULK RNASEQ	-	Files / bulk-maseq-	FASTQ	4.6 GiB	SRR9059001	HBV_pos
<input checked="" type="checkbox"/> SRR9059001_1.fastq BULK RNASEQ	-	Files / bulk-maseq-	FASTQ	4.6 GiB	SRR9059001	HBV_pos
<input checked="" type="checkbox"/> SRR9059000_2.fastq BULK RNASEQ	-	Files / bulk-maseq-	FASTQ	3.4 GiB	SRR9059000	HBV_neg

Selecting GTF Annotation Files

DRAFT Differential Expression - Salmon + DESeq2 run - 03-09-22 03:35:18

Get support Discard Run

Last update by phil_webster on Mar. 8, 2022 22:35
App: Differential Expression - Salmon + DESeq2 - Revision: 1


Task Inputs ¹ Execution Settings

Inputs

Batching [?] Off

FASTQ read files * [?]

- SRR9059002_2.fastq
- SRR9059002_1.fastq
- SRR9059001_2.fastq
- SRR9059001_1.fastq
- SRR9059000_2.fastq
- ...and 25 more items

GTF annotation * [?] 

This input is required.

No files selected

This field is required and cannot be empty.

Genome FASTA [?]

No files selected

Transcript FASTA or Salmon Index * [?]

No files selected

This field is required and cannot be empty.

App Settings

DESeq2 (#deseq2_1_26_0)

Covariate of interest * [?]

This field is required and cannot be empty.

Factor level - reference [?]

Factor level - test [?]

Output Settings

DESeq2 analysis results [?]	No value
Expression matrix genes [?]	No value
Expression matrix transcripts [?]	No value
Gene-level quantification [?]	No value
HTML report [?]	No value
HTML reports [?]	No value
Normalized counts [?]	No value
RData file [?]	No value
Report zip [?]	No value
Salmon Quant archive [?]	No value
Salmon quant log [?]	No value
Transcript-level quantification [?]	No value
pheno_out	No value

Selecting GTF Annotation Files

1. Select **GRCh38ERCC.endemble95.gtf**
2. click **Save selection**

Select files for "GTF annotation" 2

Current Project Projects Public Files

Files

Search Extension: 2 selected Sample ID: All Task ID: All Tags: All +

Copy


Name	Experimental strategy	Path	Paired-end	Extension	Size	Sample ID	induced_hbv_replication
<input type="checkbox"/> gencode.v27.annotation.gtf <small>SCRNA SUGGESTED RNASEQ SEQUENCING</small>	-	Files / reference_files	-	GTF	1.1 GiB	-	-
<input type="checkbox"/> gencode.v24.annotation.gtf <small>SCRNA RNASEQ SEQUENCING</small>	-	Files / reference_files	-	GTF	1.2 GiB	-	-
<input type="checkbox"/> Homo_sapiens.GRCh38.84.gtf <small>SCRNA RNASEQ SEQUENCING</small>	-	Files / reference_files	-	GTF	1.3 GiB	-	-
<input checked="" type="checkbox"/> GRCh38ERCC.ensembl95.gtf <small>SCRNA RNASEQ GRCH38 SEQUENCING</small>	-	Files / reference_files	-	GTF	1.1 GiB	-	-

1

Save selection

Selecting Transcript Salmon Index File

Dashboard Files Apps **Tasks** Purdue Lecture Series ⓘ Interactive Analysis Settings Notes


DRAFT Differential Expression - Salmon + DESeq2 run - 03-09-22 03:35:18 

Last update by phil_webster on Mar. 8, 2022 22:35
App: Differential Expression - Salmon + DESeq2 - Revision: 1


Task Inputs Execution Settings

Inputs


Batching ⓘ Off

FASTQ read files * ⓘ  Change selection



- SRR9059002_2.fastq
- SRR9059002_1.fastq
- SRR9059001_2.fastq
- SRR9059001_1.fastq
- SRR9059000_2.fastq
- ...and 25 more items

GTF annotation * ⓘ  Change selection

- GRCh38ERCC.ensembl95.gtf

Genome FASTA ⓘ  Select file(s)



No files selected

Transcript FASTA or Salmon Index * ⓘ  Select file(s) 

No files selected

This field is required and cannot be empty.

App Settings

 Edit parameters  Show editable ▾

DESeq2 (#deseq2_1_26_0)

Covariate of interest * ⓘ

No value

This field is required and cannot be empty.

Factor level - reference ⓘ

No value

Factor level - test ⓘ

No value

Output Settings

DESeq2 analysis results. ⓘ	No value
Expression matrix genes ⓘ	No value
Expression matrix transcripts ⓘ	No value
Gene-level quantification ⓘ	No value
HTML report ⓘ	No value
HTML reports ⓘ	No value
Normalized counts ⓘ	No value
RData file ⓘ	No value
Report zip ⓘ	No value
Salmon Quant archive ⓘ	No value
Salmon quant log ⓘ	No value
Transcript-level quantification ⓘ	No value
pheno_out	No value

Selecting Transcript FASTA/Salmon Index File

1. Sort by **Name**
2. Select **GRCh38ERCC.ensembl95.transcriptome.gentrome.salmon-1.2.0-index-archive.tar**
3. click **Save selection**

Select files for "Transcript FASTA or Salmon Index" 2

Current Project Projects Public Files

Files

Search Extension: 5 selected Sample ID: All Task ID: All Tags: All +

Copy

Name	Experimental strategy	Path	Paired-end	Extension	Size	Sample ID	induced_hbv_replication
<input type="checkbox"/> GRCh38ERCC.ensembl95.fasta <small>GRCH38</small>	-	Files / reference_files	-	FASTA	2.9 GiB	-	-
<input type="checkbox"/> GRCh38ERCC.ensembl95.transcriptome.fasta <small>GRCH38</small>	-	Files / reference_files	-	FASTA	305.7 MiB	-	-
<input checked="" type="checkbox"/> GRCh38ERCC.ensembl95.transcriptome.gentrome.salmon-1.2.0-index-archive.tar <small>SB-GENERATED GRCH38</small>	-	Files / reference_files	-	TAR	16.6 GiB	-	-
<input type="checkbox"/> SRR9058988.salmon_quant_archive.tar	-	Files / workflow_outputs / run_1	-	TAR	11.8 MiB	SRR9058988	HBV_neg
<input type="checkbox"/> SRR9058989.salmon_quant_archive.tar	-	Files / workflow_outputs / run_1	-	TAR	11.8 MiB	SRR9058989	HBV_pos

Save selection

1

Fill In App Settings

Verify that everything is set up as shown here, and the draft task is in **your own project**.

There should be a total of 30 input fastq files and the DESeq2 settings should be as shown below

Inputs

Batching [?] Off

FASTQ read files * [?]

- SRR9059002_2.fastq
- SRR9059002_1.fastq
- SRR9059001_2.fastq
- SRR9059001_1.fastq
- SRR9059000_2.fastq
- ...and 25 more items

GTF annotation * [?]

GRCh38ERCC.ensembl95.gtf

Genome FASTA [?]

No files selected

Transcript FASTA or Salmon Index * [?]

GRCh38ERCC.ensembl95.transcriptome.gentrome.salmon-1.2.0-index-arch...

App Settings

DESeq2 (#deseq2_1_26_0)

Covariate of interest * [?]

Genotype

Factor level - reference [?]

WT

Factor level - test [?]

KD

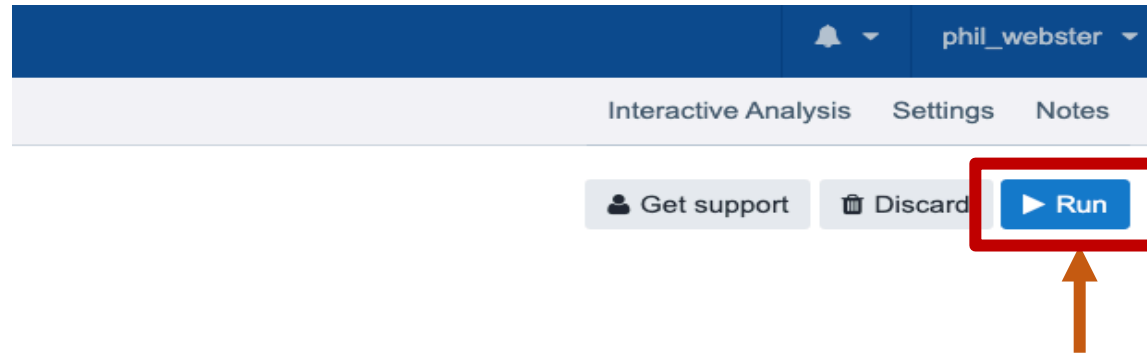
Output Settings

DESeq2 analysis results. [?]	No value
Expression matrix genes [?]	No value
Expression matrix transcripts [?]	No value
Gene-level quantification [?]	No value
HTML report [?]	No value
HTML reports [?]	No value
Normalized counts [?]	No value
RData file [?]	No value
Report zip [?]	No value
Salmon Quant archive [?]	No value
Salmon quant log [?]	No value
Transcript-level quantification [?]	No value
pheno_out	No value

Confirm only 30 files total

Fill in as seen here

Run!



The screenshot shows a software interface with a dark blue header bar containing a notification bell icon and the user name 'phil_webster'. Below the header is a light grey navigation bar with 'Interactive Analysis', 'Settings', and 'Notes'. Underneath are three buttons: 'Get support' (with a person icon), 'Discard' (with a trash icon), and 'Run' (with a play icon). The 'Run' button is highlighted with a red rectangular box, and an orange arrow points upwards towards it from below.

Output Settings

<i>DESeq2 analysis results</i> ?	<i>No value</i>
<i>Expression matrix genes</i> ?	<i>No value</i>
<i>Expression matrix transcripts</i> ?	<i>No value</i>
<i>Gene-level quantification</i> ?	<i>No value</i>
<i>HTML report</i> ?	<i>No value</i>
<i>HTML reports</i> ?	<i>No value</i>
<i>Normalized counts</i> ?	<i>No value</i>
<i>RData file</i> ?	<i>No value</i>
<i>Report zip</i> ?	<i>No value</i>
<i>Salmon Quant archive</i> ?	<i>No value</i>
<i>Salmon quant log</i> ?	<i>No value</i>
<i>Transcript-level quantification</i> ?	<i>No value</i>
<i>pheno_out</i>	<i>No value</i>

Please Contact Us With Further Questions

