

Step-by-Step screen shots for training task of gene expression analysis

Week 4 Task 1

Perform Groups comparison for Rembrandt brain cancer study using Astrocytoma samples grade 2 and 3

- Step-by-step instructions (for more details see screen shots below):
 - Log in to G-DOC at gdoc.georgetown.edu
 - Find and Select Rembrandt study in G-DOC
 - Follow the initial steps of gene expression demo and select and save the same 4 groups of samples that were saved in gene expression demo
 - Select “group comparison” option from main menu under “study options” and set up comparison of Astrocytoma grade 2 vs Astrocytoma grade 3 (make sure that those two groups were saved with the following names, for grade 2 : Rem_Astr_gr2; and for grade 3 : Rem_Astr_gr3
 - To compare those two groups: select two groups for comparison from drop down menus: for base group select Rem_Astr_gr2 and for comparison group select Rem_Astr_gr3
 - Select multiple comparison option 2 i.e. False Discovery Rate option from drop down menu for “multiple testing adjustment”
 - Select gene expression as data type
 - Run the comparison
 - When analysis shown as complete on notification page open the result by clicking on hyper link
 - In group comparison results window inspect the table of differentially expressed genes. Take a record of total number of genes in the table (shown in the right lower corner under the table.
 - Sort the table by the column “Fold Change” in descending order of fold change values
 - Find the gene name on top row of the sorted table – this is the top overexpressed gene in Astrocytoma group of grade 3 tumors as compared to grade 2.
 - Compare this gene name to the gene name of top overexpressed gene shown in the demo.

To check if this workflow was executed correctly answer two questions in the quiz:

Question 1: what was the total number of genes in the table? Enter number here: _____

Question 2: What was the name of the top overexpressed gene in the sorted table? Enter gene name here: ____

Screen shot 1: select study option “group comparison”

Georgetown Database of ...

https://gdoc.georgetown.edu/gdoc/workflows/studySpecificTools

G-DOC^{plus}

Home Studies Lists Analyses Groups Notifications Study Options Help

yg63

Study Selected!

You have picked the study: REMBRANDT*

[Change my study](#) [Help me pick another study](#)

* NCI Rembrandt Study: Molecular Analysis of Brain Neoplasia

Based upon the study you picked, here is a list of tools

Search

- [Genome Browser](#)
- [Compound/Drug Targets](#)
- [Findings](#)
- [Explore Clinical Data and Create Groups](#)
- [Gene Expression Data](#)

Analyze

- [Group Comparison](#)
- [Chromosomal Instability Index](#)
- [KM Clinical Plot](#)
- [KM Gene Expression Plots](#)
- [Classification](#)
- [HeatMap Viewer](#)

Study Selected

REMBRANDT

SEARCH


- [Genome Browser](#)
- [Compound/Drug Targets](#)
- [Findings](#)
- [Explore Clinical Data and Create Groups](#)
- [Gene Expression Data](#)

ANALYZE


- **Group Comparison**
- [Chromosomal Instability Index](#)
- [KM Clinical Plot](#)
- [KM Gene Expression Plots](#)
- [Classification](#)
- [HeatMap Viewer](#)

https://gdoc.georgetown.edu/gdoc/groupComparison/index

Screen shot 2: select baseline group




HomeStudiesListsAnalysesGroupsNotificationsStudy Options▼Help



yg63▼

Perform Group Comparison Analysis

Current Study: REMBRANDT [change study?](#)

Select a baseline group and a comparison group(s) 

Select baseline group:

Rem_Astr_g2▼

Select comparison group:

▼

p-value:

.05

Fold Change:

1.5

Statistical Method:

T-Test: Two Sample T

Multiple Comparison Adjustment:

False Discovery Rate (FDR): F

Select baseline group...

rem_gbm

Rem_astroc

list1463583422692

Rem_astr_r3

Rem_astr_gr2

Rem_Olig_gr3

Rem_Glio_gr2

Rem_Astr_g3

Rem_Astr_g2

Rem_Oli_g3

Rem_Oli_g2

test2

test1

astro

list1445614636407

gbm1

astro1

test111

bm_cn

Screen shot 3: select comparison group

Group Comparison Analysis

https://gdoc.georgetown.edu

G-DOC^{Plus} Home Studies

Perform Group Comparison

Current Study: REMBRANDT

Select a baseline group and a comparison group

Select baseline group: Rem_gbm

Select comparison group: Rem_Astr_g3

p-value: .05

Fold Change: 1.5

Statistical Method: T-Test: Two Sample Test

Multiple Comparison Adjustment: False Discovery Rate(FDR): E

Select comparison group...

- rem_gbm
- Rem_astroc
- list1463583422692
- Rem_ast_r3
- Rem_ast_r2
- Rem_Olig_gr3
- Rem_Glio_gr2
- Rem_Astr_g3
- Rem_Astr_g2
- Rem_Oli_g3
- Rem_Oli_g2
- test2
- test1
- astro
- list1445614636407
- gbm1
- astro1
- test111
- bm_cn

Screen shot 4: Select Multiple Comparison Adjustment

Group Comparison Analysis x

https://gdoc.georgetown.edu/gdoc/groupComparison/index

G-DOC^{plus}

Home Studies Lists Analyses Groups Notifications Study Options Help

yg63

Perform Group Comparison Analysis

Current Study: REMBRANDT [change study?](#)

Select a baseline group and a comparison group(s) ⓘ

Select baseline group: Rem_Astr_g2

Select comparison group: Rem_Astr_g3

p-value:
.05

Fold Change:
1.5

Statistical Method:
T-Test: Two Sample Test

- None
- Family-Wise Error Rate(FWER): Bonferroni
- False Discovery Rate(FDR): Benjamini-Hochberg**
- False Discovery Rate(FDR): E

Screen shot 5: Select Data Type “Gene Expression”

Group Comparison Analysis

https://gdoc.georgetown.edu/gdoc/groupComparison/index

G-DOC^{plus} Home Studies Lists Analyses Groups Notifications Study Options - Help

Perform Group Comparison Analysis

Current Study: REMBRANDT [change study?](#)

Select a baseline group and a comparison group(s) [i](#)

Select baseline group: Rem_Astr_g2

Select comparison group: Rem_Astr_g3

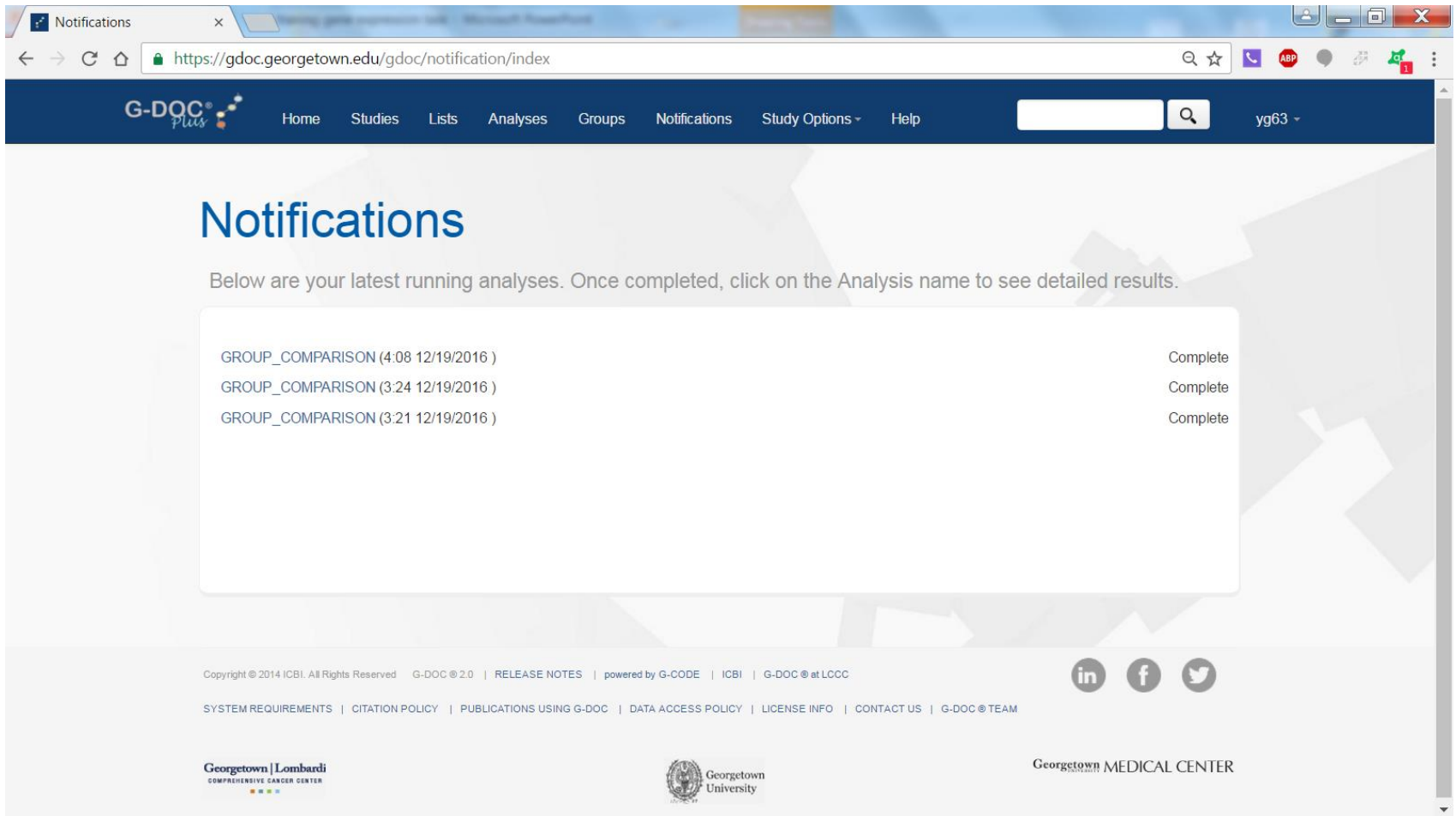
p-value:
.05

Fold Change:
1.5

Statistical Method:
T-Test: Two Sample Test

Multiple Comparison Adjustment:
False Discovery Rate(FDR): B
Select Data Type
COPY_NUMBER
GENE EXPRESSION
GENE EXPRESSION

Screen shot 6: Check Notifications for completion of analysis



The screenshot shows a web browser window with the address bar displaying <https://gdoc.georgetown.edu/gdoc/notification/index>. The page title is "Notifications". The navigation bar includes links for Home, Studies, Lists, Analyses, Groups, Notifications, Study Options, and Help. A search bar and a user profile "yg63" are also present.

Notifications

Below are your latest running analyses. Once completed, click on the Analysis name to see detailed results.

GROUP_COMPARISON (4:08 12/19/2016)	Complete
GROUP_COMPARISON (3:24 12/19/2016)	Complete
GROUP_COMPARISON (3:21 12/19/2016)	Complete

The footer contains copyright information: "Copyright © 2014 ICBi. All Rights Reserved. G-DOC © 2.0 | RELEASE NOTES | powered by G-CODE | ICBi | G-DOC © at LCCC". It also includes social media icons for LinkedIn, Facebook, and Twitter, and a list of links: SYSTEM REQUIREMENTS | CITATION POLICY | PUBLICATIONS USING G-DOC | DATA ACCESS POLICY | LICENSE INFO | CONTACT US | G-DOC © TEAM. Logos for Georgetown Lombardi Comprehensive Cancer Center, Georgetown University, and Georgetown Medical Center are displayed at the bottom.

Screenshot 7: Open analysis results of the group comparison

G-DOC^{plus}

HomeStudiesListsAnalysesGroupsNotificationsStudy Options -Help

yg63

Analysis Results

Current Study: REMBRANDT

Analysis Results

Statistical Method	TTest
Adjustment	FDR
Fold Change	1.5
Pvalue	.05
Study	REMBRANDT
Data File	REMBRANDT.Rda
Baseline Group	Rem_Astr_g2
Groups	Rem_Astr_g3

List Name:

Save Selected

View HeatMap for selected reporters

Analysis Results

Reporter ID	Gene Symbol	p-value	Fold Change	Mean Baseline (log)	Mean Group (log2)	Std Baseline	Std Group	Target Data
<input type="checkbox"/> 213462_at	NPAS2	1.573×10^{-4}	2.429	9.567	10.848	1.063	1.151	
<input type="checkbox"/> 1553991_s_at	FLJ20674	4.073×10^{-4}	2.006	8.462	9.466	0.800	1.030	
<input type="checkbox"/> 219312_s_at	ZBTB10	5.551×10^{-4}	2.190	6.365	7.496	0.948	1.188	
<input type="checkbox"/> 1553992_s_at	NBR2	5.551×10^{-4}	1.950	9.143	10.107	0.764	1.036	
<input type="checkbox"/> 1563913_at		8.893×10^{-4}	2.258	7.942	9.117	1.085	1.196	
<input type="checkbox"/> 209343_at	EFHD1	9.259×10^{-4}	2.284	8.304	9.496	0.994	1.330	
<input type="checkbox"/> 1557593_at	SPAG17	1.014×10^{-3}	2.124	9.838	10.925	0.928	1.217	
<input type="checkbox"/> 1554609_at		1.014×10^{-3}	1.683	9.057	9.808	0.662	0.829	
<input type="checkbox"/> 213819_s_at	FLOT1	1.071×10^{-3}	2.368	7.088	8.332	1.203	1.295	
<input type="checkbox"/> 204700_x_at	C1orf107	1.071×10^{-3}	1.791	7.951	8.792	0.681	0.992	
<input type="checkbox"/> 213549_at		1.108×10^{-3}	1.990	6.282	7.275	1.107	0.866	
<input type="checkbox"/> 208006_at	FOXJ1	1.503×10^{-3}	2.048	8.226	9.260	0.887	1.236	
<input type="checkbox"/> 226441_at		1.699×10^{-3}	2.053	5.790	6.827	1.188	0.940	
<input type="checkbox"/> 1555238_at	PTH2	1.748×10^{-3}	3.495	5.611	7.417	1.788	2.018	

Export results

Page 1 of 8

View 1 - 50 of

Screenshot 8: Check the table with the analysis results of the group comparison – find a total number of differentially expressed genes (lower right corner under the table)

G-DOC^{plus} Home Studies Lists Analyses Groups Notifications Study Options - Help yg63 -

Analysis Results
Current Study: REMBRANDT

Analysis Results

Statistical Method	TTest
Adjustment	FDR
Fold Change	1.5
Pvalue	.05
Study	REMBRANDT
Data File	REMBRANDT.Rda
Baseline Group	Rem_Astr_g2
Groups	Rem_Astr_g3

List Name:

Analysis Results

<input type="checkbox"/>	Reporter ID	Gene Symbol	p-value	Fold Change	Mean Baseline (log)	Mean Group (log2)	Std Baseline	Std Group	Target Data
<input type="checkbox"/>	213462_at	NPAS2	1.573×10^{-4}	2.429	9.567	10.848	1.063	1.151	
<input type="checkbox"/>	1553991_s_at	FLJ20674	4.073×10^{-4}	2.006	8.462	9.466	0.800	1.030	
<input type="checkbox"/>	219312_s_at	ZBTB10	5.551×10^{-4}	2.190	6.365	7.496	0.948	1.188	
<input type="checkbox"/>	1553992_s_at	NBR2	5.551×10^{-4}	1.950	9.143	10.107	0.764	1.036	
<input type="checkbox"/>	1563913_at		8.893×10^{-4}	2.258	7.942	9.117	1.085	1.196	
<input type="checkbox"/>	209343_at	EFHD1	9.259×10^{-4}	2.284	8.304	9.496	0.994	1.330	
<input type="checkbox"/>	1557593_at	SPAG17	1.014×10^{-3}	2.124	9.838	10.925	0.928	1.217	
<input type="checkbox"/>	1554609_at		1.014×10^{-3}	1.683	9.057	9.808	0.662	0.829	
<input type="checkbox"/>	213819_s_at	FLOT1	1.071×10^{-3}	2.368	7.088	8.332	1.203	1.295	
<input type="checkbox"/>	204700_x_at	C1orf107	1.071×10^{-3}	1.791	7.951	8.792	0.681	0.992	
<input type="checkbox"/>	213549_at		1.108×10^{-3}	1.990	6.282	7.275	1.107	0.866	
<input type="checkbox"/>	208006_at	FOX11	1.503×10^{-3}	2.048	8.226	9.260	0.887	1.236	
<input type="checkbox"/>	226441_at		1.699×10^{-3}	2.053	5.790	6.827	1.188	0.940	
<input type="checkbox"/>	1555238_at	PTH2	1.748×10^{-3}	3.495	5.611	7.417	1.788	2.018	

Export results

Screenshot 7: Sort a table with the analysis results by column “Fold Change” in descending order (by clicking once on column label);
Take a record of the name of the top overexpressed gene (top row, second column)

georgetown.edu/gdoc/groupComparison/view/1249578

G-DOC^{plus} Home Studies Lists Analyses Groups Notifications Study Options - Help yg63 -

Analysis Results

Current Study: REMBRANDT

Analysis Results

Statistical Method	TTest
Adjustment	FDR
Fold Change	1.5
Pvalue	.05
Study	REMBRANDT
Data File	REMBRANDT.Rda
Baseline Group	Rem_Astr_g2
Groups	Rem_Astr_g3

List Name:

Analysis Results

<input type="checkbox"/>	Reporter ID	Gene Symbol	p-value	Fold Change	Mean Baseline (log)	Mean Group (log2)	Std Baseline	Std Group	Target Data
<input type="checkbox"/>	209937_at	PTX2	3.358×10^{-3}	3.917	5.810	7.779	1.958	2.453	
<input type="checkbox"/>	1555238_at	PTH2	1.748×10^{-3}	3.495	5.611	7.417	1.788	2.018	
<input type="checkbox"/>	1555409_at	BAGE2	1.905×10^{-2}	3.397	8.462	10.227	2.311	2.595	
<input type="checkbox"/>	1570176_at		1.588×10^{-2}	3.295	5.637	7.357	1.905	2.707	
<input type="checkbox"/>	208765_s_at	HNRNP	2.448×10^{-3}	3.173	5.626	7.292	1.746	1.870	
<input type="checkbox"/>	203003_at	MEF2D	2.676×10^{-3}	2.693	5.726	7.156	1.362	1.760	
<input type="checkbox"/>	225783_at	UBE2F	1.235×10^{-2}	2.677	3.553	4.973	1.741	1.978	
<input type="checkbox"/>	217253_at	SH3BP2	2.848×10^{-2}	2.626	6.486	7.879	1.992	2.149	
<input type="checkbox"/>	1558093_s_at	MATR3	1.827×10^{-2}	2.616	5.168	6.556	1.587	2.221	
<input type="checkbox"/>	1555628_at	HAVCR2	2.836×10^{-3}	2.598	6.116	7.493	1.331	1.695	
<input type="checkbox"/>	208736_at	ARPC3	3.358×10^{-3}	2.587	6.055	7.427	1.373	1.701	
<input type="checkbox"/>	1558475_at		3.358×10^{-3}	2.579	6.107	7.474	1.436	1.627	
<input type="checkbox"/>	228629_s_at	LOC100128025	2.781×10^{-2}	2.568	8.185	9.546	1.952	2.077	
<input type="checkbox"/>	208672_s_at	SFR53	7.247×10^{-3}	2.568	5.561	6.922	1.598	1.718	

Explore the function of the top over-expressed gene by left-clicking on gene name in top row and selecting “View at Gene Cards” from the drop down menu

Analysis Results

Current Study: REMBRANDT

Analysis Results

Statistical Method	TTest
Adjustment	FDR
Fold Change	1.5
Pvalue	.05
Study	REMBRANDT
Data File	REMBRANDT.Rda
Baseline Group	Rem_Astr_g2
Groups	Rem_Astr_g3

List Name:

Save Selected

View HeatMap for selected reporters

Analysis Results

Reporter ID	Gene	Fold Change	Mean Baseline (log)	Mean Group (log2)	Std Baseline	Std Group	Target Data
209937_at	TM6	917	5.810	7.779	1.958	2.453	
1555238_at	PTH	495	5.611	7.417	1.788	2.018	
1555409_a_at	BAQ	397	8.462	10.227	2.311	2.595	
1570176_at		295	5.637	7.357	1.905	2.707	
208765_s_at	HN	173	5.626	7.292	1.746	1.870	
203003_at	ME	693	5.726	7.156	1.362	1.760	
225783_at	UB	677	3.553	4.973	1.741	1.978	
217253_at	SH	626	6.486	7.879	1.992	2.149	
1558093_s_at	MA	616	5.168	6.556	1.587	2.221	
1555628_a_at	HA	598	6.116	7.493	1.331	1.695	
208736_at	AR	587	6.055	7.427	1.373	1.701	
1558475_at		579	6.107	7.474	1.436	1.627	
228629_s_at	LOC	568	8.185	9.546	1.952	2.077	
		2,568	5.561	6.922	1.598	1.718	

Perform Gene Expression KM

Perform Gene Expression Search

Search in Entrez

View in UCSC Genome Browser

Search in iHOP

Search in PIR

Search in Ensembl Gene View

Search in Reactome

View at KEGG

View at QuickGO

View at GeneCards

View at String DB