Use Case 16: Differential Expression Analysis of Two Breast Cancer Cell Lines Using Cuffdiff

> Epigenome Informatics Workshop Bioinformatics Research Laboratory



mRNA-Seq

- Read mapping
 tophat
- Gene expression estimation
 - cufflinks
 - Confidence intervals
- Gene expression changes
 - Sample groups
 - cuffdiff

Use case for RNA-Seq tools

- 2 breast cancer cell lines
 - Joe Gray 51 breast cancer cell lines panel
 - Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo WL, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Gray JW. "A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes." *Cancer Cell.* 2006 Dec;10(6):515-27.

Evaluate Gene Expression Differences

Sample	Luminal/Basal	ER status	PR status	Her2/ERBB2 status
BT474	Luminal	+	-	+
HCC1143	BasalA	-	-	-

Exercise plan

- Evaluate differences between sample groups
 - Luminal vs basal breast cancer
 - Cuffdiff tool
 - Gene enrichment via GSEA/MSigDB
- Process subset of one of the cell lines
 - BT474

Evaluate Gene Expression Differences

- Goal
 - Gene expression comparison
 - Luminal vs Basal
- Tasks <u>Completed by BRL staff</u>
 - RNA-Seq received from Joe Gray and uploaded to the Genboree Workbench
 - Map reads using Tophat (via Genboree Workbench)
 - File locations (output from Tophat job)
 - Epigenome Toolset Demo Input Data -> Databases -> RNA-Seq Tool Demo -> Files
 - » TopHat-BT474_accepted_hits.bam (Luminal)
 - » TopHat-HCC1143_accepted_hits.bam (Basal)
 - Job metrics
 - 4-10 hours
 - 8 cores
- Tasks <u>To do for workshop attendees</u>
 - Run cuffdiff to find significant changes in transcript expression, splicing, and promoter use

Create an Entity List for BT474

(i)) S	System/Network ▾) 🔲 Data ▾	QC and Pre-processing ▼	Genome -	Transcriptome -	Cistrome -	Epigenome •	Metagenome -	Visua
N	elcome to the Ger	nboree Workben	ch! [Gettin	g Started]				
)ata Selector			Details				
4	nefresh	Data Filter: Select a filte	er 💌	Attribute		Value		
4	www.genboree.org	Due a Tauliat		Group		Genboree	eUser_group	
	EDACC	Drag TopHat		1474_accepted_	_nits.bam*	Into Input D	ataabase	
	Epigenome Informatics A A Epigenome ToolSet Der	Workshop (May 2012) no Input Data		Description		Template UCSC Bui	for Human Genome, ild Hg19	
	🔺 🚰 Databases			Species		Homo sap	biens	
	Dia Brain			Version		hq19		Ŧ
	Breast Breast 450K			Input Data				
	Disease Epigenor	ne	≡	* ± ¥				
	MeDIP and GSEA			TopHat-BT4	74 accepted b	its ham		
	Peak Calling Demo	0			r4_accepted_n	into.bain		
	A RNA-Seq Tool Der	no		/				
		in Database						
	Lists & Selection	ons						
	Sample Sets							
	b 🚵 Samples							
	4 📑 Files	A secondard bits have		Output Targets				
		/4_accepted_nits.bam		🛨 🐳 🗙				
	TopHat-HCC	1143_accepted_hits.bam						
	⊳ 💾 Queries							

Choose a Destination for the Entity List

🔅 System/Network ▾ 🔲 Data ▾	QC and Pre-processing ◄	Genome 🕶	Transcriptome -	Cistrome ▼	Epigenome 🗸	Metagenome -	 Visualizatio
Welcome to the Ge	nboree Workber	Ch! [Gettin	g Started]				
Data Selector			Details				
🤣 Refresh	Data Filter: Select a	filter 🔻	Attribute		Value	e	
 Www.genboree.org Atlas Tools Access EDACC Epigenome Informatic Epigenome ToolSet De Epigenomics Roadmag GenboreeUser_group Databases Databases GenboreeUser_o GenboreeUser_o GenboreeUser_o GenboreeUser_o GenboreeUser_o SonathanMill_Lab SonathanMill_Lab SonathanMill_Lab ROI Repository SonathanSes Vamin_group 	s Workshop (Ma emo Input Data o Repository database	e destinati Genboree w	Group on database in vhere to deposi Species Version Input Data TopHat-I Output Targe Cutput Targe Genbore	ato Output it the Entity BT474_accepte	Gent Targets (this f / List) UCS Homo hg19 d_hits.bam	boreeUser_group tells r_database Human Gen C Build Hg19 o sapiens	e iome,

Invoke the "Create File Entity List" Tool

🔅 System/Network 🗸	Data - QC and	Pre-proces	sing 🗸	Genome 🗸	Tran	iscriptome 🗕	Cistrome 🕶	Epigenome 🗸	Metagenome -	● Visualization •	
Welcome to	Databases	▶	ben	ch! [Getting	Start	ed]					
	🚊 Entity Lists	Þ	i⊟ Tra	ack Entity List 🕨	1						
Data Selector	Entrypoints	Þ	Ξ Fil	e Entity List 🛛 🕨	1.3	Create File E	Entity List				
参 Refresh	📄 Files	Þ	er: Seleo	ct a filter	15	Copy/Move F	File Entity List	Create File	Entity List)	
⊿ 🗊 www.genbor	. 📄 Projects	Þ		1	15	Delete/Remo	ove File Entity List	Create/Add	a new file entity list in a	a selected database.	
▷ 🛃 Atlas Tool	Atlas Tool 🔞 Samples & Sample Sets				2	Rename File	Entity List		Genhoreeliser d	atabase	
Epigenom	Tracks	Þ	2)			Description			Template for Hum	an Genome,	
Epigenom	ne ToolSet Demo Input	Data				Spacias			UCSC Build Hg19		
≥ and a generation and a second seco	eUser group	ry	i			Version	Version				
a 🔤 Databa	ises			1		Toront Date			Ŭ		
D 📄 Ger	boreeUser_database				Input Data						
	ts		/ ± ∓ X								
⊳ 🛃 Jonathan	Mill Lab		TopHat-BT474_accepted_hits.bam								
b A paithank	group		· /								
Public	citory	Invoke	e "Crea	ate File Entit	v Lis	st". which	will present t	he dialog	ue		
⊳ 📌 Targeted	Atlases	box or	n the r	next slide	·			0			
⊳ 🛃 vamin_gr	oup										
					Output Ta	rgets					
					🛨 🐳 🗙						
						🍙 Genb	oreeUser_databa	se			

Give Your First Entity List a Meaningful Name

1	Tool Settings			×				
] Data ◄		Create File	e Entity List	genome →				
he Ger	Tool Overview							
	Files of Inte	erest:		-				
org ccess	Database: Group: File:	RNA-Seq Tool D Epigenome Tool TopHat-BT474_a	emo Set Demo Input Data accepted_hits.bam	e poreeUser_g nistrator				
nformatics 'oolSet Der Roadman	Target Database: Database: GenboreeUser_databaseGroup: GenboreeUser_group							
er_group \$ reeUser_d:	Settings Entity List	Name Luminal_Brs	t_Cancer_BT474					
I _Lab oup		Submit	Cancel					
ory ases o	Name your	entity list, and th	nen click "submit"					
L			± ∓ X					
			📄 GenboreeUser_database					

Message Stating First Entity List Creation

		Details									
-	Jo	b Submission Status	×								
rg		Create File Entity List		User_gro							
cess		Job Id: wbLocal-createFileList-05FyvX-3829	oree	itor User da							
format	lics	ADDED: 1 optity LIPLs were added to the "files"-type optity list named	late	for Huma							
olSet Roadm	Den ap I	"Luminal_Brst_Cancer_BT474".	sap	iens							
r_groເ	If you have questions, please contact genboree_admin@genboree.org for										
eeUse	assistance.										
Lab		OK									
ıp				P							
y ses	The entity list will be created immediately and will appear in your destination database under "lists and Selections" \rightarrow "lists of Files" (next slide)										
L .			-								
		Output Targets									

BT474 Entity List is Now Created



Create an Entity List for HCC1143 (execute same steps as for BT474)

System/Network ▼ □ Data ▼ QC and Pre-processing ▼	Genome -	Transcriptome -	Cistrome ▼	Epigenome 🗸	Metagen	ome 🕶
Welcome to the Genboree Workber	nch! [Gettin	ng Started]				
Data Selector		Details				
Select a filter	~	Attribute		Value		
A i www.genboree.org	A	Download		Click to Downlo	oad File	
Atlas Tools Access		Group		Epigenome Tool Input Data	Set Demo	Ξ
Epigenome Informatics Workshop (May 2012)		Database		RNA-Seq Tool D	emo	
A Bepigenome ToolSet Demo Input Data		Description				
a 📮 Databases		Name		TopHat-		-
🕞 🗑 Brain						
Breast		Input Data				
▷ 🎯 Breast 450K	=	★ ∓ X				
Disease Epigenome		TopHat-HCC11	43 accepted hit	ts.bam		
MeDIP and GSEA		▲ .				
Peak Calling Demo	1					
▷ □ Lists & Selections						
Sample Sets	1	Output Targets				
Samples		+ ∓ X				
4 📑 Files		1				
TopHat-BT474_accepted_hits.bam						
⊳ 📑 mRNA.subset						
TopHat-HCC1143_accepted_hits.bam						
Queries	~					

Create the HCC1143 Entity List

1	System/Network -	🔲 Data 🗸	QC and Pre-proc	essing -	Genome -		Tran	scriptome -	Cistrome -	Epig	genome -	Metagenome -	Visua	
	Welcome to	Datas	ases Lists	kbe	nch! [Gett Track Entity Lis	ing S	tarte	ed]						
	Data Selector	Entry	points	i⊟ ctafi	File Entity List	•	119 119	Create Fil Copy/Mov	e Entity List ve File Entity Lis	st	Value			
	www.genbor Atlas Tool Atlas Tool Atlas Cool Tracks							Delete/Re Rename F	move File Entit	ty List	Genboree administra Genboree	User_group tor User_database		
	 A Epigenom 2) A Epigenome ToolSet Demo Input Data A Epigenomics Roadmap Repository A GenboreeUser_group 							Description Species Version				Template for Human Genome, UCSC Build Hg19 Homo sapiens hg19		
	Databases GenboreeUser_database Projects						Input Data							
	▷ A GMT_Tuto ▷ A GMT_Tuto ▷ A JonathanN ▷ A paithank	rial //ill_Lab group(Doen Enigenon	ne Toc	olset Demo	Inn	TopHat-HCC1143_accepted_hits.bam							
	 Open Epigenome Toolset Demo Input Da drag HCC1143_accepted_hits.bam into Input Da drag HCC1143_acc								ata r_database) into (Output 1	Fargets		
						*	Out	put Target 🐳 🗶 Genbore	s eUser_database	9				

Give Your Second Entity List a Meaningful Name

-		_								
	Tool Settings	1								
a 🕶 📃	Create File Entity List ?	igenome -								
Ger	Tool Overview									
	Files of Interest:									
s	Database:RNA-Seq Tool DemoGroup:Epigenome ToolSet Demo Input DataFile:TopHat-HCC1143_accepted_hits.bam	e < to Dowr enome To								
nation	Target Database:	-Seq Tool								
et Der	Der Database: GenboreeUser_databaseGroup: GenboreeUser_group									
	Settings	1143_acc								
(genor	Entity List Name Basal_Brst_Cancer_HCC114	-								
GSEA g Dem ol Der ations	Submit Cancel									
electio ets	Name your entity list, and then click "submit"									
t-BT47	/4_accepted_mts.bam									
.subse	Et									
II-HCC										

Message Stating Second Entity List Creation



The entity list will be created immediately and will appear in your destination database under "Lists and Selections" \rightarrow "Lists of Files" (next slide)

Run Cuffdiff to Find Expression Differences



Name The Two Samples Before Submitting

o the Too	l Settings	×	
	Detect Transcription Changes by Cuffdiff	8	
	Tool Overview		
s Acce	Input Files:	_	a
≘ Infoi ≘ Tool	File Entity Lists:Basal_Brst_Cancer_HCC1143 Luminal_Brst_Cancer_BT474		
nics Ro eUser	Output Database:		
ises	Database: GenboreeUser_databaseGroup: GenboreeUser_group		E
All Anno	Settings		ł
.ists &	Analysis Name Cuffdiff-2013-2-19-15:59:47		
Bi	Time Series		
Ξ Lι	Min Align Count 10		
Sample	Multi Read Correct 🔲		Ľ
iles	Upper Quart Norm		Ľ
Queries ts	Sample Name For: Luminal_Brst_Cancer_BT474 Lumina_Br_Cancer_BT474		
orial	Sample Name For: Basal_Brst_Cancer_HCC1143		
_group	Basal_Br_Cancer_HCC1143		F
sitory	Advanced Settings:		
A 41			
	Submit Cancel		

Message Denoting Successful Cuffdiff Submission



Cuffdiff email Notification

You will receive an email notifying you that the Cuffdiff job is complete. It will tell you where the output file is located within the Data Selector (see next slide).

Hello Genboree User,

Your job completed successfully.

Job Summary: JobID - wbJob-cuffdiff-IwKv5s-2270

Additional Info: Database: 'GenboreeUser_database' Group: 'GenboreeUser_group'

You can download result files from the 'Cuffdiff-2013-2-19-15:59:47' folder under the 'Cuffdiff' directory.

Note that files without data are stored under the 'raw' folder.

- The Genboree Team

Download Cuffdiff Results



Example Data Returned by Cuffdiff

	A	В	С	D	E	F	G
1	test_id 🖃	gene_Name	gene_id	gene 🔻	locus 💌	sample_1	sample_2 🔹
2	NM_000014	A2M	NM_000014	-	chr12:9217772-9268558	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
54	NM_000067	CA2	NM_000067	-	chr8:86376130-86393721	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
191	NM_000206	IL2RG	NM_000206	-	chrX:70327253-70331481	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
199	NM_000214	JAG1	NM_000214	-	chr20:10618331-10654694	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
401	NM_000422	KRT17	NM_000422	-	chr17:39775691-39780882	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
457	NM_000486	AQP2	NM_000486	-	chr12:50344523-50352664	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
563	NM_000598	IGFBP3	NM_000598	-	chr7:45951843-45960871	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
579	NM_000615	NCAM1	NM_000615	-	chr11:112831968-113149158	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
613	NM_000660	TGFB1	NM_000660	-	chr19:41836811-41859831	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
805	NM_000864	HTR1D	NM_000864	-	chr1:23518387-23521222	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
853	NM_000916	OXTR	NM_000916	-	chr3:8792094-8811300	Basal_Br_Cancer_HCC1143	Lumina_Br_Cancer_BT474
949	NM 001001389	CD44	NM 001001389	-	chr11:35160416-35253949	Basal Br Cancer HCC1143	Lumina Br Cancer BT474

G	Н		J	К	L	М	N	0
sample_2 🔹	status 💌	value_1 💌	value_2 💌	log2(fold_change) 💌	test_stat 🝷	p_value 💌	q_value 💌	significant 🔻
Lumina_Br_Cancer_BT474	OK	0.0284975	9.58365	8.3936	-3.76582	0.000166001	0.0179909	yes
Lumina_Br_Cancer_BT474	OK	86.5546	0.474705	-7.51044	5.13723	2.79E-07	0.000146096	yes
Lumina_Br_Cancer_BT474	OK	0.0469928	4.11807	6.45338	-3.40051	0.000672607	0.0483116	yes
Lumina_Br_Cancer_BT474	OK	56.8077	0.398662	-7.15478	4.94107	7.77E-07	0.000330952	yes
Lumina_Br_Cancer_BT474	OK	976.538	1.83901	-9.05261	5.54106	3.01E-08	2.75E-05	yes
Lumina_Br_Cancer_BT474	OK	0.0159363	3.78985	7.89368	-4.31386	1.60E-05	0.00345697	yes
Lumina_Br_Cancer_BT474	OK	453.591	7.67741	-5.88463	3.56143	0.000368835	0.031866	yes
Lumina_Br_Cancer_BT474	OK	12.3395	0.0291782	-8.72418	5.1717	2.32E-07	0.00012875	yes
Lumina_Br_Cancer_BT474	ОК	39.4051	0.561805	-6.13217	4.3299	1.49E-05	0.0033732	yes
Lumina_Br_Cancer_BT474	OK	9.28553	0.0326627	-8.1512	4.66051	3.15E-06	0.000916212	yes
Lumina_Br_Cancer_BT474	OK	16.0244	0.535301	-4.90377	3.53338	0.000410283	0.0343139	yes
Lumina_Br_Cancer_BT474	OK	111.124	0.411939	-8.07552	3.61338	0.000302228	0.0282115	yes
Lumina_Br_Cancer_BT474	OK	0.0591867	6.92481	6.87036	-4.01098	6.05E-05	0.00878971	yes

Note: Sheet sorted by the 'significant' column

Further Analysis Using GSEA/MSigDB

- Evaluate differences between sample groups
 - Luminal vs basal breast cancer
 - Cuffdiff tool
 - Gene enrichment via GSEA/MSigDB

GSEA/MSigDB

- Gene Set Enrichment Analysis
 - Subramanian, Tamayo, et al. 2005, PNAS 102, 15545-15550
 - Mootha, Lindgren, et al. 2003, Nat Genet 34, 267-273
- Molecular Signatures Database
 - Subramanian, Tamayo, et al. 2005, PNAS 102, 15545-15550
- Exposed as a web service

Register with MSigDB

http://www.broadinstitute.org/gsea/login.jsp

					login register	BROAD
Gene Set Enrichment Analysis	GSEA Home	Downloads	Molecular Signatures Database	Documentation	Contact	
Login to GSEA/	/MSigDB					
Login Click here to register to view	the MSigDB gene sets	s and/or download t	he GSEA software. This helps us trac	k and better serve our u	user communi	ty.
If you have already registered	d for GSEA or MSigDE uired.	3 please enter your	registration email address below.			
Email: *						
login						

Register with MSigDB

GSEA/MSigDB Registration and License Agreement

Instructions to obtain GSEA software and/or MSigDB gene sets. Please Read carefully.

- 1. Fill in the form below.
- 2. The software and gene sets are freely available to individuals in academic and private institutions. There are no licensing fees.
- 3. Source code is freely available.

 Read the license agreement and make sure you agree with the terms of the agreement. If so, click the 'I Agree button' at the end of the form and you will be transferred to the GSEA download page.

Items marked with * are required.

Name: *		
Email: *	(You will receive a registration notification email.)	
Organization: *		
Country: *	select a country	
Join mailing list:	notify me of GSEA updates (You will receive a confirmation email. Reply to join the list.)	
Comments:		
	ł.	
GSEA and MSigDB license agreements:	**** GSEA/MSigDB LICENSE AGREEMENT ****	A
	MASSACHUSETTS INSTITUTE OF TECHNOLOGY SINGLE USER LICENSE AGREEMENT FOR INTERNAL RESEARCH PURPOSES ONLY	-
	This Agreement is made between Massachusetts Institute of Technology with a principal address at 77	-
	Manager and the Annual Combandary MA 00100 (INTTIL) and the	

I AGREE

Login to MSigDB

http://www.broadinstitute.org/gsea/login.jsp



Next:

login

- Click 'Explore the Molecular Signatures Database (MSigDB)'
 - http://www.broadinstitute.org/gsea/msigdb/index.jsp
- Click 'Investigate Gene Sets'
 - http://www.broadinstitute.org/gsea/msigdb/annotate.jsp

Investigate Gene Sets

Gain further insight into the biology behind a gene set by using the following tools:

- compute overlaps with other gene sets in MSigDB (more...)
- display the gene set expression profile based on a selected compendium of expression data (more...)
- categorize members of the gene set by gene families (more...)



٠

Compendia expression profiles

 Human tissue compendium (Novartis) Global Cancer Map (Broad Institute) NCI-60 cell lines (National Cancer Institute)

display expression profile

Gene families

show gene families

GENE SYMBOL

Gene expression differences

Filter by "significant," column

1	1	A	В		С	D		E		F	G	H		I	J	K	L	M	N	0	
1	test_	id 🔰	gene_Nar 🔻	eene	e_id 🔻	gene	-	locus	*	sample	sample_ 💌	status	•	value_1 -	value_2 -	log2(fol ▼	test_sta 💌	p_value 🔻	q_valte ▼	significant	Л
5	NM_	0000	ACADS	NИ	0000	-		chr12:12	21:	Luminal	BasalA	OK		0.181132	9.56262	5.72229	-5.28258	1.27E-07	5.70E-06	уса	
10	NM_	0000	ADA	NN	0000	-		chr20:43	324	Luminal	BasalA	OK		0.046568	12.6184	8.08196	-5.02623	5.00E-07	1.95E-05	yes	
32	NM_	000	AR	NM	0000	-		chrX:66	76	Luminal	BasalA	OK		10.932	0.008774	-10.2831	6.59281	4.32E-11	4.09E-09	yes	
41	NM_	0000	ATP7B	NM	0000	-		chr13:52	25(Luminal	BasalA	OK		7.02049	0.417952	-4.07016	4.16209	3.15E-05	0.000777	yes	
51	NM_	000	(C3	NM	0000	-		chr19:6	6 7 1	Luminal	BasalA	OK		0.038313	48.1341	10.295	-10.0105	0	0	yes	
88	NM_	001	CYBA	NM	0001	-		chr16:8	87(Luminal	BasalA	OK		32.1228	0.168313	-7.57631	4.75509	1.98E-06	6.67E-05	yes	
91	NM_	0001	CYP1B1	NM	0001	-		chr2:38	294	Luminal	BasalA	OK		5.49719	33.7556	2.61836	-3.01935	0.002533	0.030879	yes	
195	NM_	0002	ITGA6	NM	0002	-		chr2:17	329	Luminal	BasalA	OK		3.16834	29.6013	3.22386	-3.59329	0.000327	0.006	yes	
199	NM_	0002	JAG1	NM	0002	-		chr20:10	06:	Luminal	BasalA	OK		0.530239	54.1496	6.67416	-6.70849	1.97E-11	2.02E-09	yes	
254	NM_	002	NPC1	NM	0002	-		chr18:2	10	Luminal	BasalA	OK		5.22127	31.5541	2.59535	-2.96999	0.002978	0.035004	yes	
290	NM_	003	CTSA	NM	0003	-		chr20:44	45:	Luminal	BasalA	OK		59.9431	3.58534	-4.06341	4.16583	3.10E-05	0.000766	yes	
327	NM_	0003	SOX9	NM	00034	-		chr17:7	01:	Luminal	BasalA	OK		4.77218	61.8855	3.69688	-4.09221	4.27E-05	0.001003	yes	
392	NM_	0004	HSD17B1	NM	0004	-		chr17:40	07(Luminal	BasalA	OK		27.5426	3.35674	-3.03653	3.10584	0.001897	0.024754	yes	
401	NM_	0004	KRT17	NN.	_0004	-		chr17:3	97	Luminal	BasalA	OK		0.180647	587.103	11.6662	-10.4385	0	0	yes	
403	NM_	0004	KRT5	NИ	_0004	-		chr12:52	29(Luminal	BasalA	OK		0.27407	625.903	11.1572	-10.2628	0	0	yes	
414	NM	0004	NOTCH3	<u>р.м</u>	0004	-		chr19:1	52	Luminal	BasalA	OK		11.6804	176.97	3.92134	-4.11637	3.85E-05	0.000926	yes	

Copy "official" gene symbol



Compute Overlaps for Selected Genes

Converted 701 submitted identifiers into 599 gene symbols. click here for details.

Collections	# Overlaps	# Gene Sets in	# Genes in Comparison	# Genes in Collections
	Shown	Collections	(n)	(N)
C2, C4, C5	10	5607	599	22684

Click the gene set name to see the gene set page. Click the number of genes [in brackets] to download the list of genes.

Color bar shading from light green to black, where lighter colors indicate more significant p values (< 0.05) and black indicates less significant p values (>= 0.05).

Export: Excel

Gene Set Name [# Genes (K)]	Description	# Genes in Overlap (k)	k/K	p value 🛿
NUYTTEN_NIPP1_TARGETS_DN [777]	Genes down-regulated in PC3 cells (prostate cancer) after knockdown of NIPP1 [Gene	67		0 e ⁰
SMID_BREAST_CANCER_BASAL_DN [713]	Genes down-regulated in basal subtype of breast cancer samles.	65		0 e ⁰
SMID_BREAST_CANCER_LUMINAL_B_DN [599]	Genes down-regulated in the luminal B subtype of breast cancer.	63	-	0 e ⁰
CREIGHTON_ENDOCRINE_THERAPY_RESISTANCE NCE_5 [482]	The 'group 5 set' of genes associated with acquired endocrine therapy resistance in breast tumors expressing ESR1 but not ERBB2 [Gene ID=2099, 2004].	54		0 e ⁰
SMID_BREAST_CANCER_BASAL_UP [676]	Genes up-regulated in basal subtype of breast cancer samles.	78	-	0 e ⁰
CHARAFE_BREAST_CANCER_LUMINAL_VS_BASAL SAL_DN [456]	Genes down-regulated in luminal-like breast cancer cell lines compared to the basal-like	59		0 e ⁰
	onesi			

SMID_BREAST_CANCER_BASAL_DN [713]	Genes down-regulated in basal subtype of breast cancer samles.	65	0 e ⁰
SMID_BREAST_CANCER_LUMINAL_B_DN [599]	Genes down-regulated in the luminal B subtype of breast cancer.	63	0 e ⁰
SMID_BREAST_CANCER_BASAL_UP [676]	Genes up-regulated in basal subtype of breast cancer samles.	78	0 e ⁰
CHARAFE_BREAST_CANCER_LUMINAL_VS_BASAL SAL_DN [456]	Genes down-regulated in luminal-like breast cancer cell lines compared to the basal-like ones.	59	0 e ⁰

Enrichments for gene sets differentiating luminal vs basal breast cancer cells

Acknowledgments

• BRL

- Aleksandar Milosavljevic
- Cristian Coarfa, Alan R Harris
- BRL core
 - Matt Roth, Kevin Riehle
- Genboree programming team
 - Andrew Jackson, Sameer Paithankar, Sriram Raghuram
- Former BRL members
 - Chia-Chin Wu, Arpit Tandon
- Peak Calling
 - Anshul Kundaje
 - Bob Thurman (UW)
 - Noam Shoresh (BI)
 - Martin Hirst (BCGSC)
 - Lee Daniels (NIH)
 - Wei Li (BCM)
- RNA-Seq
 - Adrian Lee, Joe Gray
 - Laising Yen, Kalpana Kannan
 - Sean McGuire, Christine Pichot