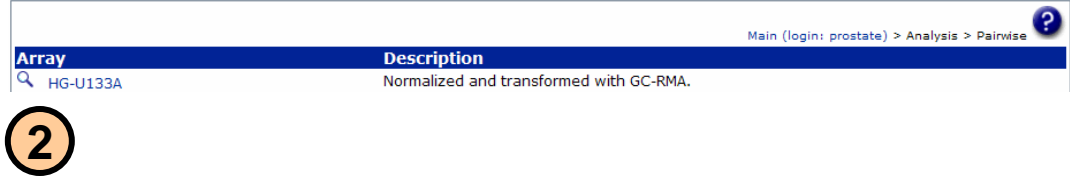
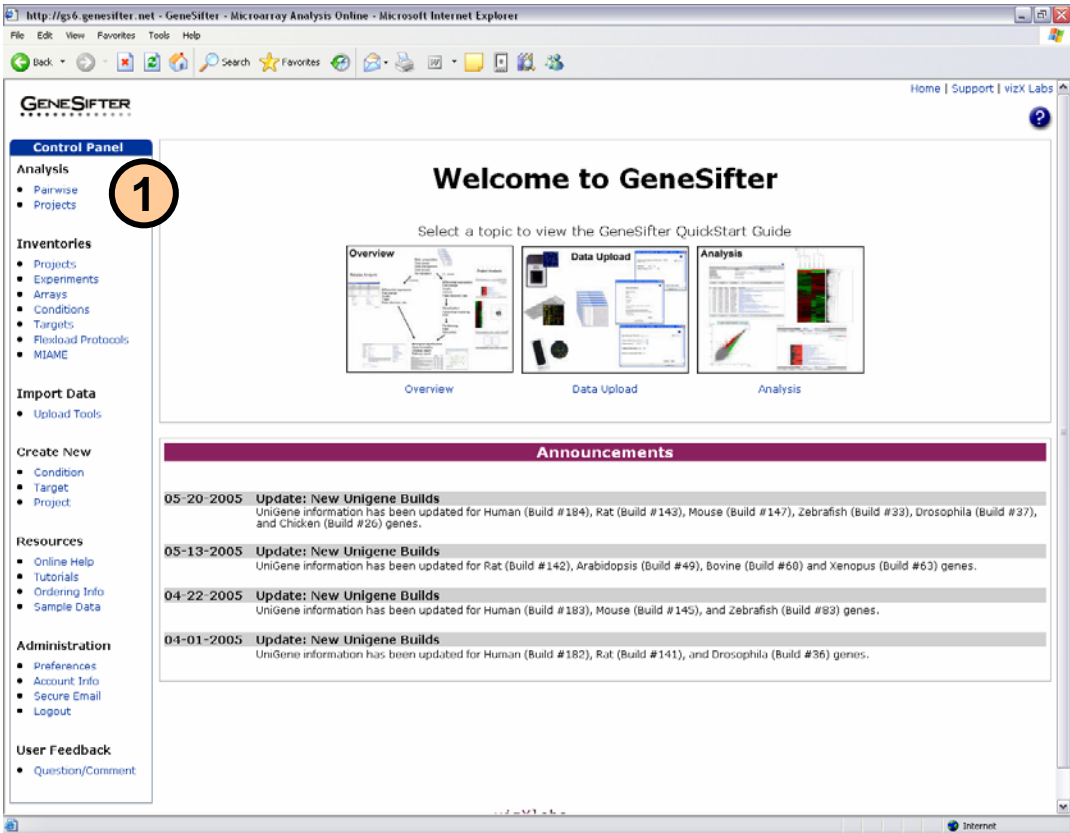


Androgen-independent prostate cancer

The following tutorial walks through the identification of biological themes in a microarray dataset examining androgen-independent prostate cancer.

Visit the GeneSifter Data Center (www.genesifter.net/web/dataCenter.html) to register for free access to the dataset.

1. From the **Control Panel** select “Pairwise” from the **Analysis** section.
2. Click on the magnifying glass icon next to “HG-U133A” to begin the analysis. The data presented here was generated using the Affymetrix® GeneChip® Human Genome U133A array. There are approximately ~22,000 transcripts represented on this array.



Androgen-independent prostate cancer

(continued)

- At the top of the page is a list of the different experiment groups contained in this analysis. We'll be comparing expression between the androgen-dependent and -independent samples. Select the ten "**Androgen-dependent**" arrays to place them in Group 1. Select the ten "**Androgen-independent**" samples for Group 2.
- Pairwise analysis combines a fold-change cutoff and comparison statistics to generate a list of differentially expressed genes. Select the following analysis parameters:

Normalization: None

Data was already normalized with GC-RMA during upload.

Statistics: Wilcoxon

Performs a Wilcoxon rank sum test for each gene that passes the fold change cutoff.

Quality (Calls): N/A

Quality calls are not generated by GC-RMA.

Check **Exclude Controls**.

Threshold: 1.5

Filters out genes with less than 1.5 fold change in expression.

Correction: Benjamini and Hochberg

Calculates a false discovery rate from the raw p-values using the method of Benjamini and Hochberg.

Data Transformation: Data Already Log Transformed.

Data was logged during GC-RMA normalization.

- Click **Analyze**.

3

Main (login: prostate) > Analysis > Arrays > Pairwise

Pairwise Analysis: HG-U133A

Group		Experiment	Target	Condition
1	2			
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45730 -1-	GSM45730	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45847 -1-	GSM45847	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45848 -1-	GSM45848	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45849 -1-	GSM45849	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45850 -1-	GSM45850	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45851 -1-	GSM45851	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45852 -1-	GSM45852	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45853 -1-	GSM45853	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45854 -1-	GSM45854	Androgen-dependent
<input checked="" type="checkbox"/>	<input type="checkbox"/>	GSM45855 -1-	GSM45855	Androgen-dependent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45856 -1-	GSM45856	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45857 -1-	GSM45857	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45858 -1-	GSM45858	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45859 -1-	GSM45859	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45860 -1-	GSM45860	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45861 -1-	GSM45861	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45862 -1-	GSM45862	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45863 -1-	GSM45863	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45864 -1-	GSM45864	Androgen-independent
<input type="checkbox"/>	<input checked="" type="checkbox"/>	GSM45865 -1-	GSM45865	Androgen-independent

Advanced Analysis Settings (RMA)

4

Normalization: None

Statistics: Wilcoxon

Quality (Calls): N/A

Exclude Controls

Show genes that are:

Up-regulated

Down-regulated

Threshold:

Lower: 1.5 Upper: None

Correction: Benjamini and Hochberg

Data Transformation:

No Transformation

Log Transform Data

Data Already Log Transformed

5

Analyze Reset

Androgen-independent prostate cancer

(continued)

6. After the analysis is performed, a gene list will be returned. The **Pairwise Analysis Results Page** lists the genes that are differentially expressed, based on the selected parameters. 785 genes passed the initial filtering criteria. The genes are sorted by fold change, and the first 50 genes in the list are displayed.
7. To filter the list using the adjusted p-value (false discovery rate), select “adjusted p” from the pull-down menu, and then click the **Search** button. Using the adjusted p-value reduces the gene list to 468 results with a false discovery rate of less than 5%.
8. To view data and a gene summary for any gene in the list, select the **Gene Name**.

[Main \(login: prostate\)](#) > [Analysis](#) > [Pairwise](#) > [Results](#)

Pairwise Analysis: HG-U133A
[[Reports: Ontology](#) | [KEGG](#) | [Scatter Plot](#)] [[Results: Export](#) | [Save](#)]

	Group 1	Group 2
Conditions:	Androgen-dependent	Androgen-independent
Experiments:	62460, 62461, 62462, 62473, 62474, 62475, 62476, 62477, 62478, 62479	62463, 62464, 62465, 62466, 62467, 62468, 62469, 62470, 62471, 62472
Significance:	1.5, Wilcoxon, Benjamini and Hochberg	
Normalization:	None	
Quality Cutoff:	N/A	
Data Transformation:	Log Transformed	

Show:
Sort By:
p Cutoff:

(468 results found)
[1 - 50] [51 - 100]

No.	Ratio	p-value	adj. p	Identifier	Gene Name
1	6.94	0.00893	0.04109	NM_014668	GREB1 protein
2	6.23	0.00893	0.04109	NM_005651	Tryptophan 2,3-dioxygenase
3	5.25	0.00150	0.02199	AF125393	RAB27A, member RAS oncogene family
4	5.08	0.00032	0.01767	NM_006705	Growth arrest and DNA-damage-inducible, gamma
5	4.90	0.01150	0.04446	AI819238	Inhibitor of DNA binding 2, dominant negative helix-loop-helix protein
6	4.87	0.01469	0.04953	NM_001085	Serpin peptidase inhibitor, clade A (alpha-1 antitrypsin), member 1
7	4.62	0.00013	0.01708	NM_017680	Asporin (LRR class 1)
8	4.48	0.00150	0.02199	NM_015675	Growth arrest and DNA-damage-inducible, beta
9	4.42	0.00288	0.02771	BF440025	Nephroblastoma overexpressed gene
10	4.42	0.00073	0.01846	AL050388	Superoxide dismutase 2, mitochondrial
11	4.39	0.00073	0.01846	NM_001855	Collagen, type XV, alpha 1
12	4.25	0.00893	0.04109	AC004010	Homo sapiens BAC clone GS1-99H8 from 7, complete sequence.
13	4.24	0.00209	0.02536	AI796581	KIAA0056 protein
14	4.08	0.00150	0.02199	AK027217	PDZ and LIM domain 5
15	4.07	0.00049	0.01830	D21254	Cadherin 11, type 2, OB-cadherin (osteoblast)
16	4.05	0.00893	0.04109	NM_001150	Alanyl (membrane) aminopeptidase (aminopeptidase N, aminopeptidase M, microsomal)
17	4.04	0.00520	0.03445	NM_001554	Cysteine-rich, angiogenic inducer, 61
18	3.83	0.01150	0.04446	NM_006931	Solute carrier family 2 (facilitated glucose transporter), member 3
19	3.78	0.00073	0.01846	BE502030	RAB27A, member RAS oncogene family
20	3.78	0.00893	0.04109	NM_024642	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase
21	3.69	0.00684	0.03647	NM_003226	Trefoil factor 3 (intestinal)
22	3.68	0.01469	0.04953	NM_001993	Coagulation factor III (thromboplastin, tissue factor)
23	3.67	0.00021	0.01708	BG253119	Dystonin
24	3.65	0.00389	0.03129	NM_005780	TRAFs and NIK-associated protein
25	3.60	0.00389	0.03129	NM_012105	Beta-site APP-cleaving enzyme 2

Androgen-independent prostate cancer (continued)

9. Selecting a gene from the list will bring up the data summary and a One-Click Gene Summary™ for the gene. The One-Click Gene Summary provides a synopsis of current UniGene and Entrez Gene information for the gene.

Go back to the gene list by clicking the **Back** button in your browser.

10. Select the **Ontology** link to view a summary of the Gene Ontology terms associated with the genes in the list. See the online help for more information about the other reports.

Note: To view page-specific help documents for any page, select the question mark icon (?) in the upper right corner of each page.

9

Main (login: prostate) > Analysis > Pairwise > Results > Gene Summary

» Gene Summary: GREB1 protein

• By Group

Group	Condition	N	Mean	SEM	SEM/Mean	Signal Mean
1	Androgen-dependent	10	9.6399	+/- 0.2897	3%	9.6399
2	Androgen-independent	10	6.8450	+/- 0.7079	10.3%	6.8450

By Target

Group	Sample	N	Mean	SEM	SEM/Mean	Signal
1	GSM45848	1	8.8594	-	-	8.85942
1	GSM45847	1	9.8246	-	-	9.82456
1	GSM45730	1	10.3107	-	-	10.3107
1	GSM45855	1	10.7046	-	-	10.7046
1	GSM45854	1	8.7320	-	-	8.73199
1	GSM45853	1	10.7173	-	-	10.7173
1	GSM45852	1	9.3042	-	-	9.30418
1	GSM45851	1	9.9504	-	-	9.95041
1	GSM45850	1	10.0727	-	-	10.0727
1	GSM45849	1	7.9234	-	-	7.92345
2	GSM45865	1	4.6837	-	-	4.68367
2	GSM45864	1	6.5770	-	-	6.57698
2	GSM45863	1	9.8884	-	-	9.88843
2	GSM45862	1	4.6329	-	-	4.63291
2	GSM45861	1	7.0269	-	-	7.02693
2	GSM45860	1	4.8884	-	-	4.88837
2	GSM45859	1	10.7343	-	-	10.7343
2	GSM45858	1	6.6568	-	-	6.65682
2	GSM45857	1	4.7933	-	-	4.79327
2	GSM45856	1	8.5685	-	-	8.56846

Androgen-independent down-regulated 6.94 fold compared to Androgen-dependent

» One-Click Gene Summary™

Probe Set ID: 205862_at
 Accession No.: NM_014668
 Cluster ID: Hs.467733
 UG Title: GREB1 protein
 Gene ID: GREB1
 Homologene: -
 Chromosome: 2
 Cytoband: 2p25.1
 Seq Count: 144
 Entrez Gene: 9687
 Gene Name: GREB1 protein
 OMIM: -
 RefSeq mRNA: NM_148903 (FASTA)
 RefSeq Prot: NP_683701 (FASTA)
 Summary: This gene is an estrogen-responsive gene in the estrogen receptor-regulated pathway, role in hormone-responsive tissues and car transcript variants encoding distinct isoform

Gene Ontologies:
 Biological Process
 • biological process unknown
 Molecular Function
 • molecular function unknown
 Cellular component
 • cellular component unknown

10

Main (login: prostate) > Analysis > Pairwise > Results

Pairwise Analysis: HG-U133A [Reports: Ontology | KEGG | Scatter Plot] [Results: Export | Save]

Conditions:	Group 1	Group 2
	Androgen-dependent	Androgen-independent
Experiments:	62460, 62461, 62462, 62473, 62474, 62475, 62463, 62464, 62465, 62466, 62467, 62468, 62476, 62477, 62478, 62479	62469, 62470, 62471, 62472
Significance:	1.5, Wilcoxon, Benjamini and Hochberg	
Normalization:	None	
Quality Cutoff:	N/A	
Data Transformation:	Log Transformed	

Show: 50 Sort By: Ratio p Cutoff: 0.05 adjusted p Search (468 results found) [1 - 50] [51 - 100]

No.	Ratio	p-value	adj. p	Identifier	Gene Name
1	6.94	0.00893	0.04109	NM_014668	GREB1 protein
2	6.23	0.00893	0.04109	NM_005651	Tryptophan 2,3-dioxygenase
3	5.25	0.00150	0.02199	AF123393	RAB27A, member RAS oncogene family
4	5.08	0.00032	0.01767	NM_006705	Growth arrest and DNA-damage-inducible, gamma
5	4.90	0.01150	0.04446	A1819238	Inhibitor of DNA binding 2, dominant negative helix-loop-helix protein
6	4.87	0.01469	0.04953	NM_001085	Serpin peptidase inhibitor, clade A (alpha-1 antitrypsin, member 1)
7	4.62	0.00013	0.00109	NM_017680	Aspirin (LRR class 1)
8	4.48	0.00150	0.02199	NM_015675	Growth arrest and DNA-damage-inducible, beta
9	4.42	0.00288	0.02771	BF440025	Nephroblastoma overexpressed gene
10	4.42	0.00073	0.01846	AL050388	Superoxide dismutase 2, mitochondrial
11	4.39	0.00073	0.01846	NM_001855	Collagen, type XV, alpha 1
12	4.25	0.00893	0.04109	AC004010	Homo sapiens BAC clone GS1-99H8 from 7, complete sequence.
13	4.24	0.00209	0.02536	A1796581	KIAA0056 protein
14	4.08	0.00150	0.02199	AK027217	PDZ and LIM domain 5
15	4.07	0.00049	0.01830	D21254	Cadherin 11, type 2, OB-cadherin (osteoblast)
16	4.05	0.00893	0.04109	NM_001150	Alanyl (membrane) aminopeptidase (aminopeptidase N, aminopeptidase M, microsomal)
17	4.04	0.00520	0.03445	NM_001554	Cysteine-rich, angiogenic inducer, 61
18	3.83	0.01150	0.04446	NM_006931	Solute carrier family 2 (facilitated glucose transporter), member 3
19	3.78	0.00073	0.01846	BE502030	RAB27A, member RAS oncogene family
20	3.78	0.00893	0.04109	NM_024642	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylglucosaminyltransferase
21	3.69	0.00684	0.03647	NM_003226	Trefol factor 3 (intestinal)
22	3.68	0.01469	0.04953	NM_001993	Coagulation factor III (thromboplastin, tissue factor)
23	3.67	0.00021	0.01708	BG253119	Dystonin
24	3.65	0.00389	0.03129	NM_005780	TRAFs and NIK-associated protein
25	3.60	0.00389	0.03129	NM_012105	Beta-site APP-cleaving enzyme 2

Androgen-independent prostate cancer (continued)

- The Ontology Report lists the Gene Ontology terms associated with the 468 genes from the pairwise results list. See the page-specific help for a complete description of the fields in the Ontology Report.
- Click on **Z-score Report**.
- The z-score report lists the biological process ontology terms that are significantly over or under-represented in the gene list (z-score greater than 2, or less than -2, respectively). Select the red arrow in the z-score column to sort the list so that the most over-represented genes are at the top.

A z-score report can also be generated for the molecular function and cellular component categories.

11

Group 1: Androgen-dependent
Group 2: Androgen-independent

[Biological Process | Cellular Component | Molecular Function]

[Ontology Report | Z-score Report]

Ontology	Genes	GO	List	Array	z-score
cellular process	262	124	138	8234	2.56 -0.98
physiological process	243	108	135	7891	-0.69 -0.17
regulation of biological process	95	48	47	2571	2.37 0.50
development	48	28	20	1532	1.55 -1.37
response to stimulus	48	20	28	1577	-0.50 0.19
biological process unknown	19	9	10	492	0.83 0.55
growth	8	5	3	151	2.01 0.25
behavior	2	0	2	171	-1.57 -0.56
interaction between organisms	2	0	2	68	-0.99 0.78
reproduction	2	1	1	213	-1.17 -1.42
viral life cycle	1	0	1	29	-0.64 0.72
pigmentation	-	0	0	5	-0.27 -0.30

pie chart labels: physiological process (13.29%), regulation of biological process (13.01%), development (6.50%), response to stimulus (6.56%), biological process unknown (2.60%), growth (1.10%), behavior (0.27%), interaction between organisms (0.27%), reproduction (0.73%), viral life cycle (0.14%), cellular process (15.89%)

12

13

Group 1: Androgen-dependent
Group 2: Androgen-independent


[Biological Process | Cellular Component | Molecular Function]



[Ontology Report | Z-score Report]

Export Report

Ontology	Genes	GO	List	Array	z-score
cellular process	262	124	138	8234	2.56 -0.98
macromolecule metabolism	117	54	63	3072	2.06 1.73
regulation of biological process	95	48	47	2571	2.37 0.50
regulation of cellular process	92	48	44	2382	2.96 0.56
cellular macromolecule metabolism	90	37	53	2229	1.19 2.75
cellular protein metabolism	88	37	51	2192	1.31 2.51
protein metabolism	88	37	51	2213	1.24 2.43
regulation of physiological process	87	45	42	2337	2.49 0.33
regulation of cellular physiological process	86	45	41	2255	2.76 0.41
biopolymer metabolism	71	44	27	1996	3.45 -1.43
cell organization and biogenesis	43	27	16	1175	2.60 -1.01
biopolymer modification	42	25	17	1190	2.20 -0.83
protein modification	41	24	17	1153	2.10 -0.69
biosynthesis	38	5	33	856	-2.14 5.05
cellular biosynthesis	34	5	29	761	-1.83 4.63
macromolecule biosynthesis	31	4	27	473	-1.05 6.85
negative regulation of biological process	28	16	12	619	2.59 0.43
negative regulation of cellular process	28	16	12	574	2.92 0.71
protein biosynthesis	28	4	24	416	-0.78 6.50
cell cycle	25	19	6	577	3.99 -1.30
establishment of protein localization	25	9	16	484	0.88 2.76
protein localization	25	9	16	499	0.78 2.63
cell adhesion	24	14	10	489	2.82 0.57
cell death	24	10	14	452	1.50 2.31
death	24	10	14	496	1.47 2.28
programmed cell death	24	10	14	429	1.69 2.52
protein transport	24	9	15	460	0.98 2.53
apoptosis	23	10	13	427	1.69 2.16
cellular localization	23	13	10	512	2.25 0.42
establishment of cellular localization	23	13	10	508	2.28 0.44
intracellular transport	23	13	10	504	2.31 0.47
morphogenesis	21	12	9	490	2.03 0.20
intracellular protein transport	19	9	10	316	2.22 2.01

Androgen-independent prostate cancer (continued)

- Return to the **Pairwise Analysis Results Page** and click on the **KEGG** link. This will bring up a z-score report for the KEGG pathway terms associated with the gene list.
- Click on the KEGG logo () for **Ribosome** to show the KEGG pathway diagram. Differentially regulated genes are shown in red.

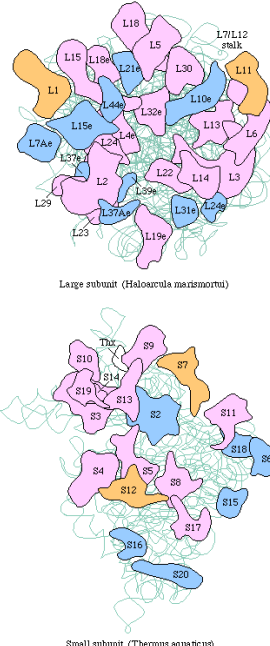



Group 1: Androgen-dependent
Group 2: Androgen-independent

Pathway	Genes	KEGG	Totals		z-score	
			List	Array		
Ribosome	16	0	16	80	-0.79	10.28
MAPK signaling pathway	9	1	8	227	-0.58	1.08
Oxidative phosphorylation	7	0	7	102	-0.90	2.93
Apoptosis	6	3	3	92	2.81	0.51
Cell cycle	6	3	3	93	2.79	0.49
Calcium signaling pathway	4	2	2	176	0.60	-1.17
Focal adhesion	4	2	2	74	1.95	0.14
Regulation of actin cytoskeleton	4	3	1	185	1.40	-1.74
Complement and coagulation cascades	3	2	1	68	2.10	-0.53
Cytokine-cytokine receptor interaction	3	0	3	236	-1.40	-1.23
Fructose and mannose metabolism	3	0	3	44	-0.58	1.88
Glutathione metabolism	3	0	3	32	-0.50	2.54
Purine metabolism	3	1	2	133	-0.01	-0.73
TGF-beta signaling pathway	3	1	2	76	0.57	0.10
Alzheimer's disease	2	1	1	22	2.05	0.63
Aminosugars metabolism	2	0	2	25	-0.44	1.80
Fatty acid metabolism	2	0	2	65	-0.71	0.33
Glycine, serine and threonine metabolism	2	0	2	36	-0.53	1.21
Inositol phosphate metabolism	2	2	0	70	2.05	-1.35
Tight junction	2	0	2	100	-0.89	-0.30
Toll-like receptor signaling pathway	2	1	1	94	0.35	-0.89
Tryptophan metabolism	2	1	1	76	0.57	-0.65
Ubiquitin mediated proteolysis	2	2	0	37	3.28	-0.97
Wnt signaling pathway	2	1	1	126	0.05	-1.23

14

RIBOSOME



Large subunit (*Halobacterium marismortui*)

Small subunit (*Thermus aquaticus*)

15

Ribosomal RNAs

Bacteria / Archaea	23S	5S	16S	
Eukaryotes	25S	5S	5.8S	18S

Ribosomal proteins

Bacteria / Archaea / Eukaryotes

EF-Tu: S10, L3, L4, L23, L2, S19, L22, S3, L16, L29, S20e, L3e, L4e, L23Ae, L8e, S15e, L17e, S3e, L10e, L35e

RpoA: S17, L14, L24, L5, S14, S8, L6, L18, S5, L30, L15, S11e, L23e, L26e, S4e, L11e, S29e, S15Ae, L9e, L32e, L19e, L5e, S2e, L7e, L27Ae

RpoC: L34e, L14e, L36, S13, S11, S4, L18e, L13, S9, S18e, S14e, S9e

EF-Tu/G: S7, S12, L30e, L7A, RpoC/B, L7L12, S5e, S23e, L7Ae, L12, L10, L1, L11, L1, L12, L10, L1, L11, L1, L12, L10, L10Ae, L12e, L1, L12, L10, L10Ae, L12e

EF-Ts: S2, S8Ae, S15, L35, L20, L34, R1, L31, L32, L9, S18, S6

Other: L28, L33, L21, L27, P0Y P0h, S16, L19, S1, S20, S21, L25

A/E: L10e, L13e, L15e, L21e, L24e, L31e, L35Ae, L37e, L37Ae, L39e, L40e, L41e, L44e

A/E: S3Ae, S5e, S8e, S17e, S19e, S24e, S25e, S26e, S27e, S27Ae, S28e, S30e, A, LX

E: L6e, L18Ae, L22e, L27e, L28e, L29e, L36e, L38e

E: S7e, S10e, S12e, S21e

Androgen-independent prostate cancer (continued)

16. Return to the **Pairwise Analysis Results Page** and click **Scatter Plot**.

17. This will bring up a scatter plot of the results. Upregulated genes are shown in red, while downregulated genes are green. The gray spots are those genes that did not pass the analysis criteria (statistics, fold change, etc.). Drag the blue box to an area, and click **Zoom** to see more detail.

18. Click on data points in the detail to bring up the One-Click Gene Summary for a specific gene.

Only a few specific aspects of the data set have been demonstrated here. Feel free to examine the data further on your own.

