Human OneArray Plus

Gene Expression Profiling Service Report

DEMO REPORT

Customer Information

Customer Name:	Phalanx Biotech
Institute:	Phalanx Biotech Group
Telephone:	886-3-5781168
Address:	6 Technology Road 5, 6th Floor Hsinchu Science Park Hsinchu 30078, Taiwan
E-mail:	twinfo@phalanxbiotech.com
PO Number:	224-2014000001
Service Code:	1H123456789
Date:	2014/06/05

I. Overview

Service Content

Item	Requisition	Note
Total RNA Sample Number	3	
Target Preparation (aRNA)	3	
Single / Dual Channel	Single	
Array Type	Human	
Array Version	HOA 6.1	
Total Array Amount	6	
Hybridization Amount	5 μg Cy5-labeled aRNA	
Hybridization Protocol	Phalanx OneArray Plus Protocol	
Scan Protocol	Agilent 0.1 XDR Protocol	

Experimental Design

Item	Control	versus	Treat
1	Control-1	->	Treatment-1
2	Control-1	->	Treatment-2

DEMO REPORT

Report Content

Final service report	Overview	Service Content
		Experimental Design
		Report Content
		Hybridization Result
	Results	Scatter Plot, Histogram and Volcano Plot
		Number of Differentially Expressed Genes
		Array QA: Clustering and PCA Analysis
		Pathway and Gene Ontology Analysis
	Materials and Methods	RNA QC Information
		Target Preparation and Hybridization
		Array Quality Control Result
		Data Analysis Overview
Data package	Service Report	1H123456789_Service Report.pdf
	Data	1H123456789_Data.xls
		1H123456789_Quality Control Result.xls
	Enrichment_analysis	Pathway Analysis
		KEGG pathway map
		Gene Ontology Analysis
	Annotation HOA6.1 release 2.0, 2013-02-01	Annotation is based on two databases: NCBI RefSeq release 57. Ensembl release 70 cDNA sequences and homo_sapiens_core_core_70_37.
	RNA QC	Agilent Bioanalyzer Results
	Array QA	Clustering and PCA Analysis
	GPR	Signal Segmented Files
	Image	Image Files

DEMO REPORT

Hybridization Result: Image (Tiff file) and Signal Segmented File (GPR file)

Item	Sample	Sample Label	Tiff File Name	GPR File Name	Chip Label
1	Control-1	C1	H001-1000000001.tif	H001-1000000001.gpr	C1_H001
2	Control-1	C1	H002-1000000002.tif	H002-1000000002.gpr	C1_H002
3	Treatment-1	T1	H003-1000000003.tif	H003-1000000003.gpr	T1_H003
4	Treatment-1	T1	H004-1000000004.tif	H004-1000000004.gpr	T1_H004
5	Treatment-2	T2	H005-1000000005.tif	H005-1000000005.gpr	T2_H005
6	Treatment-2	T2	H006-1000000006.tif	H006-1000000006.gpr	T2_H006

DEMO REPORT

II. Results

Scatter Plot

Repeatability of expression signal between technical repeats. Figures bellow contain pair wise scatter plots for each sample, including both the raw intensity (R) and normalized intensity (N) data. Only probes with P value (detected) less than 0.05 were included.

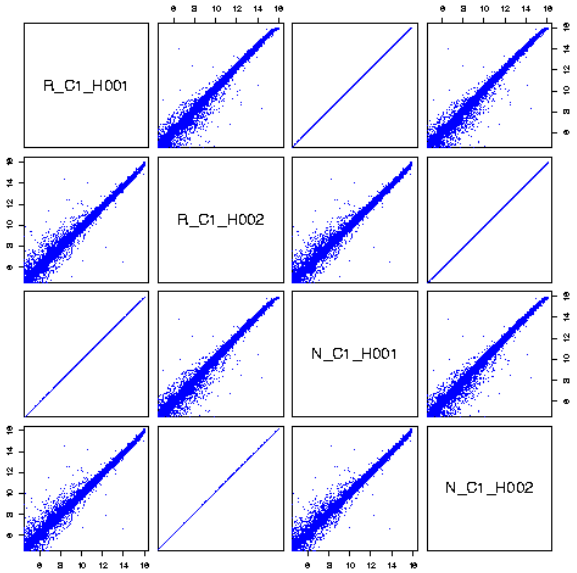

Fig 1. The figures from left to right were labeled as follows : raw data of C1_H001~C1_H002, normalized data of M_H001~M_H002.

DEMO REPORT

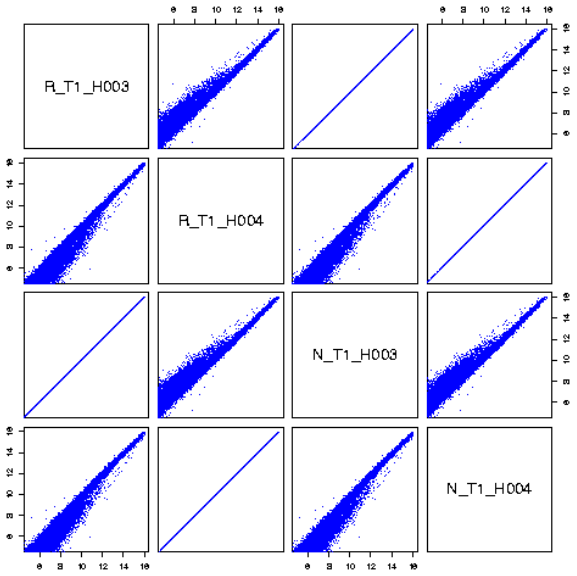

Fig 2. The figures from left to right were labeled as follows : raw data of T1_H003~T1_H004, normalized data of MIL4_H003~MIL4_H004.

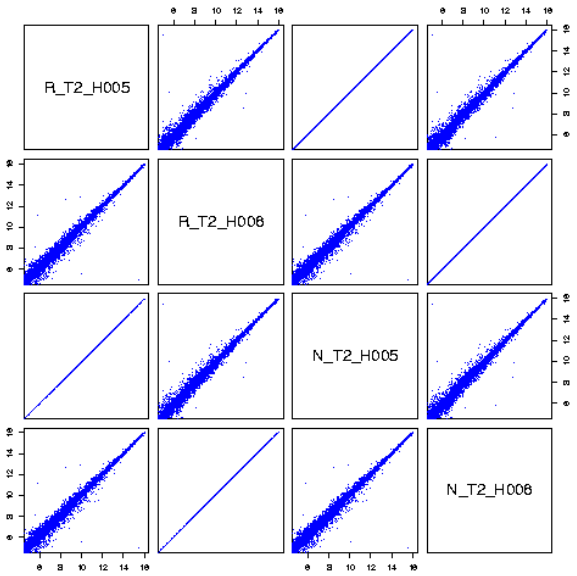

Fig 3. The figures from left to right were labeled as follows : raw data of T2_H005~T2_H006, normalized data of MIL12_H005~MIL12_H006.

DEMO REPORT

Histogram Plot

The histogram plot shows the fold change distribution of all probes excluding control and flagged probes. Fold changes were calculated by Rosetta Resolver 7.2 with error model adjusted by Amersham Pairwise Ration Builder for signal comparison of sample.

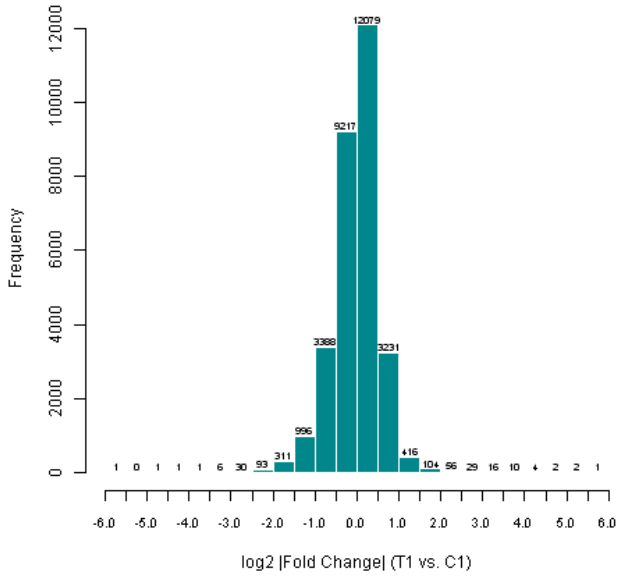

Fig 4. Histogram of log2 | fold change | (T1 versus C1).

DEMO REPORT

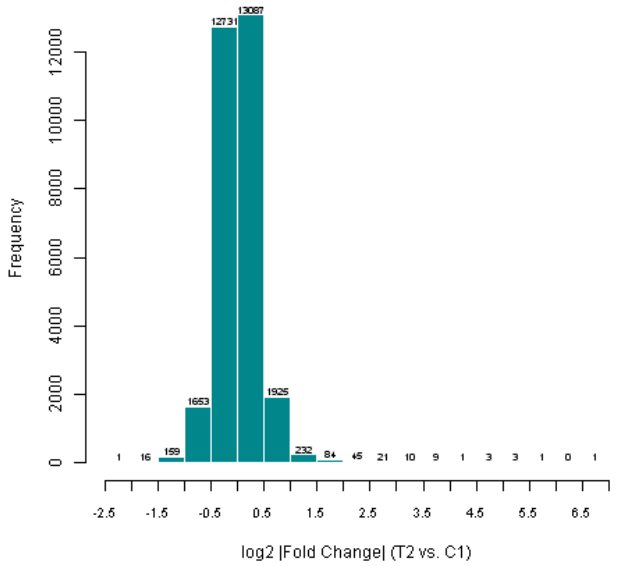

Fig 5. Histogram of log2 | fold change | (T2 versus C1).

DEMO REPORT

Volcano Plot

The volcano plot shows the distribution of differentially expressed probes according to fold-change (x-axis) and significance (negative logarithm of the P-value on the y-axis). The red dotted line is the P-value cut-off (0.05), and the green dotted line is the fold change cut-off (log2 | fold change | ≧ 1). Expression data are plotted for all probes excluding control and flagged probes.

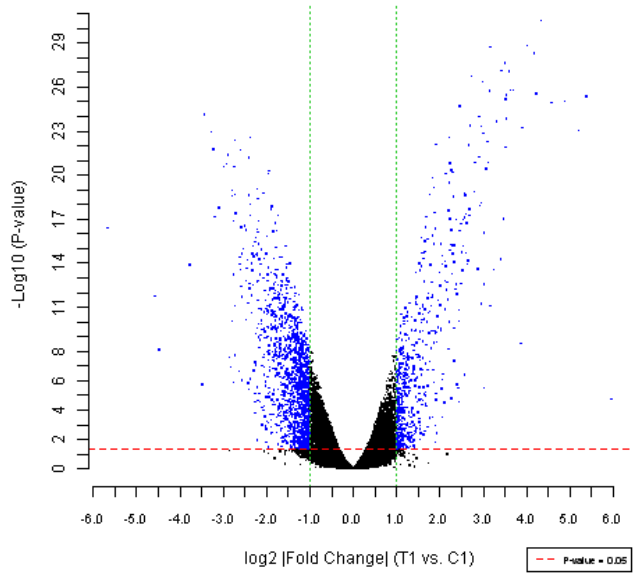

Fig 6. The volcano plot of sample T1 versus C1. Standard selection criteria to identify differentially expressed genes are established at log2 | fold change | ≧ 1 and P-value < 0.05 (Blue dots in figure).

DEMO REPORT

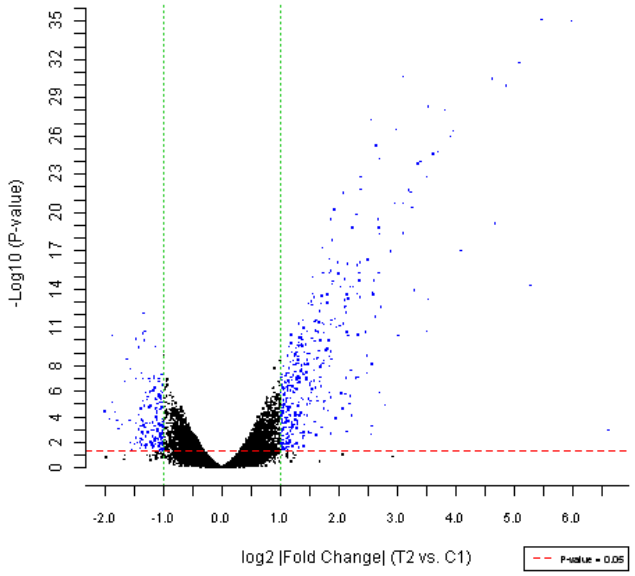

Fig 7. The volcano plot of sample T2 versus C1. Standard selection criteria to identify differentially expressed genes are established at log2 | fold change | ≧ 1 and P-value < 0.05 (Blue dots in figure).

DEMO REPORT

Number of Differentially Expressed Genes

Based on the customer's experimental design, the number of differentially expressed genes for each comparison is shown in the table below. Standard selection criteria to identify differentially expressed genes are as follows: (1) log2 |Fold change| ≥1 and P < 0.05 (2) log2 ratios="NA" and the differences of intensity between the two samples ≥1000. *Detailed gene lists can be found in the following file: 1H123456789_Data.xls.*

Item	Comparison	up-regulated	down-regulated
1	T1/C1	598	1316
2	T2/C1	397	160
Note:			

DEMO REPORT

Array QA : PCA Analysis

Principal Component Analysis (PCA) was performed to evaluate any differences among biological replicates and their treatment conditions. PCA uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components.

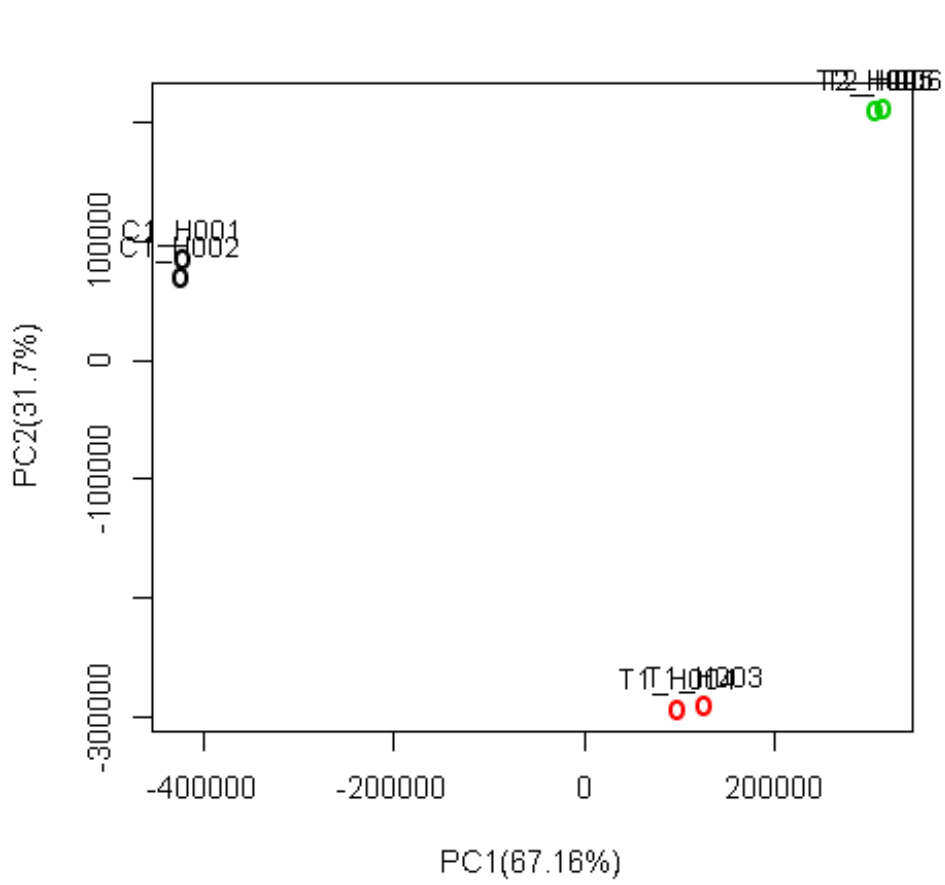

Fig 8. PCA plot

The variable of the first three principal components (PC1, PC2, PC3) for this study are 67.16%, 31.70% and 0.80% respectively. A subset of differentially expressed genes was selected for PCA analysis. A PCA plot based on all genes on the microarray can be found in the following file: *Array QA\PCA_analysis\gPCA.png*.

DEMO

Array QA : Clustering Analysis

For advanced data analysis, intensity data were pooled and calculated to identify differentially expressed genes based on the threshold of fold change and p-value. The correlation of expression profiles between samples and treatment conditions was demonstrated by unsupervised hierarchical clustering analysis. An Excel file containing the sorted expression data for the genes listed in the below heatmap can be found in the following file: *Array QA* \ *Clustering_analysis* \ *clustered_by_order.csv*.

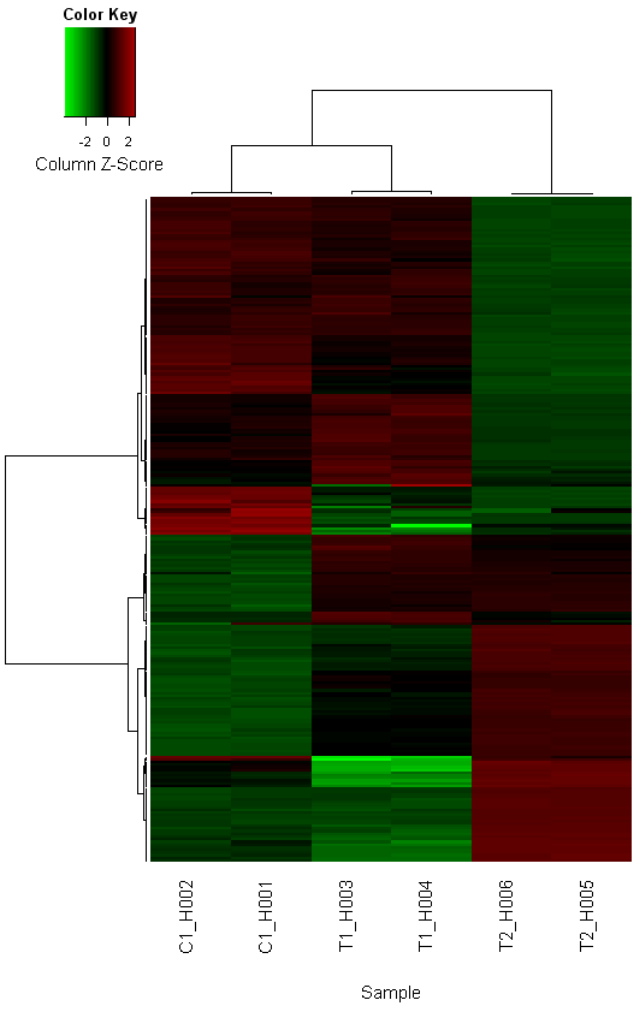

Fig 9. Clustering analysis

Clustering was performed to visualize the correlations among the replicates and varying sample conditions. Up- and down-regulated genes are represented in red and green colors, respectively. A subset of differential genes was selected for clustering analysis. An intensity filter was used for the first 250 genes where the difference between the maximum and minimum intensity values among all microarrays.

Pathway and Gene Ontology Analysis (Sample: T1 / C1)

A gene set enrichment analysis of pathways was performed using the differentially expressed gene lists as input.

For algorithms and detailed results, please review Excel files in the "Enrichment_analysis" folder.

geneset name	#genes in overlap(k)	p value
Cytokine-cytokine receptor interaction	38	1.77e-16
NOD-like receptor signaling pathway	13	1.69e-07
Toll-like receptor signaling pathway	15	1.16e-06
Apoptosis	10	0.0010
RIG-I-like receptor signaling pathway	9	0.0011
Chemokine signaling pathway	15	0.0011
Cytosolic DNA-sensing pathway	7	0.0053
Jak-STAT signaling pathway	12	0.0057
Epithelial cell signaling in Helicobacter pylori infection	7	0.0145
Pathways in cancer	17	0.0312

Table 1. Top 10 enriched pathway terms.

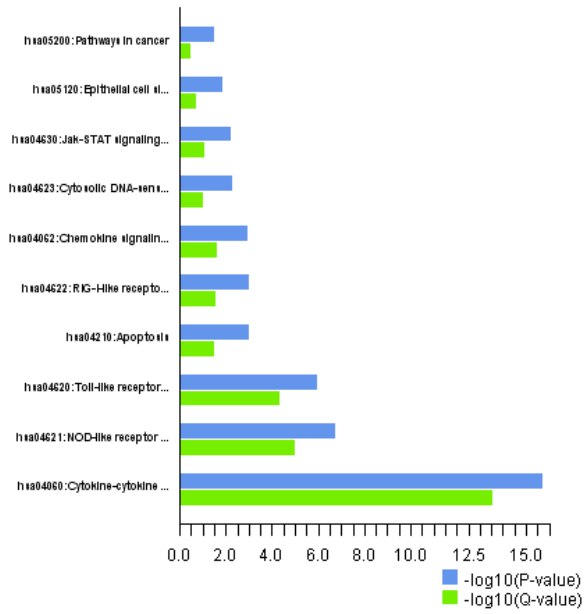

Fig 10. The top 10 enriched pathway terms are plotted on the y-axis versus a measure of significance (negative logarithm of the P-value or Q-value) on the x-axis. The Q-value is calculated by Benjamini.

DEMO REPORT

A gene set enrichment analysis of Gene Ontology (GO) terms was performed using the differentially expressed gene lists as input. *For algorithms and detailed results, please review Excel files in the "Enrichment_analysis" folder.*

geneset name	#genes in overlap(k)	p value
cytokine activity	33	9.62e-19
chemokine activity	12	4.89e-09
chemokine receptor binding	12	1e-08
tumor necrosis factor receptor binding	6	9.12e-05
growth factor activity	13	0.0003
tumor necrosis factor receptor superfamily binding	6	0.0006
carbohydrate binding	19	0.0014
cytokine receptor activity	7	0.0015
pantetheine hydrolase activity	3	0.0015
polysaccharide binding	11	0.0029

Table 2. Enrichment GO terms from GO analysis (molecular function).

geneset name	#genes in overlap(k)	p value
immune response	64	1.33e-20
defense response	58	5.82e-19
inflammatory response	41	1.04e-17
response to wounding	51	5.23e-17
regulation of cell proliferation	53	2.43e-11
positive regulation of multicellular organismal process	28	3.75e-11
positive regulation of cytokine production	17	3.83e-10
response to virus	18	9.02e-10
negative regulation of cell proliferation	31	3.41e-09
regulation of programmed cell death	49	6.29e-09

Table 3. Enrichment GO terms from GO analysis (biological process).

DEMO REPORT

geneset name	#genes in overlap(k)	p value
extracellular space	54	8.54e-15
extracellular region part	62	4.5e-13
extracellular region	83	1.52e-07
intrinsic to plasma membrane	52	2.83e-05
integral to plasma membrane	51	3.19e-05
plasma membrane part	70	0.0059
membrane raft	10	0.0066
cell fraction	39	0.0079
cell surface	16	0.0175
receptor complex	8	0.0194

Table 4. Enrichment GO terms from GO analysis (cellular component).

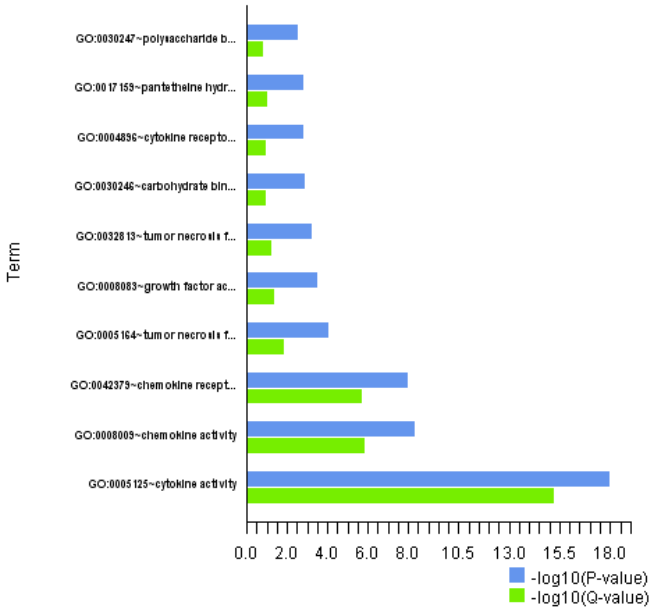

Fig 11. The top 10 enriched GO molecular function terms are plotted on the y-axis versus a measure of significance (negative logarithm of the P-value or Q-value) on the x-axis. The Q-value is calculated by Benjamini.

DRAFT REPORT

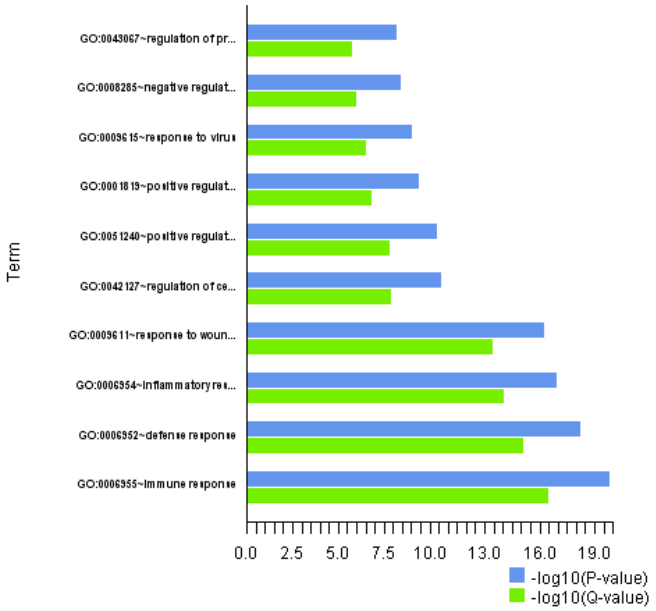

Fig 12. The top 10 enriched GO biological process terms are plotted on the y-axis versus a measure of significance (negative logarithm of the P-value or Q-value) on the x-axis. The Q-value is calculated by Benjamini.

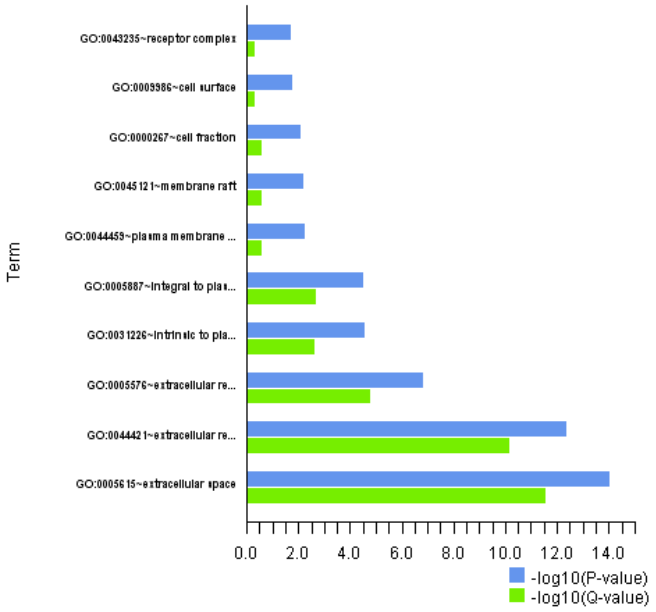

Fig 13. The top 10 enriched GO cellular component terms are plotted on the y-axis versus a measure of significance (negative logarithm of the P-value or Q-value) on the x-axis. The Q-value is calculated by Benjamini.

DEMO REPORT

III. Materials and Methods

RNA QC Information

Item	Sample Name	Sample ID	OD260/280 ≥ 1.8	OD260/230 ≥ 1.5	Total Amount	RIN > 6	QC Result (Pass or Fail)
1	Control-1	1H123456789-01A	2.00	2.00	20.00	10.00	Pass
2	Treatment-1	1H123456789-02A	2.00	2.00	20.00	10.00	Pass
3	Treatment-2	1H123456789-03A	2.00	2.00	20.00	10.00	Pass
Note: Agilent Bioanalyzer results are located in the RNA QC folder in the supplied electronic file.							
Recommended processing for failed samples:							

RNA quantity and purity was assessed using NanoDrop ND-1000. Pass criteria for absorbance ratios are established as A260/A280 ≥ 1.8 and A260/A230 ≥ 1.5 indicating acceptable RNA purity. RIN values are ascertained using Agilent RNA 6000 Nano assay to determine RNA integrity. Pass criteria for RIN value is established at > 6 indicating acceptable RNA integrity. gDNA contamination was evaluated by gel electrophoresis.

DEMO REPORT

Target Preparation and Hybridization

Item	Sample Name	Sample ID	OD260/280 > 1.8	Amplification Yield	Labeling Efficiency > 15	Result
1	Control-1	1H123456789-01A	V	V	V	Pass
2	Treatment-1	1H123456789-02A	V	V	V	Pass
3	Treatment-2	1H123456789-03A	V	V	V	Pass
Note: Labeling Efficiency: # dye molecules/per 1000 nucleotides						
Recommended processing for failed samples:						

Target preparation was performed using an Eberwine-based amplification method with Amino Allyl MessageAmp II aRNA Amplification Kit (Ambion, AM1753) to generate amino-allyl antisense RNA(aa-aRNA). Labeled aRNA coupled with NHS-CyDye was prepared and purified prior to hybridization. Purified coupled aRNA was quantified using NanoDrop ND-1000; pass criteria for CyDye incorporation efficiency at > 15 dye molecular/1000nt.

DEMO REPORT

Array Quality Control Result

Check items	Description	Spec	Result
1. Background average intensity	Cy5: B635 Median	Cy5 < 200	OK
2. Intrinsic Hybridization Controls (IHCs)	IHC probes designed for monitoring overall hybridization efficiency.	IHC > 37500, IHC CV < 15 %	OK
3. Extrinsic Target Quality Control (ETQC) with Spike-ins	ETQC probes and labeled spike-ins designed for monitoring the specificity of hybridization at low, medium and high concentrations.	High, Median, Low	OK
4. Noise	48 negative probes used in Rosetta Resolver for monitoring non-specific hybridization.	Number of negative control probe with intensity>400 should be less than 10.	OK
5. Sample integrity in processing	ITQC 4,5,6 probes designed for monitoring the sample integrity.	Number of ITQC ratios within spec (S1/S2* < 3.5) >= 2	OK
6. Technical reproducibility	Pearson's correlation coefficient between technical replicates.	R value >= 0.975	OK
Note * Two probes per gene were selected from seven consistently expressed housekeeping genes. For each gene, probes were designed from regions of 300-600 bp(S1), 900-1200 bp(S2), respectively, from each 3' end of the transcripts.			

Detail results of array quality control can be found in the file: 1H123456789_Quality Control Result.xls

DEMO REPORT

Data Analysis Overview

GPR files were loaded into Rosetta Resolver® System (Rosetta Biosoftware) to process data analysis.

1. Rosetta profile error model calculation: the error due to random factors and systematic biases are estimated by an error model.
2. Squeeze replicated probes: the repeated probes within one chip are averaged.
3. Normalize intensities: Median scaling performed on data set without flagged and control data.
4. Pearson's correlation coefficient: statistical analysis calculated on technical replicates to assess reproducibility.
5. Merge technical replicate data: Average intensity values calculated on technical replicates.
6. Pairwise ratio calculation: Probe filtering, normalization, pair-wise comparison and error-weighted modeling are performed based on customers' designated sample groups.
7. Differentially expressed gene lists: Standard selection criteria to identify differentially expressed genes are as follows: (1) log2 |Fold change| ≥ 1 and $P < 0.05$. (2) log2 ratios="NA" and the differences of intensity between the two samples >=1000. Additional adjustments available based on customer requirements.
8. PCA and Clustering analysis: Gene clustering by averagely linkage algorithm performed on selected differentially expressed gene lists after data transformation and mean centering.
9. Pathway and GO analysis: Please refer to Excel files and DAVID website for details.
   Reference: http://david.abcc.ncifcrf.gov/

DEMO REPORT

IV. Reference

The Phalanx Annotation Search System (PASS) links Phalanx probe sets of your interest to NCBI, Gene Ontology, Ensembl, and Uniprot websites. To learn more, please visit:

http://bioinformatics.phalanxbiotech.com/webAnnotation.html

DEMO REPORT