Compiled Expression

Compiled expression interface let users to perform microarray gene expression data analysis. It includes a number of data pre-treatment and analysis tools.

After open a compiled file, ExP pop-ups the following interface.

(1)   Data structure – experiment design

First thing to do is to tell ExP your data structure.

(A) Click right button on mouse over a cell, user can assign the data start column and start row.

(B) But prior to opening the compiled file, user may use excel to edit it and add a few rows to specify each of the columns, such as inter and intra groups (biological replicates), technical replicates, dye labels, cell line names, call line profiles, and then save the file into tab-delimited format. To specify the data columns, user may click right button on mouse over a selected row (any cell in the row) to assign their column specs.

(C) Also, user can perform row data plot and open a web page of a given gene either based on GenBank ID or unigene

(2) Data pre-treatment

The pretreatment includes profile-based dye-labeling bias correction and cell line alignment.

(A) Profile-based dye-labeling bias correction.

To perform profile-based dye-labeling bias correction, user needs assign technical replicate and bye-label rows.
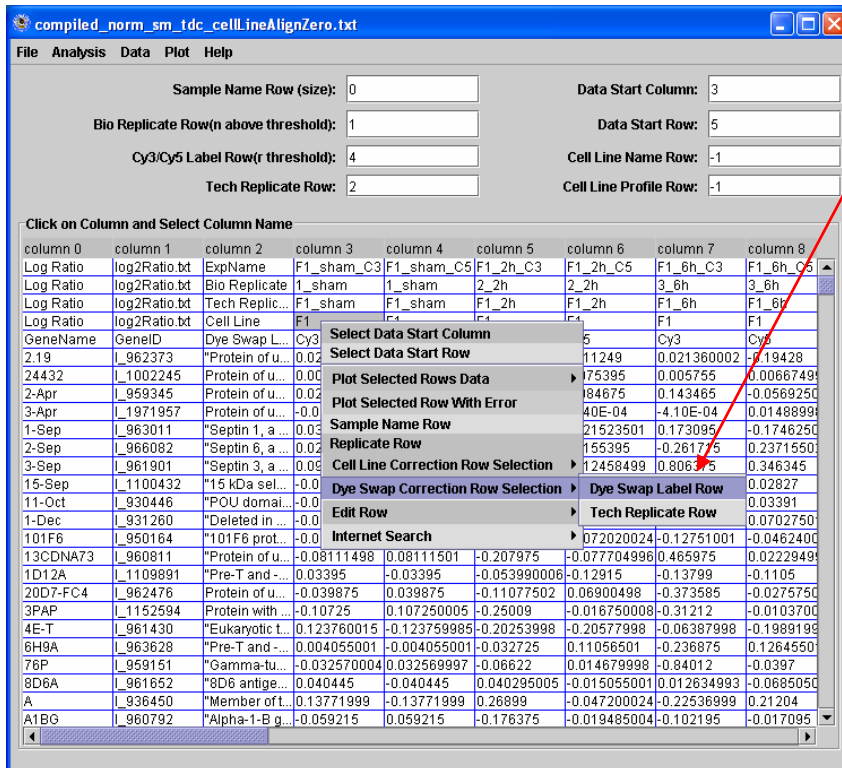
Then, user click mouse over "Dye Bias Correction" to perform the action. The dye-bias corrected result will replace the un-corrected data. User may save the result into a file.
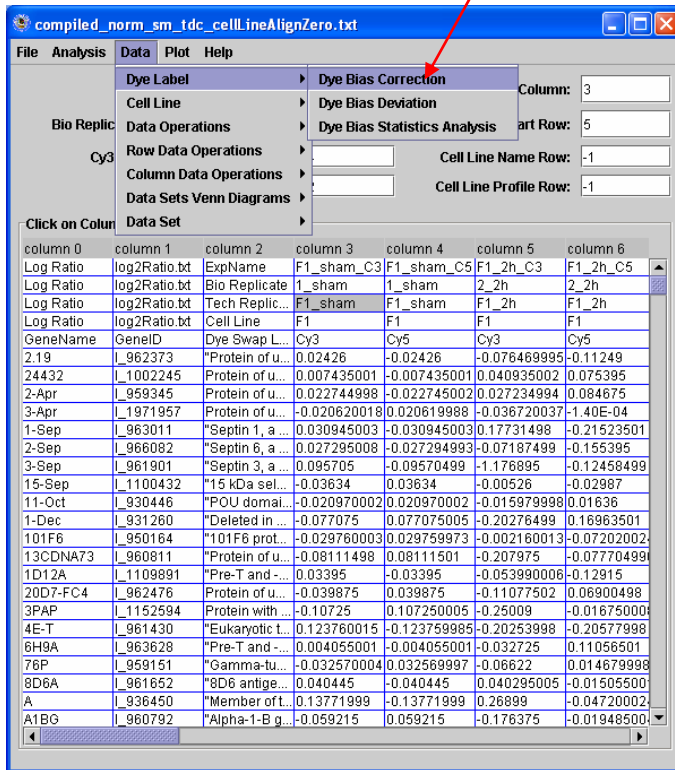
(B) Cell line alignment

If the compiled data contains multiple cell line and each cell line has a sham-treated set of data, user may apply cell line alignment. To perform cell line alignment, user needs assign cell line and cell line profile rows.

Then, user may click mouse over "Cell Line Alignment to zero" to align the averaged sham-treated state to zero. Also, user may click "Cell Line Alignment to average" to align the averaged sham-treated state to their grand average. The alignment result will replace the previous un-alignment data. User may save the result into a file.

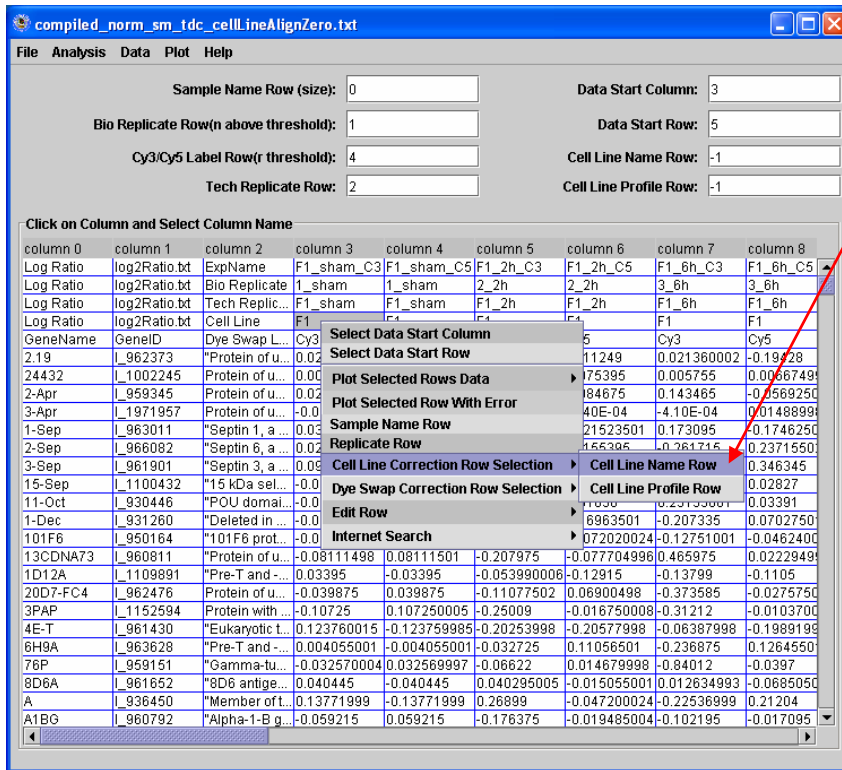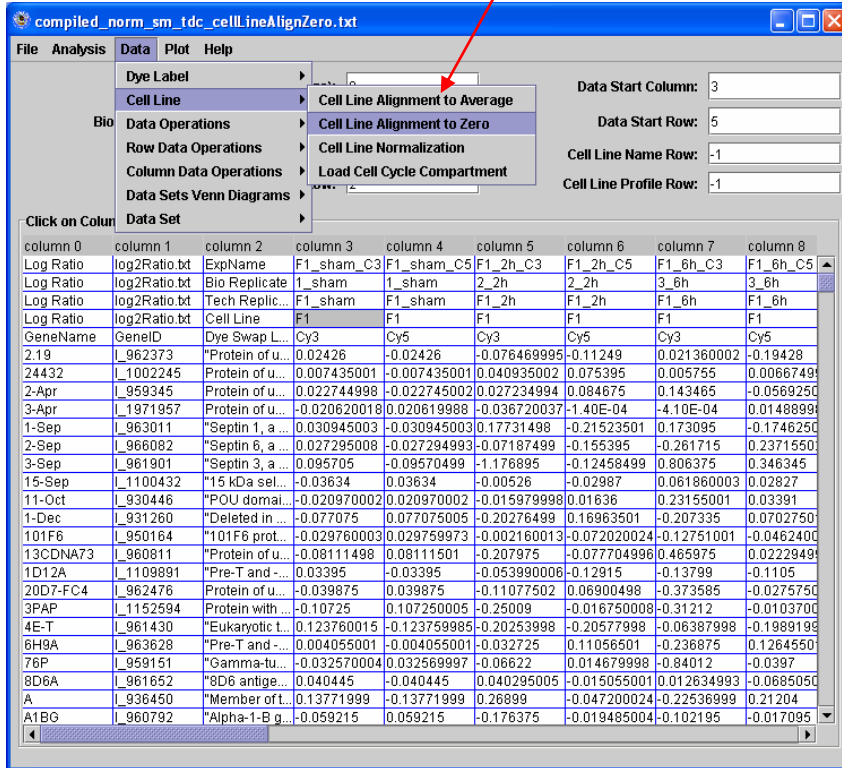compiled_norm_sm_tdc_cellLineAlignZero.txt

File   Analysis   Data   Plot   Help

Dye Label
Cell Line   ►   Cell Line Alignment to Average
Data Operations   ►   Cell Line Alignment to Zero
Row Data Operations   ►   Cell Line Normalization
Column Data Operations   ►   Load Cell Cycle Compartment
Data Sets Venn Diagrams   ►
Data Set   ►

Data Start Column: 3
Data Start Row: 5
Cell Line Name Row: -1
Cell Line Profile Row: -1

| column 0	column 1	column 2	column 3	column 4	column 5	column 6	column 7	column 8
Log Ratio	log2Ratio.txt	ExpName	F1_sham_C3	F1_sham_C5	F1_2h_C3	F1_2h_C5	F1_6h_C3	F1_6h_C5
Log Ratio	log2Ratio.txt	Bio Replicate	1_sham	1_sham	2_2h	2_2h	3_6h	3_6h
Log Ratio	log2Ratio.txt	Tech Replic...	F1_sham	F1_sham	F1_2h	F1_2h	F1_6h	F1_6h
Log Ratio	log2Ratio.txt	Cell Line	F1	F1	F1	F1	F1	F1
GeneName	GeneID	Dye Swap L...	Cy3	Cy5	Cy3	Cy5	Cy3	Cy5
2.19	I_962373	"Protein of u...	0.02426	-0.02426	-0.076469995	-0.11249	0.021360002	-0.19428
24432	I_1002245	Protein of u...	0.007435001	-0.007435001	0.040935002	0.075395	0.005755	0.0066749
2-Apr	I_959345	Protein of u...	0.022744998	-0.022745002	0.027234994	0.084675	0.143465	-0.0569250
3-Apr	I_1971957	Protein of u...	-0.020620018	0.020619988	-0.036720037	-1.40E-04	-4.10E-04	0.0148899
1-Sep	I_963011	"Septin 1, a ...	0.030945003	-0.030945003	0.17731498	-0.21523501	0.173095	-0.1746250
2-Sep	I_966082	"Septin 6, a ...	0.027295008	-0.027294993	-0.07187499	-0.155395	-0.261715	0.2371550
3-Sep	I_961901	"Septin 3, a ...	0.095705	-0.09570499	-1.176895	-0.12458499	0.806375	0.346345
15-Sep	I_1100432	"15 kDa sel...	-0.03634	0.03634	-0.00526	-0.02987	0.061860003	0.02827
11-Oct	I_930446	"POU domai...	-0.020970002	0.020970002	-0.015979998	0.01636	0.23155001	0.03391
1-Dec	I_931260	"Deleted in ...	-0.077075	0.077075005	-0.20276499	0.16963501	-0.207335	0.0702750
101F6	I_950164	"101F6 prot...	-0.029760003	0.029759973	-0.002160013	-0.072020024	-0.12751001	-0.0462400
13CDNA73	I_960811	"Protein of u...	-0.08111498	0.08111501	-0.207975	-0.077704996	0.465975	0.0222949
1D12A	I_1109891	"Pre-T and -...	0.03395	-0.03395	-0.053990006	-0.12915	-0.13799	-0.1105
20D7-FC4	I_962476	Protein of u...	-0.039875	0.039875	-0.11077502	0.06900498	-0.373585	-0.0275750
3PAP	I_1152594	Protein with ...	-0.10725	0.107250005	-0.25009	-0.016750008	-0.31212	-0.0103700
4E-T	I_961430	"Eukaryotic t...	0.123760015	-0.123759985	-0.20253998	-0.20577998	-0.06387998	-0.1989199
6H9A	I_963628	"Pre-T and -...	0.004055001	-0.004055001	-0.032725	0.11056501	-0.236875	0.1264550
76P	I_959151	"Gamma-tu...	-0.032570004	0.032569997	-0.06622	0.014679998	-0.84012	-0.0397
8D6A	I_961652	"8D6 antige...	0.040445	-0.040445	0.040295005	-0.015055001	0.012634993	-0.0685050
A	I_936450	"Member of t...	0.13771999	-0.13771999	0.26899	-0.047200024	-0.22536999	0.21204
A1BG	I_960792	"Alpha-1-B g...	-0.059215	0.059215	-0.176375	-0.019485004	-0.102195	-0.017095

Bio

Click on Colum

(C) Other data treatment
Mouse right click over a column name, user may perform a number of conversions to the data over selected column. It includes

(i)   Exp Base 2 conversion – d to 2^d;
(ii)   Log2 conversion – d to $\log_2 d$;
(iii) Minus operation – d to –d.

Also included

(iv)   String Find;
(v)   Data merge over same genes;
(vi)   Sorting
(vii) Load Data Conjunction;
(viii) Load Data Neg. Conjunction;
(ix)   Remove Color Marked Rows;
(x)   Remove Black Marked Rows;
(xi)   Mark valid rows (the row data does not contain zero);
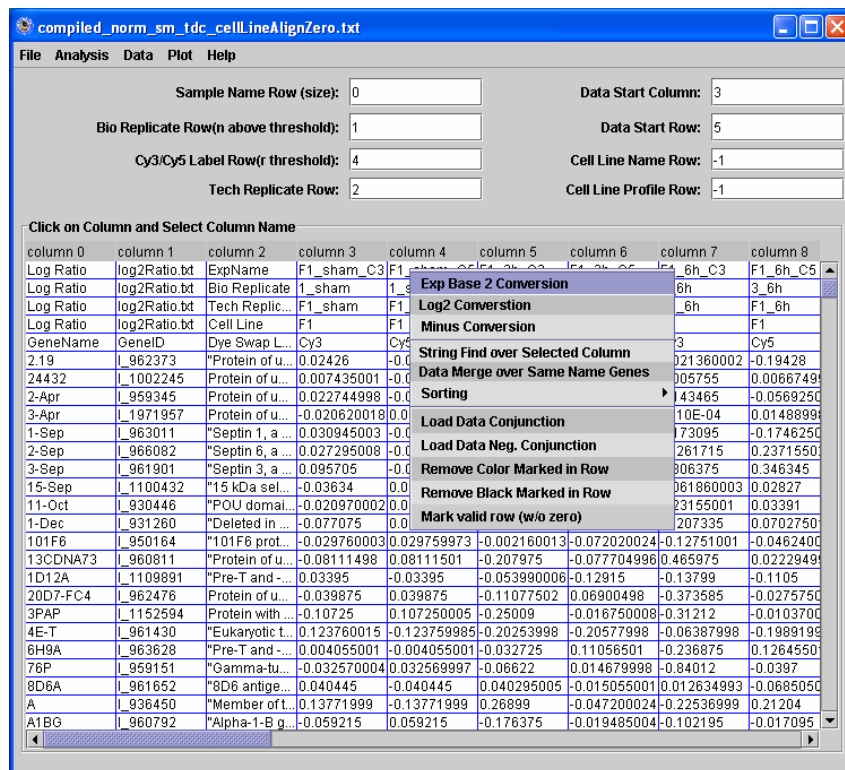
These operations may apply to all the columns, which includes
- (i)   Exp Base 2 conversion – d to 2^d;
- (ii)   Log2 conversion – d to $\log_2 d$;
- (iii)   Minus operation – d to –d.

(3) Statistical analysis

ExP let user perform a number of statistical analysis, which includes

(A) ANOVA

(B) t-test

(C) paired t-test

(D) F-test

(E) Profile signal-to-noise.

To apply all these analysis, the data must consist of a set of biological replicates, i.e. inter- and intra-group assignment.
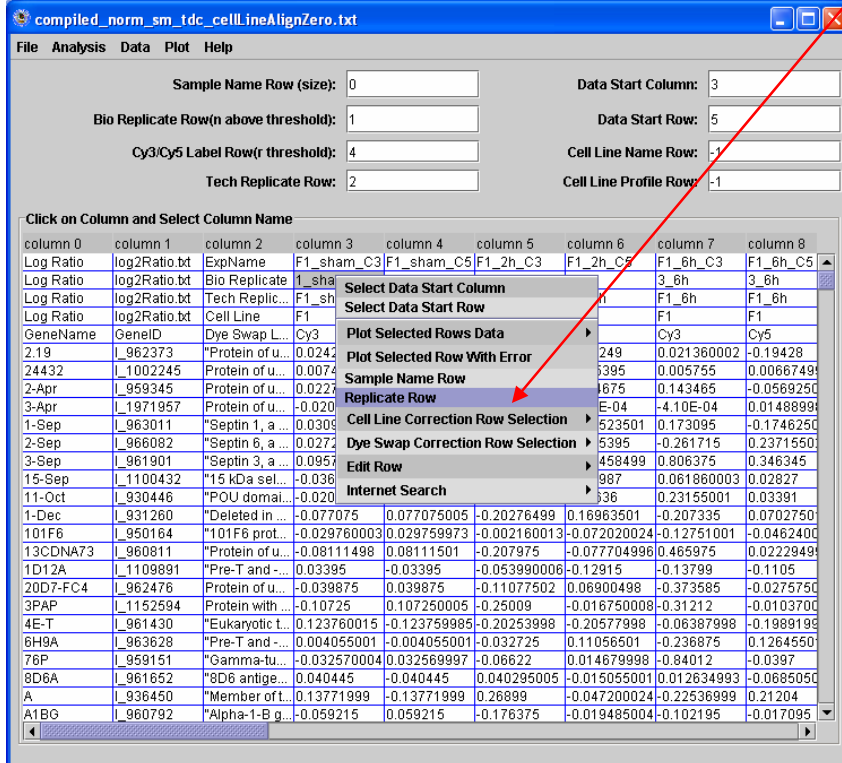
Each of these statistical analysis methods will have an output to save resulting data and their FDR estimate.

(4) High-dimensional data analysis

ExP provides a number of high-dimensional data analysis, which includes

(A)	Principal Component Analysis (PCA)
(B)	Heap Map
(C)	Hierarchical Clustering
(D)	Fuzzy ARTMAP
(E)	Relevance Analysis
(F)	Linear Discrimination Analysis
(G)	Bayesian Classification Analysis
(H)	Projection Classification
(I)	Extraction of pattern and identification of genes - EPIG