

Estudi exemple

Plataforma Bioinformatica CIBEREHD

October 2, 2009

Real-time RT-PCR is an established technique for quantifying mRNA in biological samples, this technique enables carry out two different methods of quantification gene expression based on the porpouse of the study: absolut quantification and relative quantification.

C_T package is designed for relative quantification analysis of expression pattern data measured by Real-time RT-PCR employing the ABI 7900 Sequence Detection System, and this document provides some tables and figures to show the results obtained from the analysis. To analyse Real-time RT-PCR raw data is used the C_T values; C_T is the cycle number at which the fluorecence crosses the detection threshold.

First of all, to allow comparison between the individual samples, as well as between the histological subgroups, normalization was performed. Selection of the targets for normalization is carry out by Vandesompele method. The coherence of the endogenous Ct signal is demonstrated by the constant profile of the ratio of both Ct values showed in Fig1. Subsequently, a normalization step was done employing different methods depending on the study, ($2^{-\Delta C_T}$ or $2^{-\Delta\Delta C_T}$ computation). Normalized data was analyzed employing parametric tests: ANOVA, Kruskal-Wallis and Standard 2-class unpaired t test. Thereafter, hierarchical clustering was applied to all normalized data columns and the pearson correlation distance was used as an input parameter in the clustering algorithm, which favours the grouping of probes showing similar diferential expression values across samples. Also, a barplot and boxplot of genes who present a significant diferencial expression ($p \leq 0.05$) are shown.

Data set dimensions: 96 genes vs 37 samples.

Removed genes:

No genes removed

Housekeepings:

TUBA1B-Hs00744842_sH

UBA3-Hs00234368_m1

The best housekeepings for the analysis are identified by Vandesompele method. The results obtained are:

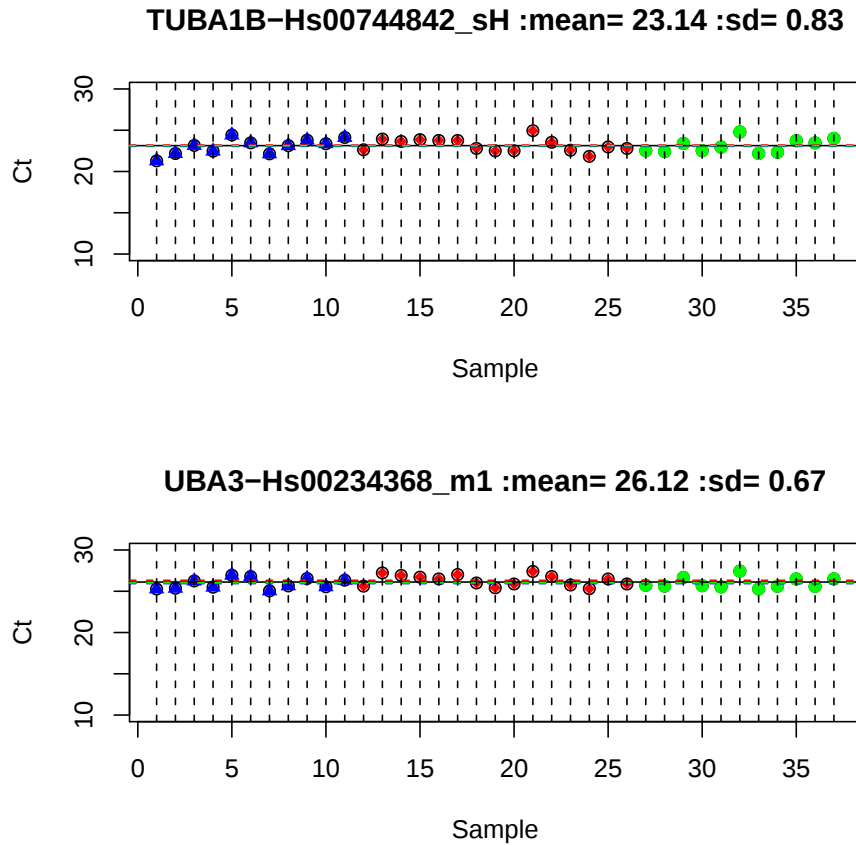


Figure 1: Housekeeping dotplot. Representation of endogenous Ct signals, each sample group is represented for a specific color. This graph show the constant profile of the ratio of both Ct values.

In this study efficiency data isn't disponible.
So it has been accepted an efficiency value of 2 (the optimal value
indicating an increasing of 2-folds per cycle).

gene	efficiency	avg Ct	min Ct	max Ct	IQR
18S-Hs99999901_s1	2	7.7	6.81	9.3	0.91
ABCB11-Hs00184824_m1	2	30.27	26.28	33.88	1.48
ABCG2-Hs01053787_m1	2	27.38	23.49	31.47	2.95
ACOX1-Hs01074241_m1	2	25.33	22.67	28.43	1.41
ACSL4-Hs01547083_m1	2	26.52	24.18	28.44	1.07
ACTB-Hs03023880_g1	2	18.41	16.68	40	1.04
ADH6-Hs00167423_m1	2	28.7	26.4	34.14	1.82
AQP3-Hs00185020_m1	2	26.49	24.62	28.75	1.68
AQP8-Hs01086280_g1	2	24.52	18.63	40	5.95
BMP4-Hs00370078_m1	2	27.24	24.95	29.59	0.88
CCL18-Hs00268113_m1	2	27.81	23.05	40	2.6
CCL20-Hs01011368_m1	2	25.86	20.65	40	2.08
CCL23-Hs00270756_m1	2	29.44	26.25	34.28	2.36
CCL8-Hs00428422_g1	2	29.45	25.78	40	2.92
CD180-Hs01069872_m1	2	27.48	25.47	29.89	1.59
CDC42-Hs00741586_mH	2	31.78	28.27	33.29	1.3
CHI3L1-Hs01072228_m1	2	27.95	22.79	33.3	4.78
COL1A2-Hs01028971_m1	2	25.18	22.55	27.93	1.97
CPB1-Hs00157026_m1	2	30.64	24.05	40	5.58
CREB1-Hs00231713_m1	2	26.42	25.08	40	0.91
CTNNA1-Hs00944792_mH	2	22.34	21.36	24.62	1.21
CXCL1-Hs00605382_gH	2	24.64	20.28	28.55	3.69
CXCL3-Hs00171061_m1	2	25.52	21.76	29.25	2.65
CXCL5-Hs00607029_g1	2	28.19	12.41	40	6.12
DEFB1-Hs00608345_m1	2	27.1	23.78	31.23	1.92
DEFB4-Hs00175474_m1	2	32.52	21.93	40	12.53
DSC2-Hs00951428_m1	2	23.05	21.55	26.09	1.24
DUSP1-Hs00610257_g1	2	24.6	20.94	27.15	1.8
EGFR-Hs01076093_g1	2	25.91	24.6	28.56	0.9
GJA1-Hs00748445_s1	2	25.3	22.99	27.24	1.33
GPR128-Hs00262184_m1	2	28.6	25.93	31.91	2.78
GREM2-Hs01934663_s1	2	28.93	26.99	31.08	1.28
GZMA-Hs00989184_m1	2	27.99	24.9	30.57	1.91
IGFBP2-Hs01040719_m1	2	23.61	21.46	27.55	1.95
IGHM-Hs00378230_g1	2	19.98	15.07	23.43	4.11
IL1B-Hs01555413_m1	2	27.93	22.36	32.02	3.55
IL1R2-Hs00174759_m1	2	25.73	23.15	28.78	1.38

IL1RN-Hs00893626_m1	2	27.41	22.67	33.36	2.92
IL6R-Hs00794121_m1	2	26.56	24.36	27.86	1.09
IL7R-Hs00902338_g1	2	27.77	25.34	30.38	1.51
IL8-Hs01567913_g1	2	29.25	21.29	40	4.37
INSL5-Hs00193884_m1	2	27.38	22.51	40	5.03
IRF6-Hs00196213_m1	2	24.21	22.76	27.9	1.11
ITGB6-Hs00168458_m1	2	27.03	23.72	29.5	1.24
JAK1-Hs00233820_m1	2	24.27	23.34	25.69	0.96
KRAS-Hs00270666_m1	2	25.52	24.33	27.52	1.05
LILRB1-Hs01848117_s1	2	28.59	26.92	30.85	1.15
LILRB2-Hs00275975_m1	2	29.51	26.03	31.69	1.56
MAP2K6-Hs00177150_m1	2	26.74	24.12	40	2.17
MBP-Hs00922788_m1	2	33.09	31.54	40	1.26
ME1-Hs01554892_m1	2	26.26	23.92	40	2.04
MEP1B-Hs00195535_m1	2	29.2	23.08	33.84	3.31
MMP1-Hs00899660_g1	2	27.96	21.43	31.2	2.29
MMP10-Hs00233987_m1	2	29.52	23.26	34.21	3.04
MMP3-Hs00968308_m1	2	28.77	21.11	35.86	3.85
MUC20-Hs00416321_m1	2	25.67	23.74	28.6	1.21
MUC4-Hs00366414_m1	2	24.74	22.36	27.79	1.65
MUC5AC;MUC5B-Hs00861588_m1	2	24.69	22.83	27.3	1.78
NFKBIZ-Hs00944732_m1	2	25.58	22.75	27.86	1.32
OSTalpha-Hs00380895_m1	2	29.72	23.42	35.34	4.72
PPAP2B-Hs00170359_m1	2	26.35	24.81	27.43	0.82
PPIC-Hs00181460_m1	2	25.19	23.6	40	1.19
PRAP1-Hs00376974_m1	2	26.19	21.97	30.53	3.82
PTGER4-Hs00168761_m1	2	26.56	24.82	28.59	1.45
PYY-Hs00373890_g1	2	23.67	20.82	29.11	2.07
REG1A-Hs00984887_g1	2	25.78	16.88	40	8.1
REG4-Hs01069973_m1	2	21.86	18.15	26.19	2.7
RHOB-Hs00269660_s1	2	26.32	22.34	40	1.55
RPLP0-like-Hs00420895_gH	2	20.52	18.97	22.41	1.01
RPS2-Hs03005117_g1	2	18.42	17.3	21.01	0.91
RPS20-Hs00828752_gH	2	21.91	20.28	40	1.15
RPS3A-Hs00832893_sH	2	26.49	25.38	28.02	1.08
RUNDC3B-Hs00289927_m1	2	30.1	27.53	32.2	1.57
S100P-Hs00195584_m1	2	22.67	20.22	26.16	2.03
SERPINB5-Hs00985283_m1	2	26.84	22.16	31.44	2.89
SLC16A1-Hs00161826_m1	2	24.59	21.11	27.62	2.11
SMAD7-Hs00998193_m1	2	26	24.16	27.64	1.19
ST6GAL2-Hs00383641_m1	2	29.75	24.73	35.05	5.38
TAPBP-Hs00542606_m1	2	23	20.51	40	1.17

TFF1-Hs00170216_m1	2	23.27	19.35	40	2.68
TFPI2-Hs00197918_m1	2	30.29	24.86	40	3.1
TGFBI-Hs00932734_m1	2	25.24	22.77	27.33	1.43
TGFBR3-Hs01114253_m1	2	28.2	26.2	30.17	0.94
TIMP1-Hs00355335_g1	2	24.04	21.19	26.7	2.17
TJP1-Hs00543824_m1	2	25.44	23.74	28.13	1.18
TLR3-Hs01551078_m1	2	28.3	25.64	31.41	1.89
TOP1-Hs00243257_m1	2	24.25	22.98	26.31	1.16
TOP2A-Hs00172214_m1	2	26.44	24.09	40	1.11
TRIM29-Hs00232590_m1	2	30.21	25.93	33.96	2.35
TSLP-Hs00263639_m1	2	29.31	25.67	40	2.02
TSLP-Hs01572933_m1	2	37.33	29.8	40	5.87
TUBA1B-Hs00744842_sH	2	23.14	21.29	24.94	1.26
UBA3-Hs00234368_m1	2	26.12	25.05	27.45	1.09
VCAN-Hs01007937_m1	2	27.76	25.41	29.95	1.36
VNN1-Hs01546812_m1	2	28.54	23.84	33.44	2.3
WNT5A-Hs00998537_m1	2	28.62	24.96	30.6	1.62

Table 1: Efficiency of each gene in Real-time RT-PCR process

Real-time RT-PCR data has been used to analyze the relative expression of one or more genes of interest to a reference gene and, in most of cases, to a reference sample.

There are different methods to calculate relative gene expression: $2^{-\Delta\Delta C_T}$ method, $2^{-\Delta C_T}$ method and Pfaffl method. The best method for analyse relative gene expression that allow have in account gene efficiency is Pfaffl method, so the method used here to calculate the relative gene expression is a variation of Pfaffl method. That permit use an especific efficiency for each gene and more than one reference sample.

Pfaffl method is defined as:

$$Ratio = \frac{(E_{target})^{\Delta C_{T,target(calibrator-test)}}}{(E_{ref})^{\Delta C_{T,ref(calibrator-test)}}}$$

The mathematical algorithm derived from Pfaffl method used is defined as:

$$Ratio = \frac{\frac{(E_{ref})^{C_{Tsample}}}{(E_{target})^{C_{Tsample}}}}{\frac{1}{n} \sum_{i=1}^n \frac{(E_{ref})^{C_{Tcalibrator_i}}}{(E_{target})^{C_{Tcalibrator_i}}}}$$

After calculation of relative gene expression data, normalized data is analysed by different statistcal method with the aim of identify the significant different expressed genes.

Depending on the number of different groups present in the study the statistical method applied will be: t-test for two groups comparasion or anova, kruskall-Wallis and t.test for more than two groups compartion.

So, in the next pages it's found the results of this analysis showed in different formats:

- Heatmaps:
 - Heatmap of all samples
 - Heatmap of significant genes (genes with p-value is equal or inferior to 0.05)
- Barplots of significant genes (genes with p-value equal or inferior to 0.05) Y axis present the relative expression (ratio) in $\log_2(E^{-\Delta C_T})$ or $\log_2(E^{-\Delta\Delta C_T})$ data, and X axis correspond to different significant genes. Each bar represent one group especified in the legend.
- Boxplot of significant genes (genes with p-value equal or inferior to 0.05) Each plot correspond to a significant gene. Y axis present the relative expression (ratio) in $\log_2(E^{-\Delta C_T})$ or $\log_2(E^{-\Delta\Delta C_T})$ data, and X axis correspond to different sample groups.

All genes

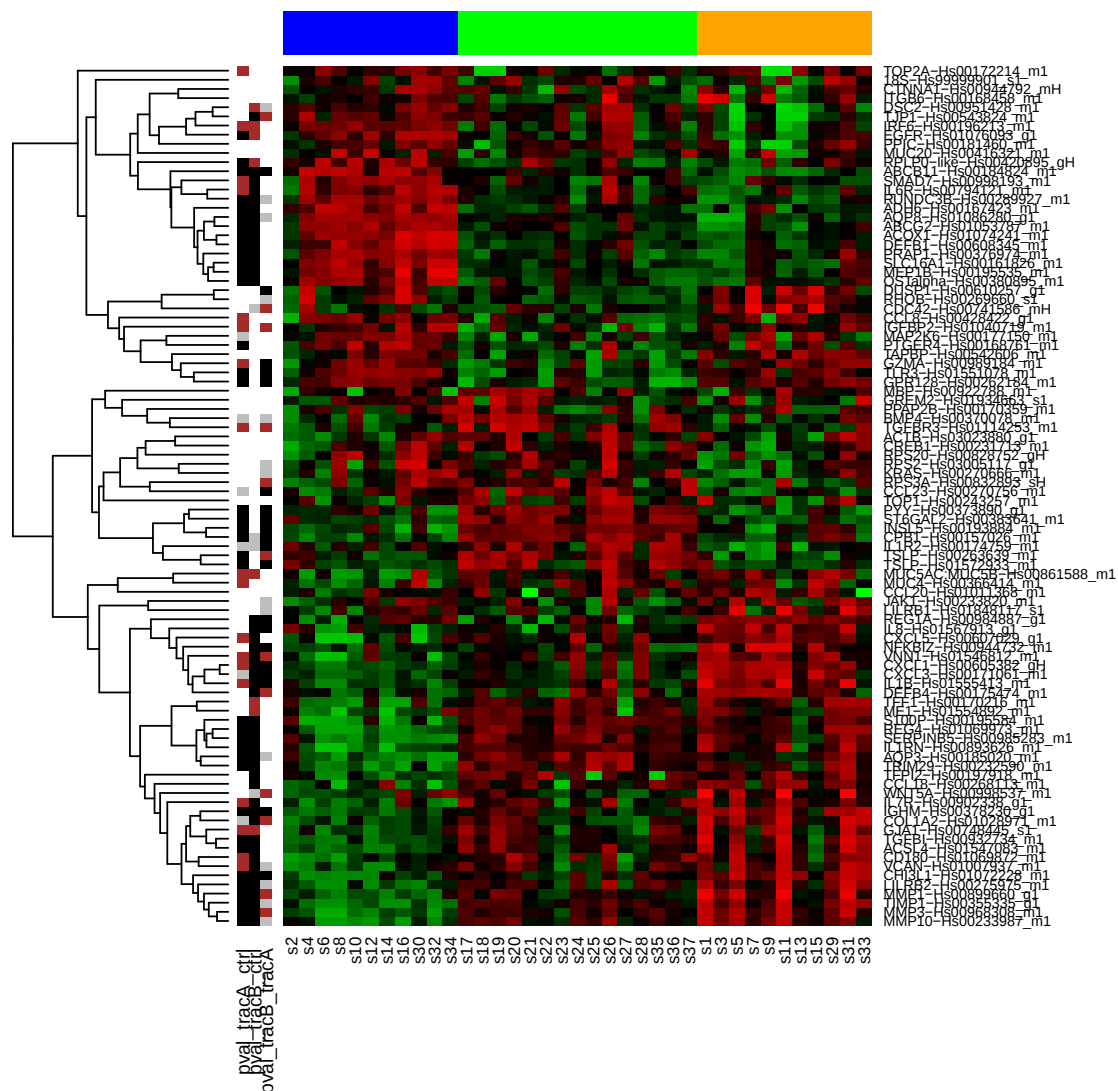
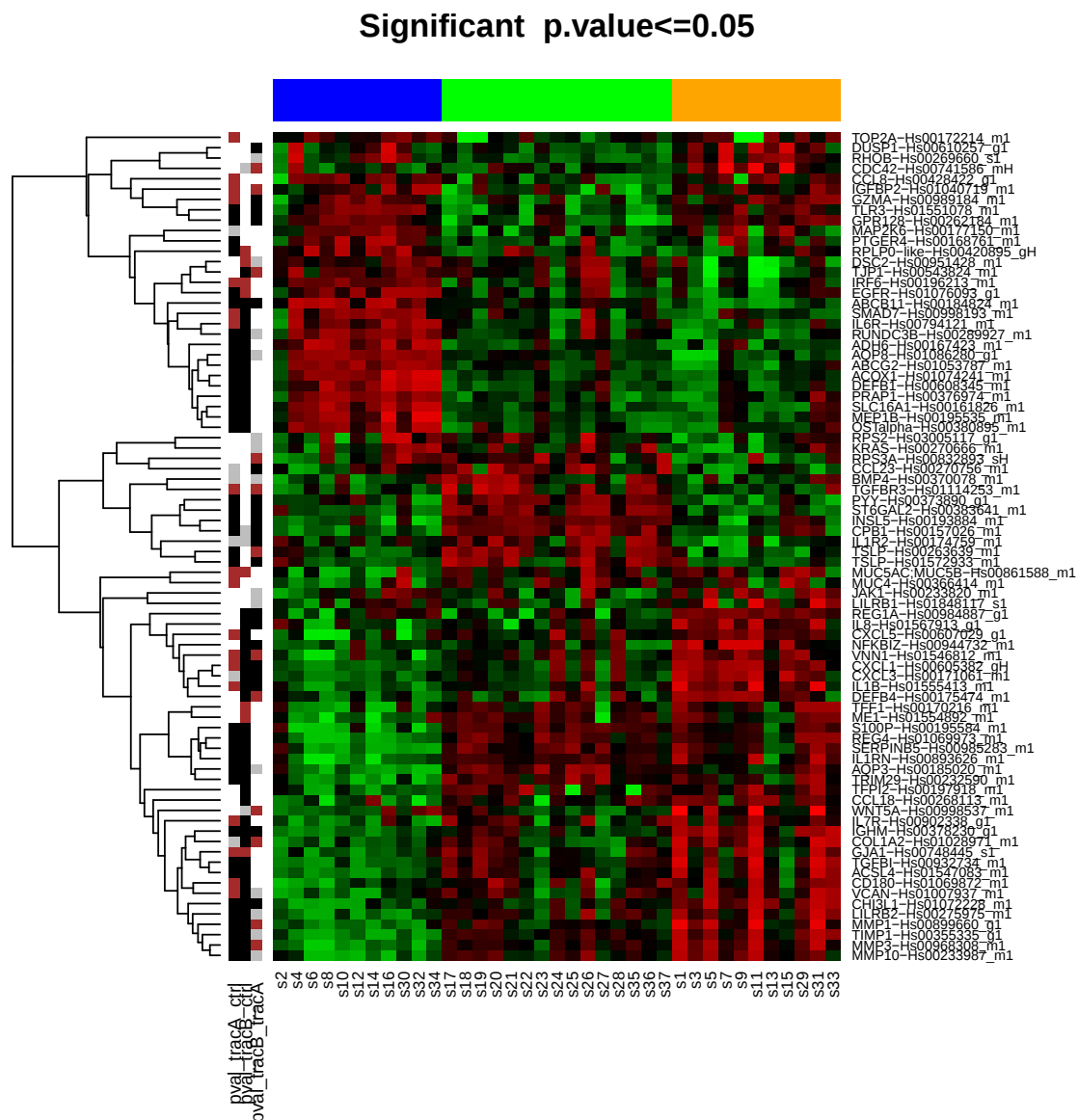
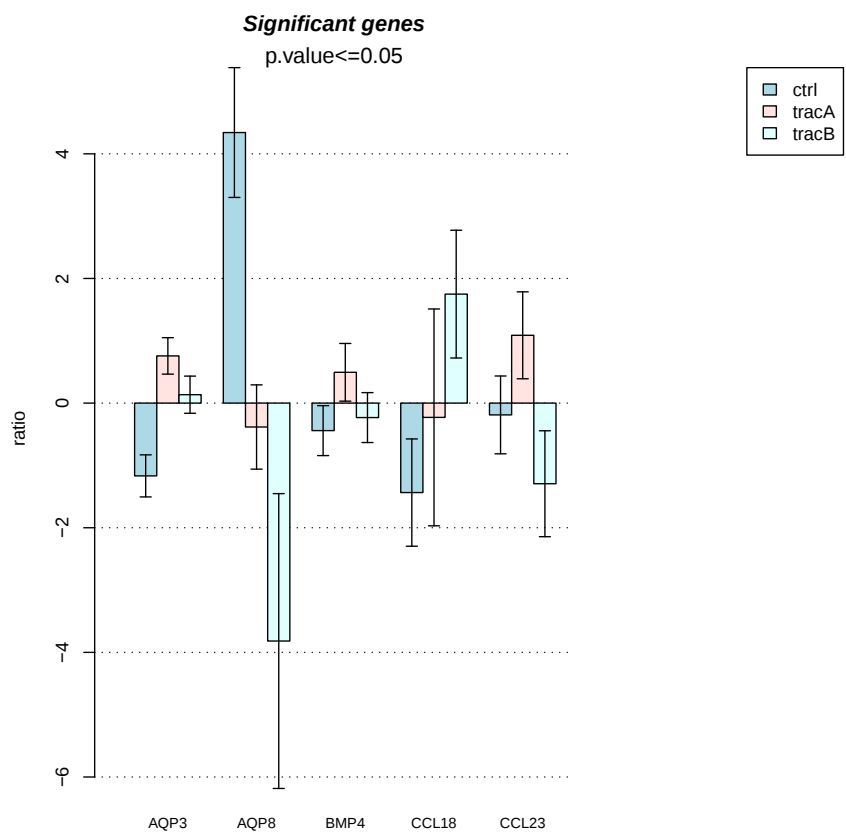
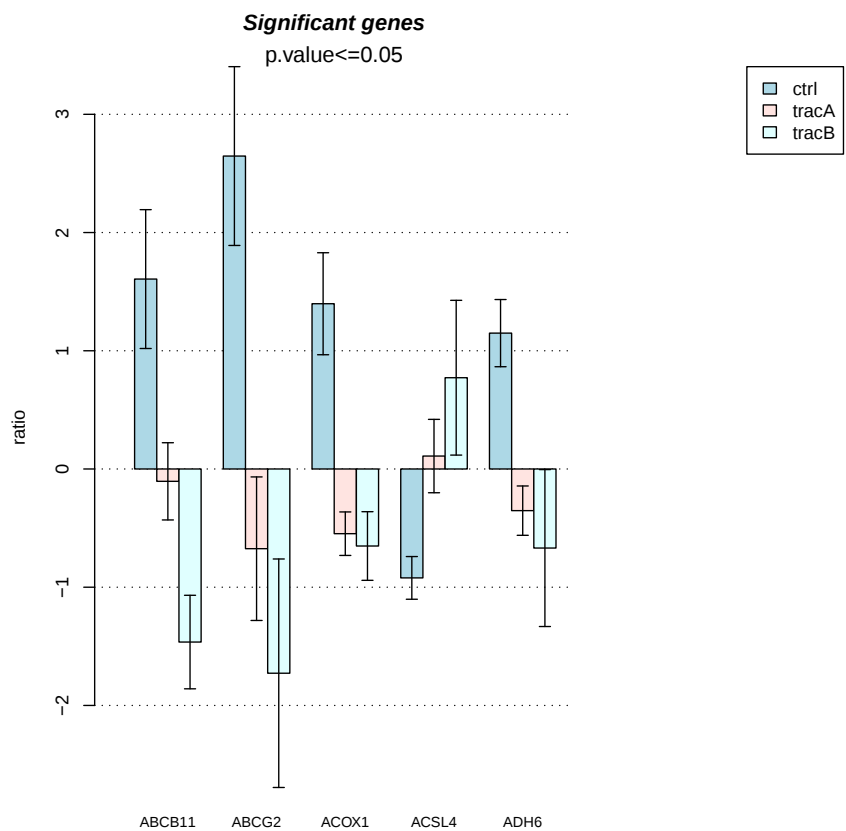
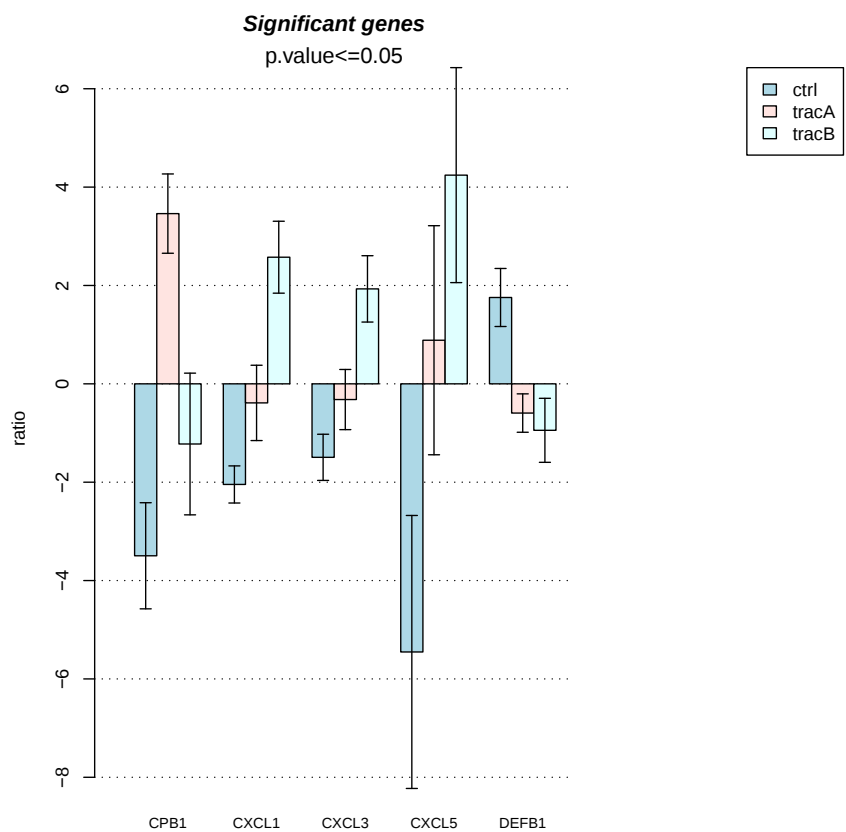
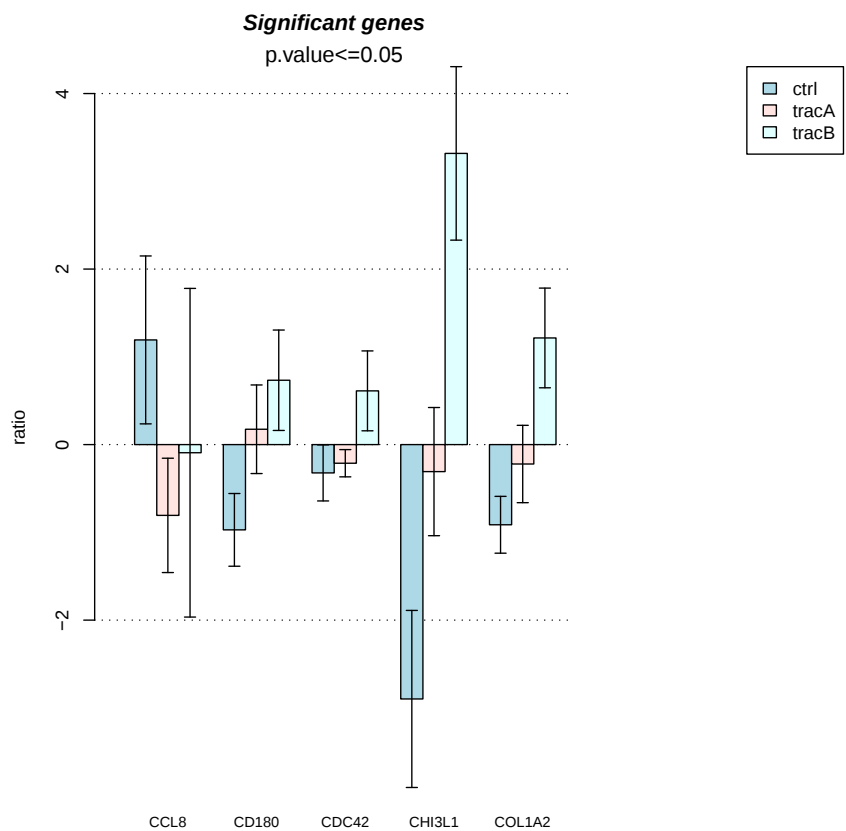
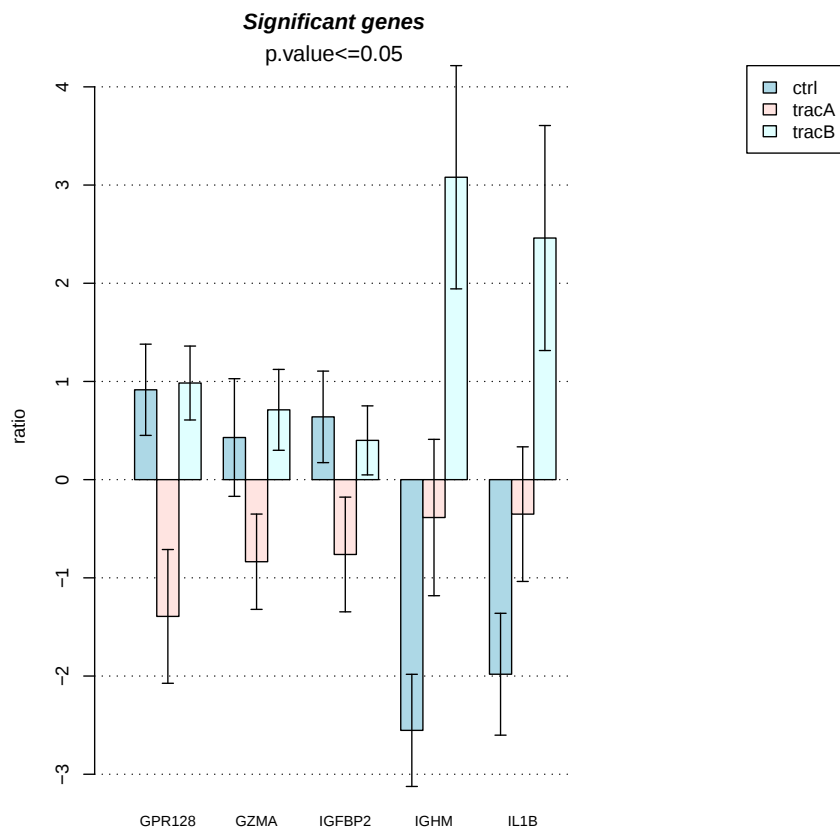
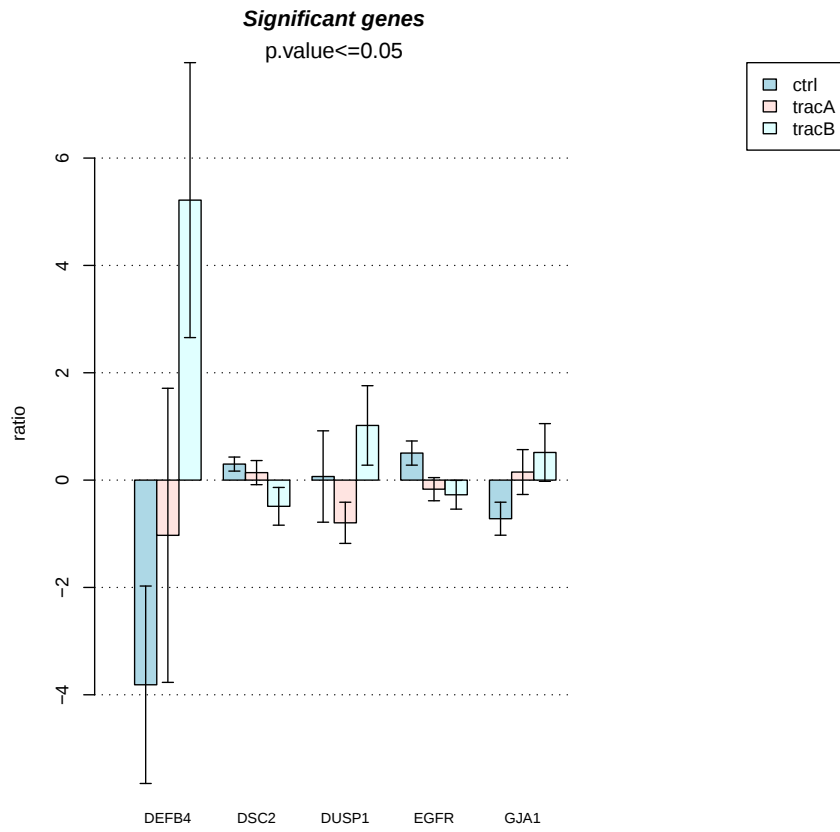


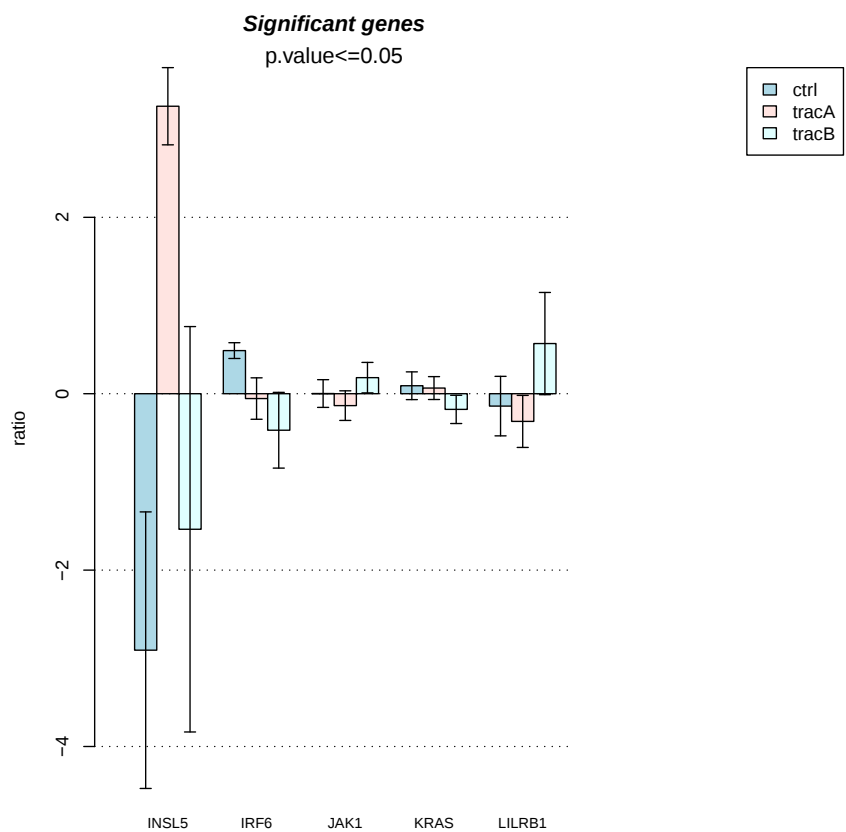
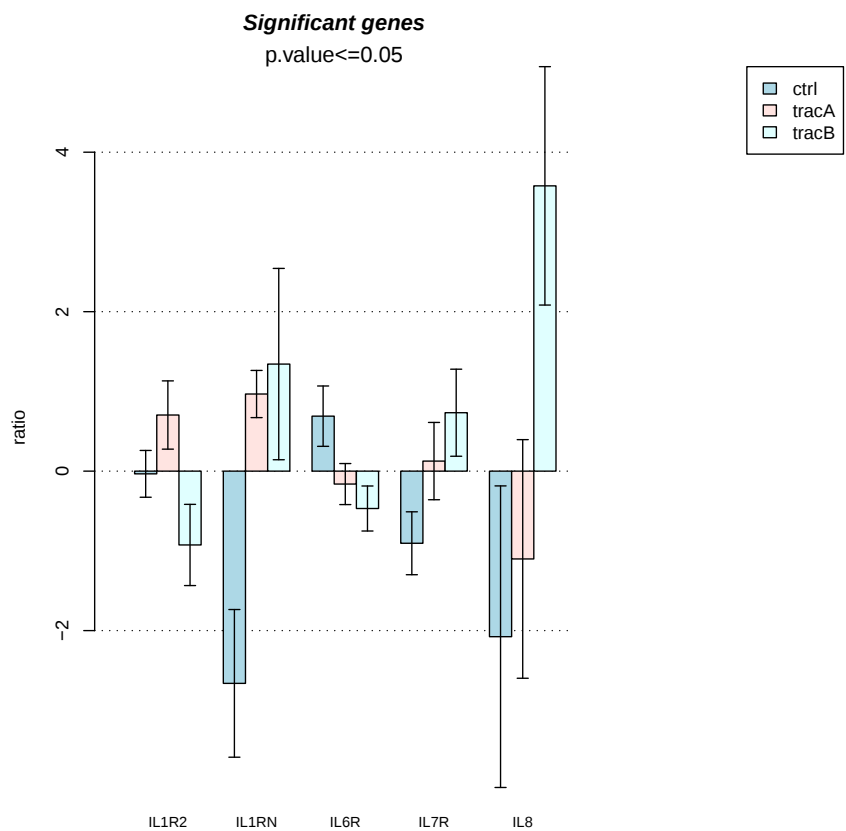
Figure 2: Heatmap of all analyzed genes. Plot colors range between green and red. Red indicates an upper regulation of gene expression and green a down regulation of gene expression. The columns in left corner indicates the significance of this genes. Genes who have a p.value <0.05 are gray, genes who have a p.value <0.01 are brown and genes who have a p.value <0.001 are black.

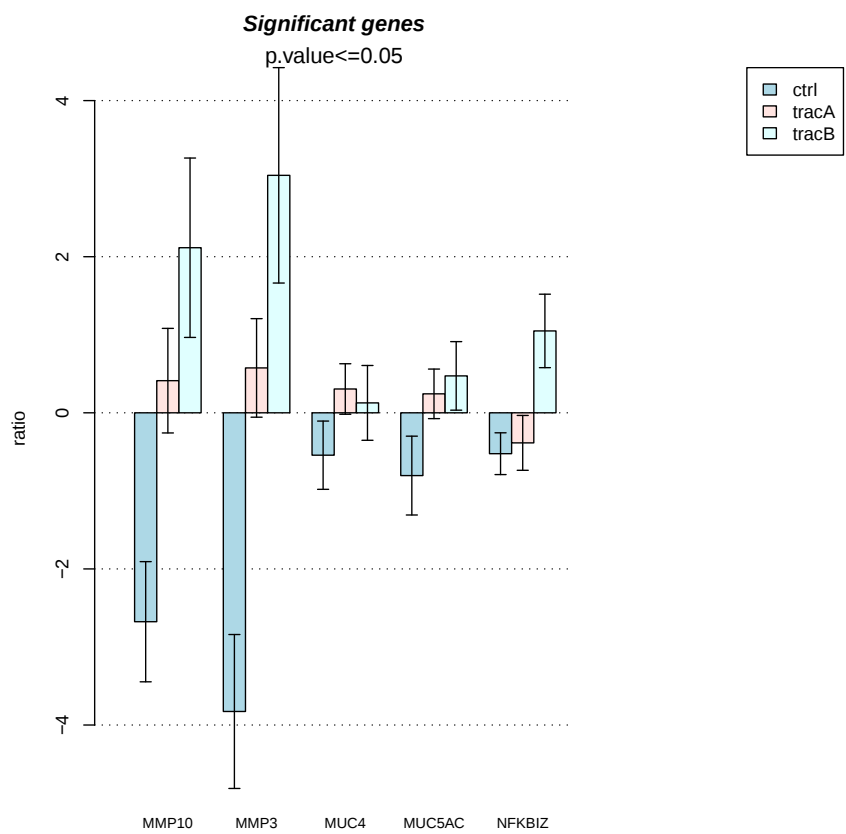
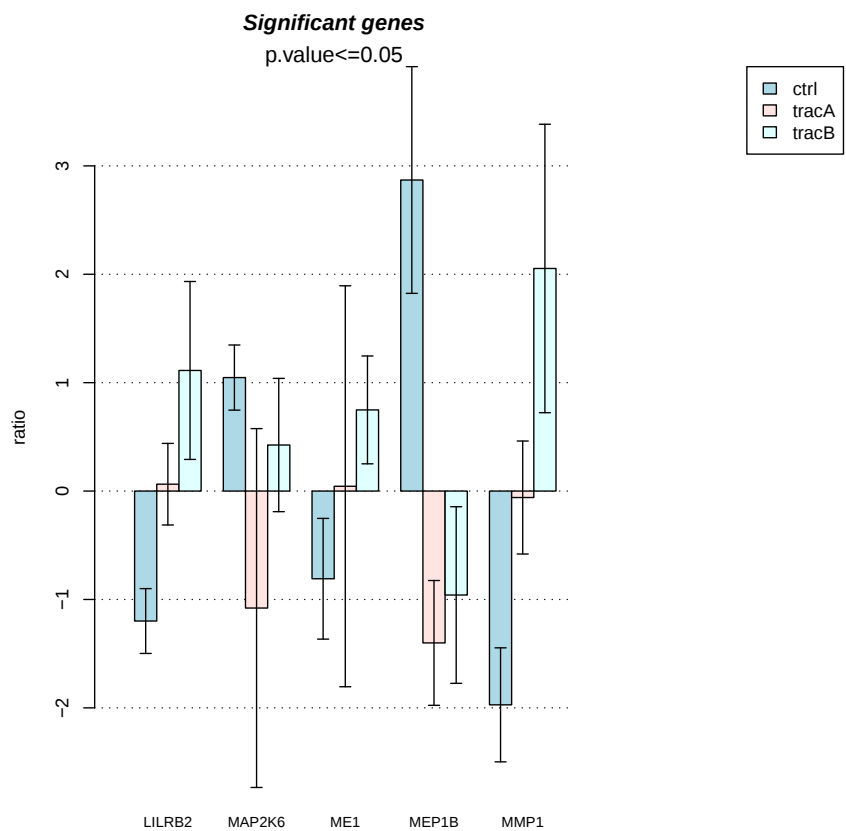


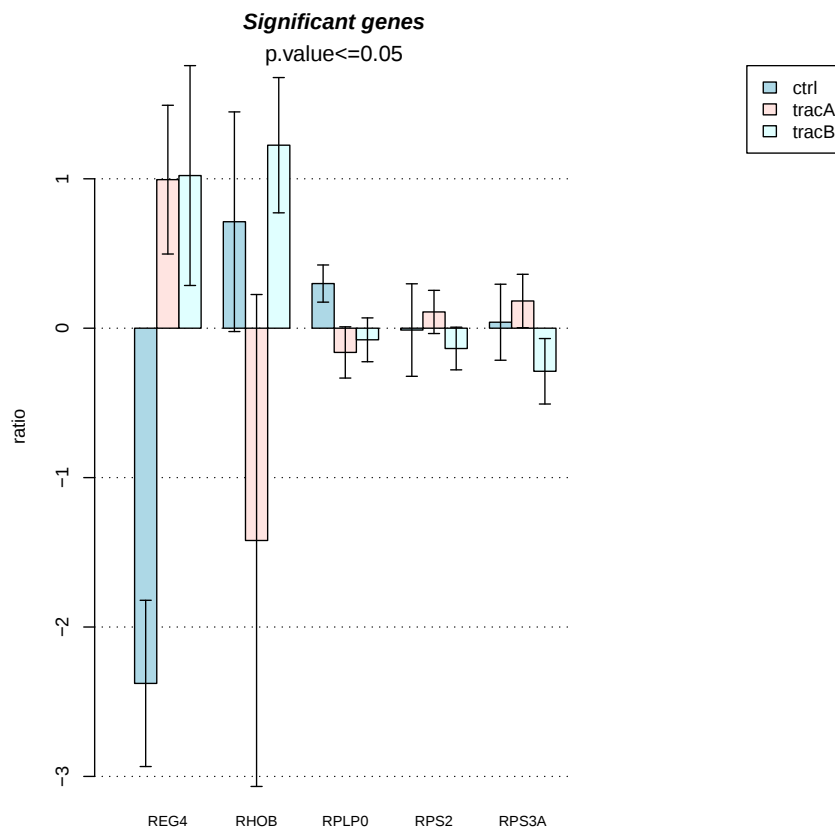
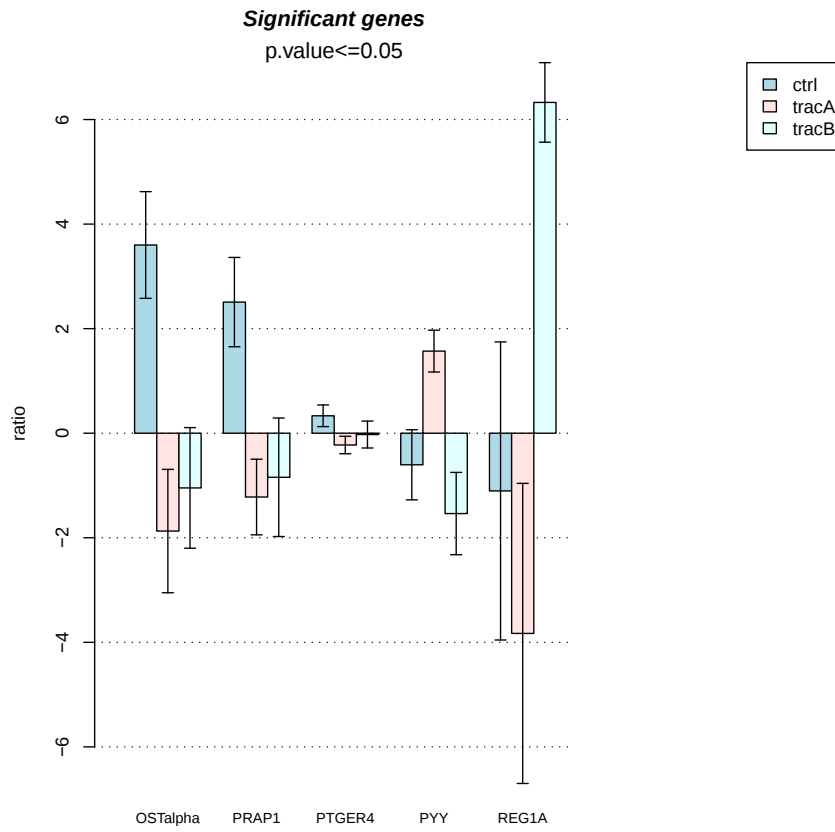


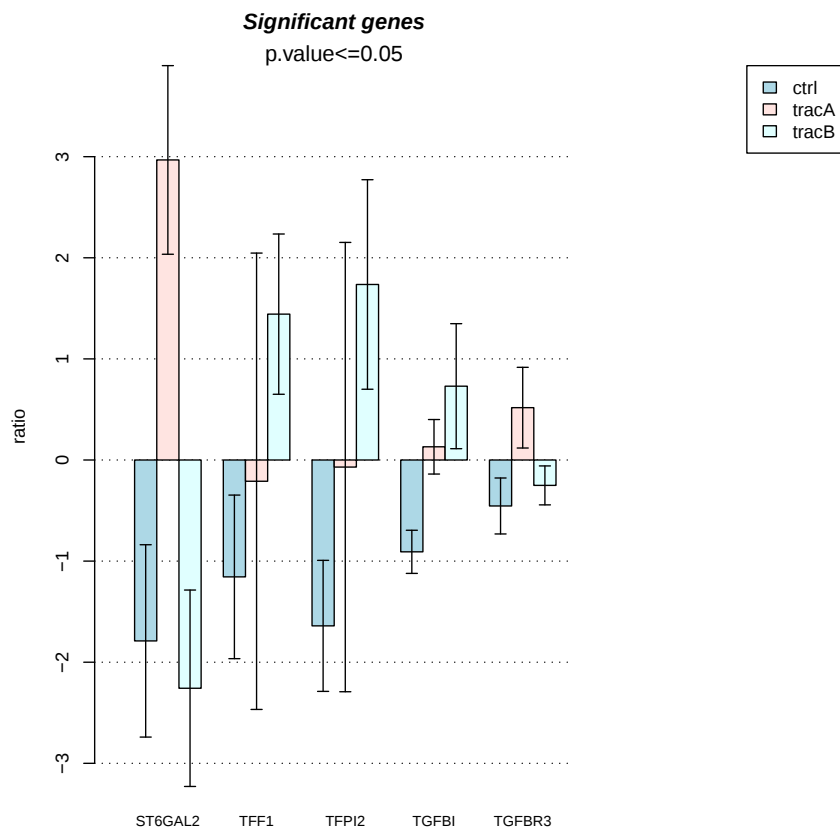
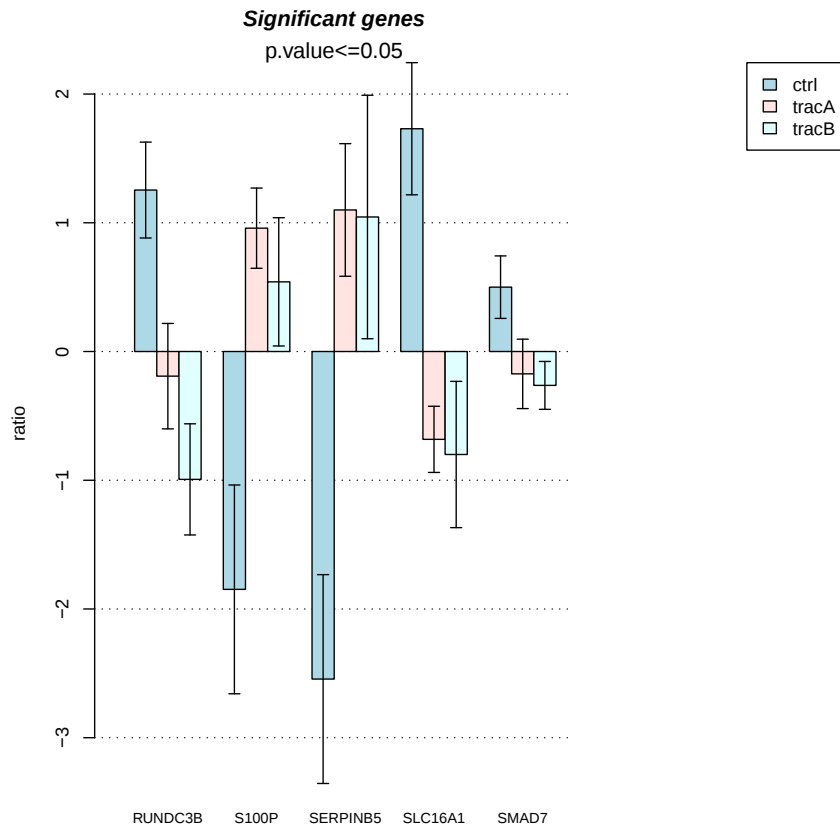


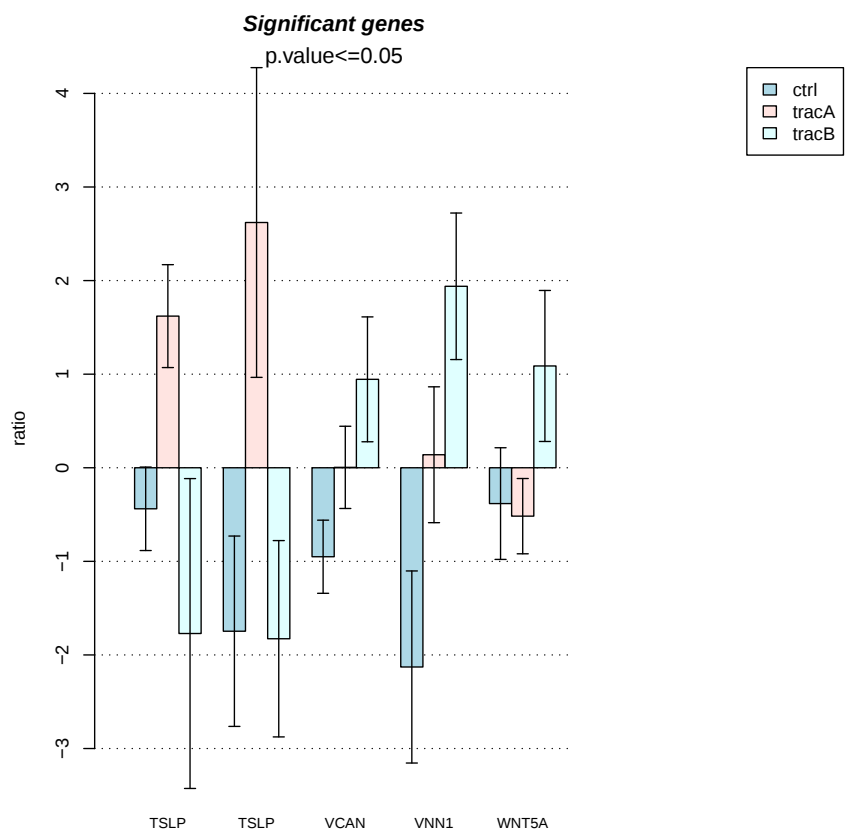
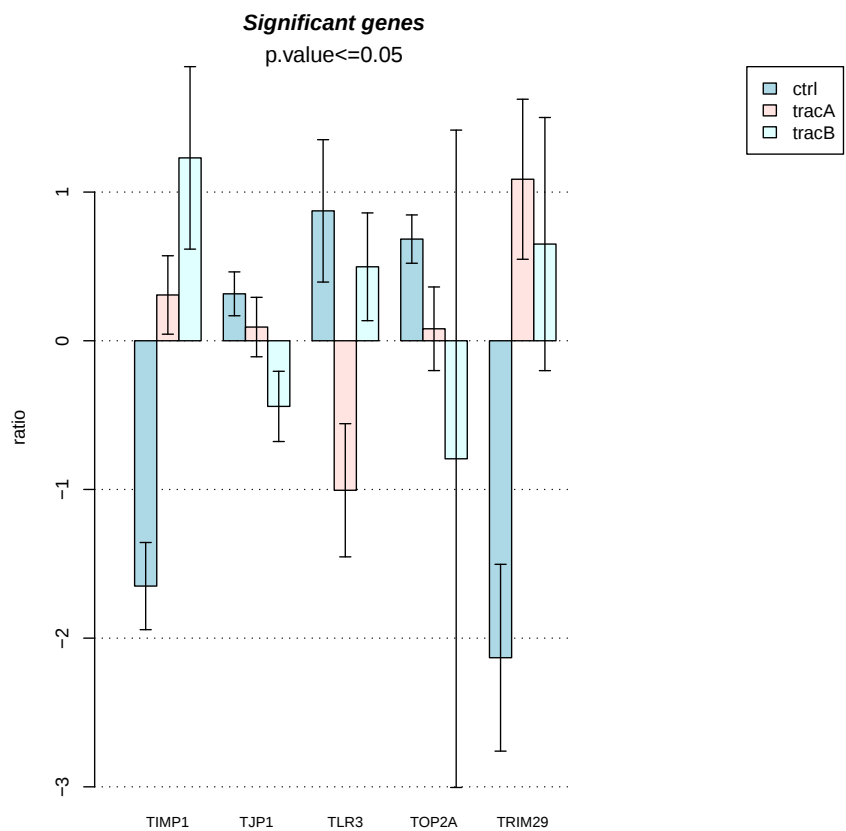




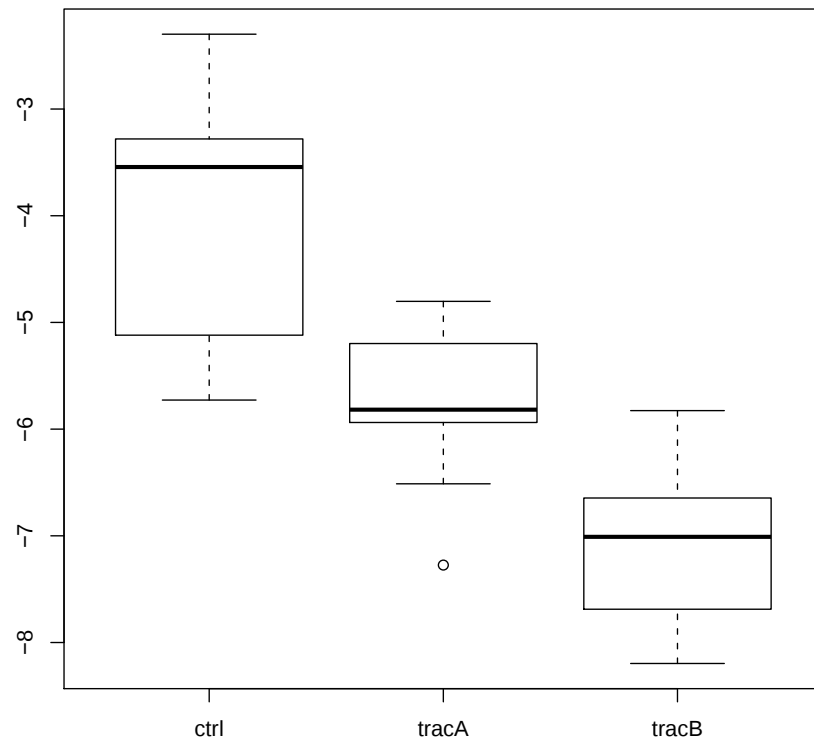




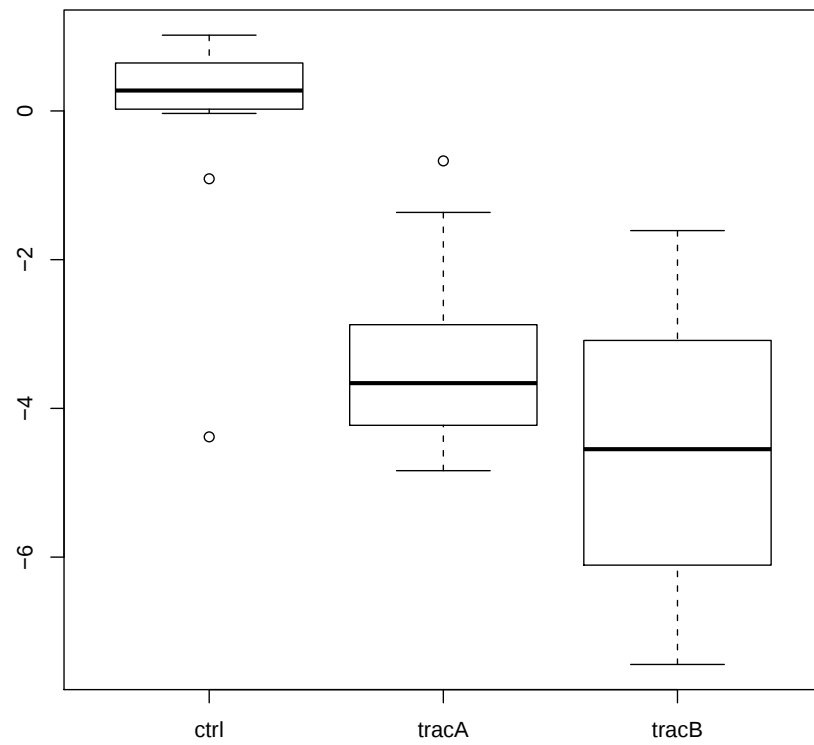




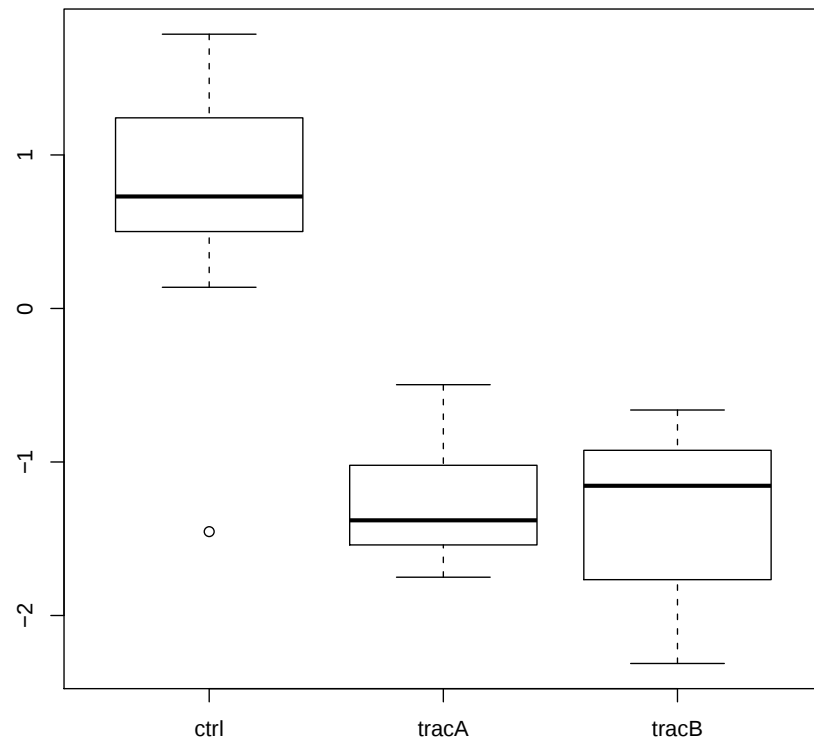
ABCB11-Hs00184824_m1 p.value=1e-06



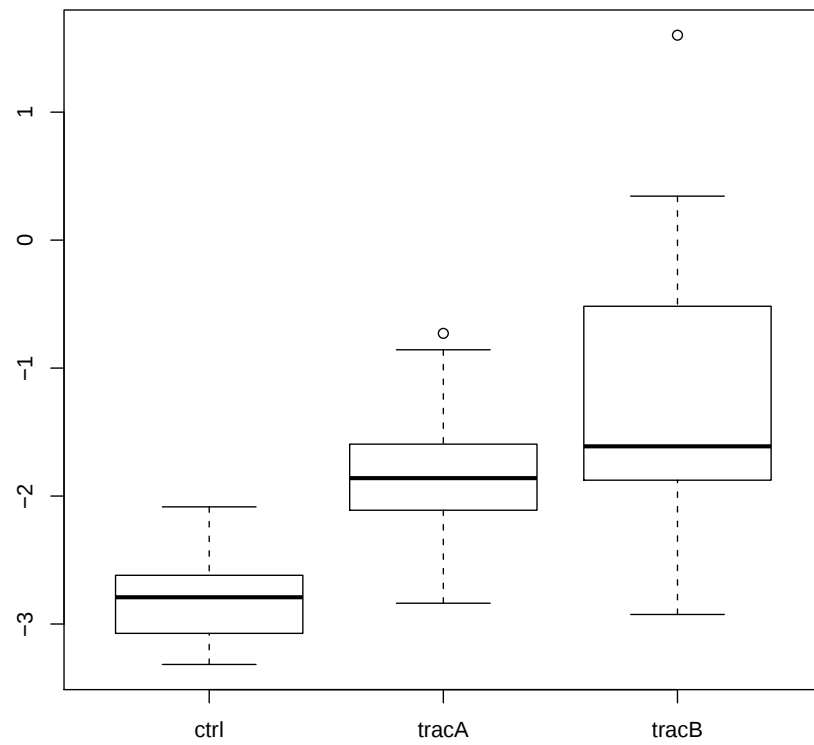
ABCG2-Hs01053787_m1 p.value=2e-06



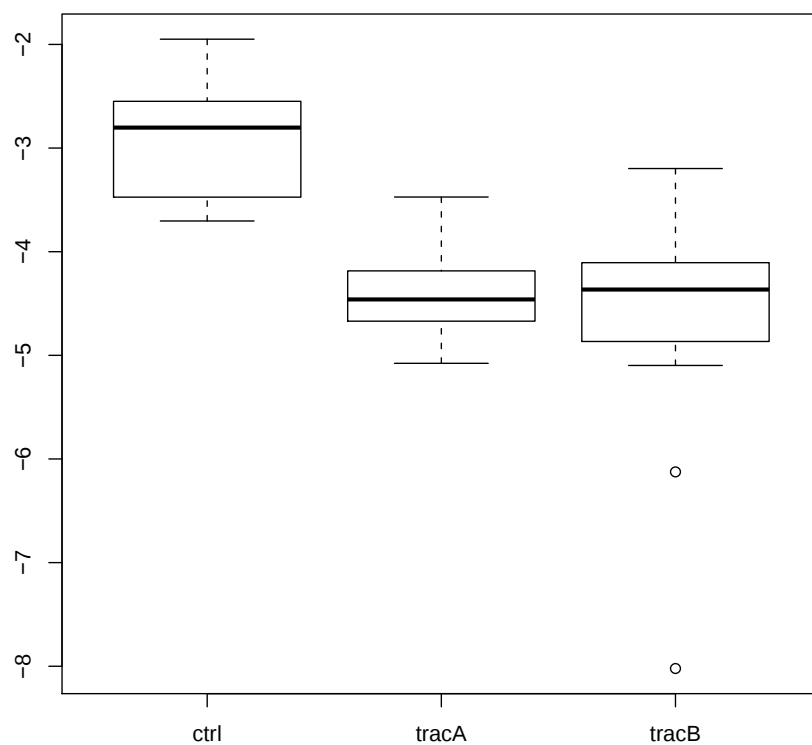
ACOX1-Hs01074241_m1 p.value=0



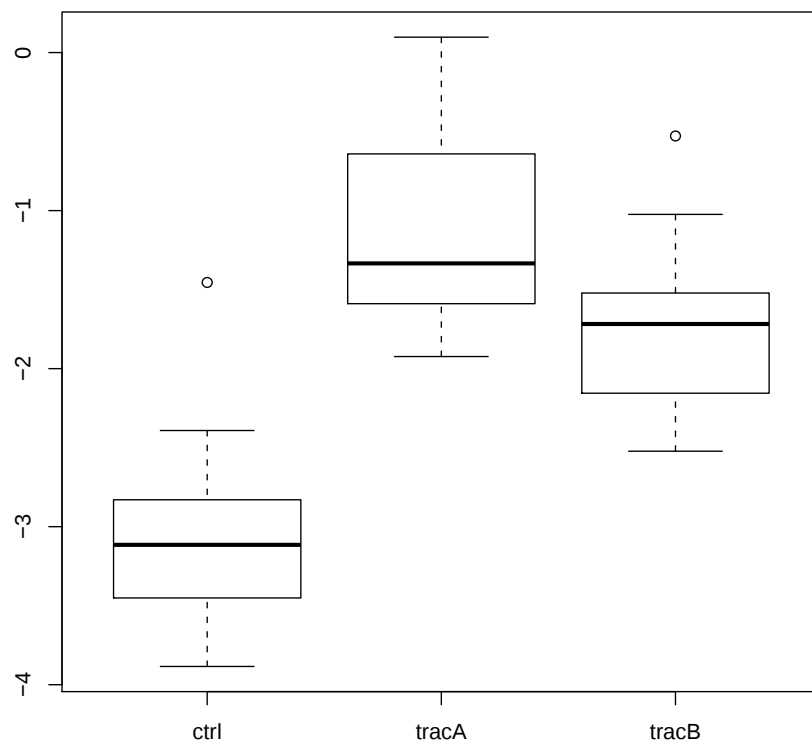
ACSL4-Hs01547083_m1 p.value=5.2e-05



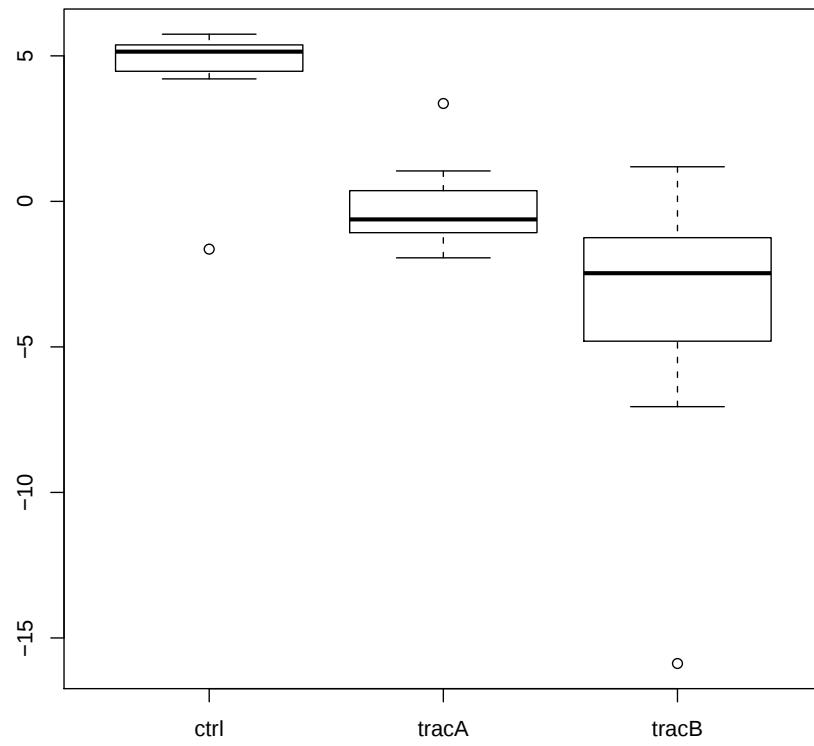
ADH6-Hs00167423_m1 p.value=0



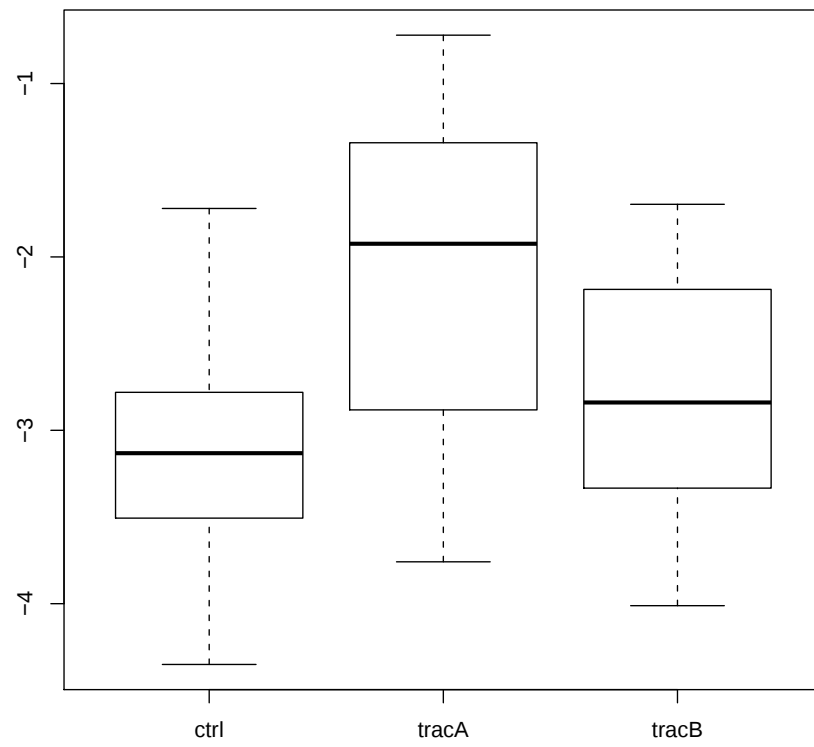
AQP3-Hs00185020_m1 p.value=0



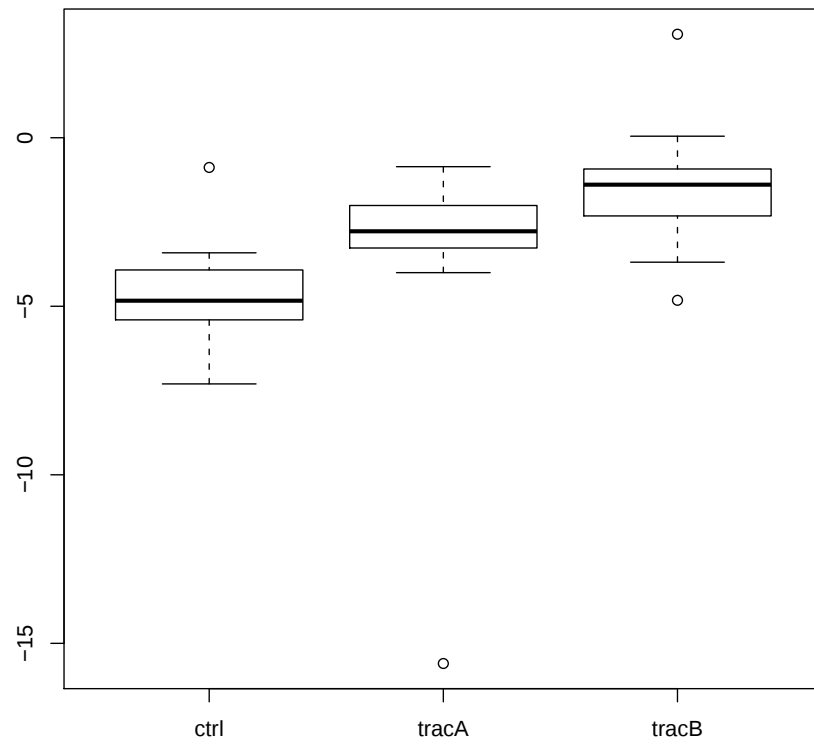
AQP8-Hs01086280_g1 p.value=0



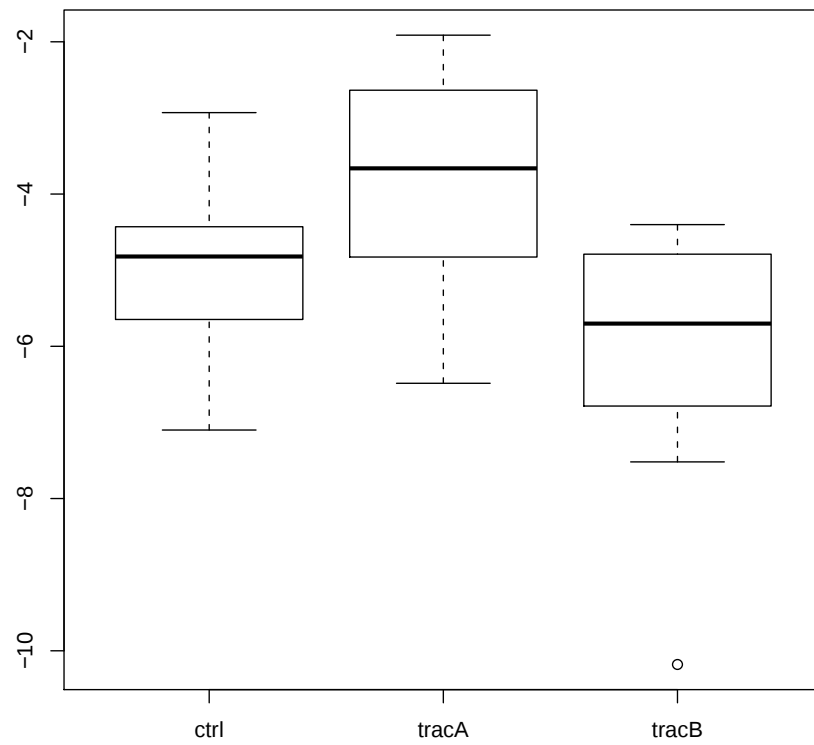
BMP4-Hs00370078_m1 p.value=0.012764



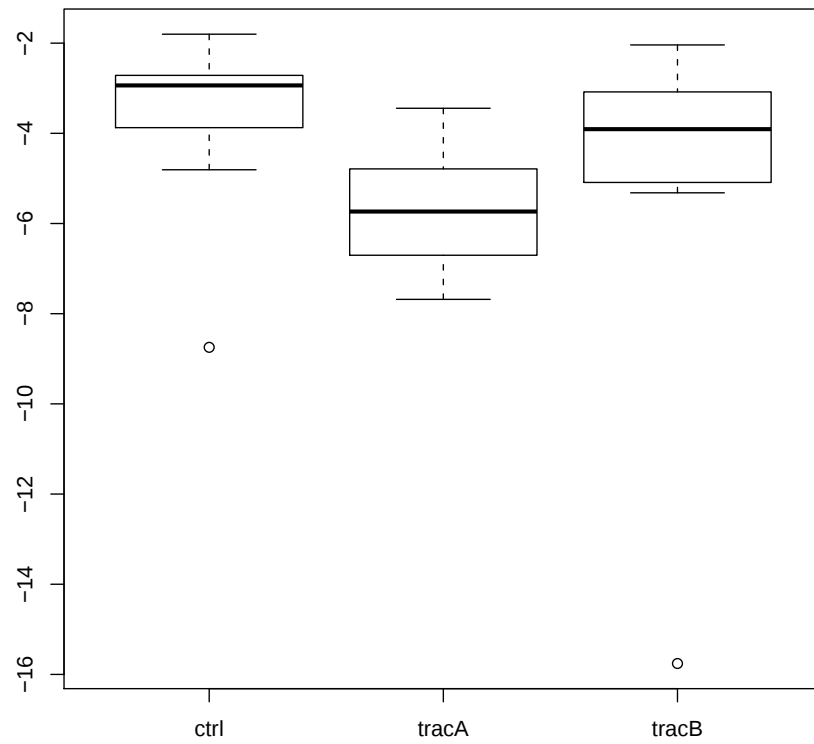
CCL18-Hs00268113_m1 p.value=8e-04



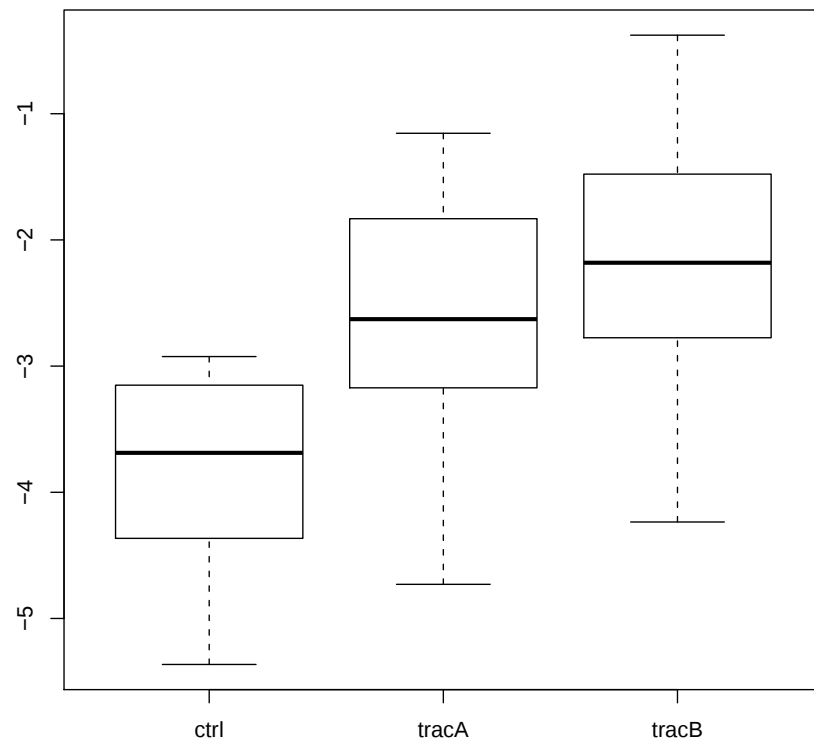
CCL23-Hs00270756_m1 p.value=0.000639



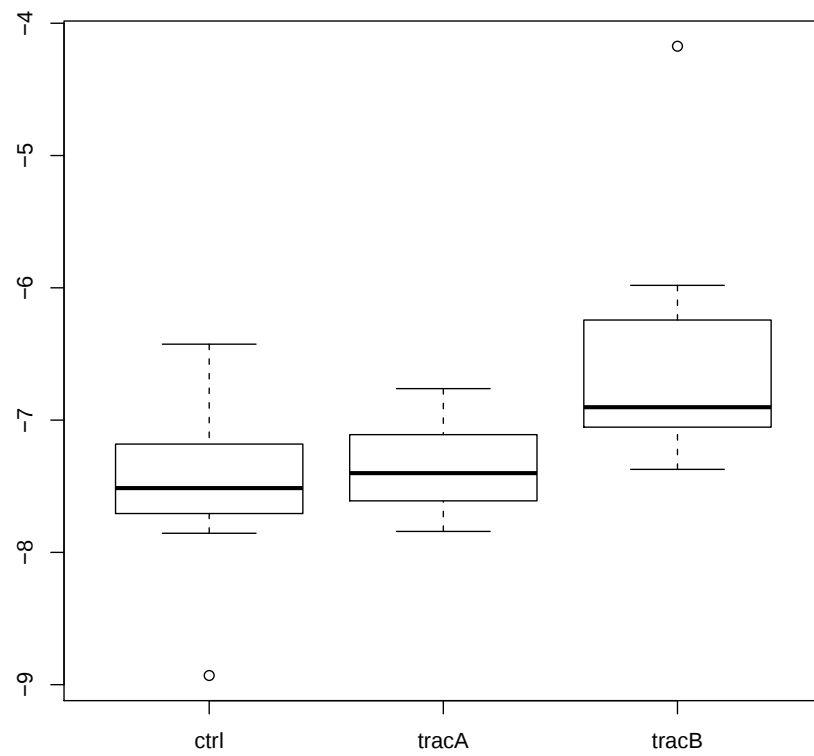
CCL8-Hs00428422_g1 p.value=0.004072



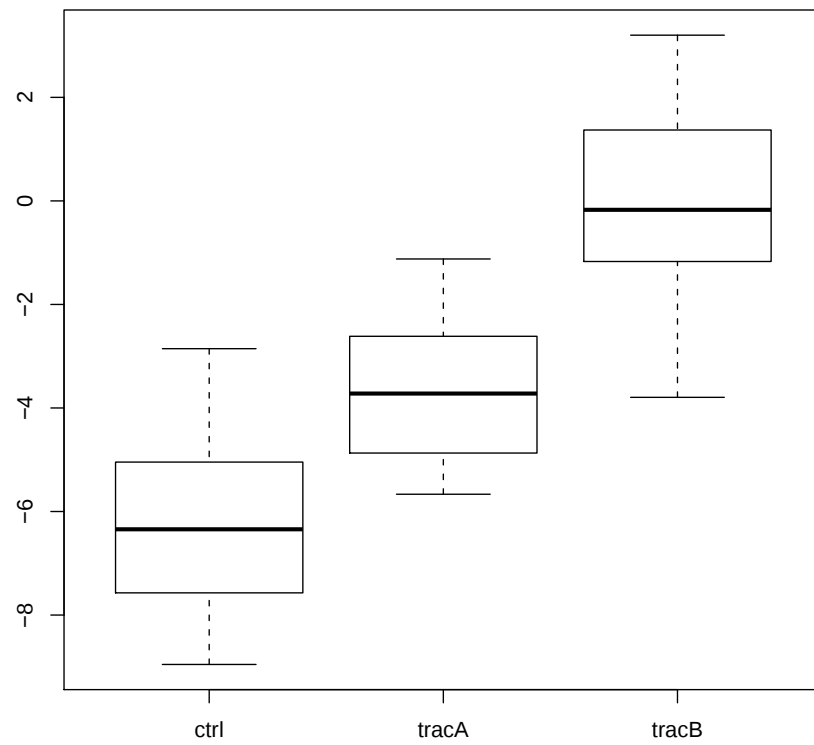
CD180-Hs01069872_m1 p.value=0.000699



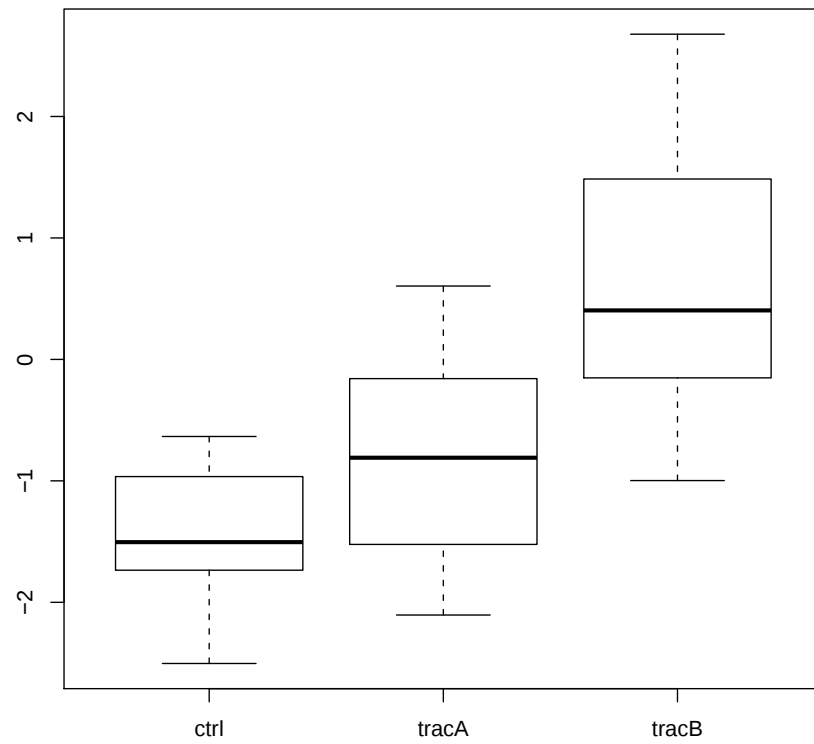
CDC42-Hs00741586_mH p.value=0.003172



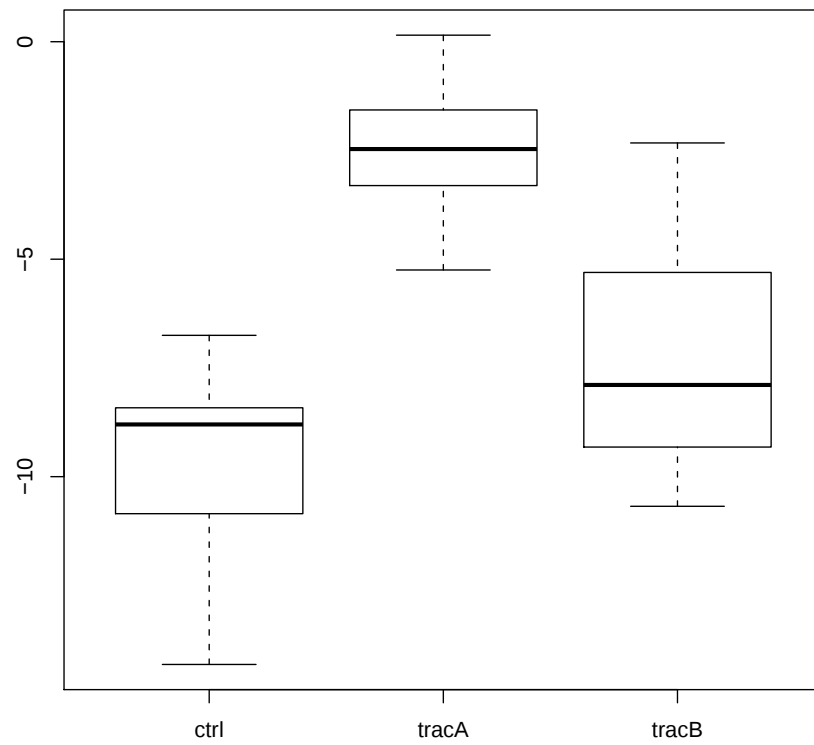
CHI3L1-Hs01072228_m1 p.value=0



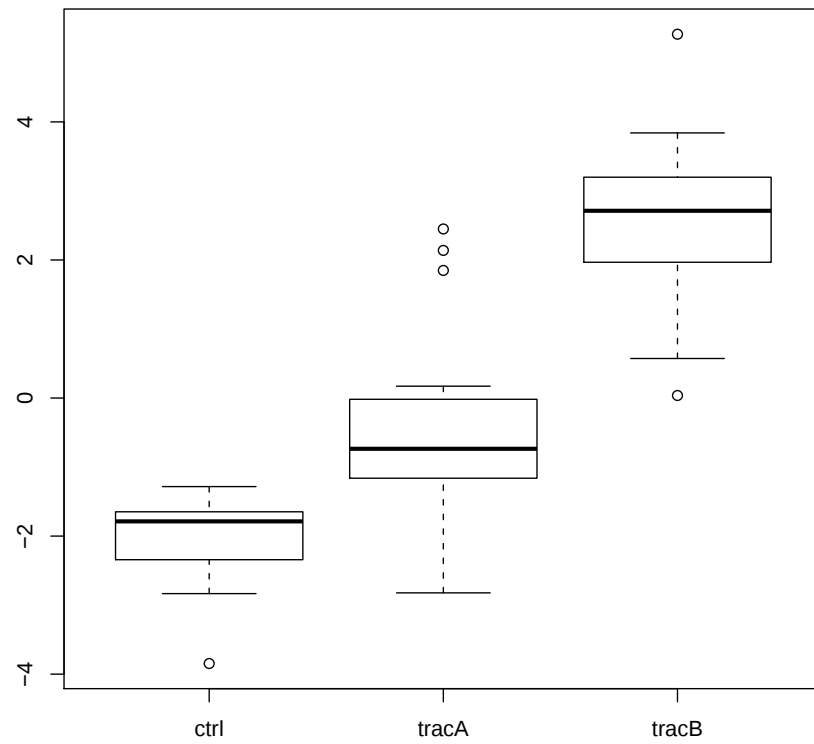
COL1A2-Hs01028971_m1 p.value=2.8e-05



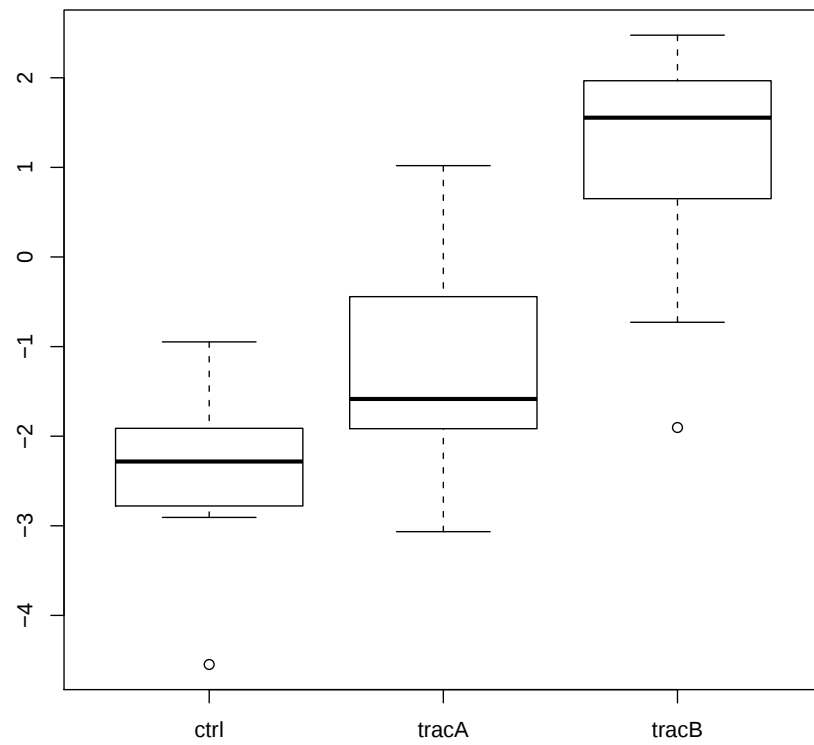
CPB1-Hs00157026_m1 p.value=0



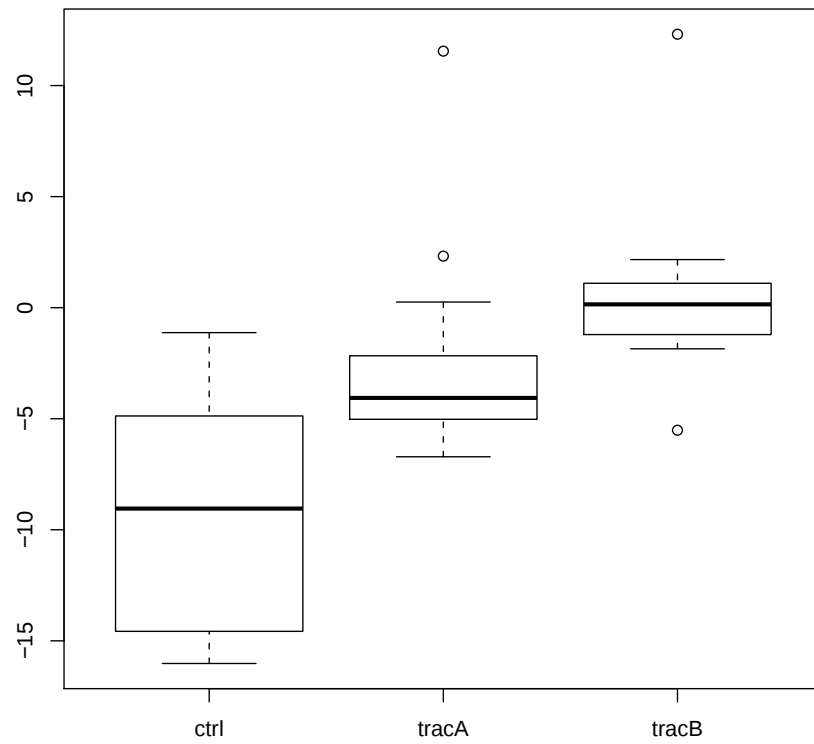
CXCL1-Hs00605382_gH p.value=0



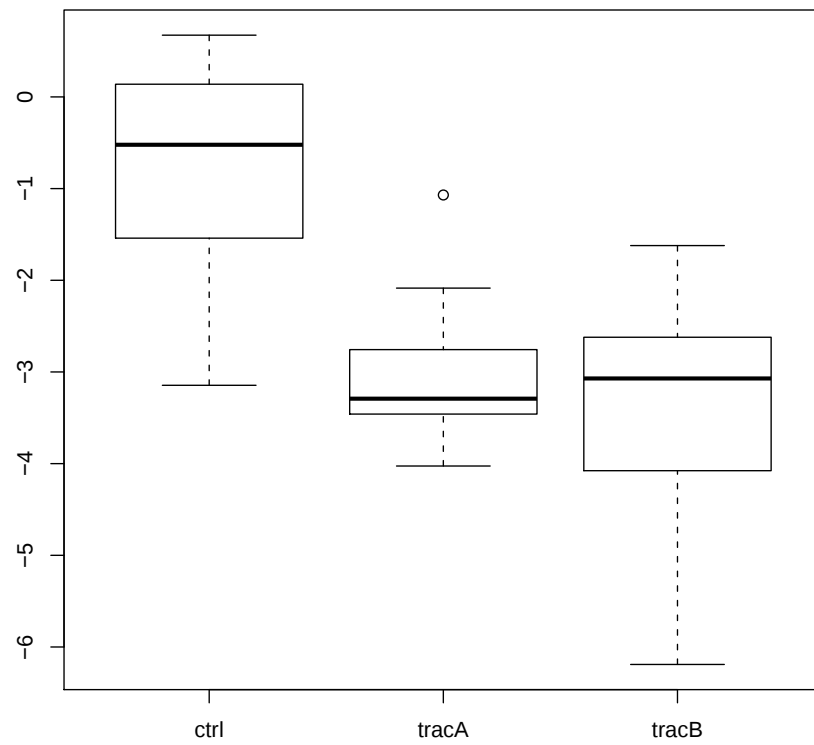
CXCL3-Hs00171061_m1 p.value=1e-06



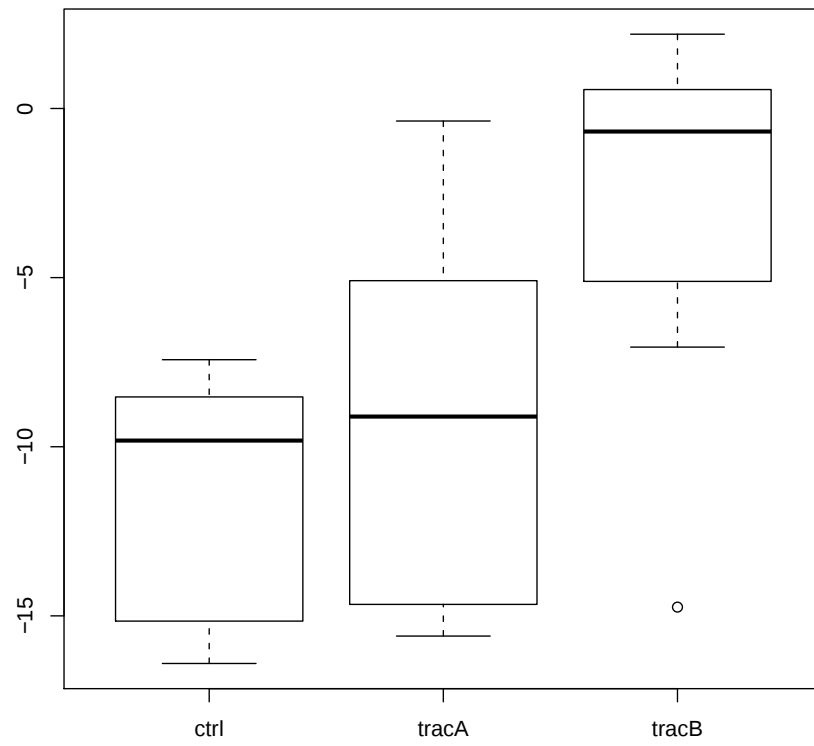
CXCL5-Hs00607029_g1 p.value=0.000194



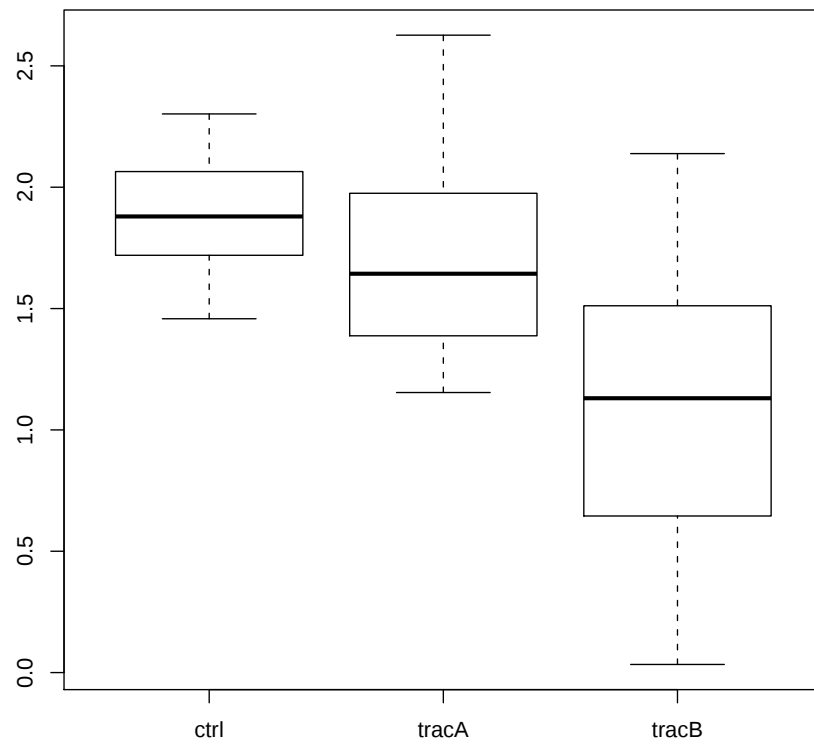
DEFB1-Hs00608345_m1 p.value=3e-06



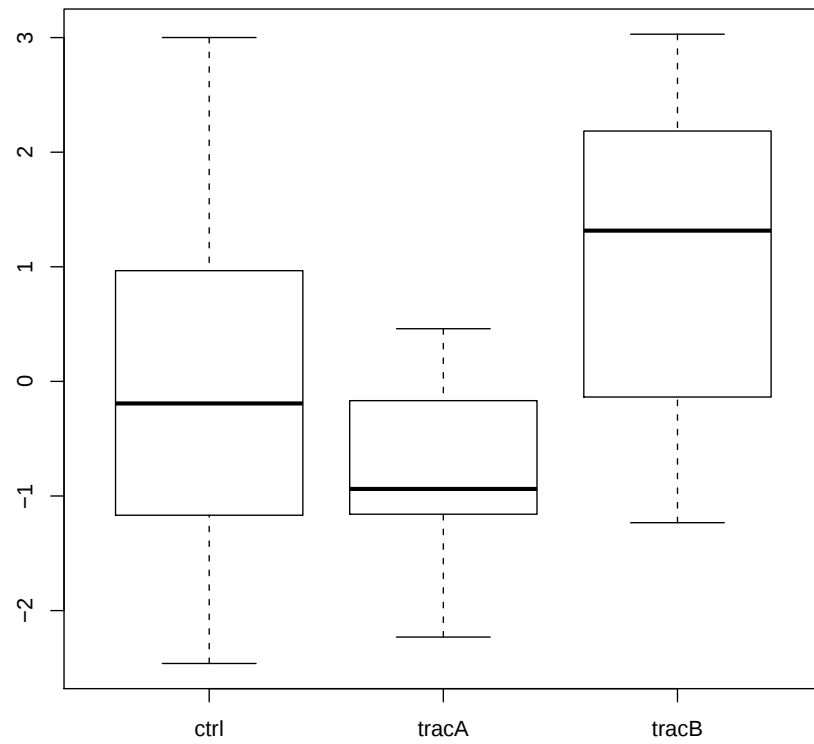
DEFB4-Hs00175474_m1 p.value=0.000123



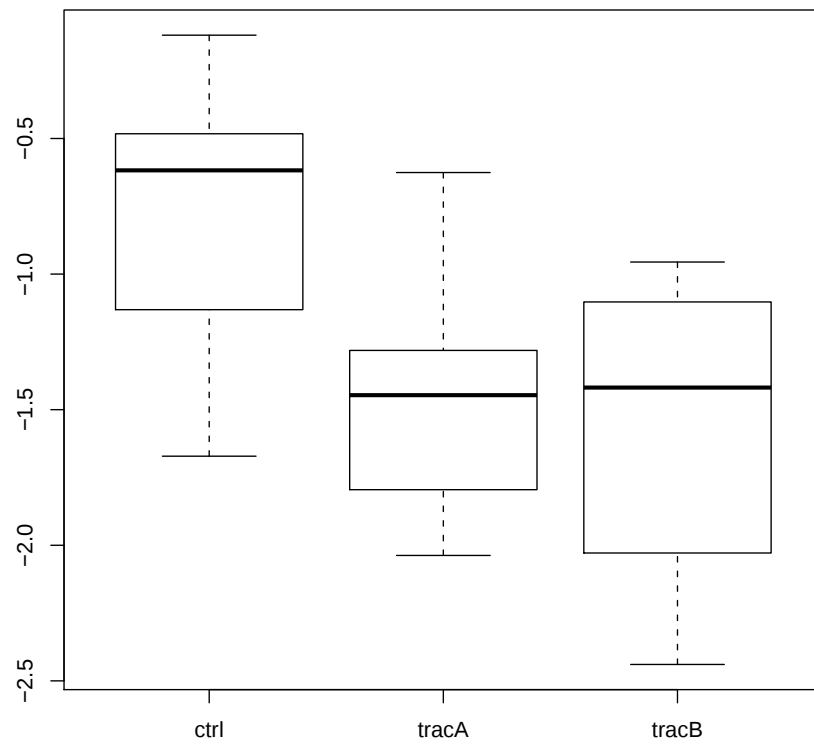
DSC2-Hs00951428_m1 p.value=0.00237



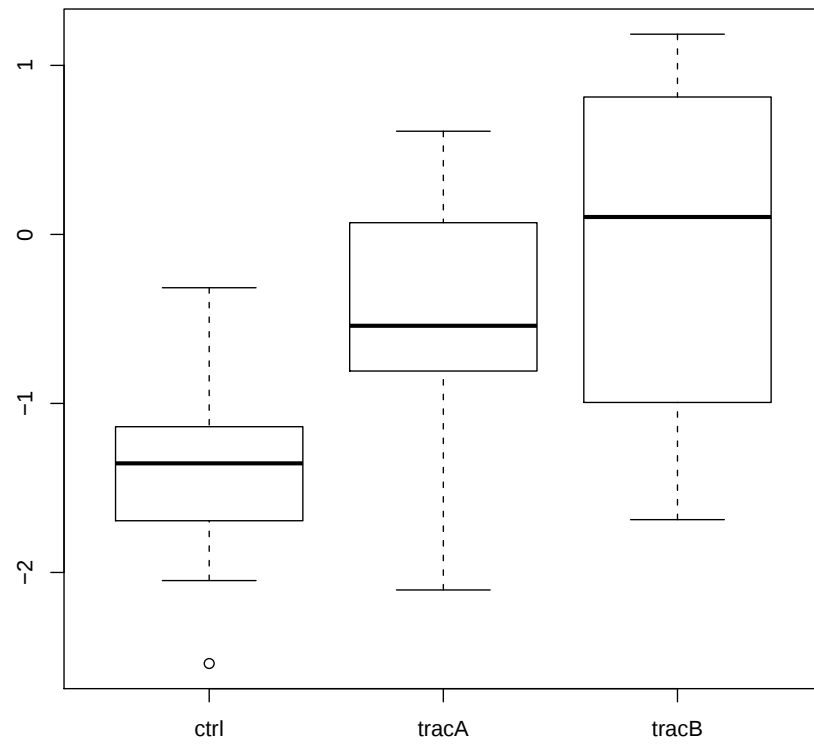
DUSP1-Hs00610257_g1 p.value=0.000435



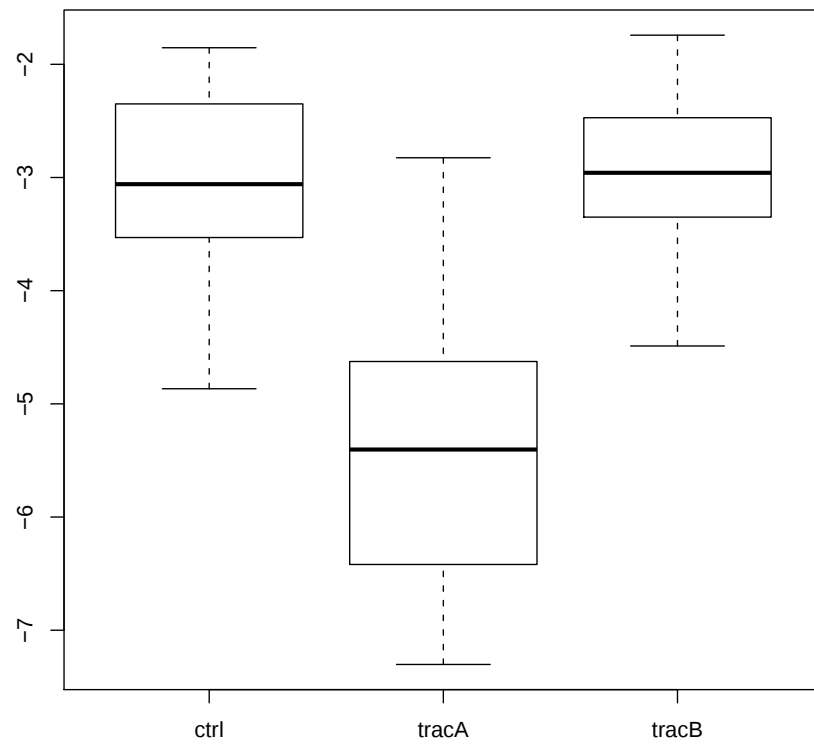
EGFR-Hs01076093_g1 p.value=0.000754



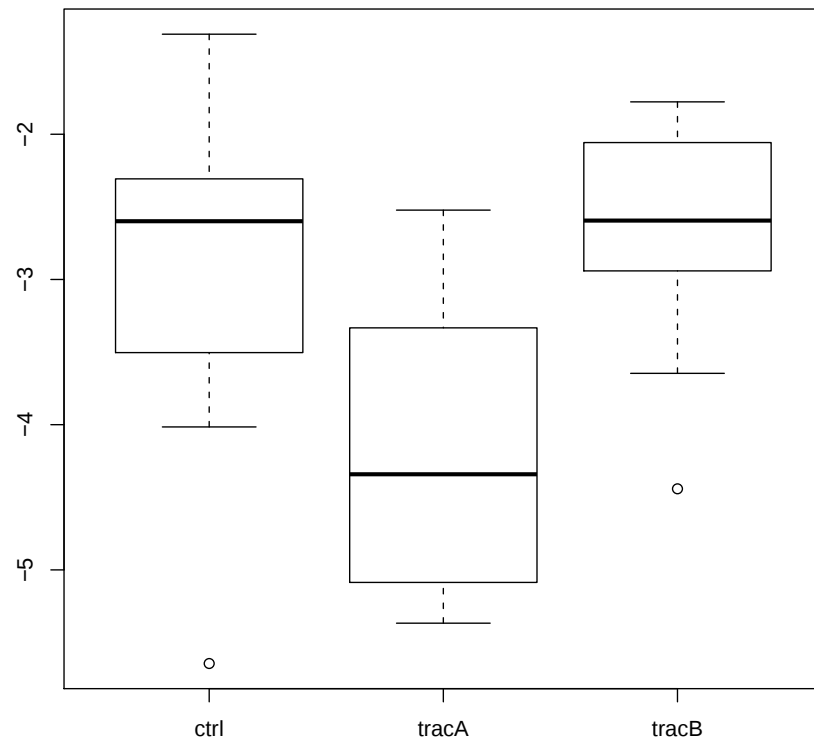
GJA1-Hs00748445_s1 p.value=0.003536



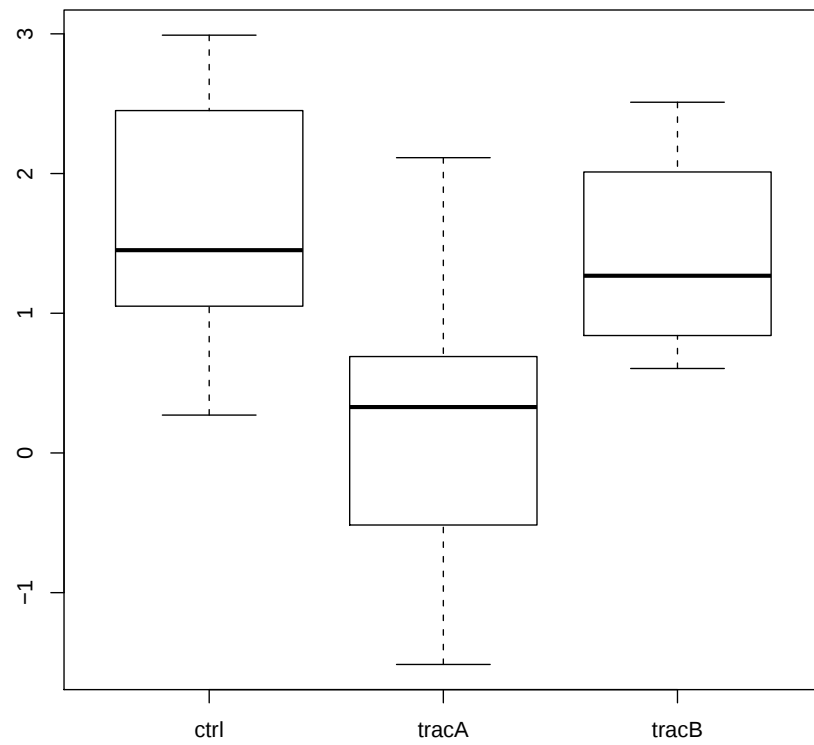
GPR128-Hs00262184_m1 p.value=2.4e-05



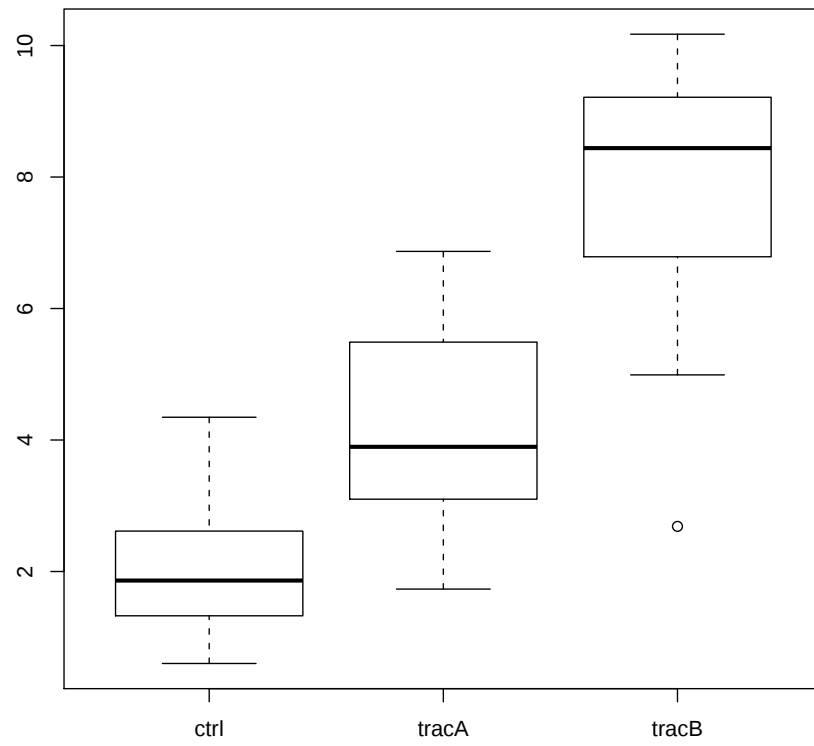
GZMA-Hs00989184_m1 p.value=0.000265



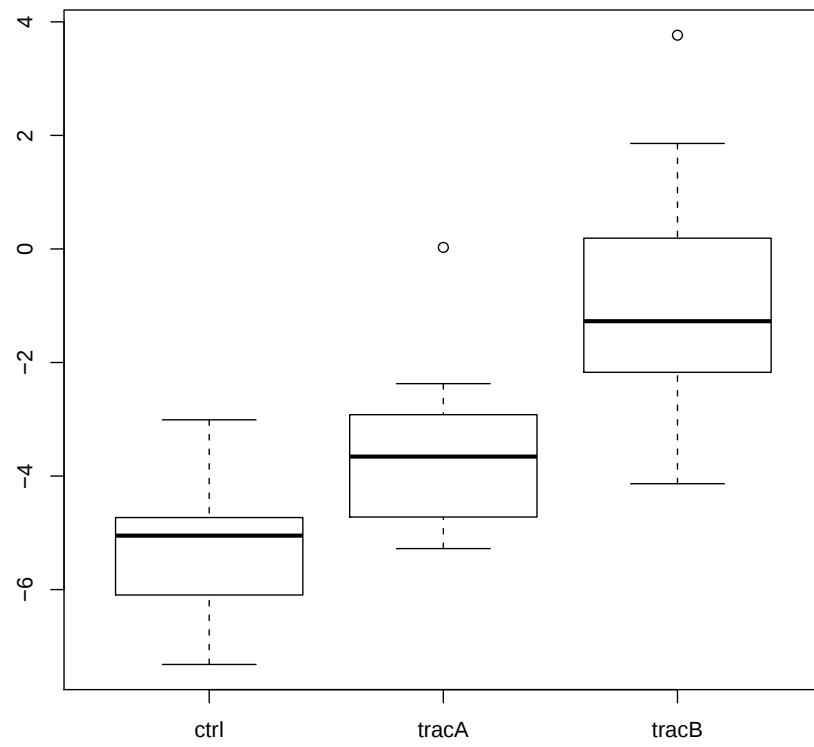
IGFBP2-Hs01040719_m1 p.value=0.003154



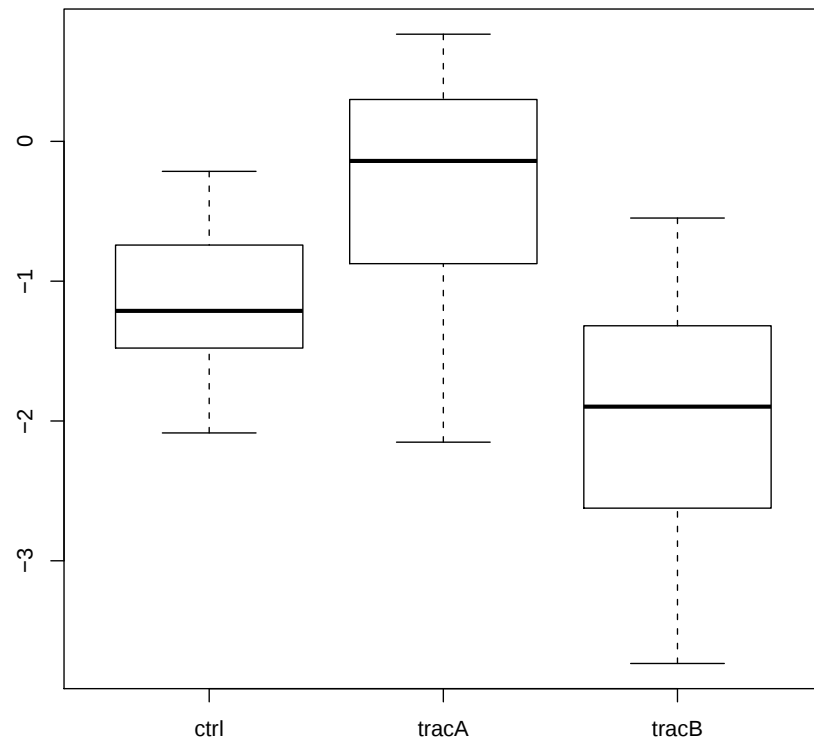
IGHM-Hs00378230_g1 p.value=0



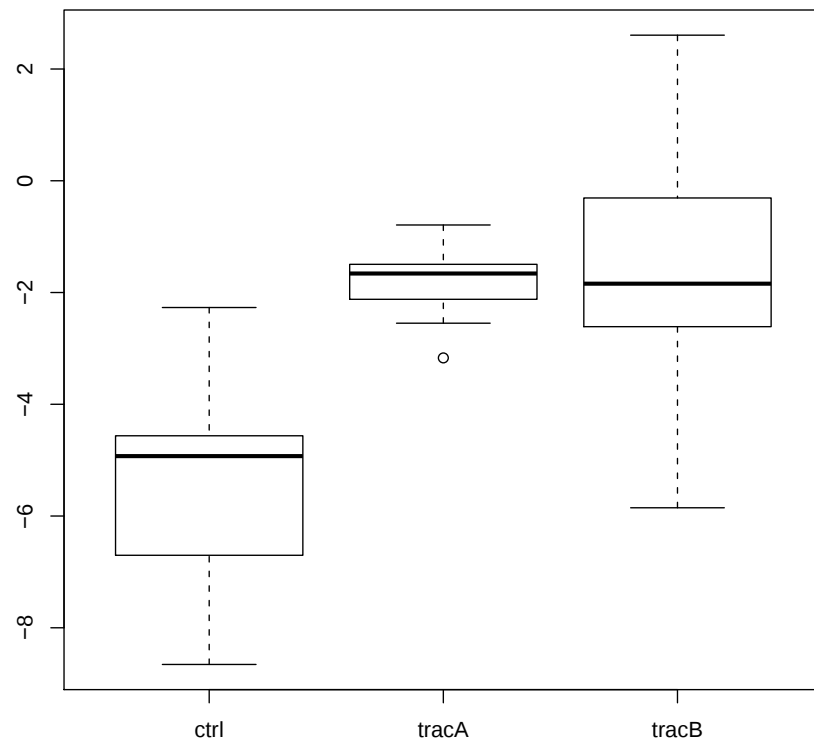
IL1B-Hs01555413_m1 p.value=1.6e-05



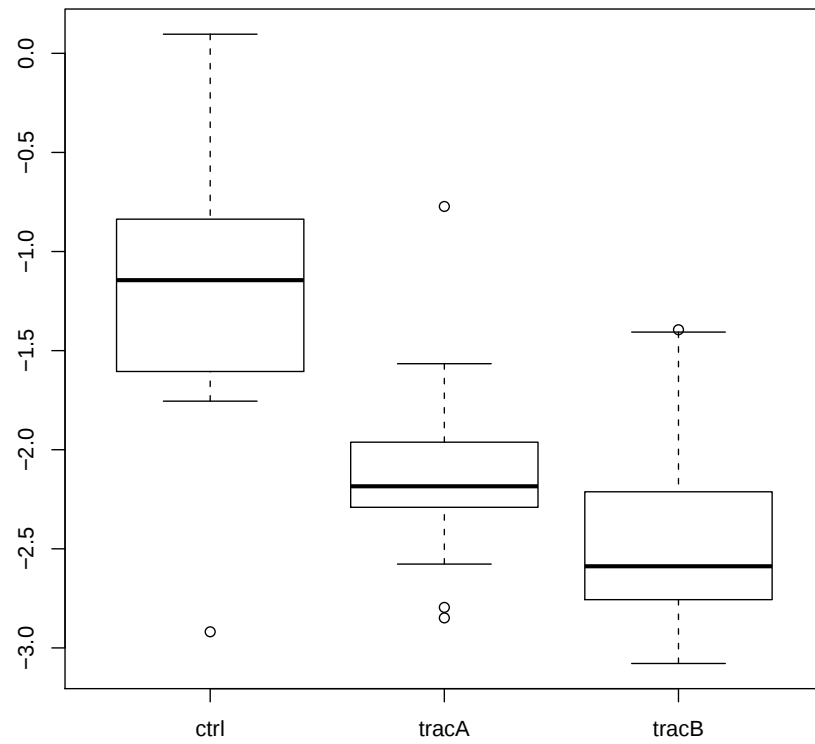
IL1R2-Hs00174759_m1 p.value=0.000178



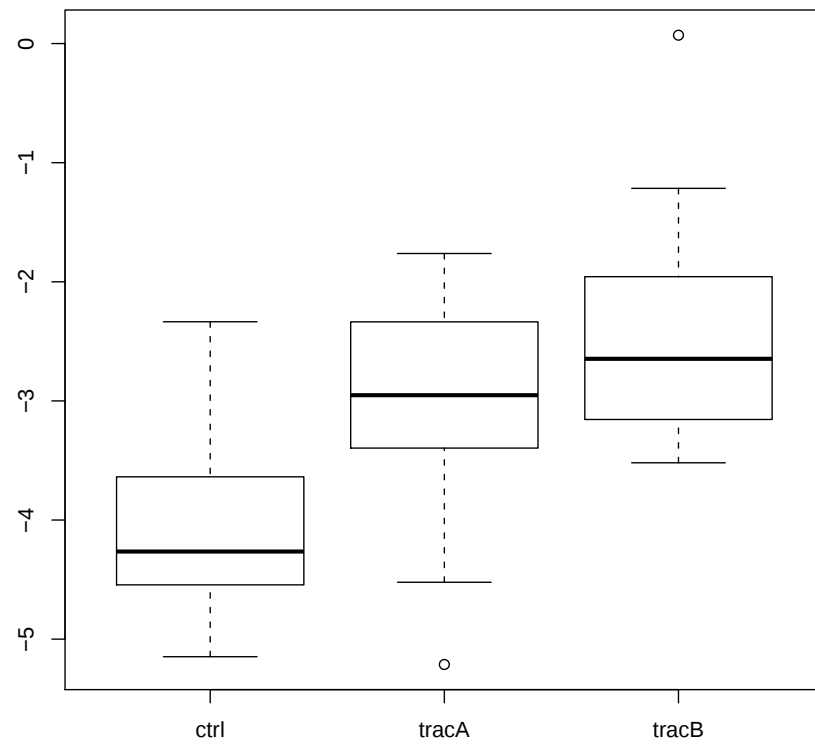
IL1RN-Hs00893626_m1 p.value=0



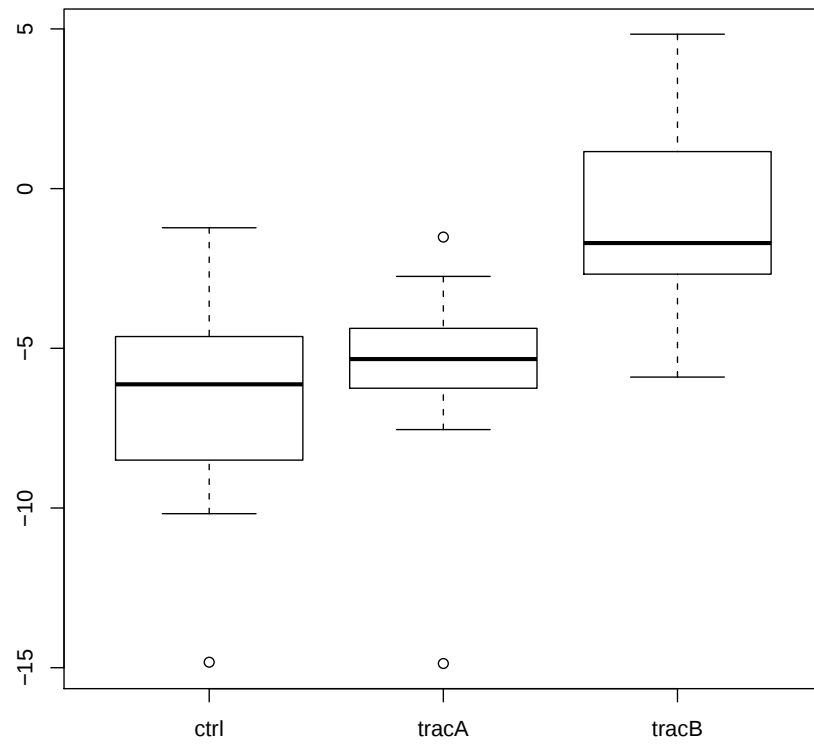
IL6R-Hs00794121_m1 p.value=6e-04



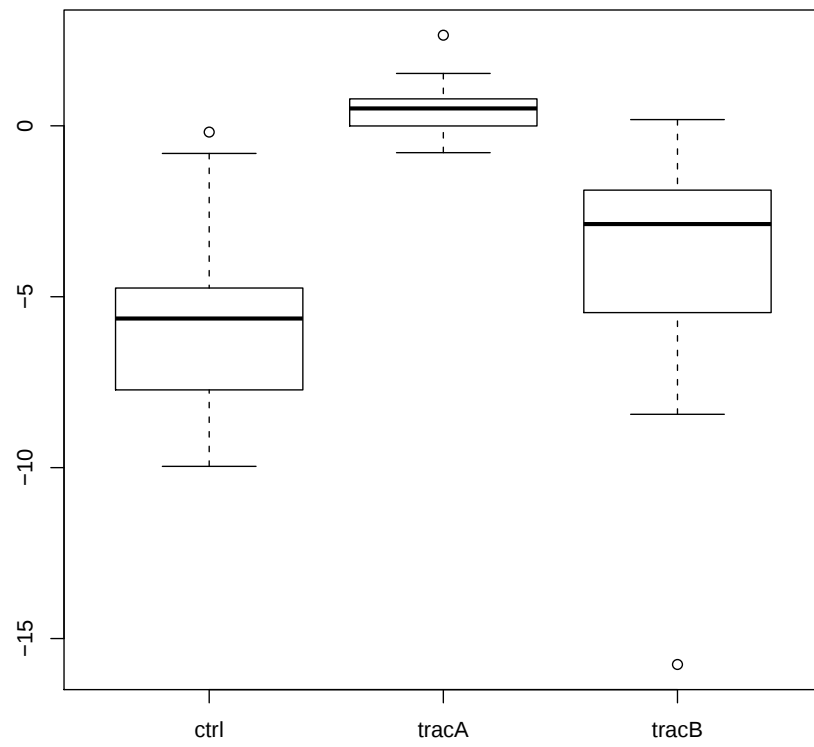
IL7R-Hs00902338_g1 p.value=0.000654



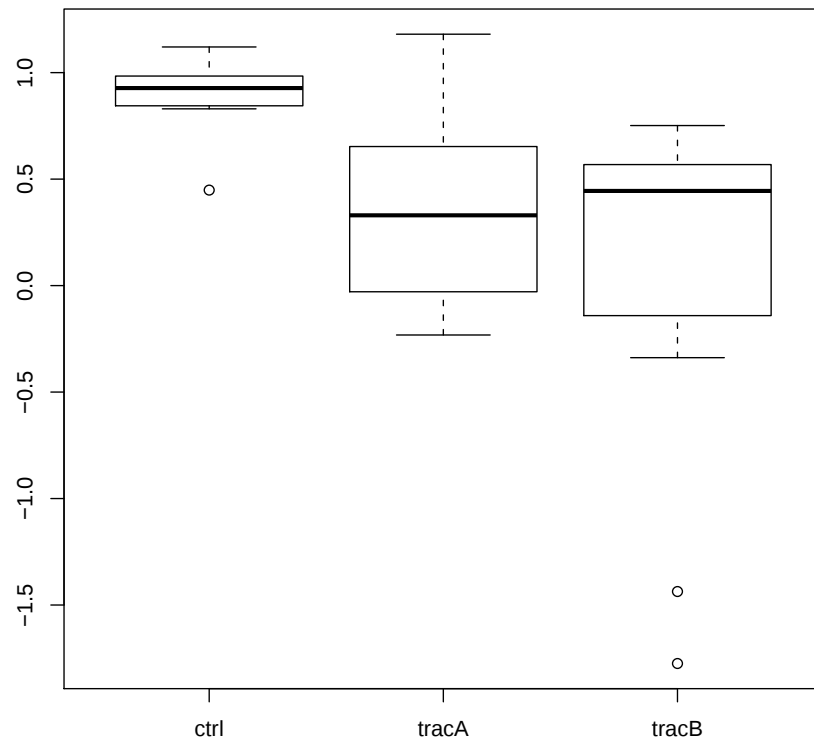
IL8-Hs01567913_g1 p.value=0.000613



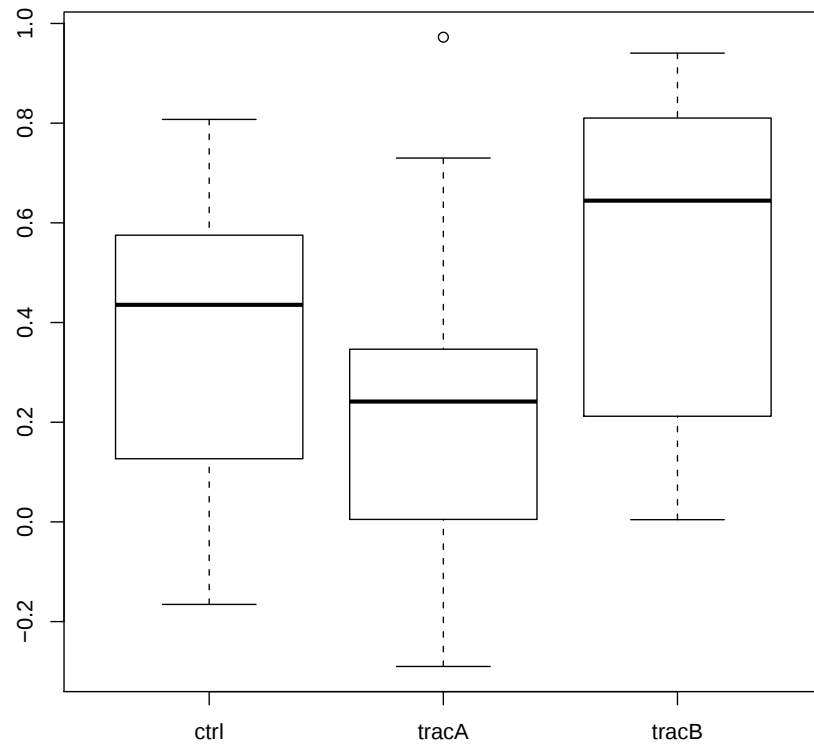
INSL5-Hs00193884_m1 p.value=0



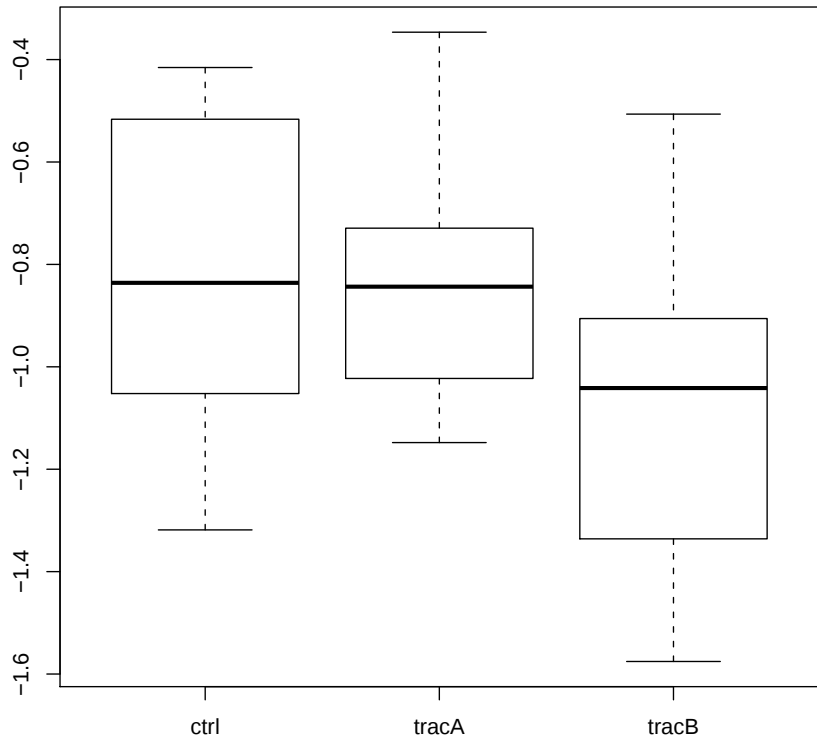
IRF6-Hs00196213_m1 p.value=0.001318



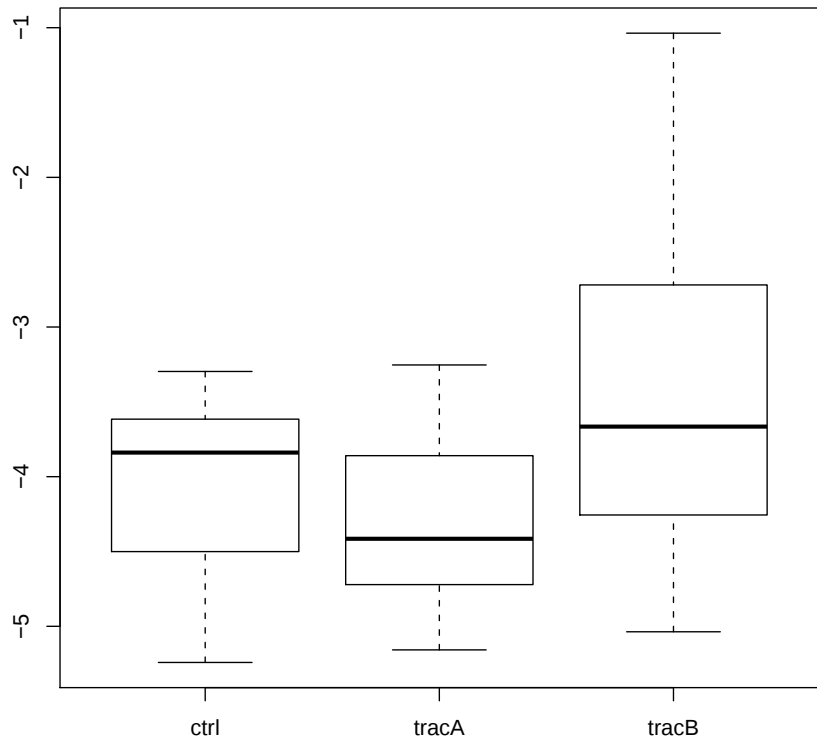
JAK1-Hs00233820_m1 p.value=0.027281



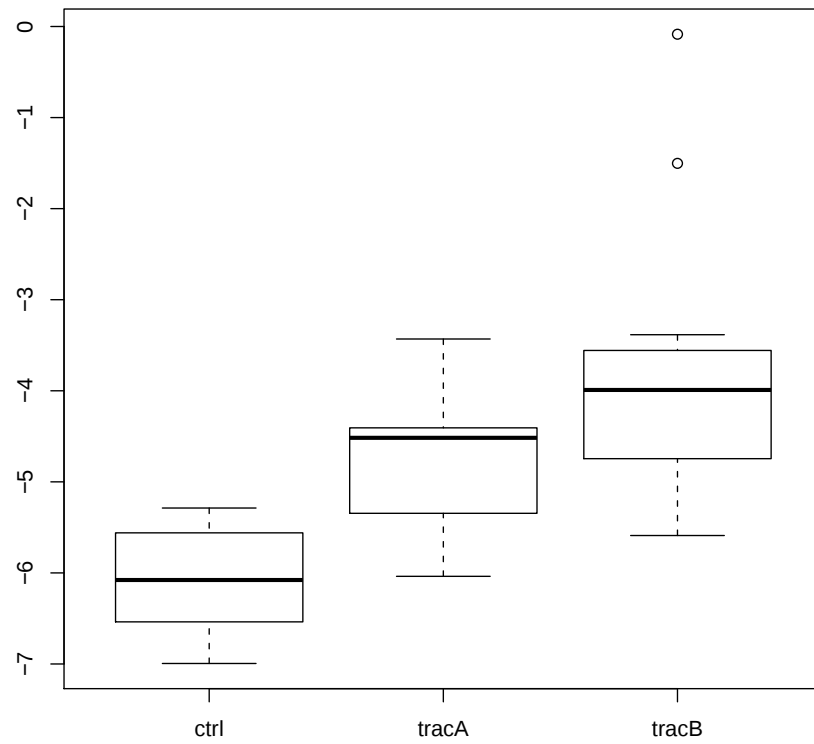
KRAS-Hs00270666_m1 p.value=0.043813



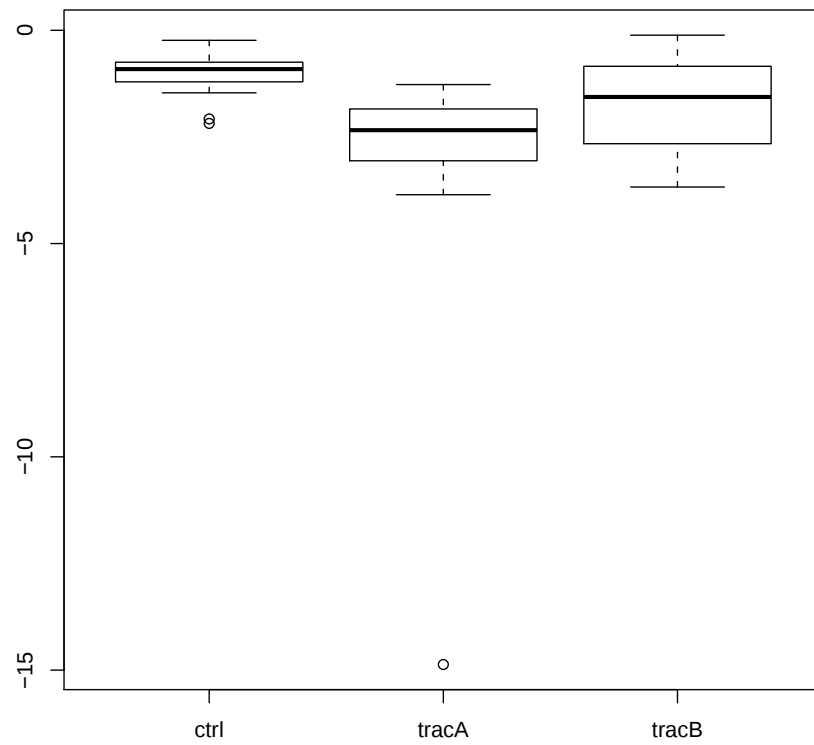
LILRB1-Hs01848117_s1 p.value=0.017737



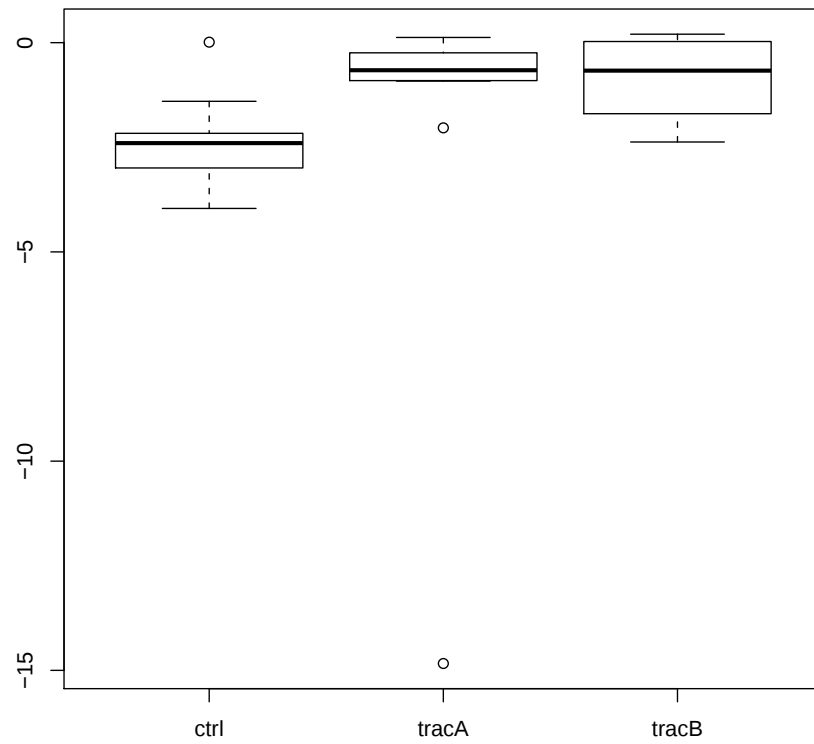
LILRB2-Hs00275975_m1 p.value=0.000117



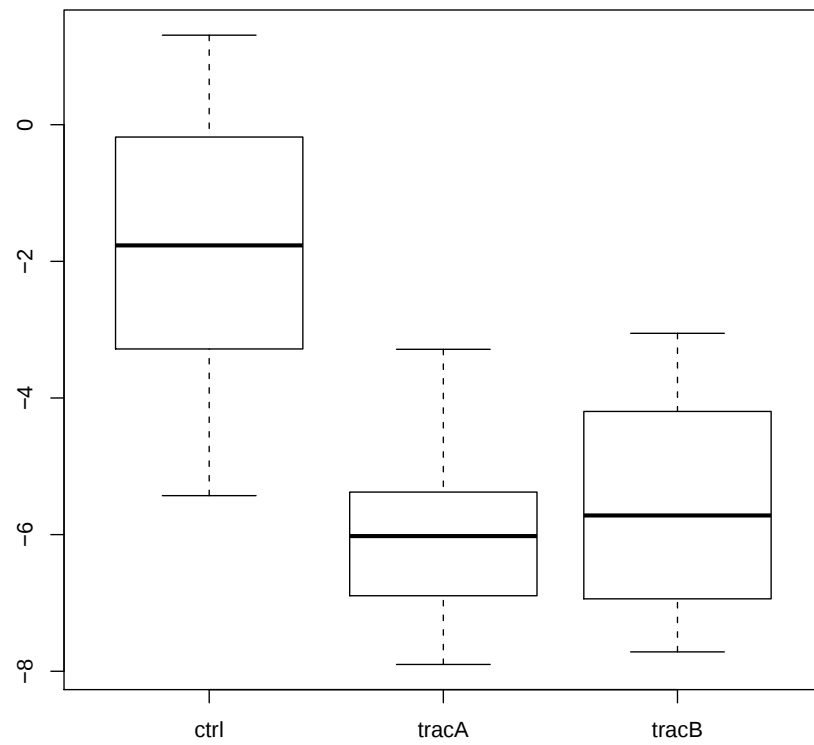
MAP2K6-Hs00177150_m1 p.value=0.047024



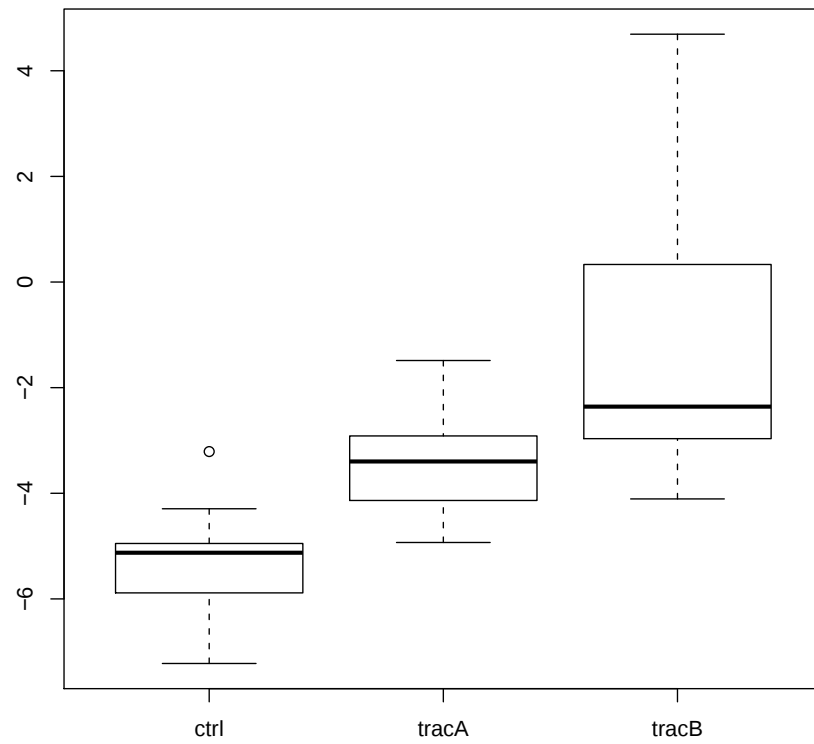
ME1-Hs01554892_m1 p.value=0.002474



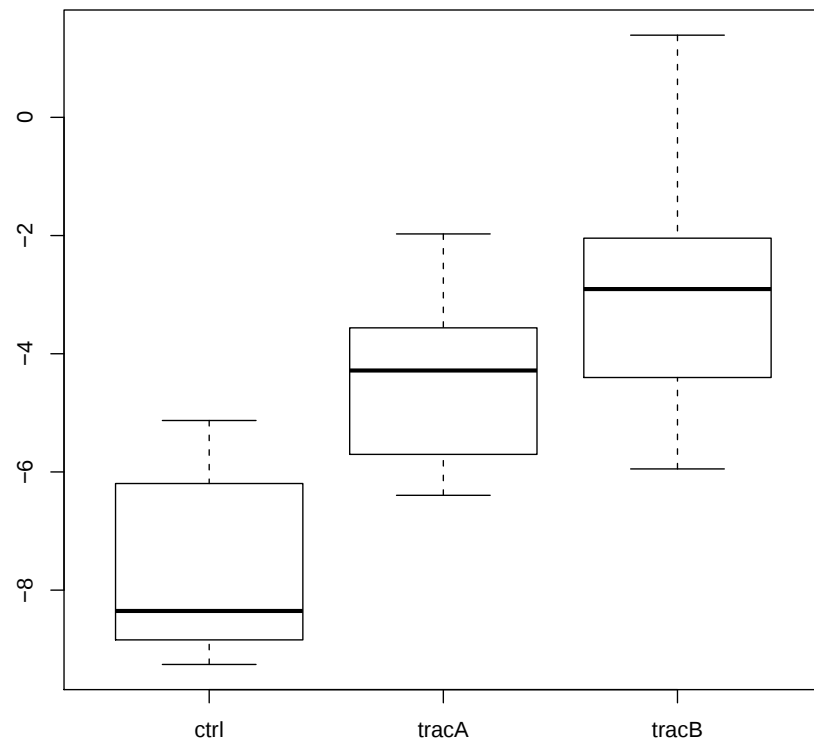
MEP1B-Hs00195535_m1 p.value=1e-06



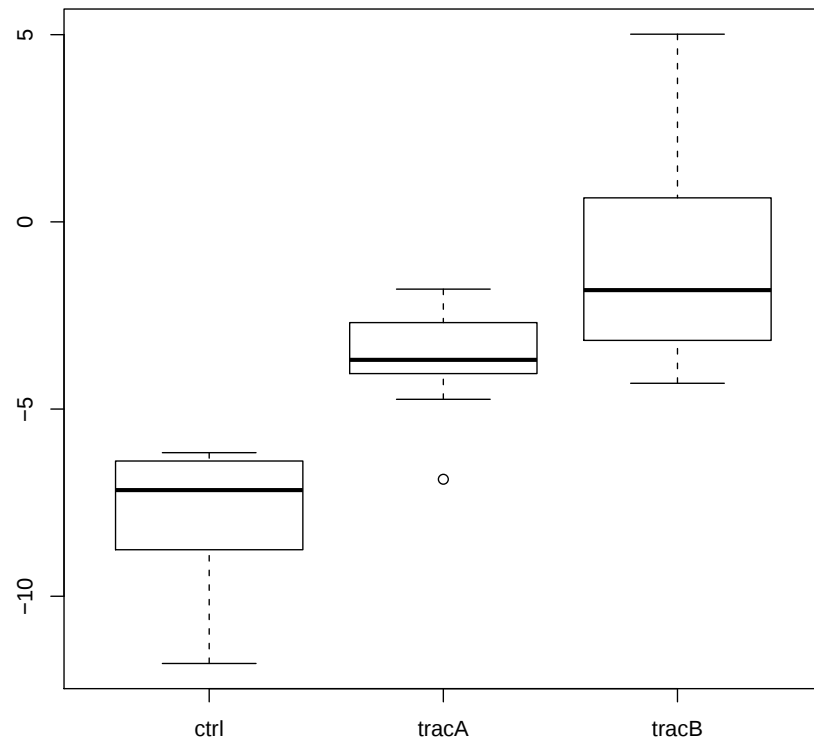
MMP1-Hs00899660_g1 p.value=0.000114



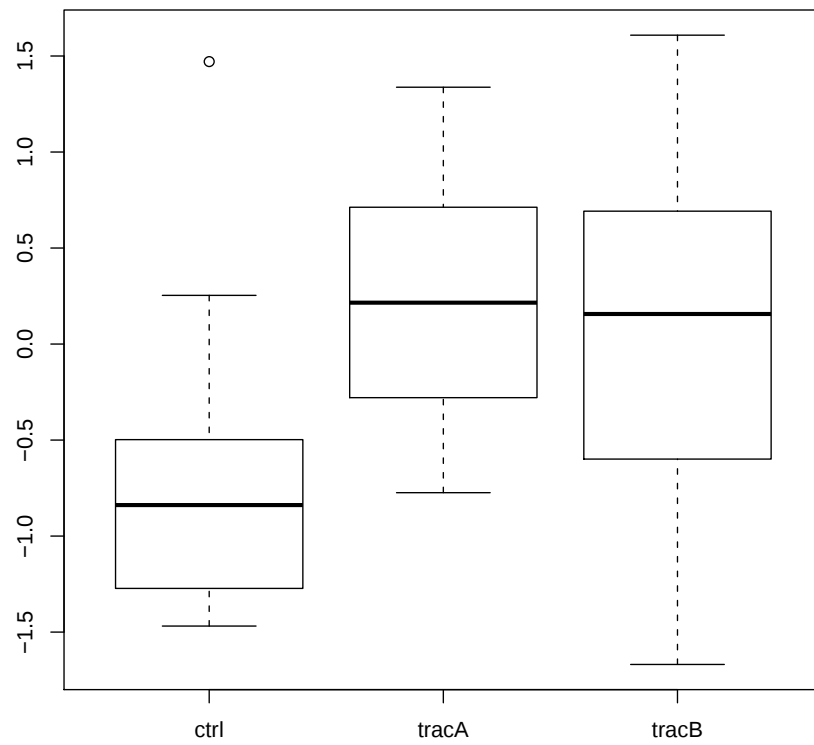
MMP10-Hs00233987_m1 p.value=1.3e-05



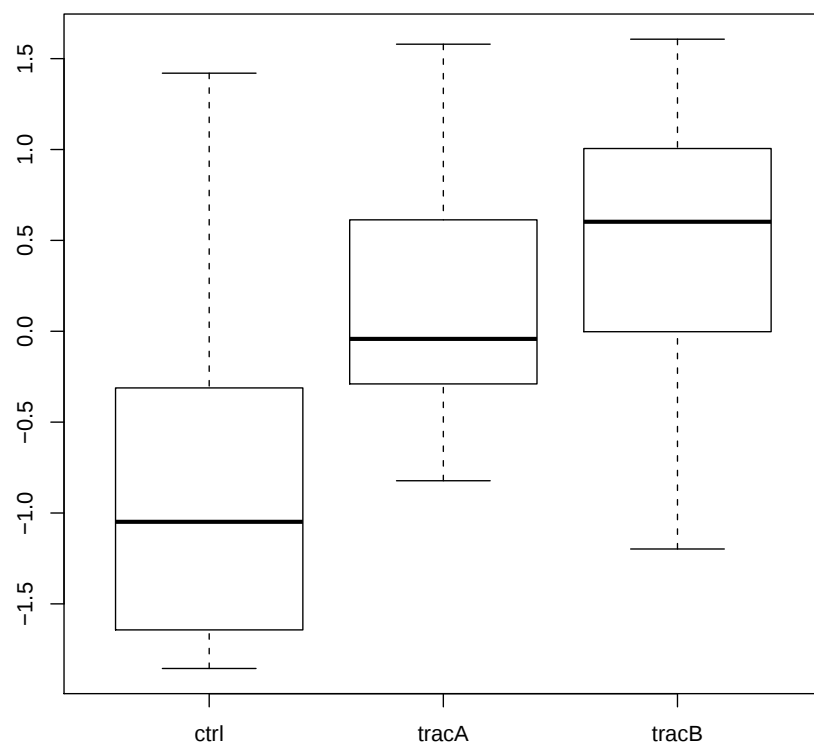
MMP3-Hs00968308_m1 p.value=0



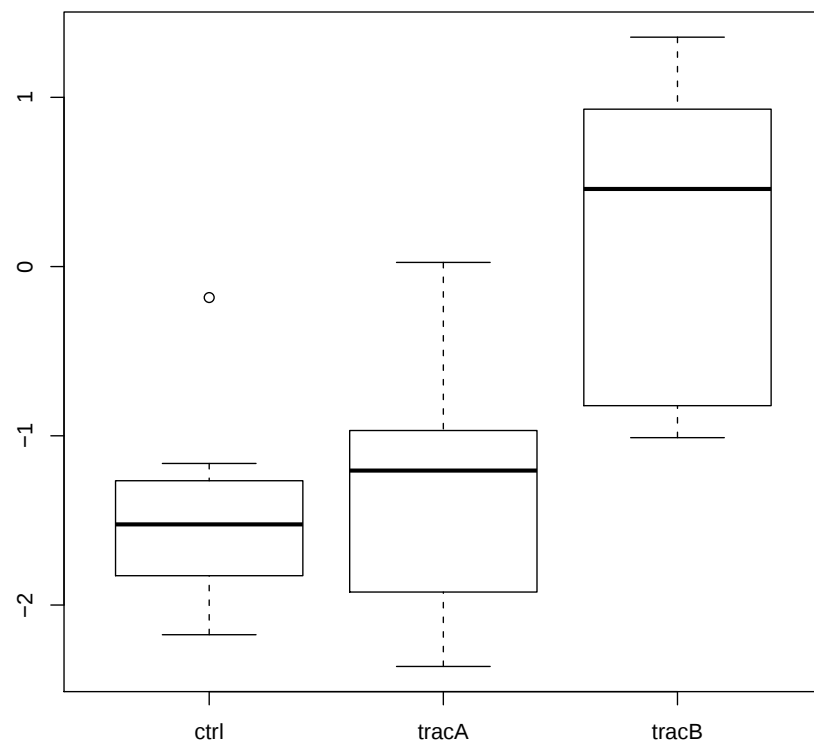
MUC4-Hs00366414_m1 p.value=0.008882



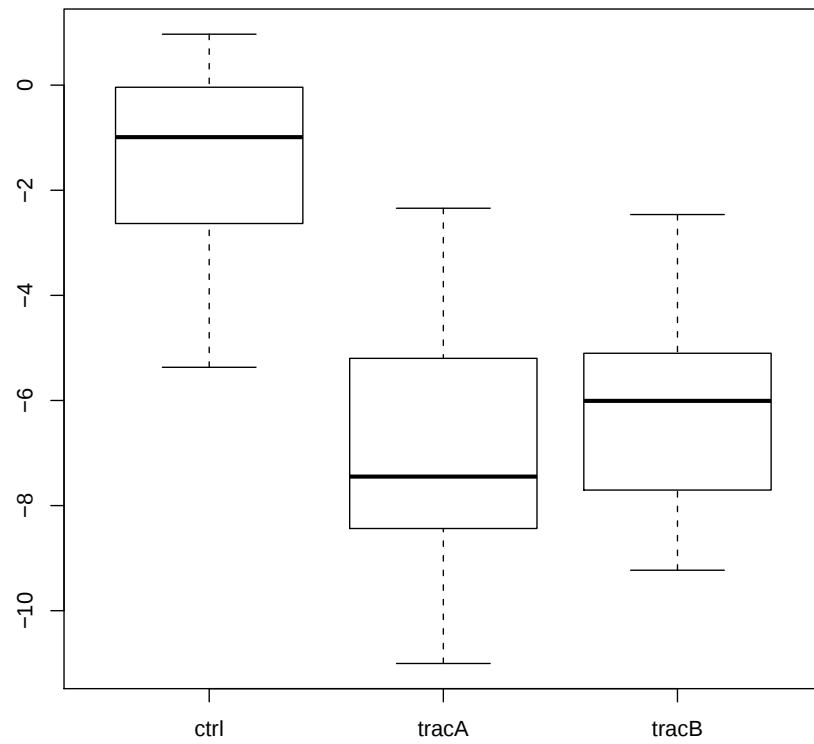
MUC5AC;MUC5B-Hs00861588_m1 p.value=0.003425



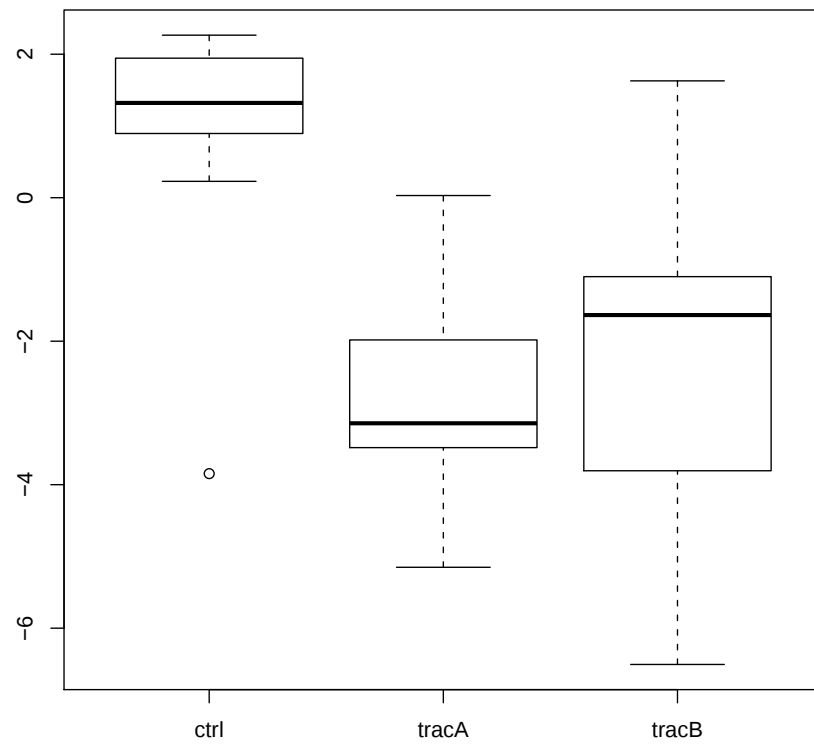
NFKBIZ-Hs00944732_m1 p.value=0.000105



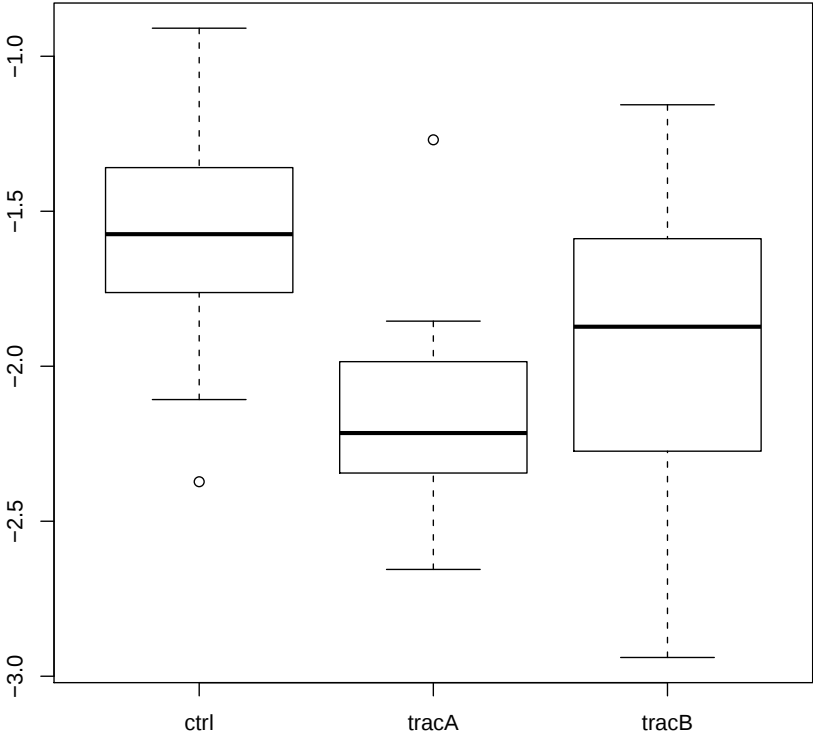
OSTalpha-Hs00380895_m1 p.value=2e-06



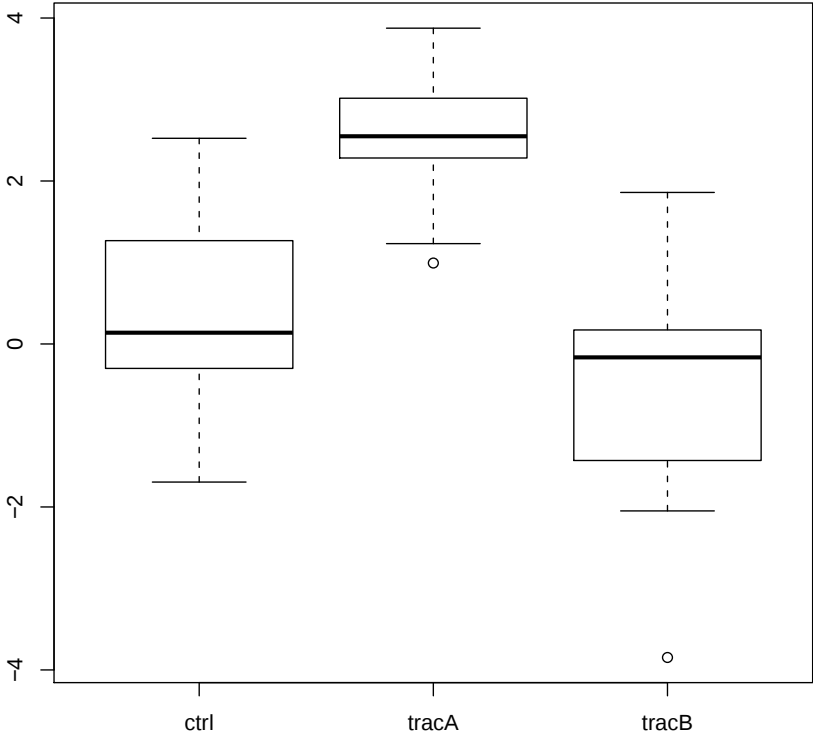
PRAP1-Hs00376974_m1 p.value=3e-06



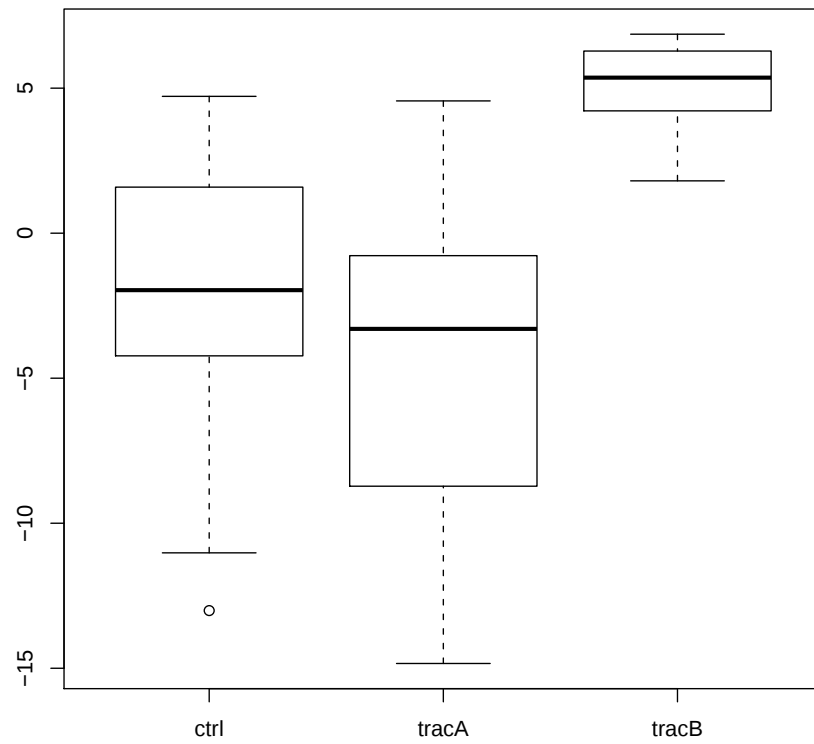
PTGER4-Hs00168761_m1 p.value=0.000864



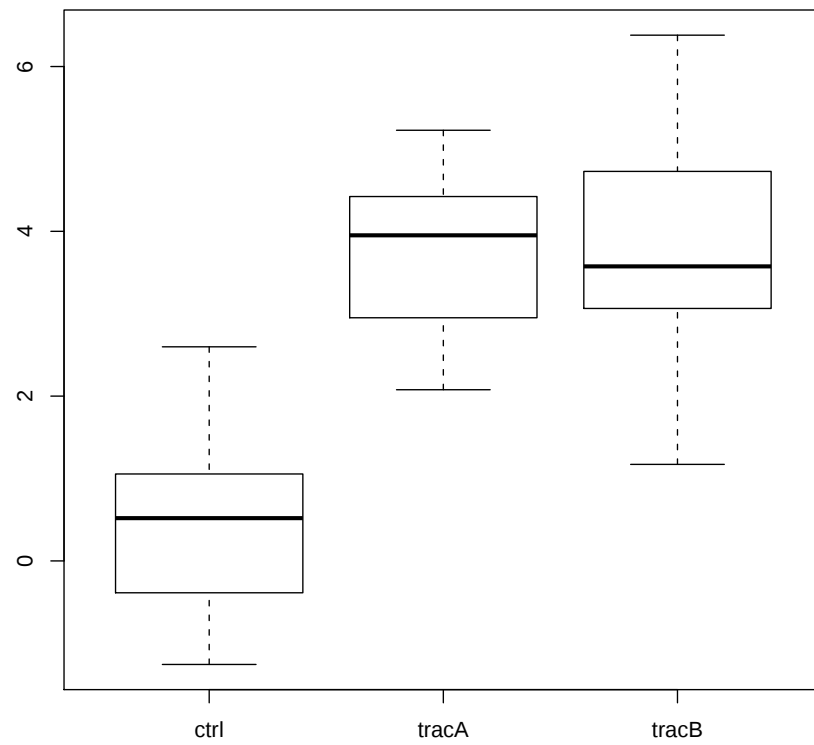
PYY-Hs00373890_g1 p.value=1e-06



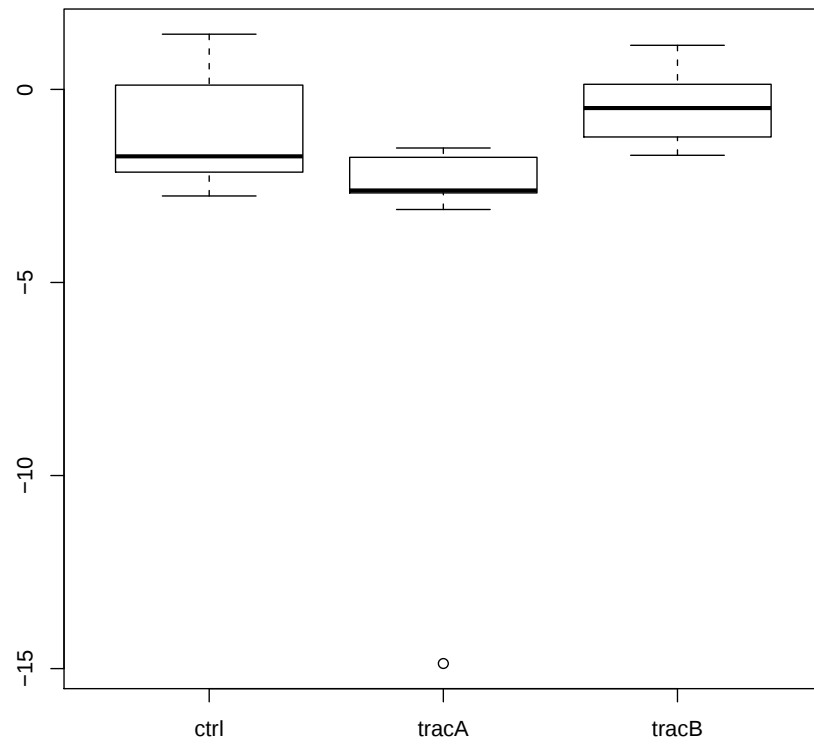
REG1A-Hs00984887_g1 p.value=7e-06



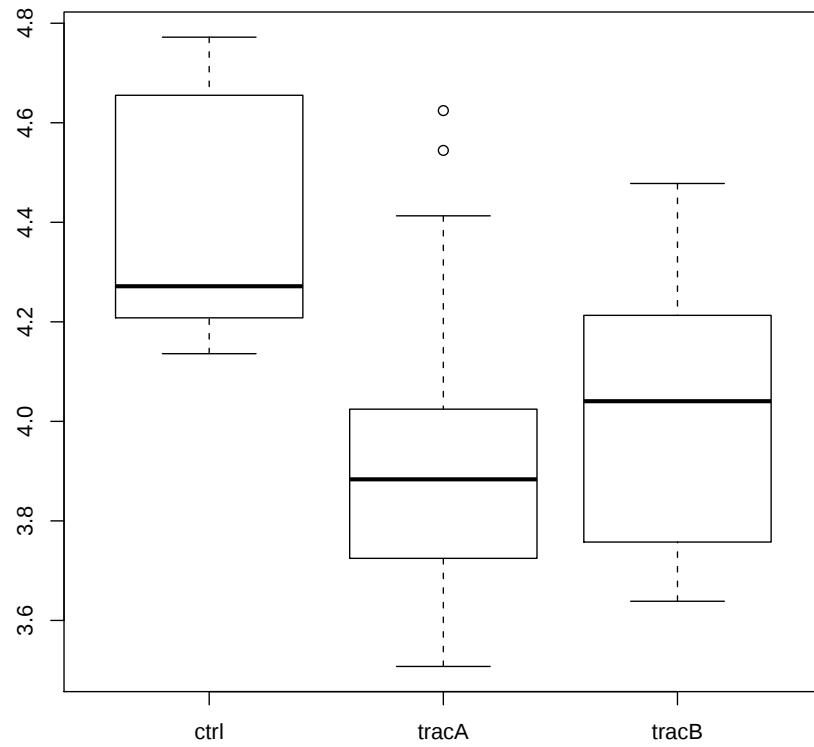
REG4-Hs01069973_m1 p.value=0



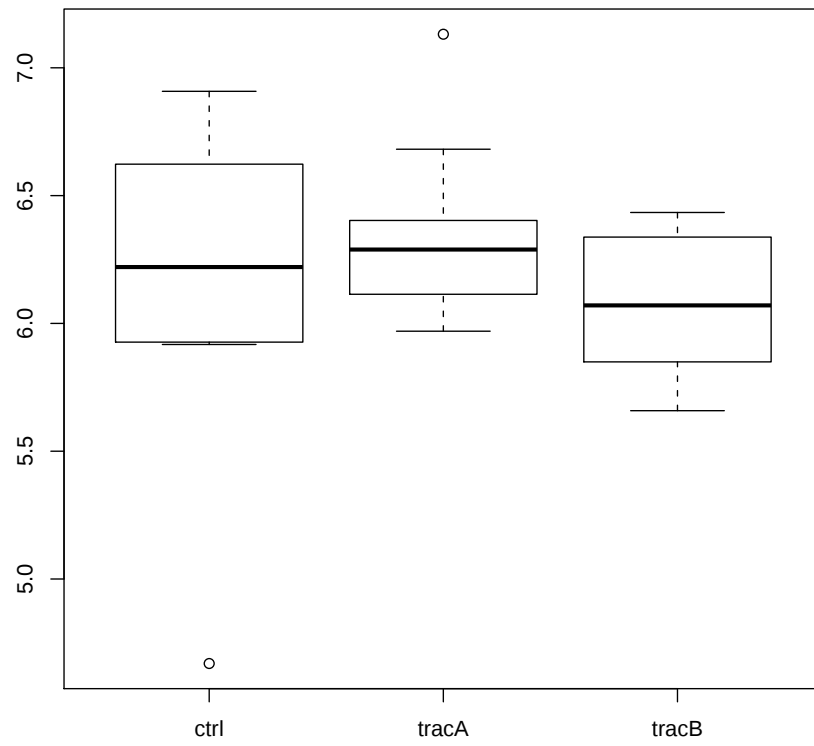
RHOB-Hs00269660_s1 p.value=0.016372



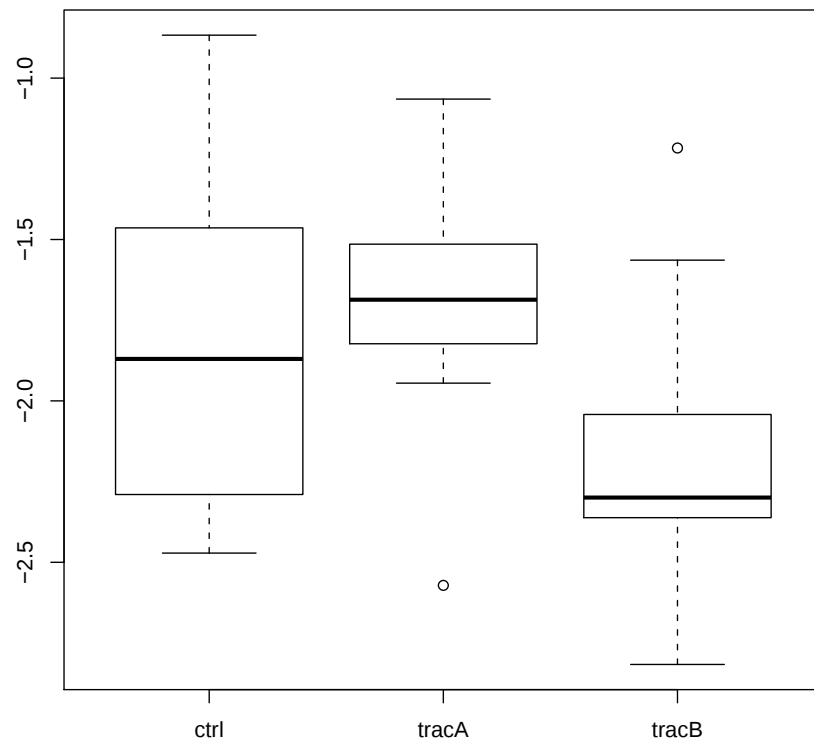
RPLP0-like-Hs00420895_gH p.value=0.000922



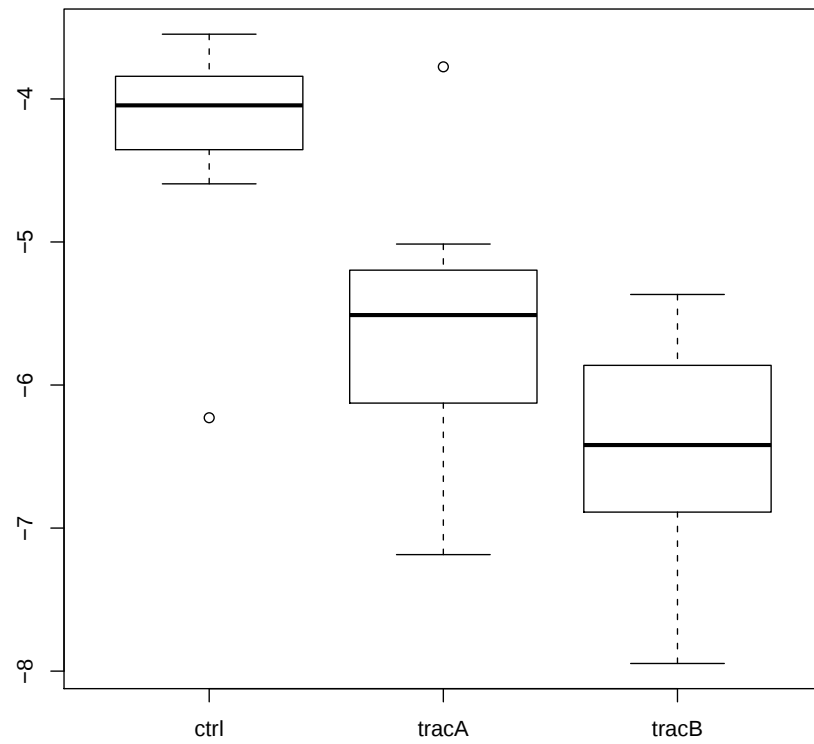
RPS2-Hs03005117_g1 p.value=0.042351



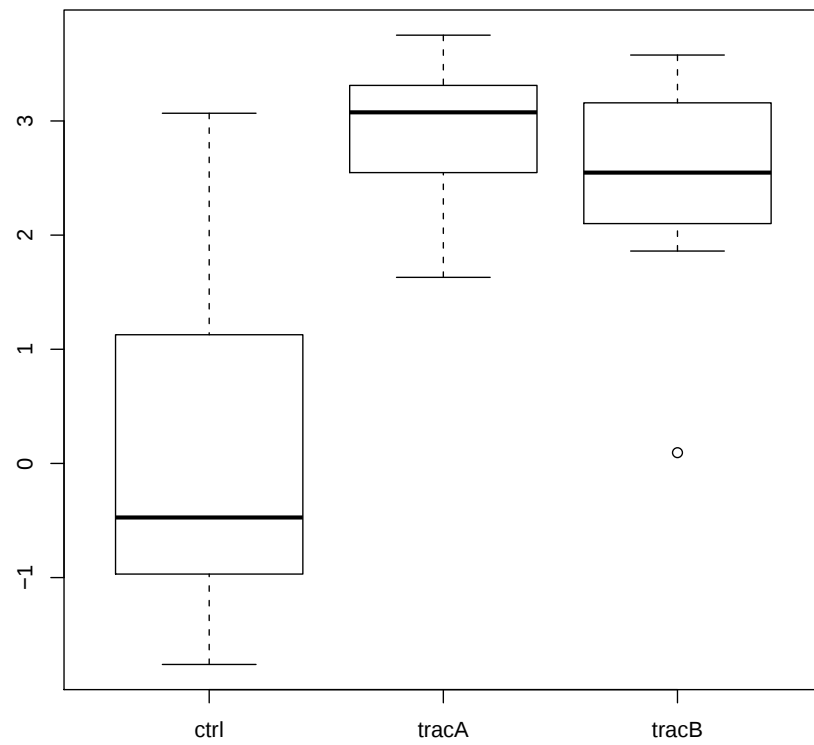
RPS3A-Hs00832893_sH p.value=0.005981



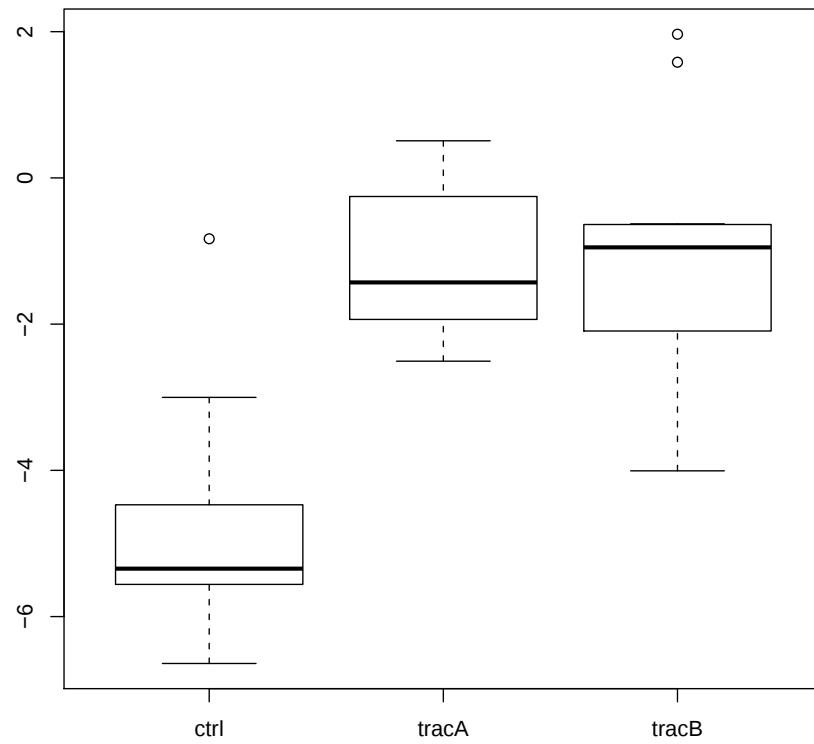
RUNDC3B-Hs00289927_m1 p.value=2e-06



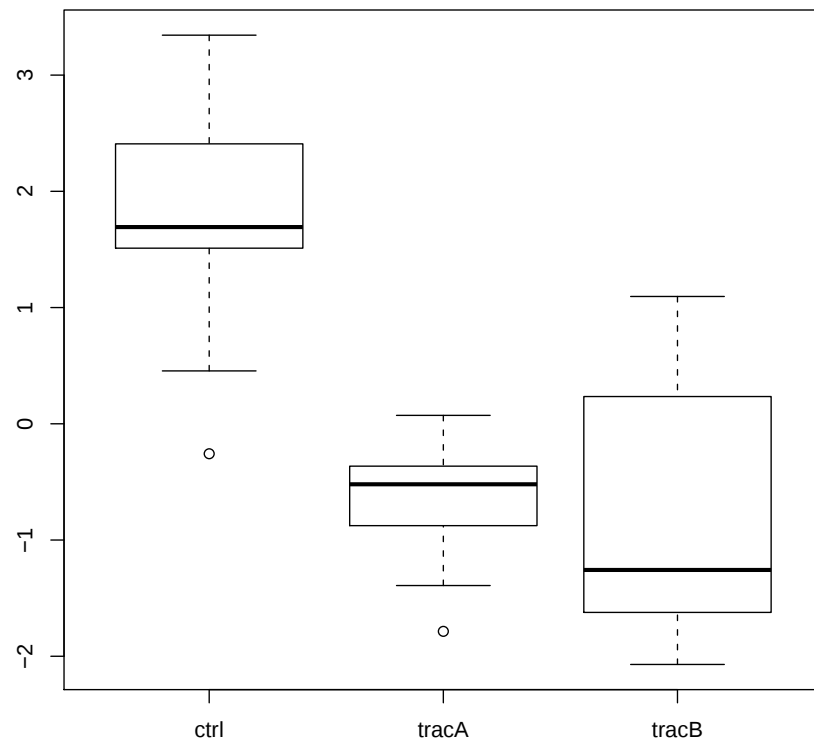
S100P-Hs00195584_m1 p.value=2e-06



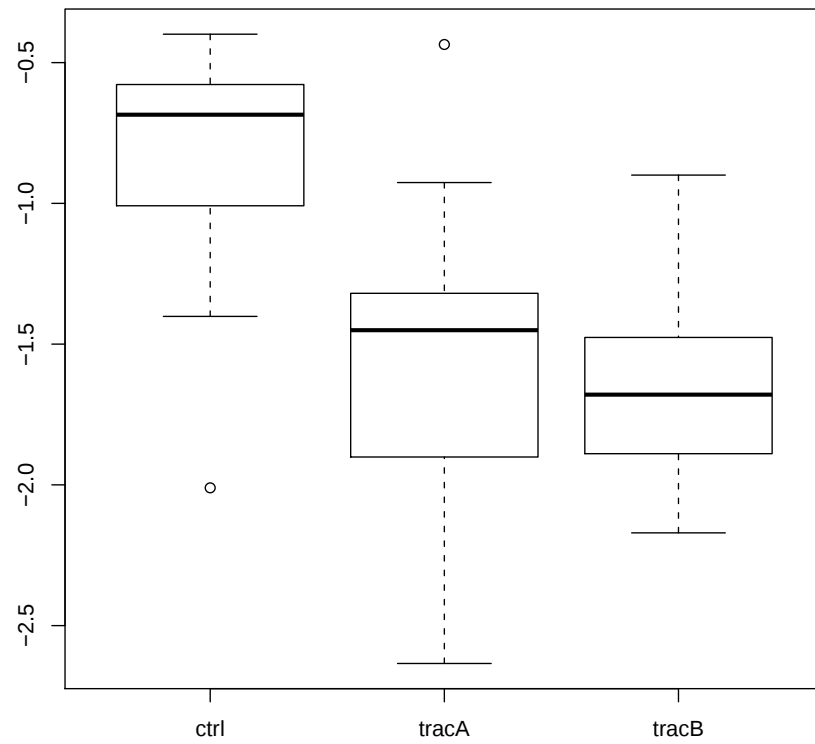
SERPINB5-Hs00985283_m1 p.value=0



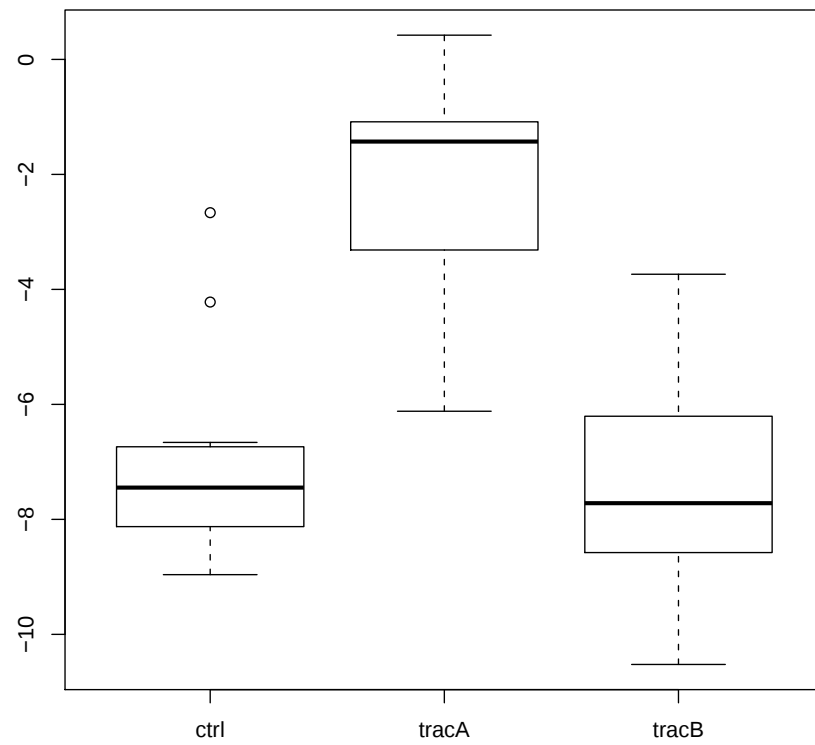
SLC16A1-Hs00161826_m1 p.value=0



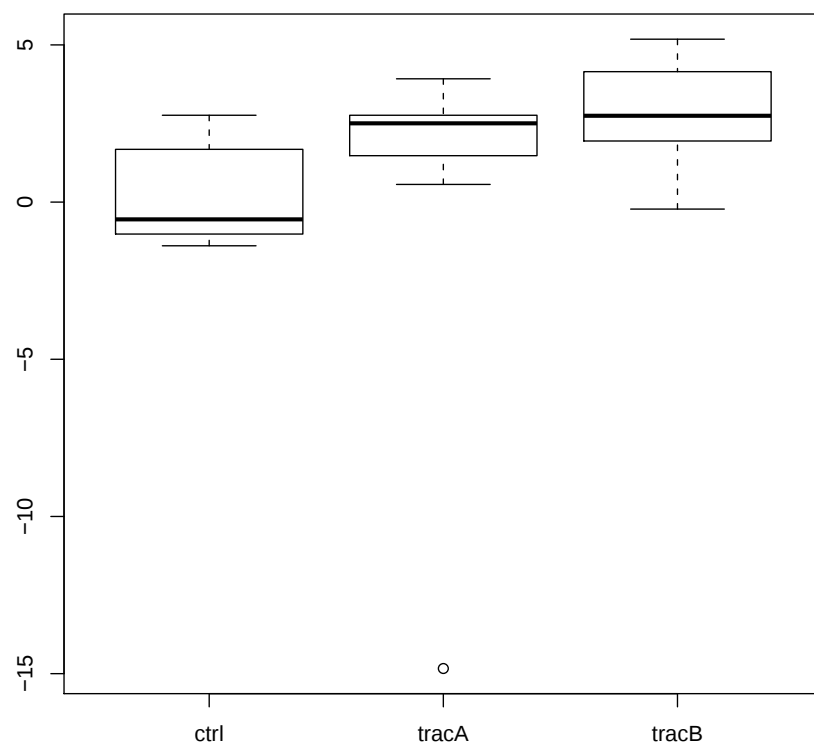
SMAD7-Hs00998193_m1 p.value=0.000508



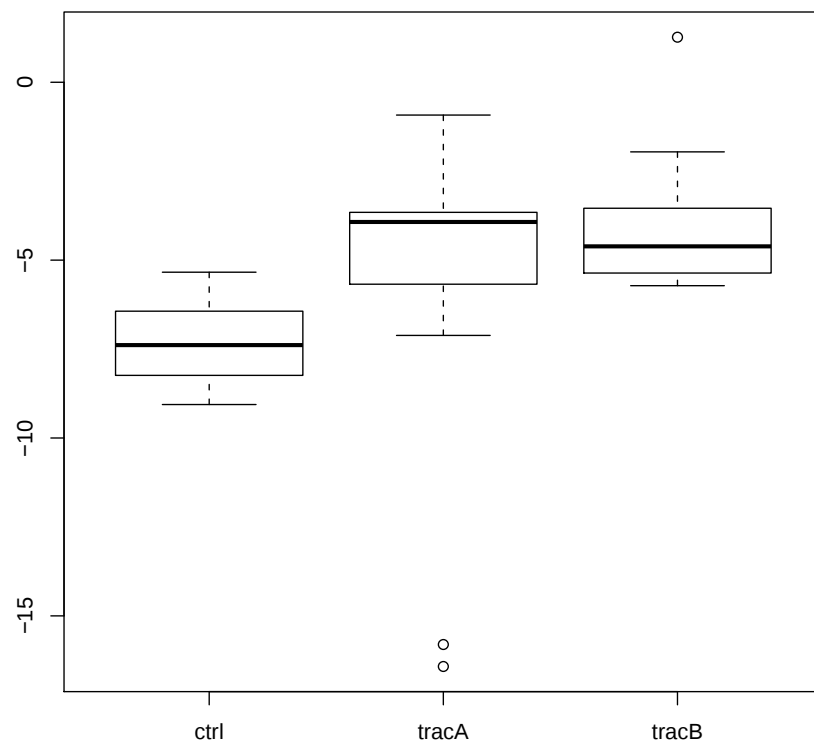
ST6GAL2-Hs00383641_m1 p.value=0



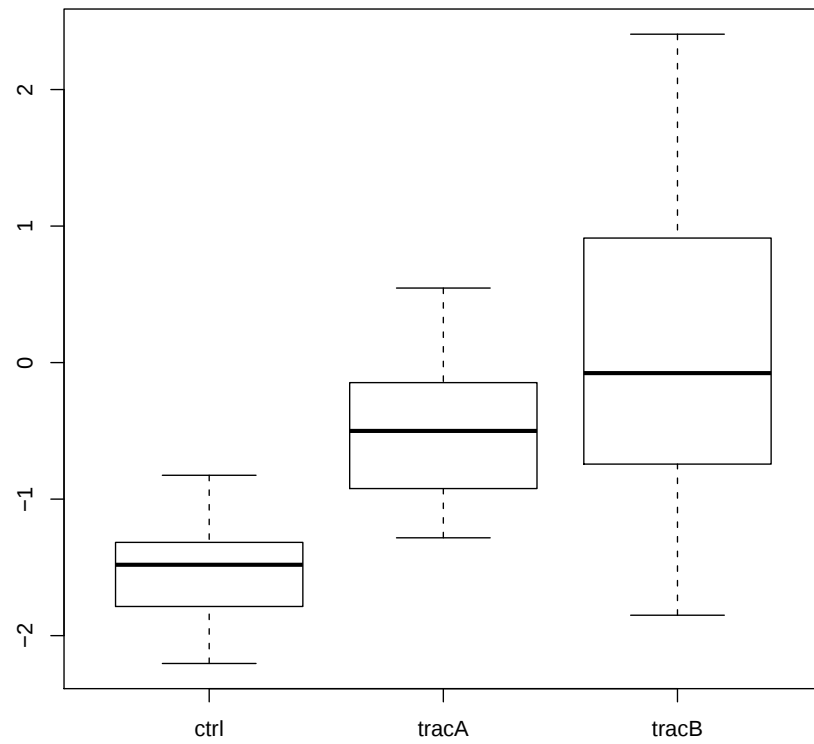
TFF1-Hs00170216_m1 p.value=0.00111



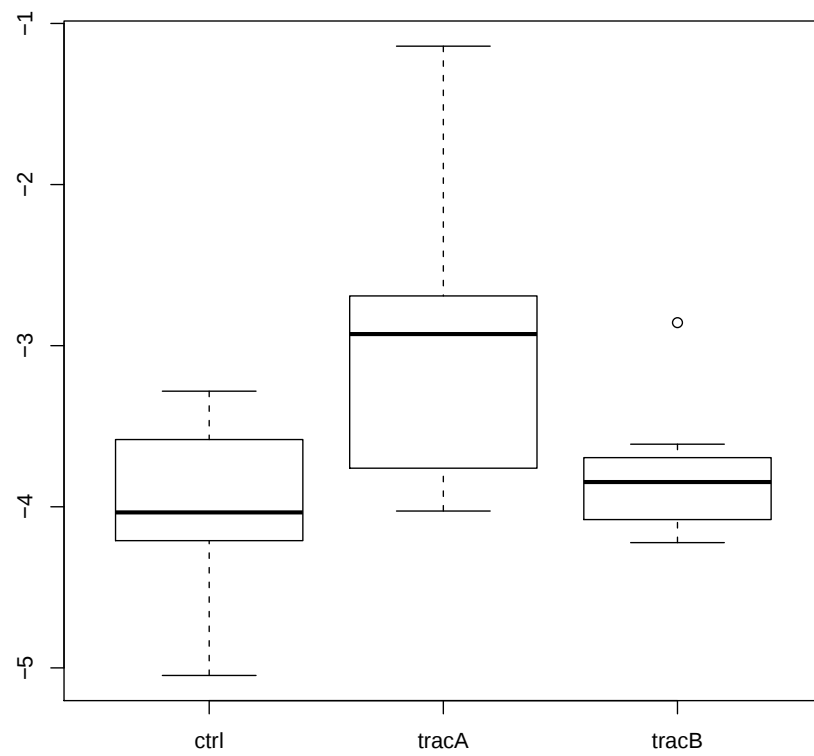
TFPI2-Hs00197918_m1 p.value=0.000181



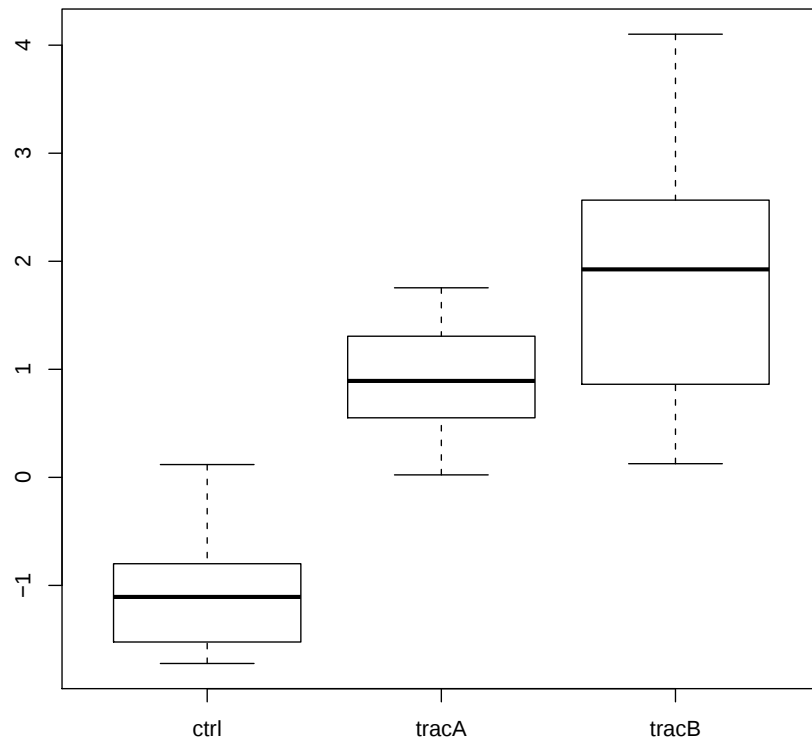
TGFBI-Hs00932734_m1 p.value=2e-05



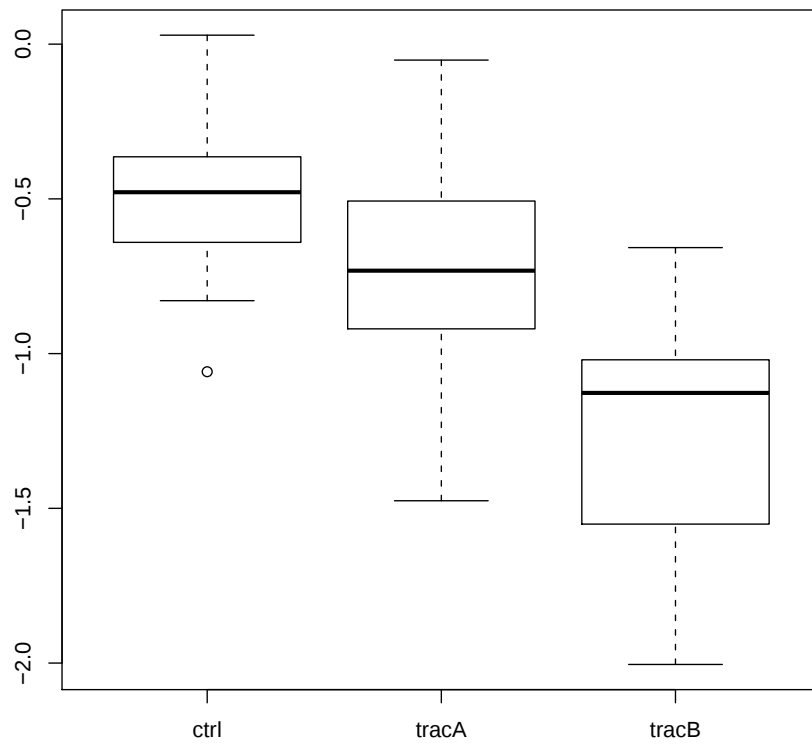
TGFBR3-Hs01114253_m1 p.value=0.001999



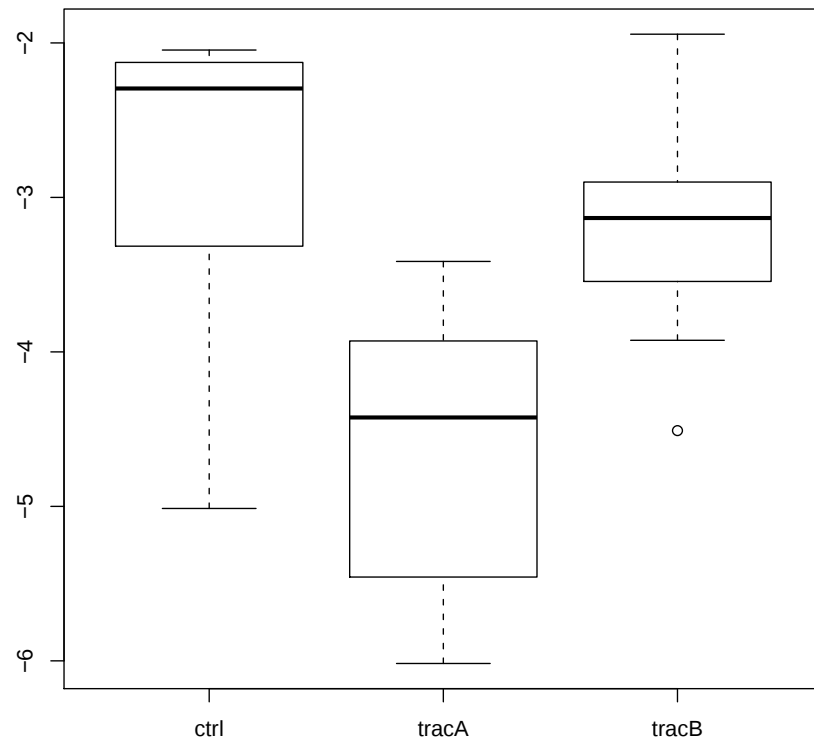
TIMP1-Hs00355335_g1 p.value=0



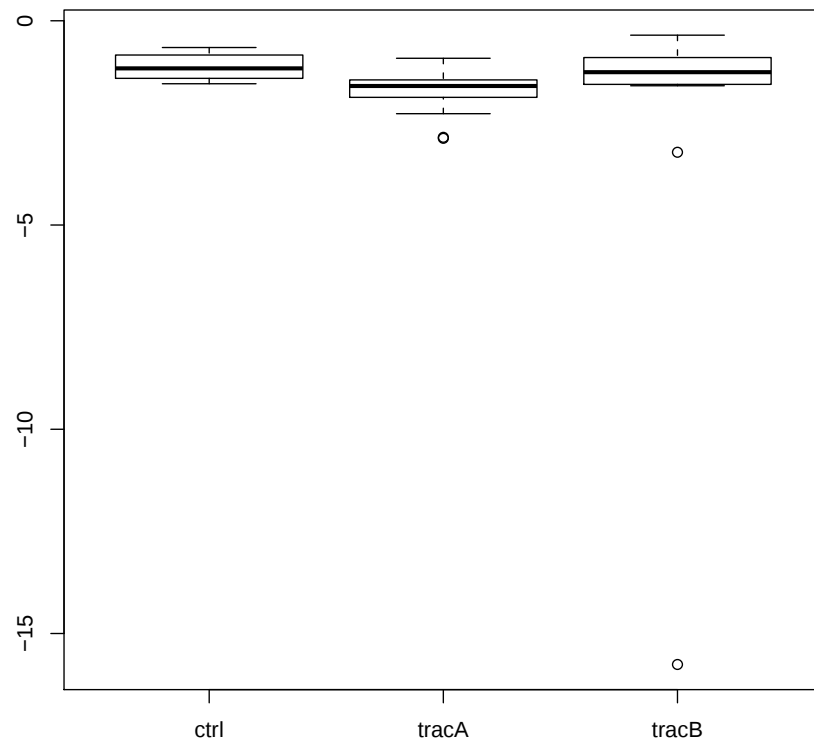
TJP1-Hs00543824_m1 p.value=0.000214



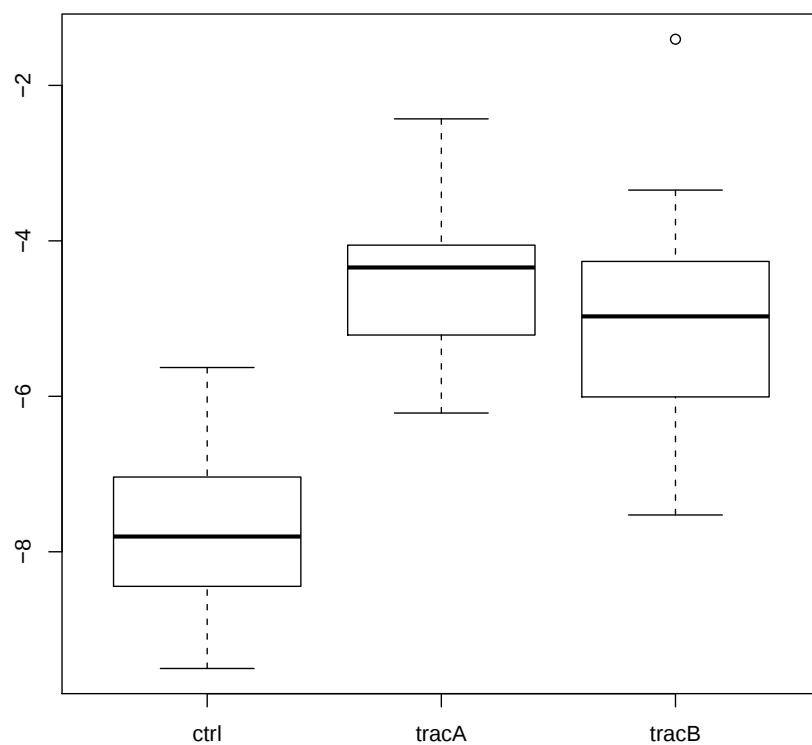
TLR3-Hs01551078_m1 p.value=3e-05



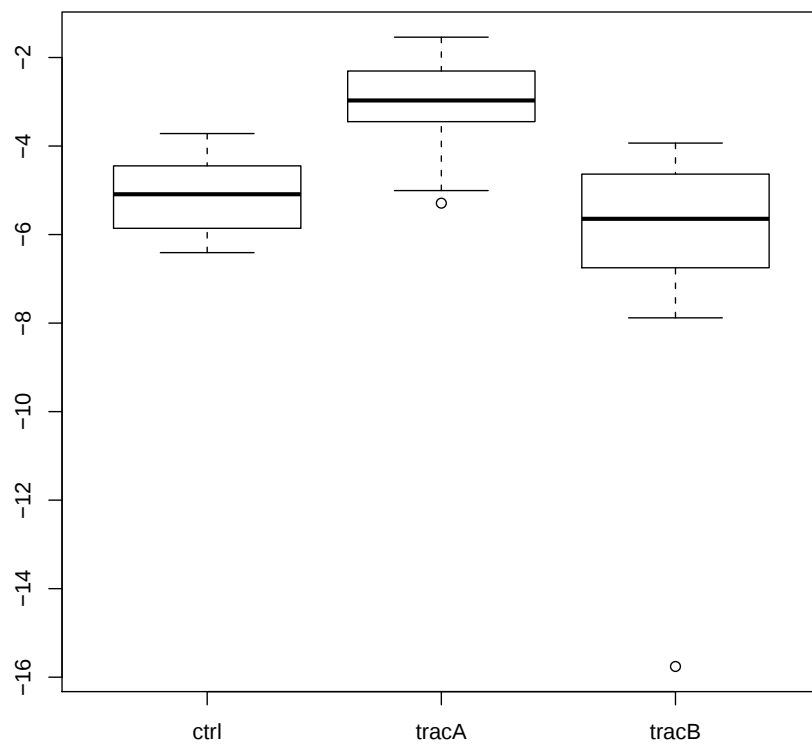
TOP2A-Hs00172214_m1 p.value=0.004041



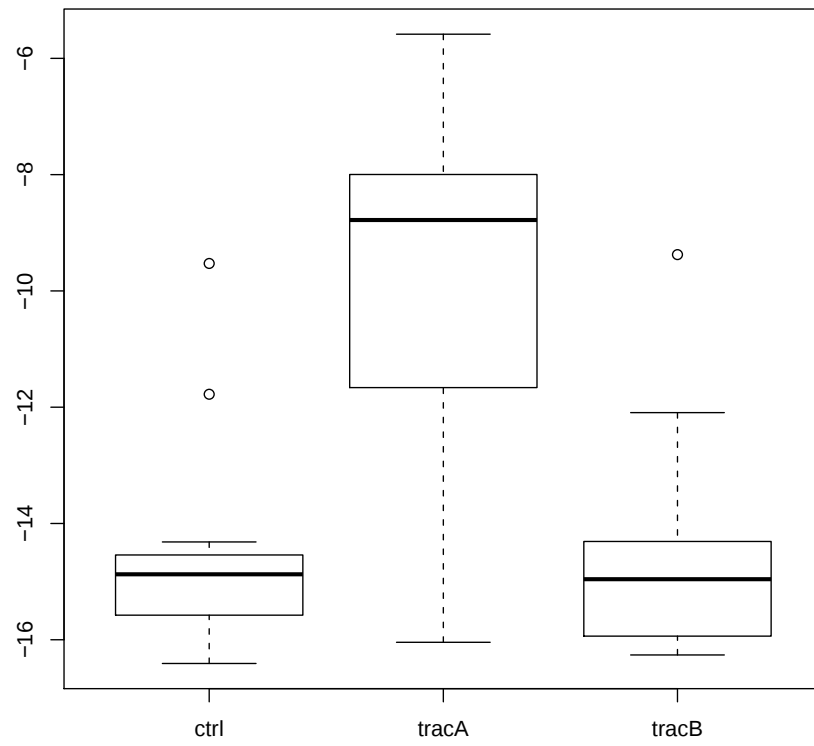
TRIM29-Hs00232590_m1 p.value=0



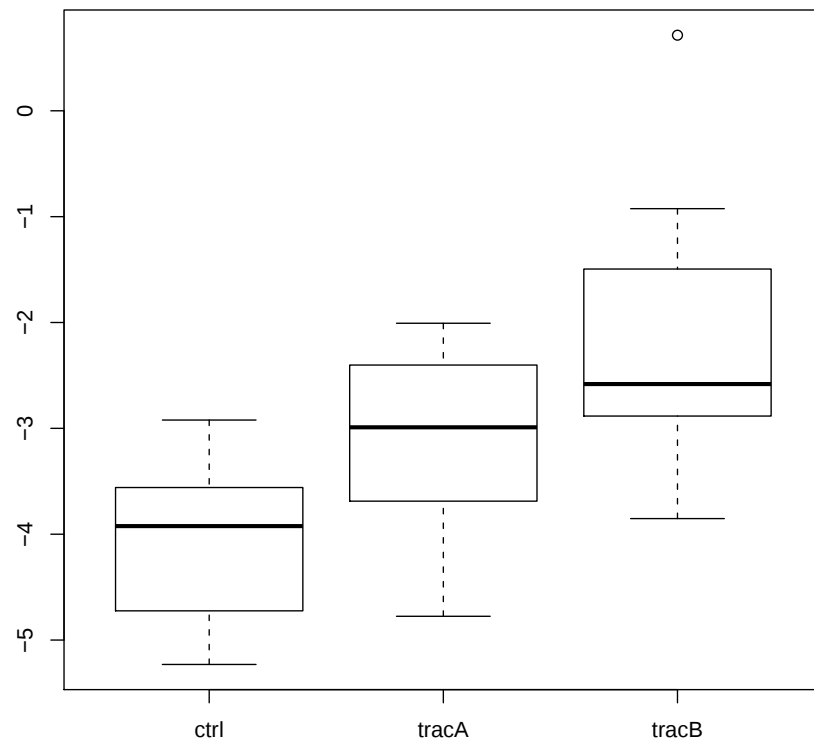
TSLP-Hs00263639_m1 p.value=3.3e-05



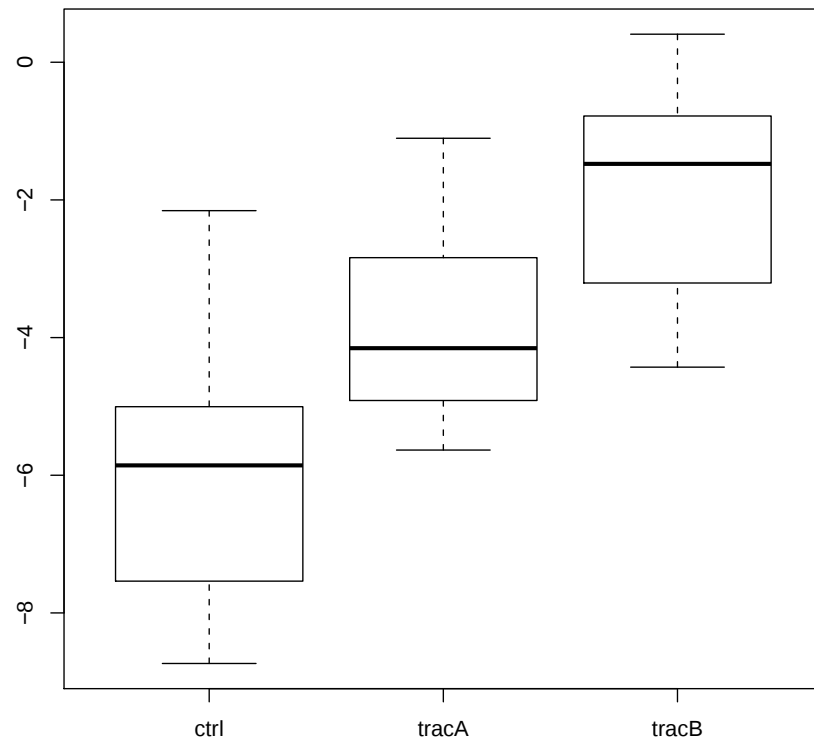
TSLP-Hs01572933_m1 p.value=0.000666



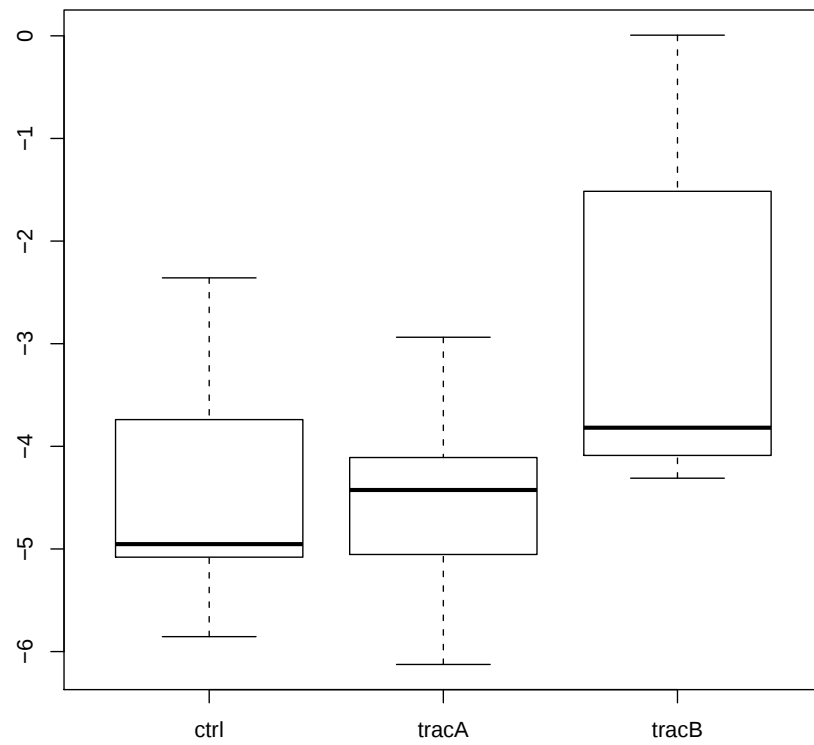
VCAN-Hs01007937_m1 p.value=0.000602



VNN1-Hs01546812_m1 p.value=4.1e-05



WNT5A-Hs00998537_m1 p.value=0.002709



References

- Kenneth J. Livak and Thomas D. Schmittgen** Analysis of Relative Gene Expression Data Using Real-Time Quantification PCR and the $2^{-\Delta\Delta C_T}$ Method. *Methods* 25, 402-408 (2001)
- Michael W.Pfaffl** A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, 29 (2001)
- Jo Vandesompele et al.** Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* (2003)
- Michael W.Pfaffl, Graham W.Horgan and Leo Dempfle** Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research* (2002).