Bioconductor Expression Assessment Tool for Affymetrix Oligonucleotide Arrays (affycomp)

June 8, 2011

Contents

In this report only assessment using the **HGU95** spike-In experiment are presetned. Figures 2,3, and 4b are therefore ommited.

```
fullname: marcelo boareto
method: RMA
nickname: RMA
competition: YES

Overall signal to ratio assessment:

        slope        R2   medianSD Null FC IQR Null FC 99.9% Rank
RMA 0.1007737 0.7506361 0.03437035  0.05449604     0.2453500  440

Expected Rank (out of 12626) for genes with fold change of 2
when all other genes are not differentially expressed
stratified by nominal concentration.

          RMA
0.25:0   1197
0.5:0.25  915
1:0.5    1180
2:1       340
4:2       195
8:4       203
16:8      203
32:16     264
64:32     381
128:64    695
```

```
256:128  1032
512:256  1756
1024:512 2761
```

Table 1 for Spike-in Data:

	RMA
Signal detect slope	0.10192586
Signal detect R2	0.75975517
AUC (FP<10)	0.40034113
AUC (FP<15)	0.45604424
AUC (FP<25)	0.51199295
AUC (FP<100)	0.63690890
AFP, call if fc>2	0.26131687
ATP, call if fc>2	1.64403292
IQR	0.13783135
Obs-intended-fc slope	0.09998496
Obs-(low)int-fc slope	0.08335309
FC=2, AUC (FP<10)	0.06554383
FC=2, AUC (FP<15)	0.06682478
FC=2, AUC (FP<25)	0.06962569
FC=2, AUC (FP<100)	0.09563200
FC=2, AFP, call if fc>2	0.07142857
FC=2, ATP, call if fc>2	0.82142857

Table 2 for Spike-in Data:

	RMA
null log-fc IQR	0.05449604
null log-fc 99%	0.17009338
null log-fc 99.9%	0.24535005
low AUC	0.07093131
med AUC	0.04262067
high AUC	0.00000000
weighted avg AUC	0.06342744
25% SD	0.02337805
Median SD	0.03437035
75% SD	0.04923020
99% SD	0.08250992
low.slope	0.08045477
med.slope	0.12634259
high.slope	0.05850531
low.R2	0.08213982

med.R2	0.34814304
high.R2	0.31519196
0.25:0	0.06287037
0.5:0.25	0.07317255
1:0.5	0.06355409
2:1	0.11027123
4:2	0.13051978
8:4	0.12929395
16:8	0.12915143
32:16	0.11964612
64:32	0.10597424
128:64	0.08365833
256:128	0.06868465
512:256	0.04898530
1024:512	0.03236327

Figure 1

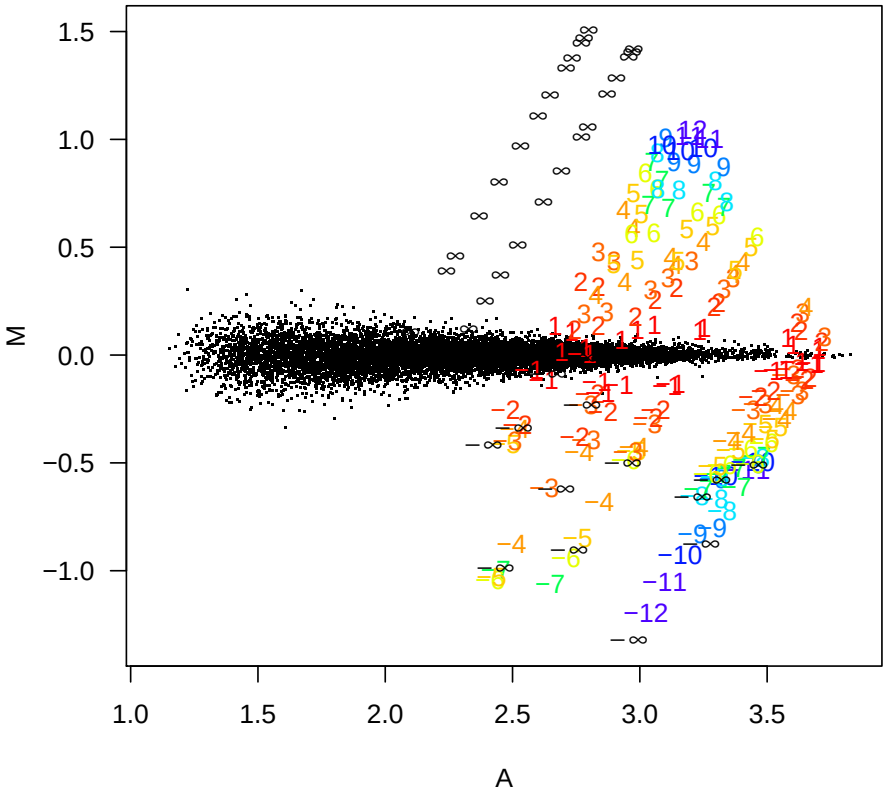

Figure 1: The MA plot shows log fold change as a function of mean log expression level. A set of 14 arrays representing a single experiment from the Affymetrix spike-in data are used for this plot. A total of 13 sets of fold changes are generated by comparing the first array in the set to each of the others. Spiked-in genes are symbolized by numbers representing the nominal $\log_2$ fold change for the gene. Non-differentially expressed genes with observed fold changes larger than 2 are plotted in red. All other probesets are represented with black dots.

Figure 1b

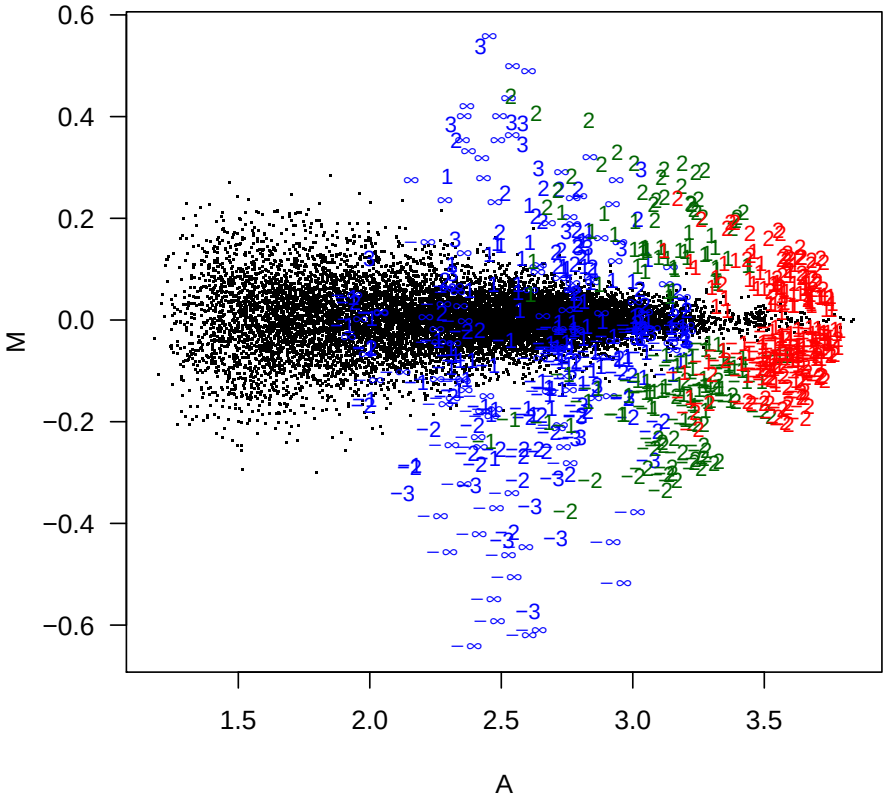

Figure 1b: The MA plot shows log fold change as a function of mean log expression level. A set of 28 arrays representing a single experiment from the Affymetrix spike-in data are used for this plot. Fold changes are generated for all possible comparisons of the the first 14 arrays and the second 14 arrays. Spiked-in genes are symbolized by numbers representing the nominal $\log_2$ fold change for the gene. Of the genes that are spiked to be differentially expressed, only genes with small nominal fold changes are shown. The colors represent four different groups: nominal concentration of genes being compared less than or equal to 2 picoMolar (blue), between 4 and 32 picoMolar (green), greater than or equal to 64 picoMolar (blue). Non-differentially expressed genes with observed fold changes larger than 2 are plotted in red. All other probesets are represented with black dots.

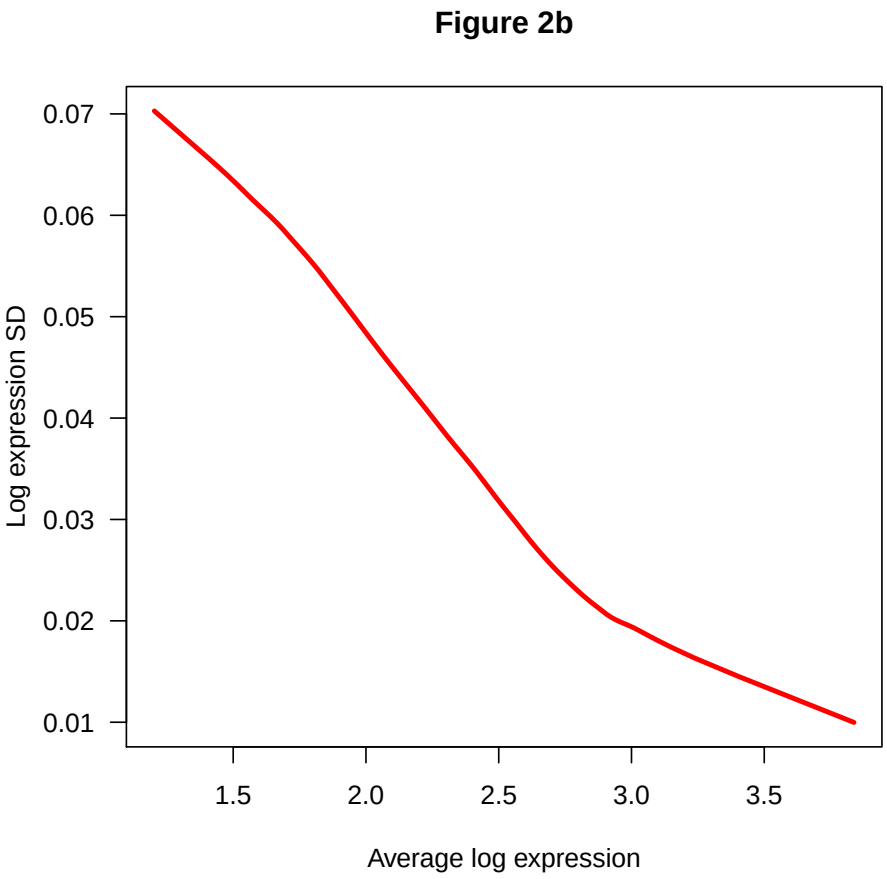

Figure 2b: For each non-spiked-in gene in the 28 arrays used in Figure 1b, we calculate the mean log expression and the observed standard deviation across the 28 replicates. The resulting scatterplot is smoothed to generate a single curve representing mean standard deviation as a function of mean log expression.

Figure 4a

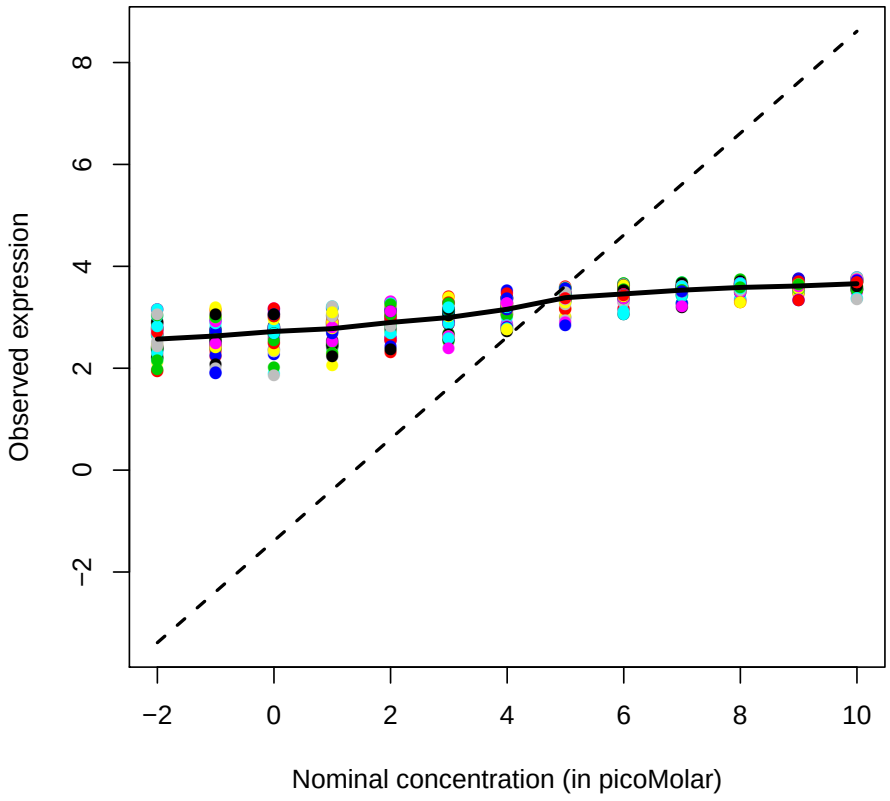

Figure 4a) Average observed log_2 intensity plotted against nominal log_2 concentration for each spiked-in gene for all arrays in Affymetrix spike-In experiment. The dashed line has the ideal slope of 1.

Figure 4c

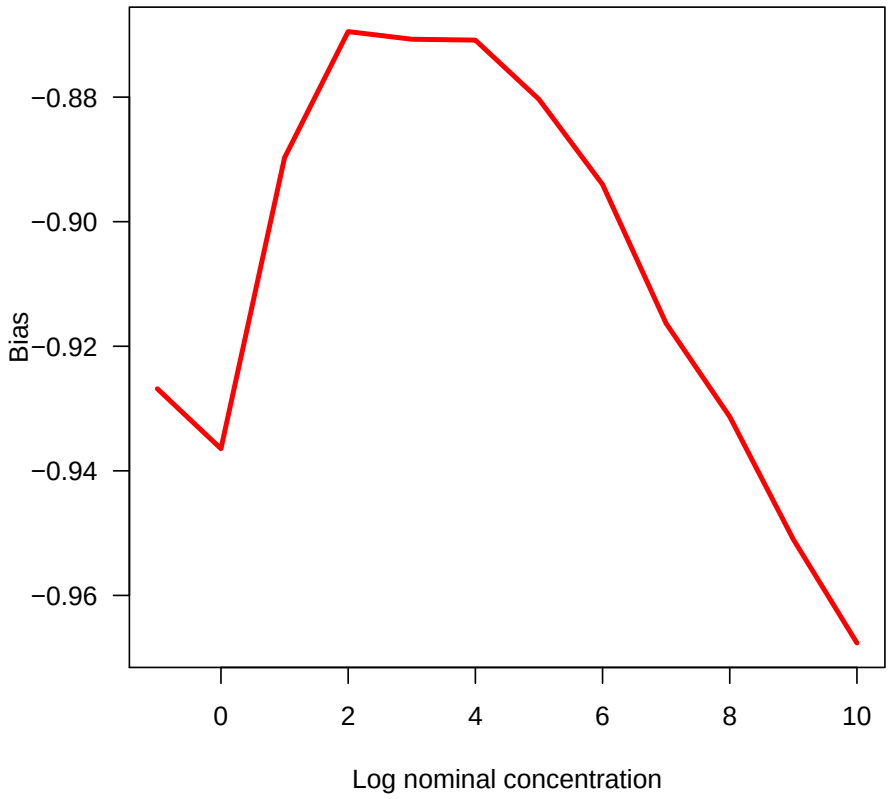

Figure 4c) Using the 28 arrays of Figure 1b, we compute local slopes. As the slopes shown in Figure 4a), the local slopes represent the expected observed log fold-change for probesets with true fold-change of 2 but they are presented as a function of the total nominal probeset concentration in the two samples being compared. In theory the local slopes should be one so we show the bias (difference between the observed local slope and one).

Figure 5a

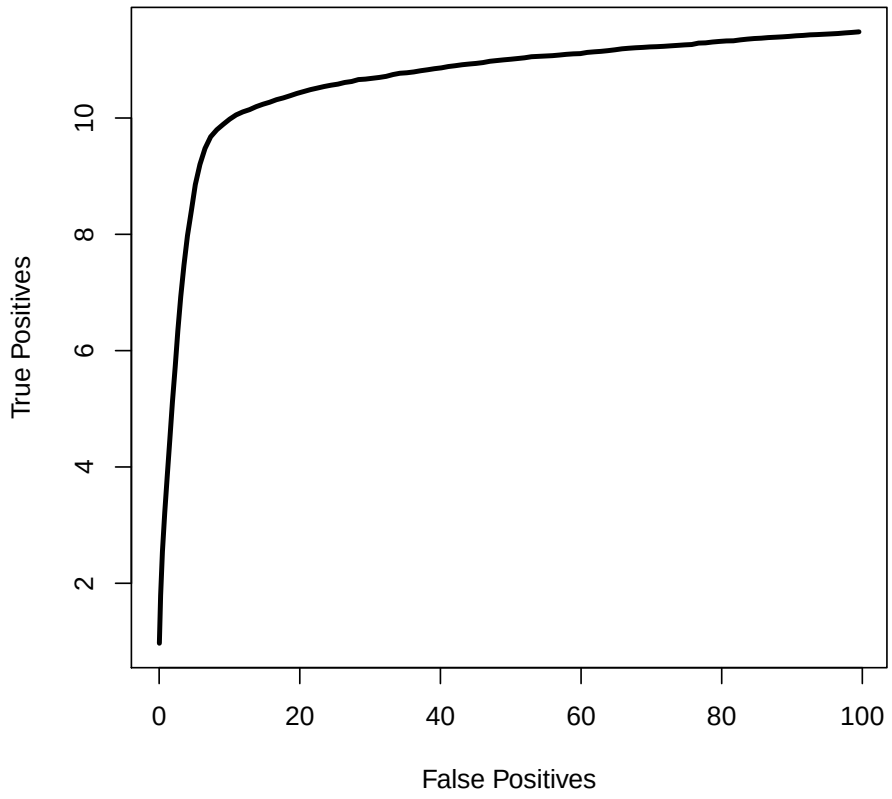

Figure 5a) A typical identification rule for differential expression filters genes with fold change exceeding a given threshold. This figure shows average ROC curves which offer a graphical representation of both specificity and sensitivity for such a detection rule. Average ROC curves based on comparisons with nominal fold changes ranging from 2 to 4096.

Figure 5b

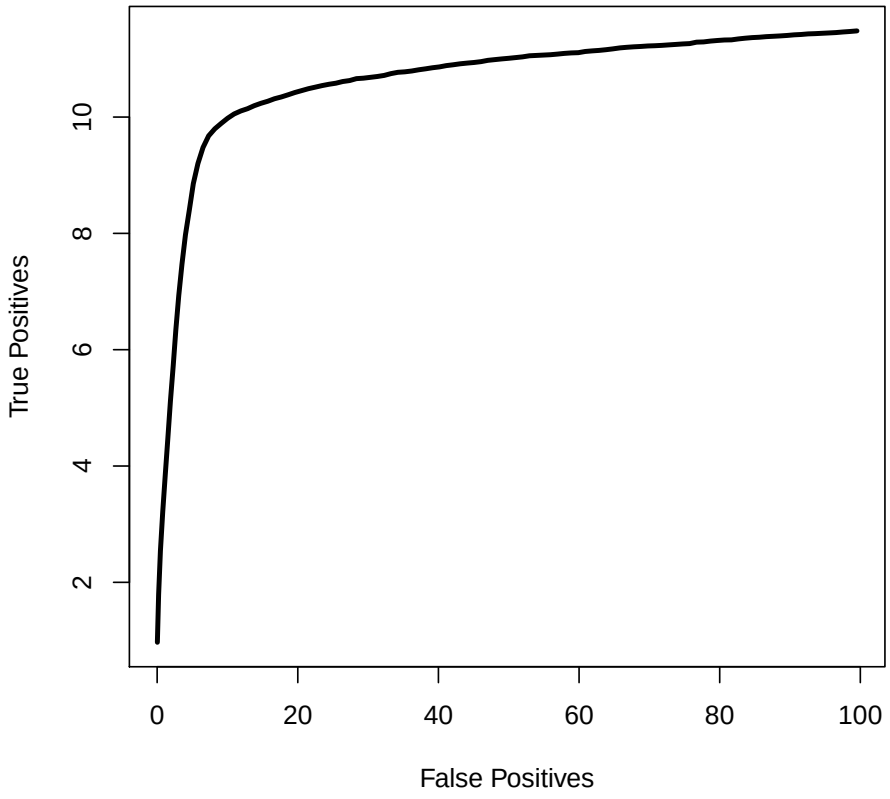

Figure 5b) As 5a) but with nominal fold changes equal to 2.

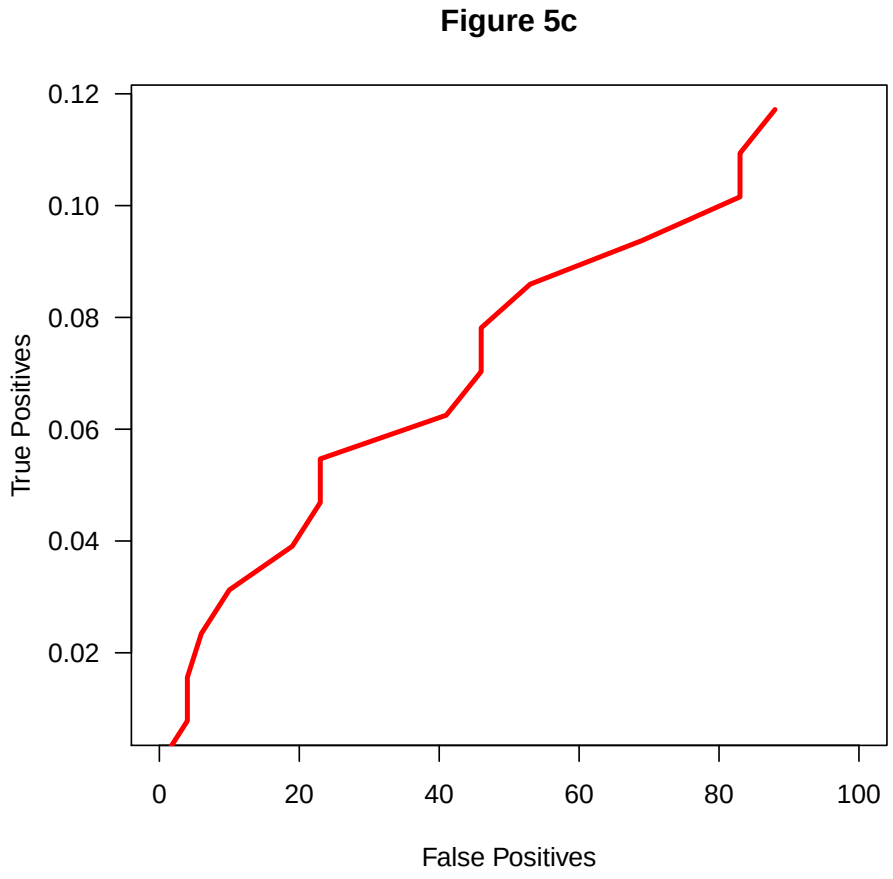

Figure 5c) As 5a) but for comparisons with both nominal concentrations at most 4 picoMolar and nominal fold changes at most 2.

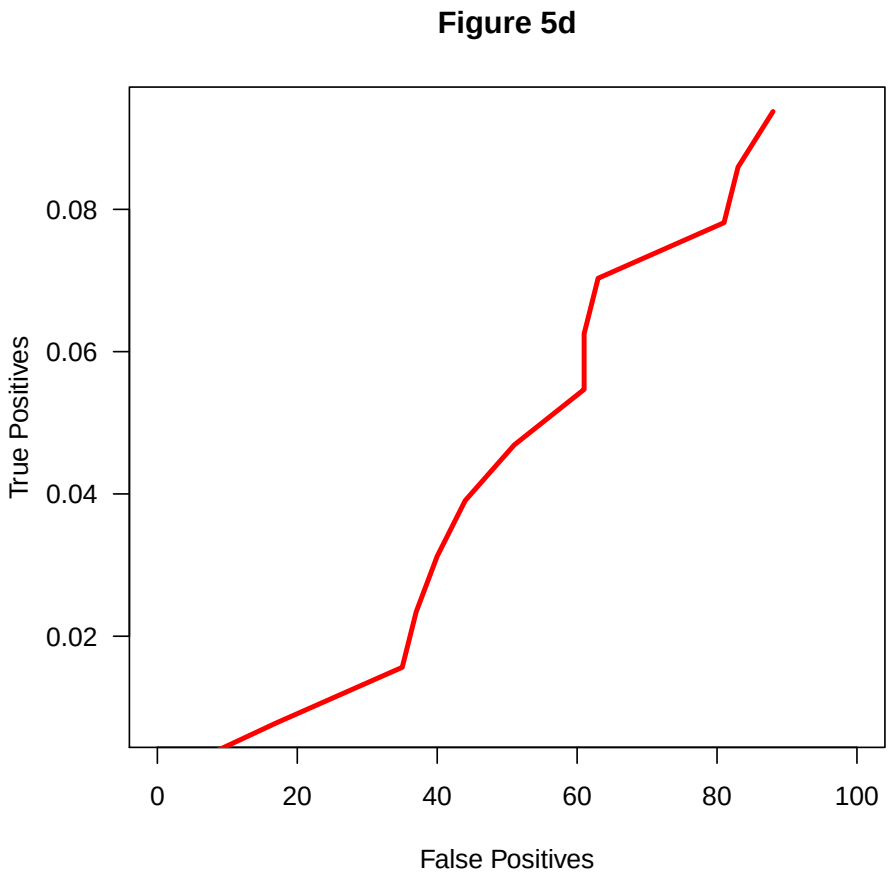

Figure 5d) As 5a) but for comparisons with both nominal concentrations between 4 and 64 picoMolar and nominal fold changes at most 2.

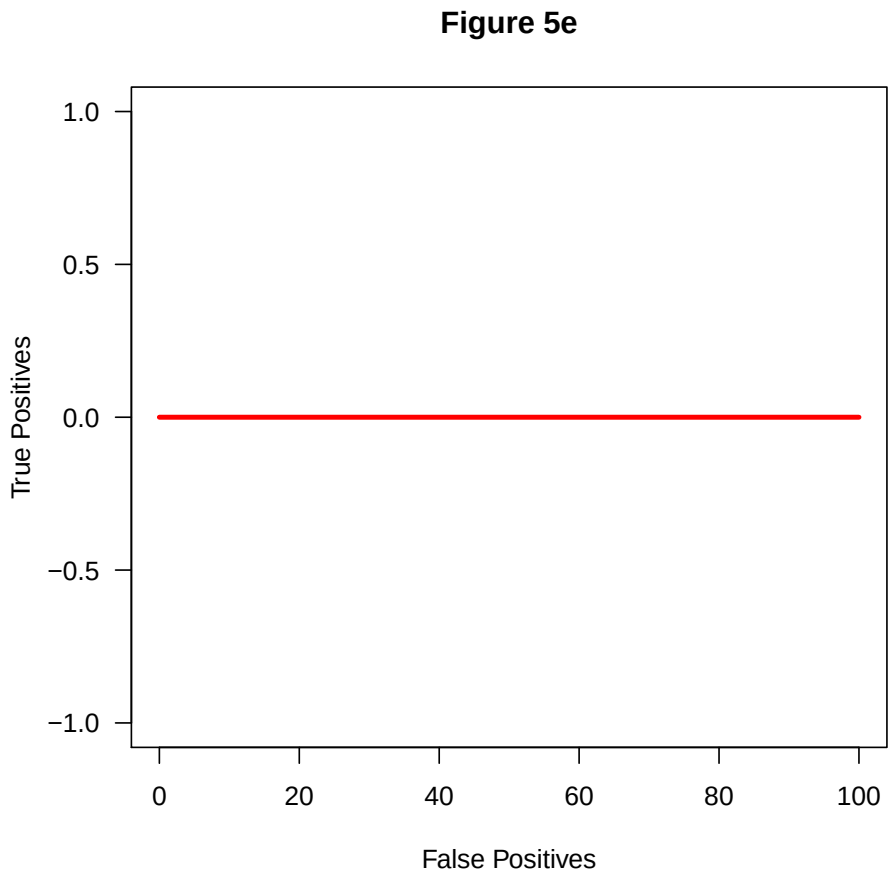

Figure 5e) As 5a) but for comparisons with both nominal concentrations at least 64 and with nominal fold changes at most 2.

Figure 6a

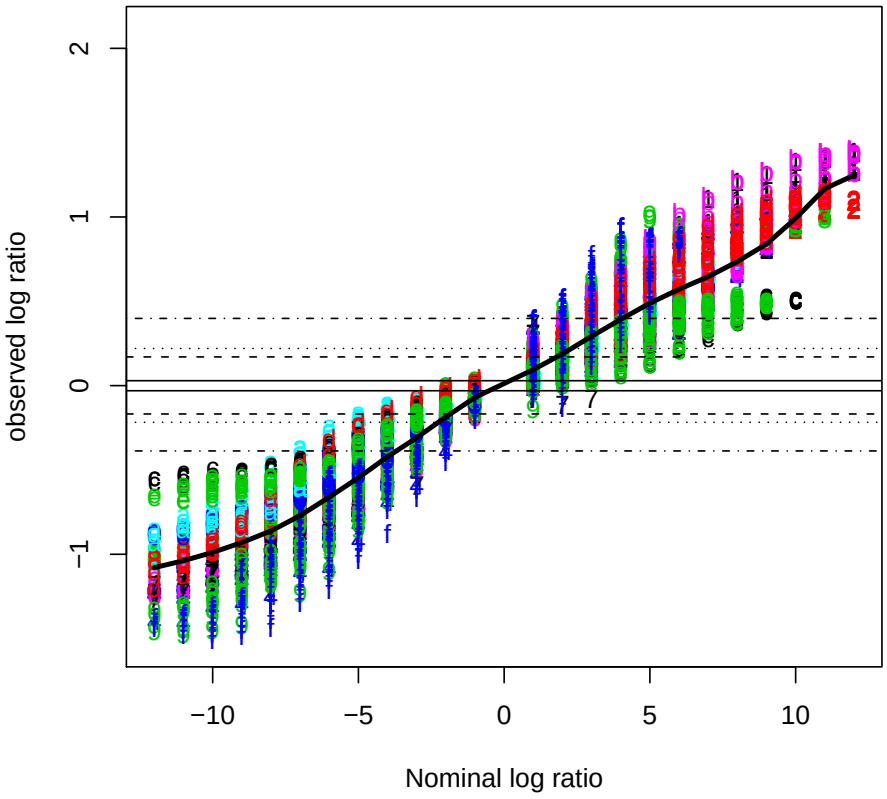

Figure 6a) Observed log fold changes plotted against nominal log fold changes. The dashed lines represent highest, 25th highest, 100th highest, 25th percentile, 75th percentile, smallest 100th, smallest 25th, and smallest log fold change for the genes that were not differentially expressed.

Figure 6b

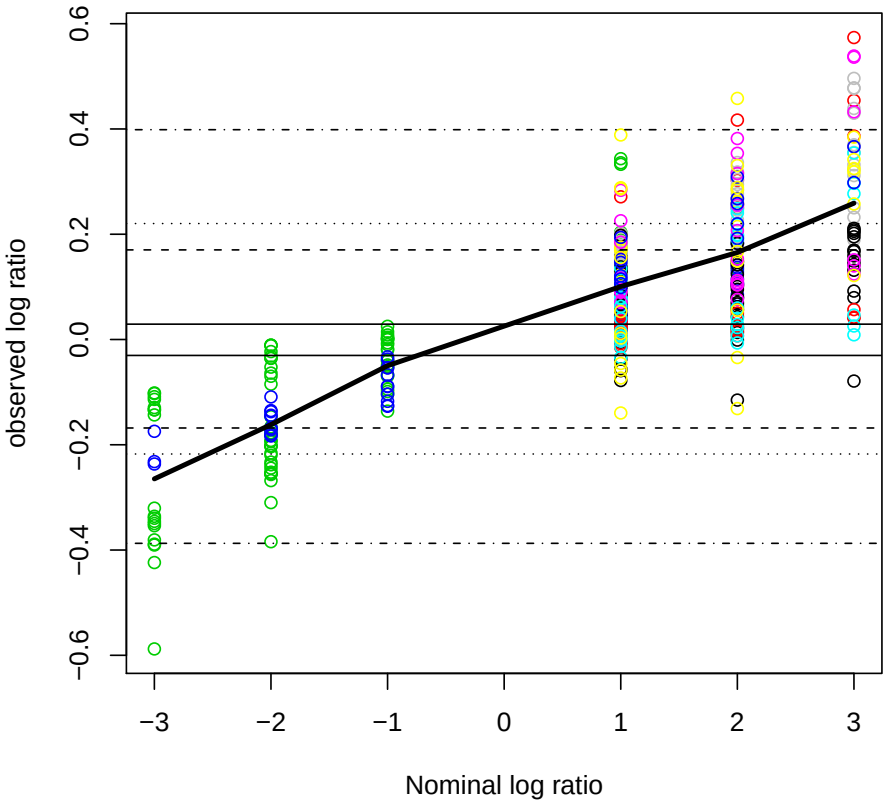

Figure 6b) Like a) but the observed fold changes were calculated for spiked in genes with nominal concentrations no higher than 2pM.