

Appendix C: Affymetrix U133A Experiment

T. M. Therneau K. V. Ballman

July 2005

This appendix contains the complete set of curves for the Affymetrix U133A spike-in data referenced in Technical Report 74 [1].

1 Data

The U133A spike-in dataset was produced by an experiment that has three technical replicates of 14 separate hybridizations of 42 spiked transcripts in a complex human background (HeLa cell line). Thirty of the spiked transcripts correspond to cDNA clones isolated from total RNAs of a lymphoblast cell line (and are not expressed in the HeLa cell line), eight of the spiked transcripts are made from artificial sequences, and the remaining four spiked transcripts are Affymetrix eukaryotic controls that are available as part of a poly A spike control kit. There are fourteen groups of three genes each. The concentration of each gene within a group is spiked at the same concentration. A cyclic latin square was used for the group concentrations. Specifically, the dosing pattern for Group 1 across 14 arrays is 0, 1/8, 1/4, 1/2, 1, 2, 4, ..., 256, and 512 pM, for Group 2 is 1/8, 1/4, 1/2, 1, ..., 512, and 0 pM, and so on until Group 14, which has a pattern of 512, 0, 1/8, 1/4, ..., 128, and 256 pM. Each pattern appears on three replicate arrays, yielding a total of 42 arrays. There are 11 probe pairs for each of 38 genes and 20 probe pairs for the four Affymetrix eukaryotic controls. Additional information about this experiment is available at the Affymetrix website, www.affymetrix.com/support/technical/sample_data/datasets.affx.

2 Plot description

We chose to use the logistic function to fit the data from this spiked-in gene expression experiment, where x represents the level of known spiked-in transcripts concentration and y is the observed intensity value produced by the array.

To fit the perfect match (PM) and mismatch (MM) probes simultaneously, we used a five-parameter version of the model

$$\log(y) = a + bf(c[\log(x) - d - e]) \tag{1}$$

where

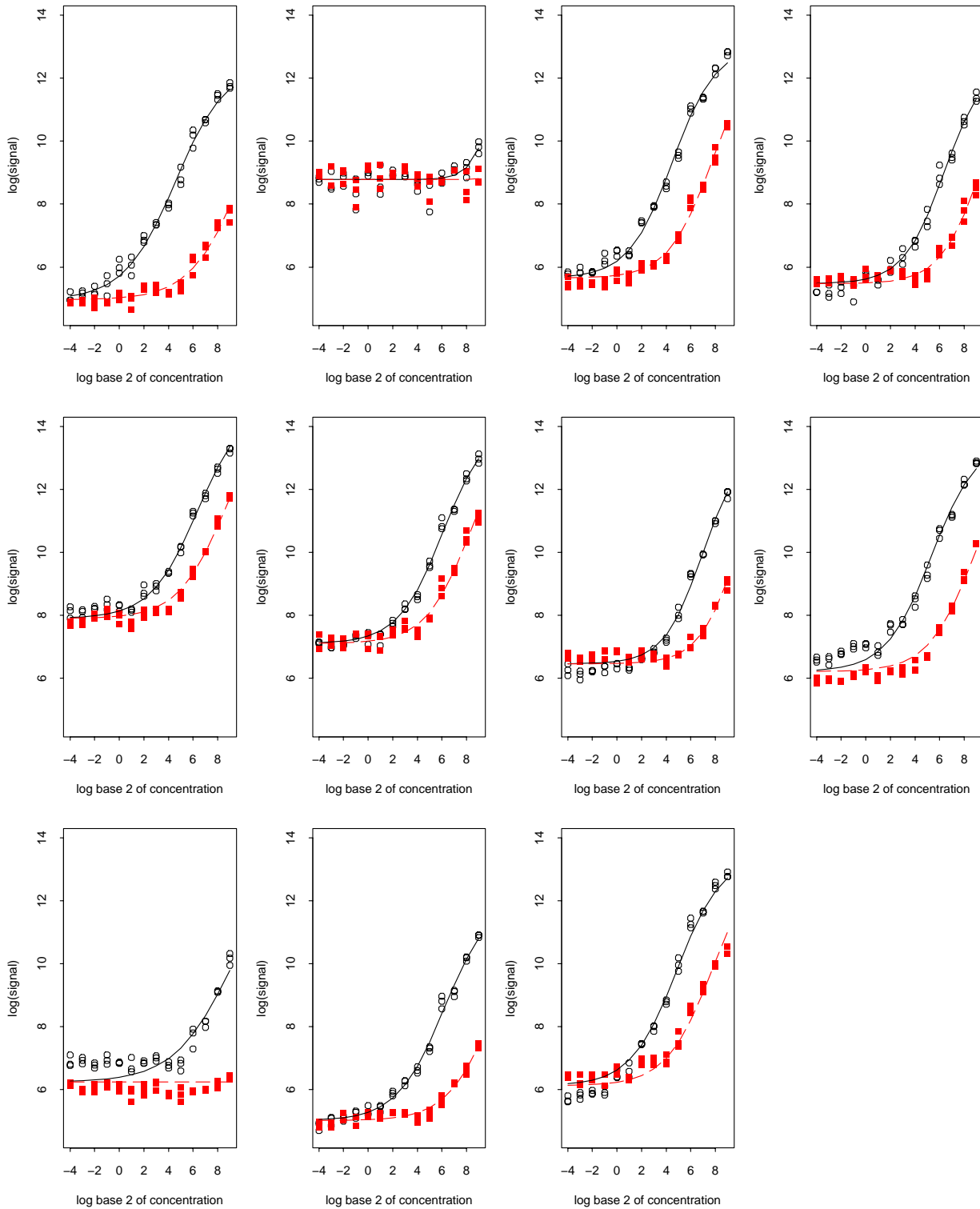
- f is the logistic function defined above,
- a is the lower threshold for the probe pair (i.e. the background level),
- b is the range of the curve,
- $a + b$ is the upper threshold for the probe pair (i.e. the saturation level),
- $bc/4$ is the slope of the curve at its inflection point,
- d is the inflection point for the PM probes, and
- $d + e$ is the inflection point for the MM probes.

All the plots are presented on a log base 2 scale. A least squares approach was used to fit the curves to the data. The observed expression values were plotted on the y -axis and the spike-in concentrations were plotted on the x -axis. A plot was made for each probe in the probeset; both the perfect match (PM) and mismatch (MM) values were plotted using different symbols. Plots for all the probe pairs within a probeset, referred to as a panel, were plotted on the same page in the order they appeared on the gene. Fitted logistic calibration curves are superimposed on the data plots.

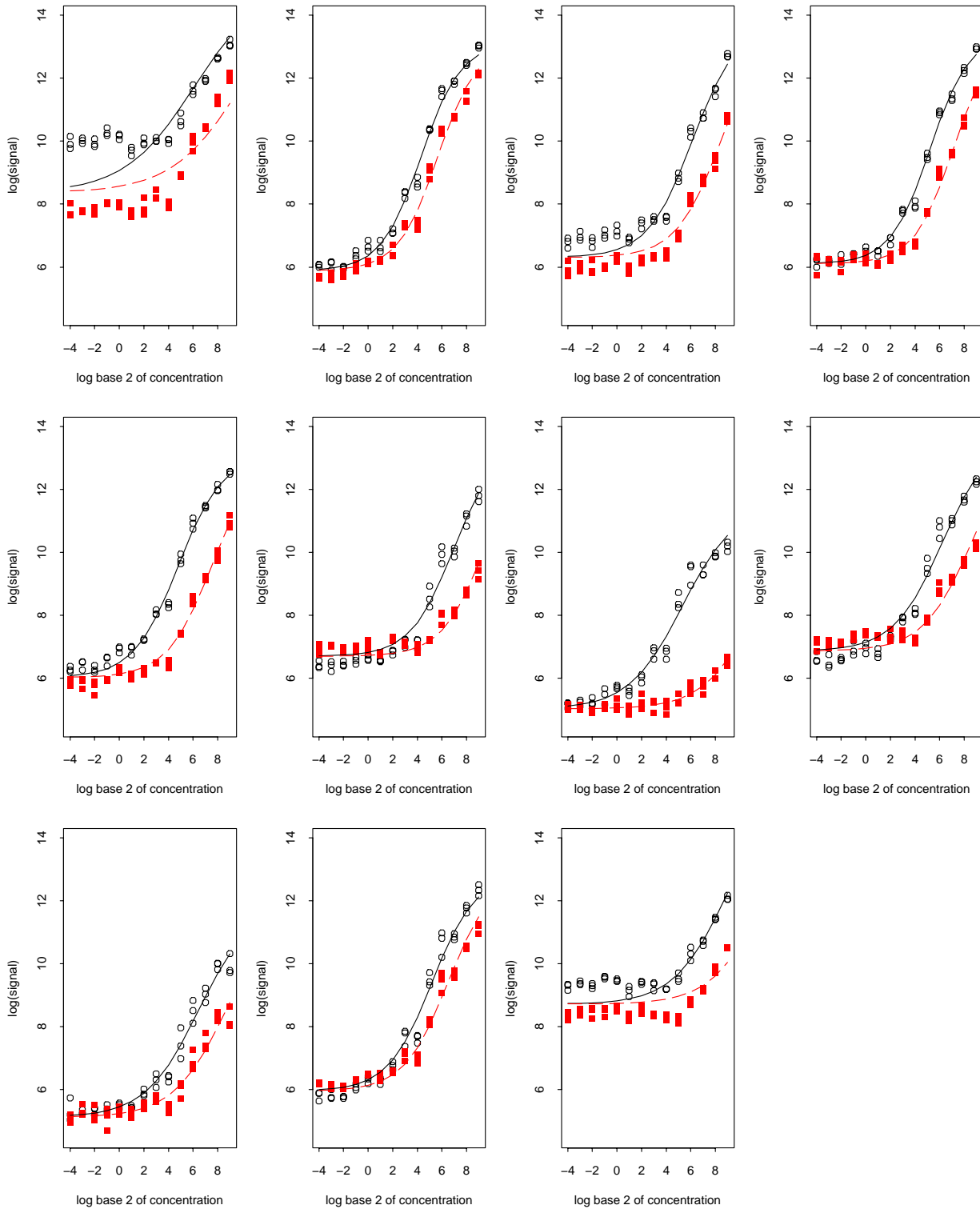
3 The plots

The first 38 pages or panels contain 11 plots (one per probeset) and the remaining 4 panels contain 20 plots yielding 498 plots. The black open circles are the values of the PM probes and the red filled squares are the values of the MM probes. The black solid line corresponds to the fitted logistic calibration curve for the PM probes and the red dashed line corresponds to the fitted logistic calibration curves for the MM probes; these curves were the results of a simultaneous fit.

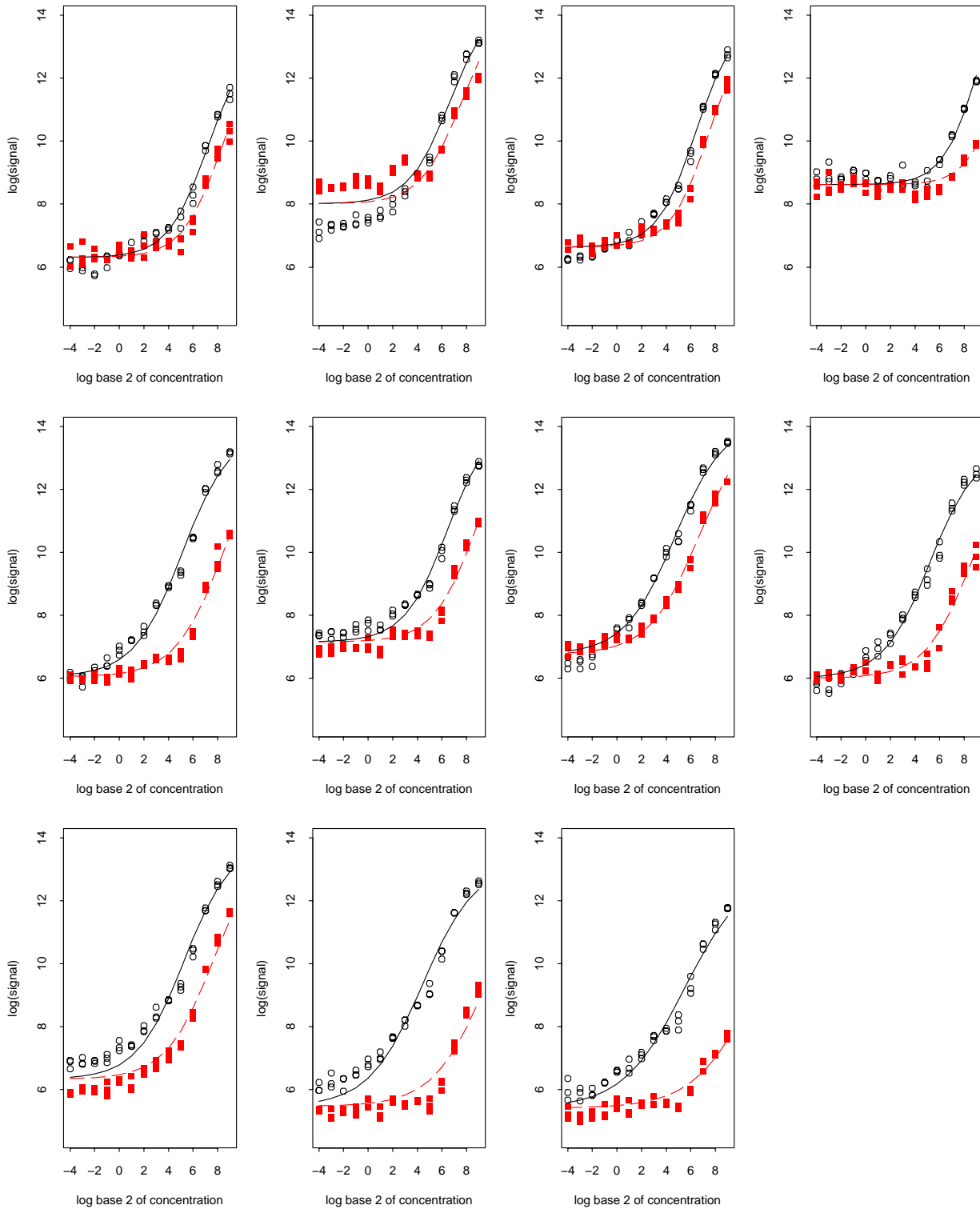
Affy U133 spikein, gene 1 (203508_at)



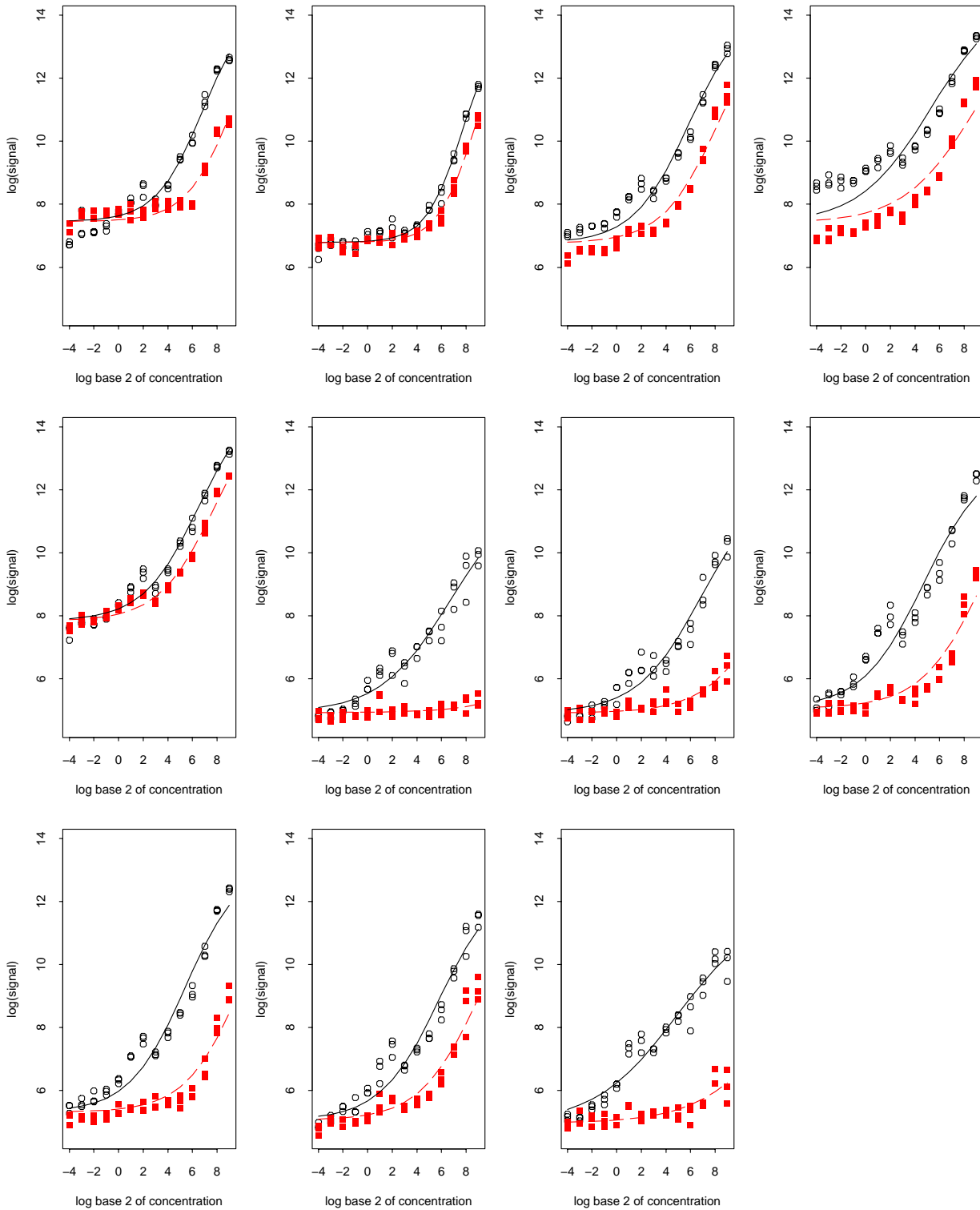
Affy U133 spikein, gene 2 (204563_at)



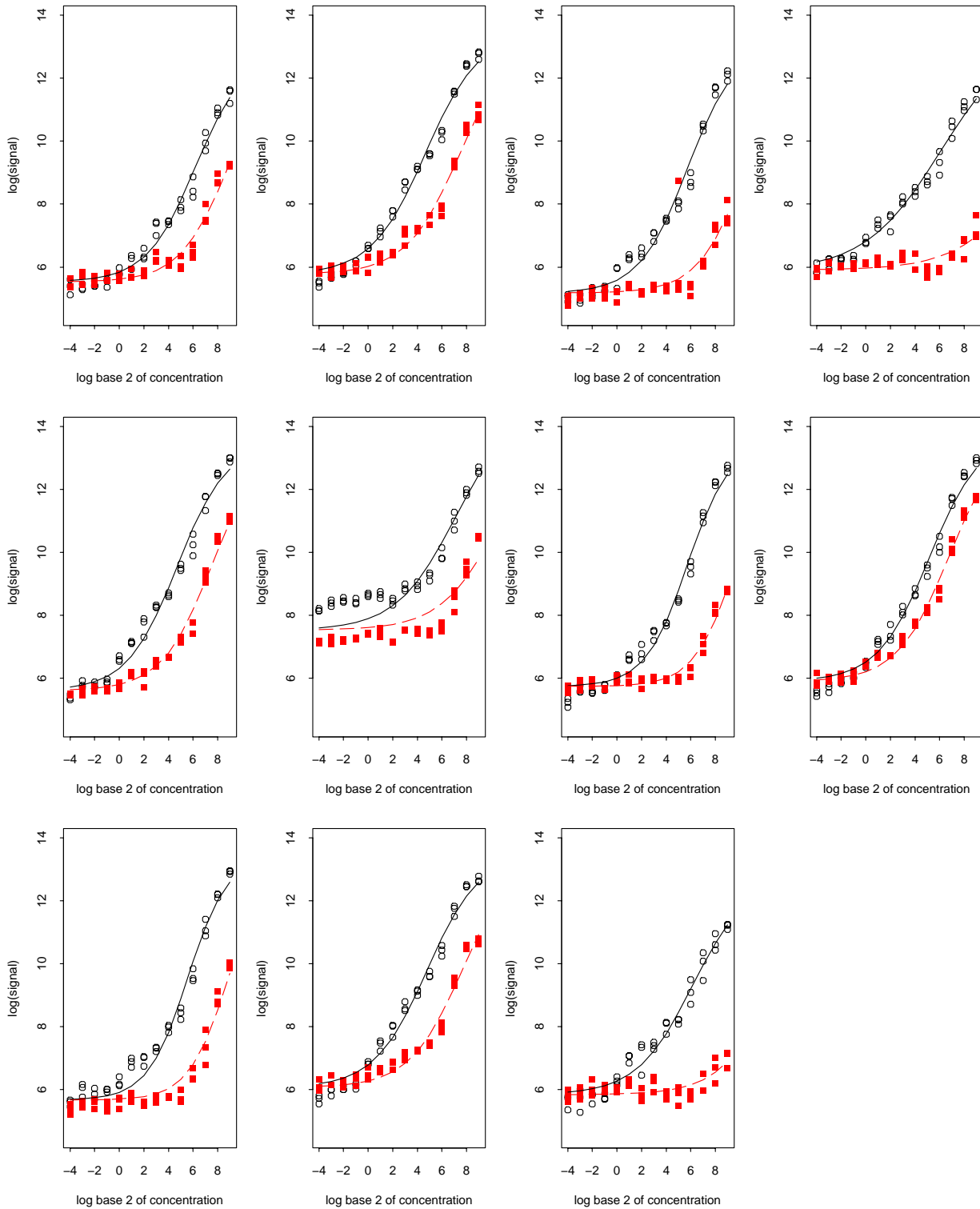
Affy U133 spikein, gene 3 (204513_s_at)



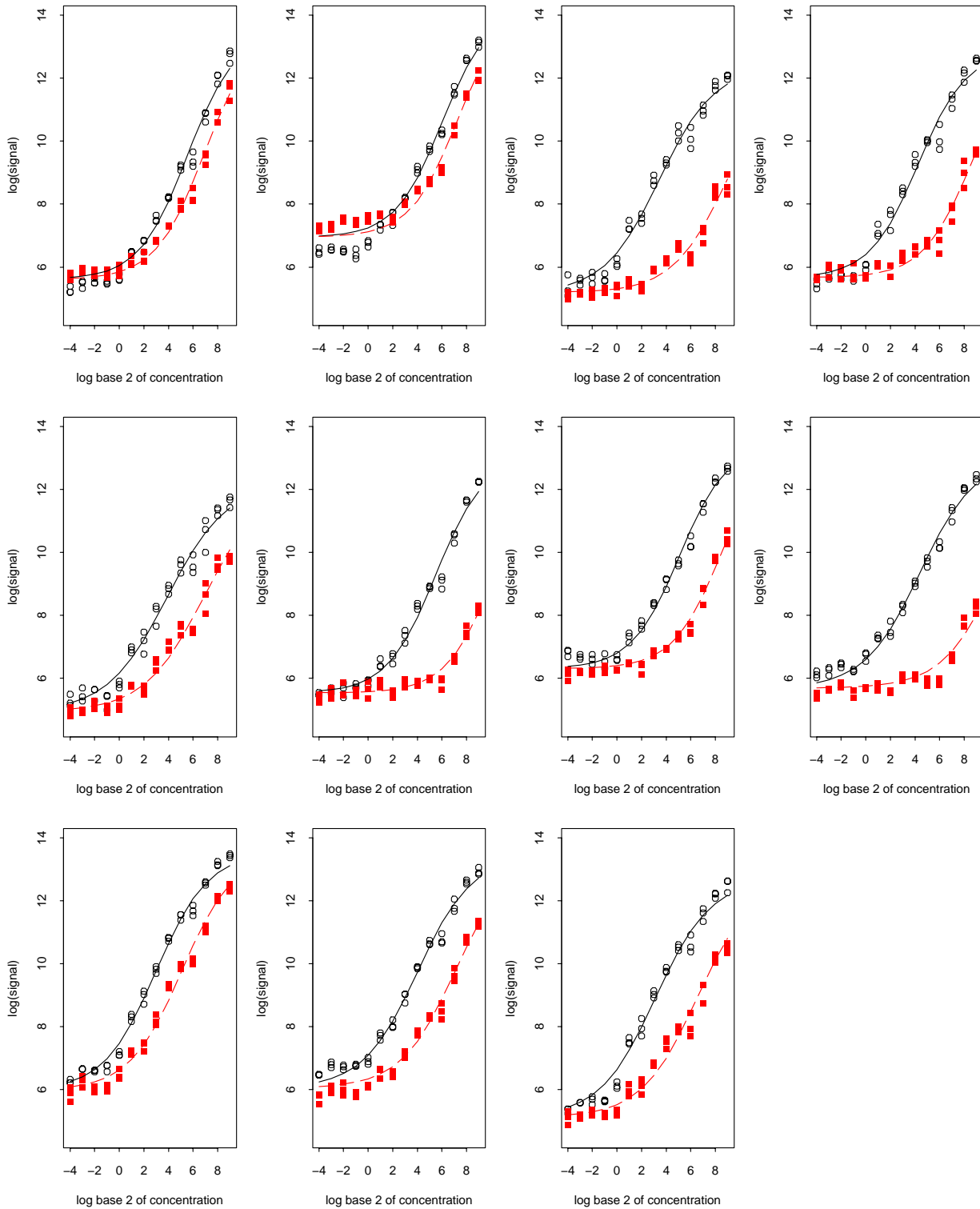
Affy U133 spikein, gene 4 (204205_at)



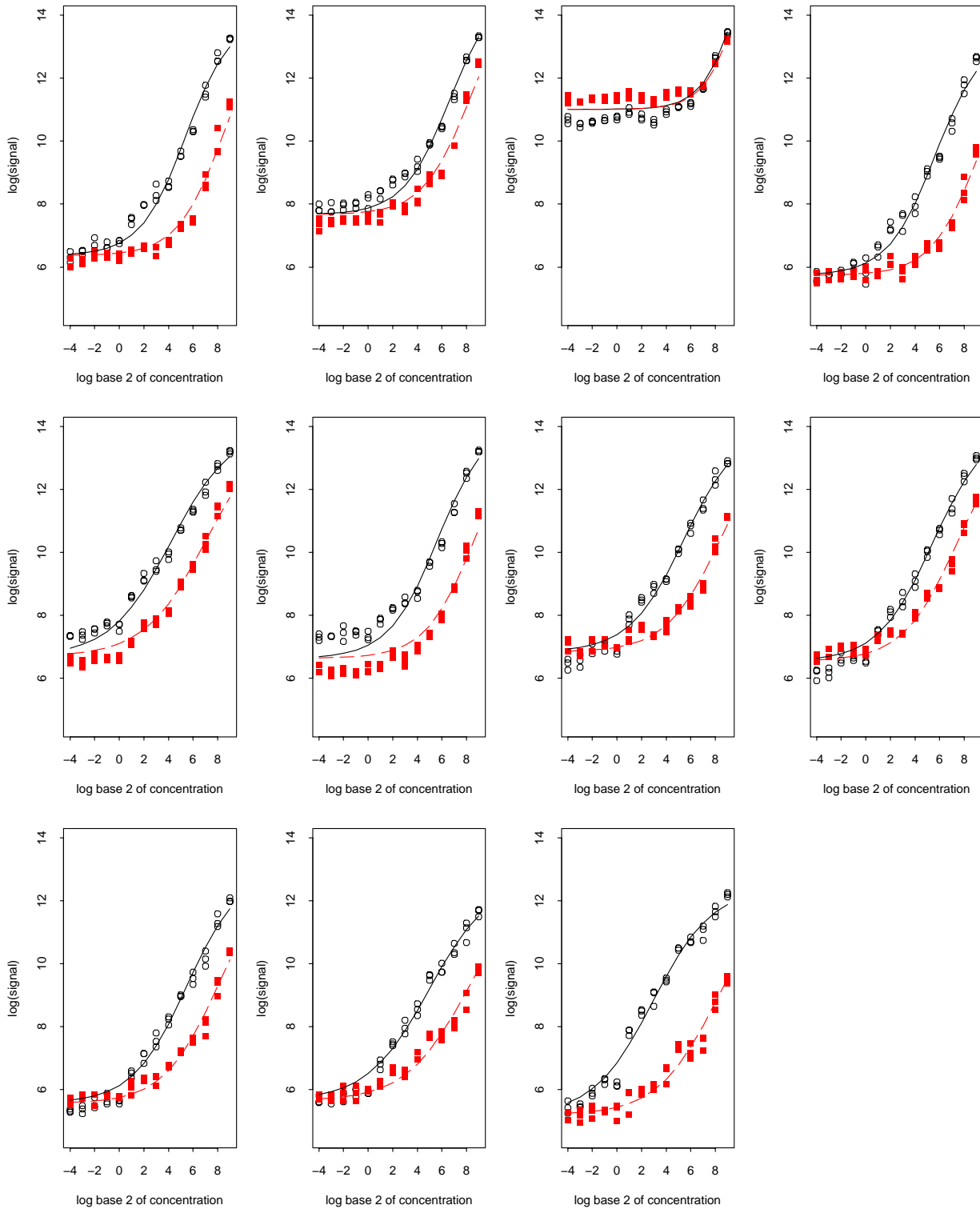
Affy U133 spikein, gene 5 (204959_at)



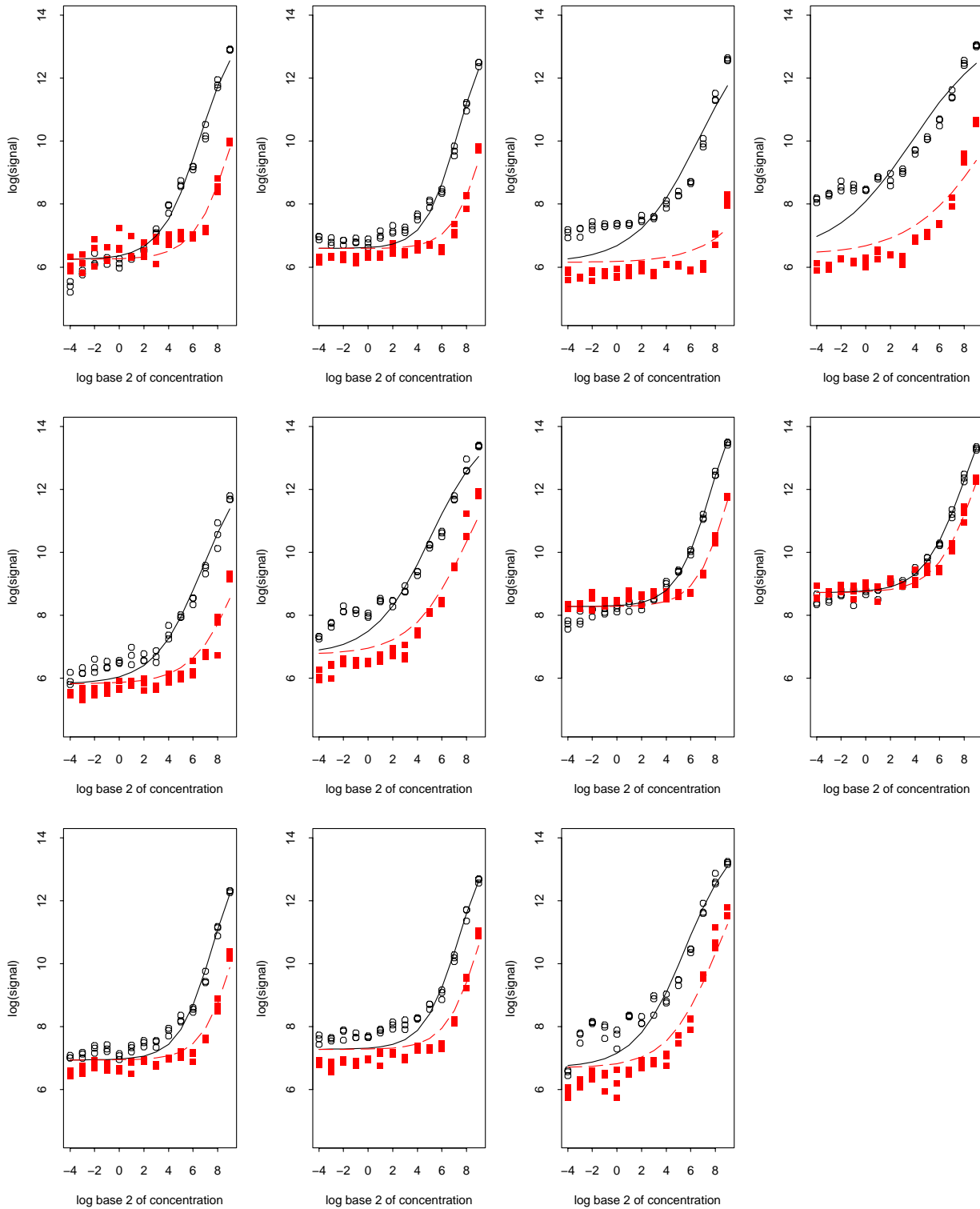
Affy U133 spikein, gene 6 (207655_s_at)



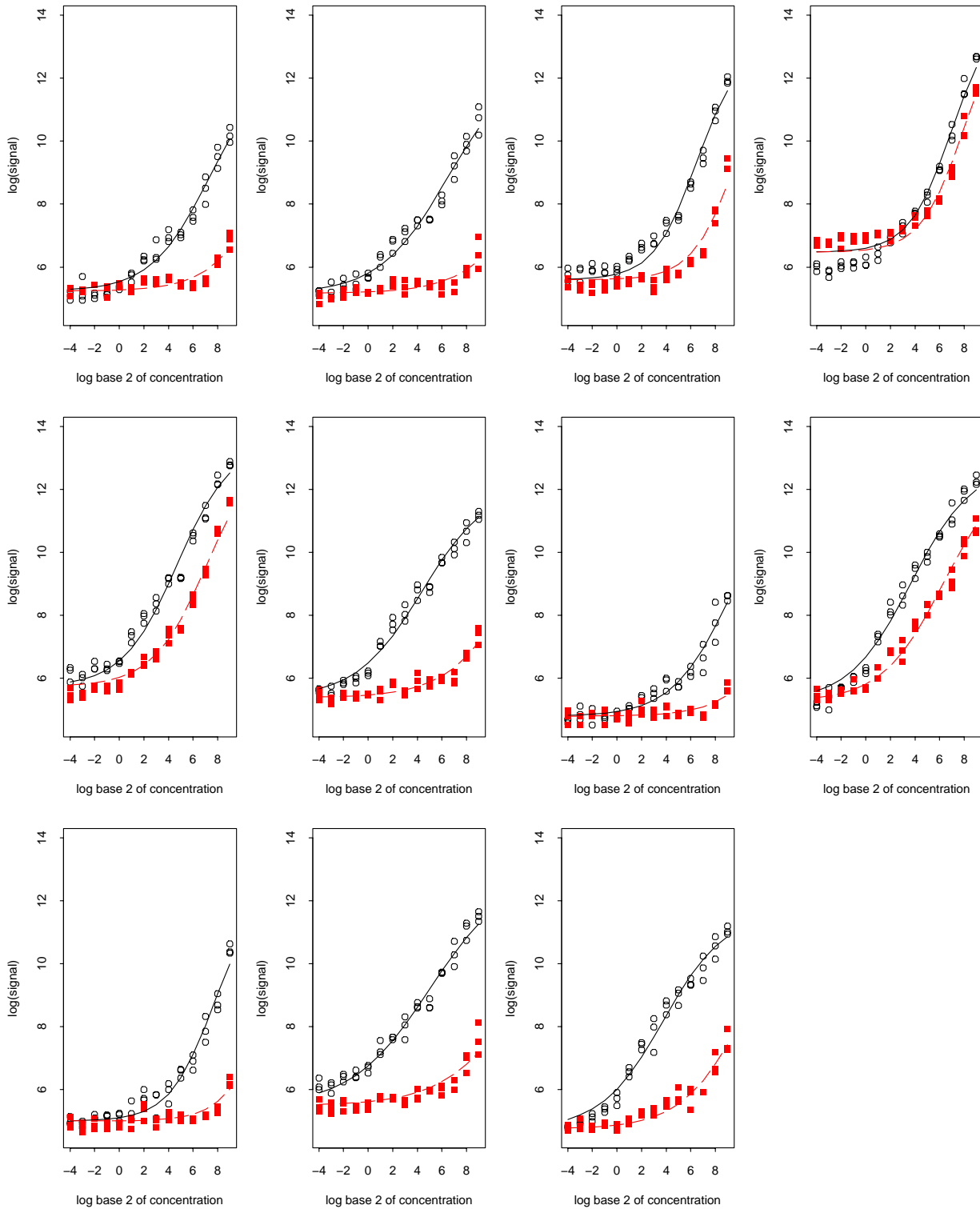
Affy U133 spikein, gene 7 (204836_at)



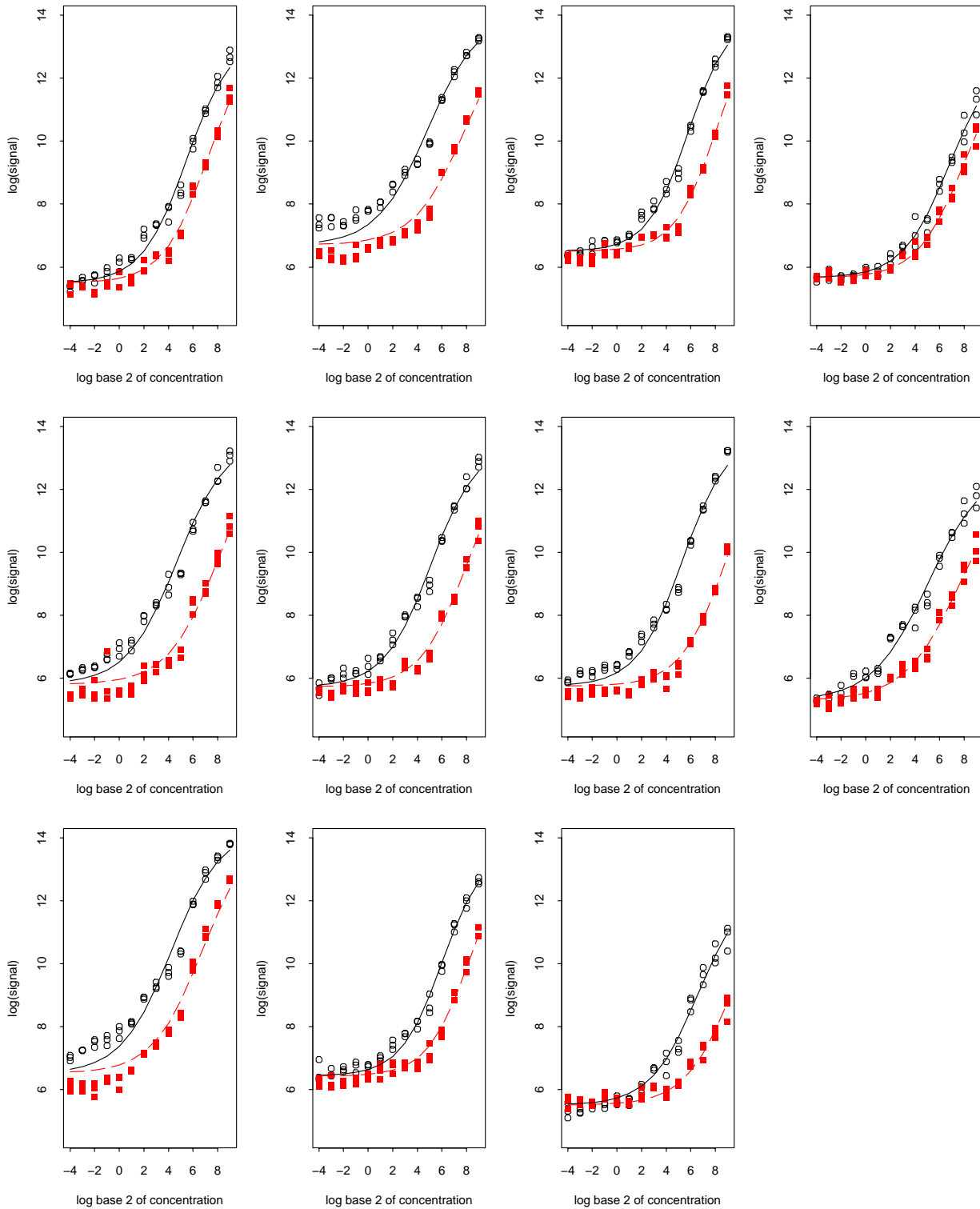
Affy U133 spikein, gene 8 (205291_at)



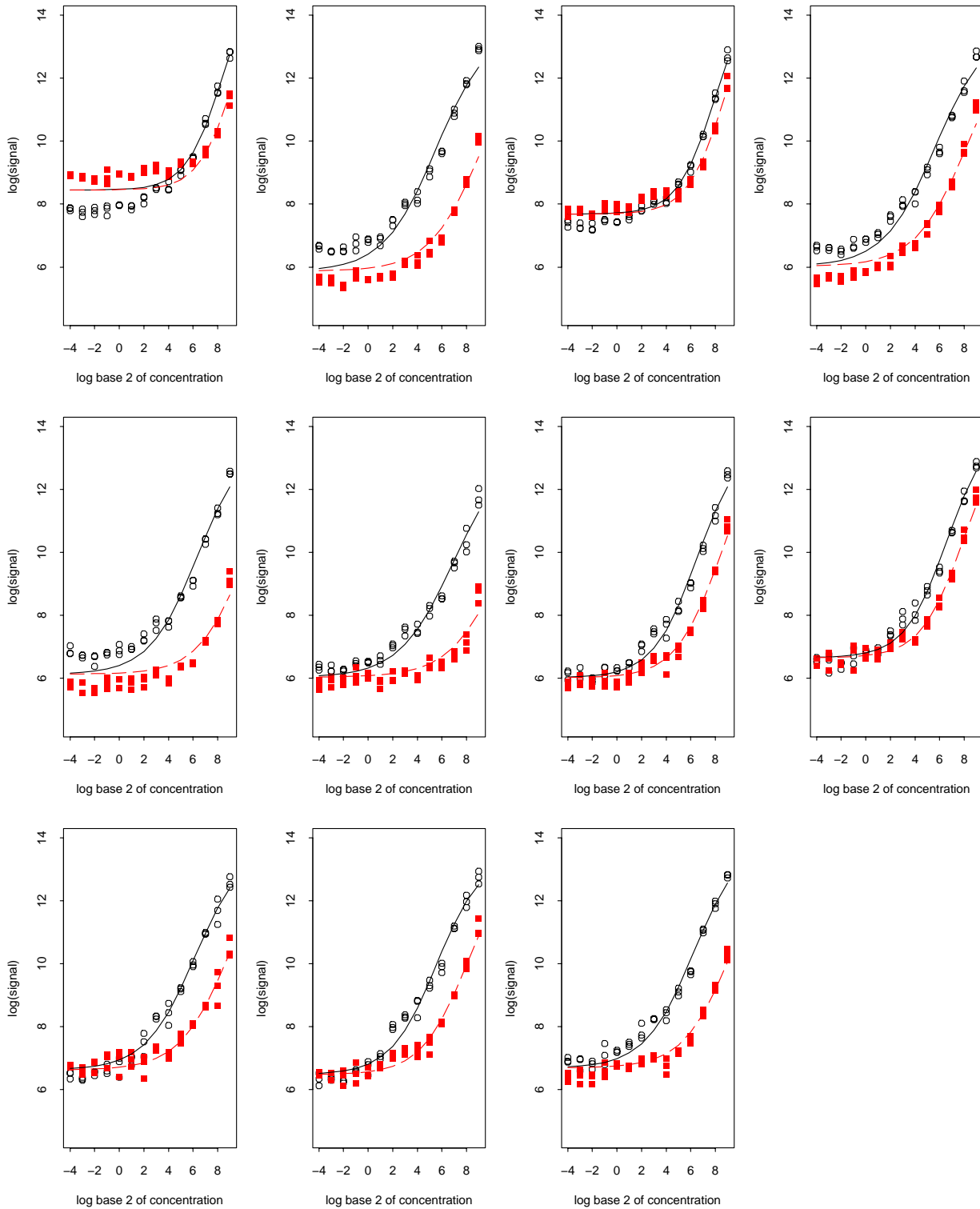
Affy U133 spikein, gene 9 (209795_at)



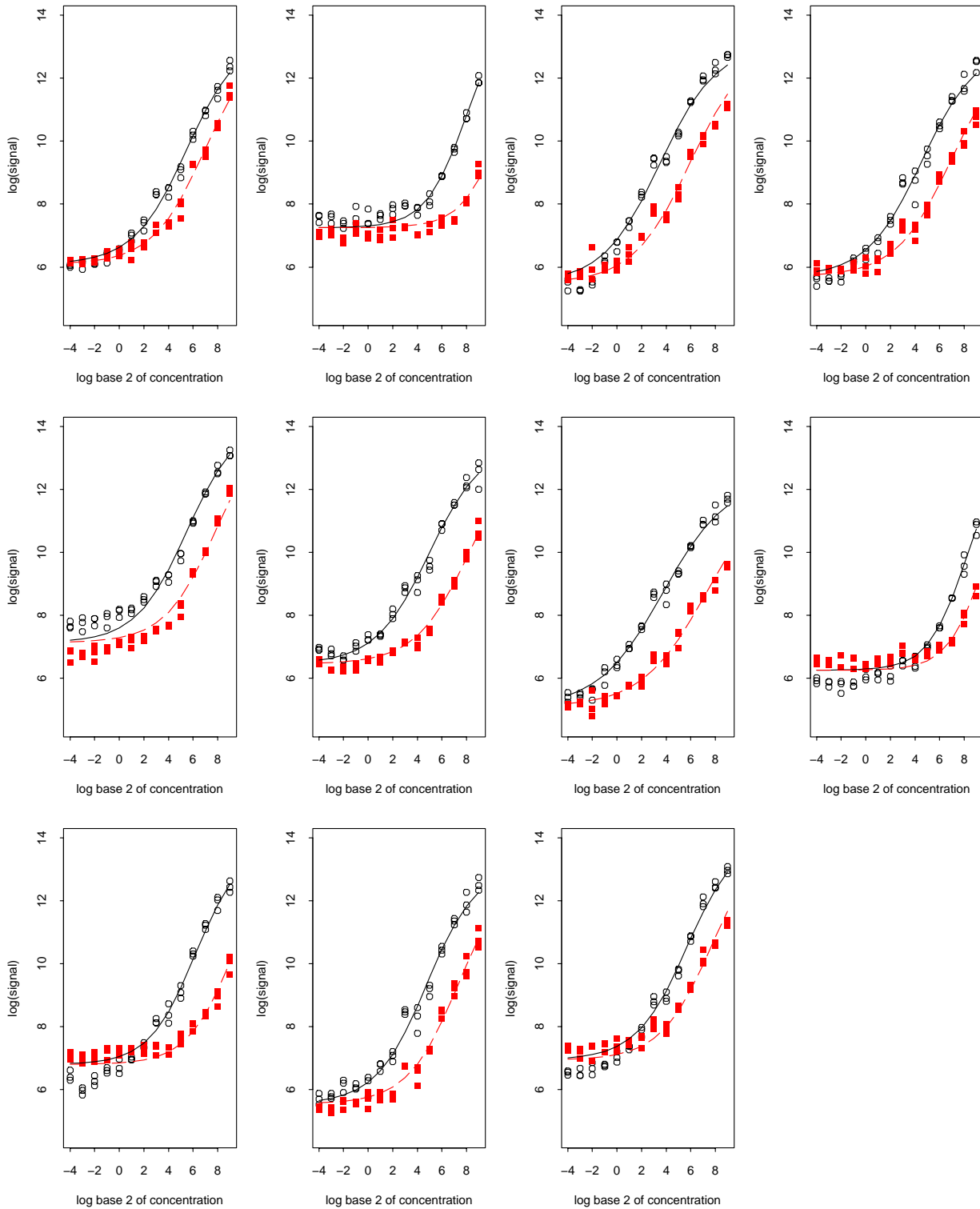
Affy U133 spikein, gene 10 (207777_s_at)



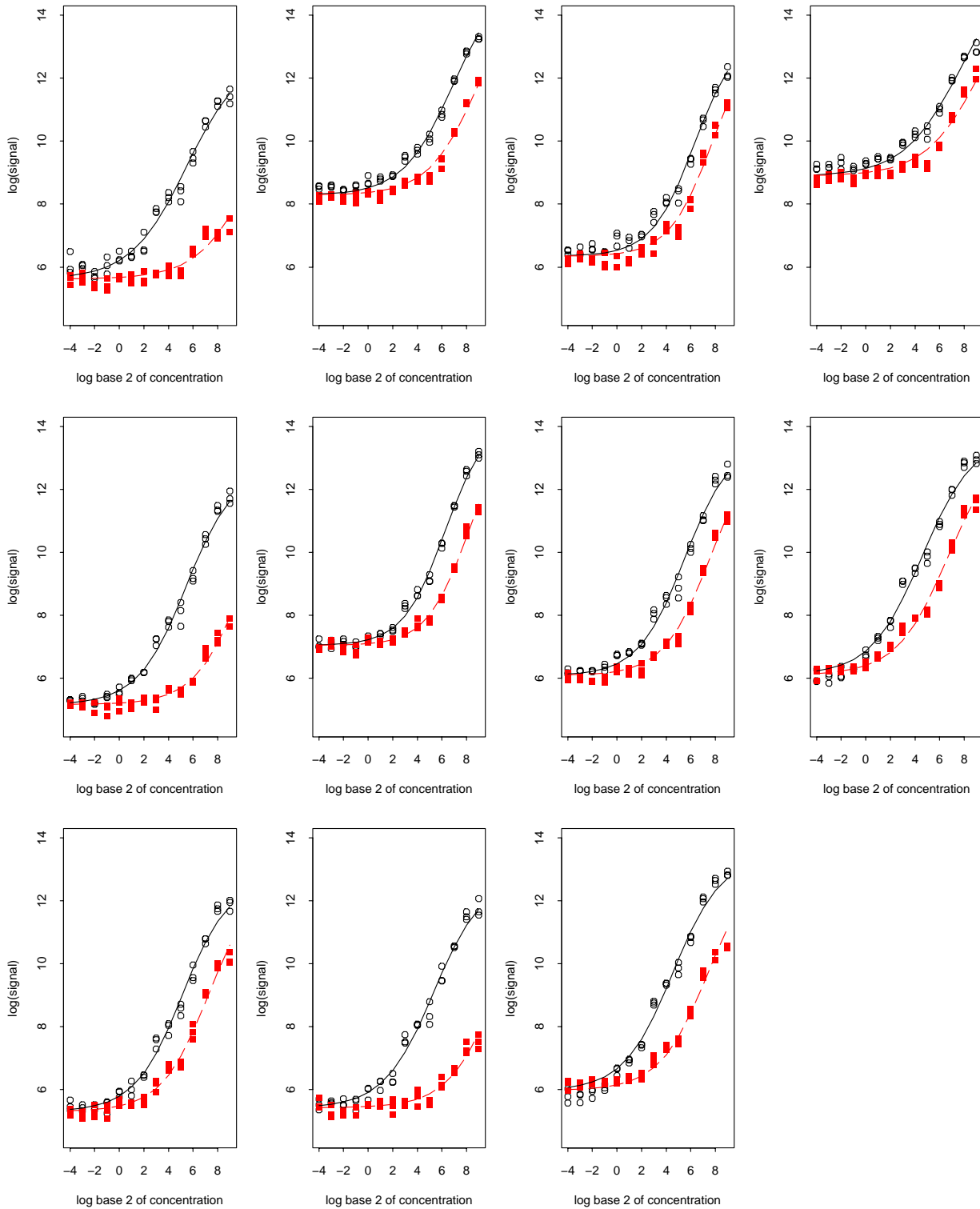
Affy U133 spikein, gene 11 (204912_at)



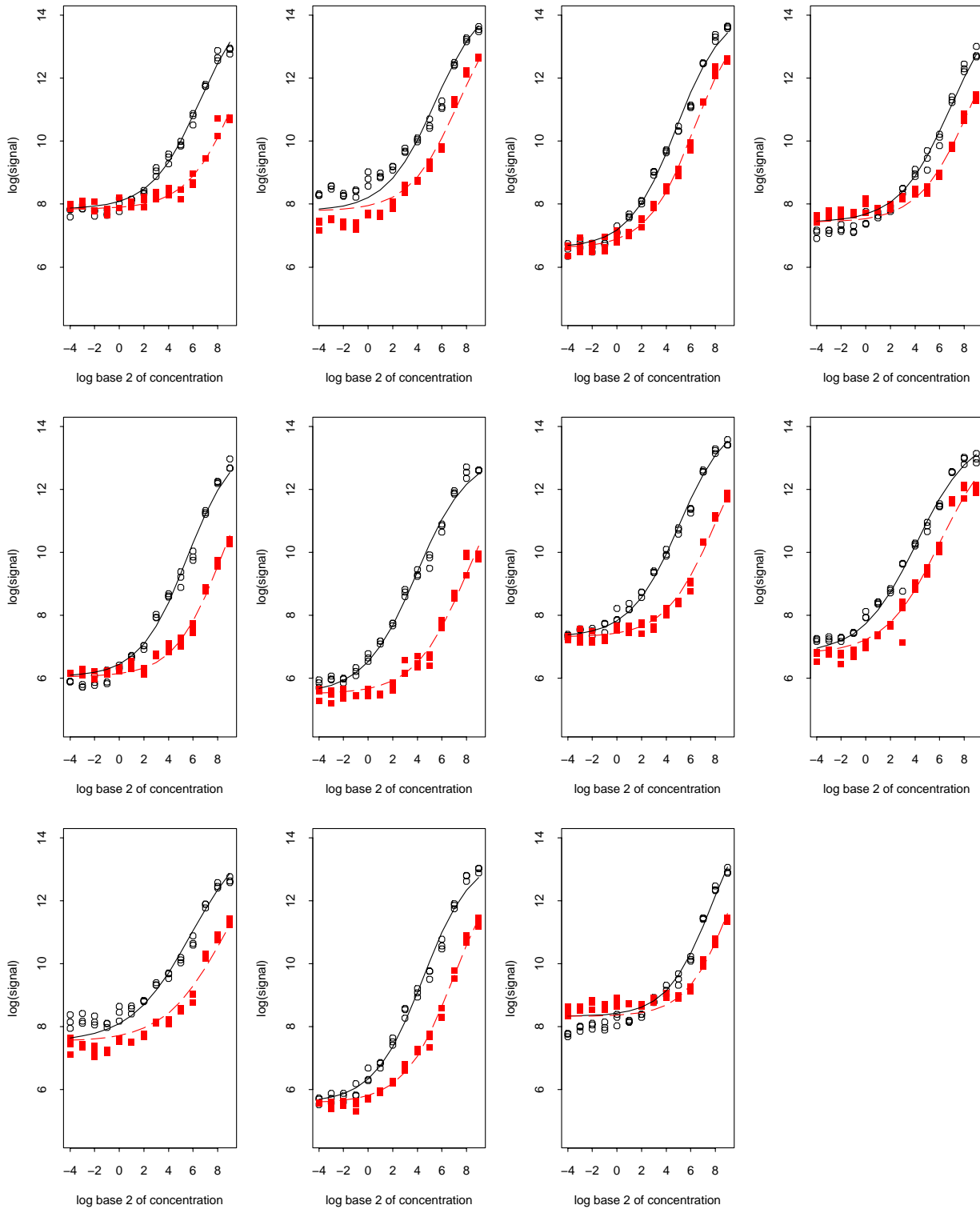
Affy U133 spikein, gene 12 (205569_at)



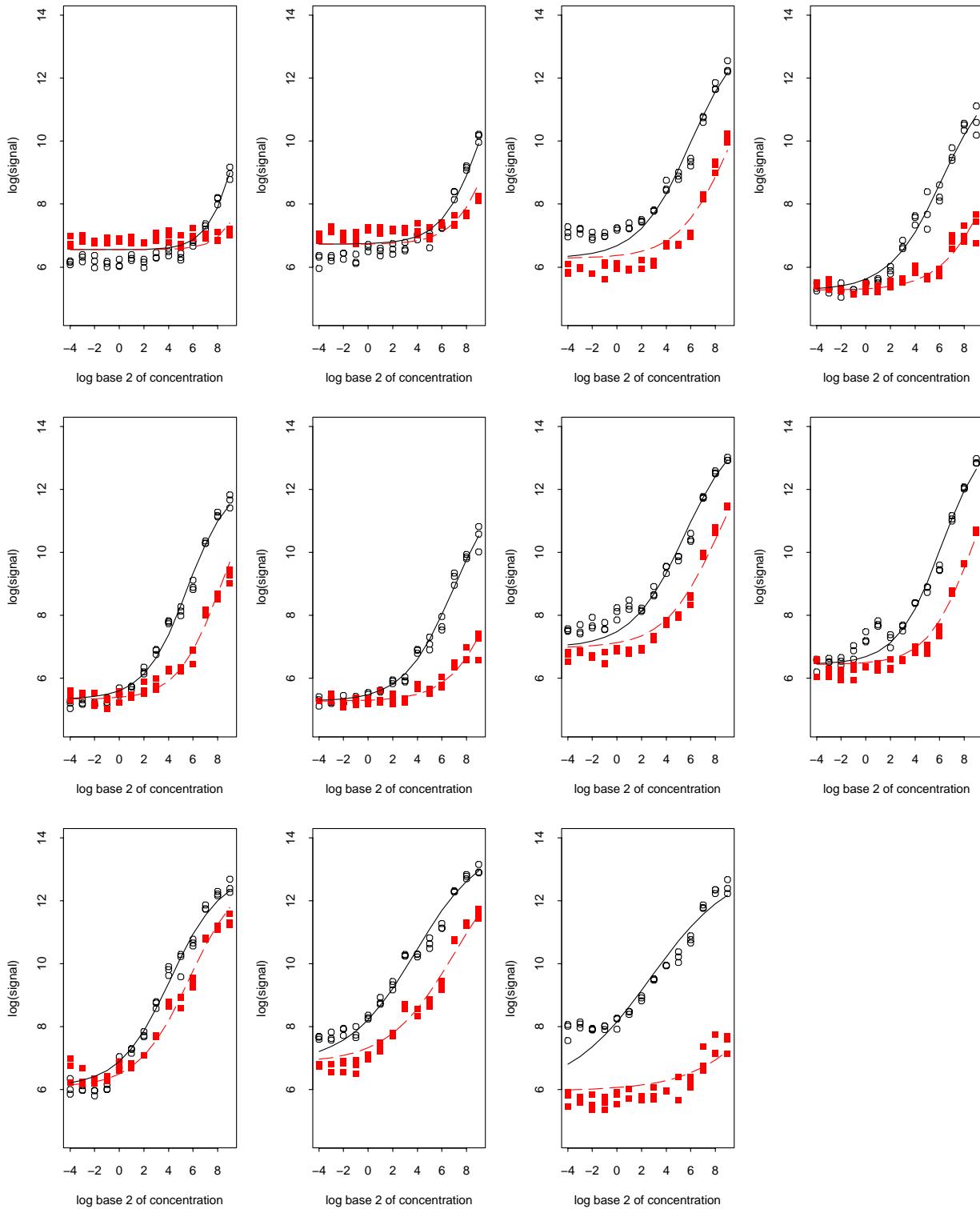
Affy U133 spikein, gene 13 (207160_at)



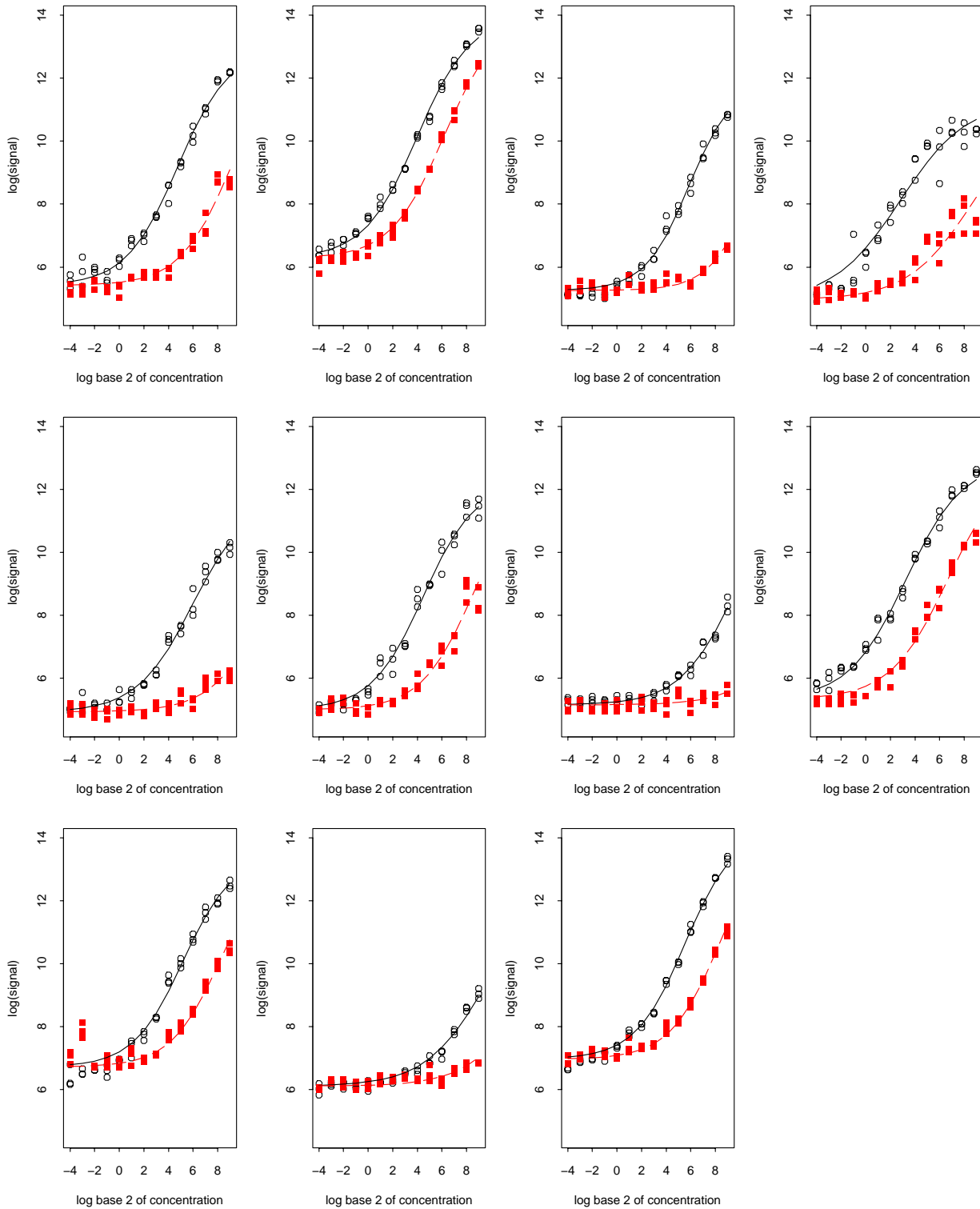
Affy U133 spikein, gene 14 (205692_s_at)



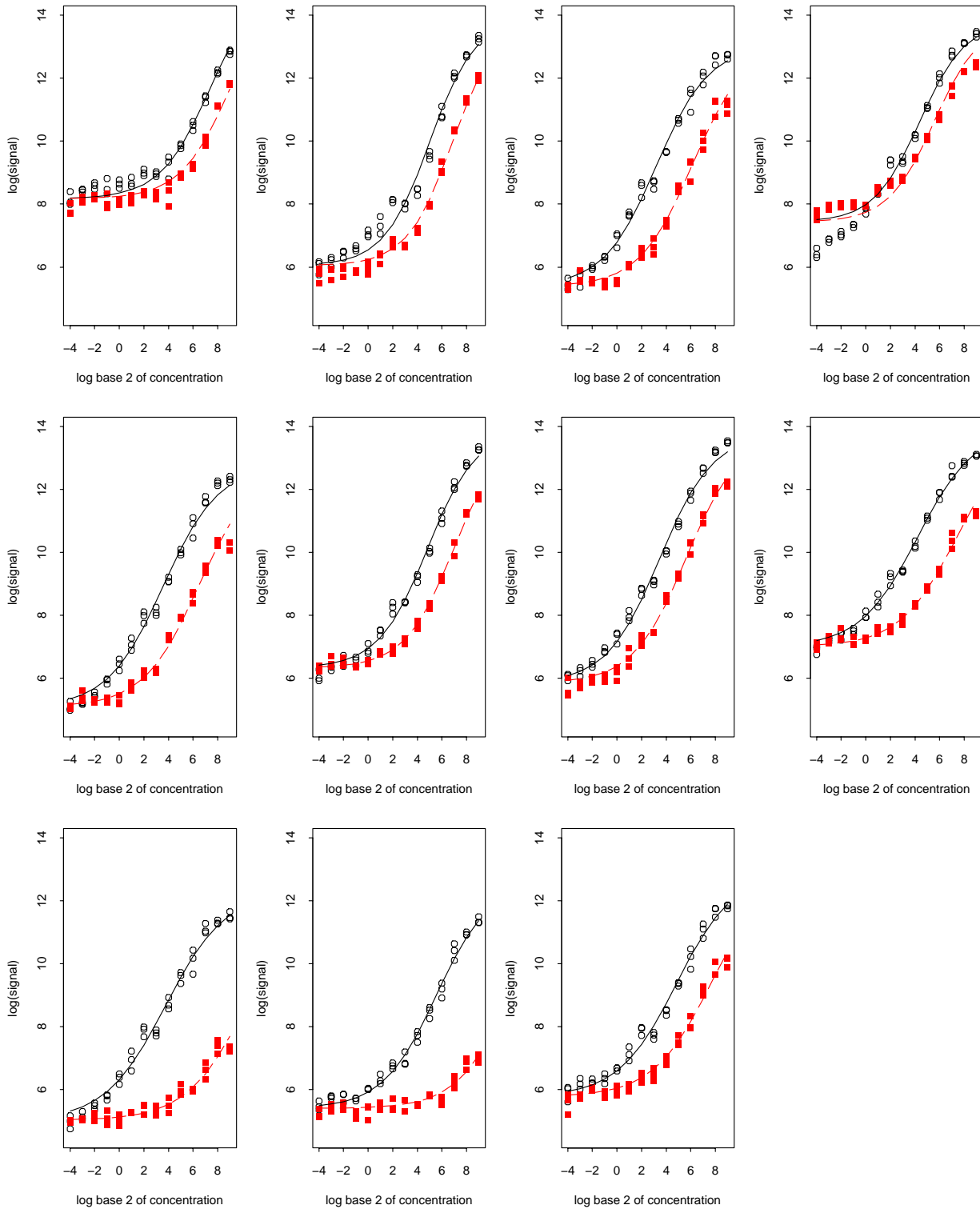
Affy U133 spikein, gene 15 (212827_at)



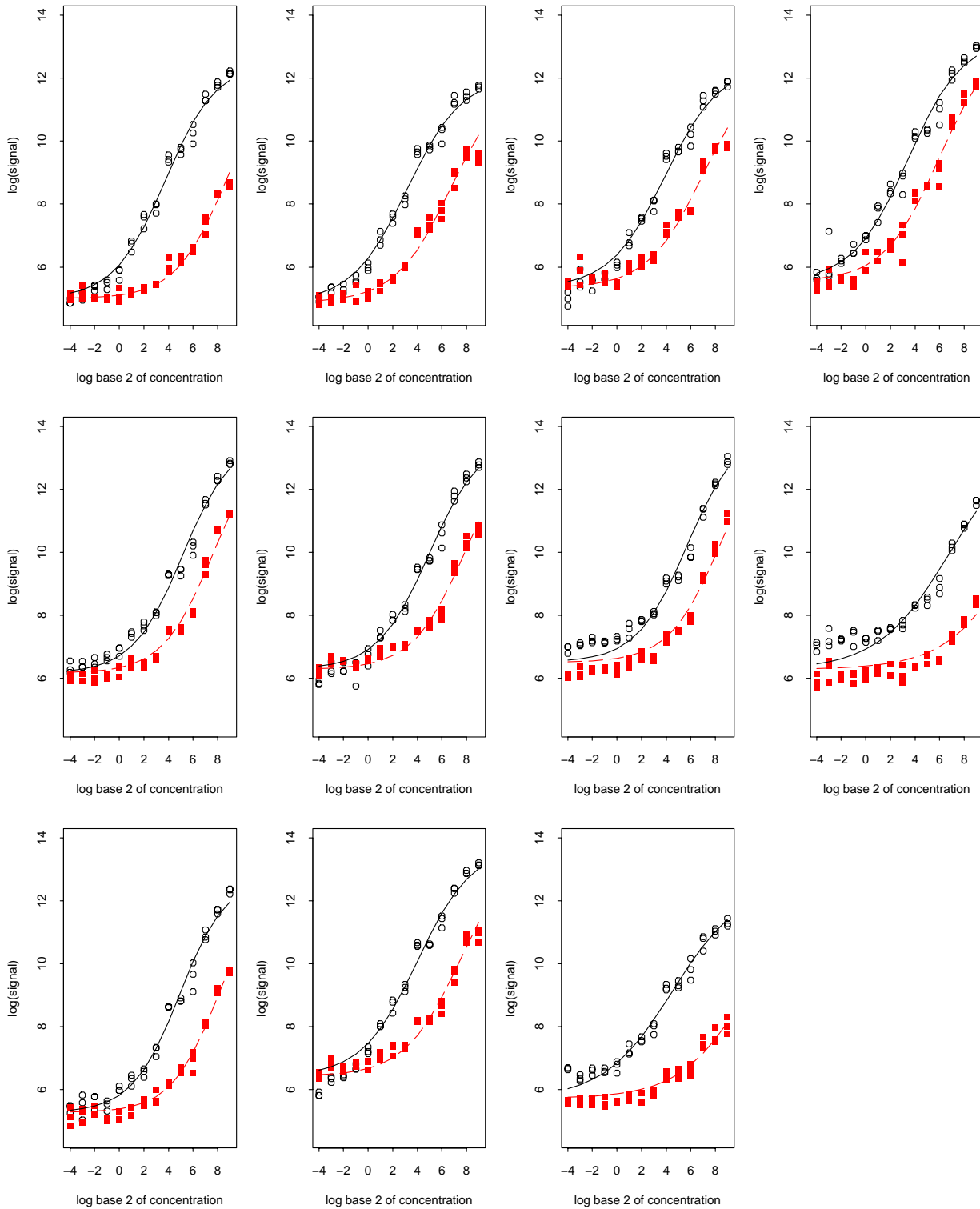
Affy U133 spikein, gene 16 (209606_at)



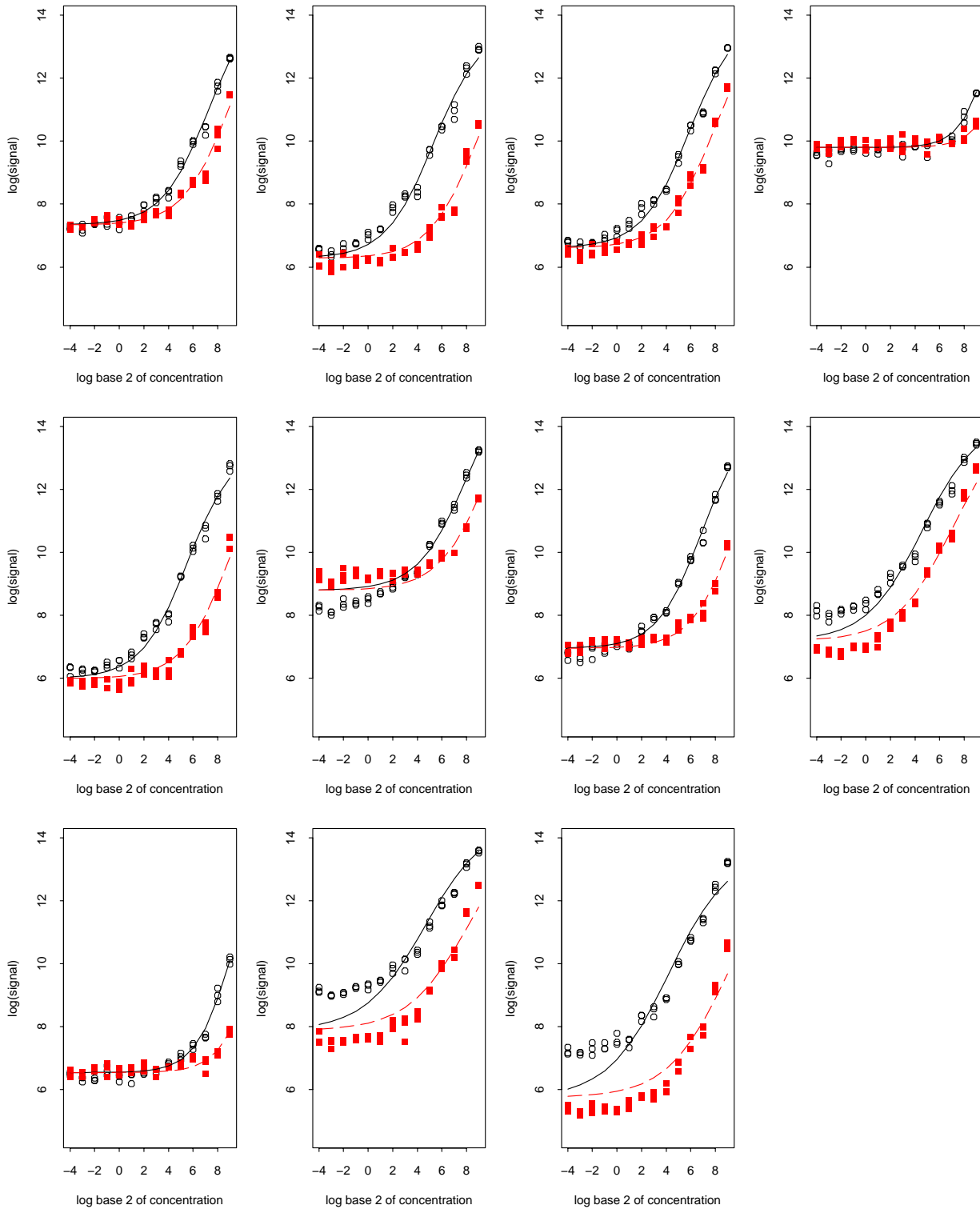
Affy U133 spikein, gene 17 (205267_at)



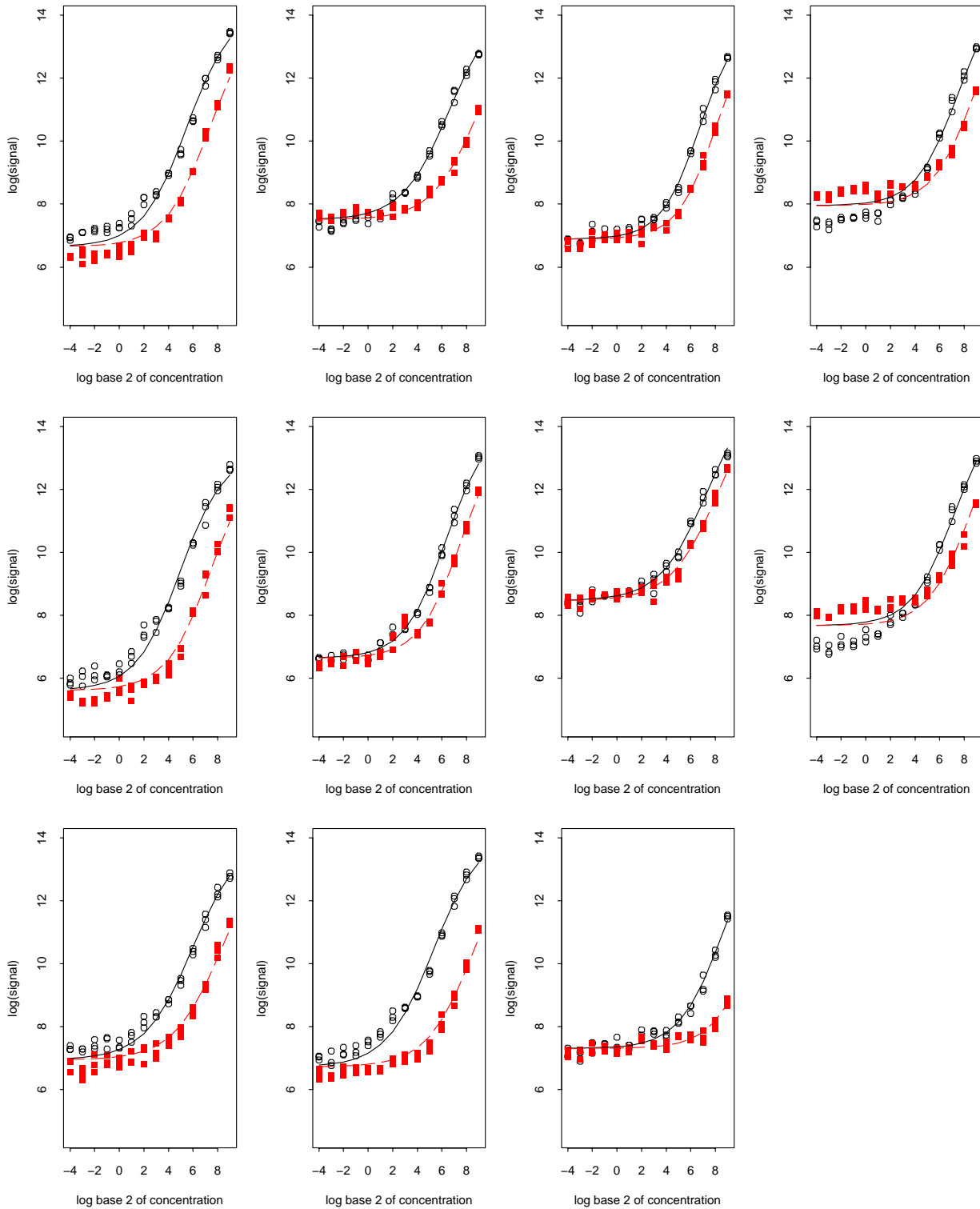
Affy U133 spikein, gene 18 (204417_at)



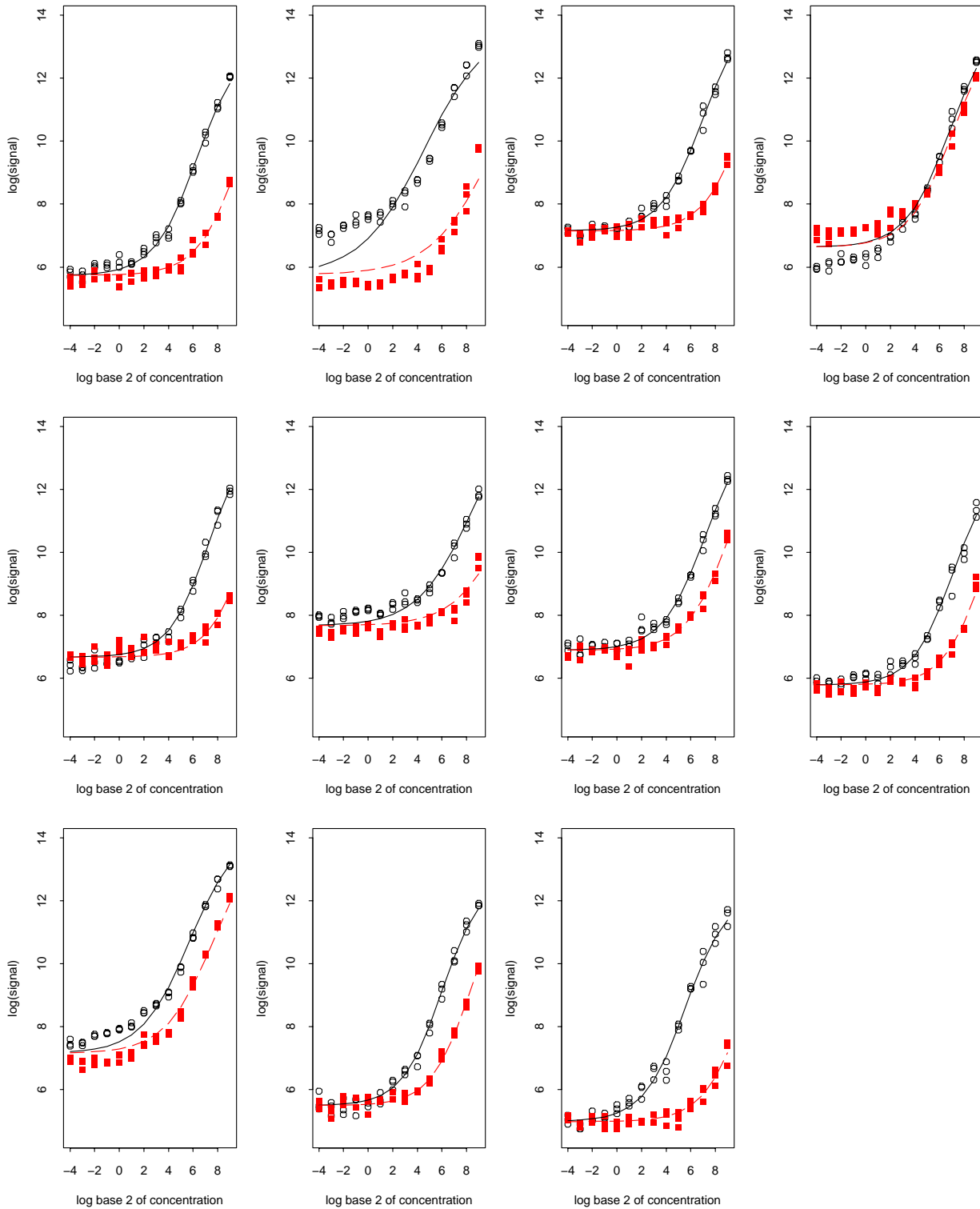
Affy U133 spikein, gene 19 (205398_s_at)



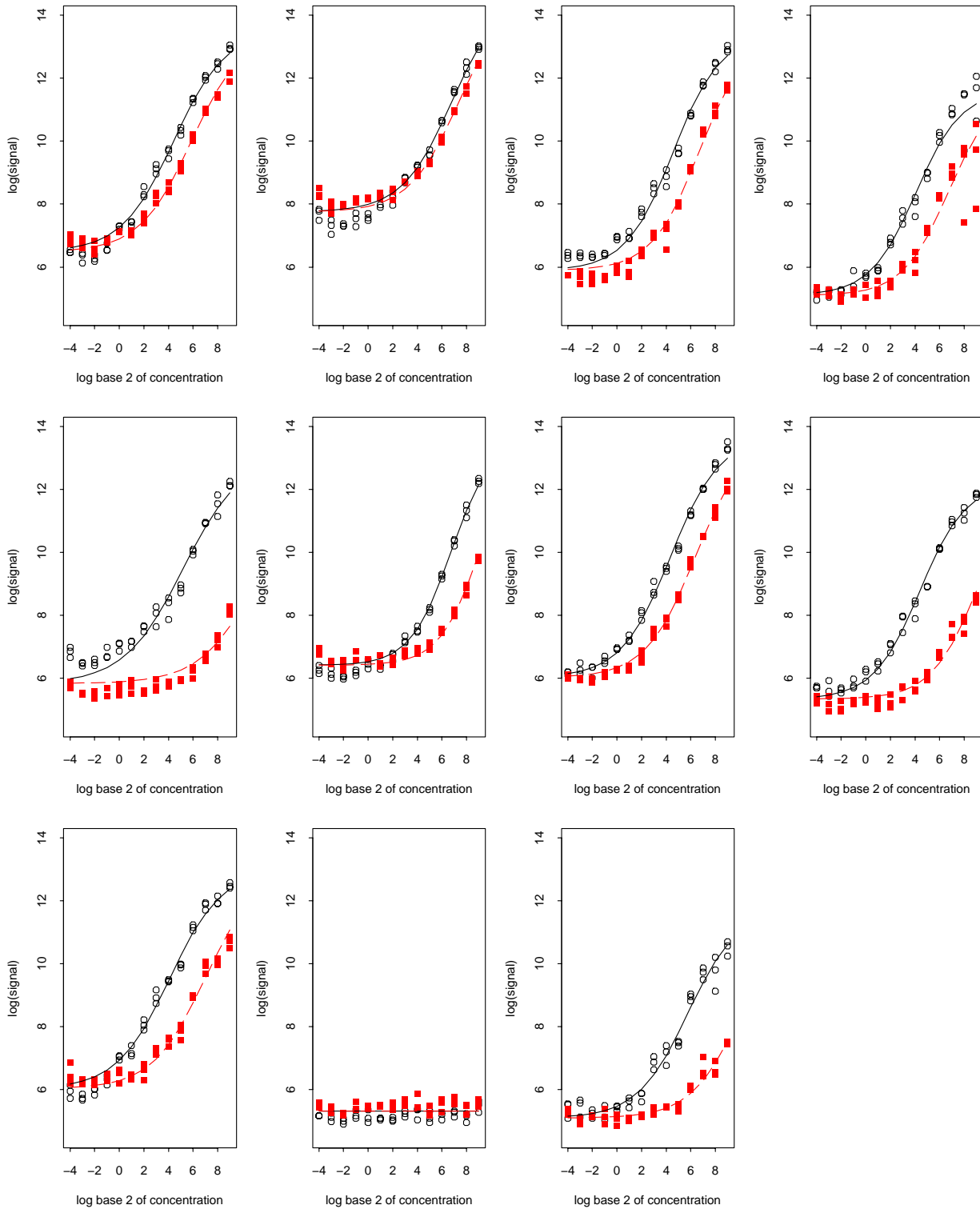
Affy U133 spikein, gene 20 (209734_at)



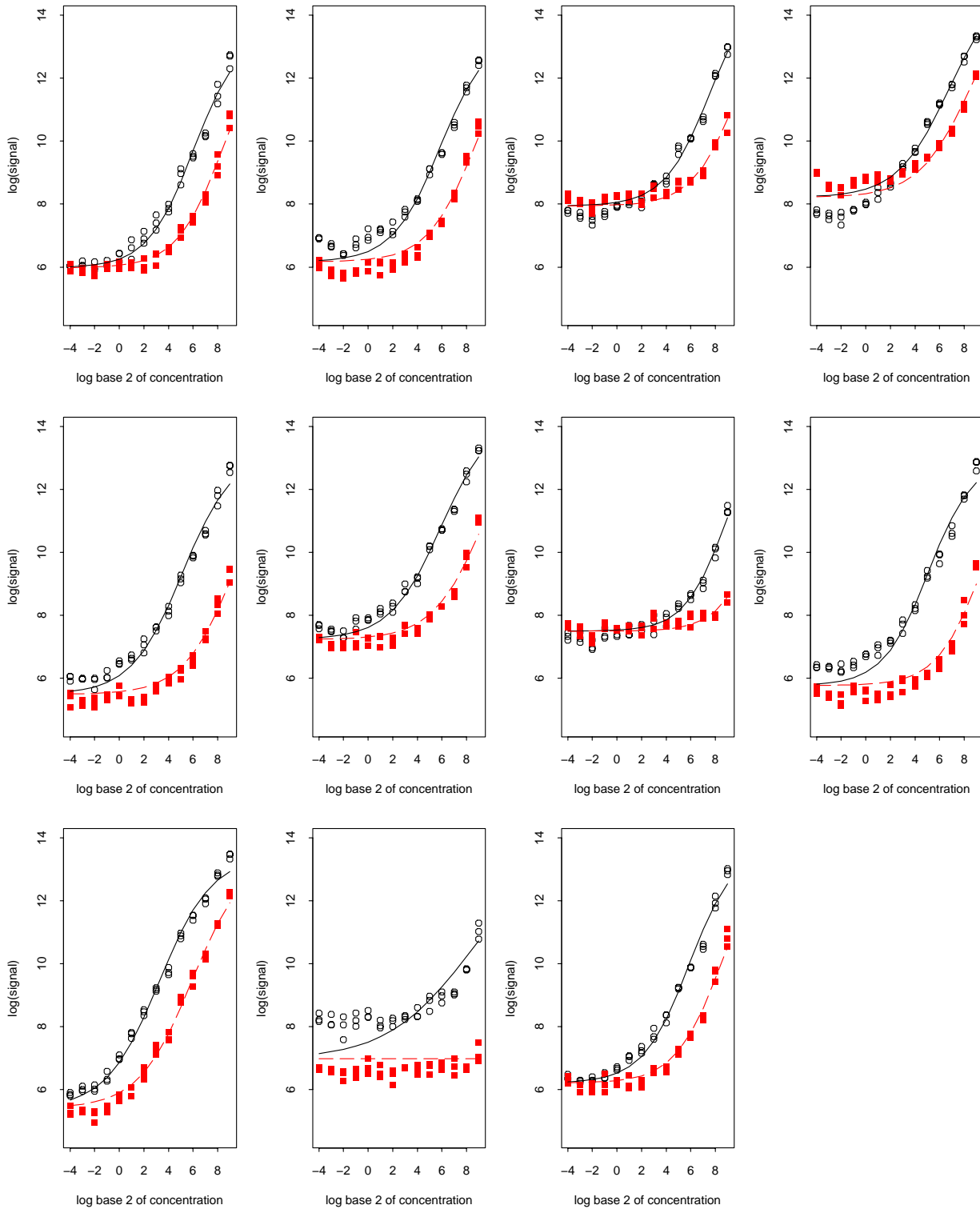
Affy U133 spikein, gene 21 (209354_at)



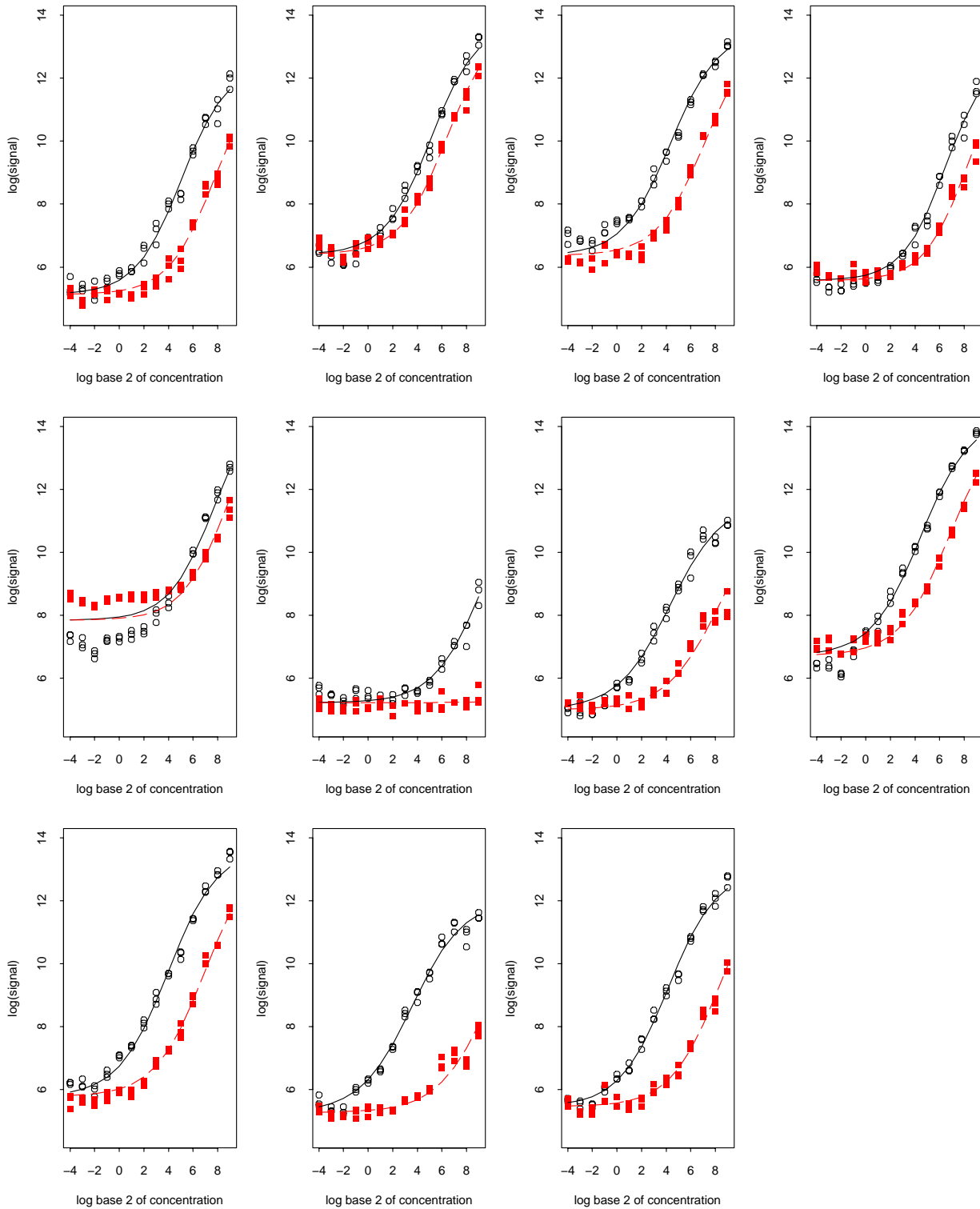
Affy U133 spikein, gene 22 (206060_s_at)



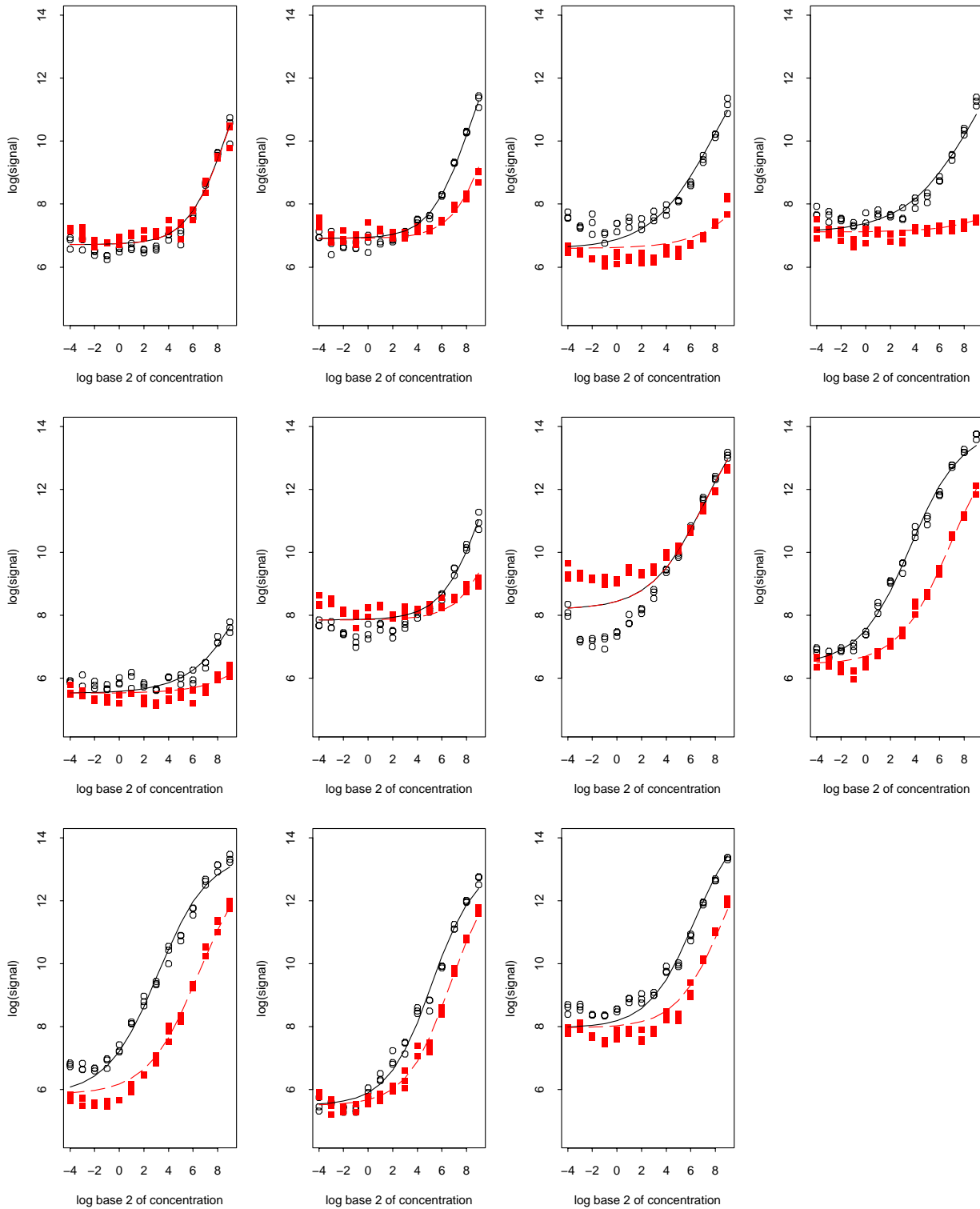
Affy U133 spikein, gene 23 (205790_at)



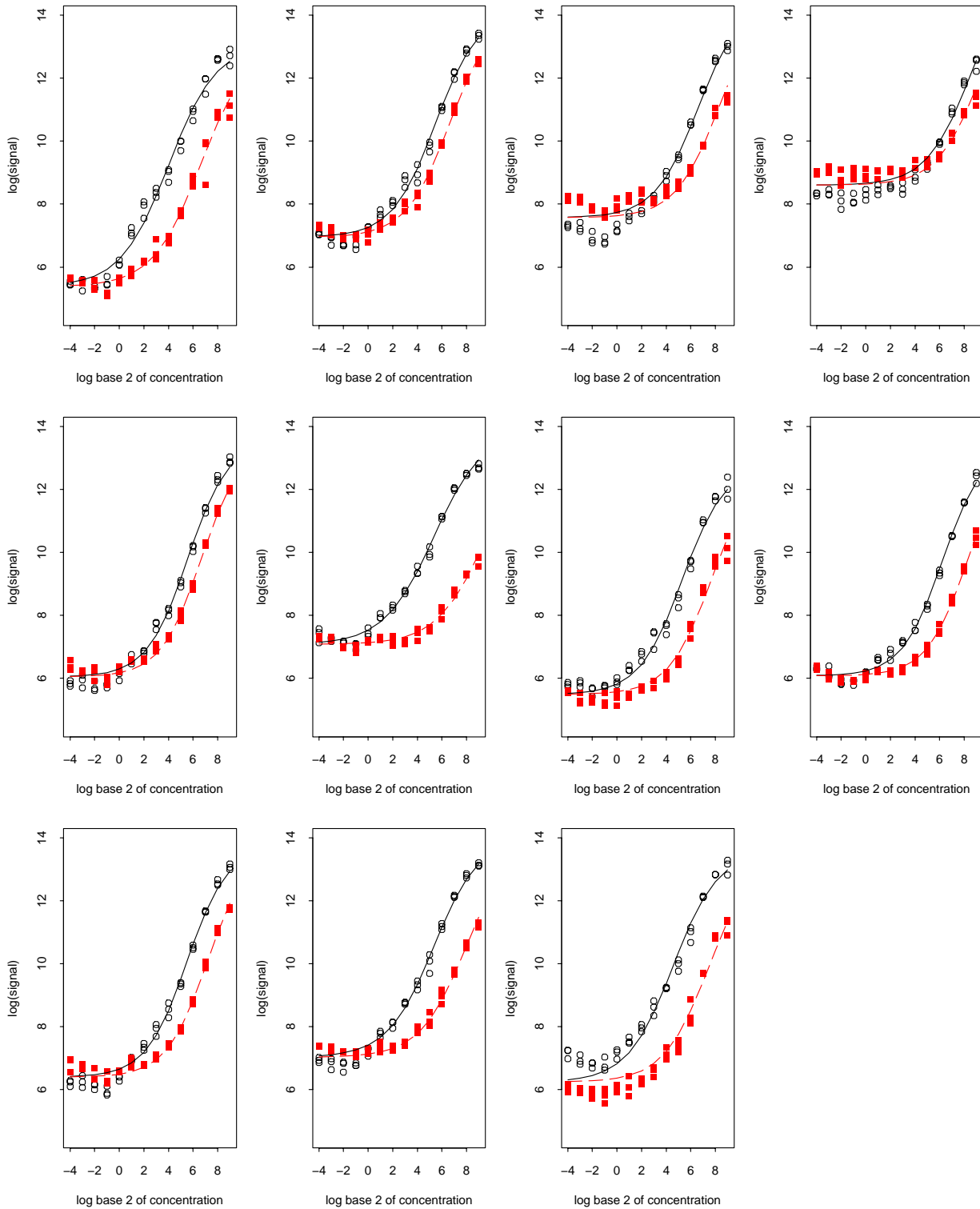
Affy U133 spikein, gene 24 (200665_s_at)



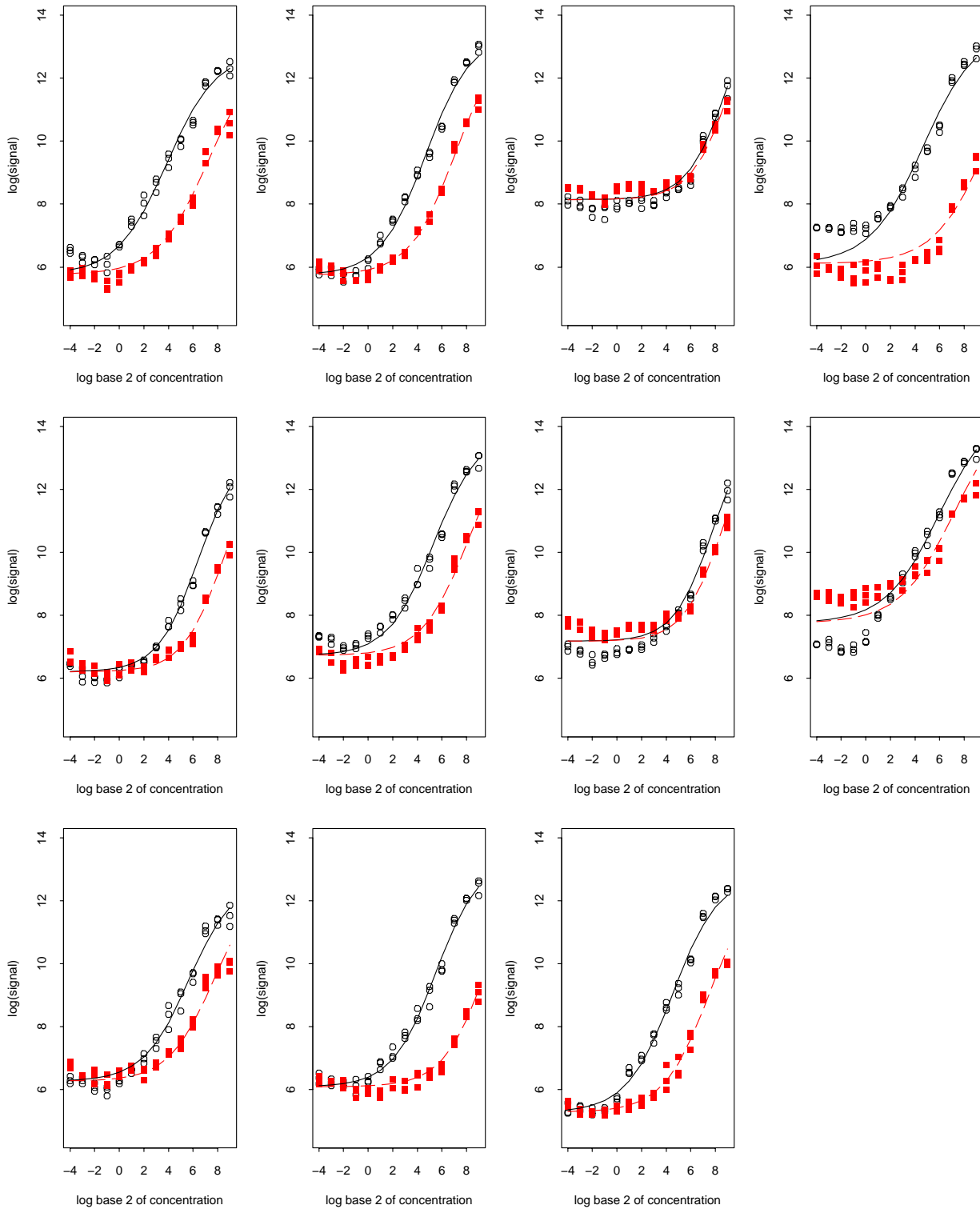
Affy U133 spikein, gene 25 (207641_at)



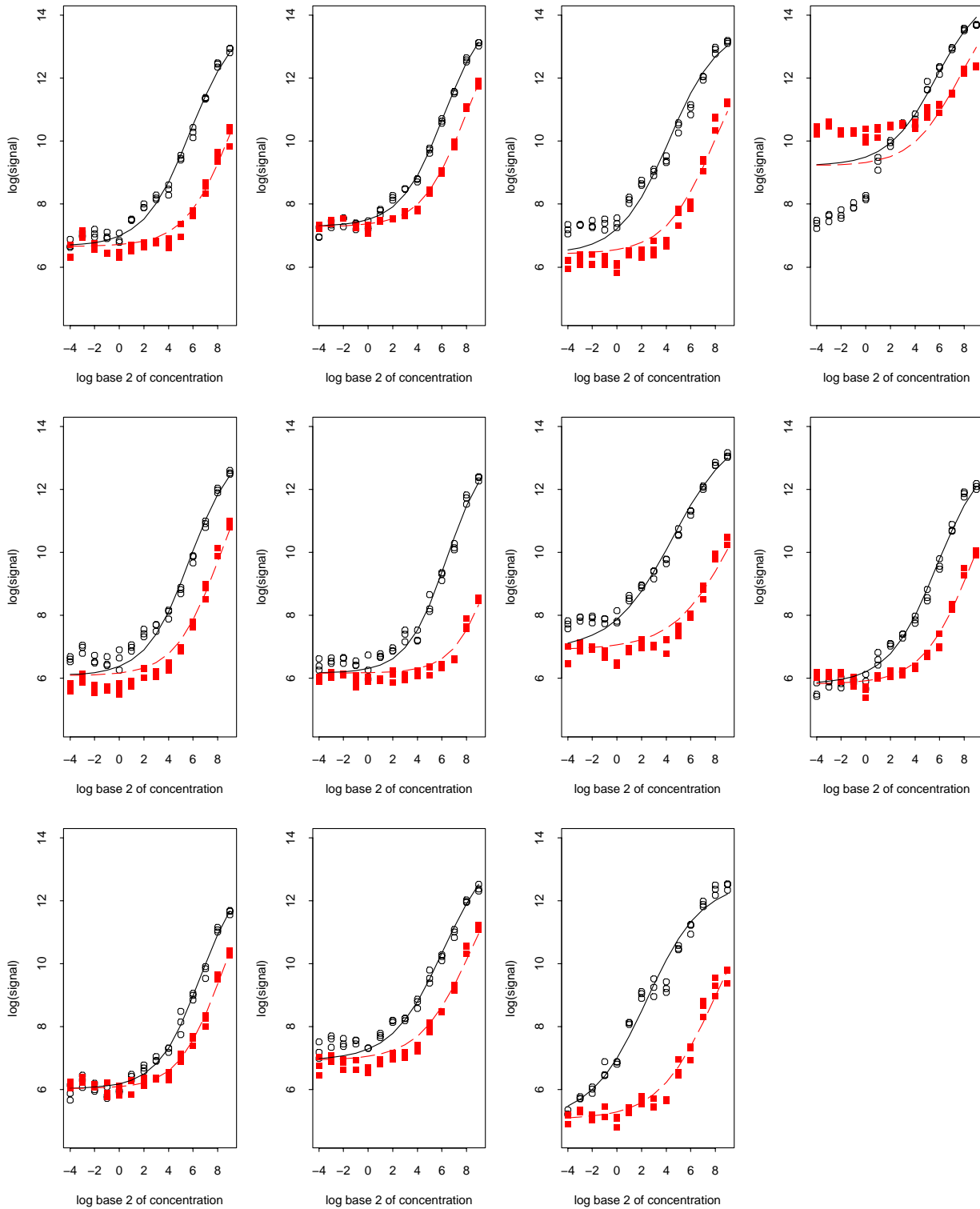
Affy U133 spikein, gene 26 (207540_s_at)



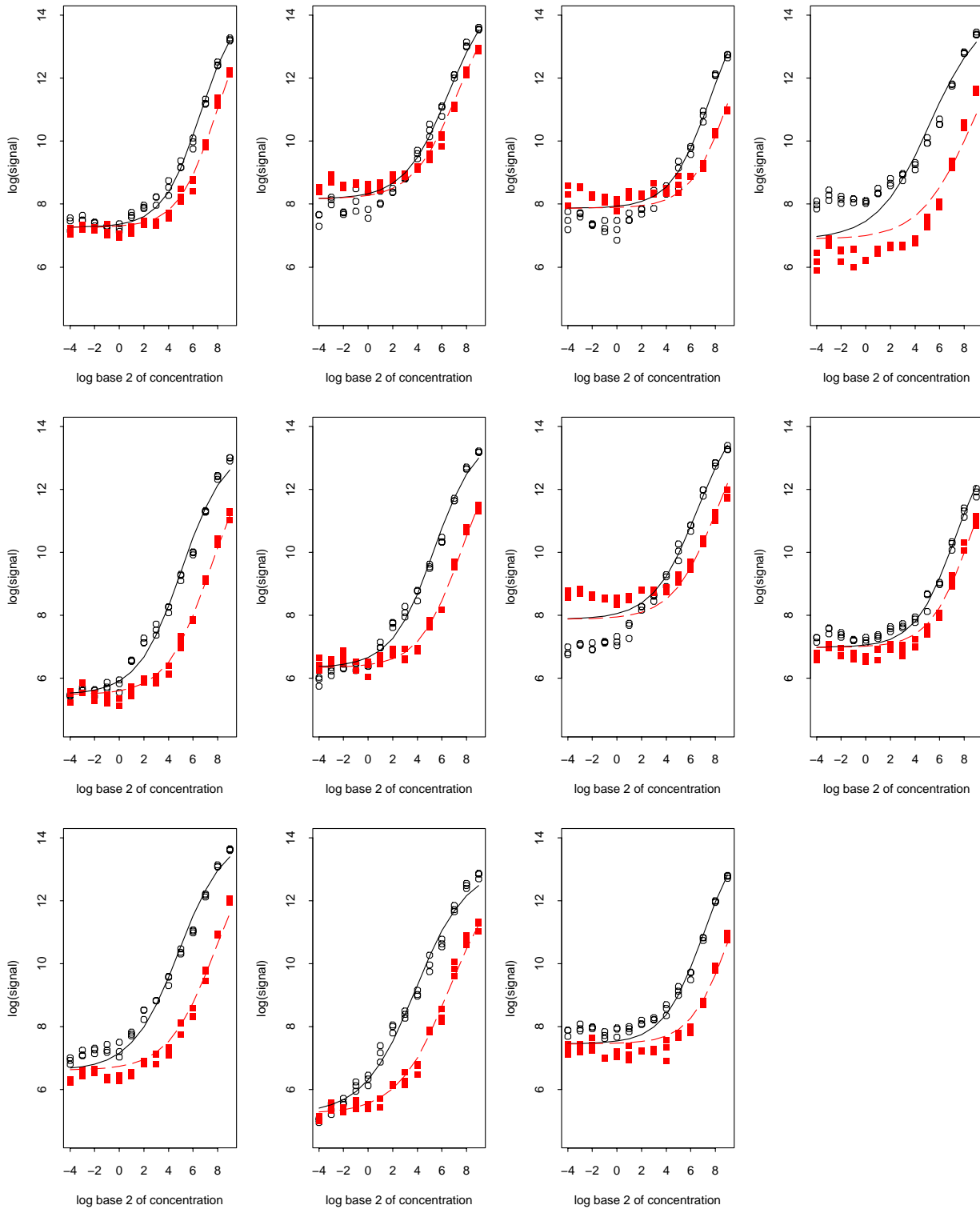
Affy U133 spikein, gene 27 (204430_s_at)



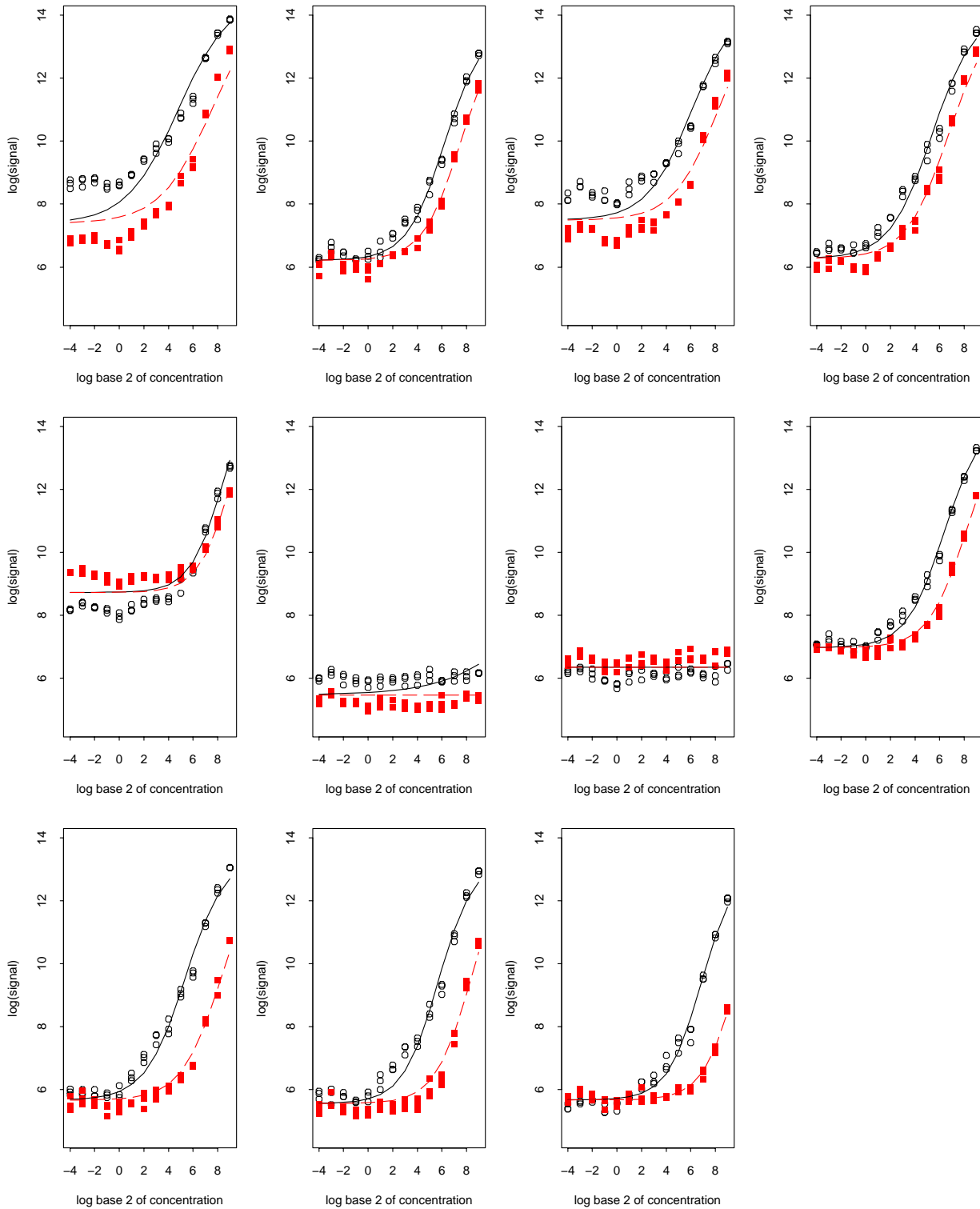
Affy U133 spikein, gene 28 (203471_s_at)



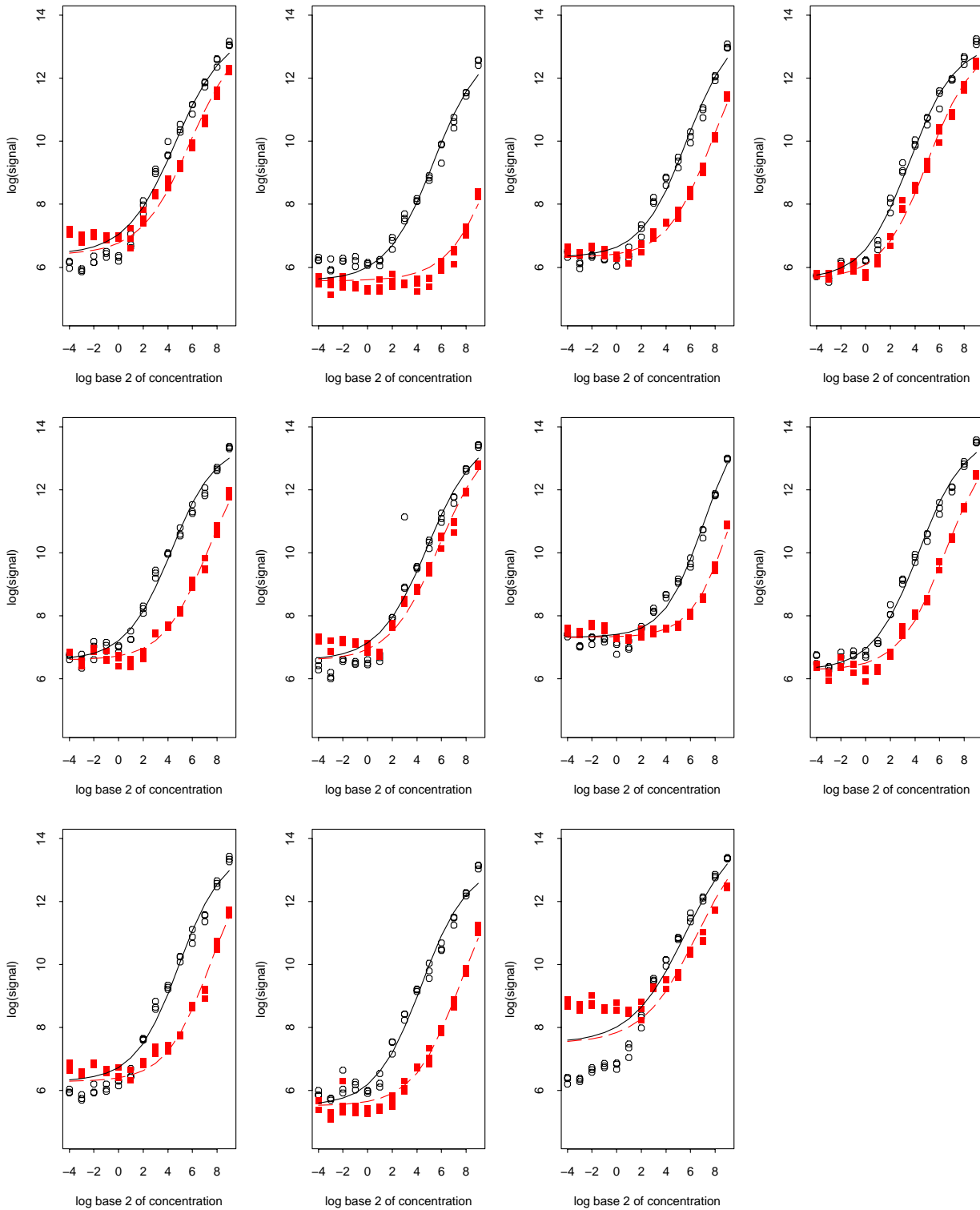
Affy U133 spikein, gene 29 (204951_at)



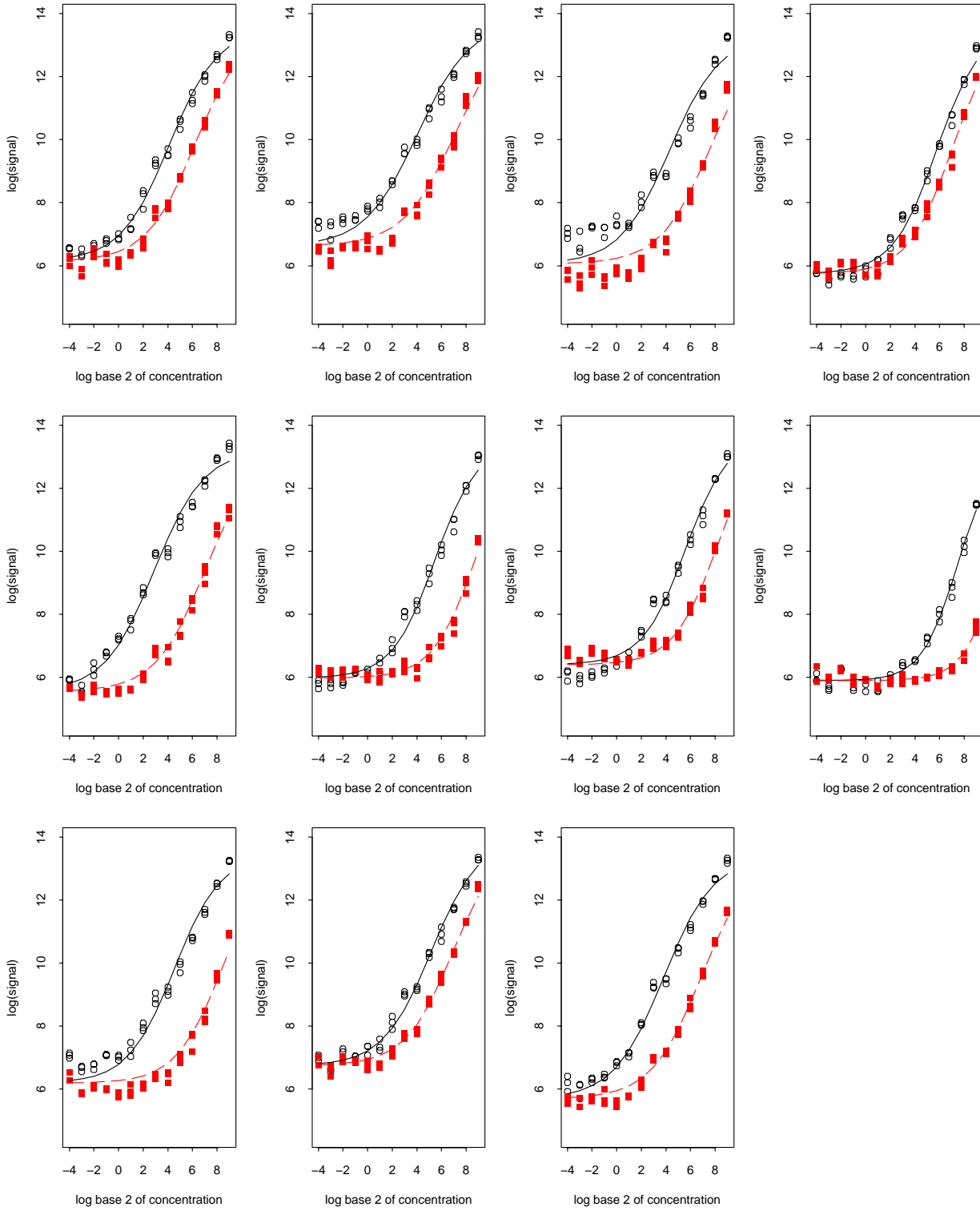
Affy U133 spikein, gene 30 (207968_s_at)



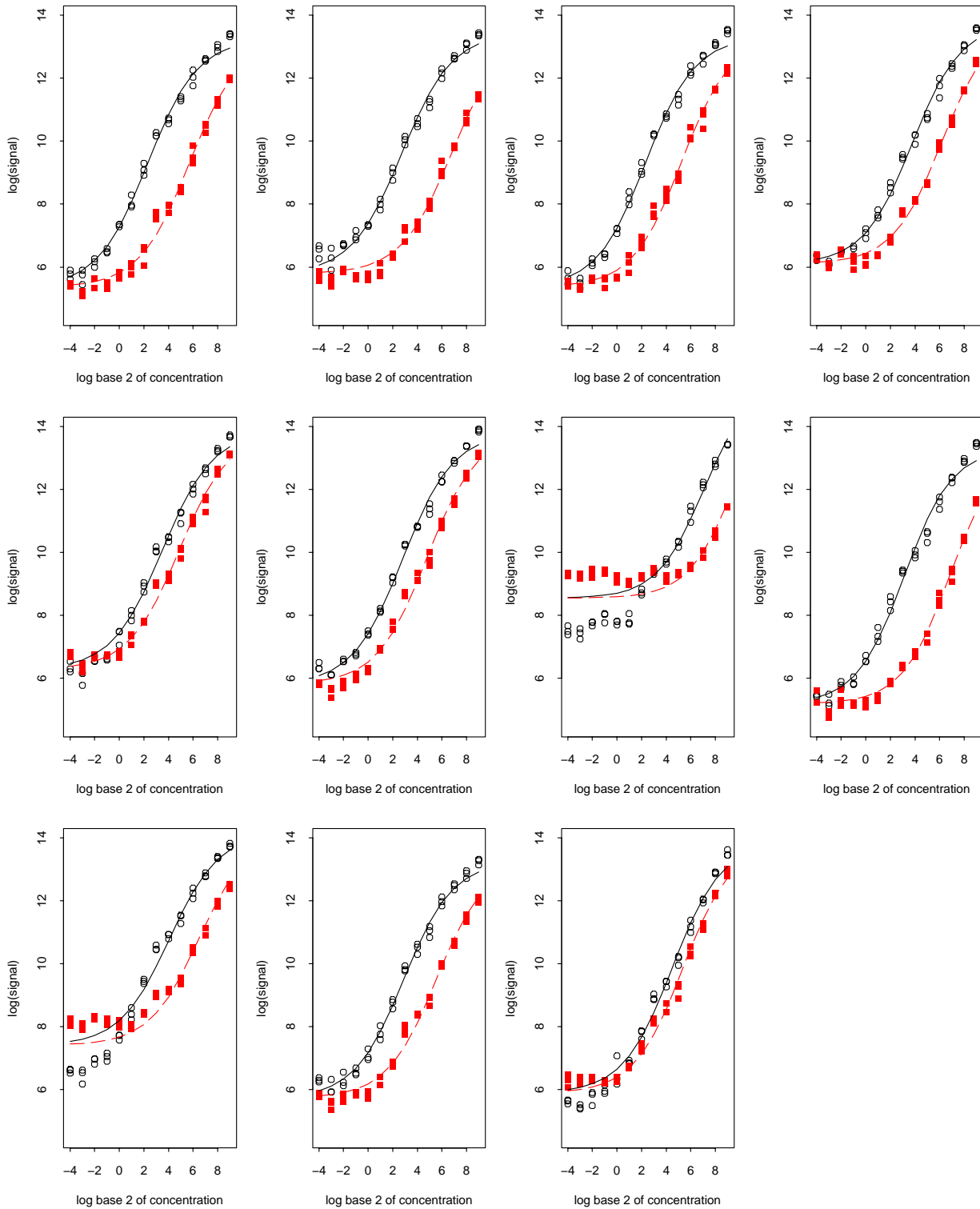
Affy U133 spikein, gene 31 (AFFX-r2-TagA_at)



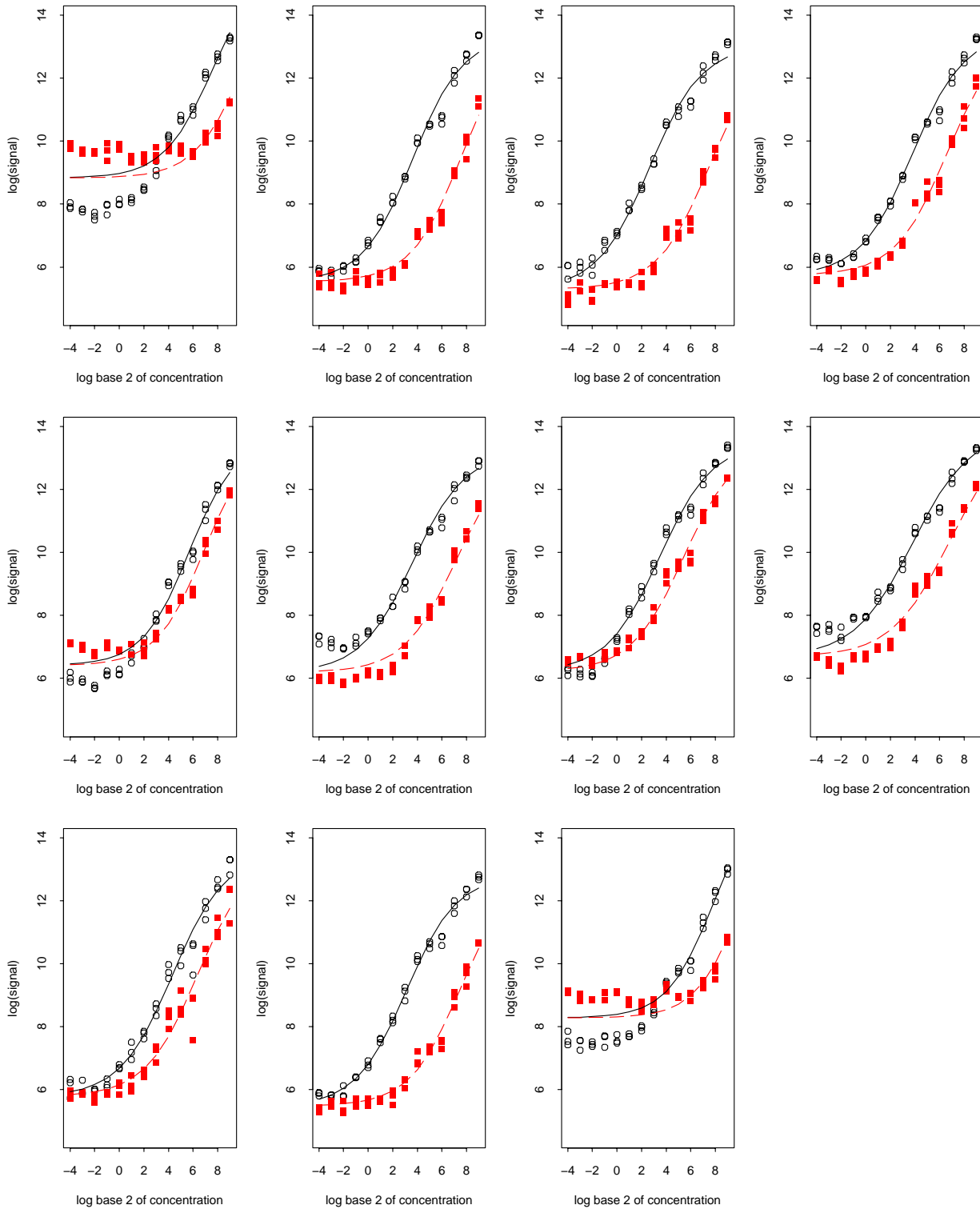
Affy U133 spikein, gene 32 (AFFX-r2-TagB_at)



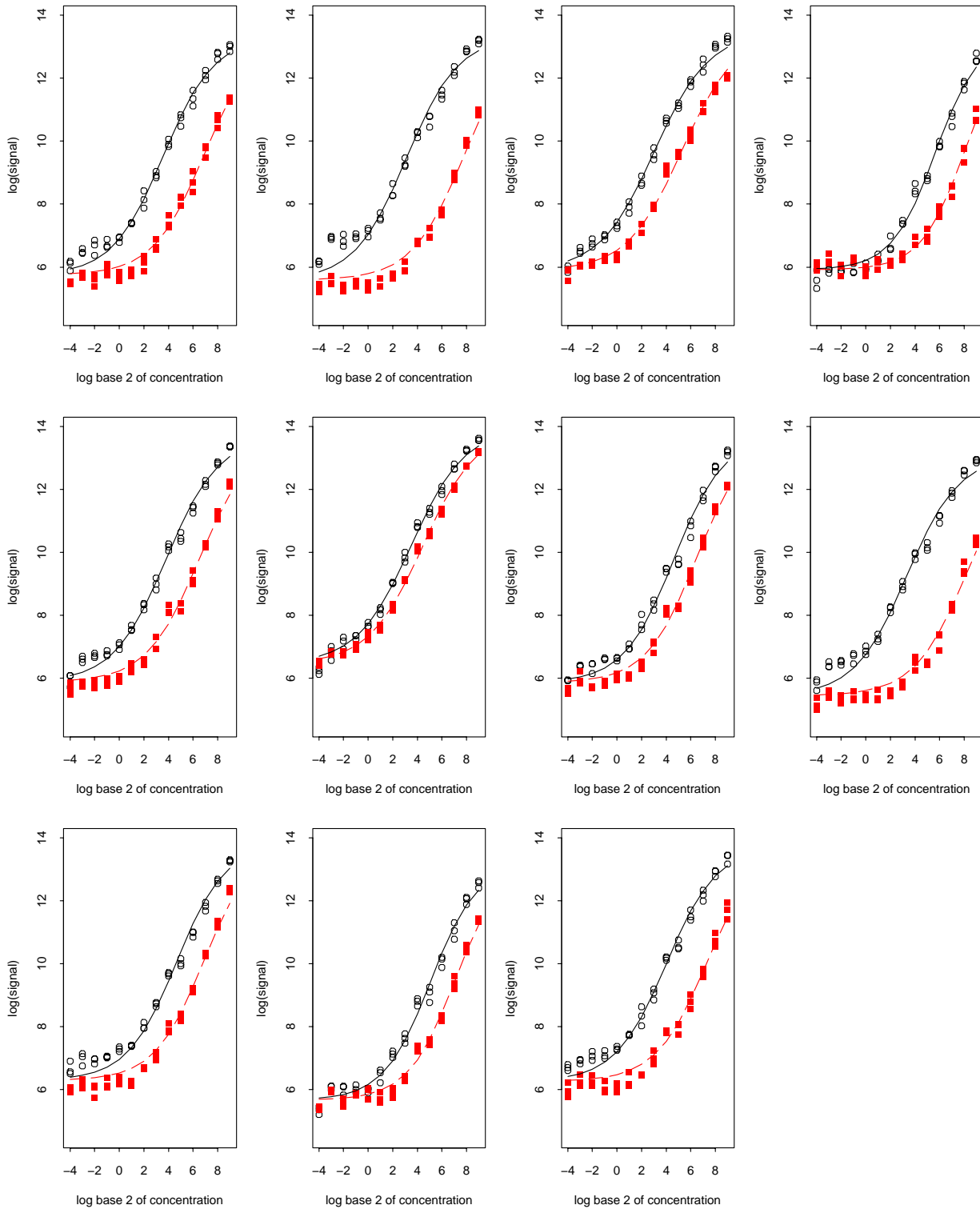
Affy U133 spikein, gene 33 (AFFX-r2-TagC_at)



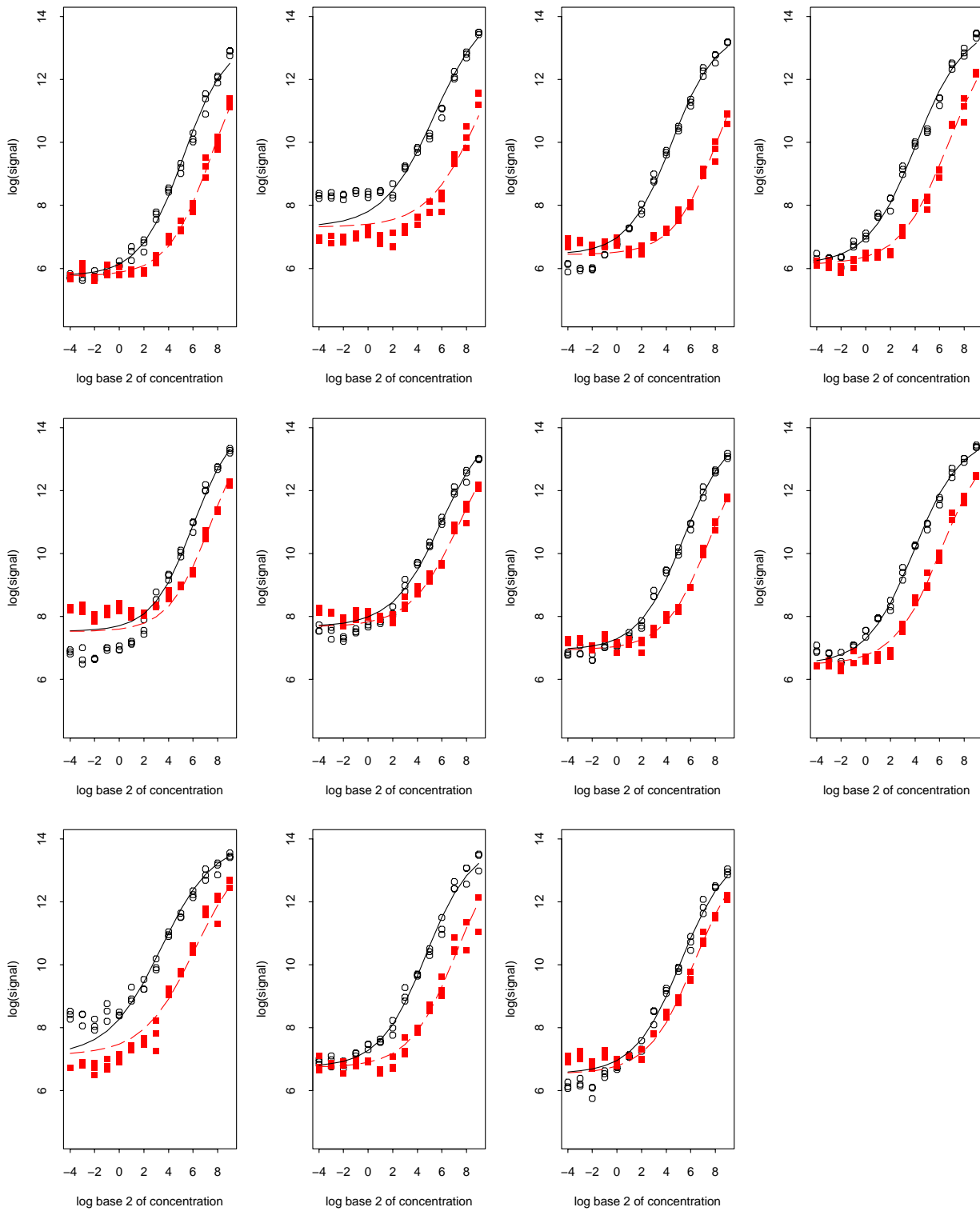
Affy U133 spikein, gene 34 (AFFX-r2-TagD_at)



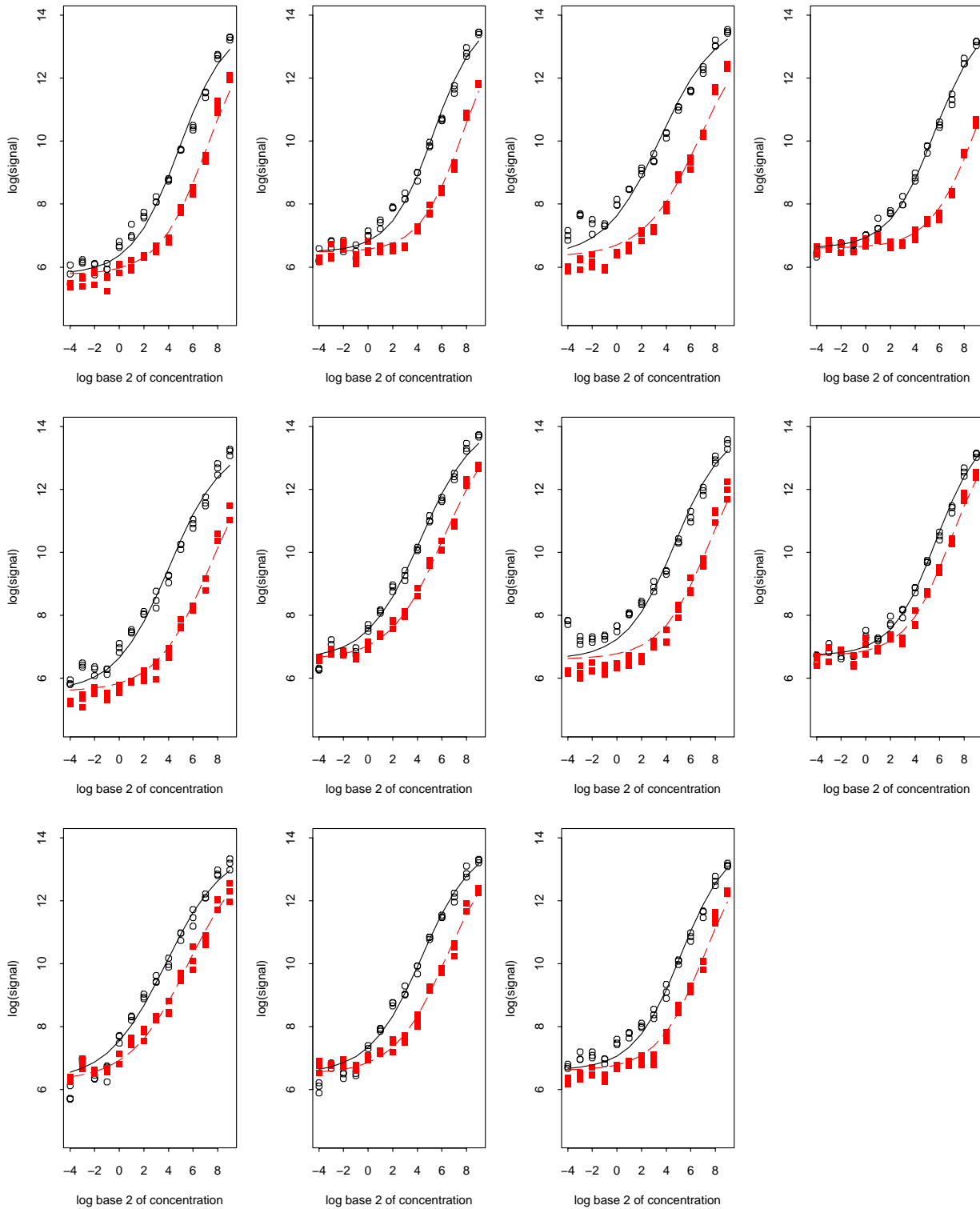
Affy U133 spikein, gene 35 (AFFX-r2-TagE_at)



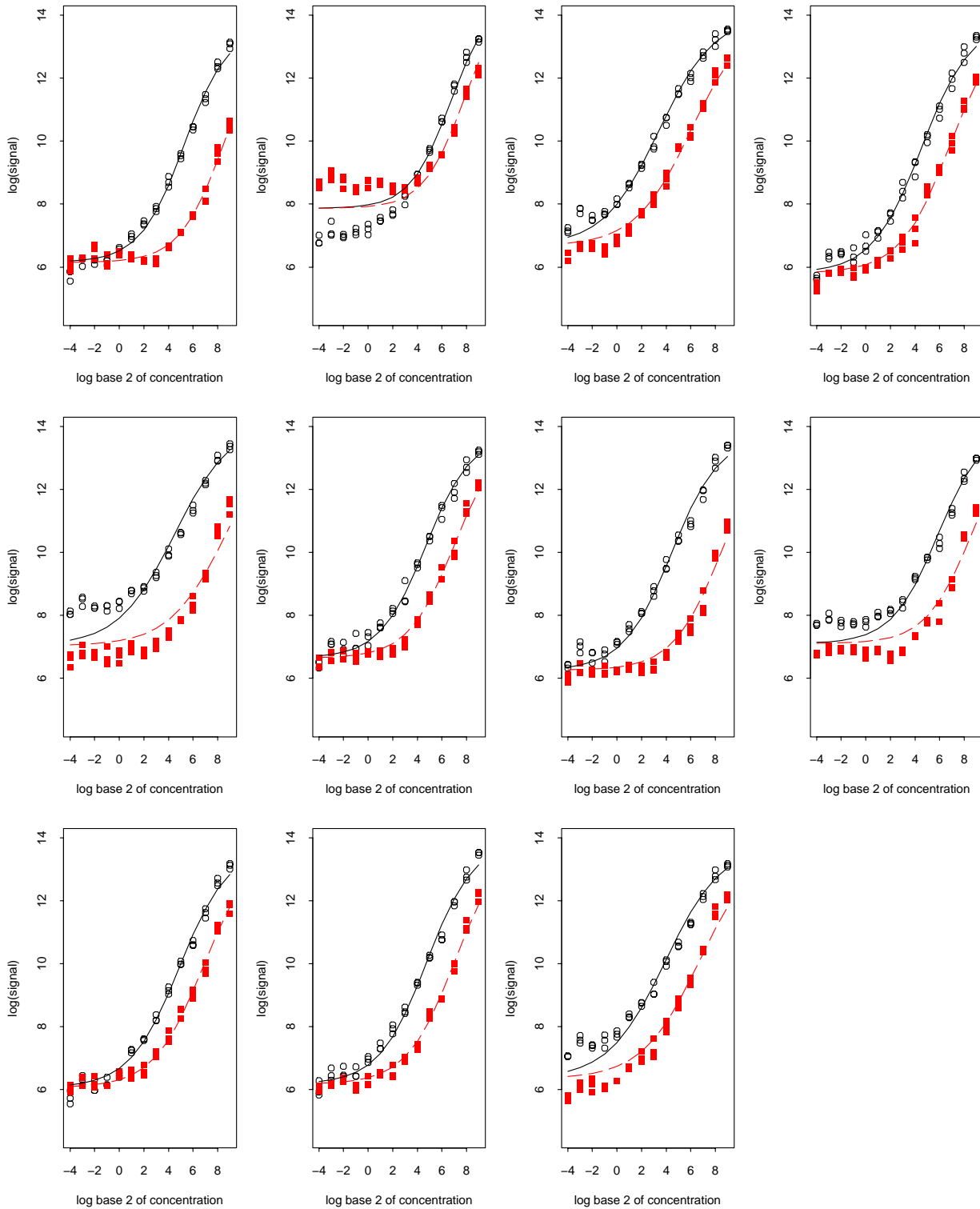
Affy U133 spikein, gene 36 (AFFX-r2-TagF_at)



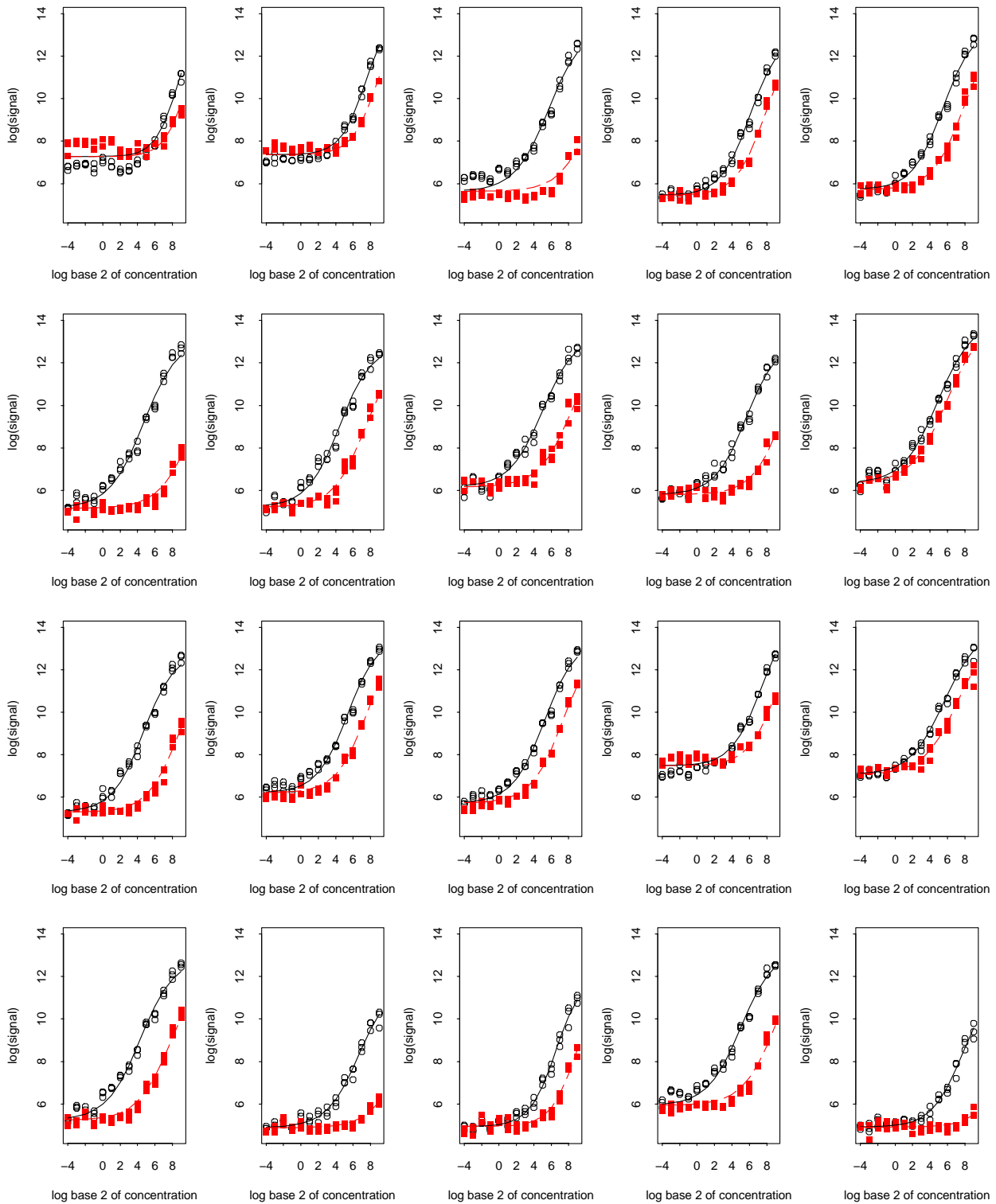
Affy U133 spikein, gene 37 (AFFX-r2-TagG_at)



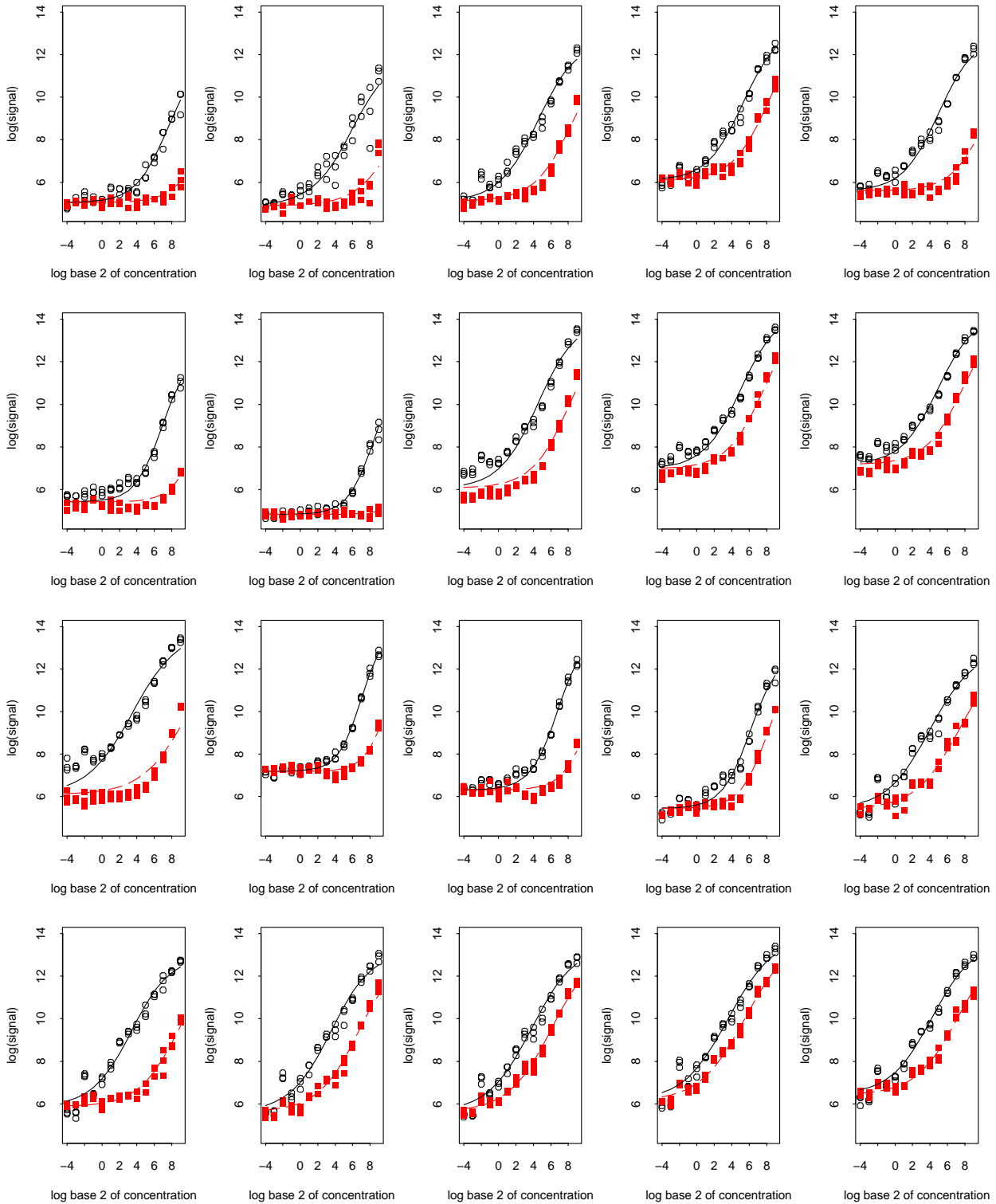
Affy U133 spikein, gene 38 (AFFX-r2-TagH_at)



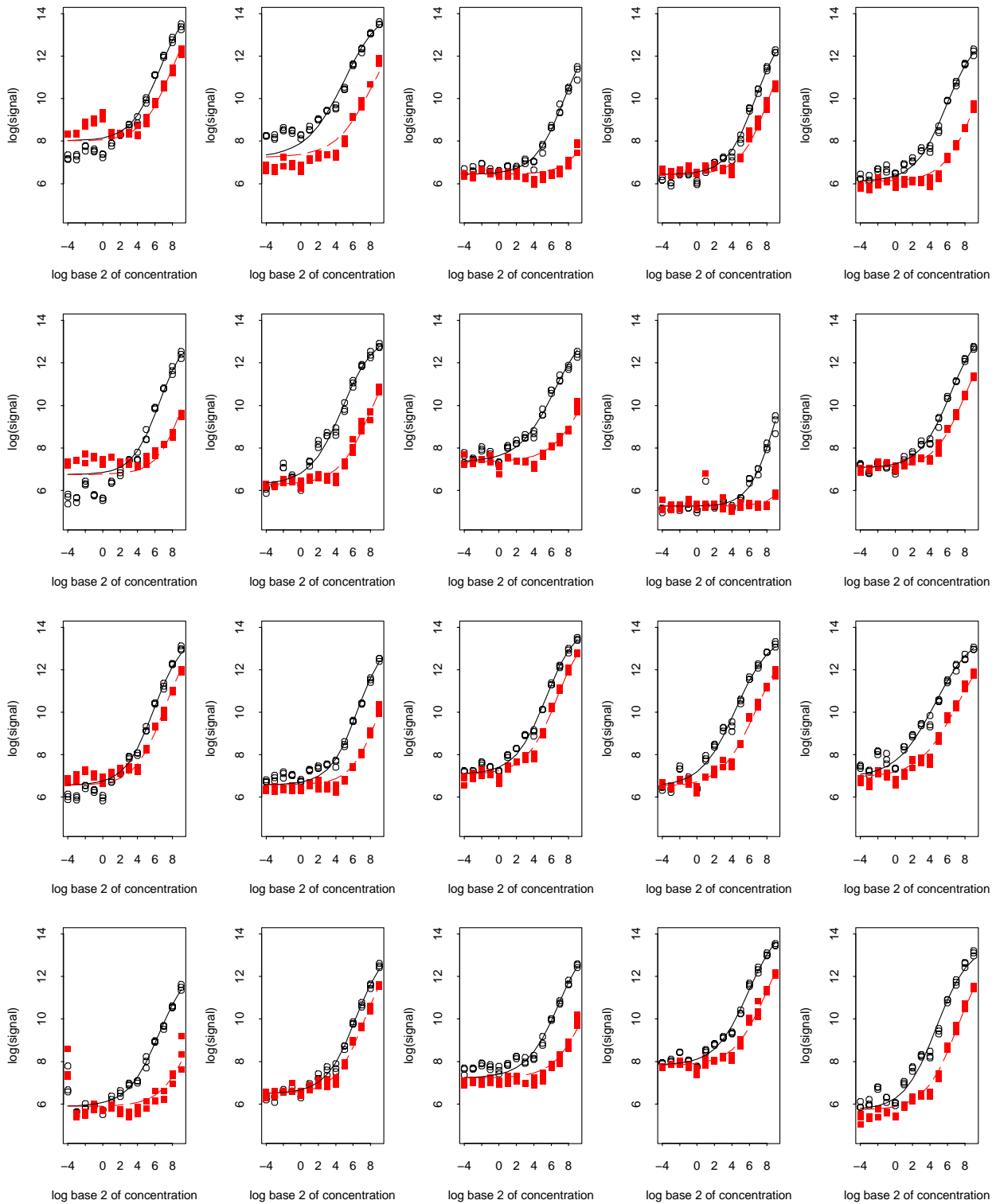
Affy U133 spikein, gene 39 (AFFX-DapX-3_at)



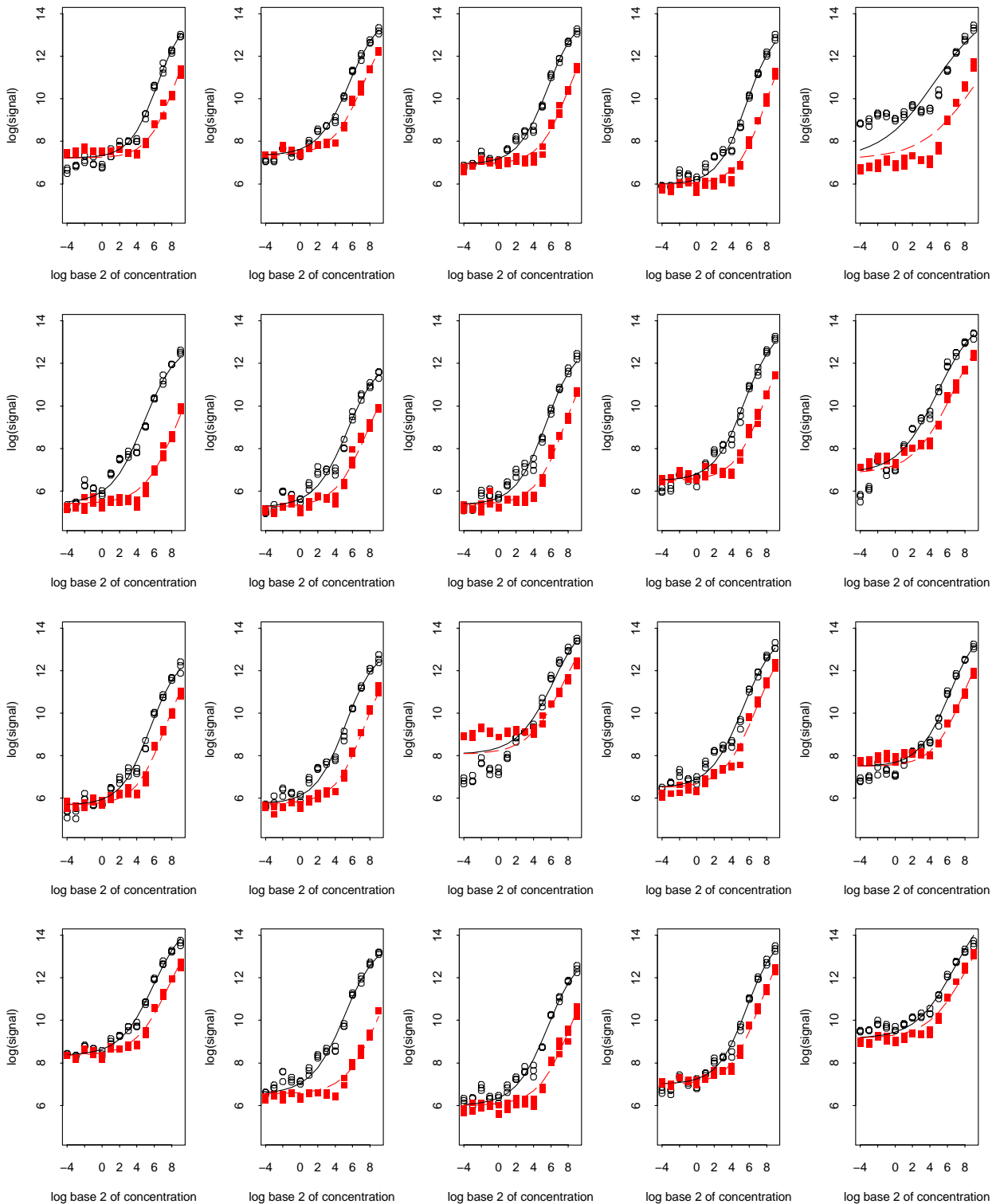
Affy U133 spikein, gene 40 (AFFX-LysX-3_at)



Affy U133 spikein, gene 41 (AFFX-PheX-3_at)



Affy U133 spikein, gene 42 (AFFX-ThrX-3_at)



References

- [1] K. V. Ballman and T. M. Therneau. A exploration of affymetrix probe-set intensities in spike-in experiments. Technical Report 74, Mayo Clinic College of Medicine, March 2005.