LemonTree performance benchmarks

1.   AUC analysis of the regulators performance using synthetic data

The following figure report on analyses that condense more than the 2 dimensions of the Cartesian plane, up to 5 variables. All variables are described in the legend and a brief discussion on the emerging trend is given. The AUC and ROC plots are produced by the R package "ROCR", v1.0-7 (Sing *et al.*, 2005) .
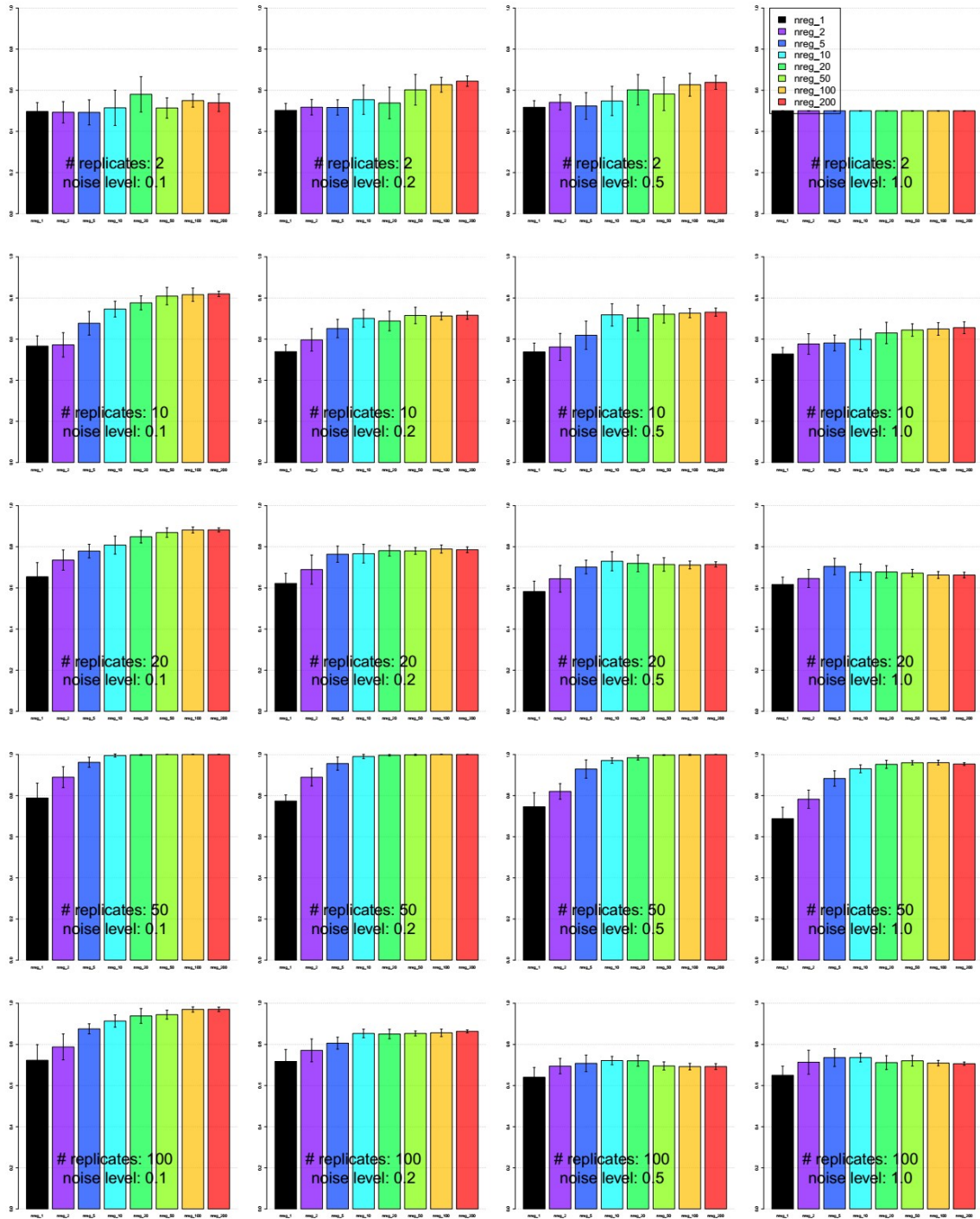
Fig S4-1. AUCs of the "random regulator lists". Rows *bn*, columns *nreplicate*. For each subfigure, *y*-axis AUC, *x*-axis, colors *nreg*.

Fig S4-1 shows that AUCs of the random assigned regulators are generally above 0.5, though lower than their non-random counterparts.

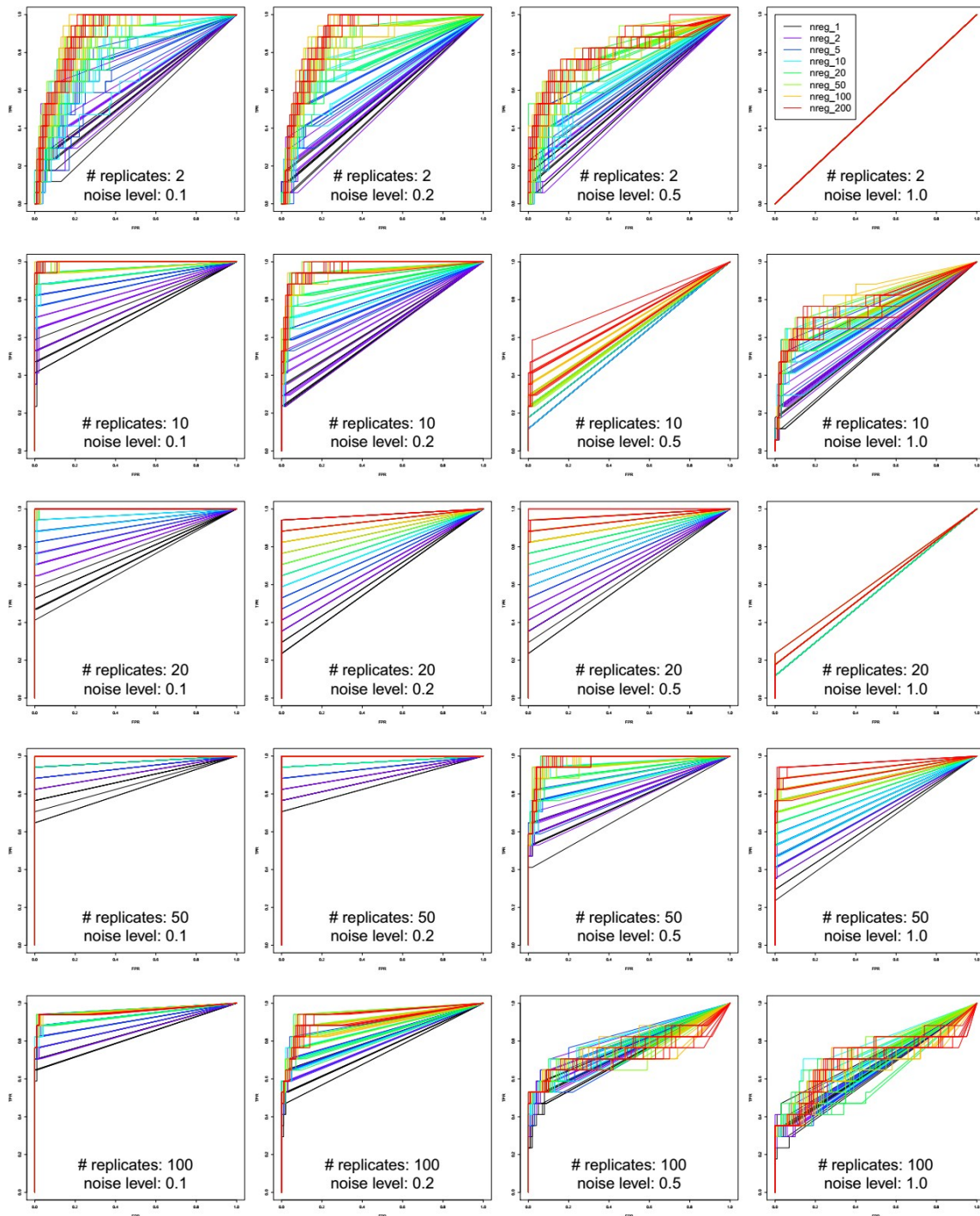## 2.   ROC analysis of the regulator performance using synthetic data

Fig S4-2. ROC curves (TPR-FPR) of regulators assignment. Rows: *bn*, columns: *nreplicate*; for each subfigure, *x*-axis FPR, *y*-axis TPR, colors *nreg*.

Fig S4-2 depicts LemonTree's decent sensitivity-specificity   performance, except for a few high noise applications.

Note, the *nreplicate* = 2, *bn* = 1.0 data set displays a linear line, because no regulatory tree was obtained, possibly due to the extremely high noise and small sample size.
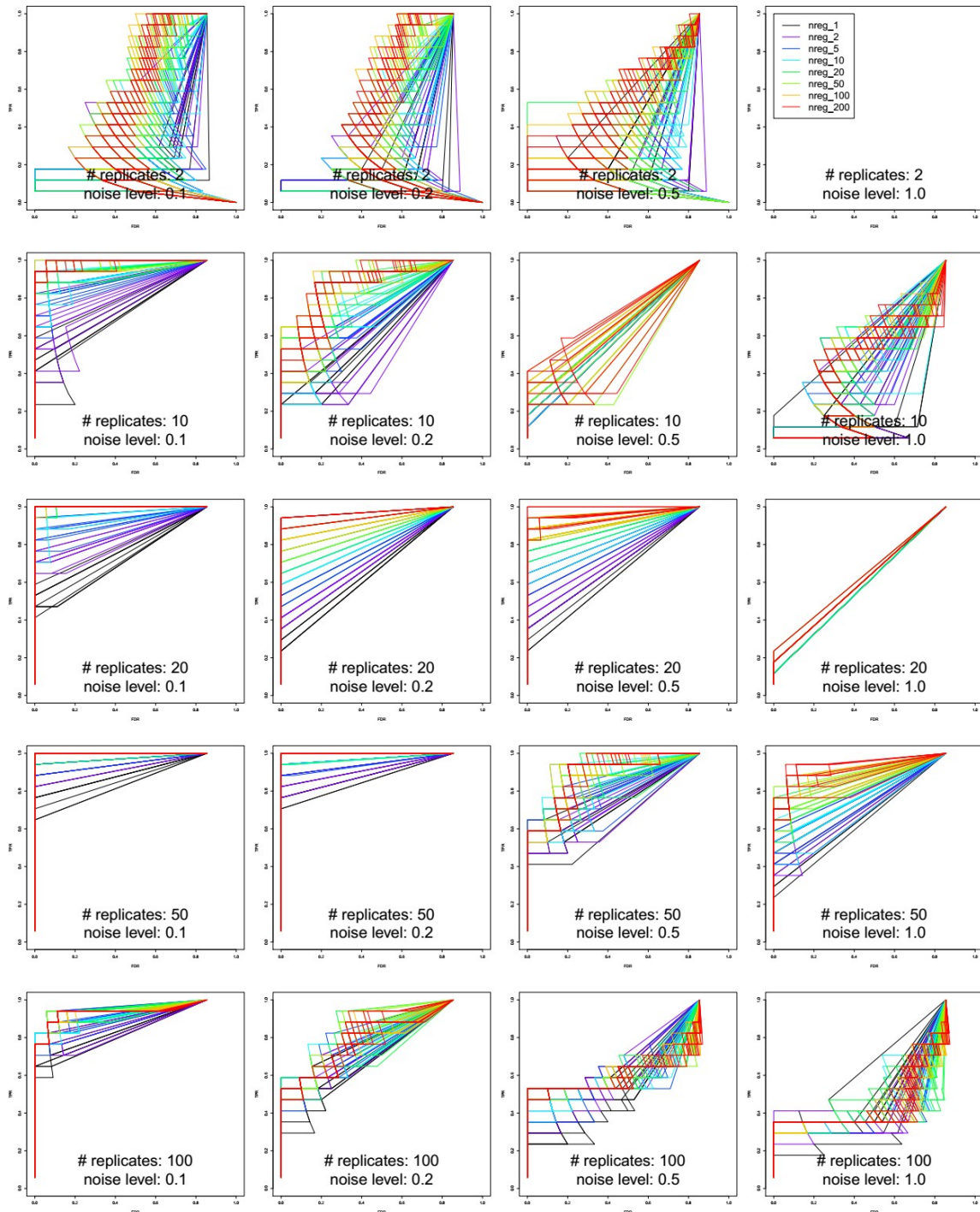

Figure S4-3. ROC curves (TPR-FDR) of regulators assignment. Rows: *bn*, columns: *nreplicate*; for each subfigure, *x*-axis FPR, *y*-axis TPR, colors *nreg*.

In contrast to generally good TPR-FPR performance, TPR-FDR curves in Fig S4-4 raises caution for LemonTree's application in small sample sized data. Nevertheless, applying LemonTree to data with 20~50 replicates and a modest noise level (*bn* < 0.2) should be fine.

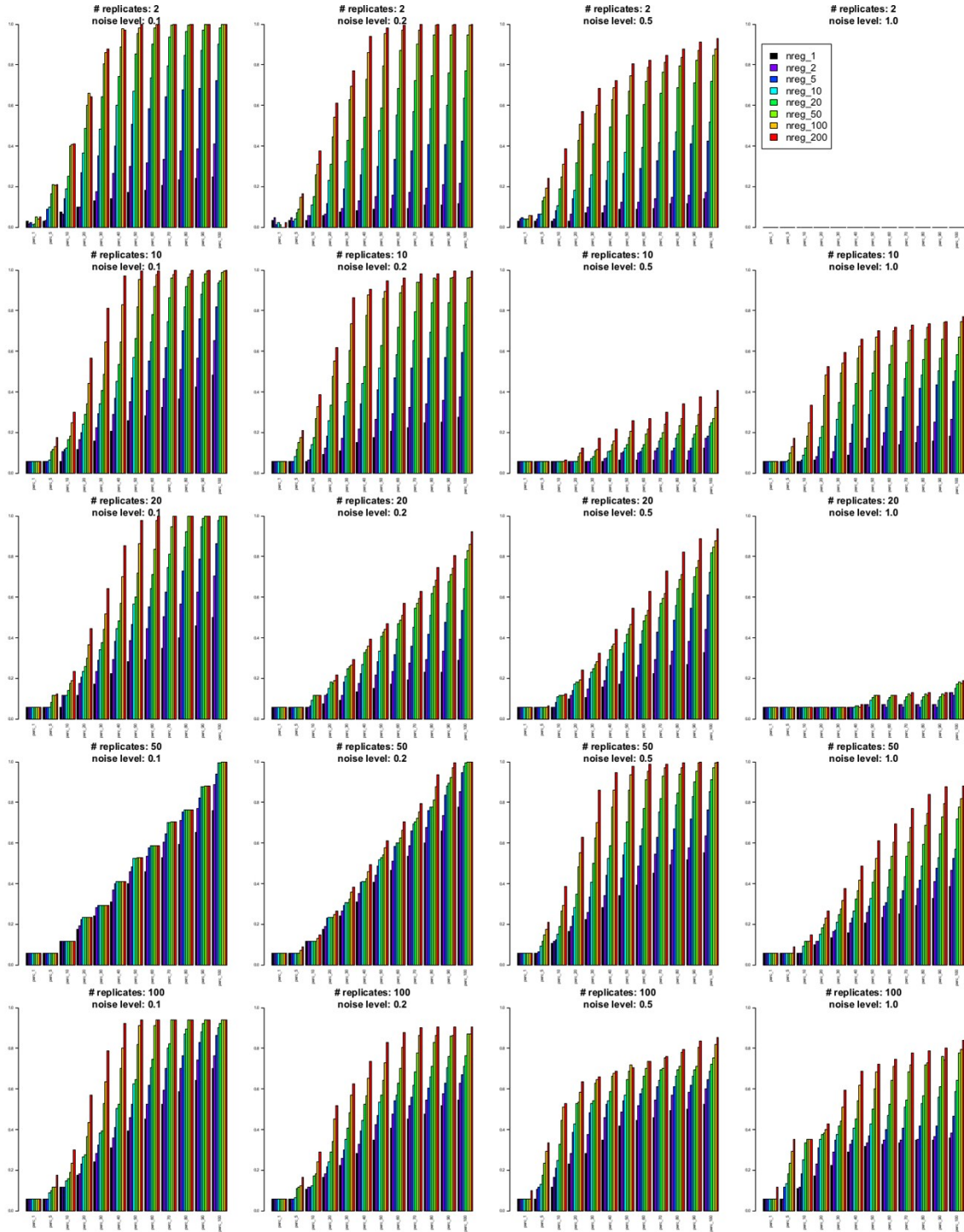3.   Dependence of TPR and FDR on *bn*, *nreplicate*, *nreg* using synthetic data

Fig S4-4. TPR dependence of regulators assignment. Rows *bn*, columns *nreplicate*; for each subfigure, *x*-axis *perc*, *y*-axis TPR, colors *nreg*.

In general, Fig S4-4 shows the trend that TPR increases with *nreg* and *perc*, while decreases with *bn,* except for a few datasets with high noise. However, LemonTree's default *perc* (1%) generally has a poor TPR (< 0.2). Therefore we recommend users to increase *perc* for the sake of a better sensitivity.
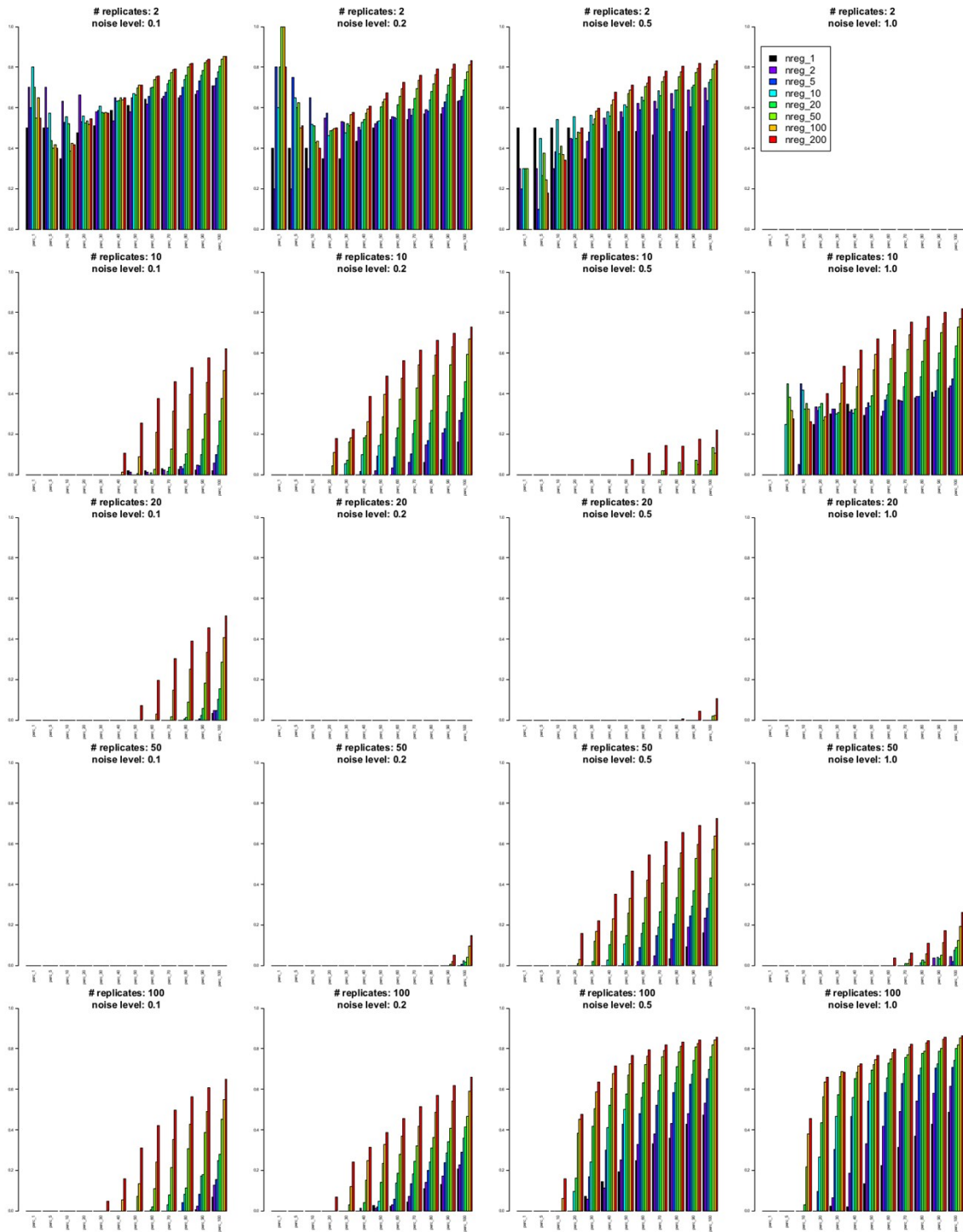
Fig S4-5. FDR dependence of regulators assignment. Rows *bn*, columns *nreplicate*; for each subfigure, *x*-axis *perc*, *y*-axis FDR, colors *nreg*.

Figure S4-5 shows the TPR-FDR trends that given the same sensitivity FDR generally increases with noise level, and decreases with sample size. Such pattern fades when the noise gets

heavier or *nreplicate* reaches 100. We note that for datasets with limited sample size (*nreplicate* = 2), FDR is constantly high, even with the most stringent *perc* (= 1%).

Reference

Sing,T. *et al.* (2005) ROCR: visualizing classifier performance in R. *Bioinforma. Oxf. Engl.*, **21**, 3940–3941.