

Figure S1: Spearman correlation between total number of contacts per bin and independent copy number estimation (COSMIC) for each of the methods compared. Left panel T47D breast cancer cell line, right panel K562 leukemia cell line. The new proposal (in blue) outperforms the rest of alternatives.

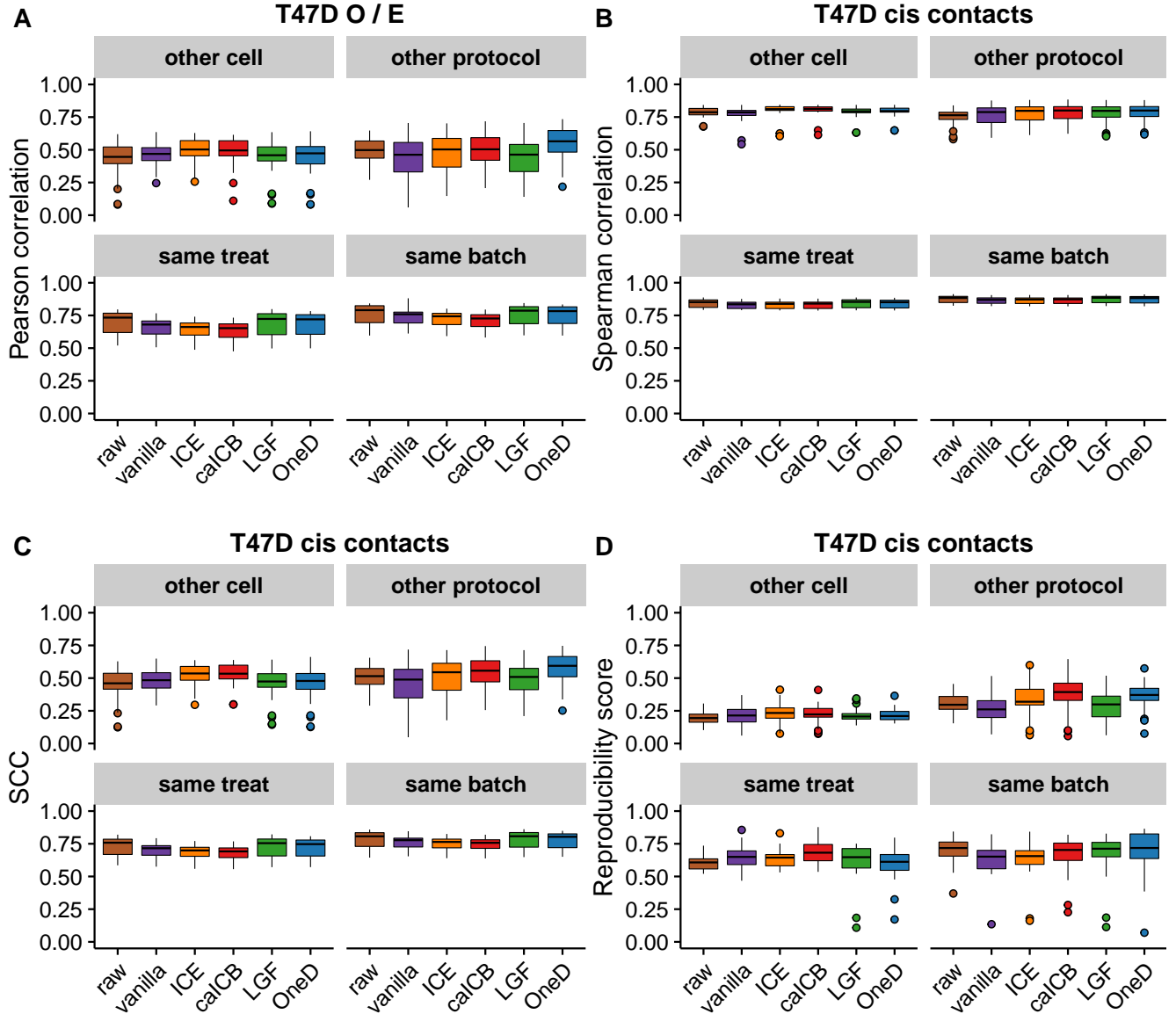


Figure S2: Boxplots representing the results of the pair-wise comparisons of the samples included in the T47D set. X axis: different correction methods. Y axis: correlation. Each panel groups pairs of samples with the corresponding characteristics (in terms of cell type, protocol, batch and treatment). A. Pearson correlation between observed over expected counts. B. Spearman correlation between observed counts. C. Stratum-adjusted correlation coefficient (SCC) between observed counts. D. Reproducibility score of observed counts.

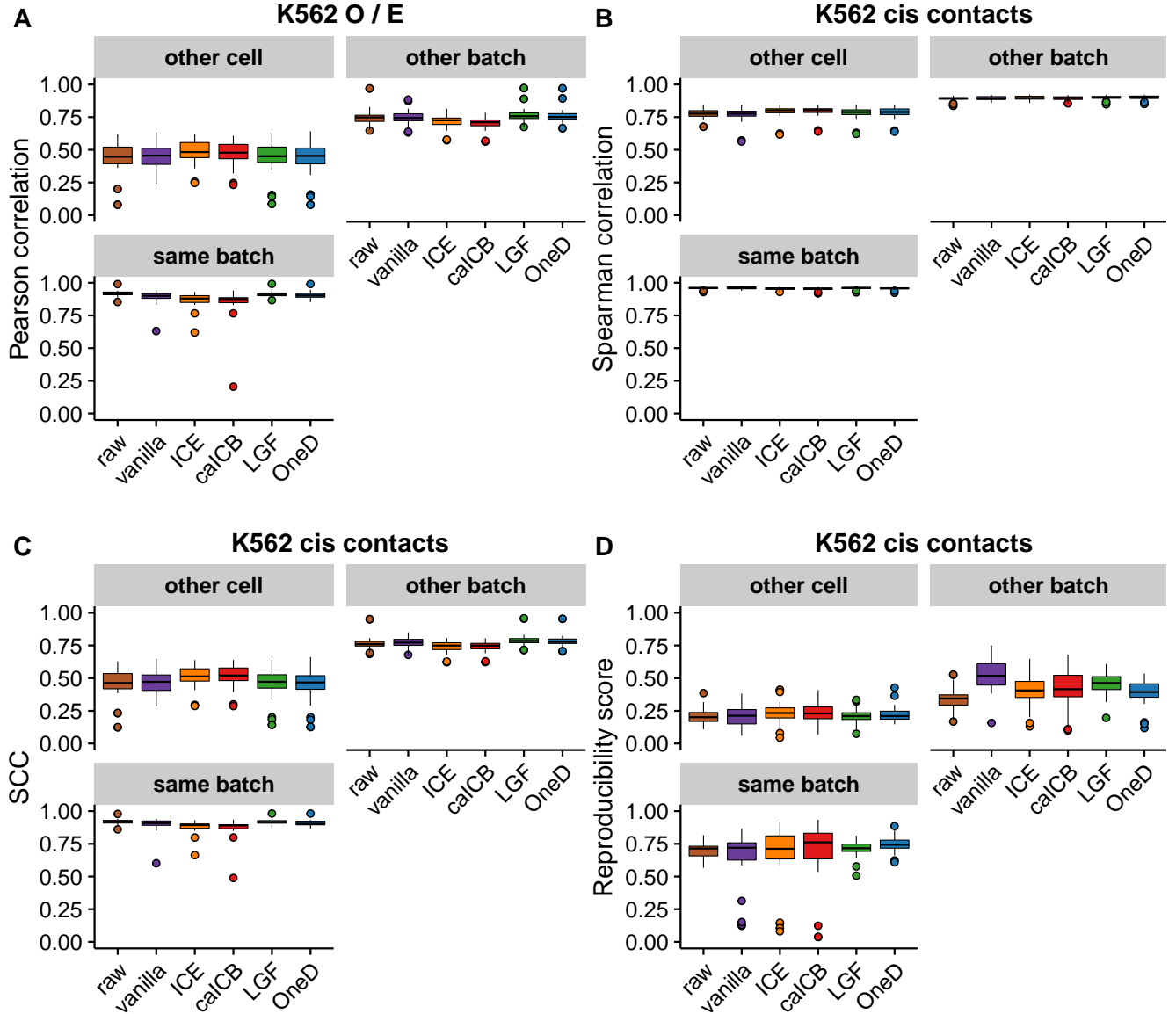


Figure S3: Boxplots representing the results of the pair-wise comparisons of the samples included in the K562 set. X axis: different correction methods. Y axis: correlation. Each panel groups pairs of samples with the corresponding characteristics (in terms of cell type, protocol, batch and treatment). A. Pearson correlation between observed over expected counts. B. Spearman correlation between observed counts. C. Stratum-adjusted correlation coefficient (SCC) between observed counts. D. Reproducibility score of observed counts.

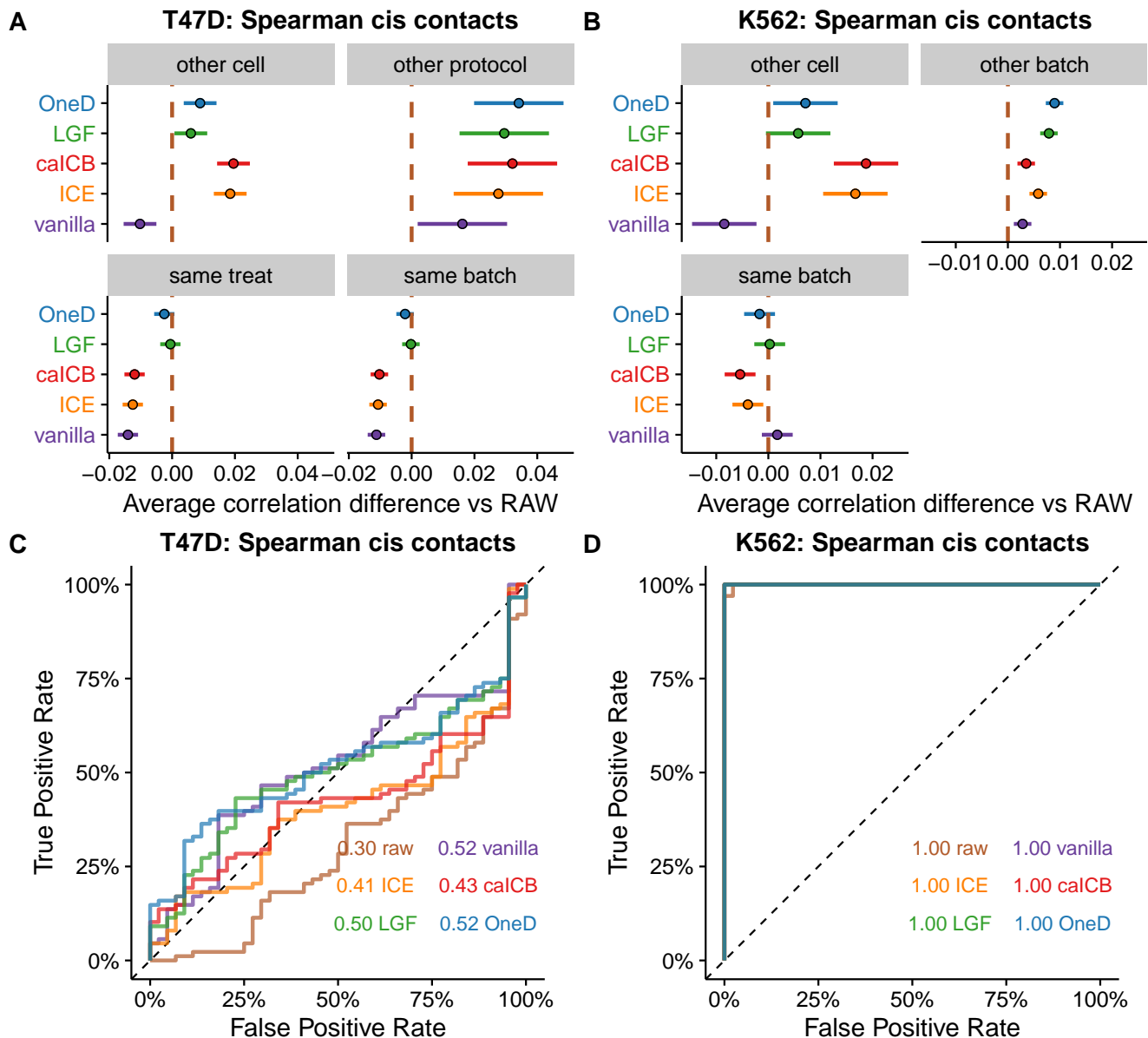


Figure S4: Results of the comparison between samples with aberrant karyotype. A and B. Average changes compared to raw on the T47D and K562 sets. The bars represent 95% confidence intervals centered on the mean difference of the correlation score between a given correction method and the raw data. The brown dashed line indicates the value of the average score on raw matrices (set to 0). C and D. ROC curves on the T47D and K562 sets. The areas under the curve are indicated in the bottom right corner. The color code is the same as panels A and B. The brown lines correspond to raw matrices. All results in this figure are based on Spearman correlations between the observed counts.

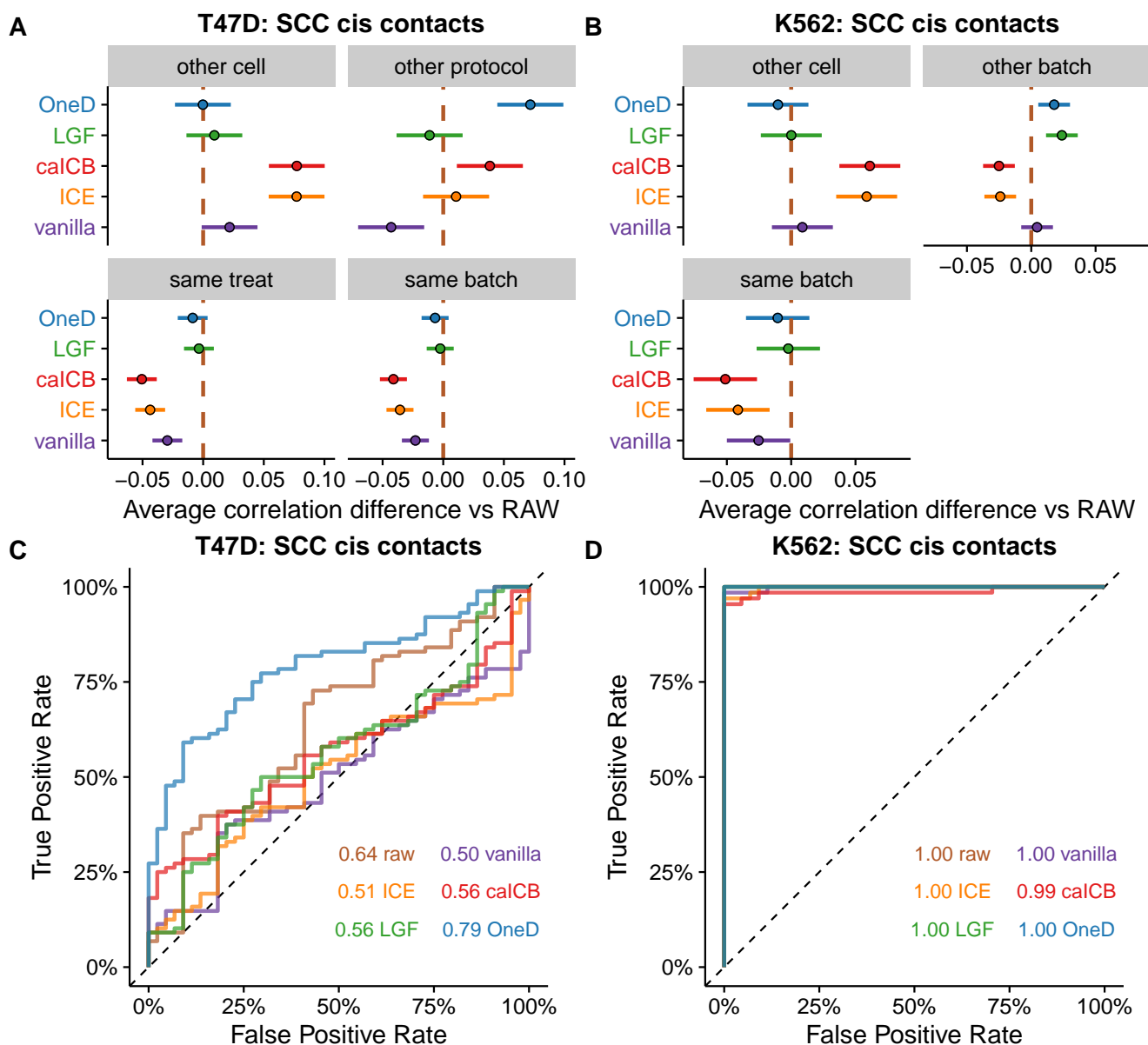


Figure S5: Results of the comparison between samples with aberrant karyotype. A and B. Average changes compared to raw on the T47D and K562 sets. The bars represent 95% confidence intervals centered on the mean difference of the correlation score between a given correction method and the raw data. The brown dashed line indicates the value of the average score on raw matrices (set to 0). C and D. ROC curves on the T47D and K562 sets. The areas under the curve are indicated in the bottom right corner. The color code is the same as panels A and B. The brown lines correspond to raw matrices. All results in this figure are based on stratum-adjusted correlations between the observed counts.

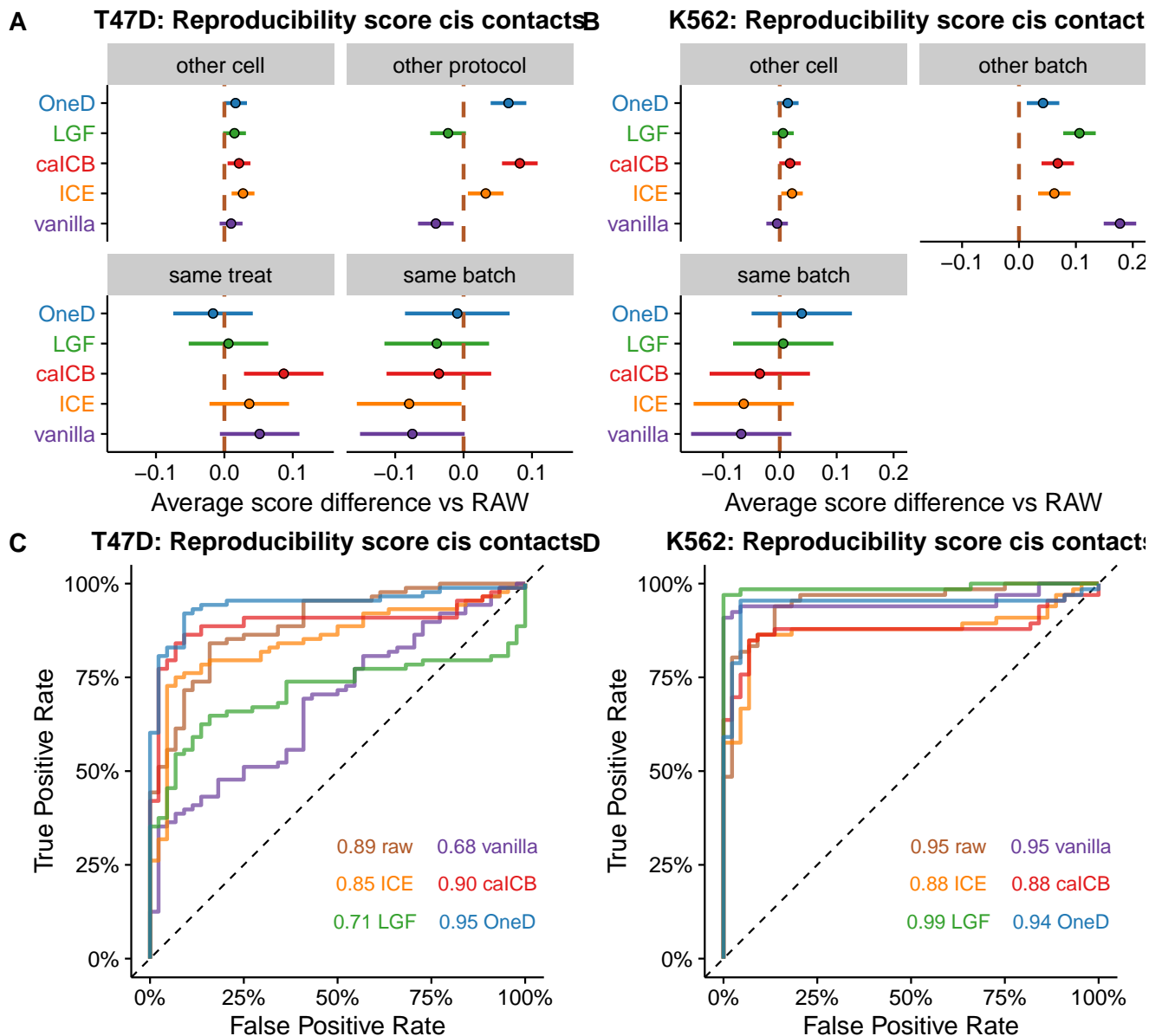


Figure S6: Results of the comparison between samples with aberrant karyotype. A and B. Average changes compared to raw on the T47D and K562 sets. The bars represent 95% confidence intervals centered on the mean difference of the correlation score between a given correction method and the raw data. The brown dashed line indicates the value of the average score on raw matrices (set to 0). C and D. ROC curves on the T47D and K562 sets. The areas under the curve are indicated in the bottom right corner. The color code is the same as panels A and B. The brown lines correspond to raw matrices. All results in this figure are based on the reproducibility score between the observed counts.

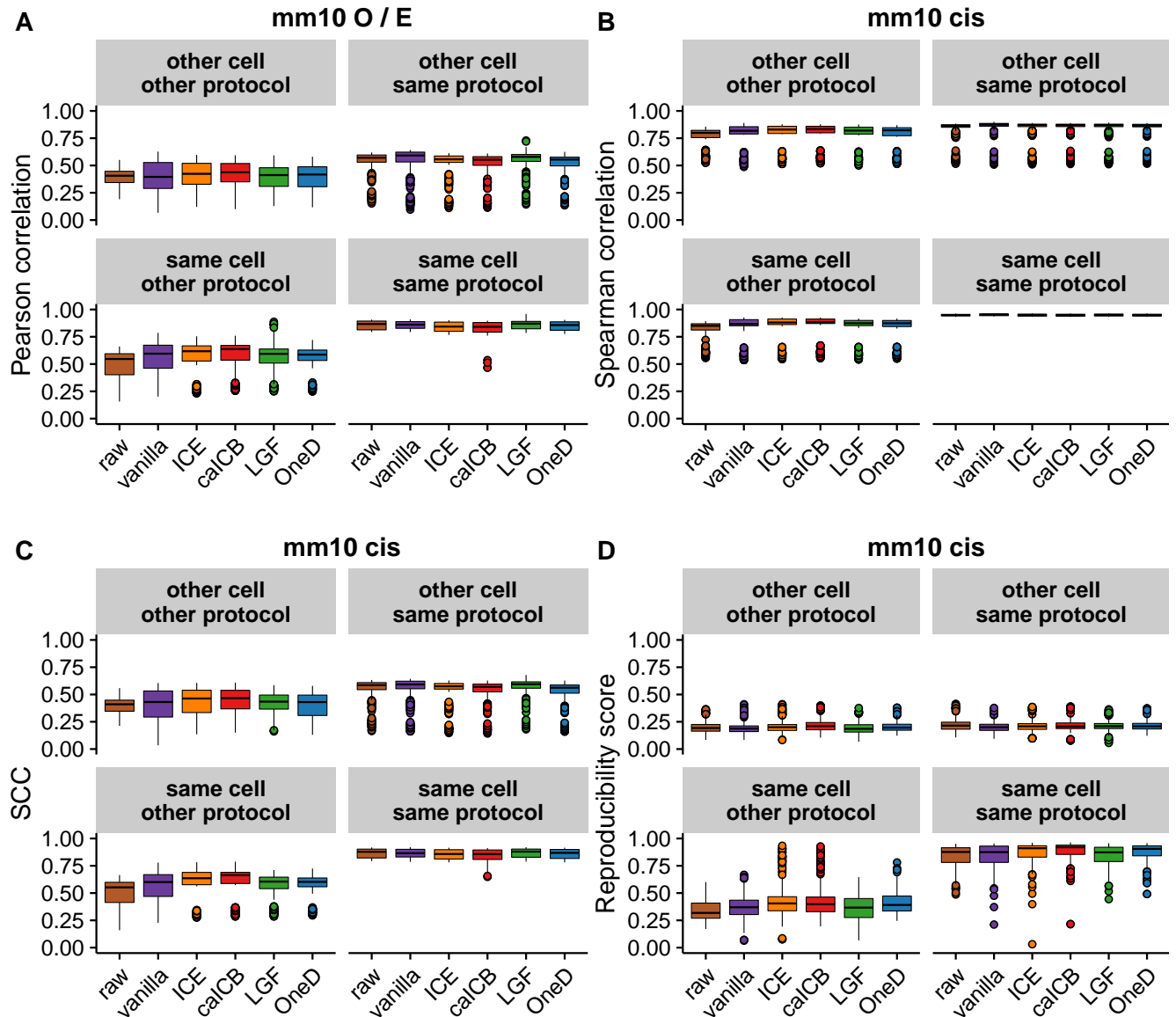


Figure S7: Boxplots representing the results of the pair-wise comparisons of the samples included in the mm10 set. X axis: different correction methods. Y axis: correlation. Each panel groups pairs of samples with the corresponding characteristics (in terms of cell type, protocol, batch and treatment). A. Pearson correlation between observed over expected counts. B. Spearman correlation between observed counts. C. Stratum-adjusted correlation coefficient (SCC) between observed counts. D. Reproducibility score of observed counts.

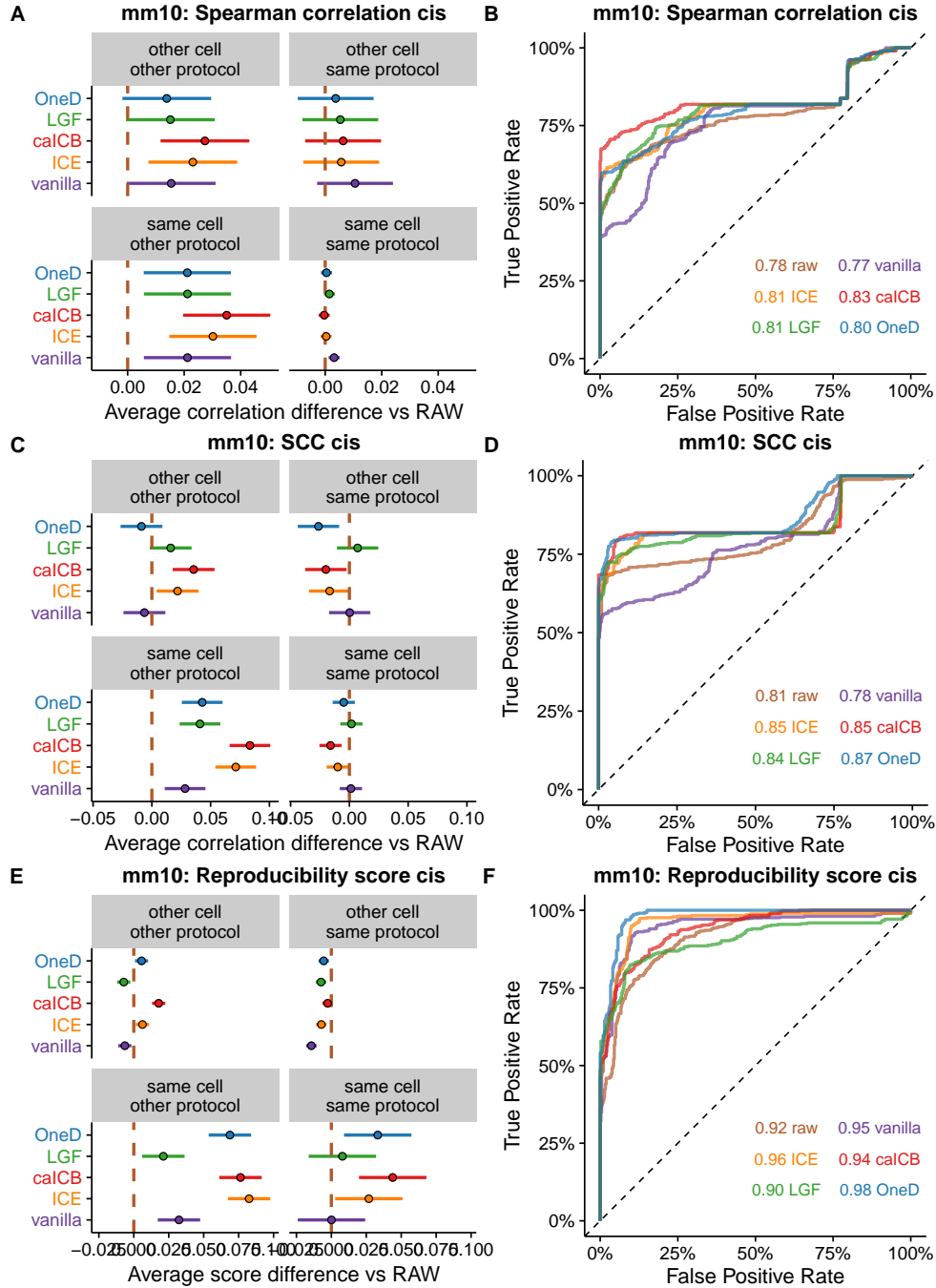


Figure S8: Results of the comparison between samples with normal karyotype. A., C. and E. Average changes compared to raw. The bars represent 95% confidence intervals centered on the mean difference of the correlation score between a given correction method and the raw data. The brown dashed line indicates the value of the average score on raw matrices (set to 0). B., D. and F. ROC curves. The areas under the curve are indicated in the bottom right corner. The color code is the same as panels A, C. B. The brown lines correspond to raw matrices. Results in panels A. and B. on Spearman correlations between the observed counts. Results in panels C. and D. on Stratum-adjusted correlation coefficient (SCC) between observed counts. Results in panels E. and F. on reproducibility score of observed counts.