
Supplementary Figures

Katayama *et al.* Guide for library design and bias correction for large-scale transcriptome studies using highly multiplexed RNAseq methods

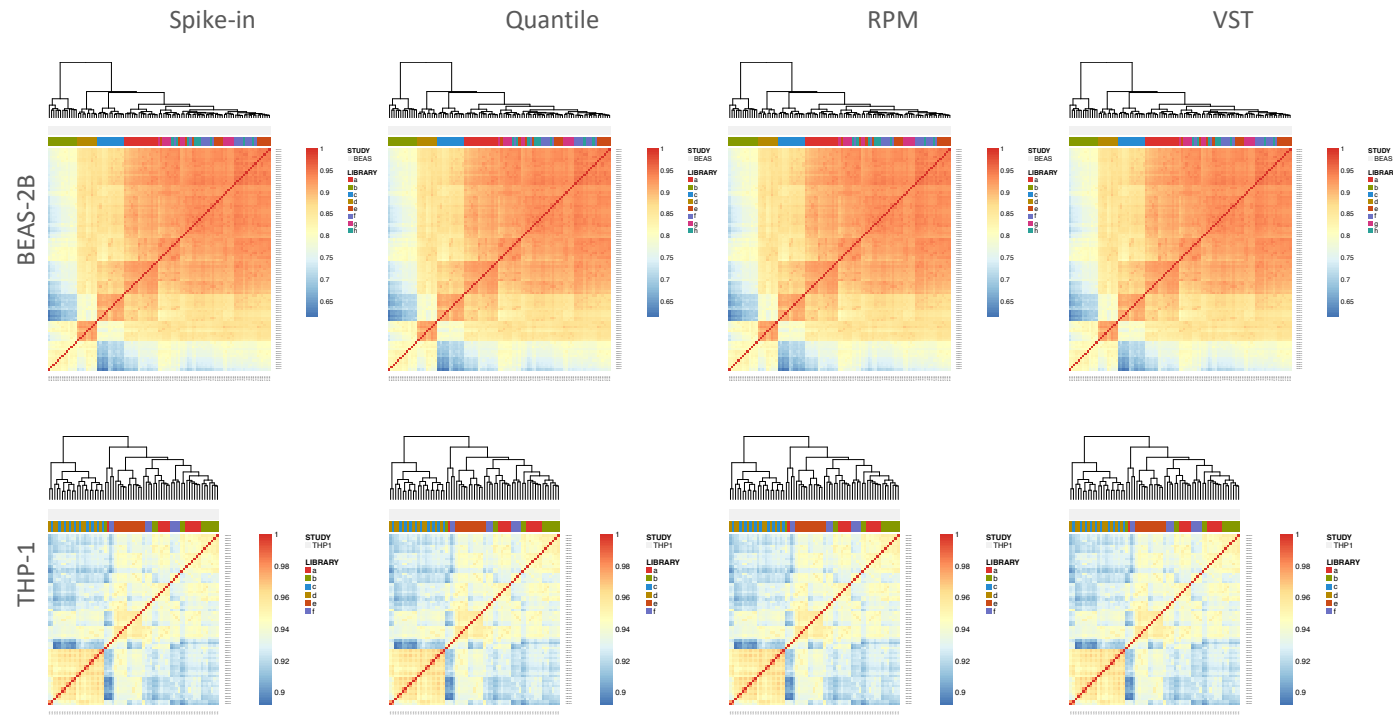


Figure S1. Spearman correlation coefficients between the technical replicates of normalized expression levels by four different normalization methods and the libraries.

The reference samples should show weak clustering, without library dependency, as they were technical replicates of the same RNA sample. However, there were some significant clusters, especially BEASb. All normalization methods transform the average and/or the variation of expression levels in each sample but do not change the ranks in each sample. Therefore, the library biases, which were represented by distances of Spearman rank correlation coefficients here, were identical between the methods.

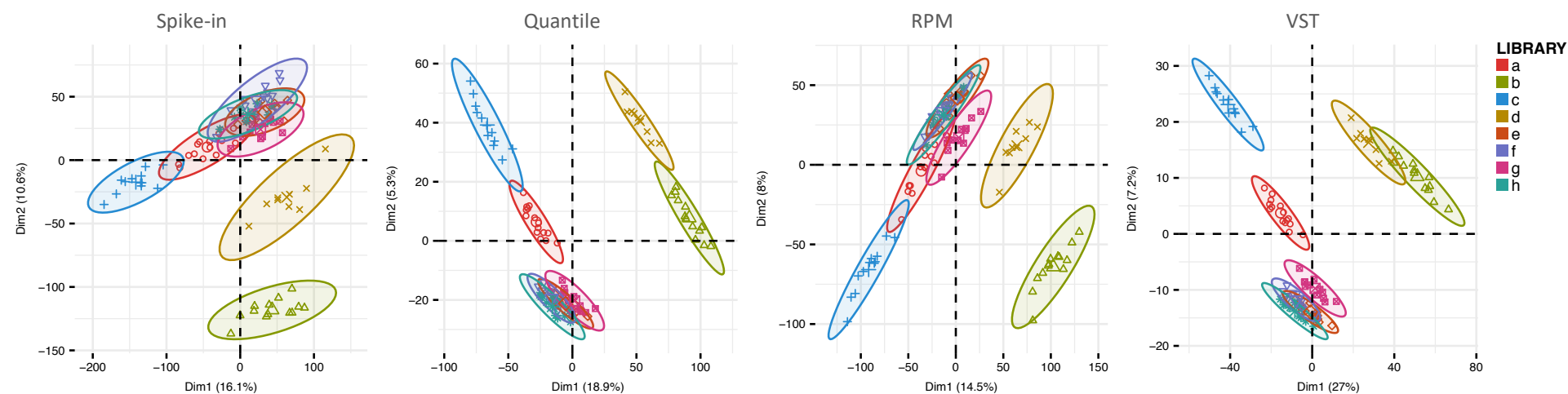


Figure S2. PCA of the BEAS-2B technical replicates of normalized expression levels by four different normalization methods.

Considering that these samples are technical replicates, they should be clustered without library dependency. Instead, they showed clustering by libraries.

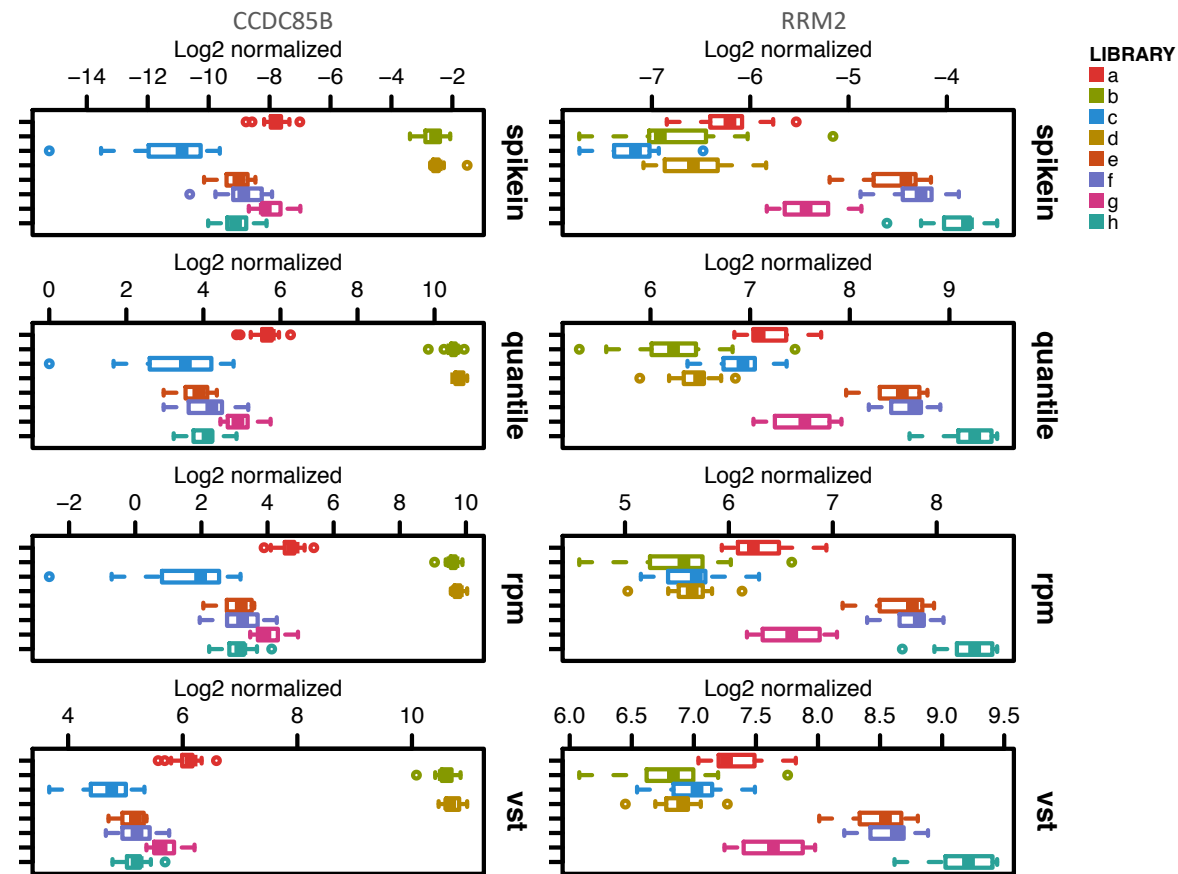


Figure S3. Normalized expression levels by the four different normalization methods of two library-biased genes.

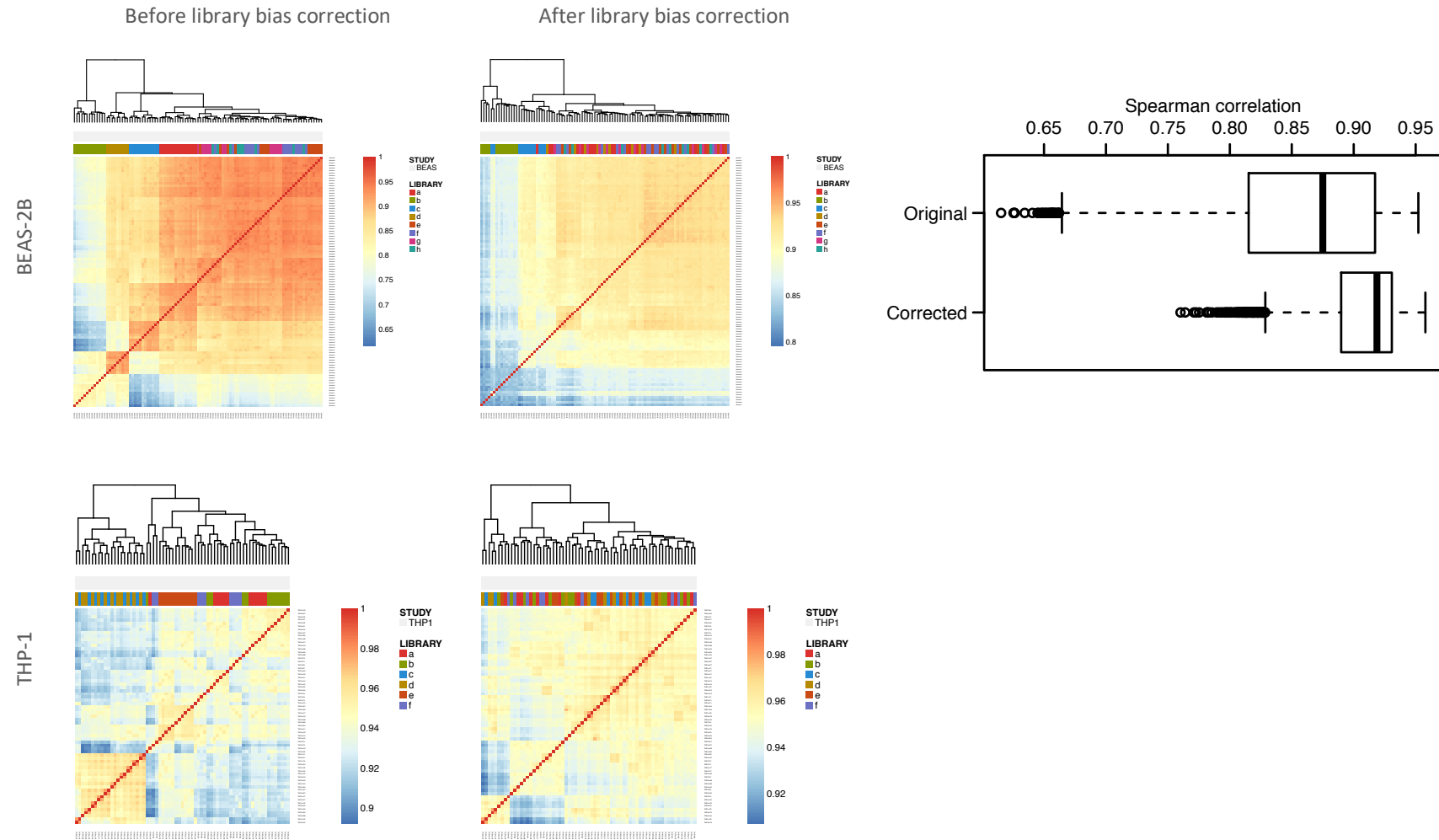


Figure S5. Spearman correlation coefficients between the technical replicates of the normalized expression levels before and after the library bias correction.

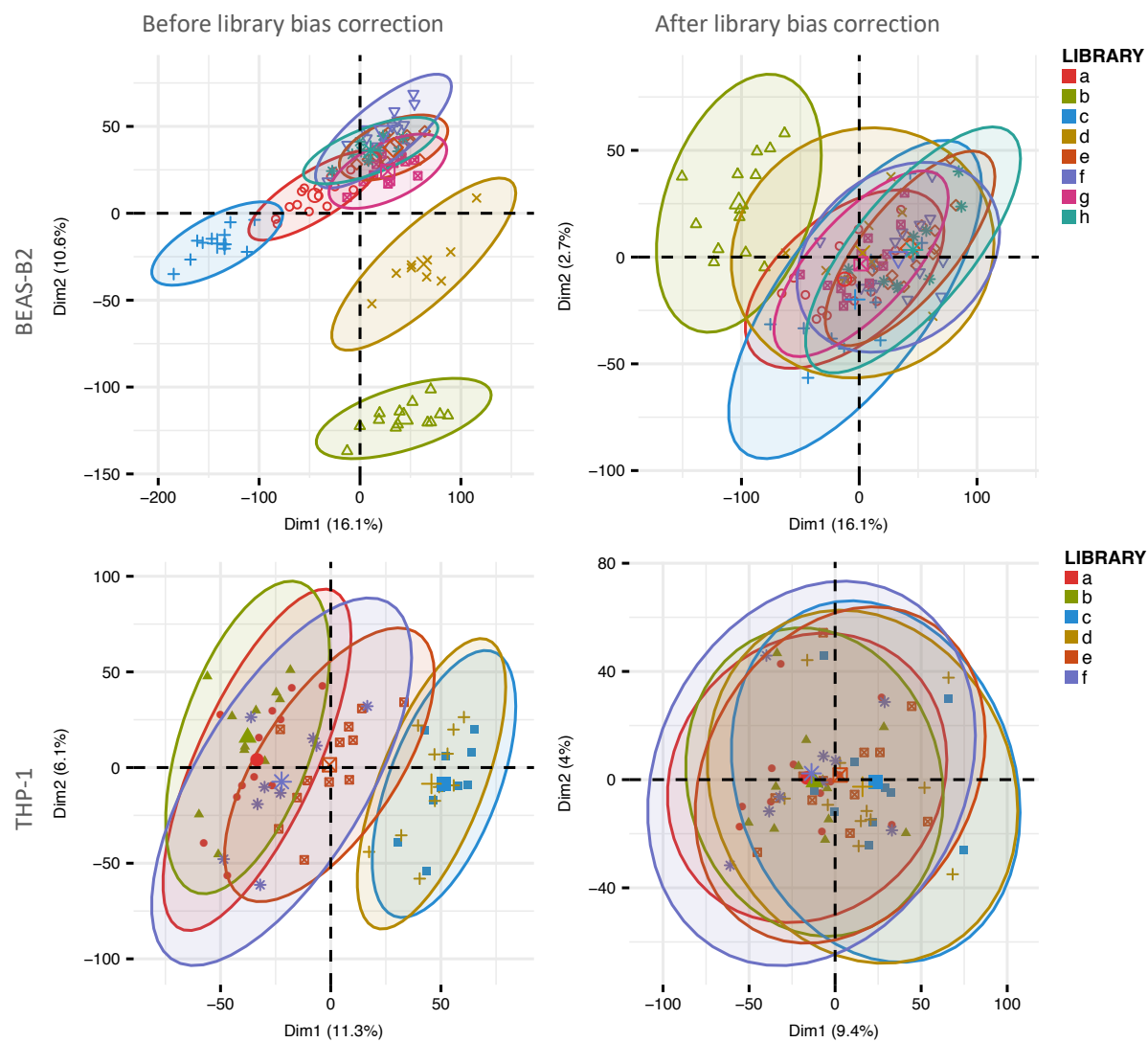


Figure S6. PCA of the technical replicates of the spike-in normalized expression levels before and after the library bias correction.

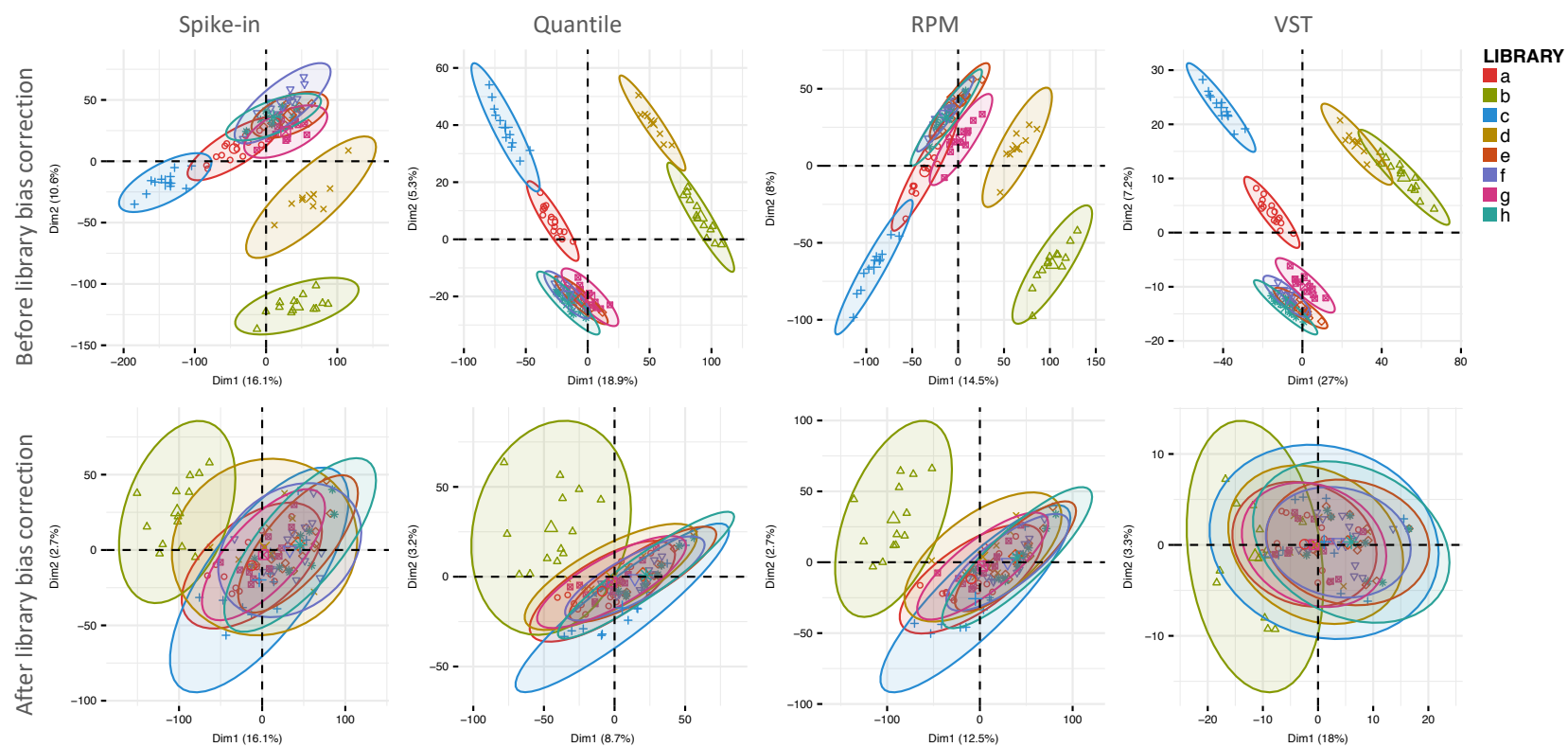


Figure S7. PCA of normalized expression levels of the BEAS-2B technical replicates by four different normalization methods, before and after the library bias correction.

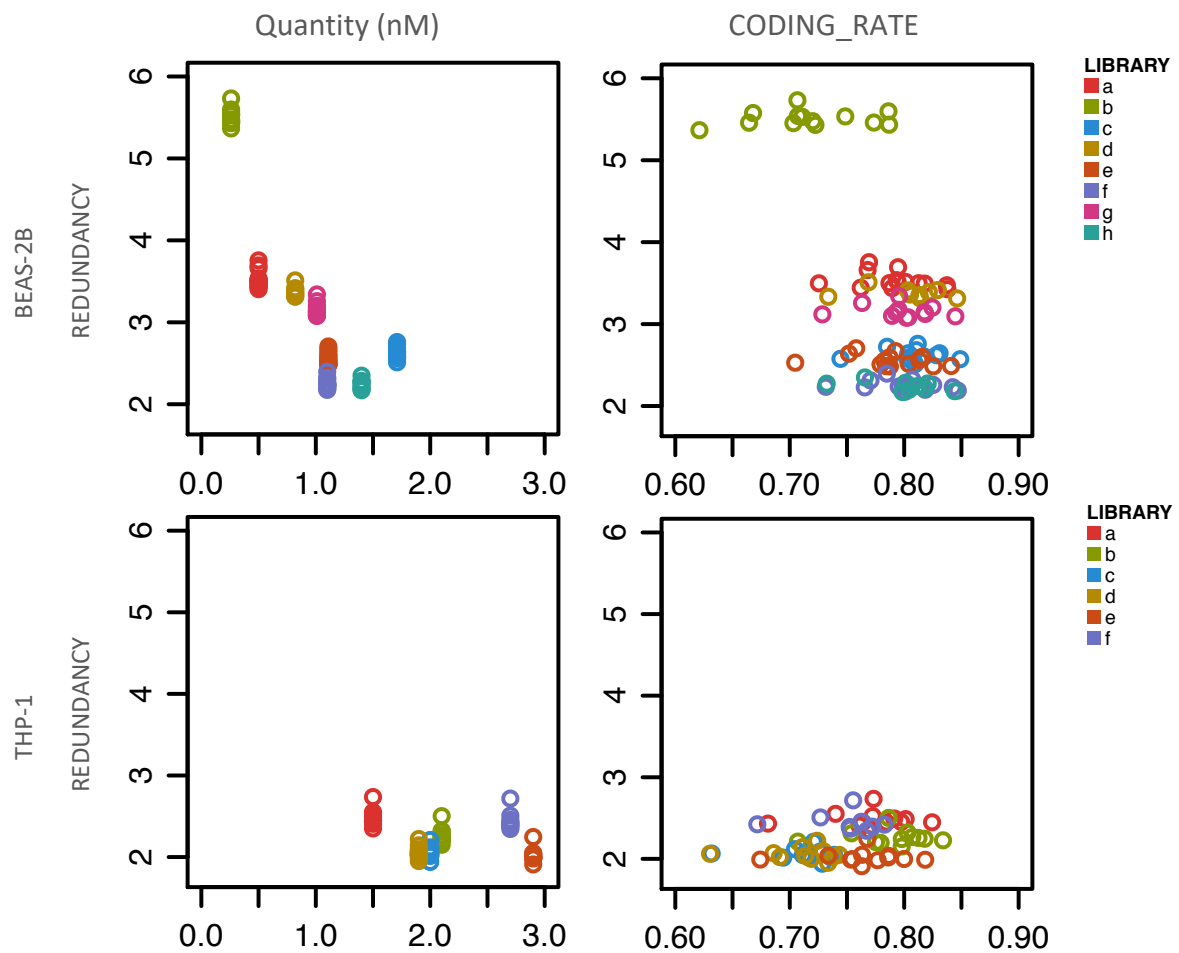


Figure S8. Quantity of the library before sequencing, and REDUNDANCY and CODING_RATE after the sequencing.

Drop of the quantity in the libraries using similar samples in a study suggests potential technical issues. (NOTE: REDUNDANCY and CODING_RATE are different by types of cells or tissues as well as the amount of input RNAs)