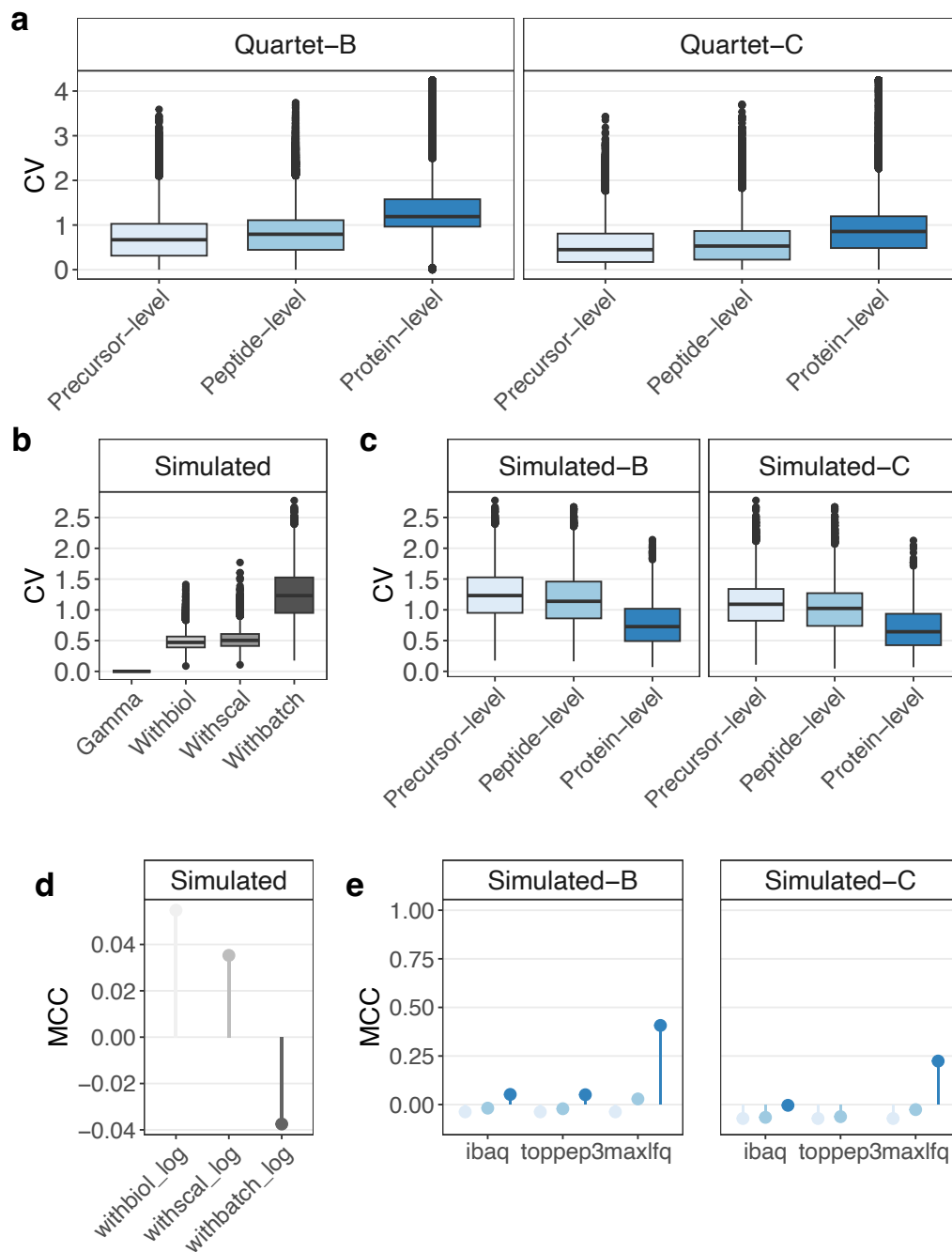
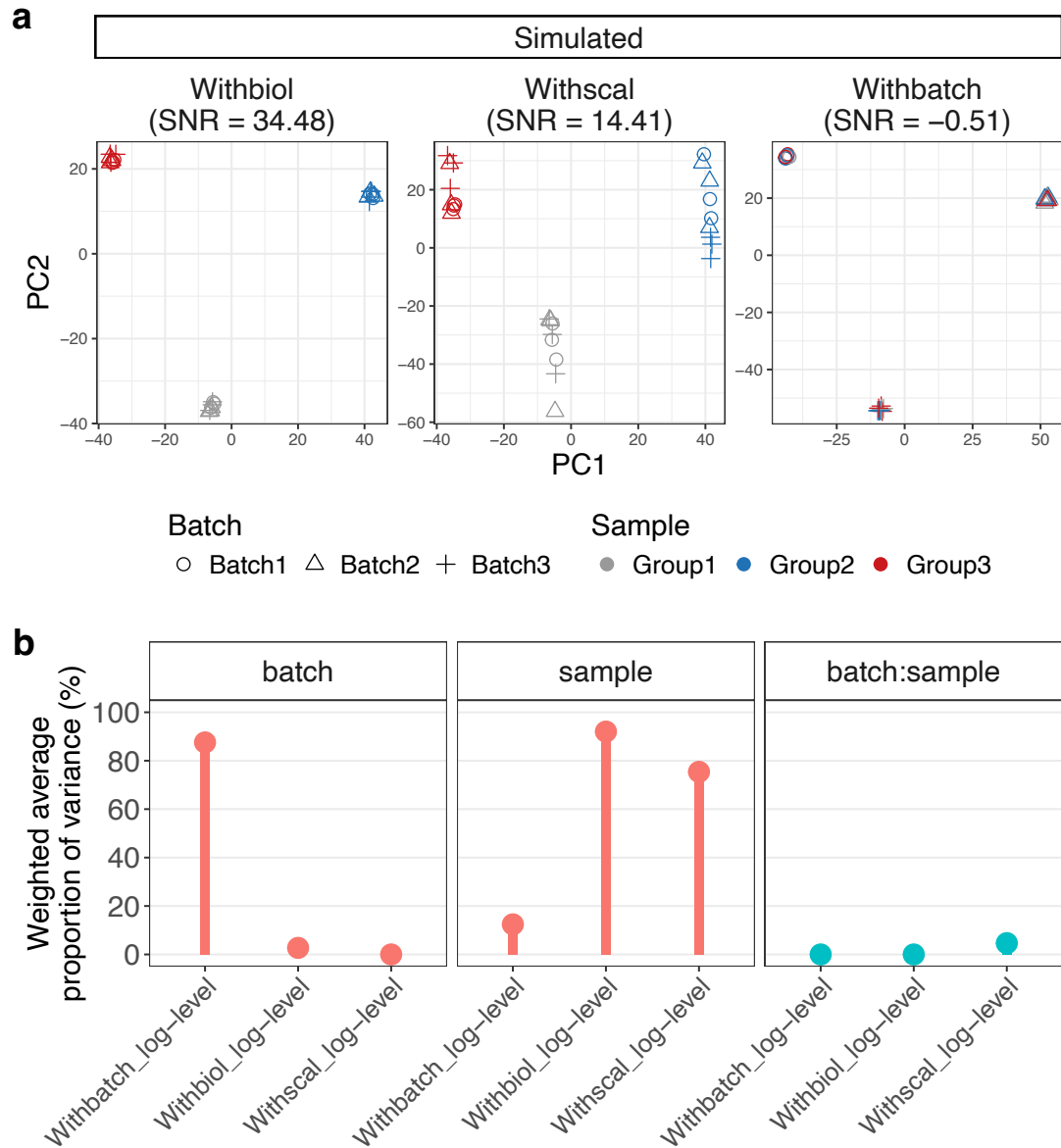


Supplementary Figures



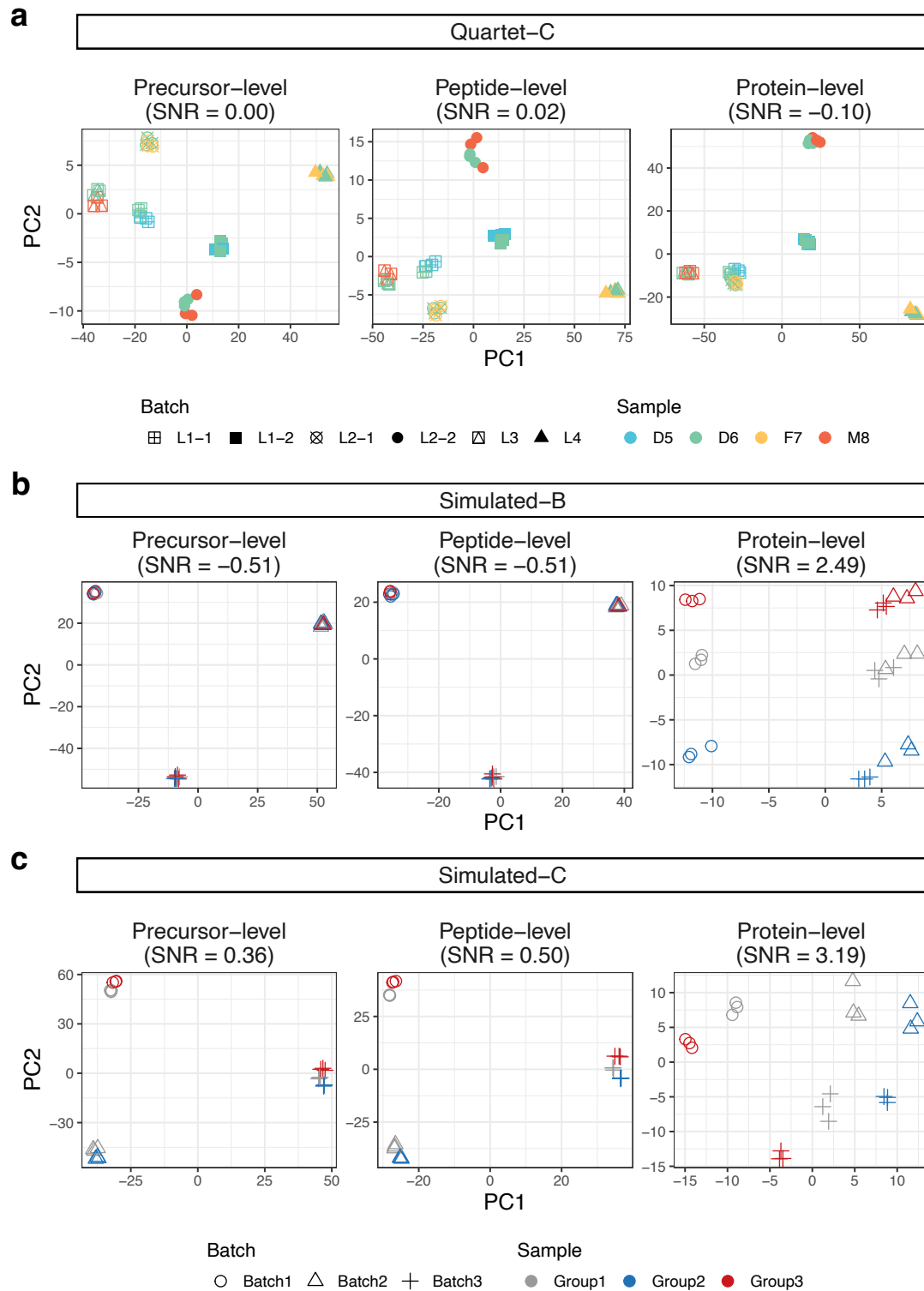
Supplementary figure 1. Feature-based quality assessment before batch-effect correction.

Box plots of the coefficient of variation (CV) (a-c) and Matthews correlation coefficient (MCC) (d-e) values across all three data levels in different designed scenarios. **a**, CV performance of Quartet datasets; colored by the data level; faceted by the scenario. **b**, CV performance of simulated datasets at precursor level; colored by the distribution or the specific factor that introduced to the data matrix. **c**, CV performance of simulated datasets at all three data levels; colored by the data level; faceted by the scenario. **d-e**, bar plots of the MCC values in simulated data; colored by the distribution or the specific factor that introduced to the data matrix (d) or the correction level (e); faceted by the scenario. For the box plots, horizontal lines indicate the median; box boundaries indicate the interquartile range (IQR); whiskers represent values within $1.5 \times$ IQR of the first and third quartiles; data points beyond the end of the whiskers are plotted individually. Detailed n statistics are provided (Supplementary Table 1). Source data are provided as a Source Data file.



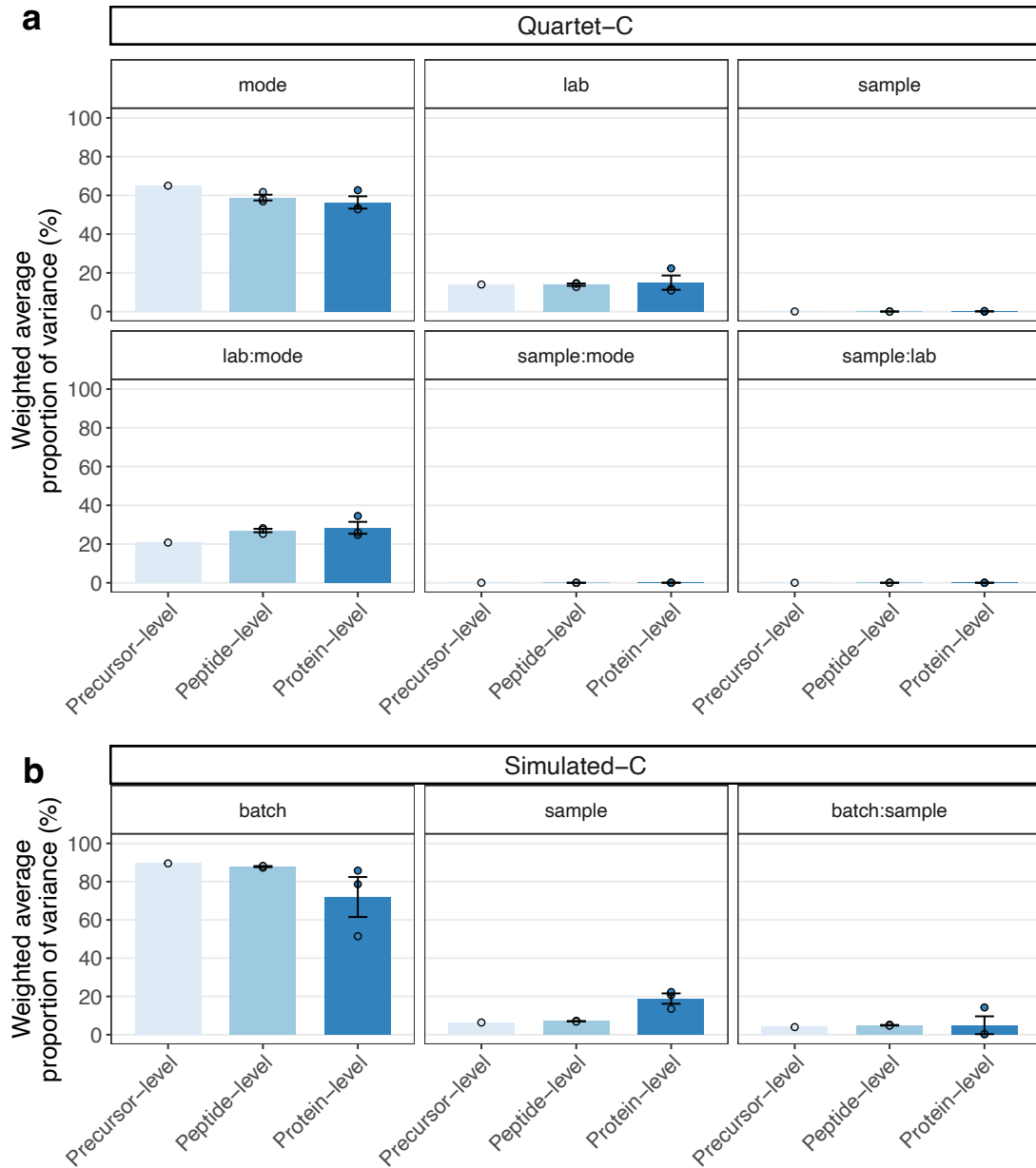
Supplementary figure 2. The stepwise introduction of factors in data simulation.

a, scatters plot of principal component analysis (PCA) in simulated datasets at raw precursor level; colored by the sample type; shaped by the batch; faceted by the distribution or the specific factor that introduced to the data matrix. **b**, bar plots of principal variance component analysis (PVCA) performance in simulated data matrix at precursor level; colored and faceted by the random effect. All data points are plotted individually. Source data are provided as a Source Data file.



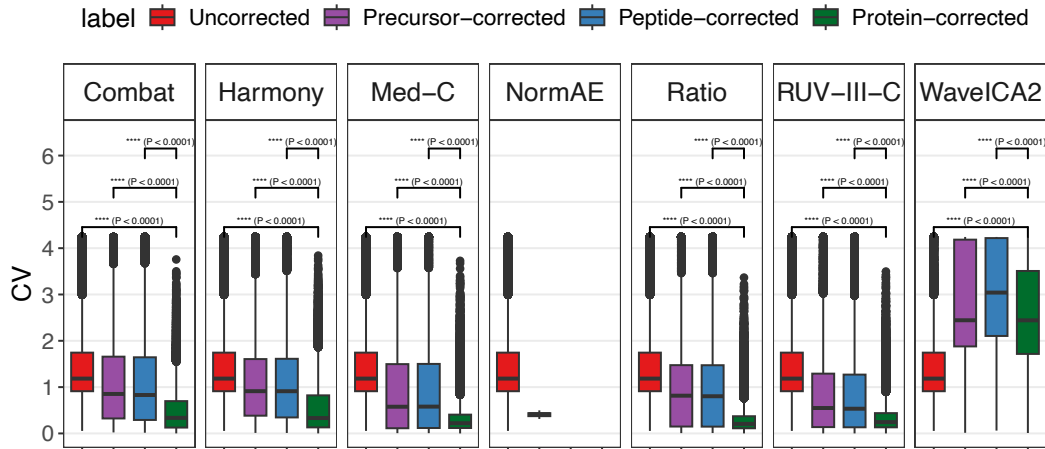
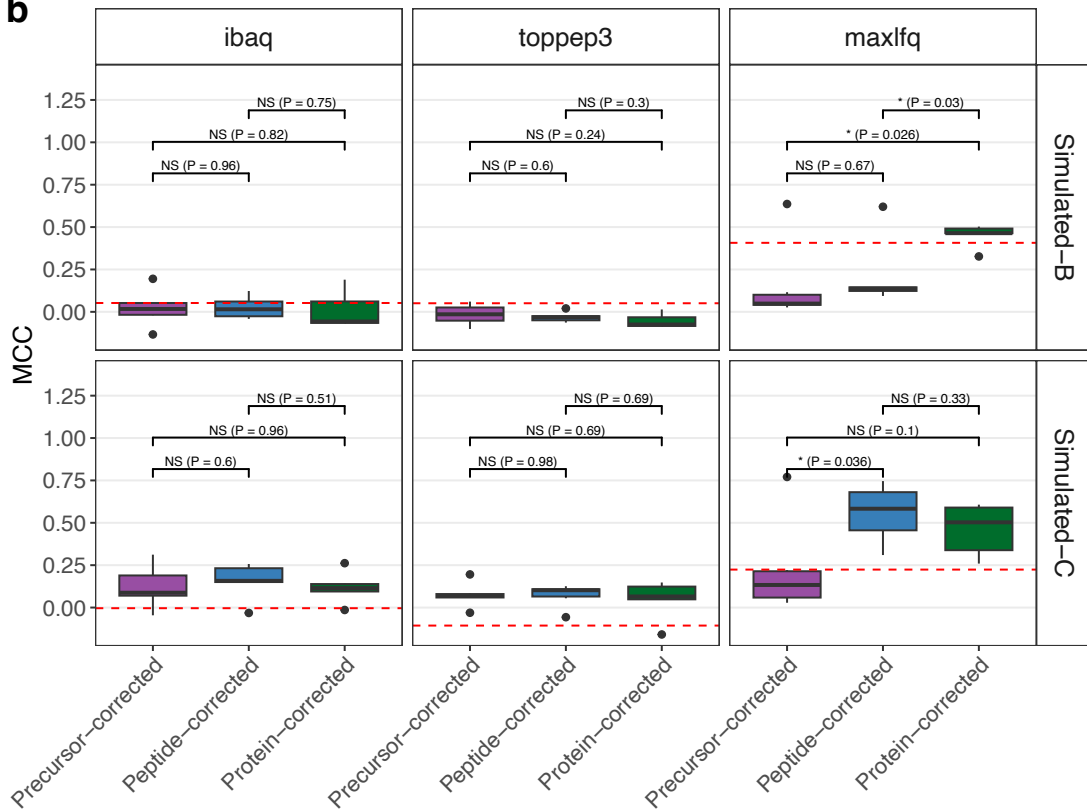
Supplementary figure 3. Signal-to-noise ratio (SNR) quality assessment before batch-effect correction.

Scatters plot of PCA in Quartet-confounded (Quartet-C) (a), Simulated-balanced (Simulated-B) (b) and Simulated-confounded (Simulated-C) (c) scenarios; shaped by the raw batch; colored by the sample type. All data points are plotted individually. Source data are provided as a Source Data file.



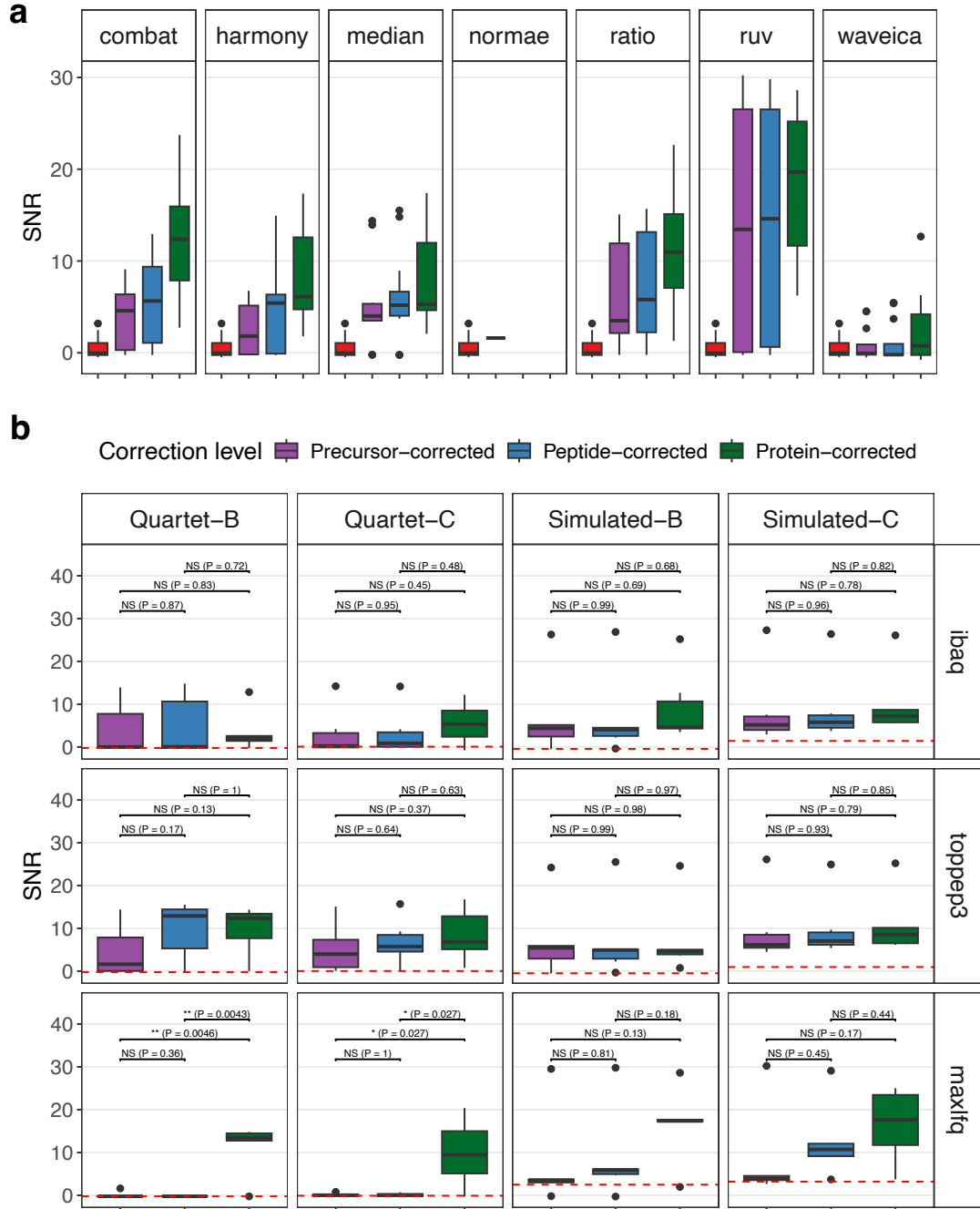
Supplementary figure 4. PVCA quality assessment before batch-effect correction.

Bar plots of PVCA performance in uncorrected data matrices across all three data levels under Quartet-C (a) and Simulated-C (b) scenarios; presented as mean values $\pm$ s.d. colored by the data level; faceted by the scenario. All data points are plotted individually. For each bar plot: Precursor-level (n=1), Peptide-level (n=3), Protein-level (n=3). All statistics are based on analysis replicates (by quantification methods (QMs)). Source data are provided as a Source Data file.

a**b**

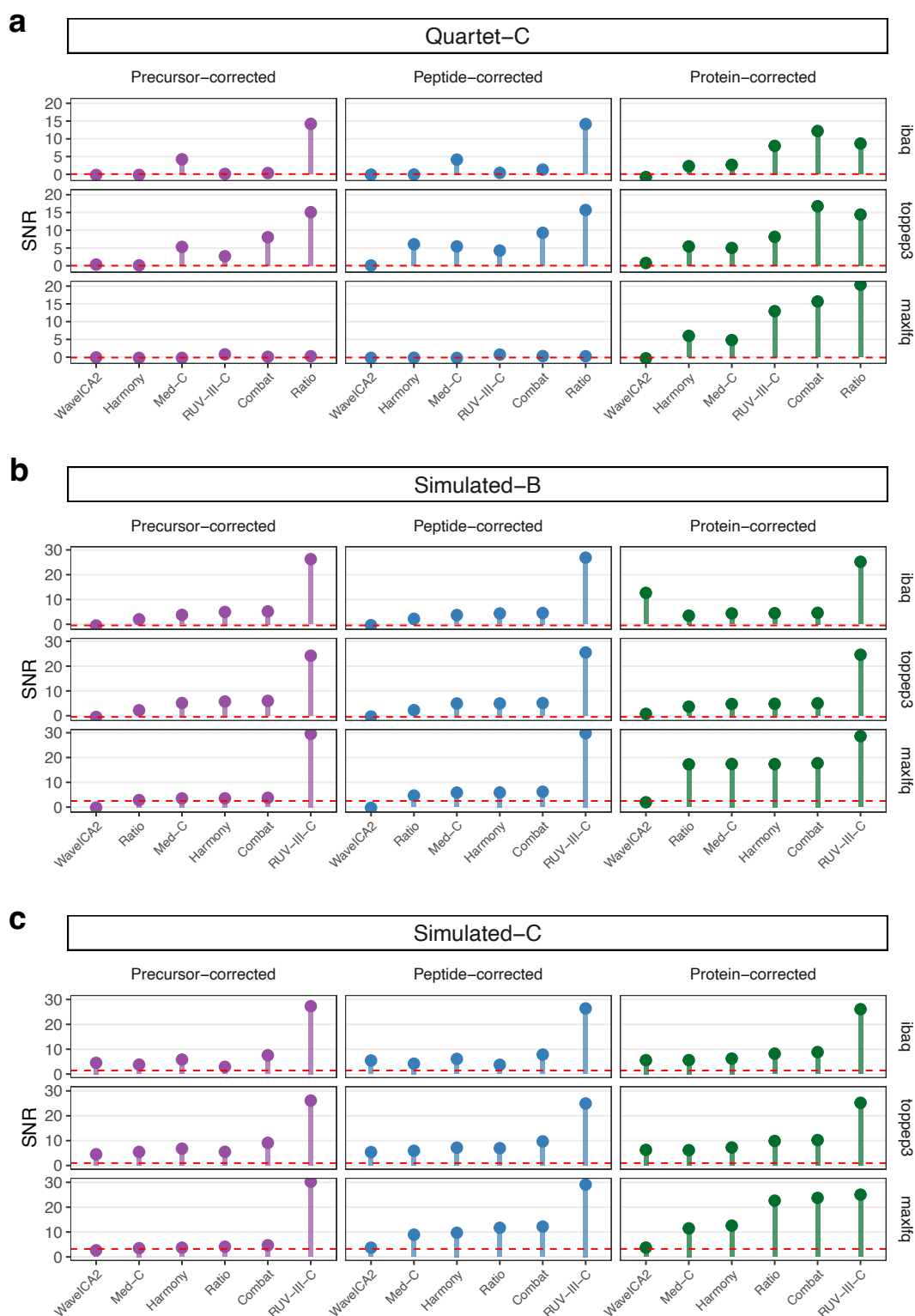
Supplementary figure 5. Comparisons of three correction levels in feature-based performance after batch-effect correction.

a, Box plots of the CV values after MaxLFQ quantification across different BECAs; colored by the correction level; faceted by the BECA; detailed n statistics are provided (Supplementary Table 2). **b**, Box plots of the MCC values of across different BECAs; colored by the correction level; faceted by the quantification method; the dashed lines represent the MCC value in the uncorrected data matrix; for each box n=6 BECAs. All statistics are based on independent biological replicates (by sample groups) and analysis replicates (by BECAs). All P values are calculated using unpaired two-tailed student's T tests, ****P < 0.0001, **P < 0.01, *P < 0.05; exact P values (when less than 0.0001) are provided above each comparison. For the box plots, horizontal lines indicate the median; box boundaries indicate the interquartile range (IQR); whiskers represent values within 1.5× IQR of the first and third quartiles; data points beyond the end of the whiskers are plotted individually. Source data are provided as a Source Data file.



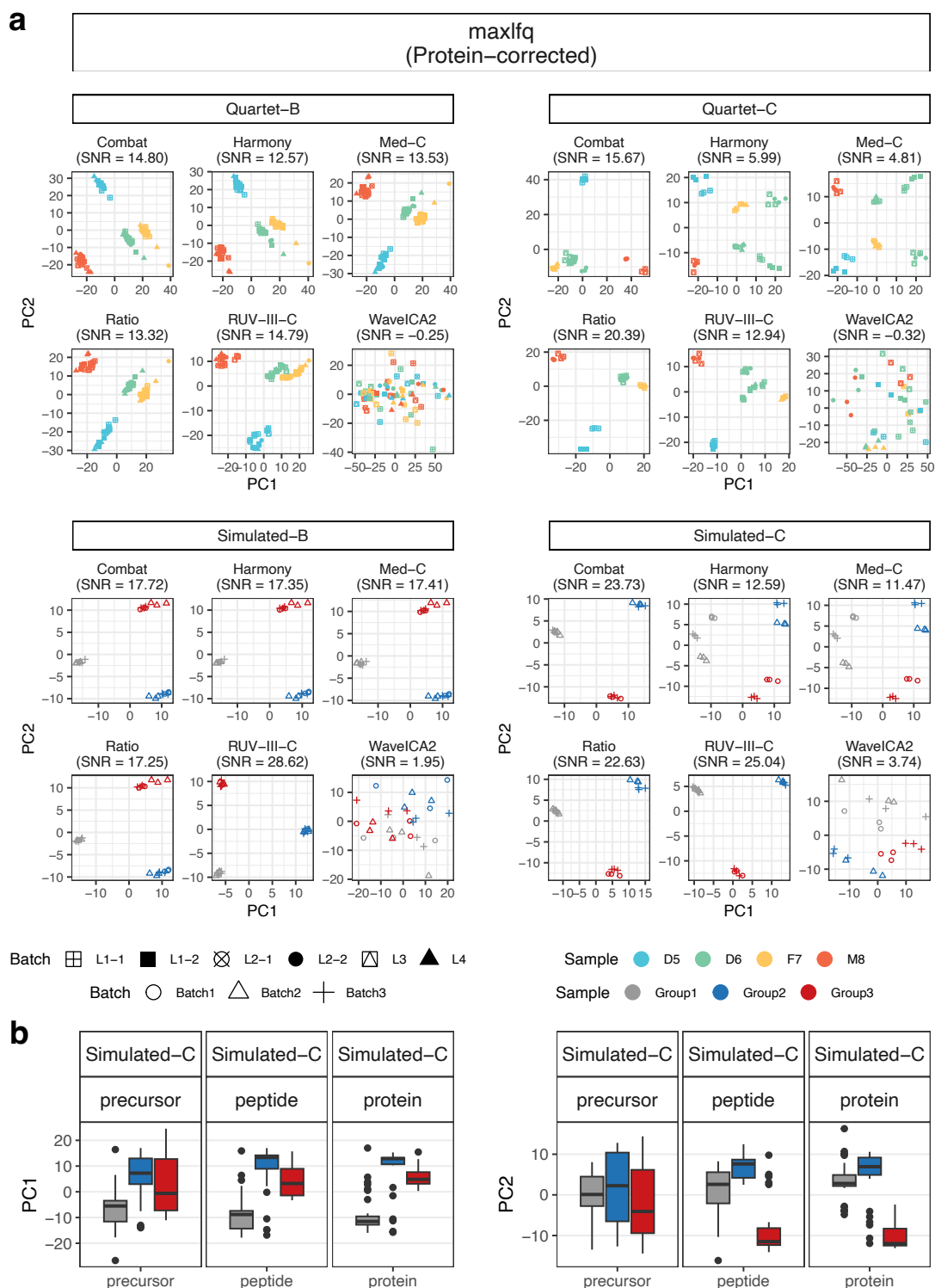
Supplementary figure 6. Comparisons of three correction levels in signal-to-noise ratio (SNR) performance after batch-effect correction.

Box plots of the SNR values of across different BECAs (a) and different scenario-quantification combinations (b); the P values are calculated using unpaired two-tailed student's T tests, ****P < 0.0001, **P < 0.01, *P < 0.05; exact P values (when less than 0.0001) are provided above each comparison; colored by the correction level; faceted by the BECA (a) or the QM (b). The dashed lines represent the SNR value in the uncorrected data matrix. For the box plots, horizontal lines indicate the median; box boundaries indicate the interquartile range (IQR); whiskers represent values within 1.5× IQR of the first and third quartiles; data points beyond the end of the whiskers are plotted individually. For each box (a): QM n=3, Scenario n=4, total n = 12 (except NormAE, QM n=3, Scenario n=1, total n = 3); for each box (b): n=6 BECAs (except Quartet-B Precursor-corrected, n=7 BECAs). Source data are provided as a Source Data file.



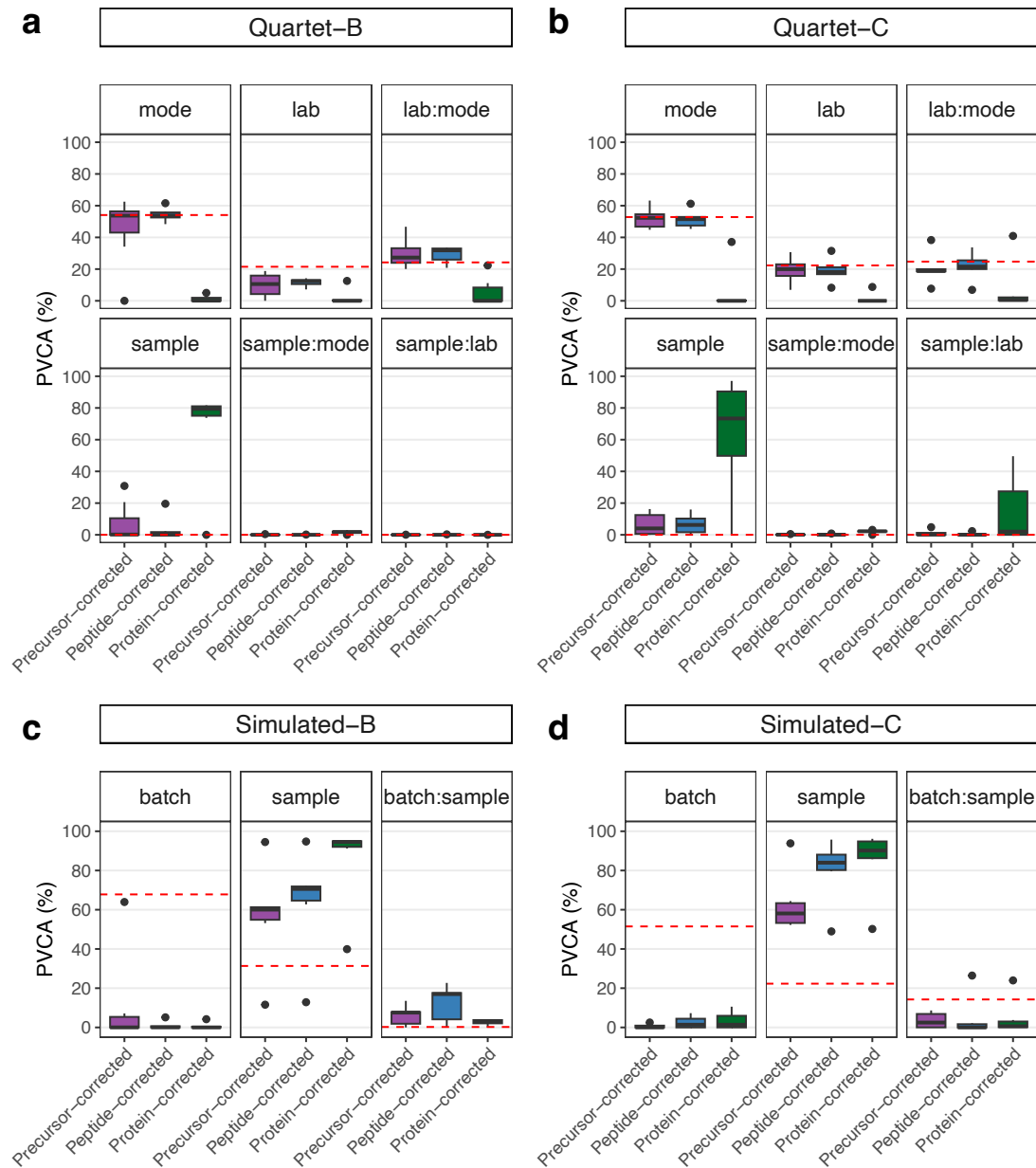
Supplementary figure 7. Comparisons of different BECA-QM combinations in SNR performance after batch-effect correction.

Bar plots of the SNR values in Quartet-C (a), Simulated-B (b) and Simulated-C (c) scenarios; colored by the correction level; faceted by the correction level (by column) and the quantification method (by row); the dashed lines represent the SNR value in the uncorrected data matrix. All data points are plotted individually. Source data are provided as a Source Data file.



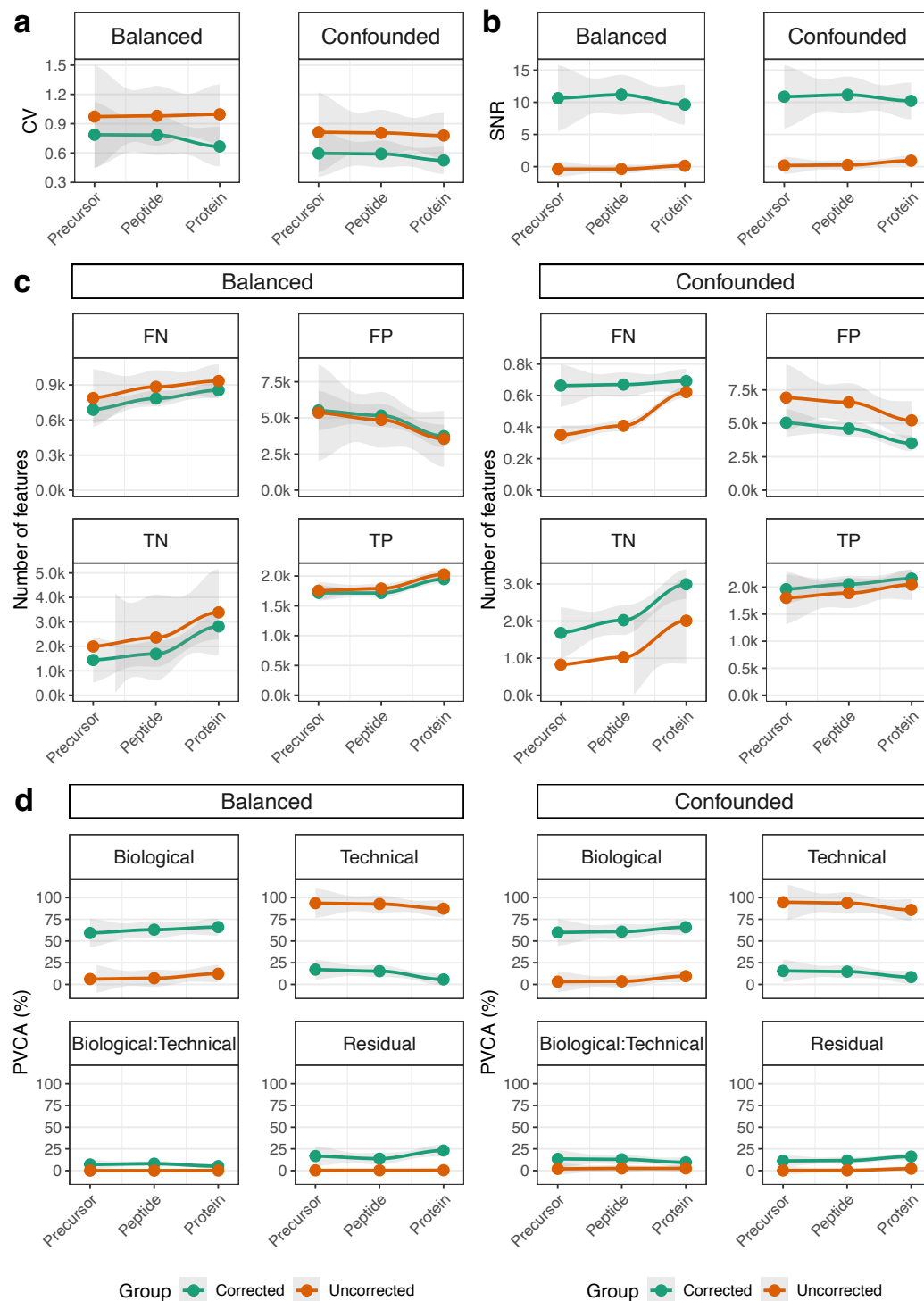
Supplementary figure 8. PCA plots of different BECA-MaxLFQ combinations after batch-effect correction.

a, scatter plots of PCA results of the corrected data matrices after MaxLFQ protein quantification; shaped by the batch; colored by the sample type. **b**, side-by-side box plots for PC-based association in Simulate-C scenario; colored by the sample type; horizontal lines indicate the median; box boundaries indicate the interquartile range (IQR); whiskers represent values within $1.5 \times$ IQR of the first and third quartiles; data points beyond the end of the whiskers are plotted individually; for each box: Group1 (n=9), Group2 (n=6), Group3 (n=6). Source data are provided as a Source Data file.



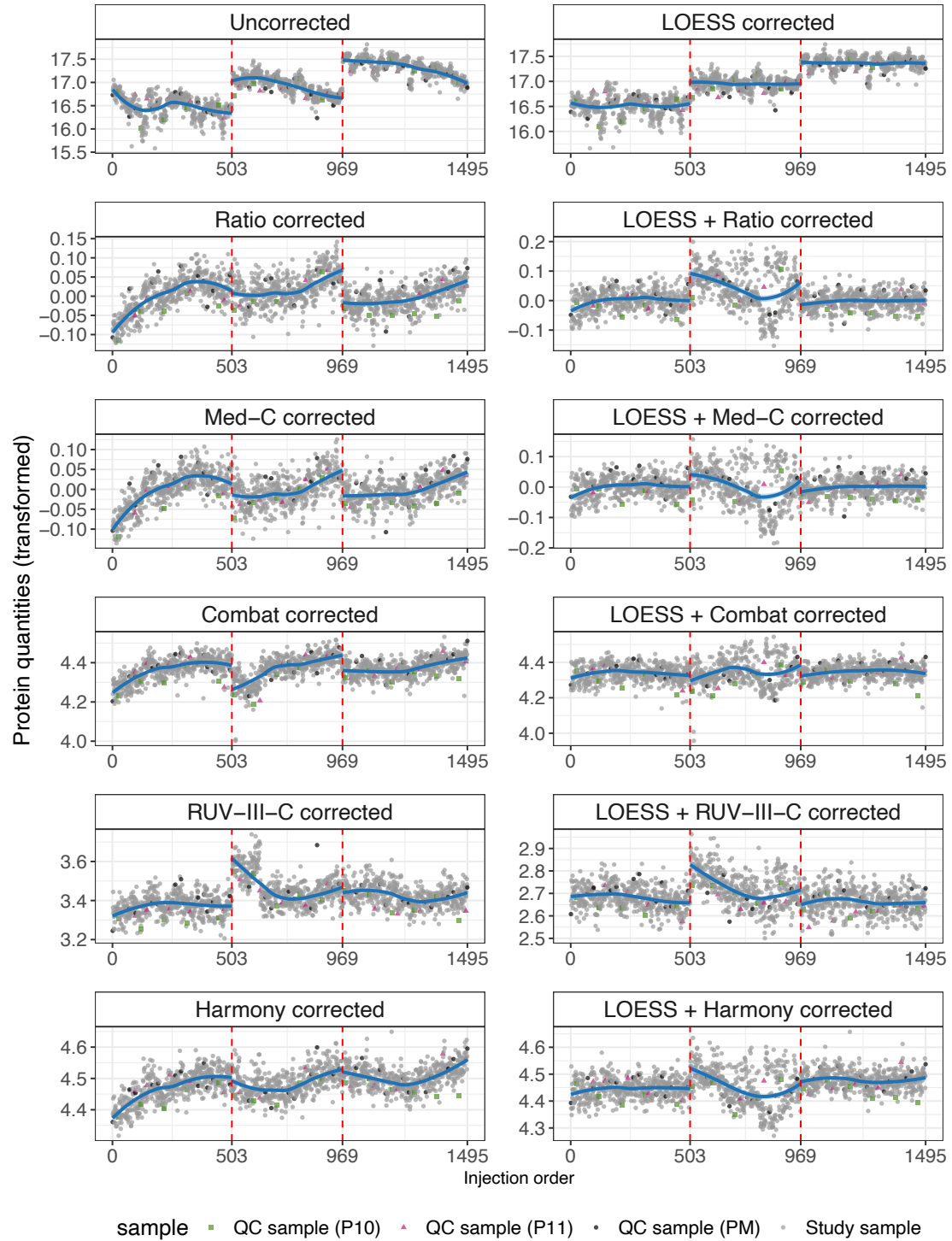
Supplementary figure 9. Comparisons of three correction levels in PVCA performance after batch-effect correction.

Box plots of the PVCA results in four scenarios; colored by the correction level; faceted by the label of random effects. The dashed red lines represent the proportion value in the uncorrected data matrix. For the box plots, horizontal lines indicate the median; box boundaries indicate the interquartile range (IQR); whiskers represent values within $1.5 \times$ IQR of the first and third quartiles; data points beyond the end of the whiskers are plotted individually. For each box: $n=6$ BECAs (except Quartet-B Precursor-corrected, $n=7$ BECAs). Source data are provided as a Source Data file.



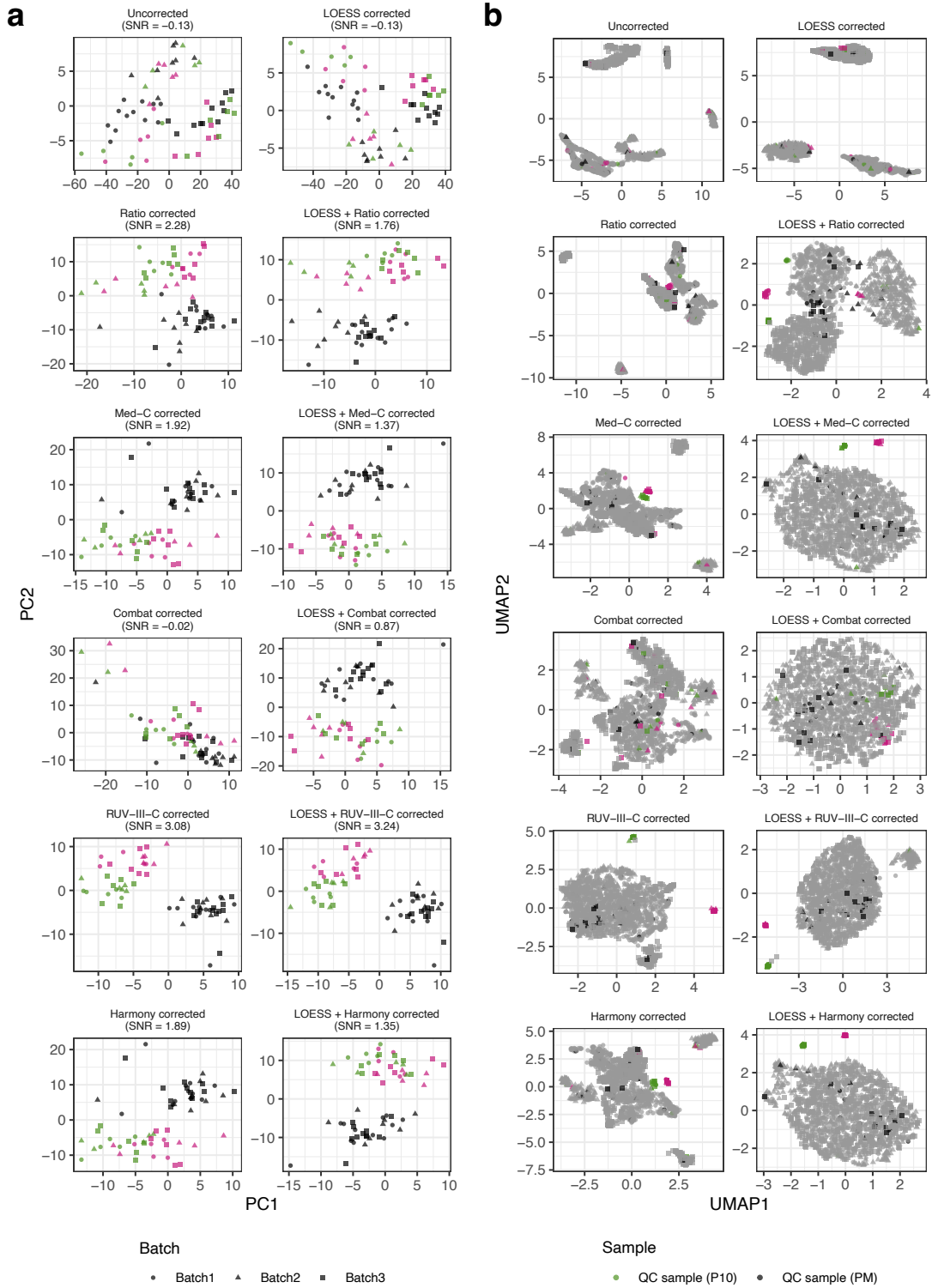
Supplementary figure 10. Comparisons of batch-effect trend at three data levels.

a-b, plots of the CV (a) and SNR (b) performances of data matrices across different data levels. **c**, plots of the number of FNs, FPs, TNs, and TPs of data matrices across different data levels. **d**, plots of the PVCA results of data matrices across different data levels. All plots are colored by before (orange) or after (green) the correction step. All plots are fitted with loess curves, with grey ribbon shaded areas indicating the standard error (se) in 95% confidence interval; the data points in the center area represents for mean values individually. All statistics are based on independent biological replicates (by sample groups) and analysis replicates (by BECAs and by QMs). Source data are provided as a Source Data file.



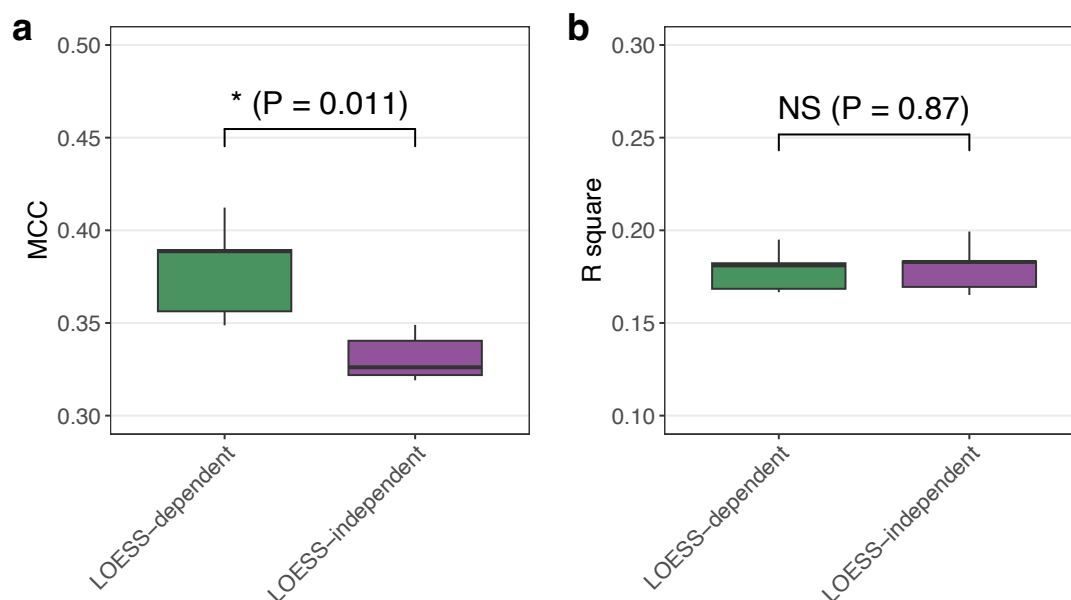
Supplementary figure 11. LOESS fitting improved the removal of MS signal drift.

Scatter plots of mean protein quantities in all samples as the injection order delays (divided by two time points of cleaning the instrument); colored by the sample type; faceted by the preprocessing of the data matrix. All plots are fitted with loess curves, with grey ribbon shaded areas indicating the standard error (se) in 95% confidence interval; the data points in the center area represents for mean values individually. Due to unusually low protein quantities, sample ExpB19 was identified as an outlier and was excluded from the visualization; but all samples were included in the batch-effect correction. Source data are provided as a Source Data file.



Supplementary figure 12. LOESS fitting improved the removal of batch effects.

a, scatter plots of PCA in the ChiHOPE datasets within 64 QC samples. Due to unusually distant from other samples in the PCA plots, an outlier sample ExpJ1 was excluded from the visualization, but the SNR value was calculated from all 64 QC samples. **b**, scatter plots of UMAP (using PCs that account for over 80% of the accumulated variance) within all 1495 injected samples (b). All plots are shaped by the batch; colored by the sample type. Source data are provided as a Source Data file.



Supplementary figure 13. LOESS fitting improved the Sex prediction performance.

a-b, box plots of MCC (**a**) and R-square (**b**) performances in the ChiHOPE datasets. All plots are filled by the group (LOESS-dependent/independent); the P values labeled above the boxes were calculated using unpaired two-tailed student's T tests, ****P < 0.0001, **P < 0.01, *P < 0.05; exact P values (when less than 0.0001) are provided above each comparison. For the box plots, horizontal lines indicate the median; box boundaries indicate the interquartile range (IQR); whiskers represent values within 1.5× IQR of the first and third quartiles; data points beyond the end of the whiskers are plotted individually. For each box: n=5 BECAs. Source data are provided as a Source Data file.

Supplementary Tables

Supplementary table 1. Detailed n statistics for uncorrected data matrices.

Scenario	Data level	Feature	Sample		Algorithm/Method		Total
			Biological group	Biological replicate	BECA	QM	
Quartet-B	Precursor-level	87,908	4	18 for D5/D6/F7/M8	-	-	284,616
	Peptide-level	60,313				3	616,973
	Protein-level	5,543				3	45,552
Quartet-C	Precursor-level	78,832		6 for D5/F7/M8; 18 for F6		-	192,985
	Peptide-level	55,600				3	446,569
	Protein-level	5,390				3	39,778
Simulated-B	Precursor-level	4,950	3	9 for group1/2/3	-	-	14,850
	Peptide-level	3,000				3	27,000
	Protein-level	300				3	2,700
Simulated-C	Precursor-level	4,950		6 for group2/3; 9 for group1		-	14,850
	Peptide-level	3,000				3	27,000
	Protein-level	300				3	2,700

Supplementary table 2. Detailed n statistics for MaxLFQ-quantified data matrices.

Scenario	Correction level	Feature	Sample		Algorithm/Method		Total
			Biological group	Biological replicate	BECA	QM	
Quartet-B	Uncorrected	5,541	4	18 for D5/D6/F7/M8	-	MaxLFQ	21,559
	Precursor-corrected	5,635			7		113,330
	Peptide-corrected	5,645			6		125,218
	Protein-corrected	5,635			6		120,292
Quartet-C	Uncorrected	5,379		6 for D5/F7/M8; 18 for F6	-		18,943
	Precursor-corrected	5,518			6		109,569
	Peptide-corrected	5,548			6		110,731
	Protein-corrected	5,551			6		103,150
Simulated-B	Uncorrected	300	3	9 for group1/2/3	-		900
	Precursor-corrected	300			6		5,100
	Peptide-corrected	300			6		5,100
	Protein-corrected	300			6		5,100
Simulated-C	Uncorrected	300		6 for group2/3; 9 for group1	-		900
	Precursor-corrected	300			6		5,100
	Peptide-corrected	300			6		5,100
	Protein-corrected	300			6		5,100

Supplementary Methods

Data Simulation

For our simulated datasets, we first assume the data contains m samples ($j = 1, \dots, m$), distributing to B batches ($B = 3$); each batch was composed of G biological groups ($G = 3$), each group was composed of three replicates ($R = 3$), i.e., the total number of samples m is 27. Based on the analysis of real-world label-free quantification data from a Quartet proteomic study, we designed the following parameters for our simulated data.

1. Data Structure Simulation

The number of total proteins, peptides, and precursors was set to 300, 3000, and 4950, respectively. We assigned each protein at least one peptide and each peptide at least one precursor, with the remaining items assigned randomly. This resulted in a simulated precursor-peptide-protein mapping relationship. We classified the proteins into three categories based on their effects: combined biological and batch effect (protein $n = 200$; peptide $n = 2,002$; precursor $n = 3,311$), pure batch effect (protein $n = 96$; peptide $n = 955$; precursor $n = 1,569$), and constant housekeeping proteins (protein $n = 4$; peptide $n = 43$; precursor $n = 70$).

2. Base Expression Simulation

The base log expression vector of size $n \times m$ was generated from a gamma distribution, which closely resembles real-world data.

$$\psi_i \sim i.i.d. \text{Gamma}(a, b)$$

where the i represents the precursors matched with the i^{th} protein; the shape ($a = 150.2$) and the rate ($b = 8.4$) parameters were determined by fitting a gamma distribution to the mean log expression values of real proteomic data.

3. Biological Effect Simulation

For proteins with a biological effect, we introduced log fold changes based on 8 differentially expression patterns (“-” represents down regulated, “+” represents up regulated):

Sample Pairs	Pattern I	Pattern II	Pattern III	Pattern IV	Pattern V	Pattern VI	Pattern VII	Pattern VIII
Group 2 / Group 1	+	-	0	0	+	+	-	-
Group 3 / Group 1	0	0	+	-	+	-	+	-

We generated the log fold changes of up-regulation from a gamma distribution:

$$\rho_{i,p}^+ \sim i.i.d. \text{Gamma}(c, d)$$

A negative sign was applied to ensure down-regulation:

$$\rho_{i,p}^- \sim i.i.d. -\text{Gamma}(c, d)$$

where the i represents the precursors matched with the i^{th} protein, the p represents the pattern; the shape ($c = 10$) and the rate ($d = 6$) parameters were determined by fitting a gamma distribution to the mean log fold changes of real proteomic data.

We then concatenated the up-regulated and down-regulated vectors of each sample pairs to obtain the biological effect matrix ρ :

$$\rho = (\rho^+, \rho^-)$$

For each sample pair, the biological effect vector was added to the base expression values. A random measurement noise was drawn from a normal distribution with a standard deviation. The expression matrix X of size $n \times m$ with class effects is then drawn from a normal distribution with mean $\psi + \rho$ and standard deviation ε ($\varepsilon = 0.5$):

$$X_{i,j,p} \sim i.i.d. \text{Normal}(\psi + \rho, \varepsilon)$$

where the i represents the precursors matched with the i^{th} protein, j represents the j^{th} sample, and the p represents the pattern.

4. Sample-Specific Scaling Factor Simulation

To introduce sample-specific effects, we generated a scaling factor vector α from a normal distribution with a mean of 0 and a standard deviation ε ($\varepsilon = 0.2$). This vector was added to the expression matrix X with the biological effects:

$$\alpha_{i,j} \sim i.i.d. Normal(0, \varepsilon)$$

where the i represents the precursors matched with the i^{th} protein, the j represents the j^{th} sample.

5. Batch Effect Simulation

Finally, the batch means β were drawn from a normal distribution with a mean of 0 and a standard deviation δ ($\delta = 2$):

$$\beta_{i,b} \sim i.i.d. Normal(0, \delta)$$

where the i represents the precursors matched with the i^{th} protein, and the b represents the b^{th} batch.

The batch-specific effects ω were then drawn from a normal distribution with mean β and a standard deviation γ ($\gamma = 0.5$):

$$\omega_{i,j,b} \sim i.i.d. Normal(\beta_{i,b}, \gamma)$$

where the i represents the precursors matched with the i^{th} protein, the j represents the j^{th} sample, and the b represents the b^{th} batch.

The resulting formula for the simulated data Y can then be represented as:

$$Y = \psi_i + X_{i,j,p} + \alpha_{i,j} + \omega_{i,j,b}$$