

1 The motivating example

[TODO: use a more representative plot]

[TODO: change the order of the panels, so that the modified and unmodified peptides are in the first row, the modified peptides without a counterpart are in the second row left column (right column empty), and the unmodified peptides without a counterpart are in the third row right column (left column empty)]

[TODO: the batches are not very clear from the plot]

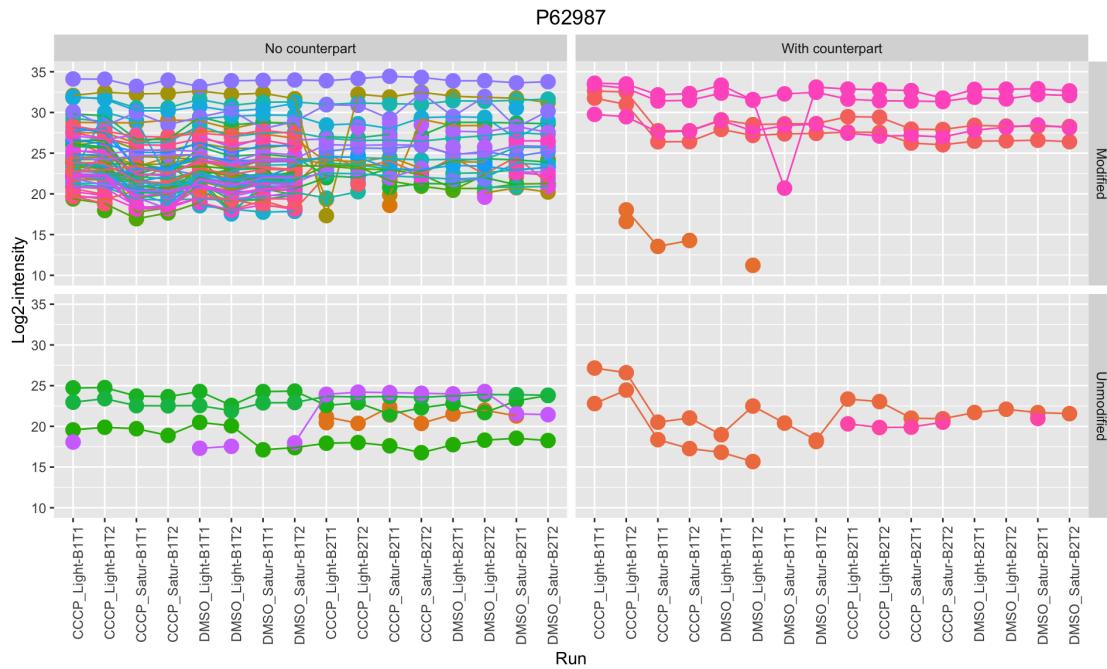


Figure S1: Profile plot of peptide features categorized based on modification and matching status.

In a typical PTM experiment, modified and unmodified forms of the same peptide can have very different quantitative characteristics such as ionization efficiency and associated variability. Observed changes in signal intensities of modified peptides may result from changes in modification as well as changes in protein abundance. To account for this confounding factor, unmodified peptides that do not contain any modified sites are used for estimation of the underlying protein abundance. Data are often acquired in multiple batches, which is another factor affecting the quantitative characteristics. As shown in the profile plot, the number of peptide features and their intensities vary across batches.

2 Overview of proposed approach and notation

[TODO: Make it more clear that this data structure reflects a particular site. Mention somewhere that we only consider peptides with one modifiable site]

[TODO: The definition of a batch is not very clear. Can you define it in a way that is not very specific to these particular experiments?]

[TODO: Why do you need stars in this table? Maybe simplify and say that this table is for unmodified peptides, and that a similar table is constructed for the modified peptides, with y^* ?]

The main objective is to detect modifications that are differentially expressed across experimental conditions.

Data structure of PTM quantification experiments. A set of fully-cleaved and/or partially-cleaved peptides with a same modification (e.g., ubiquitination) at one site are considered together. There are I conditions and J mass spectrometry runs (technical replicates) per condition, processed in B batches in the experiment. In each batch, there are $K \times L$ runs. The modification is represented by multiple spectral features (peptide ions, distinguished by their cleavage sites, mass shifts and charge states), where the number of features in Batch b is K_b . The log-intensity (base 2) of Feature k , in Run j of Condition i in Batch b is denoted by $y_{b,ijk}^*$.

Figure S2 shows an example data representation of modified peptide ions at one site. The modification measured in multiple batches is represented by different numbers of features. Unmodified peptides from the same protein are sometimes also measured, which may provide additional evidence on the underlying protein abundance. The features of unmodified peptides are denoted similarly. That is, the log-intensity (base 2) of Feature l , in Run j of Condition i in Batch b is denoted by $y_{b,ijl}$, where $l \in \{1, \dots, L_b\}$. To address the goal, statistical analysis needs to summarize values in this table using appropriate statistical models, translate the goal into a model-based quantity of interest, and draw inference (i.e., characterize the uncertainty) about the quantity.

3 Existing method: two-sample t -test

Two-sample t -test is based on the null hypothesis stating that there is no difference in mean modification abundance between Conditions i and i' against the alternative:

$$\begin{aligned} H_0 &: \mu_i^* - \mu_{i'}^* = 0 \\ H_a &: \mu_i^* - \mu_{i'}^* \neq 0 \end{aligned}$$

The test is not directly applicable to a problem with batches of data. Several *ad-hoc* approaches may be used. Three commonly used approaches are 1) t -test (no batch): ignoring batch effects when applying t -test, 2) t -test (most significant batch): applying t -test in each batch and drawing conclusions based on the most significant batch, and 3) t -test (all batch): applying t -test in each batch and drawing conclusions based on all batches. Although simple, these *ad-hoc* methods lack statistical justification. We reconstruct their meanings below and characterize their statistical properties under various forms of batch effects in Section 5. The abundance in each run is taken

		Condition 1			...		Condition I			
		Run 1	Run 2	...	Run J	...	Run 1	Run 2	...	Run J
Batch 1	Feature 1	$y_{1,111}^*$	$y_{1,121}^*$...	$y_{1,1J1}^*$...	$y_{1,I11}^*$	$y_{1,I21}^*$...	$y_{1,IJ1}^*$
	Feature 2	$y_{1,112}^*$	$y_{1,122}^*$...	$y_{1,1J2}^*$...	$y_{1,I12}^*$	$y_{1,I22}^*$...	$y_{1,IJ2}^*$

	Feature K_1	$y_{1,11K_1}^*$	$y_{1,12K_1}^*$...	$y_{1,1JK_1}^*$...	$y_{1,I1K_1}^*$	$y_{1,I2K_1}^*$...	y_{1,IJK_1}^*
...
...
Batch b	Feature 1	$y_{b,111}^*$	$y_{b,121}^*$...	$y_{b,1J1}^*$...	$y_{b,I11}^*$	$y_{b,I21}^*$...	$y_{b,IJ1}^*$
	Feature 2	$y_{b,112}^*$	$y_{b,122}^*$...	$y_{b,1J2}^*$...	$y_{b,I12}^*$	$y_{b,I22}^*$...	$y_{b,IJ2}^*$

	Feature K_b	$y_{b,11K_b}^*$	$y_{b,12K_b}^*$...	$y_{b,1JK_b}^*$...	$y_{b,I1K_b}^*$	$y_{b,I2K_b}^*$...	y_{b,IJK_b}^*
...
...
Batch B	Feature 1	$y_{B,111}^*$	$y_{B,121}^*$...	$y_{B,1J1}^*$...	$y_{B,I11}^*$	$y_{B,I21}^*$...	$y_{B,IJ1}^*$
	Feature 2	$y_{B,112}^*$	$y_{B,122}^*$...	$y_{B,1J2}^*$...	$y_{B,I12}^*$	$y_{B,I22}^*$...	$y_{B,IJ2}^*$

	Feature K_B	$y_{B,11K_B}^*$	$y_{B,12K_B}^*$...	$y_{B,1JK_B}^*$...	$y_{B,I1K_B}^*$	$y_{B,I2K_B}^*$...	y_{B,IJK_B}^*

Figure S2: Representation of the data of modified peptides at one site, measured in B batches with I conditions and J replicate runs. Modification level is quantified by multiple spectral features (peptide ions). Some spectral features can be missing (shown as $-$), either randomly in individual runs or completely from a particular form. In real practice, the number of runs can vary across conditions.

as input each of the three methods and is often estimated by sum of peak intensities, e.g., the abundance estimate for Run j of Condition i in Batch b is given by

$$y_{b,ij+} = \sum_{k=1}^{K_b} y_{b,ijk}$$

Two-sample t-test (no batch). In this approach, the modification abundance is estimated as the average over runs in all batches. For example, the estimate for Group i is given by

$$\hat{\mu}_i^* = \frac{1}{BJ} \sum_{b=1}^B \sum_{j=1}^J y_{b,ij+}.$$

This is equivalent to fitting a linear model with only an intercept, which does not account for different characteristics across batches.

Two-sample t -test (most significant batch). The modification abundance is estimated in all batches, i.e.,

$$\hat{\mu}_{b,i}^* = \frac{1}{J} \sum_{j=1}^J y_{b,ij},$$

where $\text{SE}(\hat{\mu}_{b,i}^*)$ is the standard error associated with the estimate. The test statistic in Batch b is

$$t_b = \frac{\hat{\mu}_{b,i}^*}{\text{SE}(\hat{\mu}_{b,i}^*)}$$

To summarize the evidence across batches, one strategy is to use the most significant batch. With balanced design, the batch with the largest test statistic across all batch is selected, i.e.,

$$b_{\max} = \{b \mid t_b = \max_{b'=1,\dots,B} |t_{b'}|\}.$$

The change in modification abundance is estimated as

$$\hat{\mu}_i^* - \hat{\mu}_{i'}^* = \hat{\mu}_{b_{\max},i}^* - \hat{\mu}_{b_{\max},i'}^*$$

and the test statistic for the t -test is given by

$$\frac{\hat{\mu}_i^* - \hat{\mu}_{i'}^*}{\text{SE}(\hat{\mu}_i^* - \hat{\mu}_{i'}^*)} = \frac{\hat{\mu}_{b_{\max},i}^* - \hat{\mu}_{b_{\max},i'}^*}{\left[\text{SE}(\hat{\mu}_{b_{\max},i}^*)^2 + \text{SE}(\hat{\mu}_{b_{\max},i'}^*)^2 \right]^{1/2}},$$

Two-sample t -test (all batch). A change in modification abundance is considered as statistically significant only if such change is significant in all batches. This is equivalent to take the smallest test statistic across batches for the t -test:

$$\frac{\hat{\mu}_i^* - \hat{\mu}_{i'}^*}{\text{SE}(\hat{\mu}_i^* - \hat{\mu}_{i'}^*)} = \frac{\hat{\mu}_{b_{\min},i}^* - \hat{\mu}_{b_{\min},i'}^*}{\left[\text{SE}(\hat{\mu}_{b_{\min},i}^*)^2 + \text{SE}(\hat{\mu}_{b_{\min},i'}^*)^2 \right]^{1/2}},$$

where

$$b_{\min} = \{b \mid t_b = \min_{b'=1,\dots,B} |t_{b'}|\}.$$

4 Proposed statistical modeling and inference

To characterize the observed feature intensities, different levels of variations are expressed in consideration of the following factors: modification, batch, condition, run, and feature. As the number of peptide features and their intensities are different across batches, variances of feature intensities are therefore expected to vary across batches. Also, different degrees of variability in the intensities of modified and unmodified features are present in the data; therefore, they are expressed by separate models.

[TODO: identify and present the best model that we advocate for]

4.1 Linear models and parameter estimation

The observed log-intensity of a modified peptide feature in batch b is denoted by $y_{b,ijk}^*$ and represented as

$$y_{b,ijk}^* = \psi_b^* + C_{b,i}^* + R_{b,j(i)}^* + F_{b,k}^* + (R \times F)_{b,ijk}^*,$$

where the effects of condition and feature are modeled as fixed effects:

$$\sum_{i=1}^I C_{b,i}^* = 0, \quad \sum_{k=1}^{K_b} F_{b,k}^* = 0,$$

and the effects of run and its interaction with feature are considered as random effects arising from normal distribution with mean 0:

$$R_{b,j(i)}^* = \gamma_{b,j(i)}^* \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\gamma_b^*}^2), \quad (R \times F)_{b,ijk}^* = \epsilon_{b,ijk}^* \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\epsilon_b^*}^2).$$

Similarly, the observed log-intensity of an unmodified peptide feature is denoted by $y_{b,ijl}$ and represented as

$$y_{b,ijl} = \psi_b + C_{b,i} + R_{b,j(i)} + F_{b,l} + (R \times F)_{b,ijl},$$

where the effects of condition and feature are modeled as fixed effects:

$$\sum_{i=1}^I C_{b,i} = 0, \quad \sum_{l=1}^{L_b} F_{b,l} = 0,$$

and

$$R_{b,j(i)} = \gamma_{b,j(i)} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\gamma_b}^2), \quad (R \times F)_{b,ijl} = \epsilon_{b,ijl} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\epsilon_b}^2).$$

4.1.1 Run-level summarization of feature intensities

The summarization of feature intensities at each site is carried out as in the sub-plot model of MSstats, which involves

1. Imputation of censored missing values
2. Summarization of feature intensities using Tukey's median polish (TMP), where the summarized log-intensity is denoted by $\hat{y}_{b,ij}$

4.1.2 Model-based inference of the underlying abundance

We distinguish two candidate models accounting for batch effects: 1) *per-batch model*, which assumes changes between conditions and variability in the run-level summaries are not identical across batches, and each batch needs to be modeled separately, and 2) *all-batch model*, which assumes differences between conditions are consistent across batches (i.e., no interaction effect between condition and batch) and identical variability in the run-level summarization is shared across batches. The details for each model are given below, with a summary in Table XX.

Per-batch model. Within Batch b , the abundance of the modification in each run is represented as

$$\hat{y}_{b,ij}^* = \psi_b^* + C_{b,i}^* + R_{b,j(i)}^*,$$

where $\sum_{i=1}^I C_{b,i}^* = 0$, $R_{b,j(i)}^* = \gamma_{b,j(i)}^* \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\gamma_b^*}^2)$. Similarly, the model for protein abundance in each run is given by

$$\hat{y}_{b,ij} = \psi_b + C_{b,i} + R_{b,j(i)},$$

where $\sum_{i=1}^I C_{b,i} = 0$, $R_{b,j(i)} = \gamma_{b,j(i)} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\gamma_b}^2)$. The expected values of log-abundances of the modification and protein in group i are denoted by μ_i^* and μ_i , respectively, and the values are defined as the average over the estimates in all batches:

$$\begin{aligned}\mu_i^* &= \psi^* + C_i^* = \frac{1}{B} \sum_{b=1}^B (\psi_b^* + C_{b,i}^*) \\ \mu_i &= \psi + C_i = \frac{1}{B} \sum_{b=1}^B (\psi_b + C_{b,i})\end{aligned}$$

All-batch model. The whole-plot model for the modification abundance in each run is represented as

$$\hat{y}_{b,ij}^* = \psi^* + \delta_b^* + C_i^* + R_{b,j(i)}^*,$$

where $\sum_{b=1}^B \delta_b^* = 0$, $\sum_{i=1}^I C_i^* = 0$, $R_{b,j(i)}^* = \gamma_{b,j(i)}^* \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\gamma^*}^2)$. Similarly, the model for protein abundance in each run is given by

$$\hat{y}_{b,ij} = \psi + \delta_b + C_i + R_{b,j(i)},$$

where $\sum_{b=1}^B \delta_b = 0$, $\sum_{i=1}^I C_i = 0$, $R_{b,j(i)} = \gamma_{b,j(i)} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma_{\gamma}^2)$. The expected values of log-abundances of the modification and protein in group i are defined as μ_i^* and μ_i , respectively. The values are represented as follows

$$\begin{aligned}\mu_i^* &= \psi^* + C_i^* \\ \mu_i &= \psi + C_i\end{aligned}$$

4.2 Model-based testing

Hypothesis testing on the log-abundances of the modifications in Conditions i and i' is performed to detect systematic changes in modification abundance. The hypothesis states that there is no difference in log-abundance of the modification between groups i and i'

$$\begin{aligned}H_0 : \Delta &= \mu_i^* - \mu_{i'}^* = 0 \\ H_a : \Delta &= \mu_i^* - \mu_{i'}^* \neq 0\end{aligned}$$

The difference in modification abundance, $\Delta = \mu_i^* - \mu_{i'}^*$, is estimated with the per-batch model and all-batch model as follows.

Per-batch model. The log of fold change in modification abundance is estimated as the average over batches

$$\hat{\Delta} = \frac{1}{B} \sum_{b=1}^B \left[\frac{1}{J} (\hat{y}_{b,i+}^* - \hat{y}_{b,i'+}^*) \right],$$

and the standard error of the estimate is

$$\text{SE}(\hat{\Delta}) = \left[\left(\frac{1}{B} \right)^2 \cdot \frac{2}{J} \sum_{b=1}^B \hat{\sigma}_{\gamma_b^*}^2 \right]^{1/2}$$

The test statistic $\hat{\Delta}/\text{SE}(\hat{\Delta})$ is compared against the t distribution. The degrees of freedom of the t distribution are approximated as

$$\left(\sum_{b=1}^B \hat{\sigma}_{\gamma_b^*}^2 \right)^2 / \sum_{b=1}^B \left[\frac{\hat{\sigma}_{\gamma_b^*}^4}{\text{df}(\gamma_b^*)} \right].$$

All-batch model. The log of fold change in modification abundance is estimated as the average over runs (and batches)

$$\hat{\Delta} = \frac{1}{B} \sum_{b=1}^B \left[\frac{1}{J} (\hat{y}_{b,i+}^* - \hat{y}_{b,i'+}^*) \right] = \frac{1}{BJ} (\hat{y}_{+,i+}^* - \hat{y}_{+,i'+}^*),$$

and the standard error of the estimate is

$$\text{SE}(\hat{\Delta}) = \left(\frac{2}{BJ} \cdot \hat{\sigma}_{\gamma^*}^2 \right)^{1/2}$$

The test statistic $\hat{\Delta}/\text{SE}(\hat{\Delta})$ is compared against the t distribution with degrees of freedom $\text{df}(\gamma^*)$.

4.3 Adjustment with protein abundance

Adjusted by protein abundance, the hypothesis states that there is no difference in log-abundance of the modification between groups i and i'

$$\begin{aligned} H_0 : \Delta &= (\mu_i^* - \mu_{i'}^*) - (\mu_i - \mu_{i'}) = 0 \\ H_a : \Delta &= (\mu_i^* - \mu_{i'}^*) - (\mu_i - \mu_{i'}) \neq 0 \end{aligned}$$

Details of the testing with the per-batch model and all-batch model are provided below.

Per-batch model. The log of fold change in normalized modification abundance, Δ , is estimated by

$$\hat{\Delta} = \frac{1}{B} \sum_{b=1}^B \left[\frac{1}{J} (\hat{y}_{b,i+}^* - \hat{y}_{b,i'+}^*) \right] - \frac{1}{B} \sum_{b=1}^B \left[\frac{1}{J} (\hat{y}_{b,i+} - \hat{y}_{b,i'+}) \right],$$

and the standard error of the estimate $\text{SE}(\hat{\Delta})$ is

$$\left[\left(\frac{1}{B} \right)^2 \cdot \left(\frac{2}{J} \right) \cdot \sum_{b=1}^B \left(\hat{\sigma}_{\gamma_b^*}^2 + \hat{\sigma}_{\gamma_b}^2 \right) \right]^{1/2}.$$

The test statistic $\hat{\Delta}/\text{SE}(\hat{\Delta})$ is compared against the t distribution, with degrees of freedom approximated by

$$\left[\sum_{b=1}^B \left(\hat{\sigma}_{\gamma_b^*}^2 + \hat{\sigma}_{\gamma_b}^2 \right) \right]^2 / \sum_{b=1}^B \left[\frac{\hat{\sigma}_{\gamma_b^*}^4}{\text{df}(\gamma_b^*)} + \frac{\hat{\sigma}_{\gamma_b}^4}{\text{df}(\gamma_b)} \right].$$

All-batch model. The log of fold change in normalized modification abundance Δ which is estimated as

$$\hat{\Delta} = \frac{1}{BJ} (\hat{y}_{+,i+}^* - \hat{y}_{+,i'+}^*) - \frac{1}{BJ} (\hat{y}_{+,i+} - \hat{y}_{+,i'+}),$$

and its standard error $\text{SE}(\hat{\Delta})$ is

$$\left[\frac{2}{BJ} (\hat{\sigma}_{\gamma^*}^2 + \hat{\sigma}_\gamma^2) \right]^{1/2}.$$

The test statistic $\hat{\Delta}/\text{SE}(\hat{\Delta})$ is compared against the t distribution, with degrees of freedom approximated by

$$(\hat{\sigma}_{\gamma^*}^2 + \hat{\sigma}_\gamma^2)^2 / \left(\frac{\hat{\sigma}_{\gamma^*}^4}{\text{df}(\gamma^*)} + \frac{\hat{\sigma}_\gamma^4}{\text{df}(\gamma)} \right).$$

5 Computer simulation

The proposed statistical approaches were evaluated using computer simulation. In particular, their properties under batch effects and adjustment with respect to unmodified peptides were evaluated.

5.1 Batch effects

The proposed statistical approaches were evaluated and compared with those approaches based on two-sample t -test (Section 3) using computer simulation. Two batches of data were generated, where various forms of batch effects were simulated, including difference in signal intensities across batches, difference in variability across batches, and interaction between batch and condition (i.e., change between condition affected by batch). The following interaction scenarios were considered in the simulation: 1) no interaction between condition and batch, i.e., same change across conditions in both batches, 2) positive interaction as 25% greater change in the batch of higher level, and 3) negative interaction as 25% lower change in the batch of higher level. In real experiment, multiple conditions are often compared together. Whereas t -test uses measurements from the two conditions being compared, the proposed models leverage measurements in all conditions for the inference of the underlying abundance. To highlight this distinction, multiple conditions of data were generated, where only one condition was simulated with a systematic change. Below are the parameters used in the simulation:

- Mean of log-intensity: 25
- Increased intensity level of Batch 2 versus Batch 1: 0, 1, 2
- Standard deviations of log-intensity:
 - Identical variances across batches: 0.2 in both Batch 1 and Batch 2
 - Different variances across batches: 0.2 and 0.3 in Batch 1 and Batch 2, respectively
- Difference between conditions in Batch 1: 0.5, 0.75, 1
- Number of replicates: 2, 3, 5
- Number of conditions: 2, 3, 4
- Number of realizations: 500

[TODO: comment on which combination of parameters was observed most often in the experimental dataset]

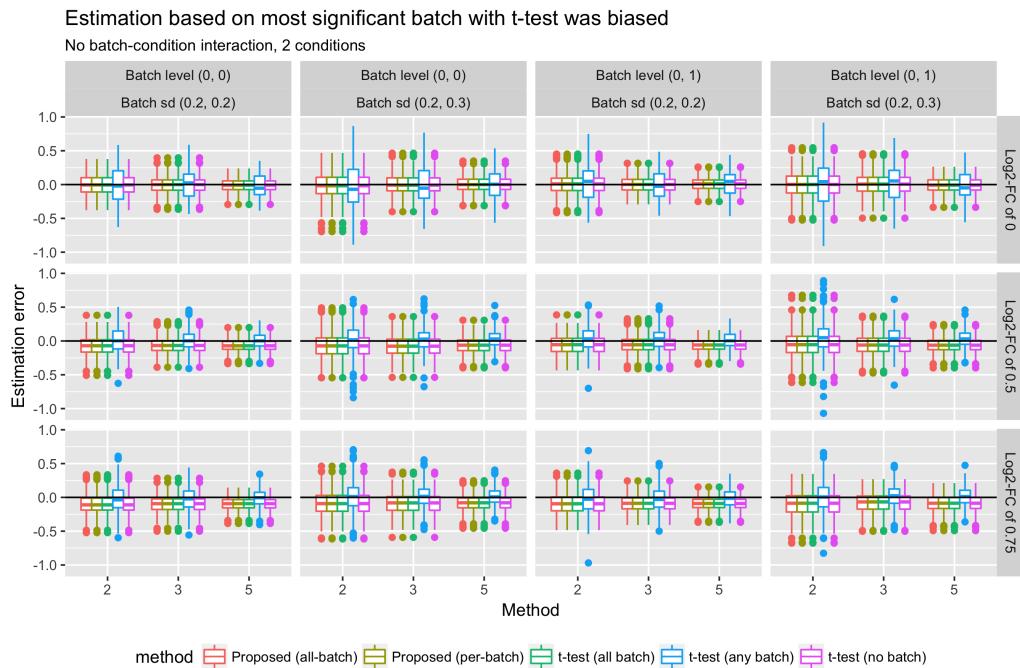


Figure S3: Estiamtion based on the most statistically significant batch with *t*-test was highly variable and frequently biased. The observation is consistent in all the simulated scenarios.

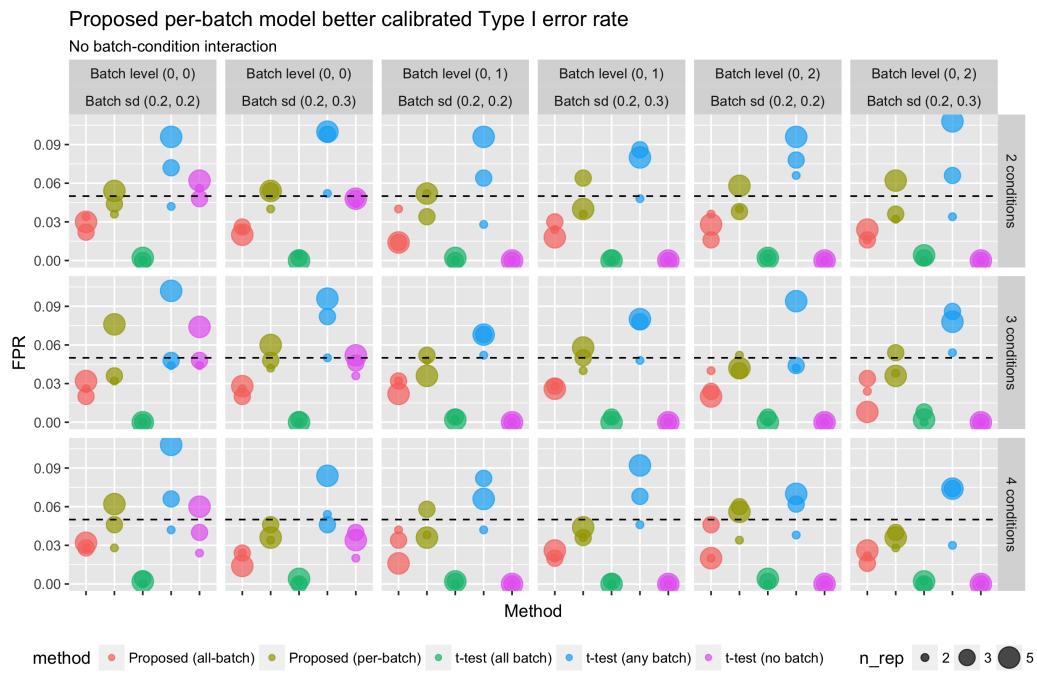


Figure S4: The proposed per-batch model better calibrated Type I error rate. Similar results were observed in the cases with positive and negative interactions.

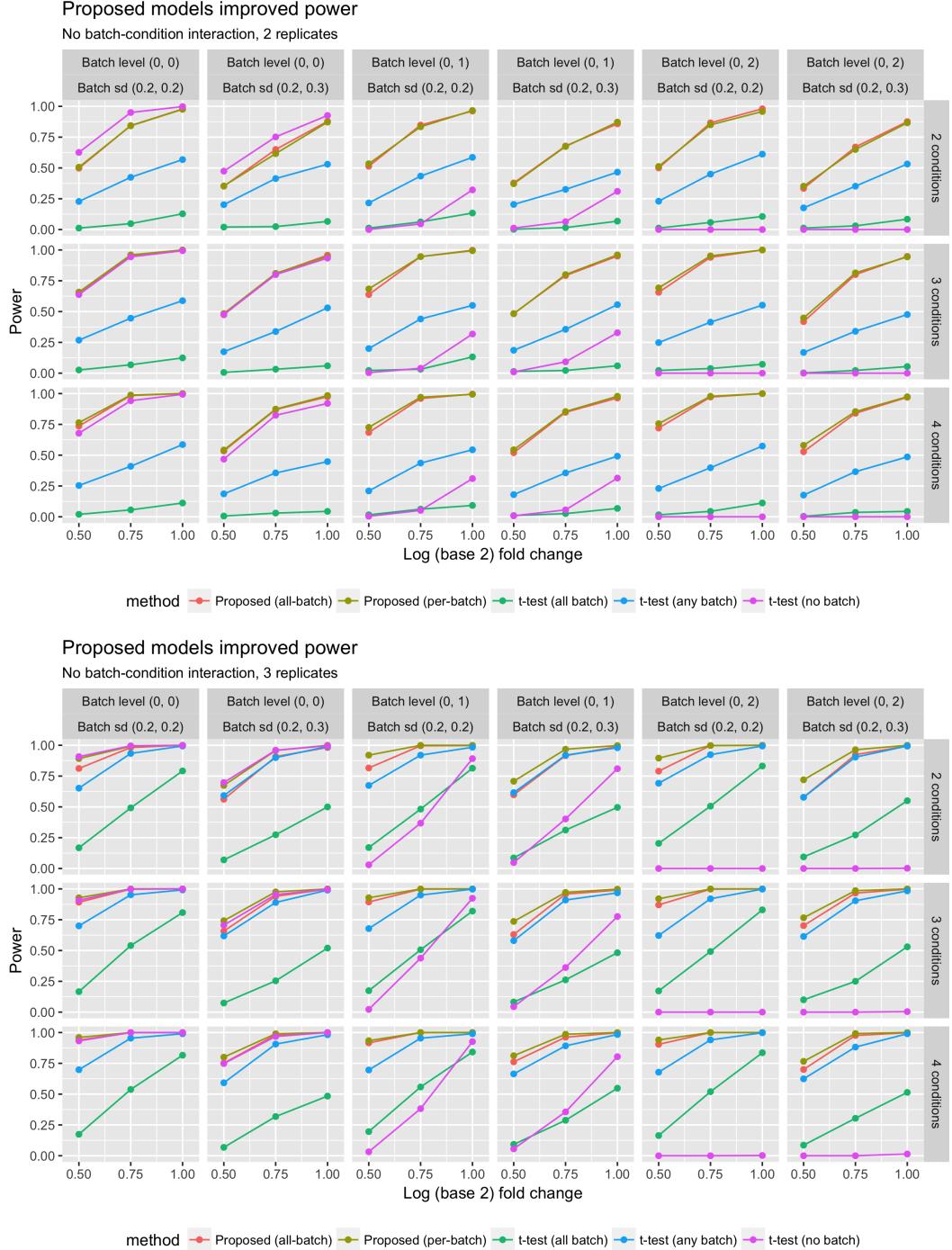


Figure S5: The proposed model improved power with small sample sizes in almost all the scenarios, under various forms of batch effects including difference in intensity level (higher on the right) and difference in variability. Using *t*-test while ignoring batch effect gave similar performance in special cases with no difference in intensity level between batches, but its performance dramatically decreased in general cases. The proposed methods gave consistently improved performance compared with other methods by properly characterizing batch effects and leveraging all available information.

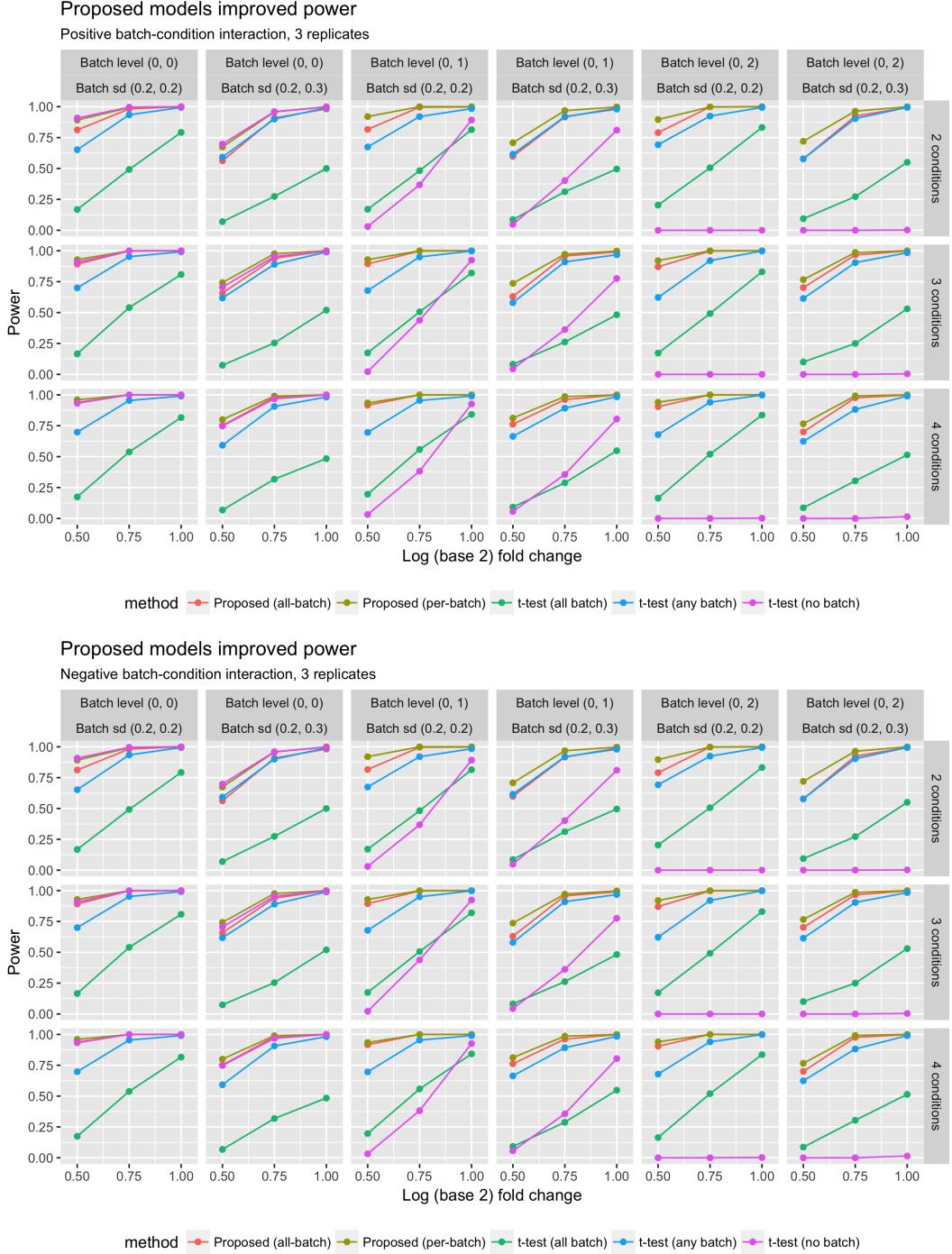


Figure S6: With 3 replicates and under the scenarios with either positive or negative interaction, the proposed models consistently improved power over other methods. Similar results were observed with 2 replicates. As in those cases with no interaction, the approaches using two-sample *t*-test did not benefit from data in additional conditions. Also, ignoring batch effects as applying *t*-test (no batch) lost power dramatically.

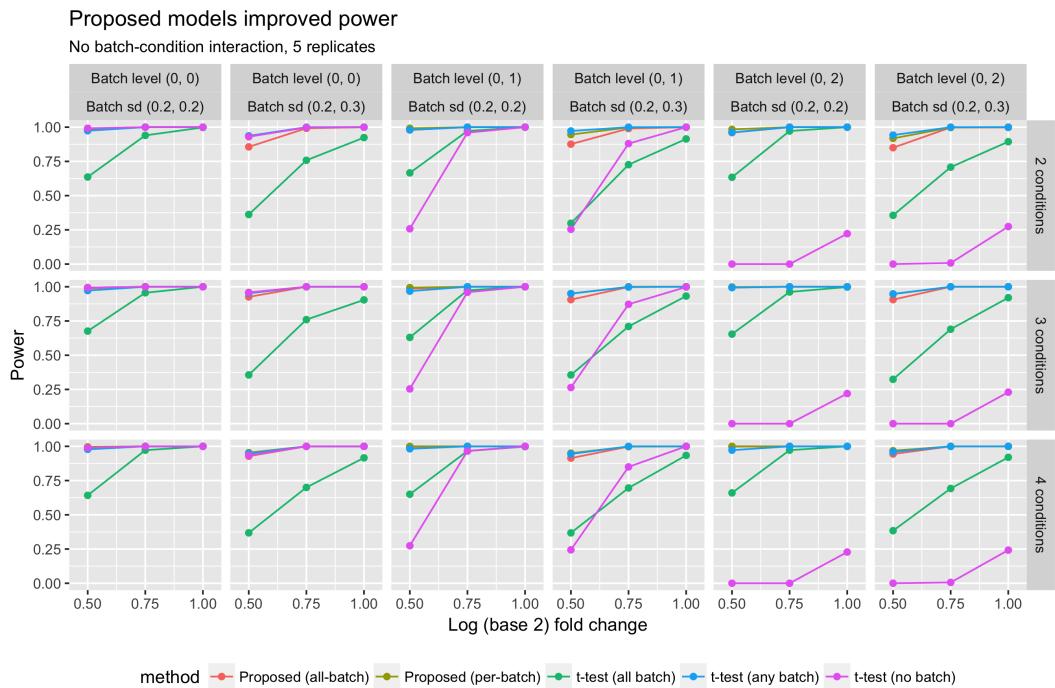


Figure S7: Ignoring batch effects lost power. The negative impact was only partially reduced by increasing the sample size to 5. Similar patterns were observed in other simulated scenarios.

5.2 Protein-level adjustment

5.2.1 One site per protein

Differential levels of modified peptides may be due to changes in modification or change in protein abundance. Adjustment with respect to unmodified peptides can correct for the confounding factor, with a cost of increased uncertainty. We evaluated such adjustment by computer simulation. Below are the parameters used in the simulation:

- Mean of log-intensity: 25
- Standard deviations of log-intensity of modified peptides:
 - Identical variances across batches: 0.2 in both Batch 1 and Batch 2
 - Different variances across batches: 0.2 and 0.3 in Batch 1 and Batch 2, respectively
- Standard deviations of log-intensity of unmodified peptides: 0.2 in both batches
- Difference in modification level between conditions: 0, 0.5
- Difference in protein level between conditions: 0, 0.5
- Number of replicates: 2, 3, 5
- Number of conditions: 2, 3, 4
- Number of realizations: 500

5.2.2 Multiple sites per protein

Expression levels of modification sites are adjusted based on the abundance of their originating protein. Since the same reference is used for all the sites, it introduces correlation between estimates and test statistics for those sites. This may cause issues in controlling false discovery rate (FDR). We investigated the property by simulating data of two conditions and 1000 proteins with different numbers of modification sites. A fraction (50% or 80%) of the 1000 proteins had no changes between conditions while systematic changes were simulated for the rest of the proteins. Multiple testing correction was performed using the Benjamini and Hochberg's method. Performances of the considered approaches were assessed by their power and FDR. The power was calculated as the fraction of proteins with adjusted p -values < 0.05 among those proteins with true differences between conditions, while the FDR was calculated as the fraction of proteins with adjusted p -values < 0.05 among the proteins with true differences.

[TODO: Storey's method?] [TODO: Proportion of null doesn't seem to be a big factor]

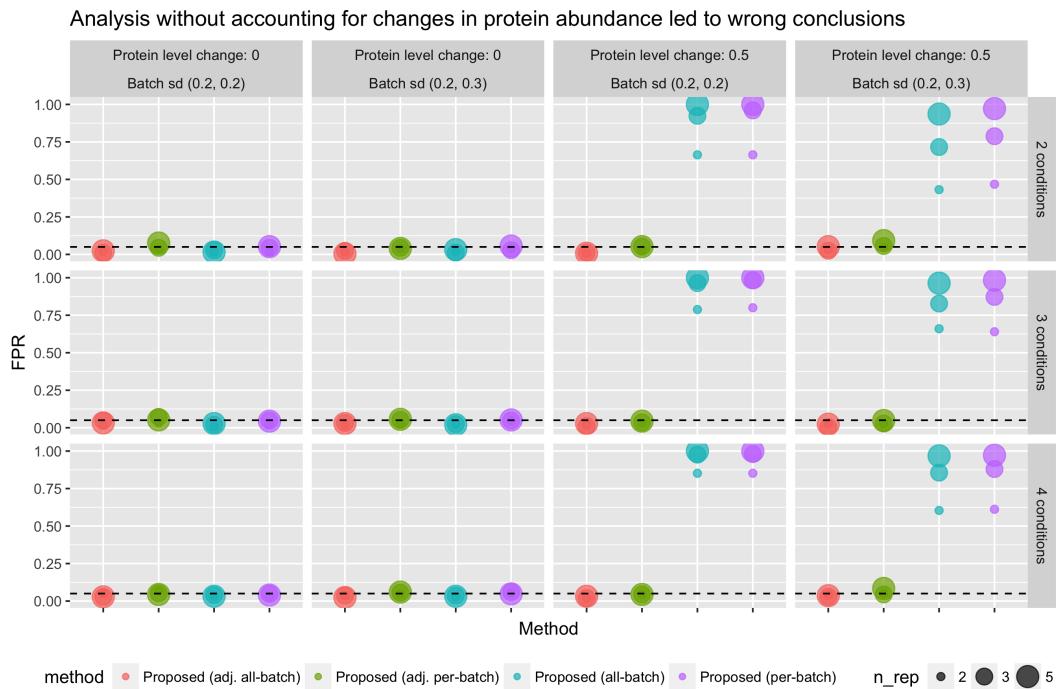


Figure S8: All the proposed methods well calibrated the Type I error rate when there was no change in protein abundance (leftmost two columns). When the modification changes were entirely due to the changes in protein abundance across conditions, analysis without accounting for the protein-level changes resulted in off-target, high false positive rates.

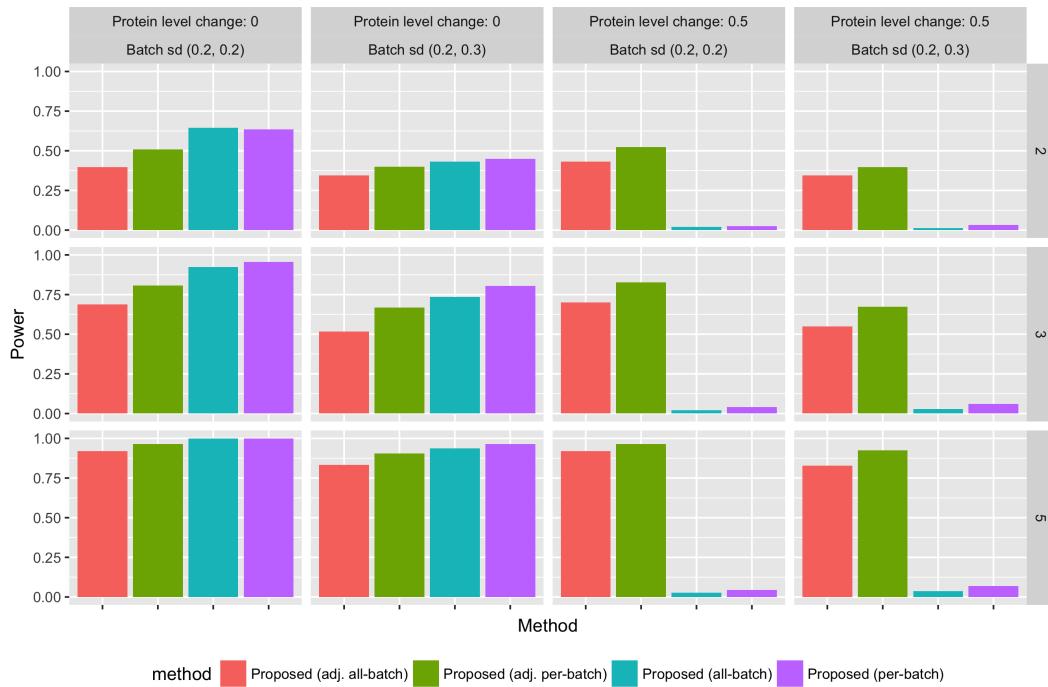


Figure S9: Adjustment with respect to unmodified peptides increase uncertainty in the estimation and testing. When there was no change in protein abundance (leftmost two columns), the approaches with no adjustment resulted in slightly higher power. However, their performances decreased significantly when such confounding indeed existed (rightmost two columns).

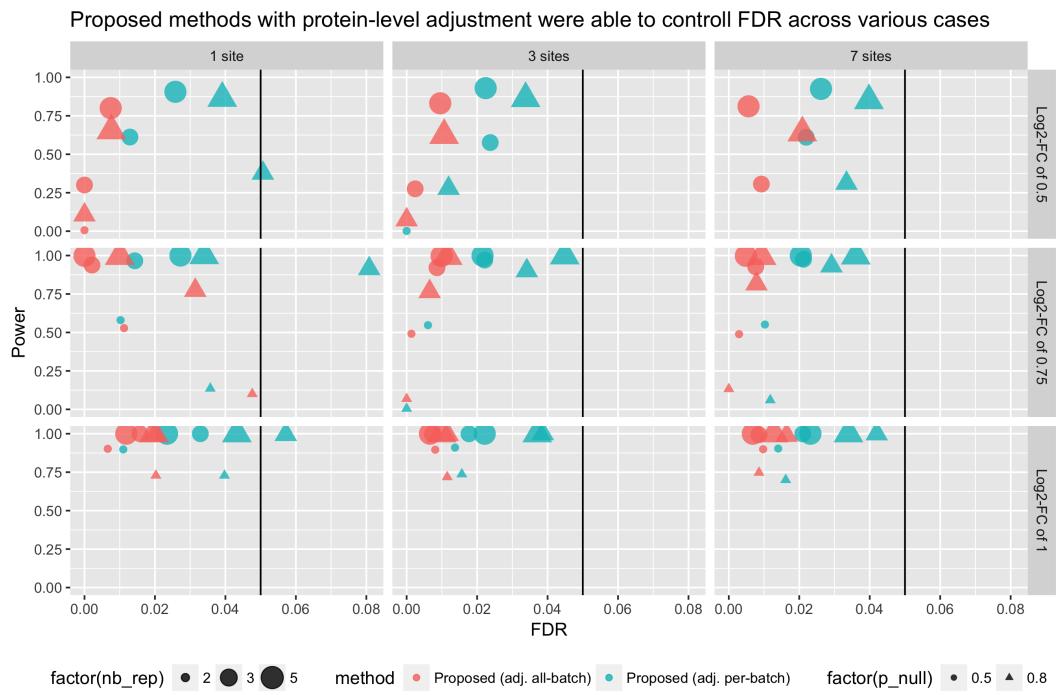


Figure S10: Correlation between test statistics across sites due to protein-level adjustment did not affect the expected level of FDR under the considered scenarios. The Benjamini-Hochberg procedure was used to control the FDR.

6 Appendix

This section contains full details of some plots that we are trying to simplify. It is only for the reference during the preparation and will be removed later.