Supplementary Information:

Wang et al.

Accurate Identification of Peptide Deamidation and Citrullination from Global Shotgun Proteomics Data Using a Dual-search Delta Score Strategy

Table of Contents

Supplementary Tables in Microsoft Excel Spreadsheets:

Table S1. Final list of unique deamidated/citrullinated peptides combined from same-seq and diff-seq identifications.

Table S2. Unique deamidated/citrullinated peptides identified from same-seq spectra.

Table S3. Unique deamidated/citrullinated peptides identified from diff-seq spectra.

Table S4. Deamidated/citrullinated PSMs identified from same-seq spectra.

Table S5. Deamidated/citrullinated PSMs identified from diff-seq spectra.

Table S6. N#Q#:to-DE type deamidated PSMs from diff-seq spectra.

Table S7. High-mass-error (< -10ppm) high-Delta-Score (~4) PSMs from diff-seq spectra.

Table S8. Citrullinated PSMs identified from benchmarking dataset from (Lee, Wang et al. 2018) by our dual-search Delta-Score strategy. Conditions: Delta Score >0.5, mass error < 5ppm, MSGF+ SpecEvalue<1e-10 & same-seq spectra. PSMs with more than one citrullination sites are excluded from results.

Supplementary Figures:

Figure S1-S6: Same-sequence spectra with different positive Delta Score

Figure S7: A same-sequence spectrum with negative Delta Score
Figure S8: A different-sequence spectrum with positive Delta Score

Figure S9: PSMs with negative Delta Score values annotated as “Valid” in the original paper is likely “Invalid”. The spectra were manually annotated by (Lee, Wang et al. 2018).

Figure S10: Relationship between Delta Score derived from MSGF+ SpecEvalue and (A) Delta Score derived from X!Tandem Evalue or (B) Delta Score derived from Comet Evalue. Conditions: mass error<5, SpecEvalue <1e-10, only 1-site modified same-seq spectra were considered.
Figure S1. A same-sequence spectrum which Delta Score equals to +0.04. Peak labeled red is deamidated fragment while peak labeled blue is anti-deamidated fragments. Dash lines indicate theoretical m/z of identified peptides.
**Figure S2.** A same-sequence spectrum which Delta Score equals to +0.50. Peaks labeled red are deamidated fragments. Dash lines indicate theoretical m/z of identified peptides.
Figure S3. A same-sequence spectrum which Delta Score equals to +1.00. Peak labeled red are deamidated fragments while peaks labeled blue are anti-deamidated fragments. Dash lines indicate theoretical m/z of identified peptides.
Figure S4. A same-sequence spectrum which Delta Score equals to +2.00. Peaks labeled red are citrullinated fragments. Dash lines indicate theoretical m/z of identified peptides.
Figure S5. A same-sequence spectrum which Delta Score equals to +4.01. Peaks labeled red are deamidated fragments. Dash lines indicate theoretical m/z of identified peptides.
Figure S6. A same-sequence spectrum with positive Delta Score and its mass error greater than 5ppm. Peaks labeled red are deamidated fragments while peaks labeled blue are anti-deamidated fragments. Dash lines indicate theoretical m/z of identified peptides.
Figure S7. Left: A same-sequence spectrum with negative Delta Score -0.62. Peaks labeled blue are anti-deamidated fragments. Dash lines indicate theoretical m/z of identified peptides. Right: MS1 of left spectrum. $^{13}$C-ion is picked as monoisotopic ion.
Figure S8. A diff-sequence spectrum with positive Delta Score. Peaks labeled red are deamidated fragments while peaks labeled blue are anti-deamidated fragments. Dash lines indicate theoretical m/z of identified peptides.
(A) SpectrumID: 3791
Annotated as "valid"
Delta Score: -1.5489863
Mass Error: -0.193754 pm
SpecEvalue: 5.18E-17

(B) SpectrumID: 3285
Annotated as "valid"
Delta Score: -1.162041
Mass Error: 0.523298 ppm
SpecEvalue: 9.94E-13
Figure S9. PSMs with negative Delta Score values annotated as “Valid” are likely “Invalid”. The spectra were manually annotated by (Lee, Wang et al. 2018).
Figure S10. Relationship between Delta Score derived from MSGF+ SpecEvalue and (A) Delta Score derived from X!Tandem Evalue or (B) Delta Score derived from Comet Evalue. Conditions: mass error<5, SpecEvalue <1e-10, only 1-site modified same-seq spectra were considered.

Reference Cited