Contribution ID: **29**   Type: **Poster**

Decoy spectrum for accurate proteoform level FDR estimation in top-down proteomics

False Discovery Rate (FDR) estimation is critical in proteomics to minimize false positive (FP) identifications. Traditional FDR estimation via the Target-Decoy Approach (TDA) mainly accounts for FPs due to fragment mass coincidences but overlooks those introduced by errors in precursor mass determination. In top-down proteomics (TDP), where spectral deconvolution assigns precursor masses, such errors are frequent and can significantly distort FDR estimates.

We introduce the concept of a decoy spectrum, which simulates false positives resulting from precursor mass errors during deconvolution. This is implemented within FLASHDeconv, which generates decoy masses mimicking typical deconvolution errors. MS2 spectra assigned to these decoy precursor masses constitute the decoy spectra. Combined with traditional target-decoy database search, this allows us to compute a more accurate refinedFDR, accounting for both types of false positives.

We evaluated this approach using an E. coli K-12 LC-MS/MS dataset. Deconvolution was performed using FLASHDeconv, and proteoform identification was conducted using TopPIC. At 5% TDA-FDR, TopPIC identified 79 decoy hits. However, an additional 86 decoy spectrum hits were found, increasing the refinedFDR to 8%. Similarly, at 1% TDA-FDR, refinedFDR was 3.25%, nearly doubling the estimated false positives.

To validate these findings, we analyzed mass shifts in identifications at 1% TDA-FDR. About 80% of shifts in the <1% refinedFDR group matched known modifications in UniMod, compared to only 58% in the >1% refinedFDR group. This suggests that many hits with high refinedFDR are likely due to incorrect precursor masses. Our results demonstrate that decoy spectra provide a more complete model of FPs in TDP and significantly improve FDR estimation, enhancing the reliability of proteoform identification and paving the way for more sensitive TDP workflows.

User consent

yes

Authors: Mrs FEROZ, Ayesha (Eberhard Karls Univeristy of Tuebingen Germany); JEONG, Kyowon (University of Tübingen)

Co-author: Prof. KOHLBACHER, Oliver (Applied Bioinformatics, Department for Computer Science, University of Tübingen)

Presenter: Mrs FEROZ, Ayesha (Eberhard Karls Univeristy of Tuebingen Germany)

Session Classification: Poster Session 1