SUPPLEMENTAL MATERIAL

A full-featured search engine for negative electron transfer dissociation

Nicholas M. Riley, Marshall Bern, Michael S. Westphall, and Joshua J. Coon

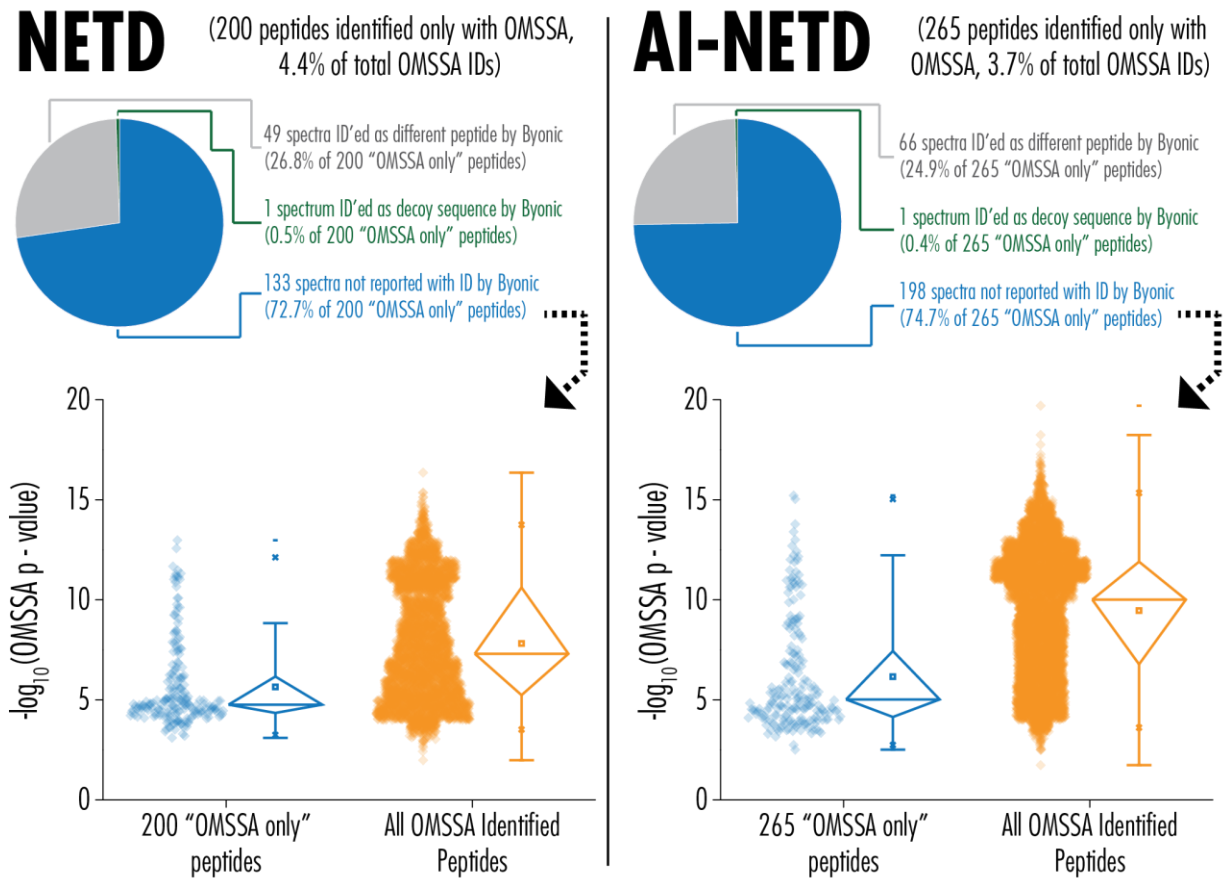
Journal of Proteome Research

Supplemental Figure 1. Peptides identified only by OMSSA are low scoring, low-quality spectra.

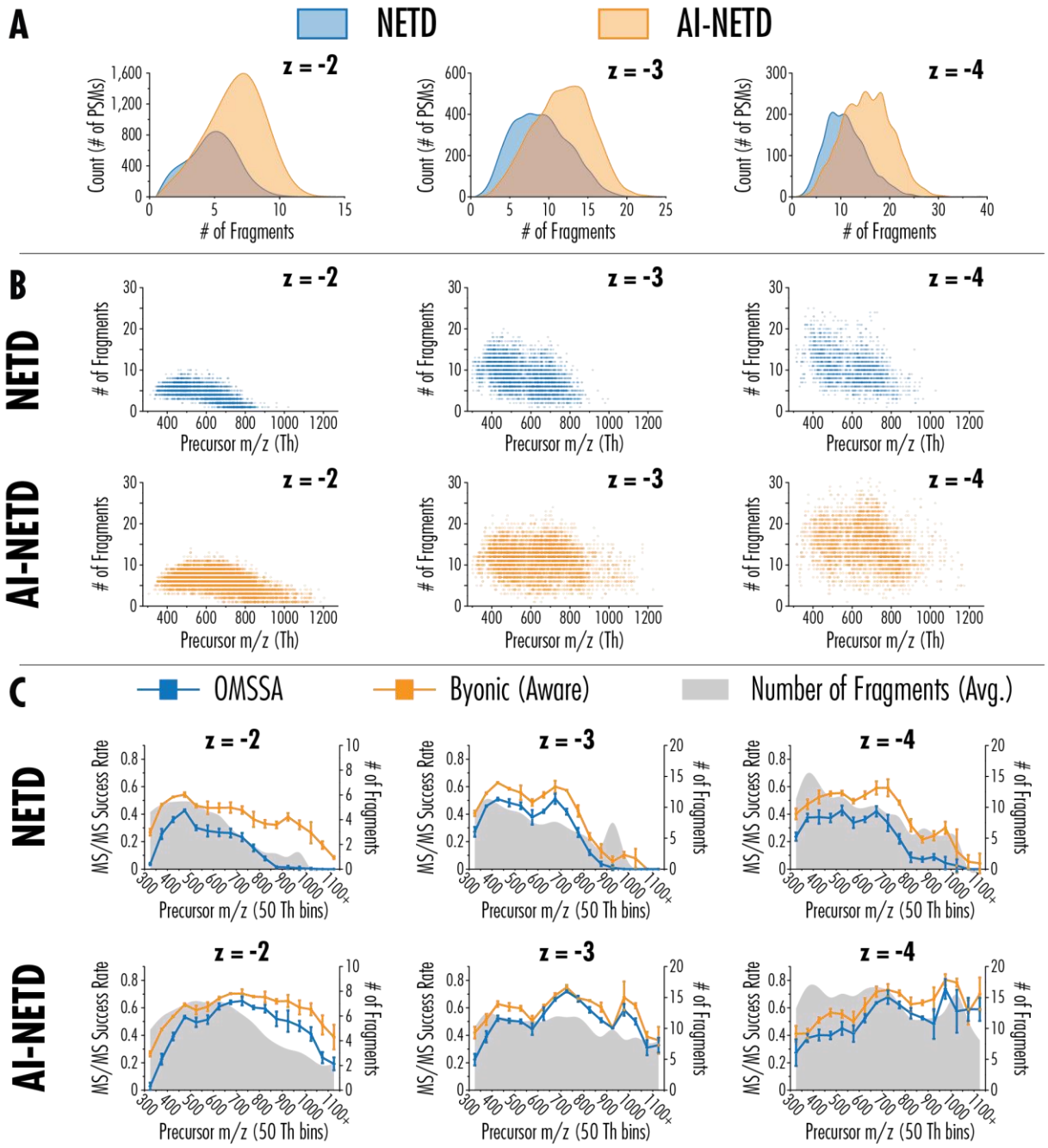
Supplemental Figure 2. AI-NETD generates more product ions than NETD, which explains why Byonic affords greater gains in identification for lower quality NETD spectra.

Supplemental Figure 3. Examining Protein Identifications with OMSSA and Byonic.

Supplemental Figure 4. Summary of the gains afforded by Byonic for unique peptide, PSM, and protein identifications.

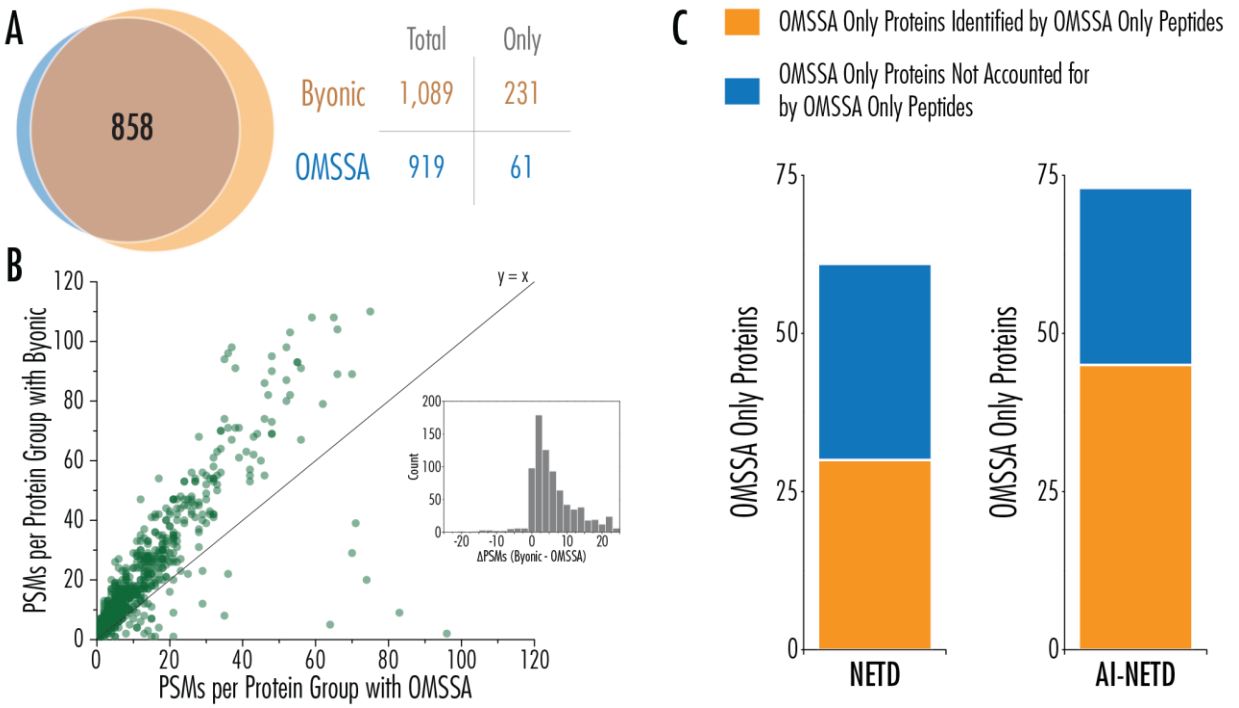


Supplemental Figure 1. Peptides identified only by OMSSA are low scoring, low-quality spectra. Approximately 25% of 200-265 spectra identified as peptides by OMSSA that were not sequenced by Byonic were assigned a different peptide identification by Byonic (grey section of pie graph). Only one spectrum identified as a peptide by OMSSA was classified as a decoy peptide by Byonic for both NETD and AI-NETD (green section of pie graph). Of the approximate three-fourths of the "OMSSA only" peptides (blue section of pie graph), the p-values returned from OMSSA (blue box plots) were far less confident than that the p-values seen for the OMSSA identifications on the whole (gold box plots), indicating that these "OMSSA only" peptides were sequenced using low quality spectra that likely did not pass Byonic's scoring/FDR thresholds. Importantly, greater than 95% of peptides sequenced with OMSSA were also identified with Byonic for both NETD and AI-NETD.

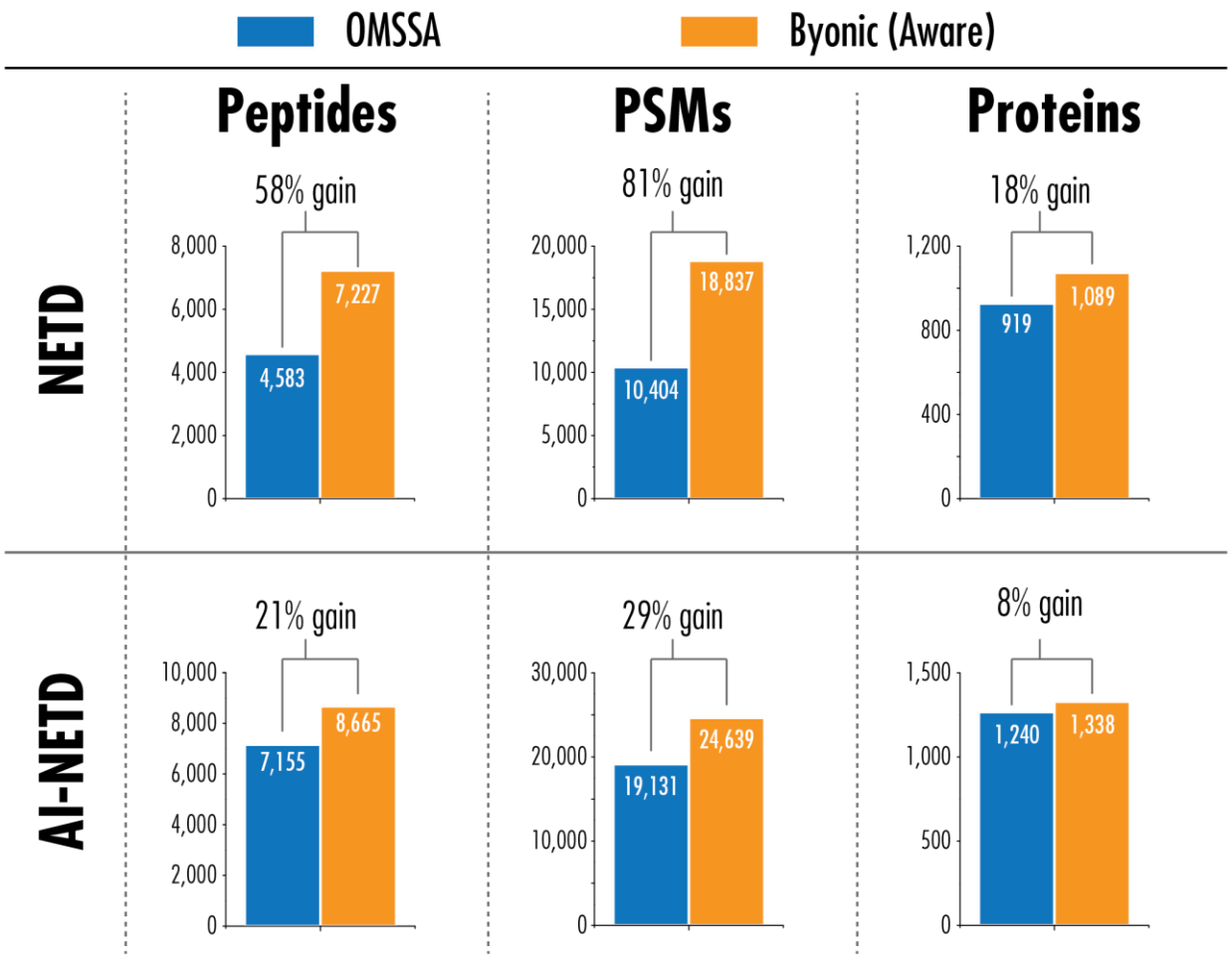


Supplemental Figure 2. AI-NETD generates more product ions than NETD, which explains why Byonic affords greater gains in identification for lower quality NETD spectra. (A) Histograms show the number of fragments (summed total of *a*-type and *x*-type ions) generated by NETD and AI-NETD for precursors with *z* = -2, -3, and -4 charge states, where AI-NETD generally produces more sequence informative fragment ions for all charge states. **(B)** Scatter plots show dependence of product ion generation on precursor *m/z*. AI-NETD clearly generates more fragment ions than NETD and extends the *m/z* range across which precursors can be successfully fragmented for all three charge states shown. Note, each point in the scatter plot is at 20% opacity, so density of color is indicative of the number of precursors at that region of the plot. **(C)** MS/MS success rates using OMSSA (blue line) and Byonic (gold line) search

algorithms are plotted as a function of precursor m/z for precursors with *z* = -2, -3, and -4 charge states. In grey, the average number of total a- and x-type fragment ions produced from a precursor is shown for the given precursor *m/z* bin. When high numbers of product ions are generated, *e.g.,* AI-NETD *z* = -3 and *z* = -4, the disparity between Byonic and OMSSA success rates is small. Byonic significantly outperforms OMSSA, however, where the number of generated product ions is low, *e.g.,* NETD *z* = -2. This supports the hypothesis that the more sophisticated scoring and 2D-FDR approaches of Byonic have more to offer to the (generally) lower quality NETD spectra, explaining why Byonic helps NETD identifications proportionally more than it helps AI-NETD.



Supplemental Figure 3. (A) The overlap in identified proteins from batched triplicate NETD analyses using OMSSA and Byonic. **(B)** For the protein groups identified by both programs, the number of PSMs per protein group is plotted with Byonic on the y-axis and OMSSA on the x-axis (for those with 120 or fewer PSMs per protein). The inset histogram shows the distribution of ΔPSMs, which is the difference of PSMs identified with Byonic minus PSMs identified with OMSSA, for all 858 protein groups shared between the two algorithms. A positive number indicates more PSMs for a given protein group with Byonic. **(C)** To investigate why 61 and 73 proteins were identified only with OMSSA ("OMSSA Only Proteins") for NETD and AI-NETD analyses, respectively, the number of proteins that were identified solely based on OMSSA Only Peptides from Figure 2 are shown in as gold portions of the bar graphs. Those OMSSA only proteins that could not be accounted for based on identification of OMSSA only proteins are shown in blue in the stacked bar graphs. These proteins, which account for only 2-3% of all OMSSA protein identifications, are likely from differences in protein grouping algorithms between the two platforms.



Supplemental Figure 4. Summary of the gains afforded by Byonic for unique peptide, PSM, and protein identifications.