Cross-run hybrid features improve the identification of data-independent acquisition proteomics

Yachen Liu,[†,‡] Longfei Mei,[†] Chenyu Liang,[‡] Chuan-Qi Zhong,[¶] Mengsha Tong,[*,‡,¶] and Rongshan Yu[*,†,‡,§]

†*School of Informatics, Xiamen University, Fujian, 361000, China*

‡*National Institute for Data Science in Health and Medicine, Xiamen University, Fujian, 361102, China*

¶*School of Life Sciences, Xiamen University, Xiamen, Fujian, 361102, China*

§*Aginome Scientific, Xiamen, Fujian, 361005, China*

E-mail: mstong@xmu.edu.cn; rsyu@xmu.edu.cn

Supplementary

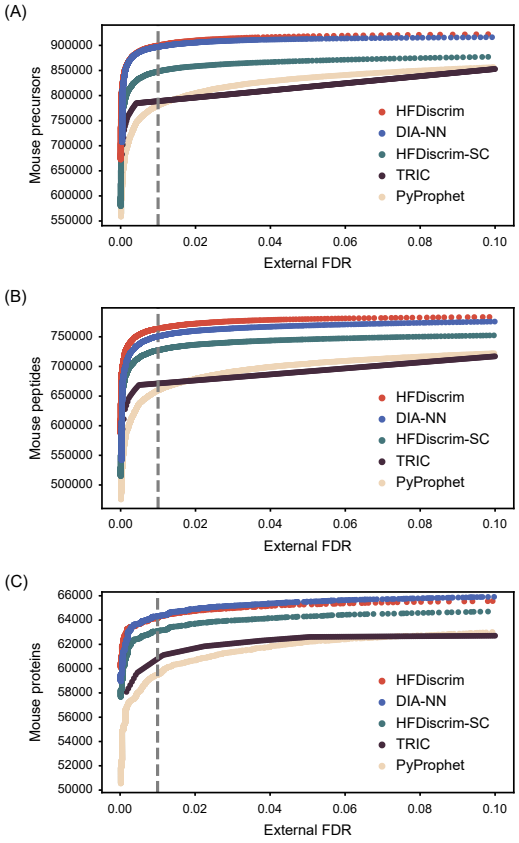

Figure S1. Identification performance on all 10 samples of the MCB dataset (DIA-based spectral library). The FDR was estimated with the 'internal' target-decoy method and with the 'external' method using mixing arabidopsis and mouse samples for generating the library and using only mouse sample in the DIA runs. (A) The number of mouse precursors identified across all 10 samples of the MCB dataset at different external FDRs was plotted. (B) Same as in (A), but reports the number of mouse peptides. (C) Same as in (A), but reports the number of mouse proteins.

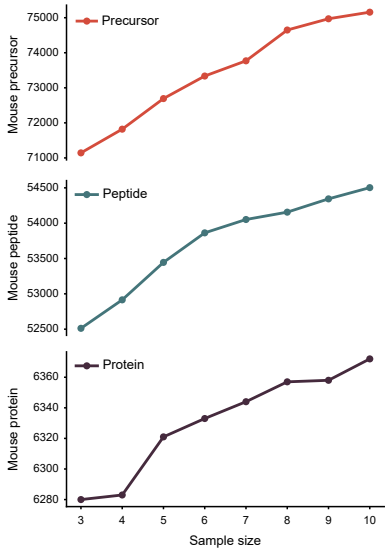

Figure S2. Impact of sample size on HFDiscrim performance. The figure shows the number of precursors, peptides, and proteins identified in the `BGS_D_D180420_S416-newPrep-DIA-D-S1-1_MHRM_R01_T0` sample at an external FDR of 0.01. The x-axis represents the sample size, and the y-axis shows the corresponding number of identified mouse precursors, peptides, and proteins.

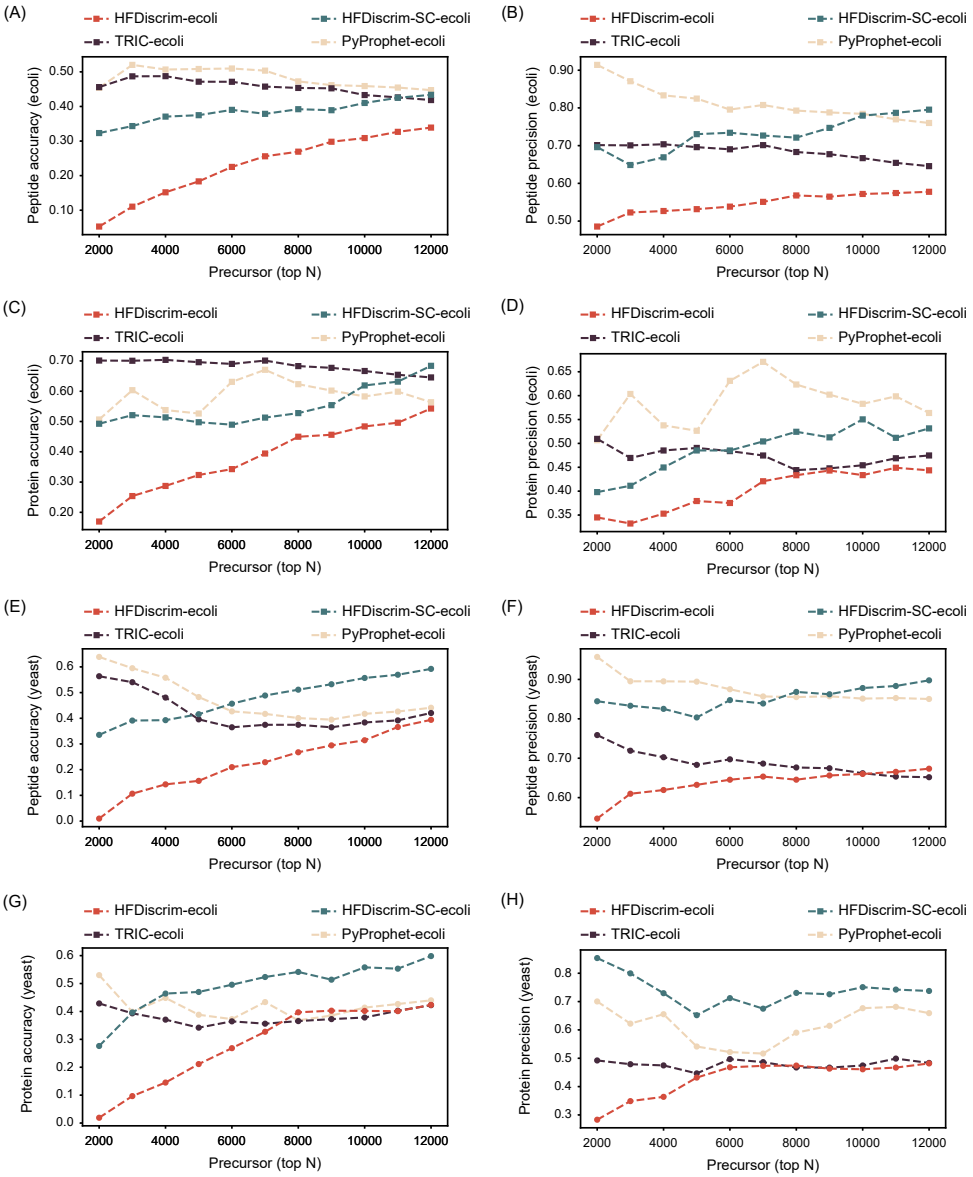

Figure S3. Quantification of HYE110 dataset that mixing proteomes from three species in defined ratio with HFDiscrim and the other tools for DIA. (A) E.coli peptide accuracy, (B) E.coli peptide precision, (C) E.coli protein accuracy, (D) E.coli protein precision, (E) Yeast peptide accuracy, (F) Yeast peptide precision, (G) Yeast protein accuracy and (D) Yeast protein precision as functions of the top N precursors identified. The x-axis represents the top N precursors, while the y-axes show the corresponding accuracy or precision values.