Supporting Information

Full-featured, real-time database searching platform enables fast and accurate multiplexed quantitative proteomics.

Devin K. Schweppe[1*], Jimmy K. Eng[2], Qing Yu[1], Derek Bailey[3], Ramin Rad[1], Jose Navarrete-Perea[1], Edward L. Huttlin[1], Brian K. Erickson[1], Joao A. Paulo[1]*, Steven P. Gygi[1]*

1 – Department of Cell Biology, Harvard Medical School, Cambridge, MA
2 – University of Washington Proteomics Resource, Seattle, WA
3 – Thermo Scientific LSMS, San Jose, CA
* – Corresponding Authors

Correspondence should be directed to:
Devin K Schweppe
devin_schweppe@hms.harvard.edu
Joao A Paulo
joao_paulo@hms.harvard.edu
Steven P Gygi
steven_gygi@hms.harvard.edu

Department of Cell Biology
Harvard Medical School
240 Longwood Ave
Boston, MA

Contents
Figures

Figure S1 – Duty cycle for RTS-MS3 and SPS-MS3 workflows.

Figure S2 – Histograms comparing mass error of the theoretical matched peptide MH+ versus instrument reported precursor mass or the Orbiter-corrected monoisotopic MH+.

Figure S3 – Real-time false discovery rate (rtFDR) to separate target and decoy matches and calculate false discovery rates (FDRs) in real-time.

Figure S4 – Convergence of the LDA Score threshold.

Figure S5 – Hyper standard quantitative comparison.

Figure S6 – Scatter plots for the three cell line binary comparisons.

Figure S7 – Western blotting analysis of Endoglin (ENG).

Figure S8 – Western blotting analysis of ALPK2.

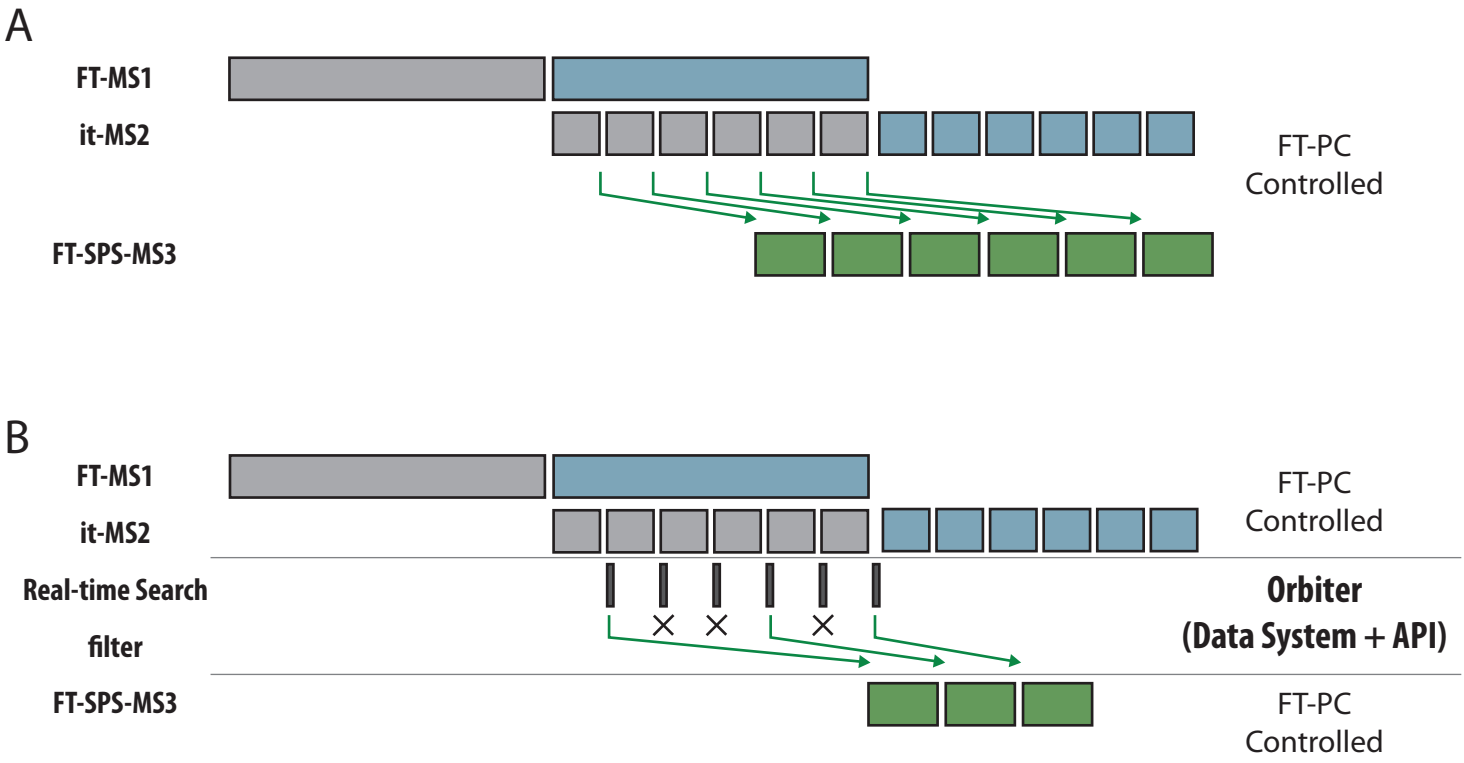


Figure S1
Duty cycle for RTS-MS3 and SPS-MS3 workflows. **A.** SPS-MS3 workflow. Precursors are identified in high-resolution MS1 spectrum and targeted for fragmentation and subsequent analysis in low resolution MS2 scans (ion trap [it] MS2). SPS ions are determined based on the most intense ions in the MS2 scans and for every MS2 spectrum ions are selected for further fragmentation and analysis in a third high resolution MS3 scan. **B.** RTS-MS3 workflow. Instead of selecting SPS ions for every MS2 scan, RTS-MS3 scans are only triggered when all filters (search score, rtFDR, etc) are passed. Thereby only a subset of MS2 scans trigger SPS-MS3 scans for quantitation, saving time for further MS1 and MS2 scans.

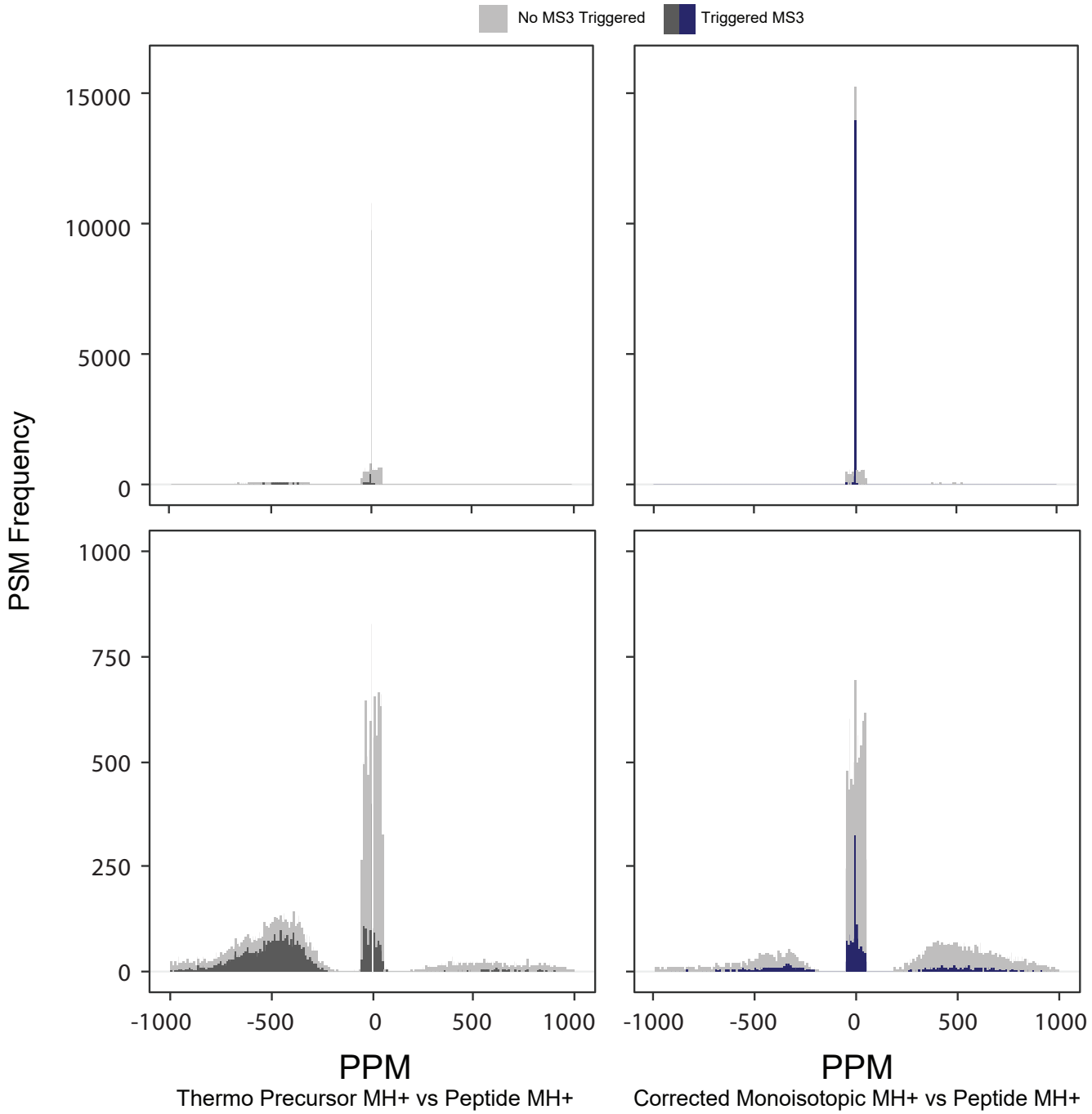


Figure S2

Histograms comparing mass error of the theoretical matched peptide MH+ versus instrument reported precursor mass (left panels) or the Orbiter-corrected monoisotopic MH+ (right panels). Light grey bars indicate PSMs that did not trigger RTS-MS3 scans. Dark grey and blue bars indicate PSMs that triggered RTS-MS3 scans. Bottom panels show a zoomed in version of the top panel (maximum PSM frequency of 1000); bins with frequencies greater than 1000 are not shown in these panels. Histogram bin width is 5 ppm. Monoisotopic peak correction resulted in 23% more PSMs within +/-50ppm of the theoretical peptide MH+.

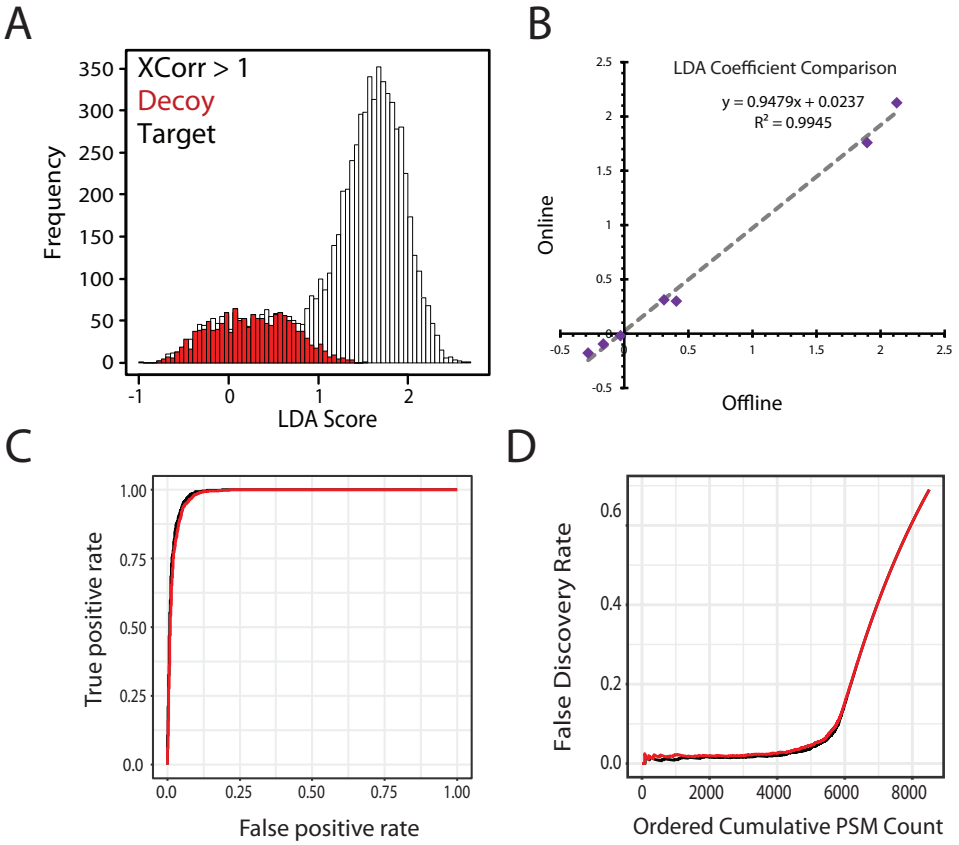


Figure S3

Real-time false discovery rate (rtFDR). The rtFDR framework uses multidimensional linear discriminant analysis (LDA) to separate target and decoy matches and calculate false discovery rates (FDRs) in real-time with instrument scan acquisition. We observed real-time performance commensurate with offline calculation[1]. **A.** Separation of target and decoy peptides by LDA score based on seven peptide spectral match metrics: XCorr, deltaCorr, missed cleavages, charge state, absolute ppm error, peptide length, and the fraction of ions matched. **B.** LDA coefficients compared from offline (R, MASS implementation) and online (C# Accord) for the same TKO sample run. These data were highly correlated. **C.** Receiver operator curve for offline (black) and online LDA scoring (red). **D.** Cumulative FDR curves for offline (black) and online LDA scoring (red).

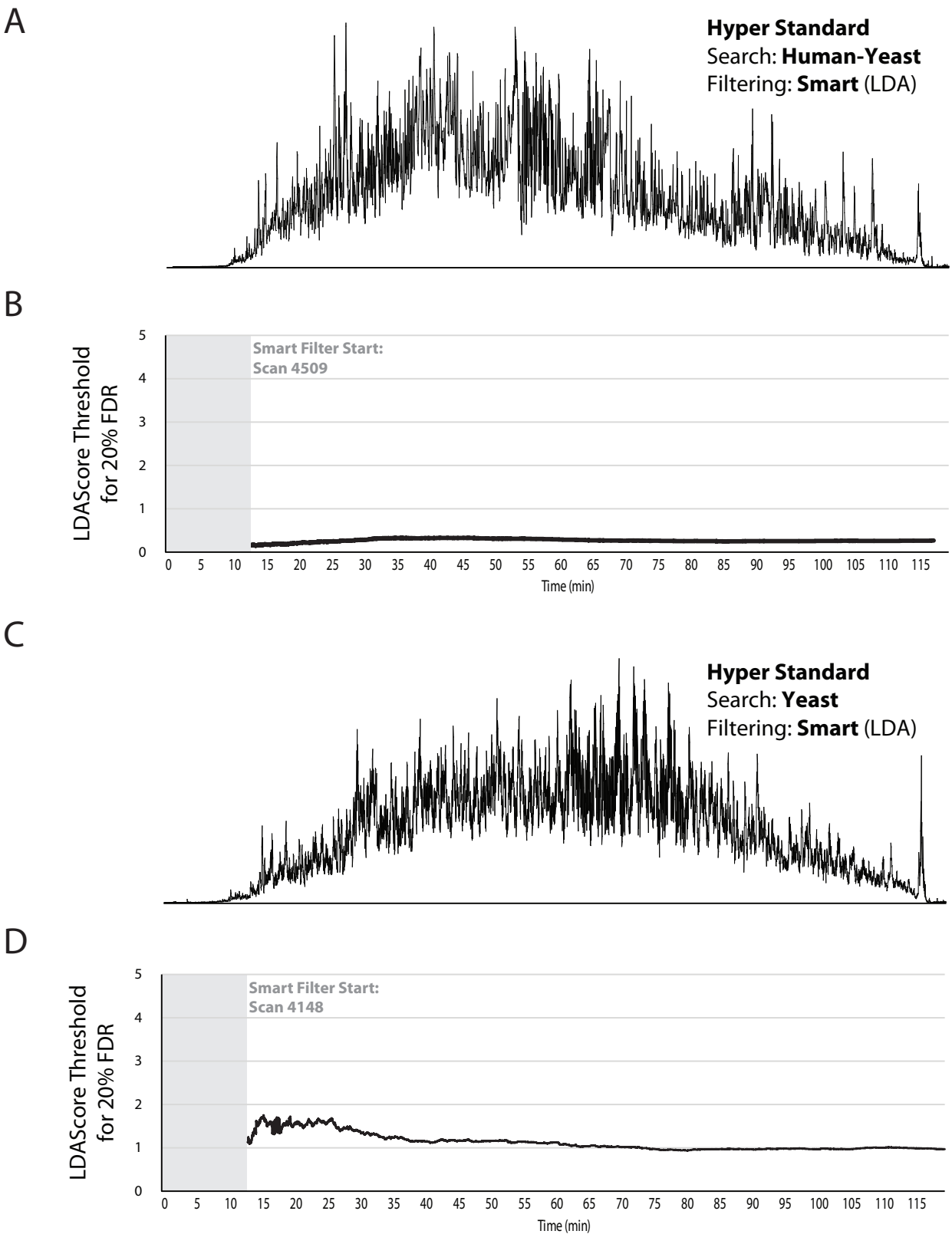
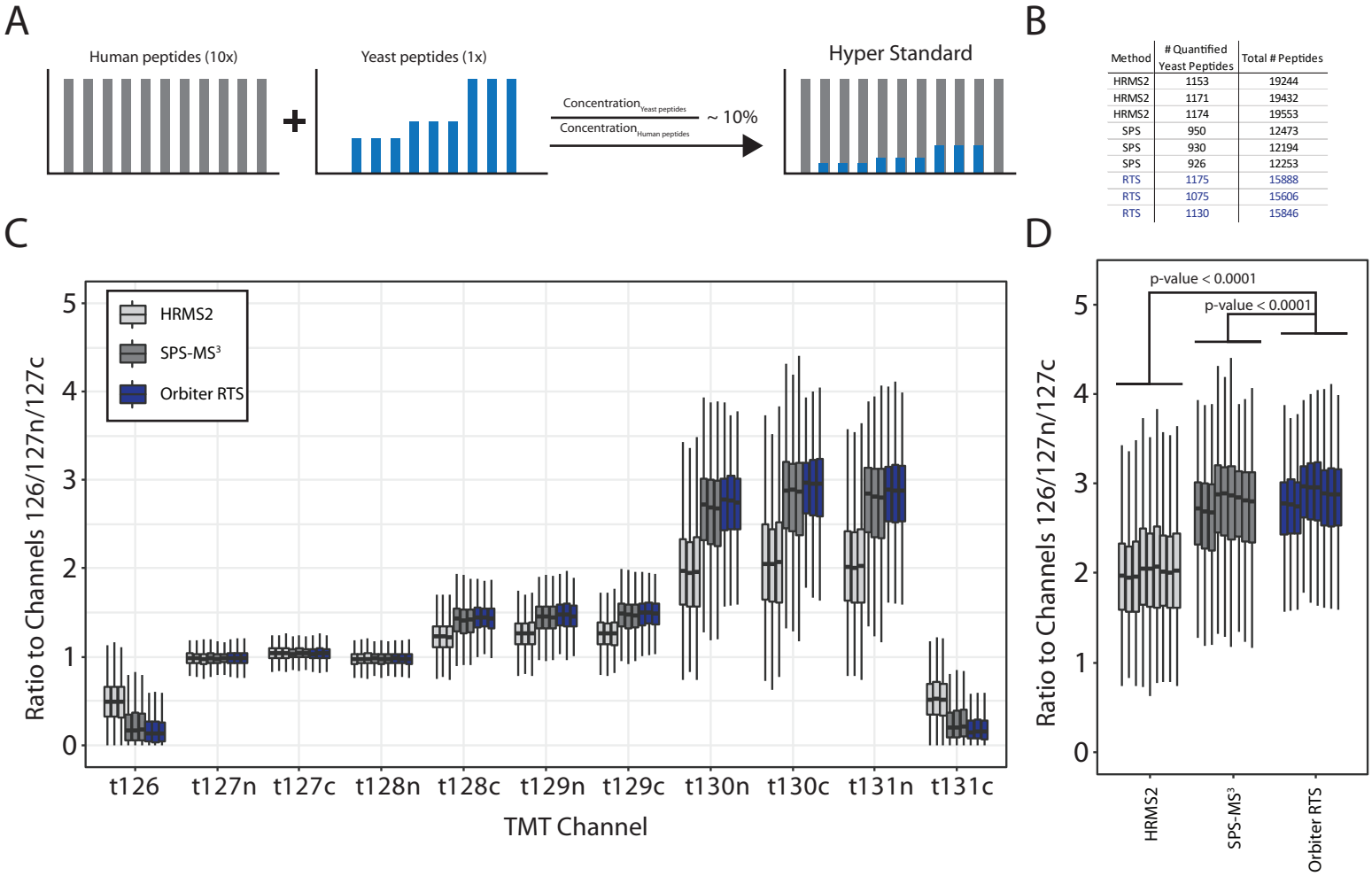


Figure S4

Convergence of the LDA Score threshold. **A.** Total ion chromatogram of the MS1 scans for Hyper standard eluting over a 120minute gradient. RTS used a concatenated human-yeast database. **B.** LDA Score threshold for a 20% FDR over the course of the run in **A**. **C.** Total ion chromatogram of the MS1 scans for Hyper standard eluting over a 120minute gradient. RTS used a only the yeast database. **D.** LDA Score threshold for a 20% FDR over the course of the run in **C**. The LDA Score threshold converged with either databased searching method, but converged and stabilized much faster when searching the two proteome Hyper standard with a two-proteome database.



Method	# Quantified Yeast Peptides	Total # Peptides
HRMS2	1153	19244
HRMS2	1171	19432
HRMS2	1174	19553
SPS	950	12473
SPS	930	12194
SPS	926	12253
RTS	1175	15888
RTS	1075	15606
RTS	1130	15846

Figure S5

Hyper standard quantitative comparison. **A.** The Hyper standard is a two proteome mixture consisting of human and yeast peptides mixed at an approximately 10:1 ratio. The yeast peptides are distributed across TMT channels to generate multiple comparable ratios to assess quantitative accuracy (1x, 1.5x, 3x). **B.** Quantified yeast and total peptides for the three technical replicates of each method test. **C.** Boxplot of ratios for quantified yeast peptide ratios for the three technical replicates of each method. **D.** The three-fold change ratios for the RTS quantification were significantly increased (i.e. closer to 3) compared to HRMS2 and SPS-MS3 (ANOVA/Tukey).

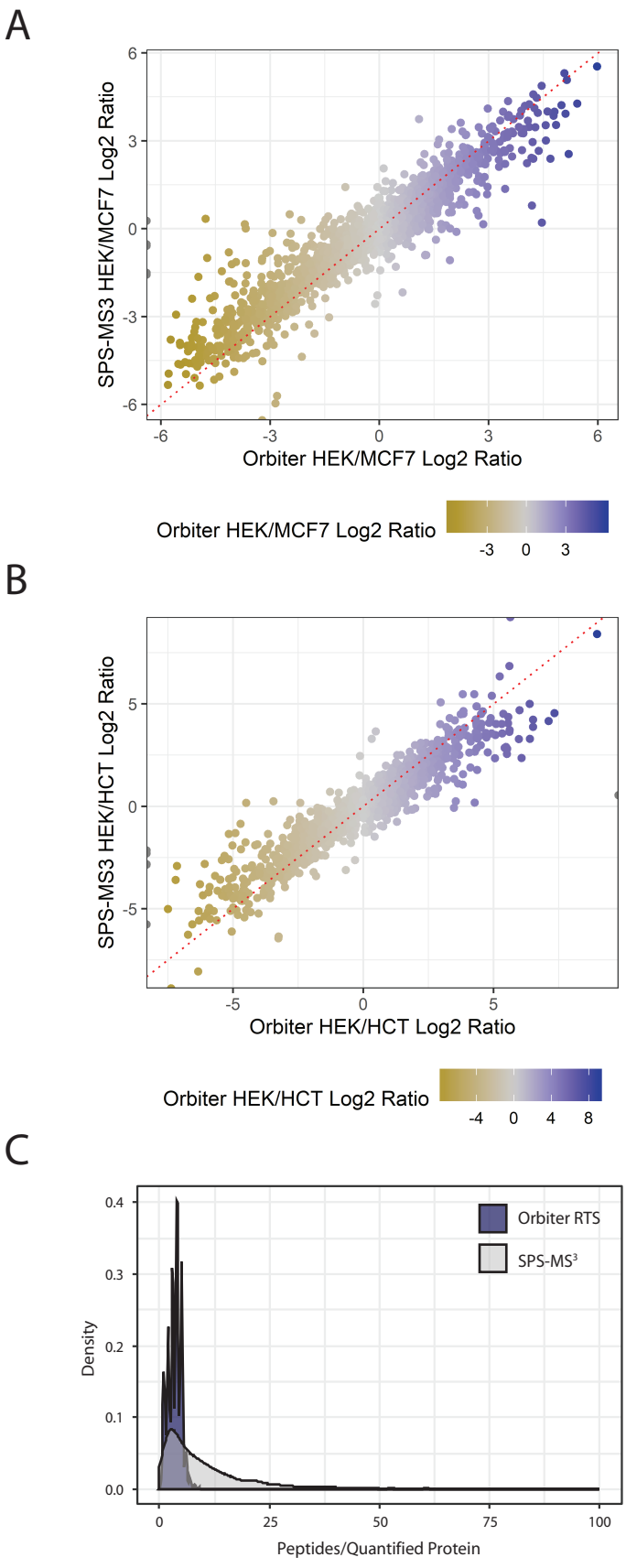


Figure S6

Three cell line comparison of SPS-MS3 and RTS acquisitions. Scatter plots for the three cell line binary comparisons of **A** HEK293T (HEK) versus MCF7 and **B** HEK293T (HEK) versus HCT116 (HCT). The red lines indicates the x = y regression. C. Quantified peptides per protein for the SPS-MS3 (grey density plot) and Orbiter RTS (blue density plot) analyses. Orbiter RTS uses a protein closeout to limit to number of peptides per run that are quantified.

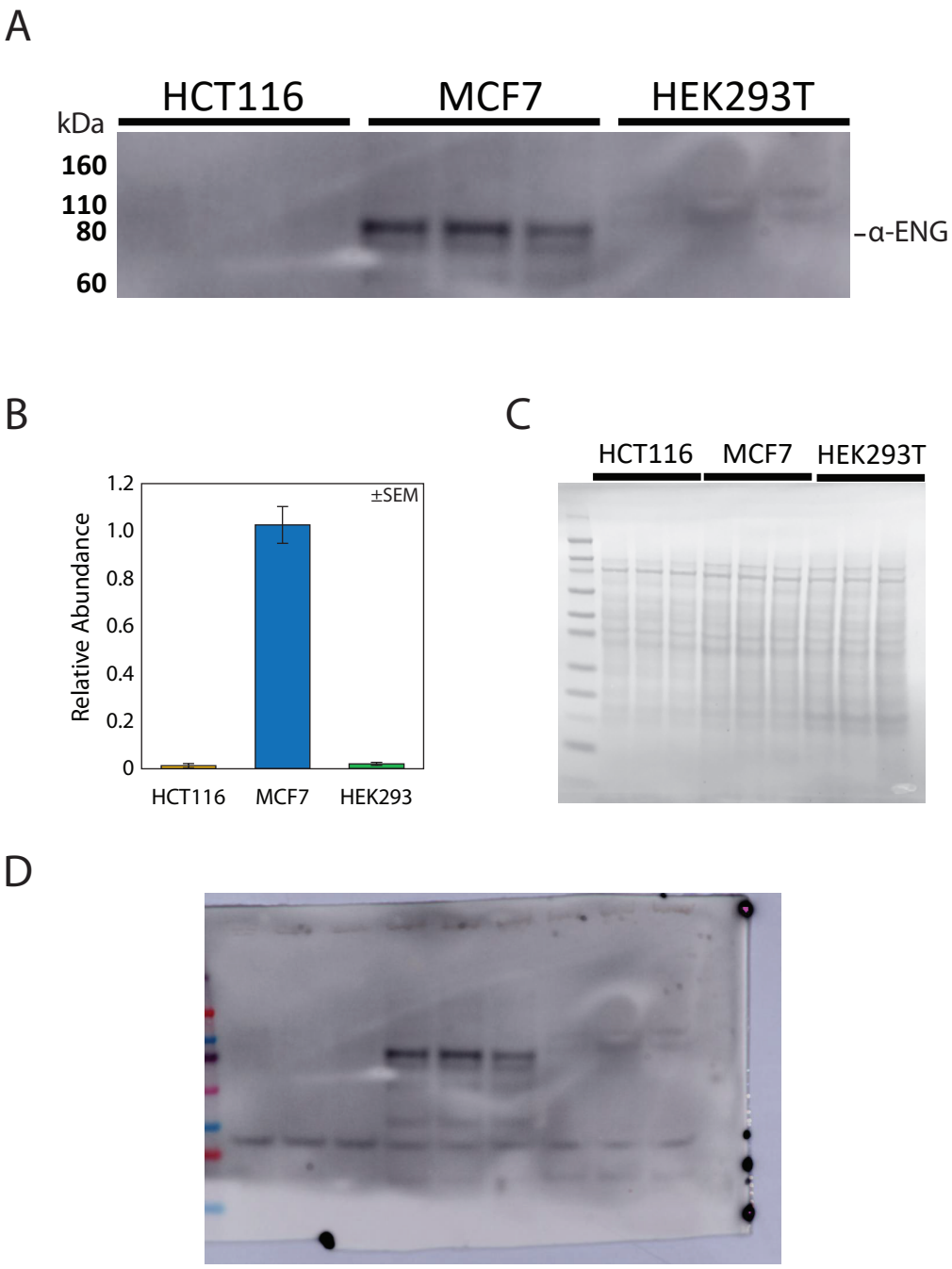


Figure S7

Western blotting analysis of Endoglin (ENG) abundance across three cell lines. **A.** Western blot for ENG revealed strong expression in MCF7 cells as seen in the RTS-MS3 analysis but not the SPS-MS3 analysis (n=3). Molecular weights (kDa) are shown. **B.** Quantitation of relative abundance of ENG across the cell lines in **A**. Error bars represent standard error of the mean (SEM) for the quantified lanes. **C.** Ponceau S staining of the PVDF membrane used in **A** demonstrating equal loading of each cell line. **D.** Image of the entire blotted membrane in **A**.

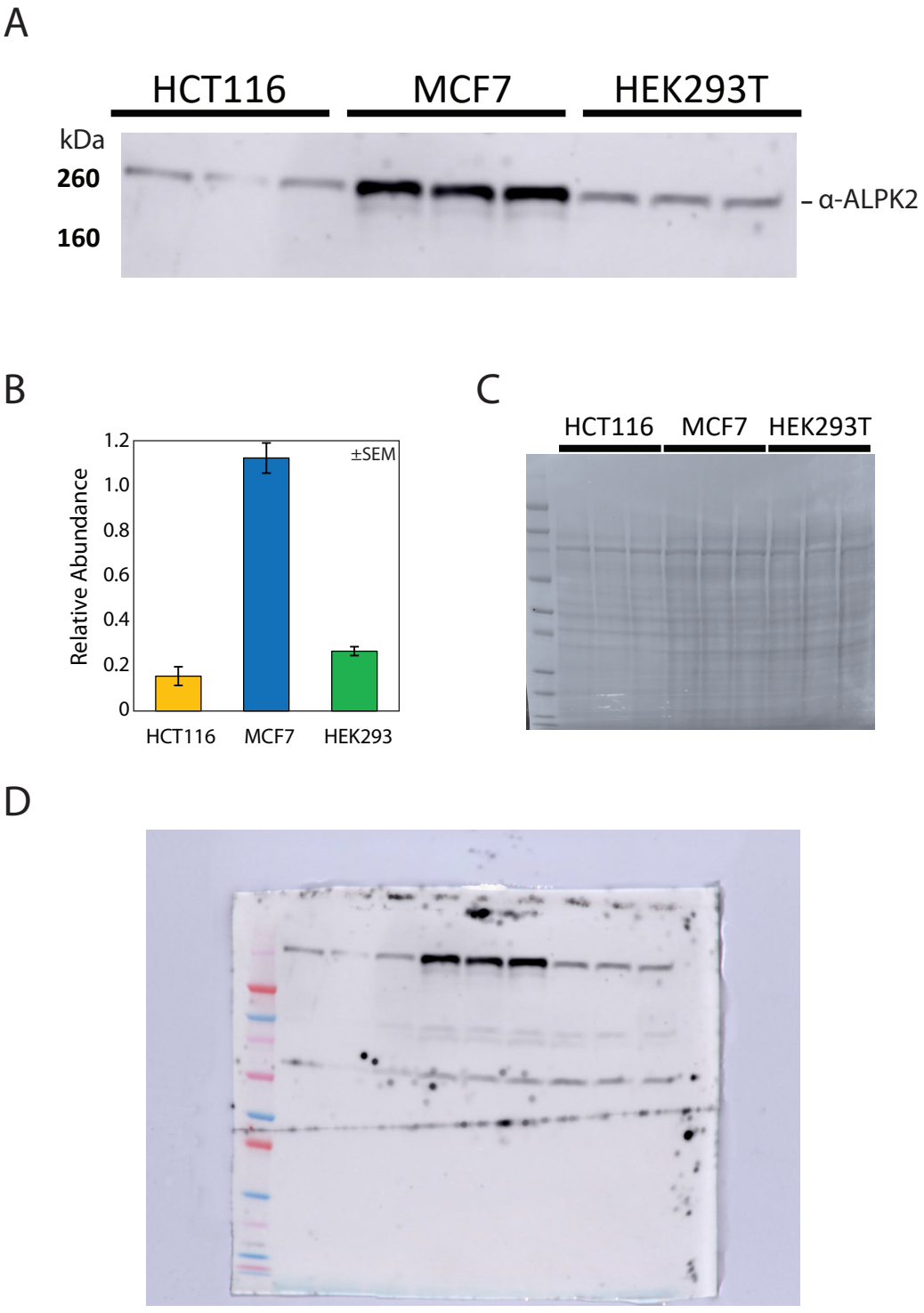


Figure S8
Western blotting analysis of ALPK2 abundance across three cell lines. **A.** Western blot for ALPK2 revealed strong expression in MCF7 cells as seen in the RTS-MS3 analysis but not the SPS-MS3 analysis (n=3). Molecular weights (kDa) are shown. **B.** Quantitation of relative abundance of ALPK2 across the cell lines in **A**. Error bars represent standard error of the mean (SEM) for the quantified lanes. **C.** Ponceau S staining of the PVDF membrane used in **A** demonstrating equal loading of each cell line. **D.** Image of the entire blotted membrane in **A**.