1 Supplementary Methods

2 LC-MS Data Analysis Exploris 480 MS (supplementary information)

3 The main report with peptide quantifications was used for downstream analysis in R (v4.1). Raw

4 precursor intensities were log-transformed and replicates of each spiked-in ratio were normalized

5 using global median log-intensities normalization to correct for the technical variation among

6 replicates. Ambiguous peptides shared by human and bovine proteins were removed from protein

7 quantification to avoid mixing the signals from the two organisms. Log-transformed normalized

8 precursor intensities were aggregated into protein group log-intensities using MaxLFQ (1) from R

9 package "iq" v1.9.6 (2).

10 We first assessed the precision performance of different workflows by calculating the coefficient

11 of variation (CV). That is, $CV = \frac{\sigma}{\mu}$ where σ is the standard deviation, and μ is the mean of protein

12 MaxLFQ intensities across the four replicates of each spike-in ratio.

13 We then evaluated the accuracy of the workflow by estimating how the measured protein fold

14 changes between pairs of spiked-in ratios deviate from the expected ones. We focused on bovine

15 proteins since their expected fold changes span from 1.2 to 100 (in comparison to 1.01–2 for human

16 proteins). The spiked-in ratios for bovine proteins (1, 0.5, 0.4, 0.3333, 0.1, 0.01; samples with no

17 bovine proteins were excluded) produce 15 pairwise comparisons with the following expected fold

18 changes: 1/0.5 (2), 1/0.4 (2.5), 1/0.3333 (3), 1/0.1 (10), 1/0.01(100), 0.5/0.4 (1.25), 0.5/0.3333

19 (1.5), 0.5/0.1 (5), 0.5/0.01 (50), 0.4/0.3333 (1.2), 0.4/0.1 (4), 0.4/0.01 (40), 0.3333/0.1 (3.33),

20 0.3333/0.01 (33.33), and 0.1/0.01 (10).

21

22 Compression of dynamic range at the NP-protein interface can lead to systematic compression or

23 inflation of the measured fold changes. To further improve quantification accuracy, one can correct

24 these systematic deviations. We performed a linear fit between the expected and measured log

25 fold-changes using the least square method for each precursor:

26 $$\log_2 \widetilde{FC}_i \sim \beta \log_2 FC_i + \varepsilon_i, \quad i = 1 \dots 15,$$

27 where $\log_2 FC_i$ and $\log_2 \widetilde{FC}_i$ represent the expected and measured $\log_2$ fold changes of the i-th

28 comparison, respectively. The estimated linear fit parameter $\hat{\beta}$ was then used to calculate the

29 corrected $\log_2$ fold change:

30 $$\log_2 \widehat{FC}_i = \frac{\log_2 \widetilde{FC}_i}{\hat{\beta}}$$

31 The errors between the expected and either measured or corrected fold-changes were calculated as

32 $\frac{1}{15}\sum_i |\log_2 FC_i - \log_2 \widetilde{FC}_i|$ and $\frac{1}{15}\sum_i |\log_2 FC_i - \log_2 \widehat{FC}_i|$, respectively **(Supplementary Figure**

33 **5)**.

34 Finally, we evaluated the quality of protein intensities by matrix-matched calibration curve

35 approach (3). For each analyte, we estimated the observed noise floor and linear intensity response

36 to concentration by curve fitting. Limits of detection (LoD) and limits of quantification (LoQ)

37 were estimated as the "spike-in ratio" (percentage of undiluted abundance) above which the

38 predicted intensity response exceeds the observed noise floor by two standard deviations, and the

39 concentration above which the coefficient of variation (CV) of intensity response (estimated by

40 bootstrapping) falls below 20%, respectively. We required the background noise to be estimated

41 from at least one concentration point (4 replicates), and the linear range to be estimated from at

42 least two concentration points (8 replicates). To account for large/small steps at the extremes of

43   our dilution range we employed a modified method that initially estimates noise from the lowest

44   two concentrations and linear response by regression on the remaining points. This initial fit was

45   refined by subsequent curve fitting and bootstrapping steps that matched the published

46   implementation from (3).

47   **Cohort Study (supplementary information)**

48   The sample size that allows for detecting protein regulation at a given threshold ($\log_2 FC$) with

49   given power (sensitivity) β, while also controlling the false discovery rate (FDR) (4,5) was

50   calculated as:

51

$$n = \frac{2(z(1-\alpha/2)+z(\beta))^2}{\left(\frac{\log_2 FC}{\sigma}\right)^2}, \text{ where } \alpha = \frac{\beta\, FDR}{1+\pi_0(1-FDR)},$$

52   σ is the median standard deviation of $\log_2$ intensities for all proteins, π_0 is the ratio of non-

53   regulated to regulated proteins in the data ($\pi_0 = 99$ was used), and $z(x)$ is the inverse cumulative

54   density function of the standard normal distribution. We calculated the required sample size once

55   without adjusting for known batch effects and setting σ to the standard deviation of median

56   normalized intensities for proteins present in 3 or more NP-specific samples, and once adjusting

57   for known batch effects such as plate and LC-MS instrument using the residual variance from the

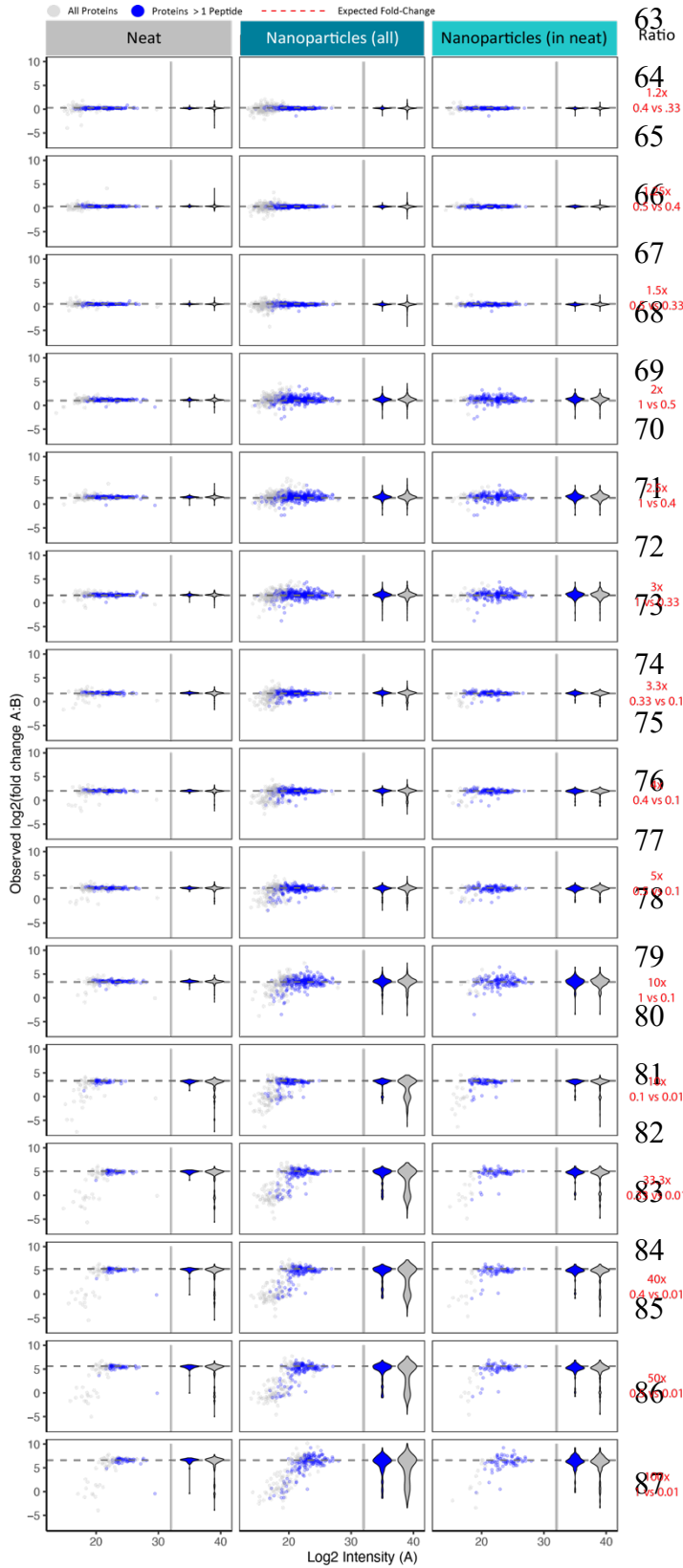58   following linear mixed model (in lme4 notation):

59          $Abundance \sim 1 + (1|Plate) + (1|Instrument)$.

60

61   **Supplementary Figures**

62

Supplementary Figure 1. Protein Quantification Accuracy of Neat Plasma Digestion and Proteograph Workflows. Spiked-in ratios for bovine proteins (1, 0.5, 0.4, 0.3333, 0.1, 0.01, and 0), producing 15 pairwise comparisons for small (A), medium (B), and large (C) fold changes. We skipped the ratio 0 since it had no bovine proteins. The pairwise comparisons were labeled as the expected fold-changes of the bovine proteins, i.e., 1 vs 0.5(2), 1 vs 0.4(2.5), 1 vs 0.3333(3), 1 vs 0.1(10), 1 vs 0.01(100), 0.5 vs 0.4(1.25), 0.5 vs 0.3333(1.5), 0.5 vs 0.1(5), 0.5 vs 0.01(50), 0.4 vs 0.3333(1.2), 0.4 vs 0.1(4), 0.4 vs 0.01(40), 0.3333 vs 0.1(3.33), 0.3333 vs 0.01(33.33), and 0.1 vs 0.01(10). X-axis denotes the log2 intensities. Y-axis is denoting fold changes of bovine proteins with dotted line indicating the expected fold change. For each comparison (neat, NP workflow all, shared), we illustrate all proteins as well (grey) as well as those quantified with more than 1 peptide (blue) as a scatter plot as well as violin plot. Data shown here are based on IP10.
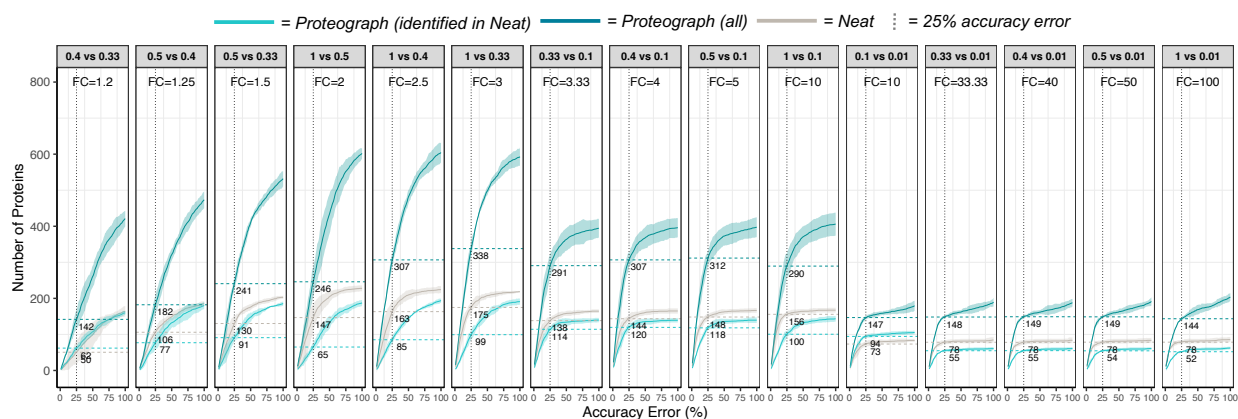
Supplementary Figure 2. Quantitative Accuracy Performance of Proteograph Workflow in Comparison to Neat Digestion Workflow. Each panel represents one fold-change of bovine proteins. X-axis is the % accuracy error, i.e., the difference between the observed and expected fold-change divided by the expected fold-change, and Y-axis is the number of bovine proteins identified at a given accuracy threshold. The horizontal dashed lines indicate bovine proteins reported at 25% threshold. Proteograph workflow with proteins identified in neat workflow is colored in light teal, Proteograph workflow with all proteins is colored in dark teal, and neat digestion workflow is colored in grey. Proteograph workflow with all proteins has demonstrated higher protein identification than neat digestion workflow with certain accuracy. Data shown here are based on IP10.

98
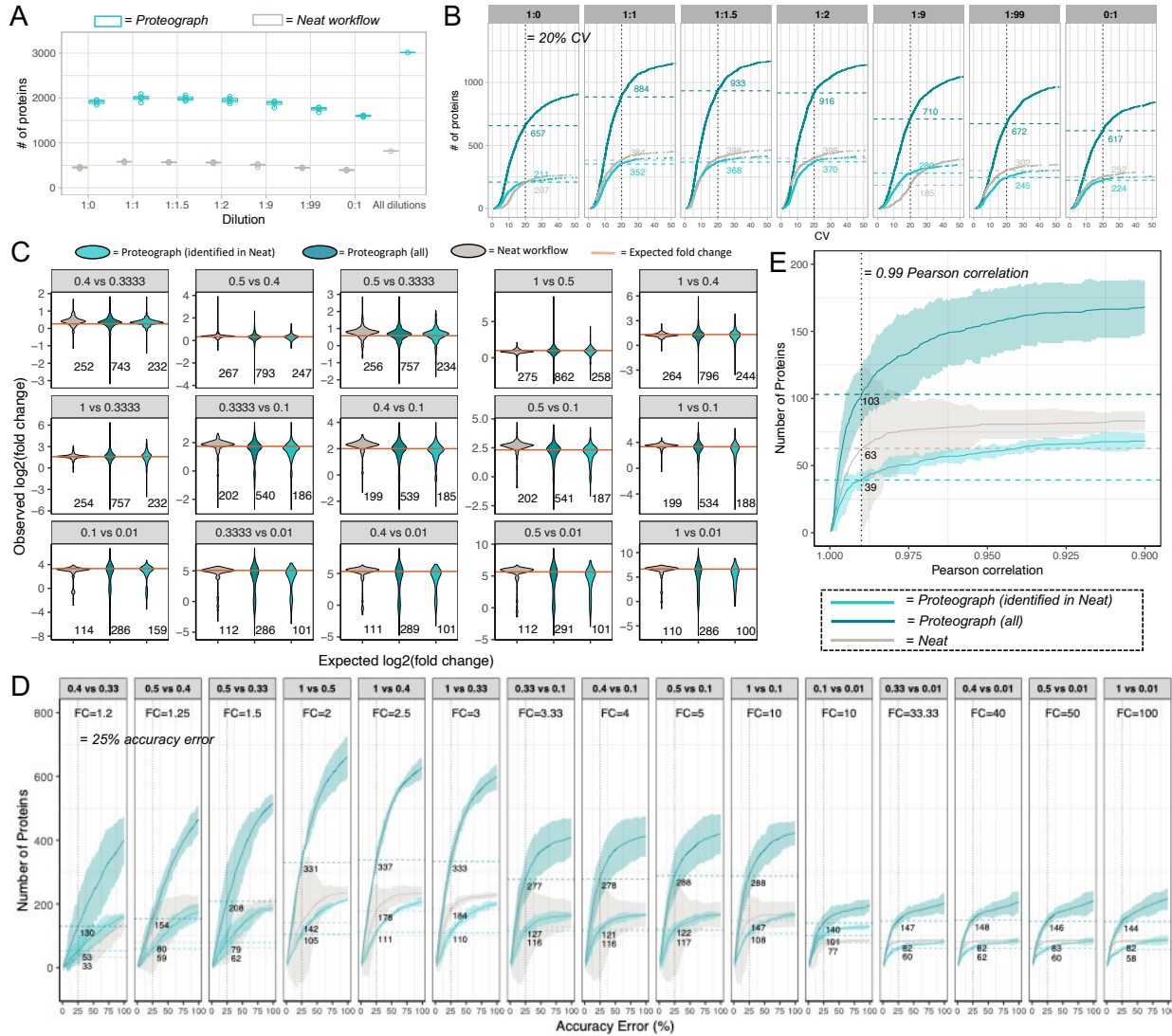
Supplementary Figure 3. **Quantification Performance of Neat and Proteograph Workflows for Another Human Plasma Sample PC6.** All the data processing and statistical analysis steps are the same as those on the previous IP10 sample. (A) The number of proteins quantified at each ratio. Proteograph workflow is colored in teal and neat workflow is colored in grey. (B) Number of Proteins Identified at a given CV Threshold for each Spiked-in ratio. X-axis is the CV calculated across four replicates, and the Y-axis is the number of proteins with a CV lower than the given threshold. NP-workflow with proteins identified in neat workflow is colored in light teal, Proteograph workflow with all proteins is colored in dark teal, and neat workflow is colored in grey. Only proteins quantified in all four replicates are counted. (C) Fold change accuracy of Neat and NP- Workflows. X-axis is the 15 comparisons and y-axis is observed fold changes. The orange dots connected by the orange line, as well as the orange numbers,

108   indicate the expected fold changes of bovine proteins. The distribution of observed fold changes is shown by each box

109   plot. The barplots at the top panel show the corresponding number of proteins summarized by each box. (D) Number

110   of Proteins Identified at a given accuracy error for each expected fold change. Each panel represents one fold change.

111   X-axis is the % accuracy error, i.e., the difference between the observed and expected protein fold change divided by

112   the expected fold change, and Y-axis is the number of proteins identified at a given accuracy threshold. The horizontal

113   dashed lines indicate proteins reported at 25% threshold. Proteograph workflow with all proteins has demonstrated

114   higher protein identification than neat digestion workflow with certain accuracy.

115

116
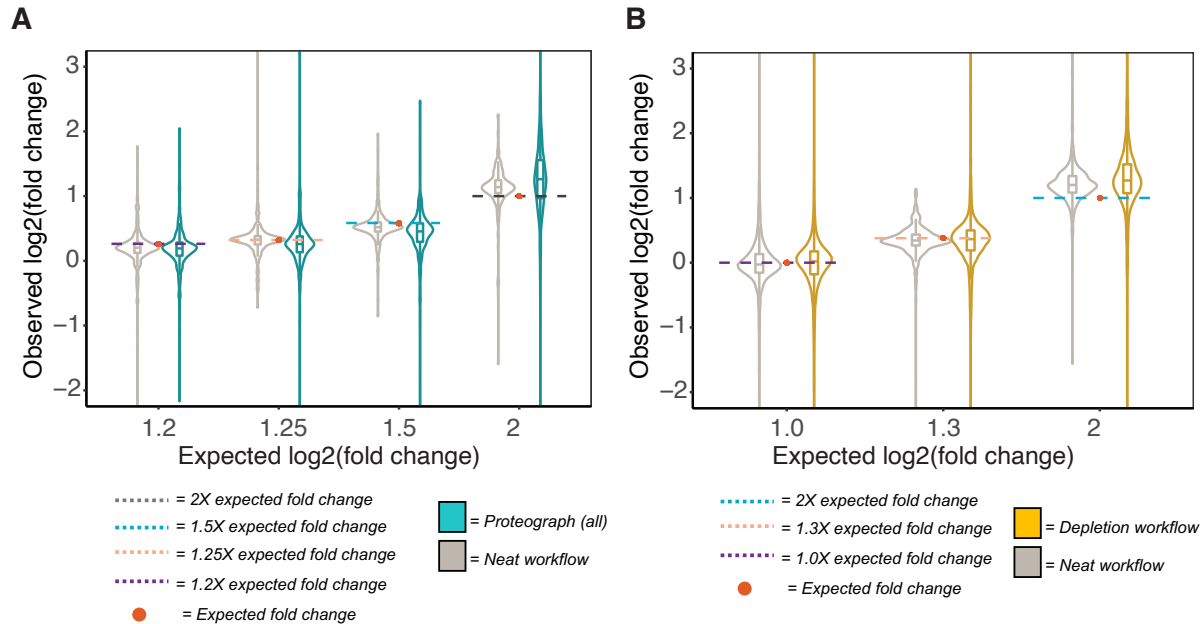
117

118   **Supplementary Figure 4. Quantification Accuracy Across Lower Fold Change Ranges.** (A) Proteograph

119   workflow quantitation accuracy across 1.2-2 fold-change. Quantification data produced by NP-workflow shows

120   comparable accuracy with those from neat workflow at low fold change of 1.2-2 fold. Data shown here are based on

121   IP10. (B) The performance of depletion workflow is compared with neat digestion workflow here across similar low

122   fold changes of 1.0-2 fold-change (6). Although this data is generated on different plasma sample and LC-MS

123   workflow, overall we observe similar accuracies with Proteograph workflow.

124

125
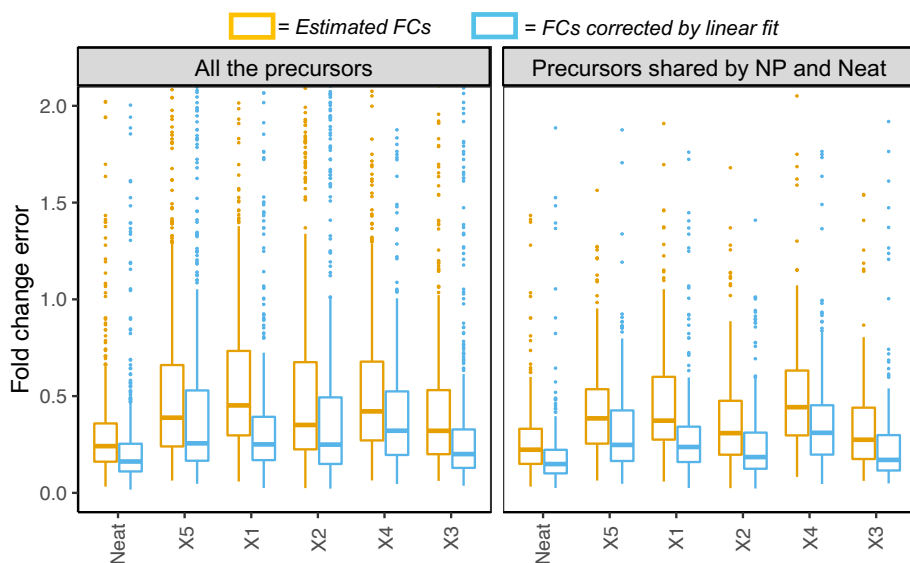
Supplementary Figure 5. Quantification Accuracy Correcting for Systematic Shifts in Quantification. For each precursor, we performed a linear fit correction between the measured log FC and the expected log FC of bovine precursors. The fold change error (y-axis) was then calculated as the mean absolute difference between the measured or corrected log FCs and the expected log FCs. The x-axis is the Neat workflow and the 5 NPs used by Proteograph workflow. After correction by linear fit, the fold change error can be significantly reduced.

131

132

133 **Supplementary References**

134

135 1. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M. Accurate
136 Proteome-wide Label-free Quantification by Delayed Normalization and
137 Maximal Peptide Ratio Extraction, Termed MaxLFQ*. Mol Cell
138 Proteomics. 2014;13:2513–26.

139 2. Pham TV, Henneman AA, Jimenez CR. iq: an R package to estimate
140 relative protein abundances from ion quantification in DIA-MS-based
141 proteomics. Bioinformatics. 2020;36:2611–3.

142 3. Pino LK, Searle BC, Yang H-Y, Hoofnagle AN, Noble WS, MacCoss MJ.
143 Matrix-Matched Calibration Curves for Assessing Analytical Figures of
144 Merit in Quantitative Proteomics. J Proteome Res. 2020;19:1147–53.

145 4. Jung S-H. Sample size for FDR-control in microarray data analysis.
146 Bioinformatics. 2005;21:3097–104.

147 5. Lachin JM. Introduction to sample size determination and power
148 analysis for clinical trials. Control Clin Trials. 1981;2:93–113.

149 6. Tognetti M, Sklodowski K, Müller S, Kamber D, Muntel J, Bruderer R,
150 et al. Biomarker Candidates for Tumors Identified from Deep-Profiled
151 Plasma Stem Predominantly from the Low Abundant Area. J Proteome
152 Res. 2022;21:1718–35.

153

154