HYPERsol: high-quality data from archival FFPE tissue for clinical proteomics

Dylan M. Marchione[1], Ilyana Ilieva[2], Kyle Devins[2], Danielle Sharpe[2], Darryl J. Pappin[3,4], Benjamin A. Garcia[1], John P. Wilson[4*], and John B. Wojcik[2*†]

1. Epigenetics Institute, Department of Biochemistry & Biophysics, University of Pennsylvania, Philadelphia PA
2. Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia PA
3. Cold Spring Harbor Laboratory, Cold Spring Harbor NY
4. ProtiFi, LLC, Huntington NY john@protifi.com

*Corresponding authors (jbwojcik5@gmail.com, john@protifi.com)

†Current address: Bristol-Myers Squibb, Lawrenceville NJ

Table of Contents

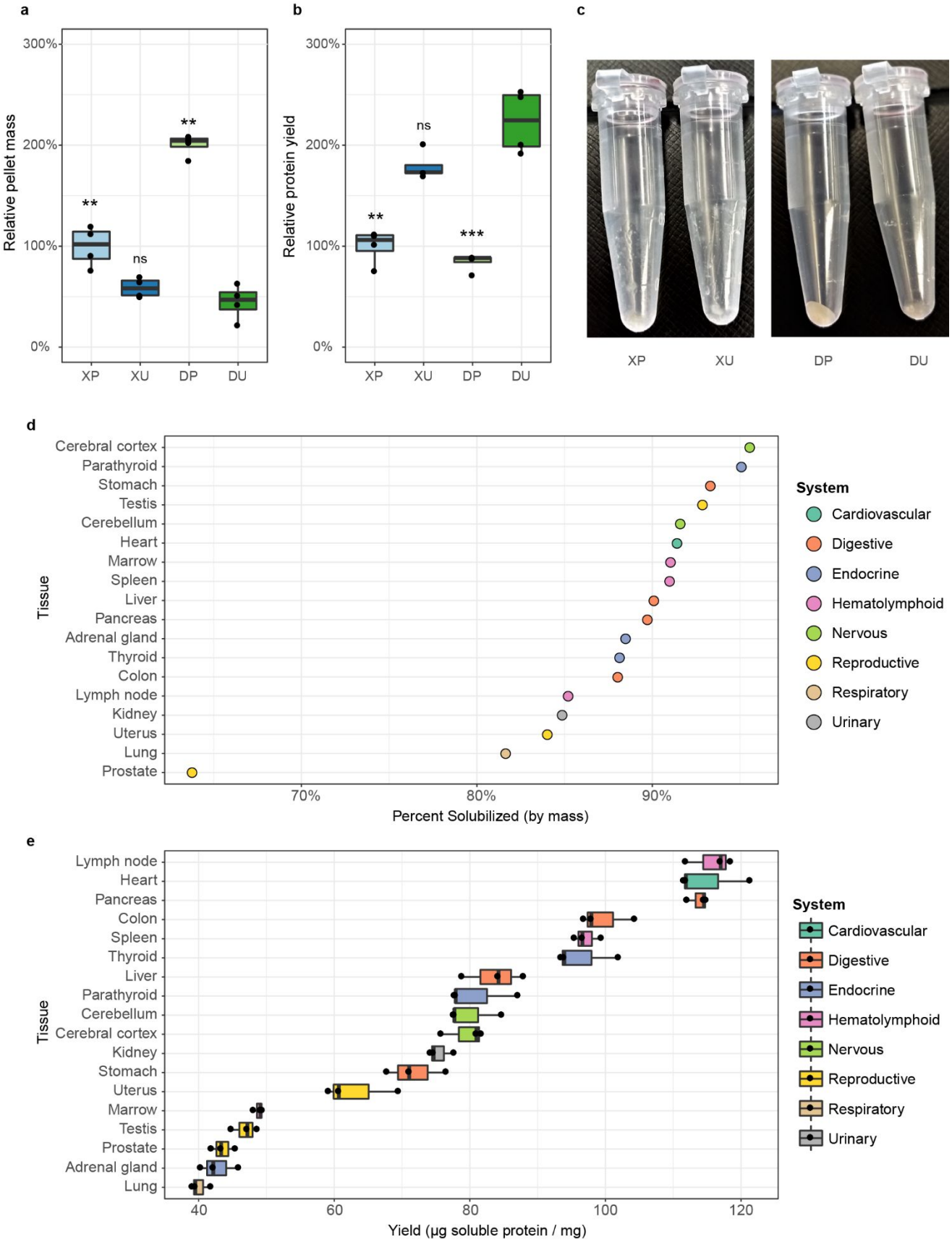
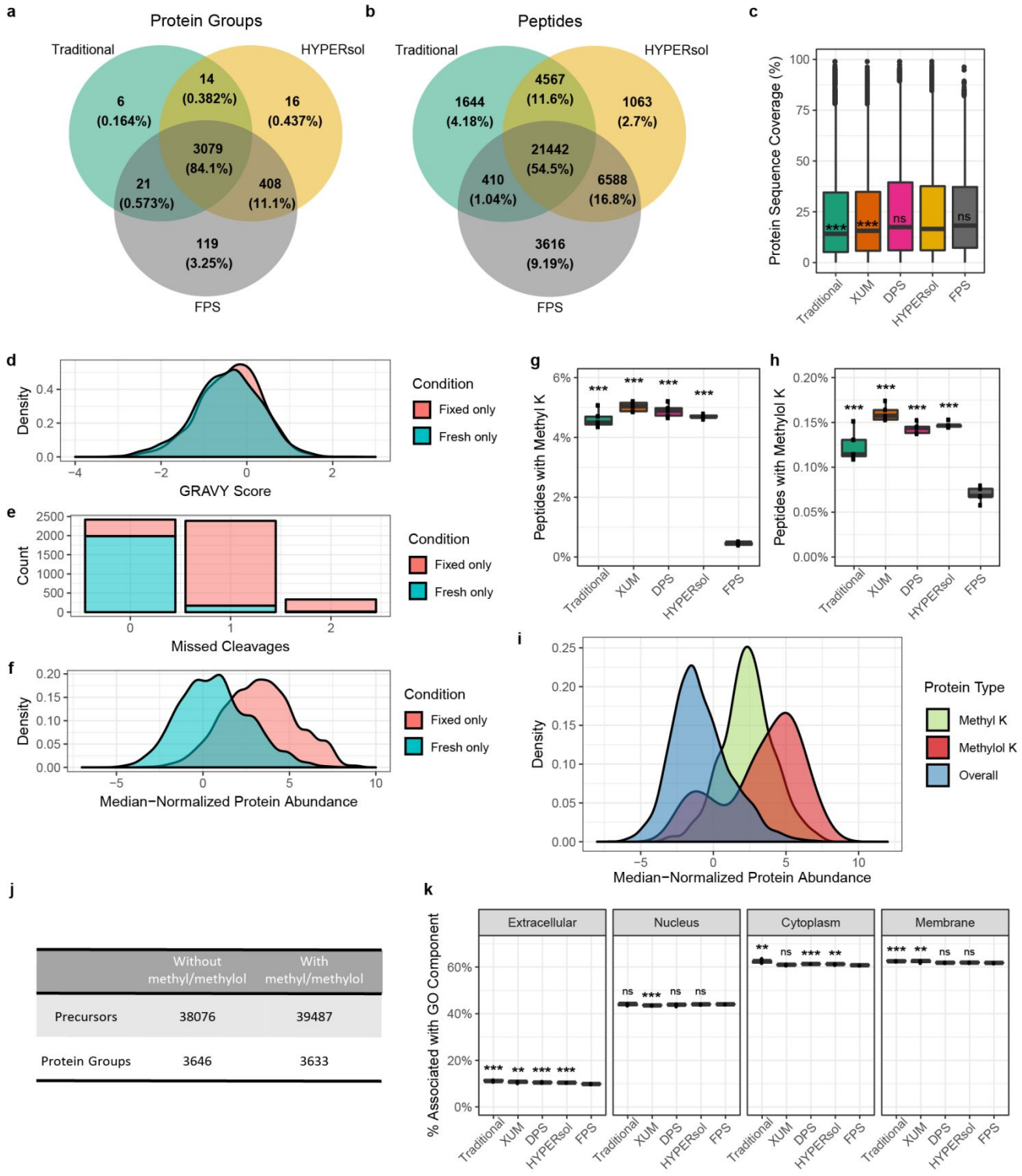

a
b
c
Relative pellet mass
Relative protein yield
300%
200%
100%
0%
**
ns
***
XP
XU
DP
DU
d
Tissue
Cerebral cortex
Parathyroid
Stomach
Testis
Cerebellum
Heart
Marrow
Spleen
Liver
Pancreas
Adrenal gland
Thyroid
Colon
Lymph node
Kidney
Uterus
Lung
Prostate
70%
80%
90%
Percent Solubilized (by mass)
System
Cardiovascular
Digestive
Endocrine
Hematolymphoid
Nervous
Reproductive
Respiratory
Urinary
e
Yield (µg soluble protein / mg)
40
60
80
100
120

Figure S1. Direct solubilization combined with ultrasonication efficiently solubilizes FFPE samples. a). Bar graph of relative residual pellet mass, normalized to initial pellet mass. XP: xylene-ethanol, probe. XU: xylene-ethanol, ultrasonication. DP: direct, probe. DU: direct, ultrasonication. b). Bar graph of relative protein yields, normalized to initial pellet mass. For bar graphs, means ± standard deviations are shown (n = 4), and asterisks indicate statistical significance compared to HYPERsol with Welch's two-tailed t-test and $p < 0.05$ = *, $p < 0.01$ = **, and $p < 0.001$ = ***. c). Representative images of residual material after each combination of conditions. d). Scatter plot depicting the extent of solubilization of 18 FFPE human tissue samples (n = 1). e) Scatter plot depicting protein yield per milligram of FFPE across 18 human tissue samples (n = 1). Error bars depict the standard deviation from 3 technical replicates of the BCA assay.



	Without methyl/methylol	With methyl/methylol
Precursors	38076	39487
Protein Groups	3646	3633

Figure S2. Peptides from FFPE samples contain missed cleavages and modifications. a) Venn diagram illustrating the overlap among detected proteins in Traditional, HYPERsol, and FPS. b) Venn diagram illustrating the overlap among detected proteins in Traditional, HYPERsol, and FPS. c) Tukey boxplot depicting average sequence coverage across conditions. d-f) Density plots depicting grand average of hydropathy (GRAVY) scores (d), missed cleavage counts (e), and median-normalized protein abundance (f) associated with peptides that were

unique to either FFPE samples (pink) or flash-frozen samples (blue). g-h) Tukey boxplot depicting the fraction of peptides containing methyl lysine (g) or methylol lysine (h) across experimental conditions. i) Density plot depicting median-normalized protein abundance of methylated proteins (green), or methylolated proteins (red), relative to the entire dataset (blue). j) Table illustrating the effect of including methyl and methylol lysine as variable modifications on spectral library size in Spectronaut. k) Tukey boxplot depicting the fraction of identified proteins associated with each Gene Ontology (GO) Component term. The conditions were: Traditional: xylene-ethanol, probe, methanol-chloroform. XUM: xylene-ethanol, ultrasonication, methanol-chloroform. DPS: direct, probe, S-Trap. HYPERsol: direct, ultrasonication, S-Trap. FPS: flash-frozen, probe, S-Trap. For box plots, $n = 5$. In panel c) asterisks indicate statistical significance when compared to HYPERsol with Welch's two-tailed t-test and $p < 0.05$ = *, $p < 0.01$ = **, and $p < 0.001$ = ***. In panel k), conditions were compared against FPS rather than HYPERsol.