Fully automated sample processing and analysis workflow for low-input proteome profiling

Yiran Liang[1], Hayden Acor[1], Michaela A. McCown[2], Andikan J. Nwosu[1], Hannah Boekweg,[2] Nathaniel B. Axtell[1], Thy Truong[1], Yongzheng Cong[1], Samuel H. Payne[2], Ryan T. Kelly[1]*

1. Department of Chemistry and Biochemistry, Brigham Young University, Provo, UT, 84602

2. Department of Biology, Brigham Young University, Provo, UT, 84602

* Email: ryan.kelly@byu.edu

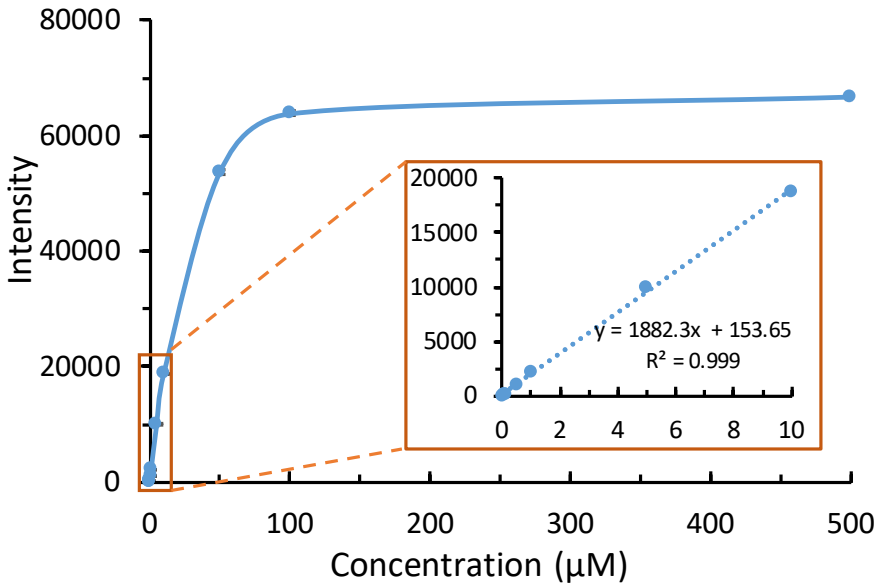

Figure S1: Calibration curve for 10 µL fluorescein samples using Synergy H4 Hybrid plate reader. R^2 = 0.999 for 0–10 µM.

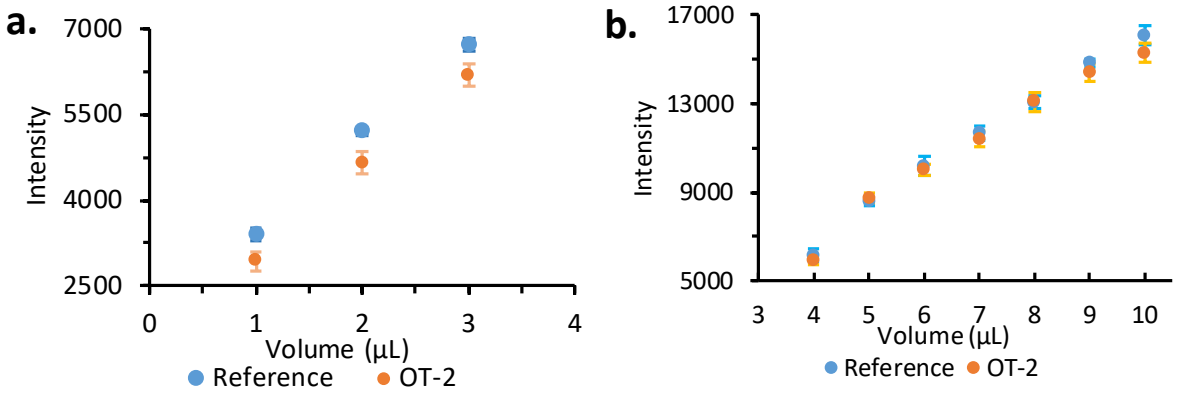

Figure S2: Accuracy and reproducibility of OT-2 pipetting compared with manual pipetting: (a) 1–3 µL, (b) 4–10 µL. Error bars represent ±1 standard deviation for 12 replicates. Corresponding data are shown in Table S1.

Table S1: Reproducibility and accuracy of OT-2 robot pipetting from 1 to 10 µL. For accuracy, the manual pipetting was treated as the reference (100%). 12 replicates were obtained for each volume. This result is similar with the manufacturer specifications. [1]

Volume (µL)		10	9	8	7	6	5	4	3	2	1
CV (%)	Manual	2.6	1.2	2.1	2.8	4.3	2.0	4.4			
	OT-2	2.8	2.8	3.2	3.3	2.8	2.4	2.3	3.3	4.3	5.6
Accuracy (%)	OT-2	95.1	97.4	100.3	98.1	98.3	102	95.7	92.1	89.3	85.9

Table S2: The effect of pipetting speed and changing tips. 10 µL of 10 µM NaFl was used as sample.

		Intensity			
	Pipetting speed*	Average	STDEV	CV (%)	Accuracy (%)
Single tip	Slow	17699	703	4.0	101.8
	Fast	17405	277	1.6	100.1
	Manual	17385	195	1.1	
Change tip	Slow	17002	366	2.1	97.8
	Fast	16762	450	2.7	96.4

“Single tip”: all the replicates were pipetted using only one tip.

“Change tip”: change pipette tip for every replicate.

“Slow”: the flowrate of aspirating and dispensing were set as 50 µL/s.

“Fast”: default setting was used. The flowrate of aspirating was 150 µL/s and the flowrate of dispensing was 300 µL/s.

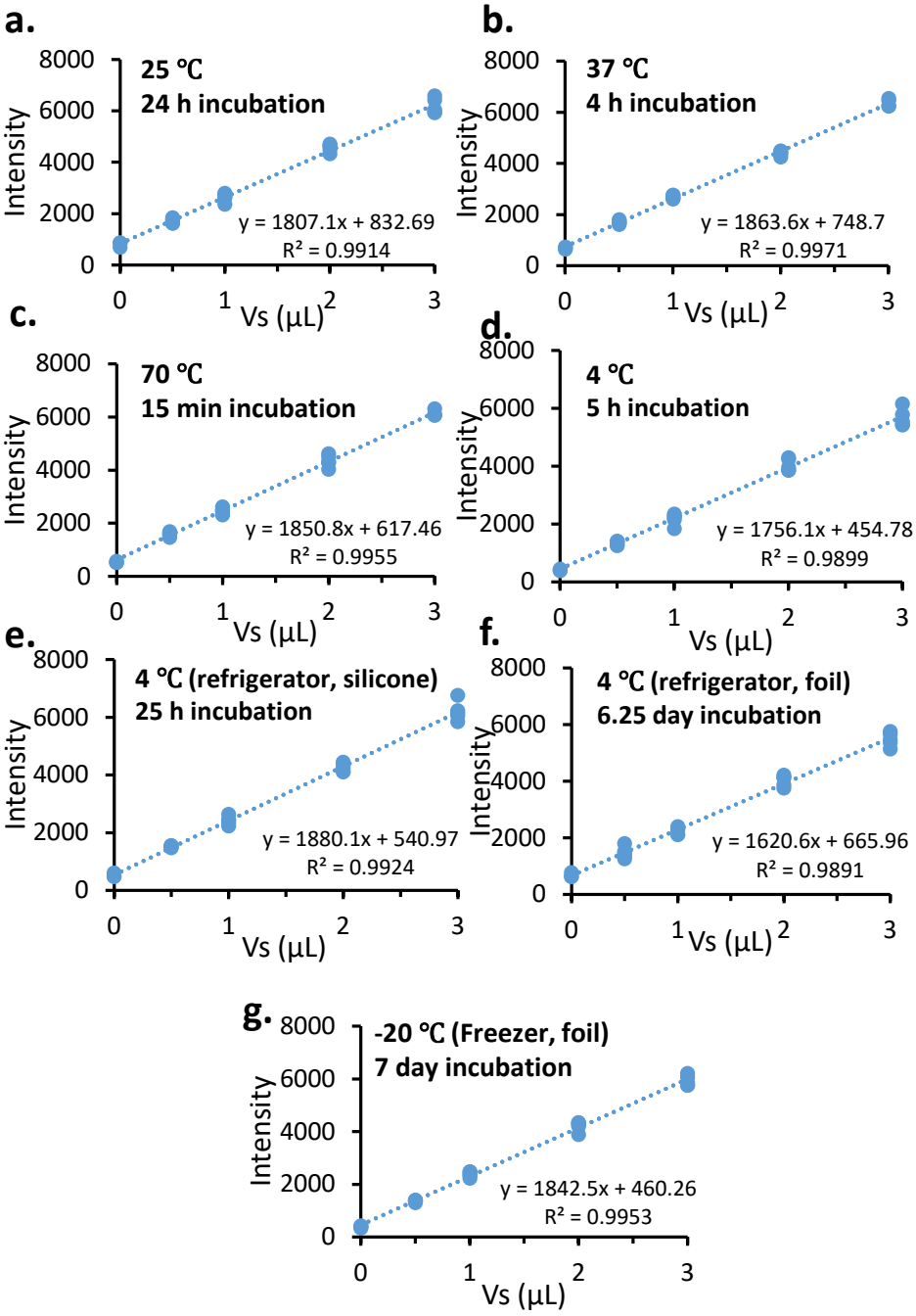

Figure S3: Determination of evaporation rates from well plates by standard addition at: (a-d) various temperatures using the OT-2 temperature control module and silicone sealing mat; (e) 4 ℃ in a refrigerator using a silicone sealing mat; (f) 4 ℃ in a refrigerator using an adhesive foil sealing mat; (g) -20 ℃ in a freezer using an adhesive foil sealing mat.

Table S3: The number of cells sampled and corresponding peptide and protein IDs in each replicates of ~150 and ~500 HeLa cells. The cells were counted from the imaged taken using a microscope.

			MSFragger		MaxQuant	
	Cell Counts	File Name	Peptide ID	Protein ID	Peptide ID	Protein ID
	141	HeLa_150cell_I12	20441	2715	15722	2260
	137	HeLa_150cell_I16	22550	2791	17990	2482
	160	HeLa_150cell_I18	18936	2532	13471	2083
Average	146		20642	2679	15728	2275
	466	HeLa_500cell_J5	36471	3307	23385	3446
	520	HeLa_500cell_K16	31416	3280	28885	3119
	540	HeLa_500cell_K8	35700	3455	27424	3314
Average	509		34529	3347	26565	3293

Table S4: The accuracy of FAC-sorting. Fluorescent beads were used with a target of 145.

	Replicates										Average	STDEV	CV(%)	Accuracy
Counts	133	125	126	130	134	138	135	127	124	121	129.3	5.5	4.3	89%

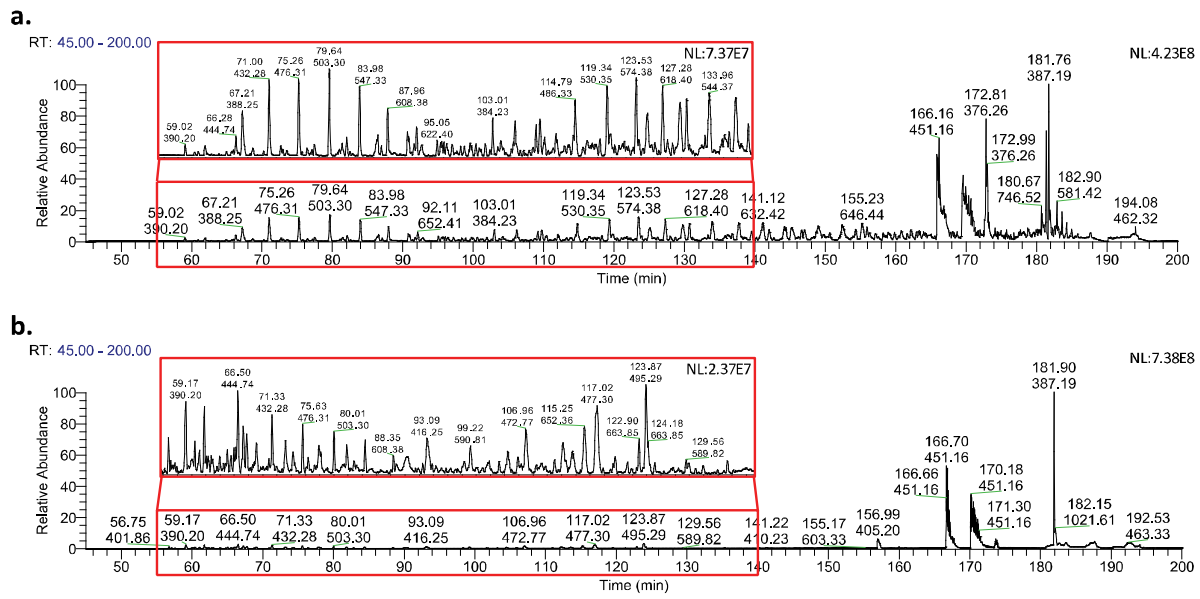

Figure S4: Chromatograms of 2 ng of HeLa protein digest standard. The HeLa digest solution was stored in (a) a fresh 384 well plate and (b) a pre-washed well plate.

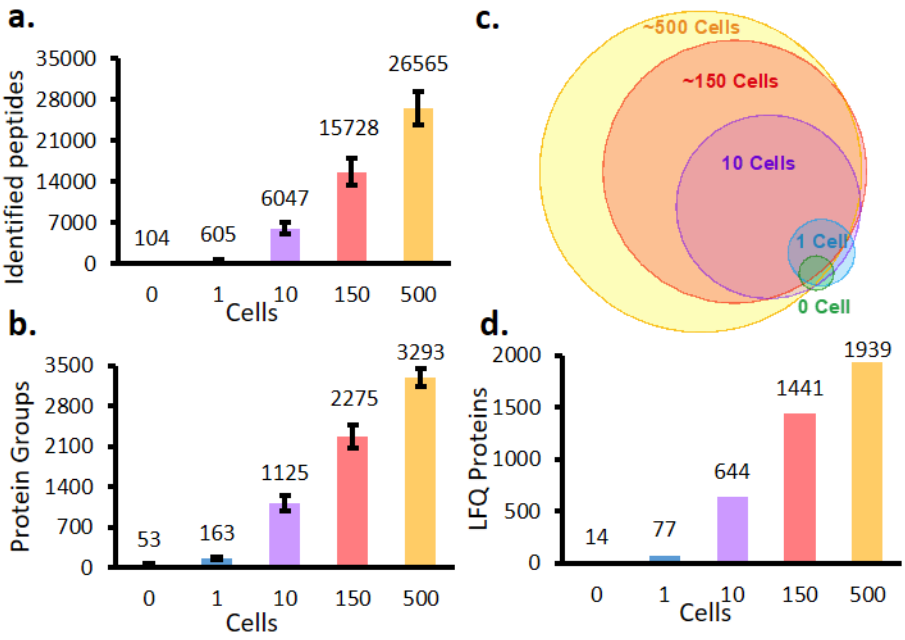

Figure S5: Number of peptides (a) and proteins (b) identified by MS/MS from 0 (supernatant from cell suspension), 1, 10, ~150, and ~500 HeLa cells using MaxQuant. (c) Overlap of protein groups identified in 2 of 3 replicates from various sample loadings. (d) Label free quantifiable proteins based on identification in 2 of 3 replicates and at least 2 peptides per protein group.

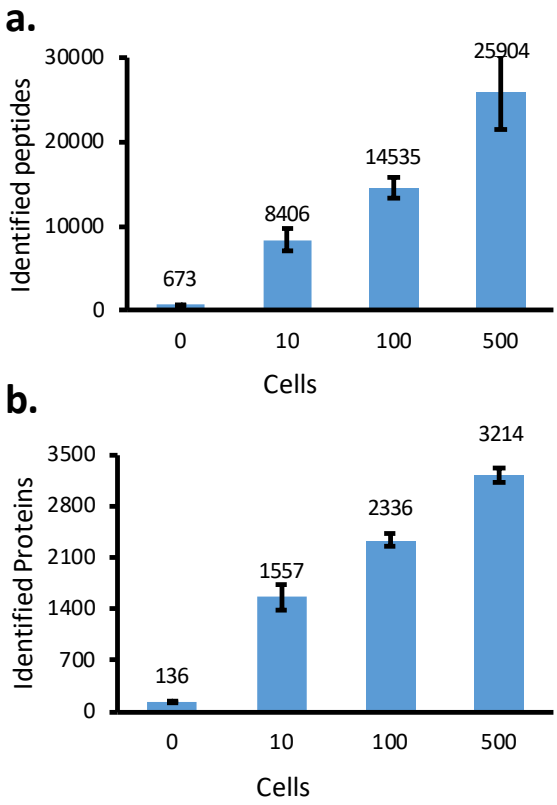

Figure S6: MS/MS identified peptides (a) and protein groups (d) from 0, 10, 100 and 500 HeLa cells isolated by FACS. For blanks, a cell-free droplet was sorted and dispensed into the well. Actual cell numbers are likely lower given the measured deposition efficiency of ~90% for this FACS system. Analysis were performed using MSFragger without MBR, and identified proteins are listed in Supplemental Table E1.

Table S5: Gene set enrichment analysis using a cell-type classifier ProteomicsDB[2] of identified proteins from B and T lymphocytes.

Cell Type	Cell Line ID	Adjusted P-value
Lymphoblastoid	BTO:0000773 X129.126 HM11.GM18552	1.12E-80
Hematopoietic MOLT-4	BTO:0000873 P003361	7.96E-72
Breast HCC-202	PDB:200003 HCC202 R2-1	9.93E-72
Breast MCF-7	BTO:0000093 HCC202 R2-2	8.83E-72
Hematopoietic JURKAT	BTO:0000661 Jurkat R2	8.96E-65
Hematopoietic PBMC	BTO:0001008 81719202114 replicate01-4	8.18E-63
Bone U2-OS	BTO:0001938 CYTO BioRep1 heavy	2.88E-60
Hematopoietic CCRF-CEM	BTO:0000736 P003197	1.28E-57

Table S6: Gene set enrichment using Gene Ontology Molecular Function (GOMF) annotations of identified proteins from B and T lymphocytes.

Molecular function	GO term	Adjusted P-value
RNA binding	(GO:0003723)	9.41E-148
cadherin binding	(GO:0045296)	3.39E-52
mRNA binding	(GO:0003729)	6.96E-19
actin filament binding	(GO:0051015)	4.79E-14
translation factor activity, RNA binding	(GO:0008135)	2.10E-13
actin binding	(GO:0003779)	4.92E-13
nucleoside-triphosphatase activity	(GO:0017111)	9.24E-13
purine ribonucleoside triphosphate binding	(GO:0035639)	3.34E-12
translation initiation factor activity	(GO:0003743)	4.10E-11
NADH dehydrogenase (ubiquinone) activity	(GO:0008137)	2.65E-10
NADH dehydrogenase (quinone) activity	(GO:0050136)	2.41E-10
double-stranded RNA binding	(GO:0003725)	6.38E-09
protein homodimerization activity	(GO:0042803)	6.08E-09
purine ribonucleoside binding	(GO:0032550)	6.82E-09
GTP binding	(GO:0005525)	3.17E-08
guanyl ribonucleotide binding	(GO:0032561)	5.20E-08
peptidase activity, acting on L-amino acid peptides	(GO:0070011)	2.03E-07
MHC class II protein complex binding	(GO:0023026)	1.99E-07
exopeptidase activity	(GO:0008238)	2.93E-07
snRNA binding	(GO:0017069)	3.19E-07

Table S7: Gene set enrichment using Wikipathways annotations of identified proteins from B and T lymphocytes.

Pathway	Gene or Protein	Adjusted P-value
Cytoplasmic Ribosomal Proteins W	P477	9.65E-67
mRNA Processing W	P411	2.60E-38
Proteasome Degradation W	P183	5.41E-32
Electron Transport Chain (OXPHOS system in mitochondria) W	P111	1.92E-27
Metabolic reprogramming in colon cancer WP	4290	7.24E-25
Translation Factors W	P107	9.78E-17
Parkin-Ubiquitin Proteasomal System pathway WP	2359	3.18E-16
Oxidative phosphorylation W	P623	1.12E-15
Nonalcoholic fatty liver disease WP	4396	9.05E-15
TCA Cycle (aka Krebs or citric acid cycle)	WP78	2.27E-14
Glycolysis and Gluconeogenesis W	P534	2.50E-12
VEGFA-VEGFR2 Signaling Pathway WP	3888	8.07E-12
Pathogenic Escherichia coli infection WP	2272	5.00E-11
B Cell Receptor Signaling Pathway	WP23	9.12E-11
TCA Cycle and Deficiency of Pyruvate Dehydrogenase complex (PDHc) WP	2453	9.63E-11
Amino Acid metabolism WP	3925	1.15E-08
Ciliary landscape WP	4352	3.65E-08
Ebola Virus Pathway on Host WP	4217	6.06E-08
Regulation of Actin Cytoskeleton	WP51	7.12E-08
Prolactin Signaling Pathway WP	2037	9.14E-08

Reference

(1) Opentrons. Opentrons Electronic Pipettes. https://opentrons.com/publications/Opentrons-Pipette-White-Paper.pdf (accessed Aug 7, 2020).

(2) Schmidt, T.; Samaras, P.; Frejno, M.; Gessulat, S.; Barnert, M.; Kienegger, H.; Krcmar, H.; Schlegl, J.; Ehrlich, H. C.; Aiche, S.; et al. ProteomicsDB. *Nucleic Acids Res.* **2018**, *46* (D1), D1271–D1281. https://doi.org/10.1093/nar/gkx1029.