Cell Reports Methods, Volume *4*

Supplemental information

Droplet-based proteomics reveals CD36 as a marker for progenitors in mammary basal epithelium

Matthew Waas, Amanda Khoo, Pirashaanthy Tharmapalan, Curtis W. McCloskey, Meinusha Govindarajan, Bowen Zhang, Shahbaz Khan, Paul D. Waterhouse, Rama Khokha, and Thomas Kislinger

Supplemental Figures

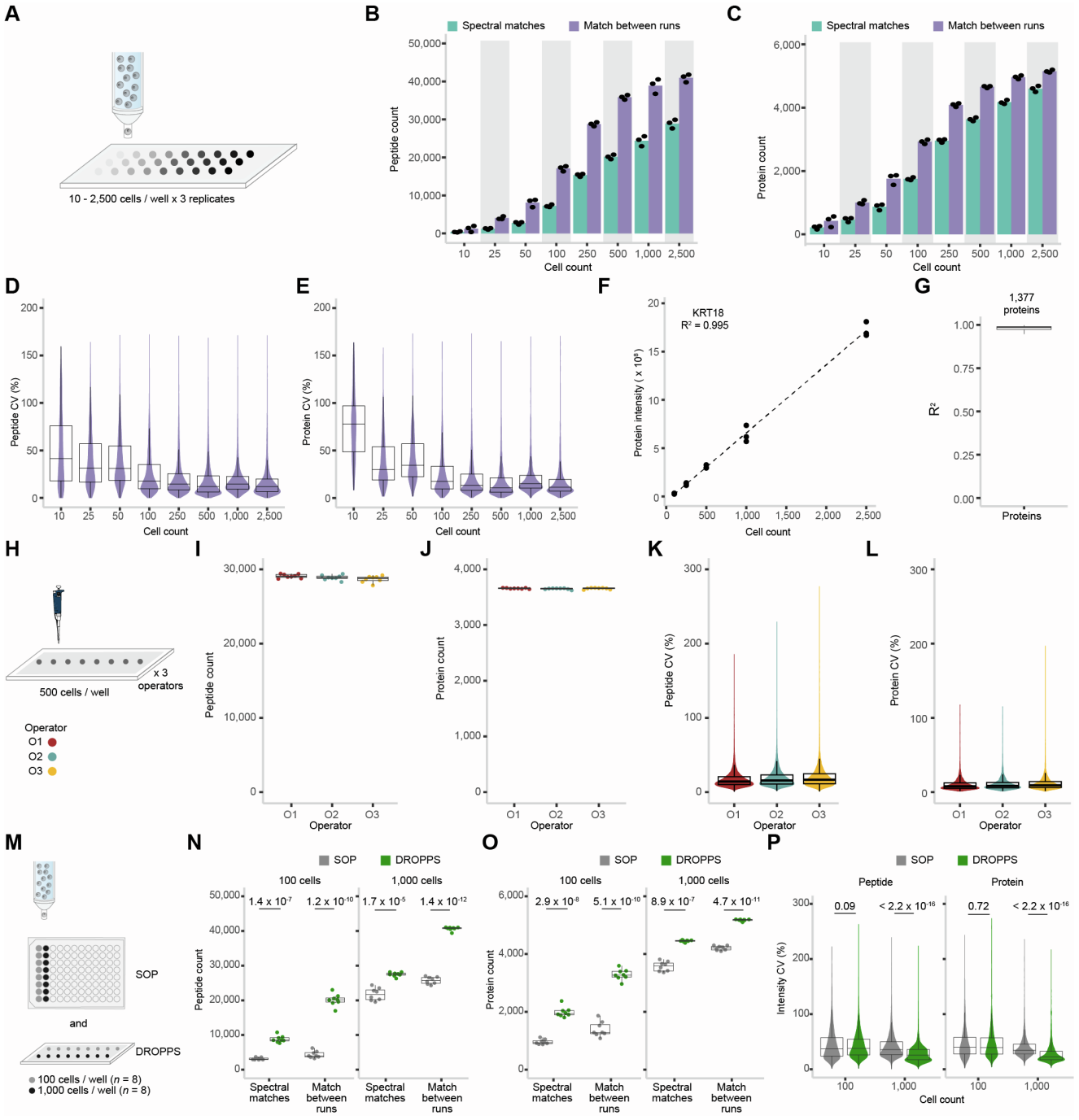

Figure S1. Evaluation of DROPPS performance across a range of starting cell numbers and among operators, related to Figure 1. (A) Cell number titration workflow where 10 – 2,500 MCF10A cells were deposited per well (n = 3 each) by fluorescent-activated cell sorting and subsequently processed using DROPPS. (B,C) The (B) peptide counts and (C) protein counts of the individual runs with and without "match between runs" where columns represent the mean values. (D,E) Distribution of intensity CVs at the (D) peptide and (E) protein level. (F) KRT18 protein intensity as a function of input cell number. (G) Distribution of R^2 calculated from modeling the protein intensity as a function of cell number. (H) Schematic of operator experimental design where 500 MCF10A cells were deposited per well (n = 8 each) by three operators and subsequently processed using Dropps. (I,J) The (I) peptide counts and (J) protein counts of the individual samples prepared by each operator. (K,L) Distribution of (K) peptide intensity or (L) protein intensity CVs. (M) Schematic of comparison between SOP-MS[19] and DROPPS (n = 8 for each cell number for each method). (N,O) The (N) peptide counts and (O) protein counts of the individual runs with and without "match between runs" Welch's t-test p-values are shown. (P) Distribution of intensity CVs at the peptide and protein level. Wilcoxon rank sum test p-values are shown. Boxplots show the median, interquartile ranges, and 95% confidence interval estimate.

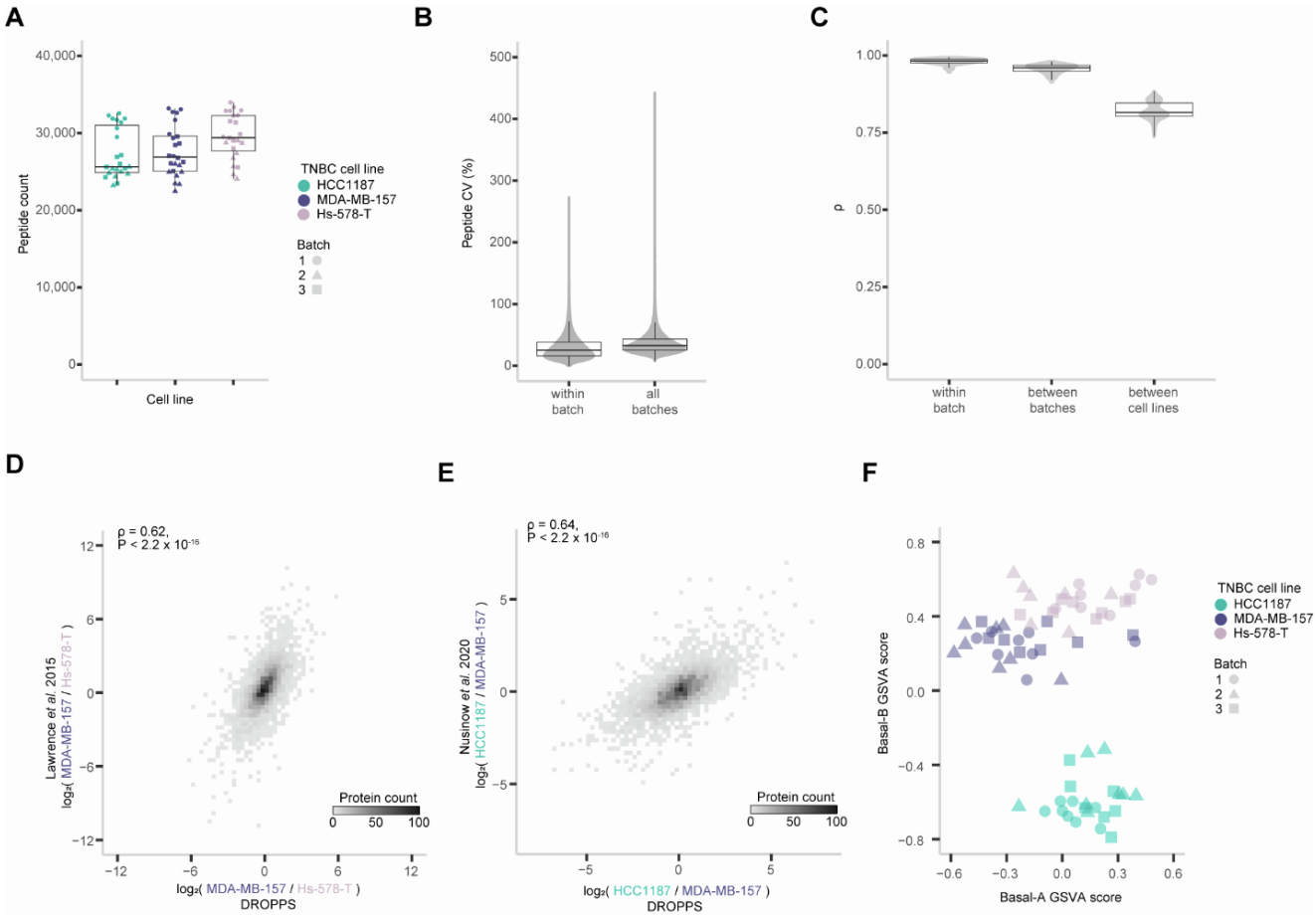

Figure S2. Evaluation of DROPPS for profiling TNBC cell lines differences across batches, **related to Figure 2.** (A) The peptide counts of the individual runs. (B) Violin plots of peptide intensity CVs for runs within a batch or for combining batches. (C) Spearman's ρ between runs of the same cell line within batch, runs of the same cell line between batches, and between all other runs. (D,E) $\log_2$ fold-changes between cell lines, using data acquired by DROPPS compared to data acquired by (D) Lawrence *et al.*[14] and (E) Nusinow *et al*.[13] shown with Spearman's ρ. (F) GSVA values for Basal A and Basal B TNBC subtypes. Boxplots show the median, interquartile ranges, and 95% confidence interval estimate. TNBC: triple negative breast cancer.

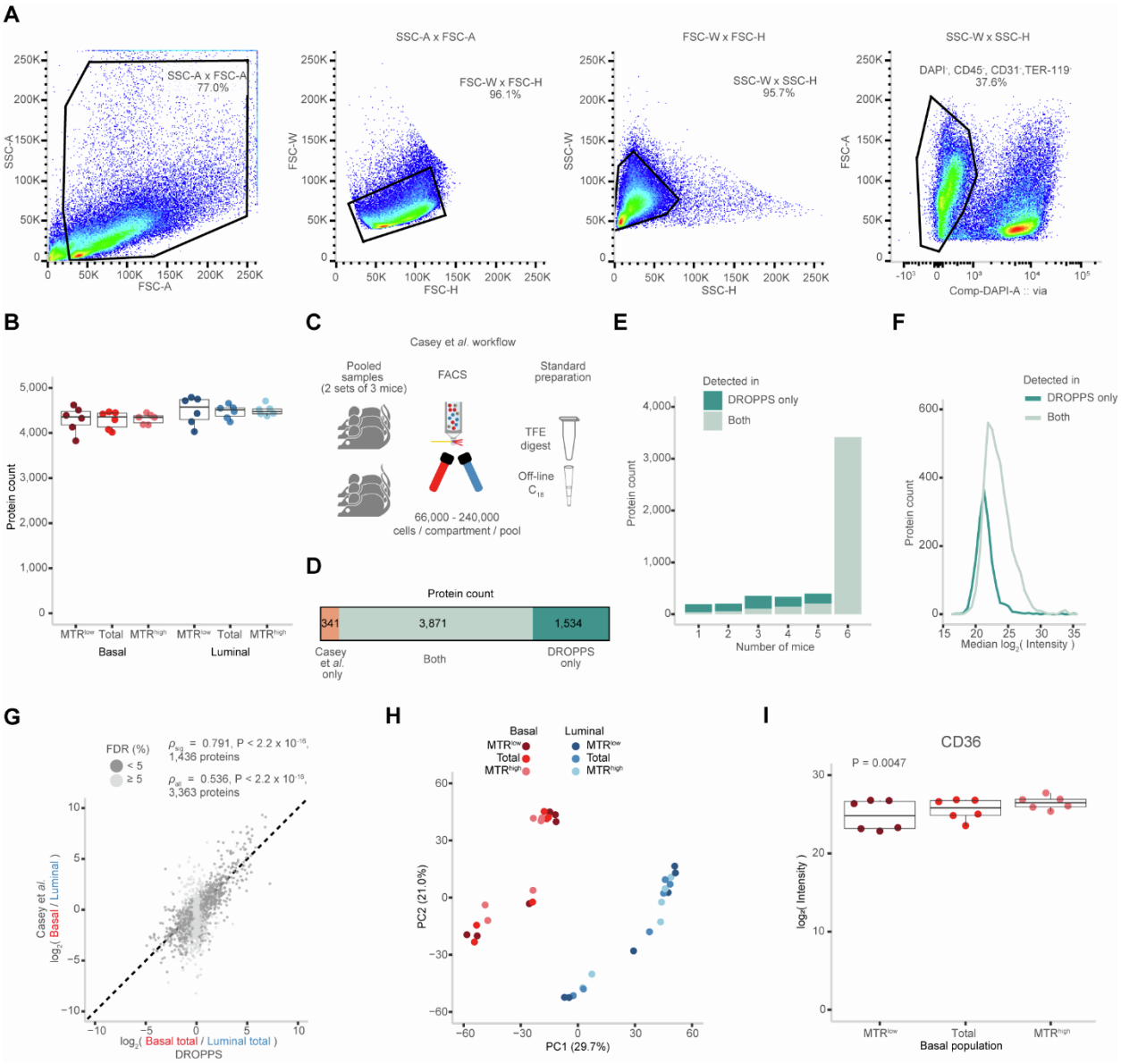

Figure S3. Proteomic profiling of primary mammary epithelial cells sorted by mitochondrial potential, related to Figure 3. (A) Gating scheme used for FACS purification of epithelial subpopulations. (B) The protein counts of the individual runs shown (n = 6 per group) (C) Schematic of experimental workflow used in previous study to sort mammary epithelial compartments. (D) Stacked bar plot showing the protein counts of shared, DROPPS-unique, and Casey *et al.*-unique proteins. (E) Stacked column plot depicting a histogram of the number of individual mice and the set of proteins that were detected using DROPPS including the overlap with our previous dataset (F) Density plot depicting the distribution of $\log_2$(intensities) for proteins detected using DROPPS and the overlap with our previous dataset. (G) Dot plot depicting the fold change of the total basal and total luminal populations compared to the fold changes from Casey *et al*. where points are colored by significance from the current study. The Spearman's correlations and significances are shown for all and significantly different proteins along with $y = x$ for improved visualization. (H) First two principal components of PCA. (I) The $\log_2$(intensities) of CD36 with repeated measures ANOVA p-value. Boxplots show the median, interquartile ranges, and 95% confidence interval estimate. MTR: MitoTracker Red.

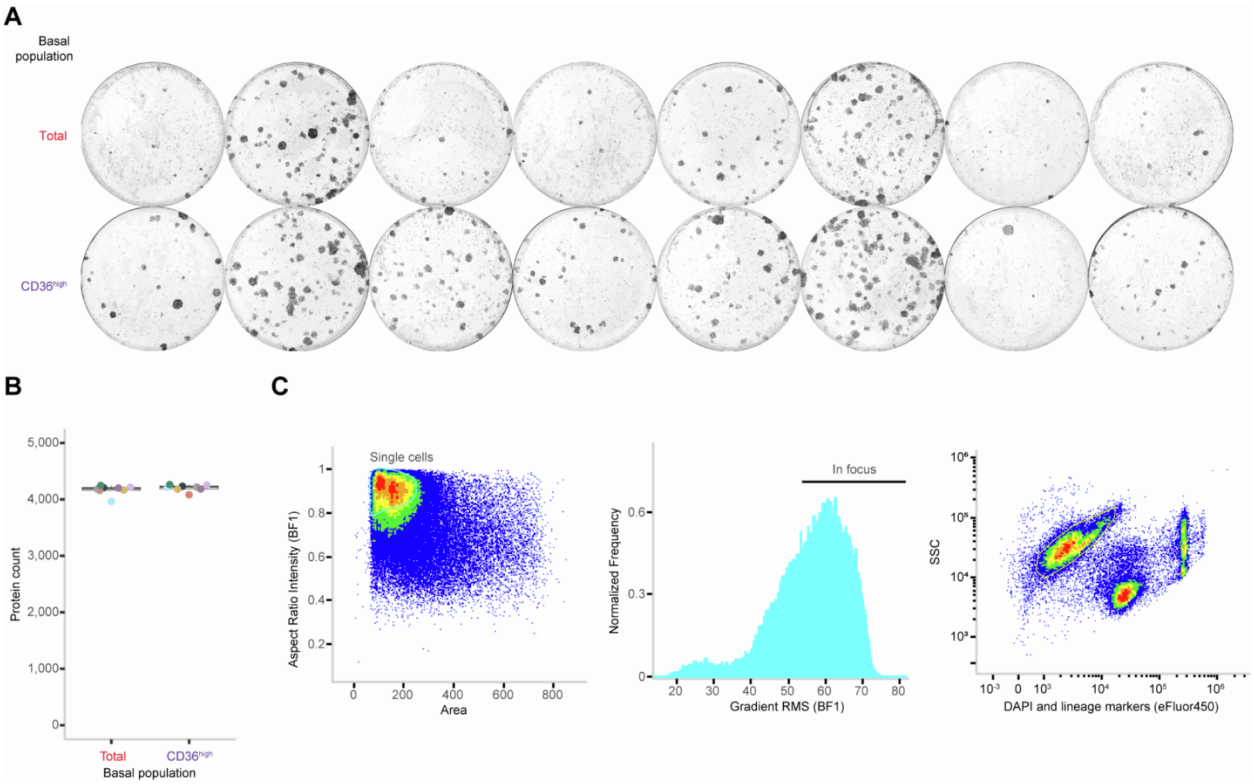

Figure S4. Investigating the phenotype of CD36^{high} basal mammary epithelial cells, related to Figure 4. (A) CFC assay images of basal mammary epithelial cells sorted by total or CD36^{high} basal populations. (B) The protein counts of the individual runs colored by mouse (*n* = 8). (C) Gating scheme used to analyze mammary epithelial cells by imaging flow cytometry Boxplots show the median, interquartile ranges, and 95% confidence interval estimate.