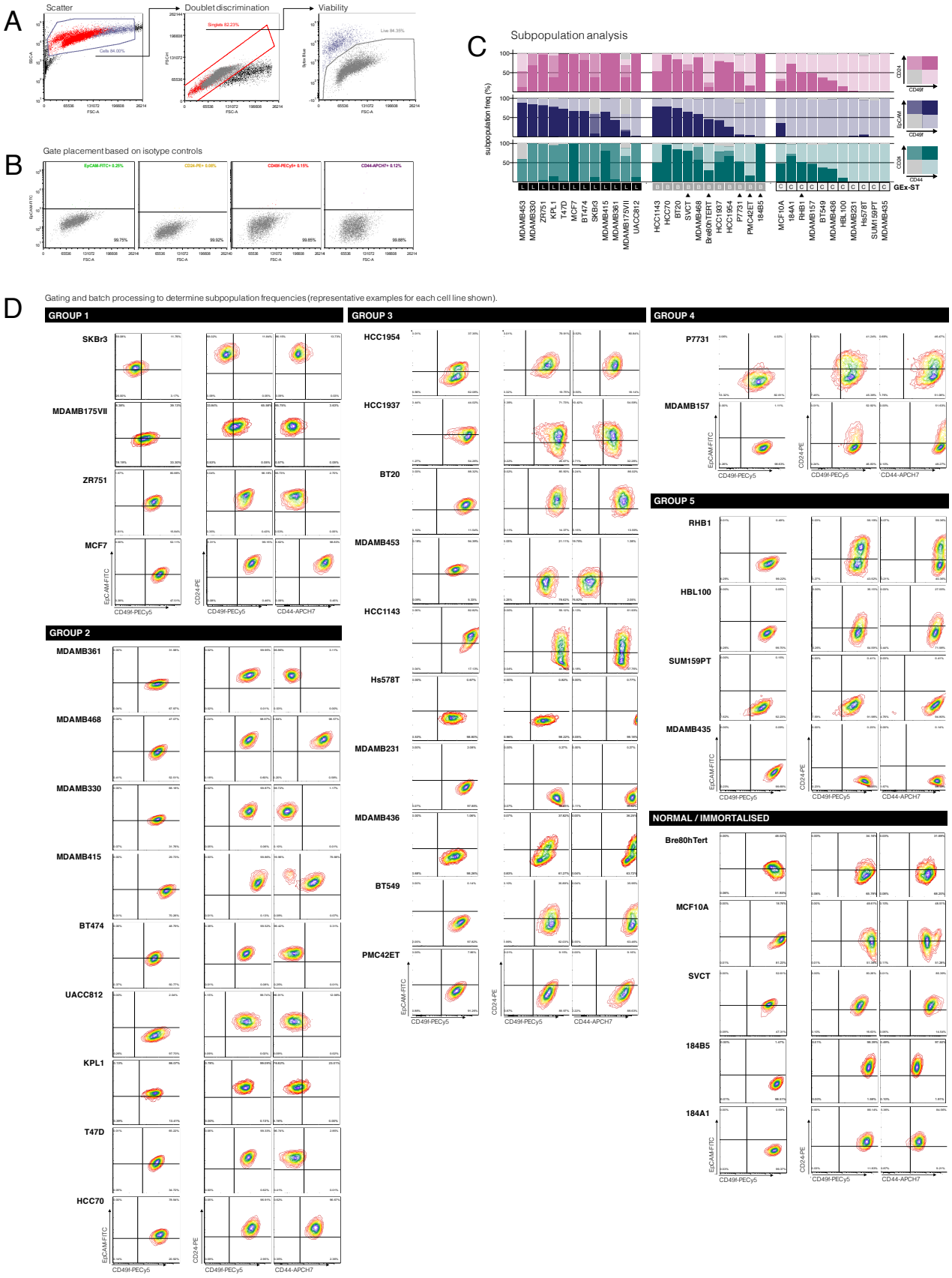
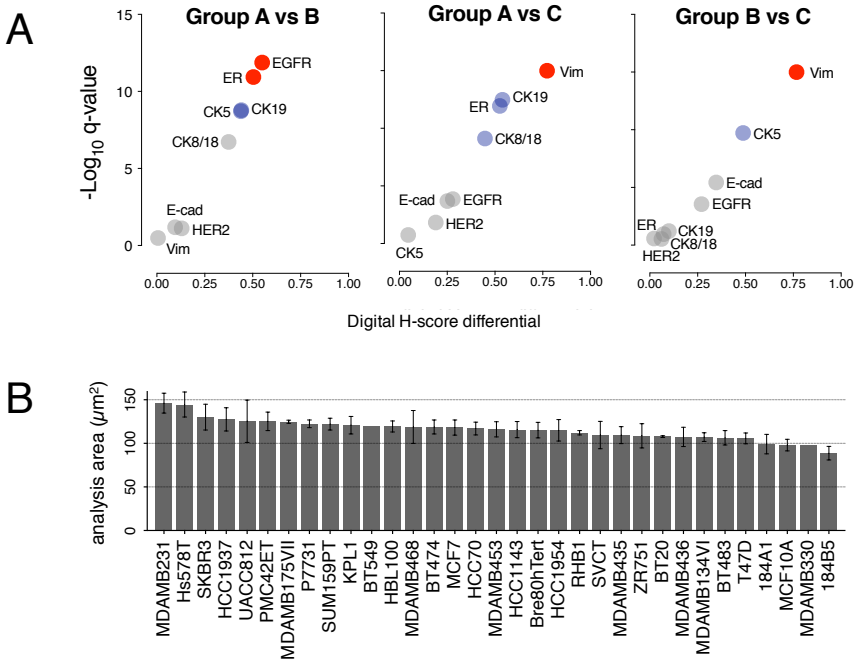


Figure-S1. Supporting data for Figure-1: *"Breast cancer cell line differentiation states assessed by multiplexed flow cytometry"*. Flow cytometry analysis was performed as previously described [6]. Briefly, debris, cell doublets and dead cells were excluded from further analysis (A). Isotype controls were used for gate placement (B), before batch analysis to automate calculation of subpopulation frequencies according to criteria in Fig-1. (C) Population frequencies according to multiplexed analysis of CD24/CD49f, EpCAM/CD49f and CD24/CD44. (D) Representative plots for each cell line.



<u>Figure S2</u>. Supporting data for Figure-2, "*Mapping heterogeneity in cell line cultures using digital immunohistochemistry (IHC) analysis of lineage markers*". (A) Results of a two-way analysis of variance (ANOVA) test demonstrating the markers that strongly (red) or moderately (blue) discriminate IHC clusters as a whole. Markers coloured grey are variable. The -Log10 q-values (corrected for multiple comparisons) are plotted against the H-score differential for each marker, in each pair. (B) Average (+/- standard deviation) area of the regions of interest (ROI) identified by the definiens algorithm, which was based on initially identifying haematoxylin-stained nuclei then dilating this area to define the ROI in which staining is quantified.

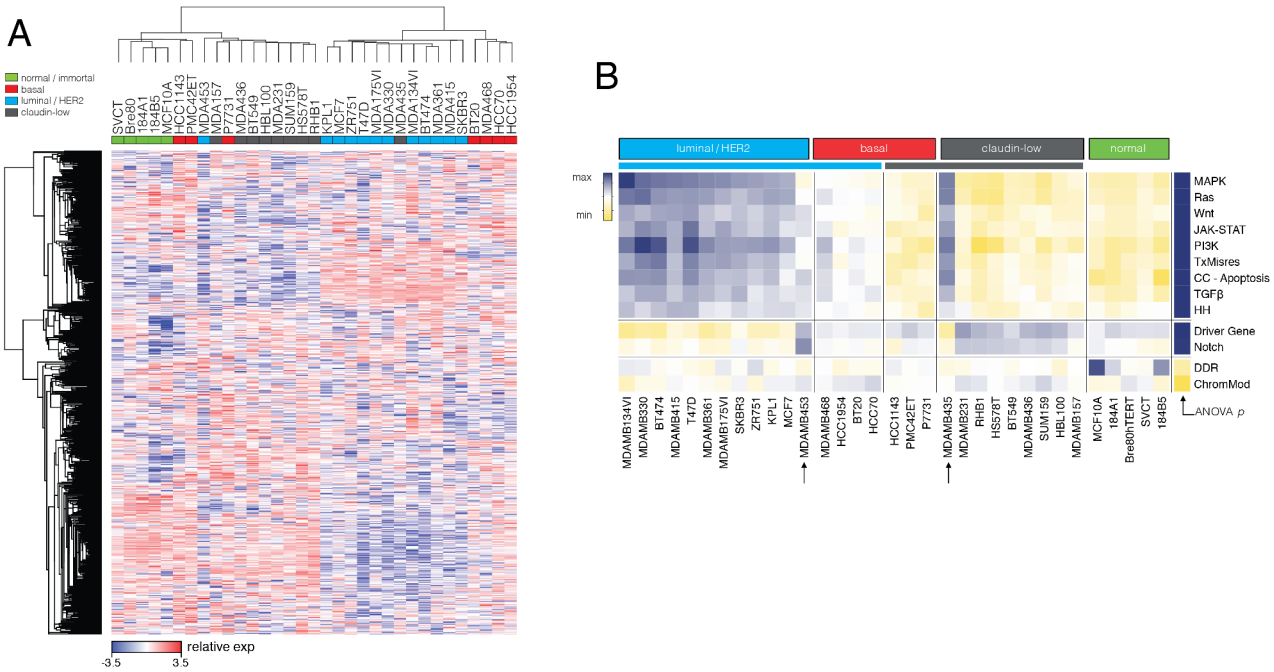


Figure-S3. Supporting data for Fig-3: "*Pan-cancer pathway profiling by digital mRNA counting identifies new breast cancer cell line subgroups*". (A) Heatmap and cluster dendrogram showing five main subgroups produced by unsupervised analysis of nCounter® PanCancer pathways expression data from the cell line panel prior to removal of MDAMB435 and MDAMB453, which were identified as outliers after considering expression of individual pathways. Note that outlier removal did not alter the subgroups as shown in Fig-3A. The coloured tile bar indicates the most robust subgrouping based on analysis of pathways individually (Fig-3B/S3B). (B) Heatmap showing pathway scores for each cell line. Lines classified as 'basal-like' based on whole transcriptome profile or surrogate cytokeratin expression segregated into two groups in this analysis – 4/7 were more similar to the luminal/HER2 group, and 3/7 ('basal-mes') exhibited more similarity to the claudin-low and normal groups according to their expression of cancer pathway genes. Arrows: discordant pathway expression for MDAMB435 and MDAMB453 with respect to expression subgroup. Abbreviations: CC, cell cycle; ChromMod, chromatin modification; DDR, DNA damage repair; Tx, transcriptional.