

Cell Reports Medicine, Volume 3

Supplemental information

**Single-cell proteomics defines the cellular
heterogeneity of localized prostate cancer**

Laura De Vargas Roditi, Andrea Jacobs, Jan H. Rueschoff, Pete Bankhead, Stéphane Chevrier, Hartland W. Jackson, Thomas Hermanns, Christian D. Fankhauser, Cedric Poyet, Felix Chun, Niels J. Rupp, Alexandra Tschuebunin, Bernd Bodenmiller, and Peter J. Wild

Supplemental figures and legends:

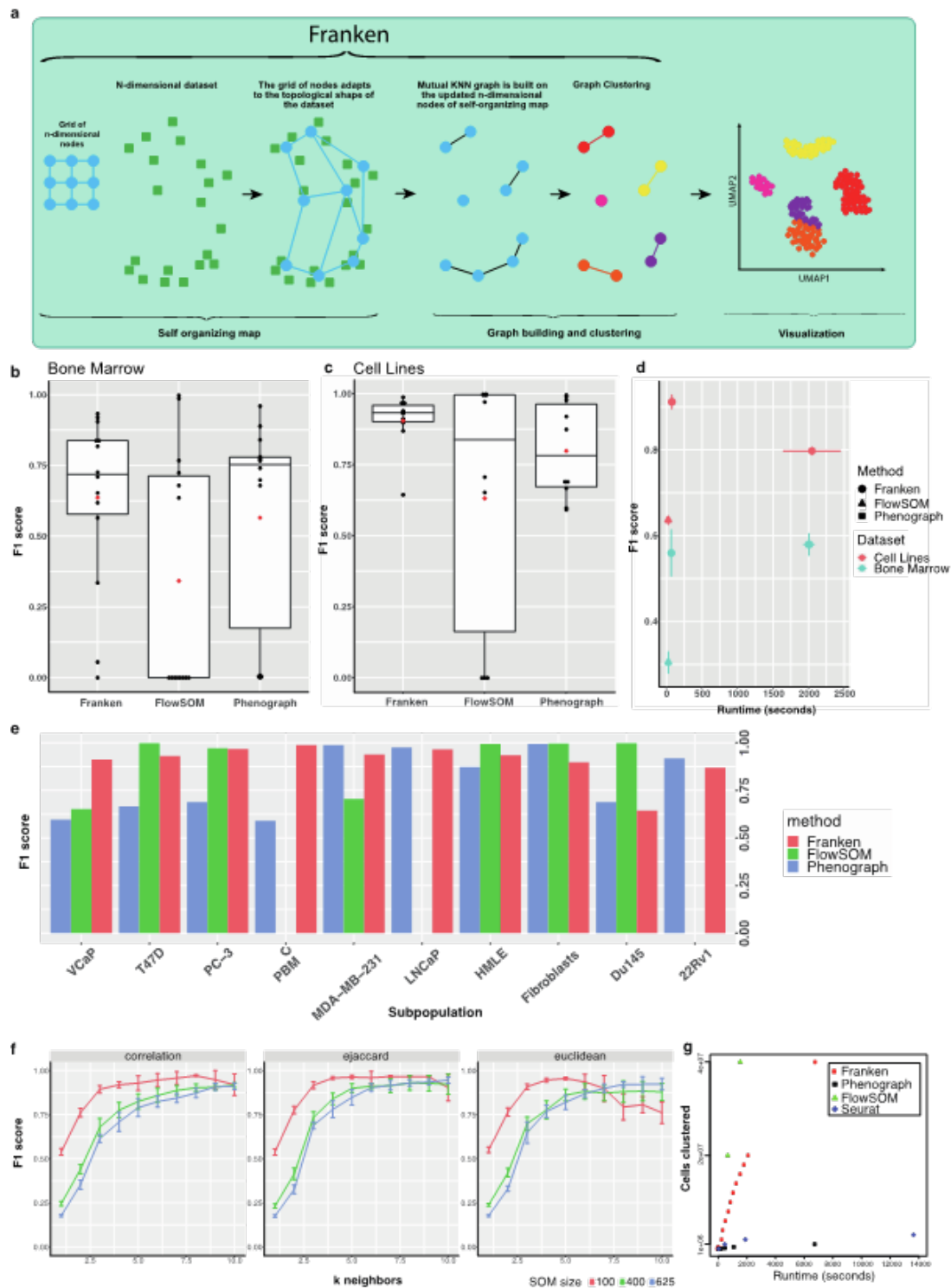


Figure S1. High-dimensional clustering with Franken combines performance and speed. (a) Franken involves fitting a large SOM to the data, building a mutual k -nearest neighbor graph on the n -dimensional SOM nodes, and applying the Walktrap algorithm. A UMAP 2-D map can be used for visualization. (b-c) Performance of Franken, PhenoGraph, and FlowSOM in clustering of b) bone marrow and c) cell lines as measured by F1 scores. (d) Average runtimes of Franken, PhenoGraph, and FlowSOM applied to the bone marrow (blue) and cell line (pink) datasets 10 times and plotted against mean F1 score. FlowSOM and PhenoGraph were run using their default parameters in their R and Matlab implementations, respectively (e) subset-specific evaluation of the different clustering approaches on the cell lines dataset. (f) F1 values for

Franken tested on cell lines dataset for different SOM sizes as well as number of k neighbors and distance measures in mutual nearest neighbor graph. The Tanimoto similarity (also known as extended Jaccard; ejaccard) is the recommended measure for Franken. The default k-nearest neighbors is $kn=6$. As Franken builds a mutual k-nearest-neighbor graph, too few neighbors can lead to poor results, however Franken proved to be very stable a minimum number of neighbors of at least $kn = 3$. (g) Franken, PhenoGraph and FlowSOM were run on a series of synthetic datasets of increasing size up to 40 million cells and runtimes were recorded for each. These consisted of simulated gaussians in 10 dimensions. We also included a state-of-the-art clustering technique used for single-cell RNA sequencing data (Seurat) according to Duo et al. 2018. PhenoGraph and Seurat could not be run on the largest datasets and were therefore run on the largest computationally feasible set (1 million for PhenoGraph and 3 million for Seurat). Franken was able to cluster 40 million cells in the same time necessary for PhenoGraph to cluster 1 million cells. Seurat performed slightly faster than PhenoGraph but attempting to cluster 3 million cells with Seurat took twice as long as it would take to cluster 40 million cells with Franken. FlowSOM was faster than all methods, however its F1-score performance was far inferior compared to Franken and PhenoGraph when applied to all benchmarking datasets (Figure 2b-d). Franken still provided comparable speed to FlowSOM and was at most three times slower for the largest 40 million dataset. Related to STAR methods' *Franken pipeline* section.

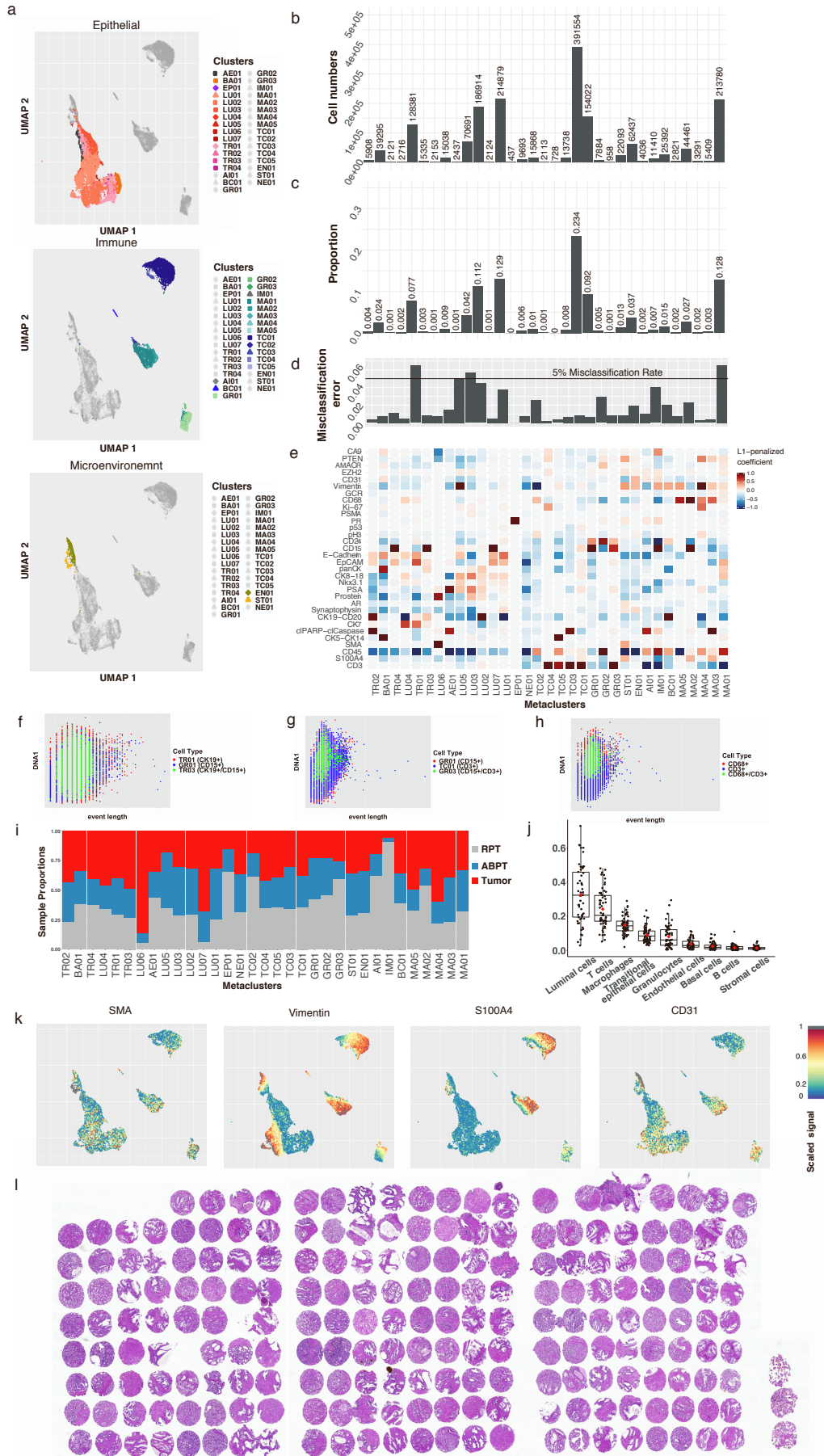


Figure S2. Logistic regression classification confirms metaclustering resolution. Metacluster labels given in b-e are all given in panel (d). (a) UMAP 2-D representation colored by only epithelial, immune or microenvironment cell subsets (b) Number of cells in each metacluster (c) Proportion of metacluster (d) Misclassification error for each metacluster resulting from logistic regression classification. (e) Coefficients from logistic regression performed with LASSO regularization. Biaxial plots of cell subsets confirming co-expression of markers is not a result from doublet formation for (f) CD15+/CK19+ (g) CD3+/CD15+ and (h) CD3+/CD68 cells. (i) Relative proportion of cells in a metacluster from a tumor-sample (red), adjacent benign prostatic tissue-sample (ABPT; blue) and from a random prostatic tissue (RPT; grey) (j) Average proportion (normalized by total number of cell in a patient) of cell types across all 58 patients in cohort (k) Expression of Stromal (SMA, S100A4 and Vimentin) and endothelial (CD31) markers (l) H&E of TMA from patient cohort analysed with CyTOF. The TMA contains two BPH and two tumor regions from all patients in the selected cohort was generated as previously described (Mortezavi et al. 2011). For TMA construction, representative tumor areas of the second and third slice of radical prostatectomy specimens were chosen, as close as possible to the area of tissue sampling for mass cytometry. Related to Figure 2.

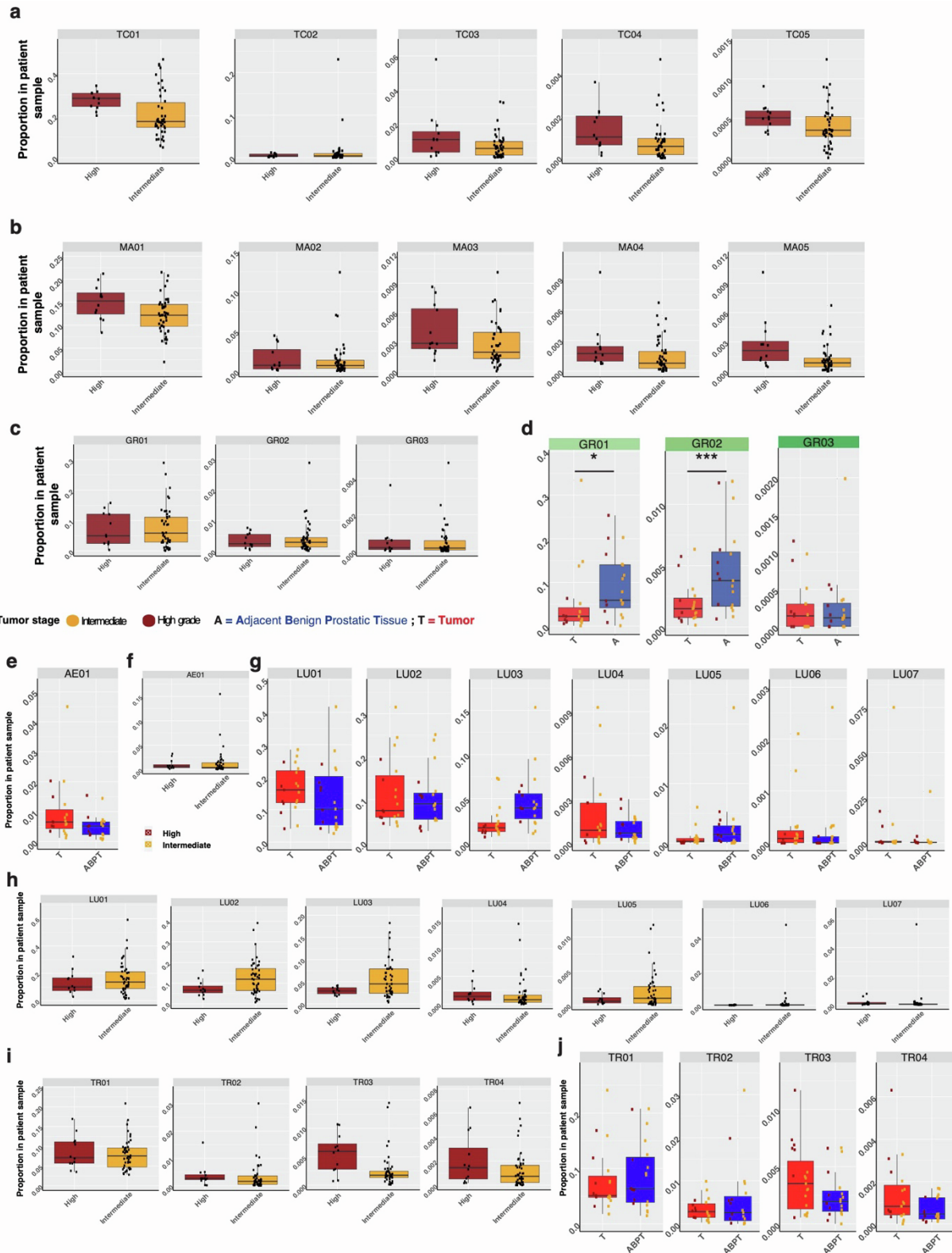


Figure S3. Immune and epithelial metacluster comparison between grades and tumor/non-tumor groups. Proportions (normalized by total number of cells in a patient) of (a) T cell, (b) macrophages and (c) granulocytes metaclusters across all 58 patients in cohort (Intermediate N = 46 and high grade N = 12) (d) Proportion (normalized by total number of cells in a patient) of cells from tumor and adjacent benign prostatic tissue samples for each granulocyte metacluster (N=17). Dots are colored by disease severity (intermediate vs high grade). Paired tumor/ABPT samples were analysed with a two-sided Wilcoxon signed rank paired test and

unpaired intermediate/high grade samples were analysed with a two-sided Wilcoxon rank sum test (also known as a Mann-Whitney-Wilcoxon). Related to Figure 3. (e) Comparison of proportions (f) Proportion (normalized by total number of cells in a patient) of luminal cells metaclusters stratified by (g) tumor and adjacent benign prostatic tissue samples (N=17) and (h) intermediate and high grade patient samples (for combined tumor/ABPT; Intermediate N = 46 and high grade N = 12). (i) same as (h) for transitional epithelial metaclusters. (j) proportion (normalized by total number of cells in a patient) of cells from tumor and adjacent benign prostatic tissue samples for each transitional epithelial metacluster. Dots are colored by disease severity (intermediate vs high grade). Paired tumor/ABPT samples were analysed with a two-sided Wilcoxon signed rank paired test and unpaired intermediate/high grade samples were analysed with a two-sided Wilcoxon rank sum test (also known as a Mann-Whitney-Wilcoxon). Related to Figures 3 and 4.

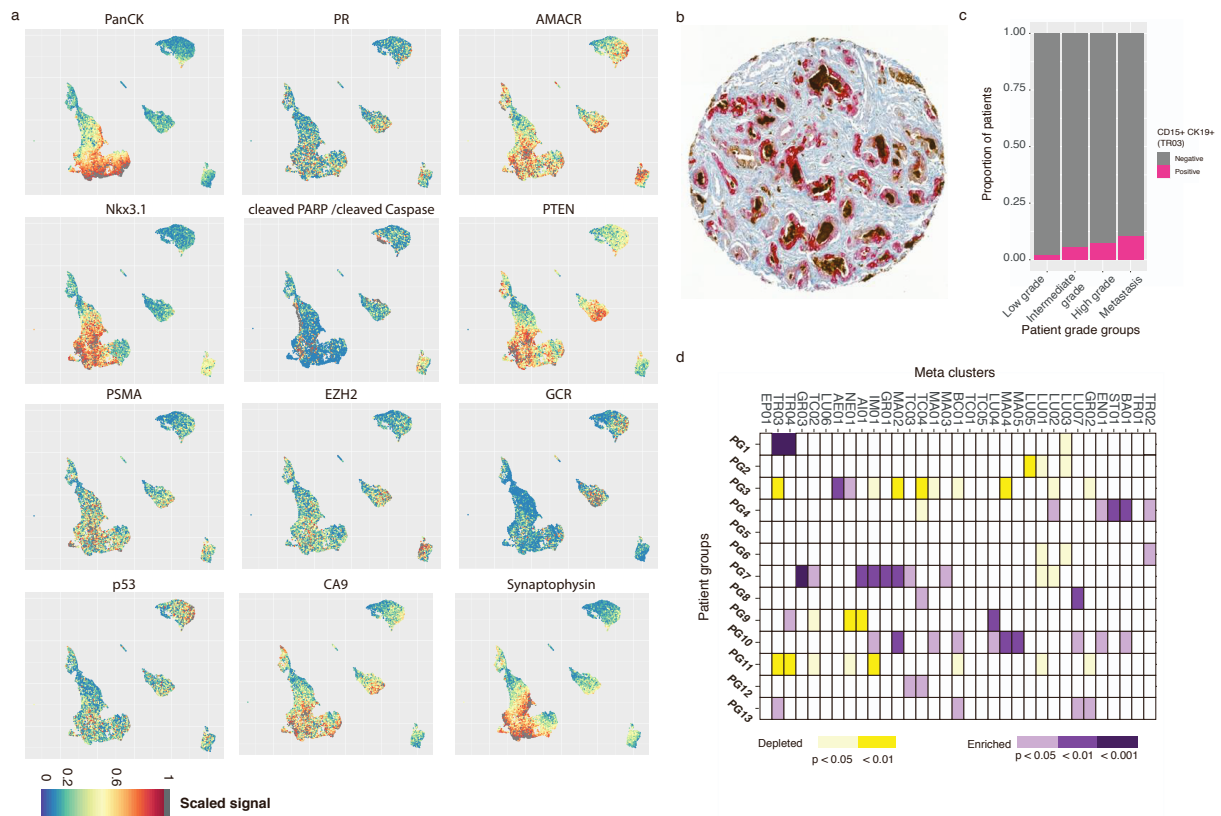


Figure S4. UMAP 2-D representation of cells across all 58 patients and analysis of CD15 prostate epithelial cells and patient groups. (a) 400 cells were sampled from each patient and projected using dimensionality reduction with UMAP. The normalized expression of each marker is shown for each cell. Related to Figure 4. **(b)** Sample core from TMA stained with CD15 (brown) and CK19 (red). The diameter of each spot is exactly 0.6 mm **(c)** The TMA was graded as positive or negative in the presence or absence (respectively) of double CD15 and CK19 positivity. The bar plot represents the proportion of patients positive for the double staining. From low grade (Gleason 6 or lower) to metastatic patients the proportions are (Positive/Negative): 1/47 (low grade); 11/188 (intermediate grade); 7/88 (high grade) and 4/34 (metastasis). **(d)** Summary of significance testing showing which cellular metaclusters are enriched or depleted in a patient group. Only significant p-values below of at least 0.05 are shown. Related to Figure 4.

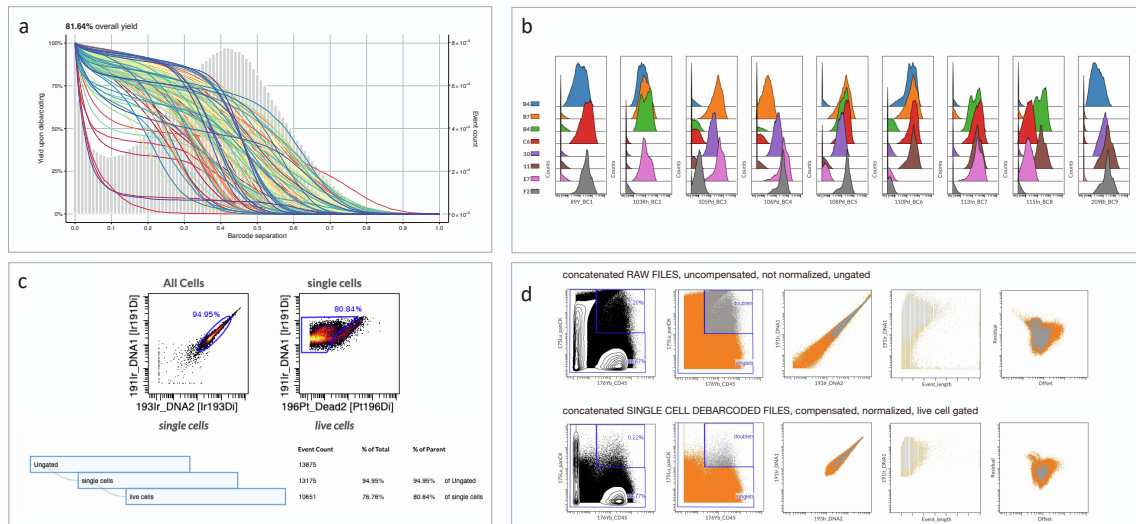


Figure S5. Quality control of CyTOF samples. (a) Yield plots of doublet free sample debarcoding performed with CATALYST. Histogram in the back shows distribution of barcode population separations. Lines in the foreground show cell yields by sample as a function of the applied separation cut-off. Left y-axis corresponds to cell yield in percent; right y-axis shows the total number of cells. All measured samples are displayed. Three lines that correspond to the samples P61_I, P61IV, & empty show a poor yield upon debarcoding. They were excluded with other samples with too few cells or poor viability subsequently. (b) Example of barcoding signatures after doublet free debarcoding. Histograms show staining intensities of barcoding reagent on 8 randomly selected samples. (c) Example (P38_N) for manual gating approach on live single cells in Cytobank to exclude intra well doublets and dead cells. (d) Dot plots show effectiveness of the chosen gating strategy to remove immune-epithelial can only be removed by exclusion non existing CD45+/panCK+ clusters from the analysis. Related to STAR methods' *Mass cytometry data analysis* section.

	Intermediate vs High	Tumor vs ABPT	Intermediate vs High Tu
AE01	0.299667932	0.003917694	1
AI01	0.947145257	0.567760468	0.122947744
BA01	0.783528041	0.921714783	0.490207806
BC01	0.11411	0.541217804	0.141652594
EN01	0.430598236	0.708557129	0.730861244
EP01	0.066186256	0.711997146	0.366052516
GR01	0.917029458	0.044559479	0.332195797
GR02	0.962233422	0.000419617	0.629796416
GR03	0.915765168	0.798272981	0.770263521
IM01	0.520189906	0.257926941	0.368409325
LU01	0.588859896	0.395462036	0.266593958
LU02	0.036087507	0.465316772	0.730861244
LU03	0.097218186	0.000164032	0.02128483
LU04	0.286692639	0.395462036	0.832332166
LU05	0.25706893	0.025821686	0.004760342
LU06	0.0124254	0.530158015	0.028507805
LU07	0.028625801	0.097942043	0.915768036
MA01	0.032658761	0.09551239	0.162362792
MA02	0.842335487	0.168792725	0.836593489
MA03	0.060665363	0.515277863	0.730861244
MA04	0.033903965	0.000419617	0.002884886
MA05	0.012244233	0.014068604	0.000773994
NE01	0.430598236	0.123188019	0.890655784
ST01	0.327307712	0.921714783	0.836593489
TC01	0.005867147	0.066287994	0.055551646
TC02	0.857185684	0.168792725	0.836593489
TC03	0.082336603	0.022987366	0.141652594
TC04	0.046835508	0.00202179	0.028692568
TC05	0.034337657	0.332065582	0.013720799
TR01	0.754548216	0.738090515	0.783258279
TR02	0.2187601	0.241252899	0.490207806
TR03	0.003438092	0.09551239	0.026175063
TR04	0.175943059	0.087429047	0.534911812

Table 1. All statistical results from the Wilcoxon test results shown in figures 3 and 4 and supplemental figure 3.