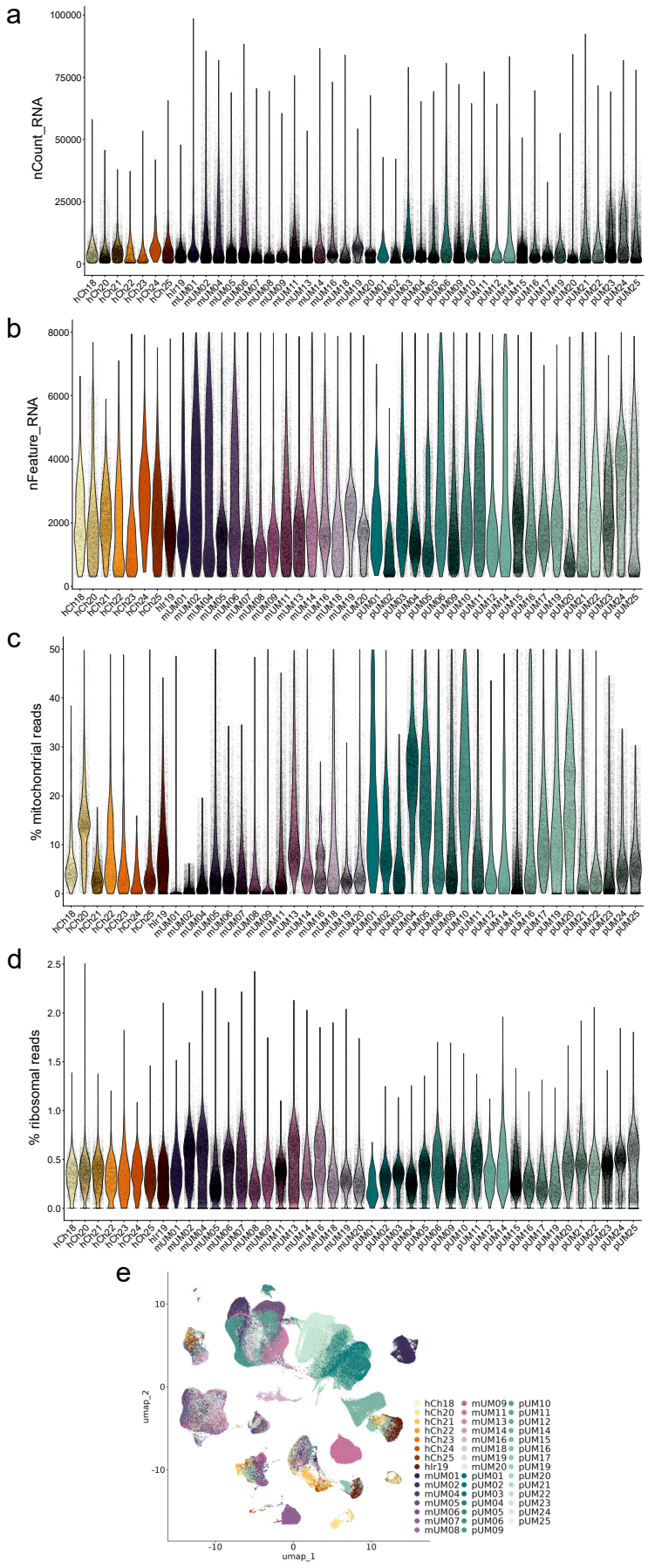
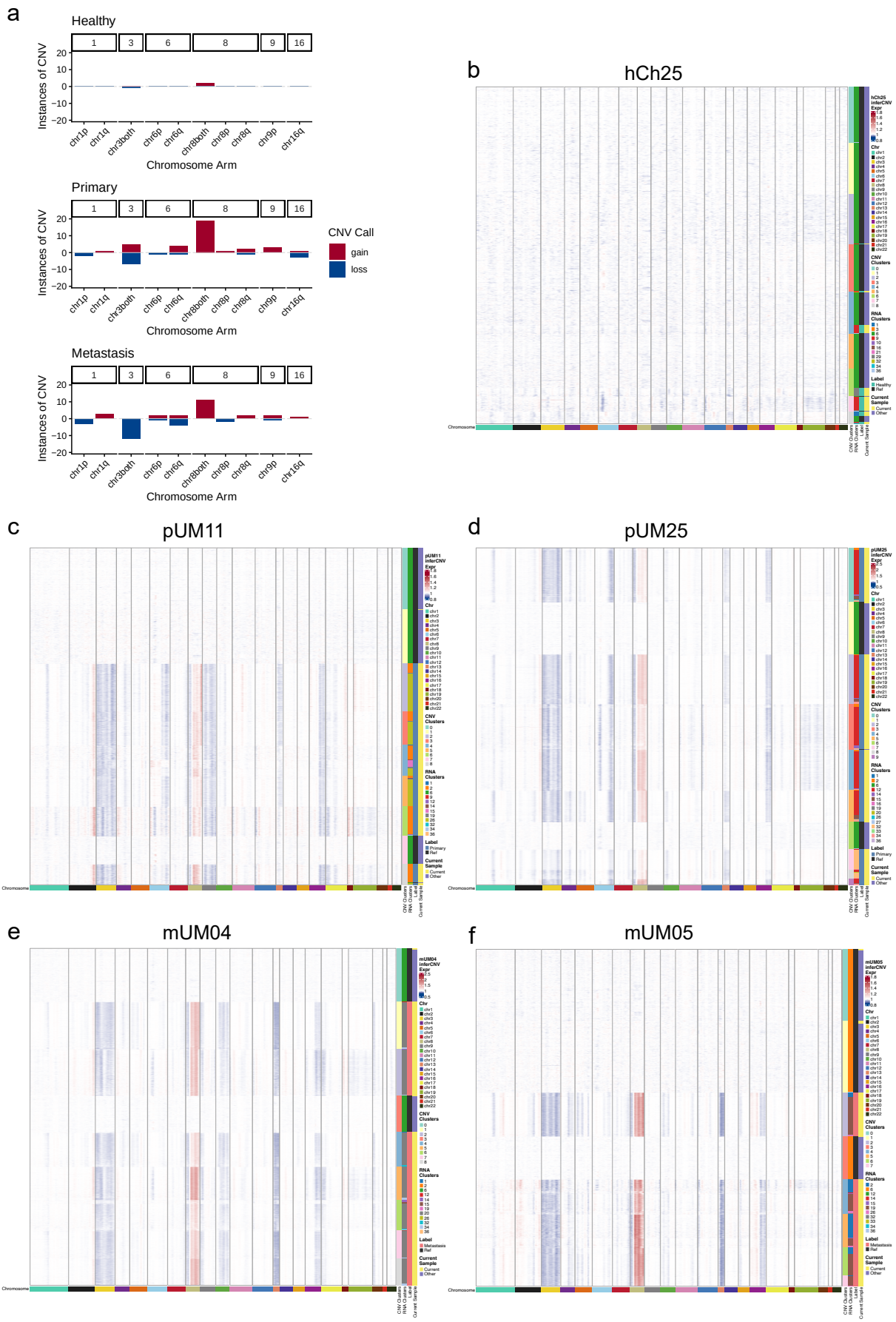
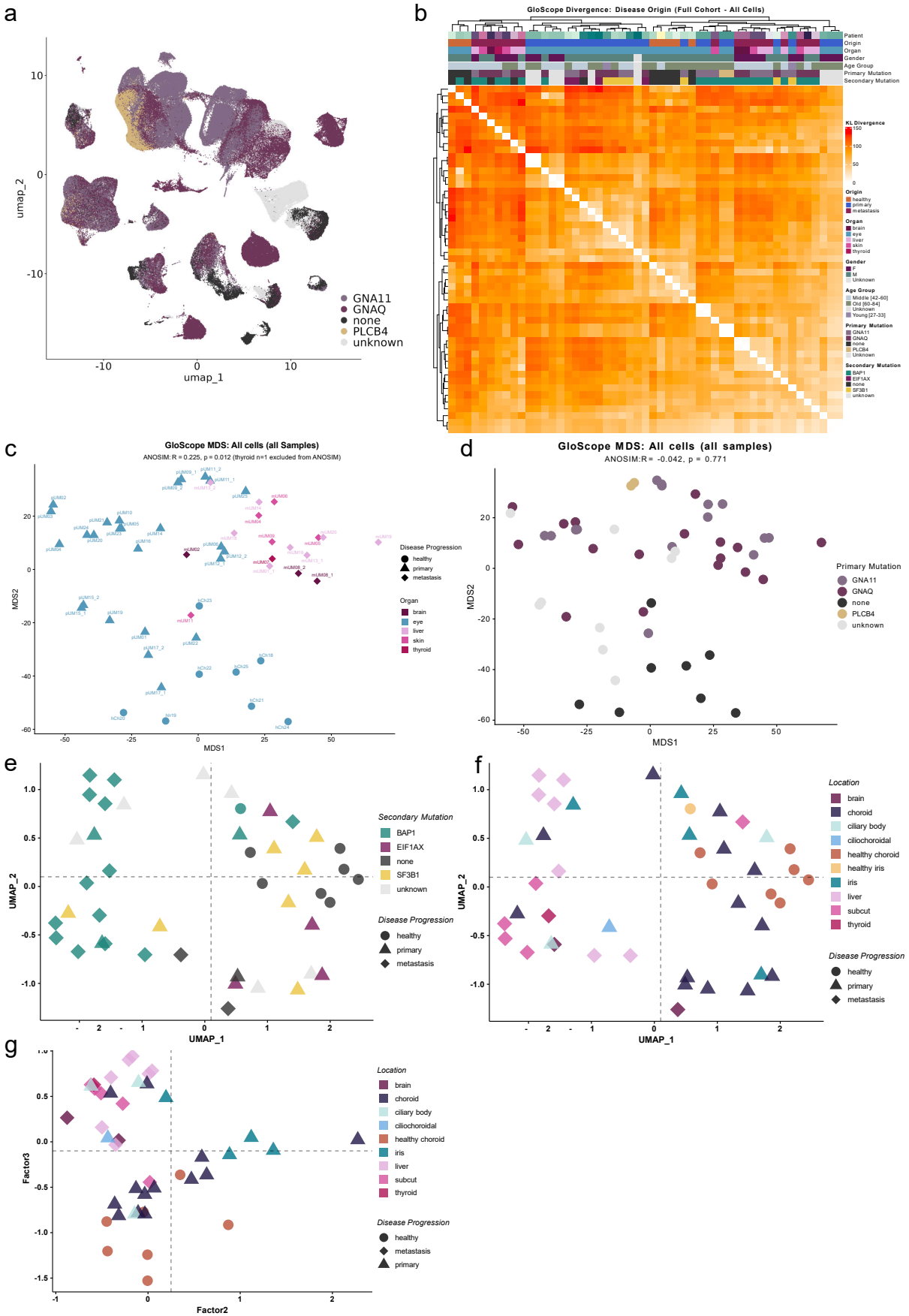
“A comprehensive single cell atlas of healthy ocular tissue and uveal melanoma reveals mutation- and uveal tract–associated cell states”



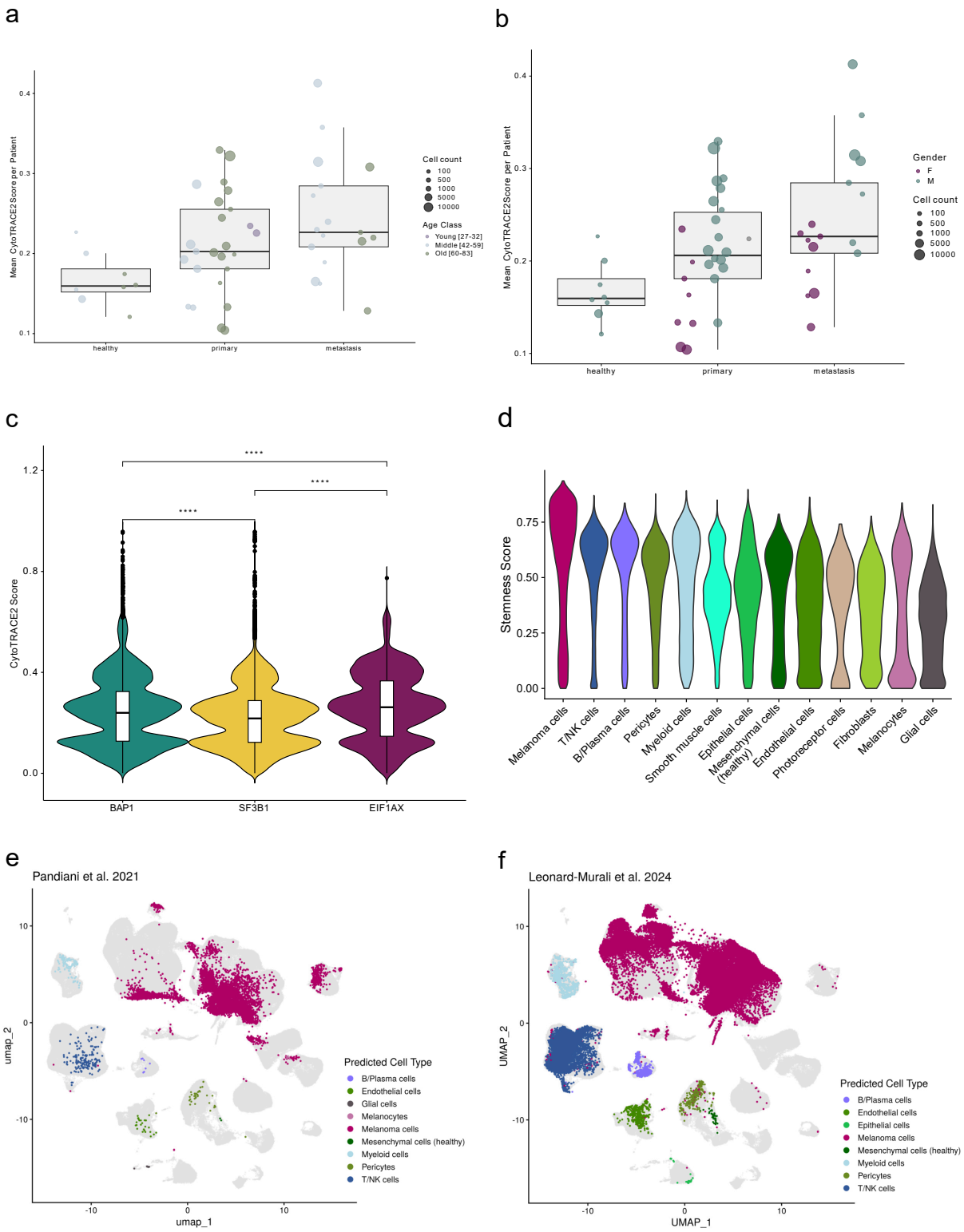
Supplementary Figure 1. a Violin plot showing number of detected transcripts (nCount_RNA). **b** Violin plot showing number of detected genes (nFeature_RNA). **c** Violin plot showing percentage of mitochondrial reads. **d** Violin plot showing percentage of ribosomal reads. **e** UMAP colored by tumor and tissue samples.
UMAP: Uniform Manifold Approximation and Projection.



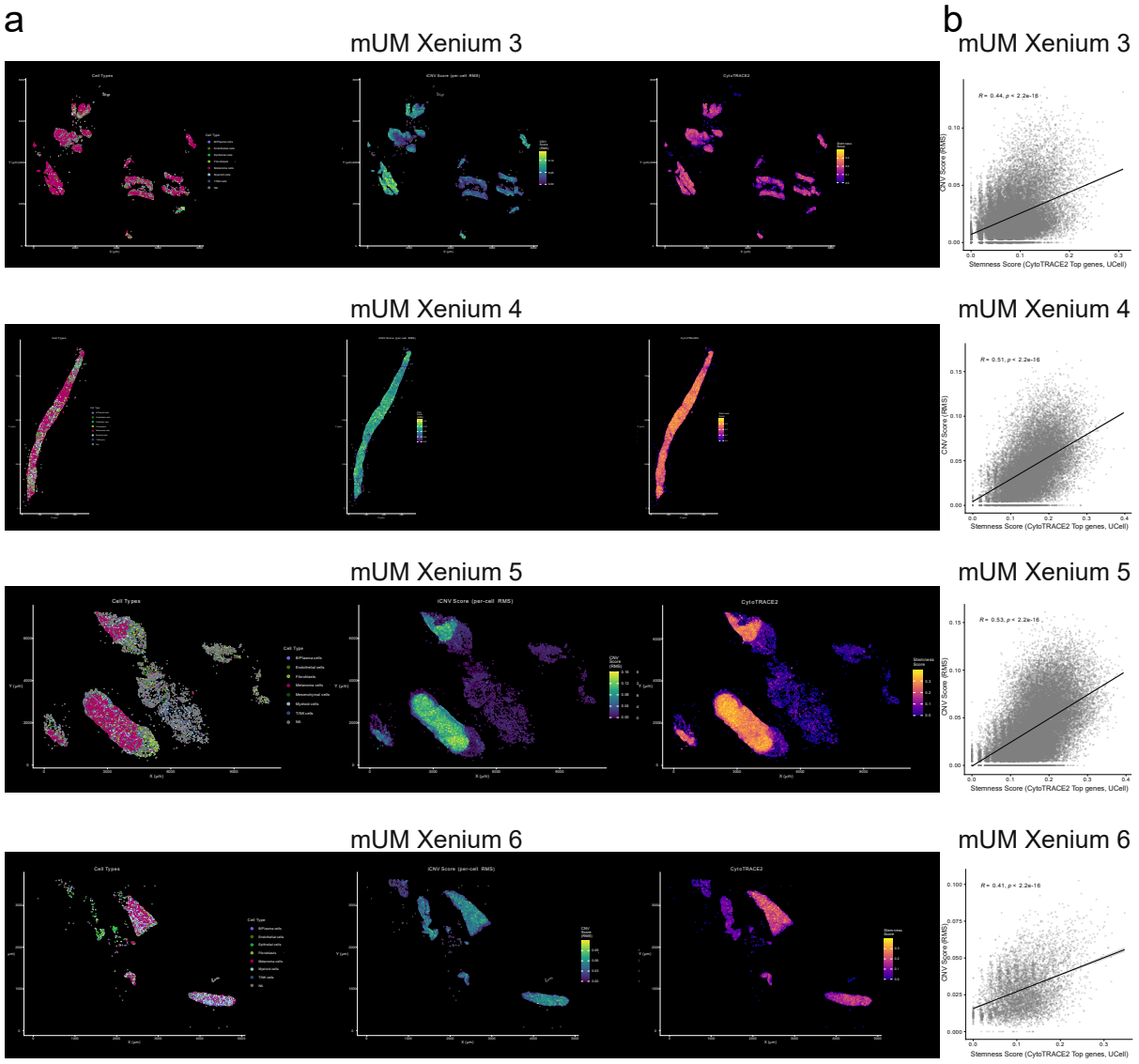
Supplementary Figure 2. a Summary bar plot showing inferred large CNV instances for most common chromosomal aberrations in uveal melanoma split by healthy tissue, primary and metastatic tumors. **b** Heatmap showing inferred large CNV for healthy melanocytes. **c** Heatmap showing inferred large CNV for a primary tumor. **d** Heatmap showing inferred large CNV for a primary tumor. **e** Heatmap showing inferred large CNV for a metastatic tumor. **f** Heatmap showing inferred large CNV for a metastatic tumor.
CNV: copy number variation.



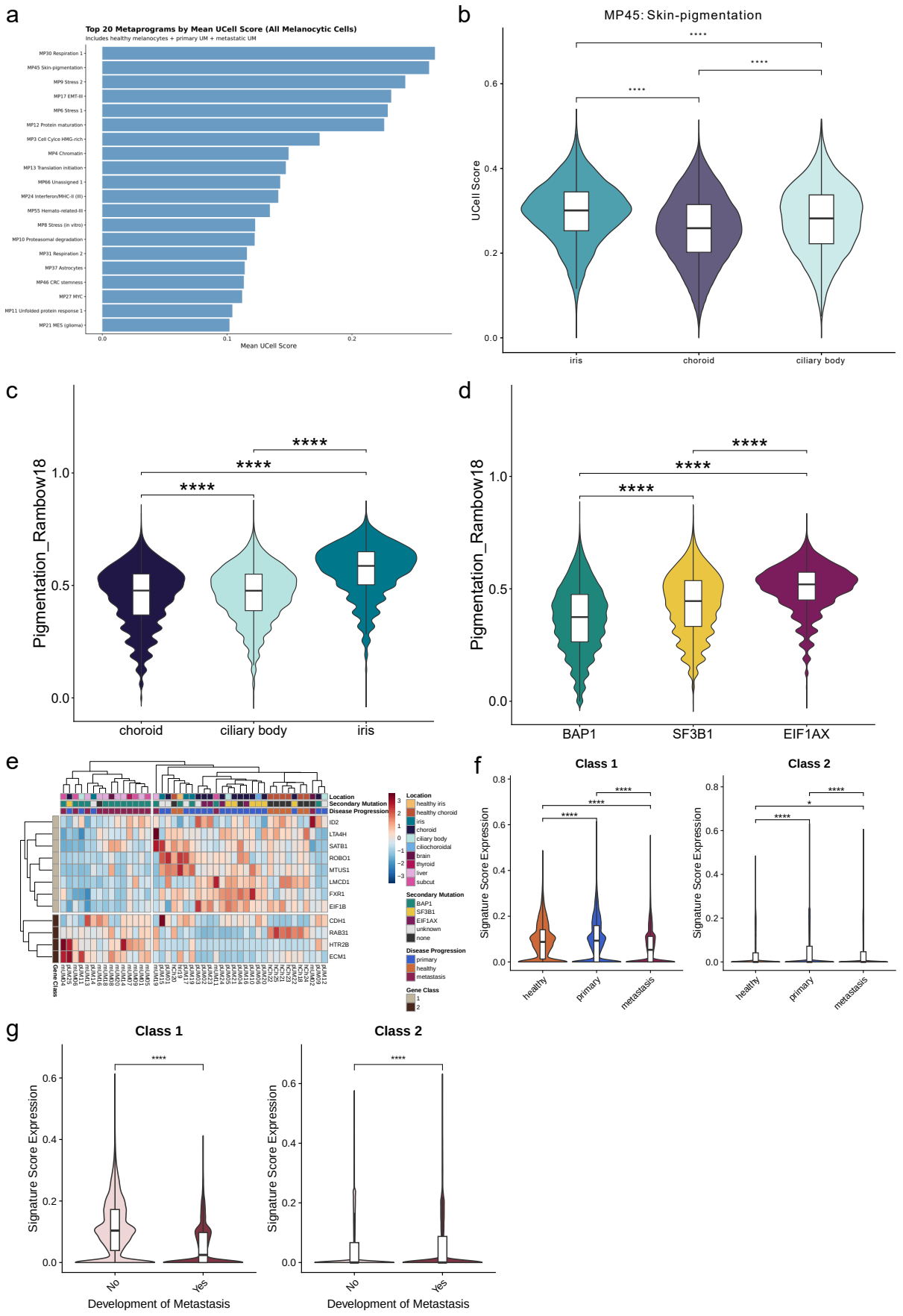
Supplementary Figure 3. a UMAP colored by primary driver mutation status. **b** GloScope divergence heatmap. **c** Projection of samples showing the distribution of samples based on GloScope-derived transcriptional patterns across organs of origin (brain, eye, liver, skin, thyroid). **d** Association between GloScope-derived transcriptional patterns and primary mutations in all samples. **e** UMAP plot showing the relationship of disease progression (healthy tissue, primary tumors, and metastases) and secondary mutation (*BAP1*, *SF3B1*, *EIF1AX*) as inferred by multi-cellular factor analysis. **f** UMAP plot showing the relationship of disease progression (healthy tissue, primary tumors, and metastases) and location (brain, choroid, ciliary body, ciliochoroidal, healthy choroid, iris, liver, subcut., thyroid) as inferred by multi-cellular factor analysis. **g** Scatter plot showing the relationship between inferred latent factors 2 and 3. Samples are annotated by clinical variables such as tissue or tumor location (brain, choroid, ciliary body, ciliochoroidal, healthy choroid, iris, liver, subcut., thyroid) and disease progression (healthy tissue, primary tumors, and metastases).
UMAP: Uniform Manifold Approximation and Projection, MDS: multidimensional scaling, subcut: subcutaneous.



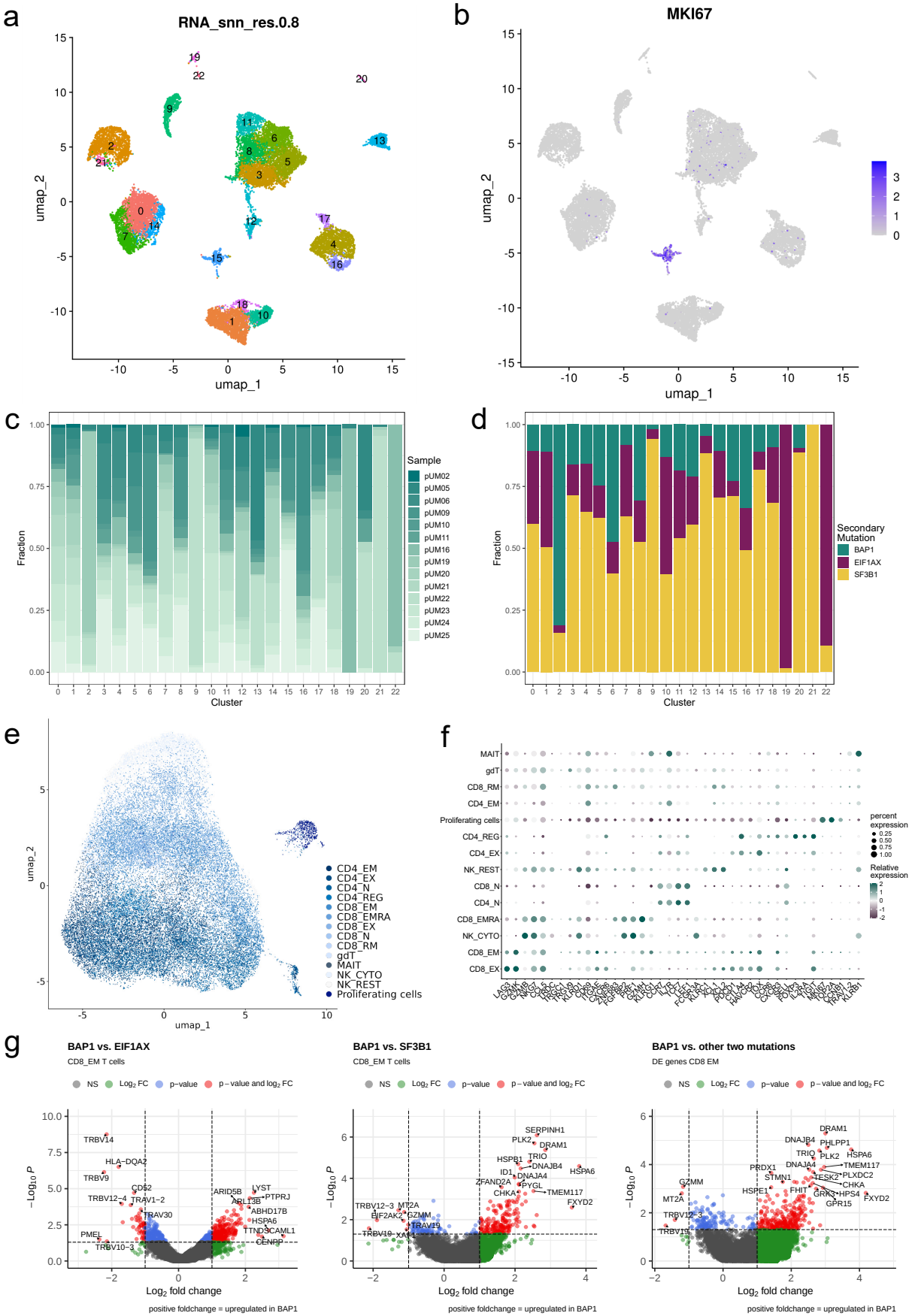
Supplementary Figure 4. a Patient-level CytoTRACE2 developmental potential scores across age class (young, [27-32], middle [42-59], old [60-83]). **b** Patient-level CytoTRACE2 developmental potential scores across gender (male, female). **c** CytoTRACE2 developmental potential of melanoma cells stratified by secondary mutation. **d** Validation of melanoma developmental potential in full dataset. **e** UMAP of external dataset assigning cell type labels based on clustering with our scAtlas (grey). **f** UMAP of external dataset assigning cell type labels based on clustering with our scAtlas (grey).
UMAP: Uniform Manifold Approximation and Projection.



Supplementary Figure 5. a Single-cell spatial transcriptomics of metastatic lesions showing spatial distribution of melanoma cells, inferred CNV scores, and CytoTRACE2 stemness scores. **b** Scatter plots showing correlation between stemness score and CNV intensity.
CNV: copy number variation.

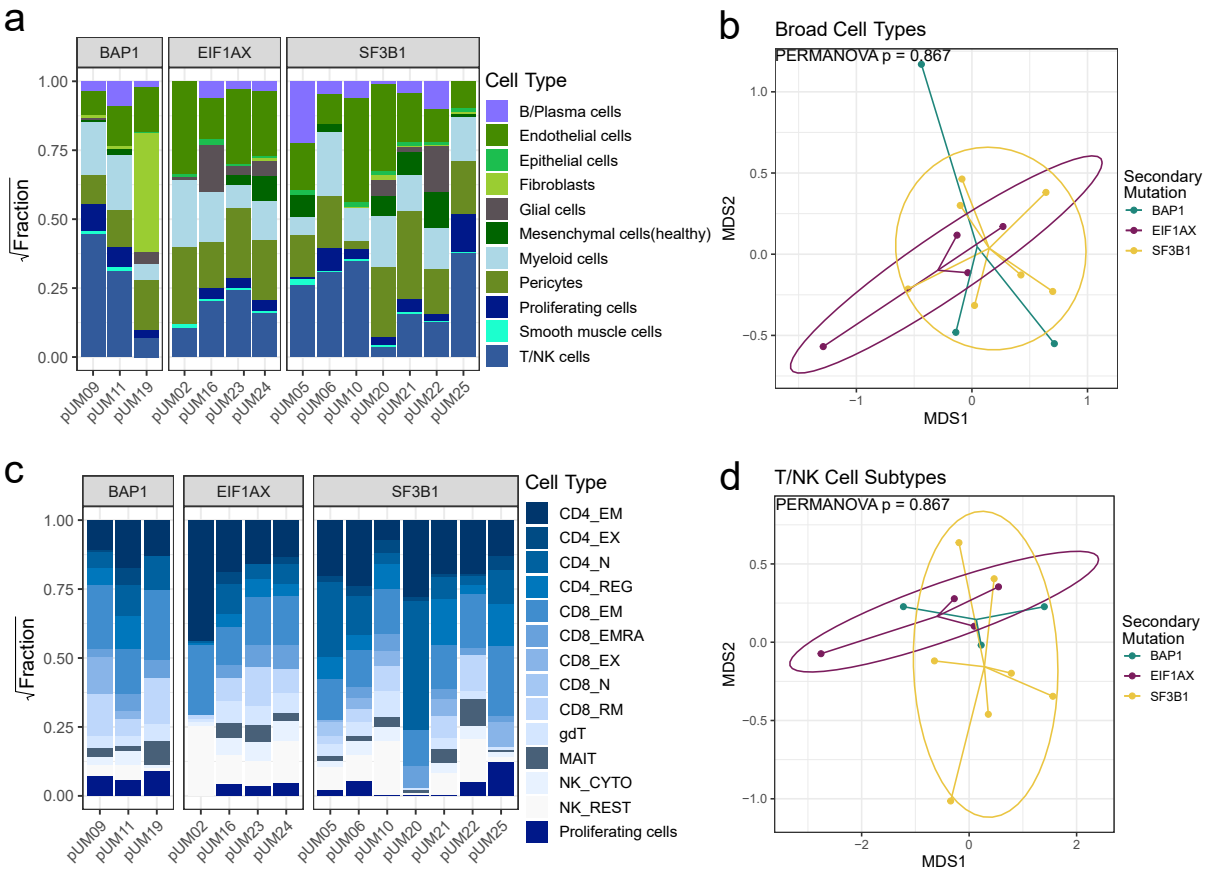


Supplementary Figure 6. a Bar plot of top 20 Metaprograms expressed in melanocytic cells. **b** Violin plot showing distribution of skin-pigmentation metaprogram (MP45) across uveal locations. **c** Expression of the Pigmentation_Rambow18 gene signature across uveal locations. **d** Expression of the Pigmentation_Rambow18 gene signature across secondary mutations. **e** Unsupervised hierarchical clustering of pseudobulk expression profiles separating genes by predicted class[57]. **f** Violin plot showing distribution of class 1 and class 2 genes across disease progression (healthy tissue, primary tumor, metastatic tumor). **g** Violin plot showing distribution of class 1 and class 2 genes across progression to metastatic disease (yes, no).
*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001

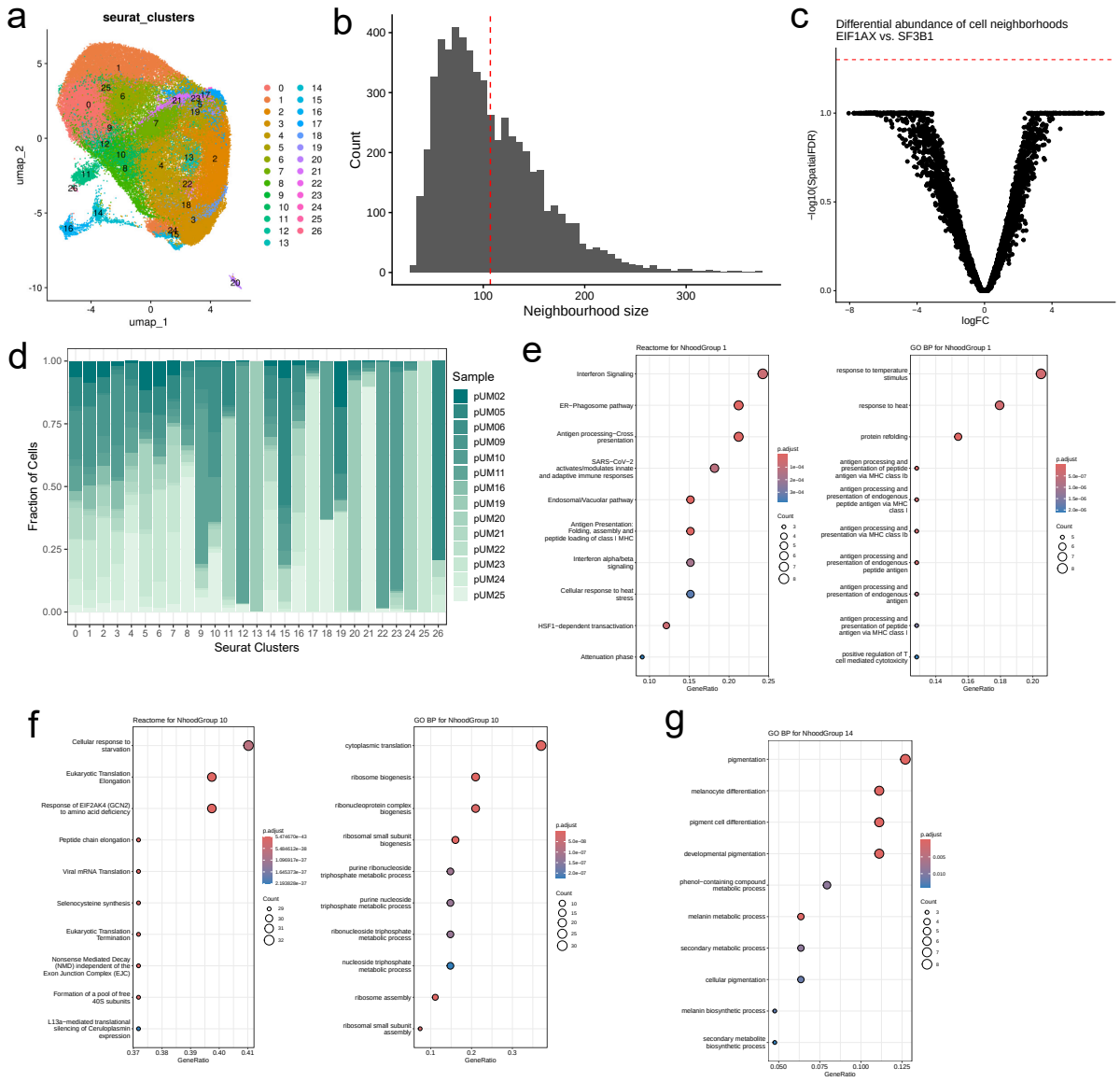


Supplementary Figure 7. a UMAP of all non-malignant cells (n=17'603) from primary tumors with known mutation in *BAP1, SF3B1* or *EIF1AX* colored by cluster at resolution 0.8. **b** Feature plot showing expression of the proliferation marker gene *MKI67*. **c** Bar plot showing compositions of Seurat clusters colored by sample. **d** Bar plot showing compositions of Seurat clusters colored by secondary mutation. **e** UMAP showing all T/NK cells (n=38'656) with more refined cell typing. **f** Dot plot of marker genes used for refined cell typing of T/NK cells. **g** Volcano plots of pseudobulk differential gene expression analysis results between CD8+ effector memory T cells of all three secondary mutations in pairwise comparisons

UMAP: Uniform Manifold Approximation and Projection, snn: shared nearest neighbor, DE: differentially expressed, CD4_EM: CD4+ effector memory T cells, CD4_EX: CD4+ experienced T cells, CD4_N: CD4+ naive T cells, CD4_REG: CD4+ regulatory T cells, CD8_EM: CD8+ effector/memory T cells, CD8_EMRA: CD8+ recently activated effector/memory T cells CD8_EX: CD8+ experienced T cells, CD8_N: CD8+ naive T cells, CD8_RM: CD8+ resident memory T cells, gdT: gamma-delta T cells, NK_CYTO: cytotoxic natural killer cells, NK_REST: resting natural killer cells.



Supplementary Figure 8. a Bar plot showing compositions corrected by sccomp for individual samples colored by major cell type. **b** NMDS plot of dissimilarity between secondary mutations based on major cell types. **c** Bar plot showing compositions of T/NK cells corrected by sccomp for individual samples colored by T/NK cell subtypes. **d** NMDS plot of dissimilarity between secondary mutations based on T/NK cell subtypes.
NMDS: Non-metric Multidimensional Scaling, PERMANOVA: Permutational Multivariate Analysis of Variance, CD4_EM: CD4+ effector memory T cells, CD4_EX: CD4+ experienced T cells, CD4_N: CD4+ naive T cells, CD4_REG: CD4+ regulatory T cells, CD8_EM: CD8+ effector/memory T cells, CD8_EMRA: CD8+ recently activated effector/memory T cells CD8_EX: CD8+ experienced T cells, CD8_N: CD8+ naive T cells, CD8_RM: CD8+ resident memory T cells, gdT: gamma-delta T cells, NK_CYTO: cytotoxic natural killer cells, NK_REST: resting natural killer cells.



Supplementary Figure 9. a UMAP of all malignant cells (n=93’377) from primary tumors with known mutation in *BAP1, SF3B1* or *EIF1AX* colored by cluster at resolution 0.8. **b** Histogram showing size distribution for all 5’755 neighborhoods identified by Milo. The dashed vertical line indicates the mean neighborhood size of 107 cells. **c** Volcano plot of results from differential abundance testing of neighborhoods between *EIF1AX* and *SF3B1*. **d** Bar plot showing compositions of Seurat clusters colored by sample. **e-g** Dot plots of overrepresentation analysis for NhoodGroups 1 (**e**), 10 (**f**) and 14 (**g**), using Reactome and Gene Ontology gene sets. No Reactome terms were significantly overrepresented for NhoodGroup 14.
UMAP: Uniform Manifold Approximation and Projection, FDR: false discovery rate, FC: fold change, Nhood: neighborhood