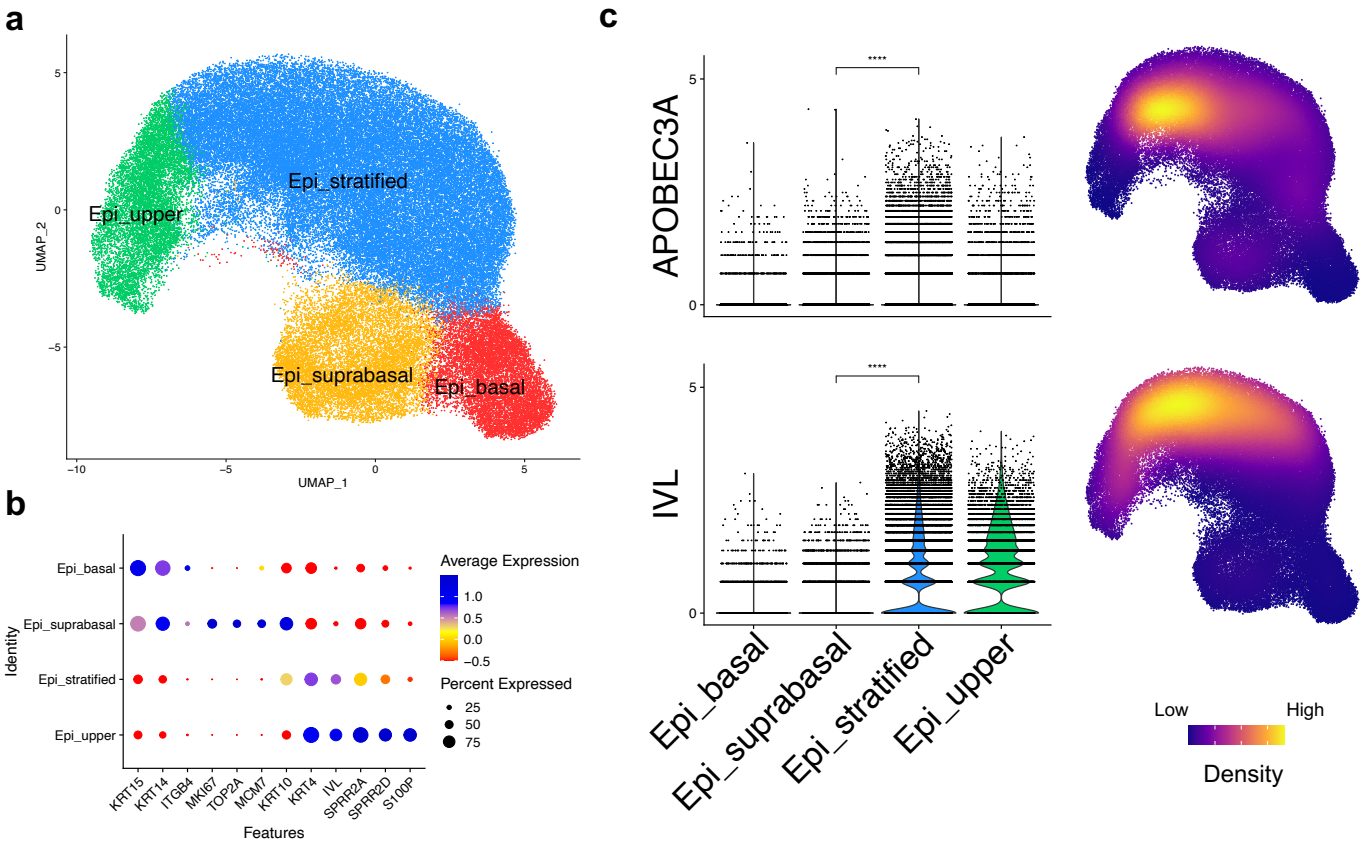
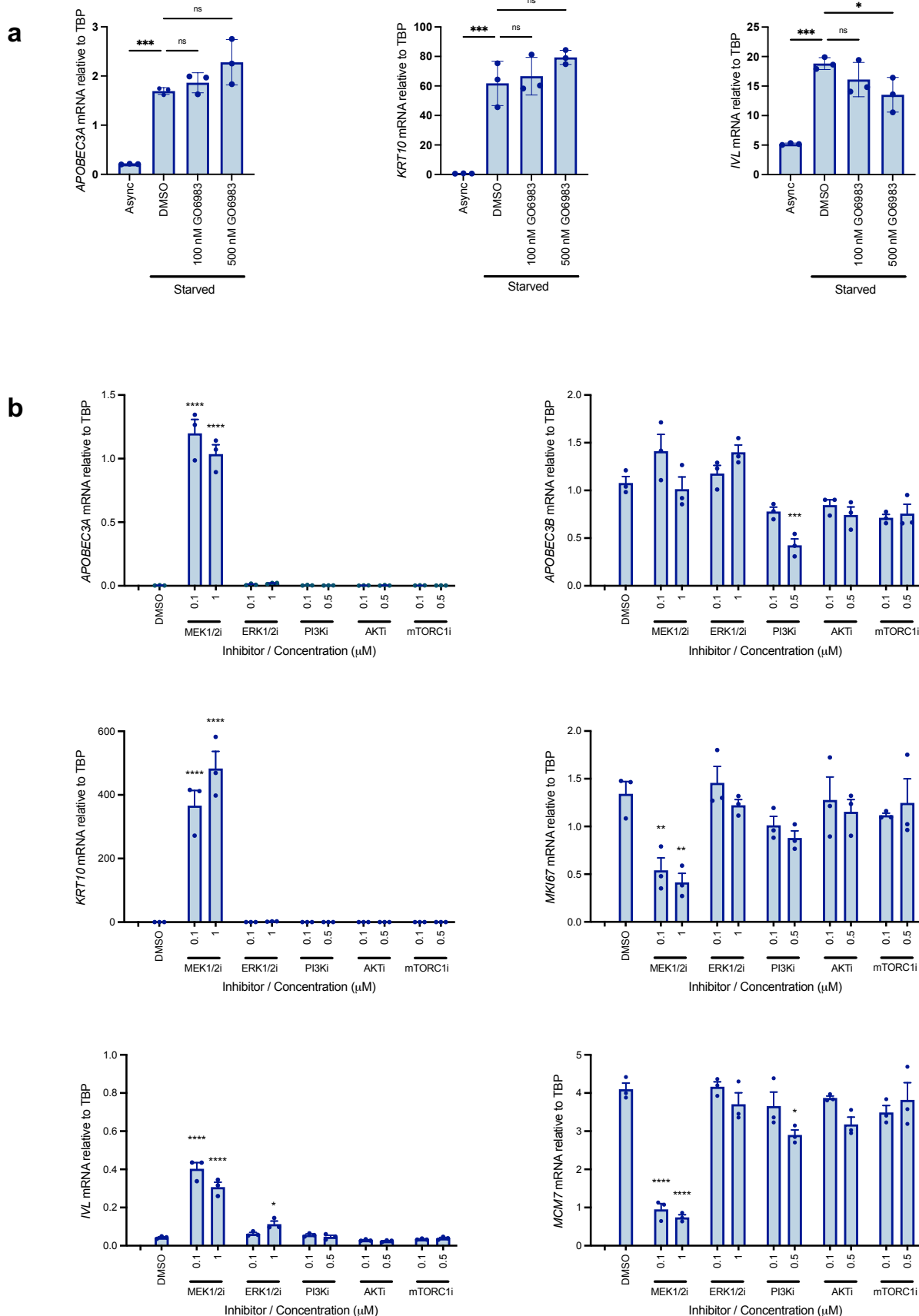
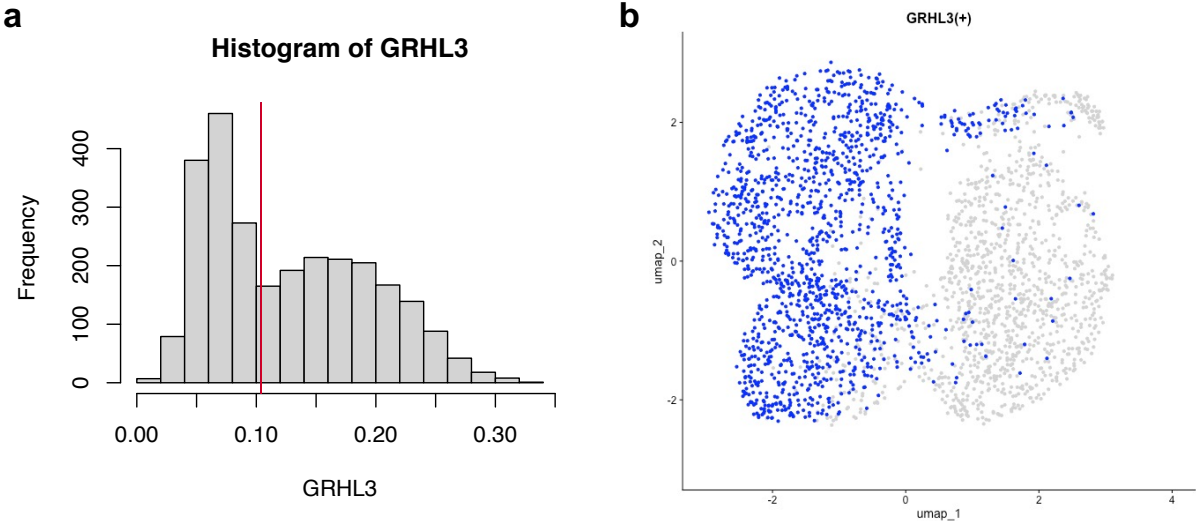




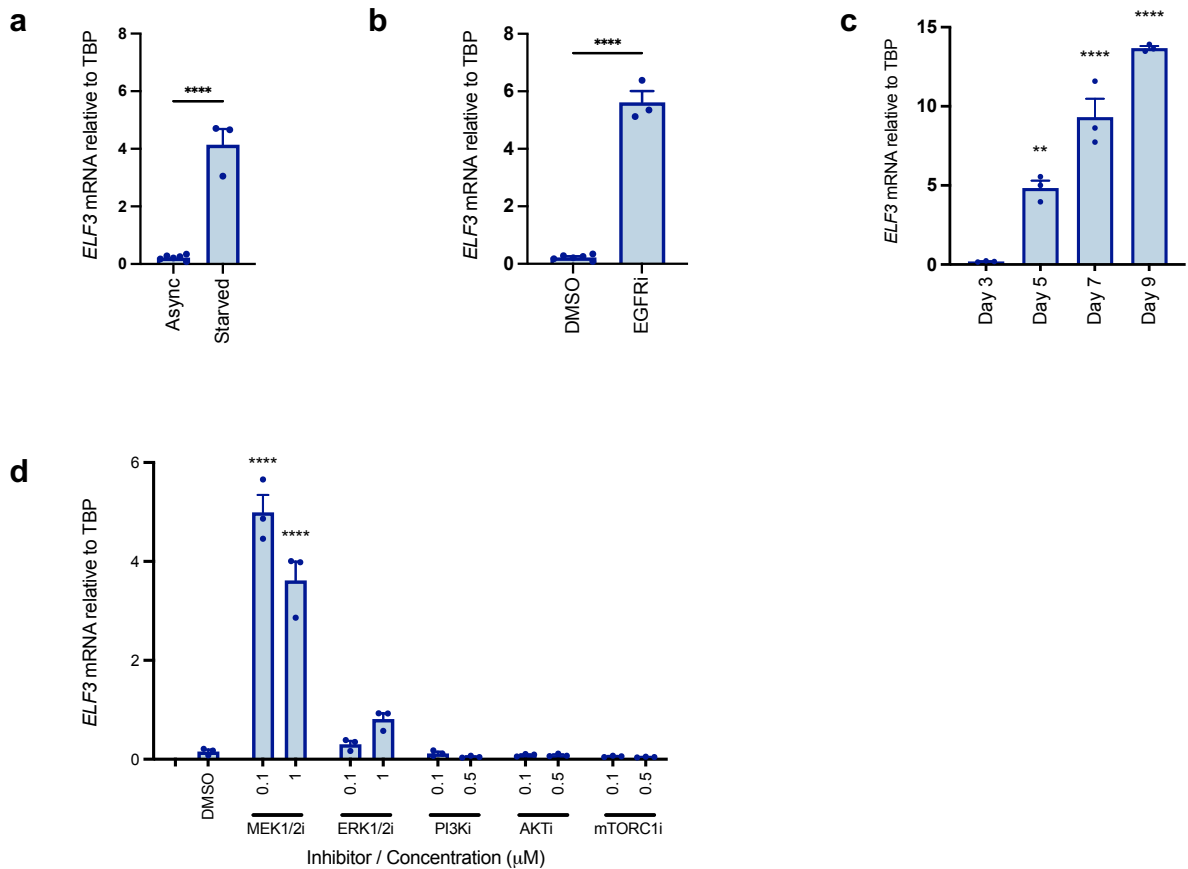
Supplementary Figure 1. *APOBEC3A* is predominantly expressed in differentiating ('epi-stratified') oesophageal epithelial cells **a)** UMAP projection of 80,489 oesophageal epithelial cells from healthy tissue. **b)** Dotplot of marker genes used to identify epithelial phenotypes. **c)** Violin plots of gene expression in the oesophageal epithelial cell states (left, **** = p-value < 0.0001, Wilcoxon's Rank Sum Test) and UMAP projections of the density of gene expression in individual oesophageal epithelial cells (right).



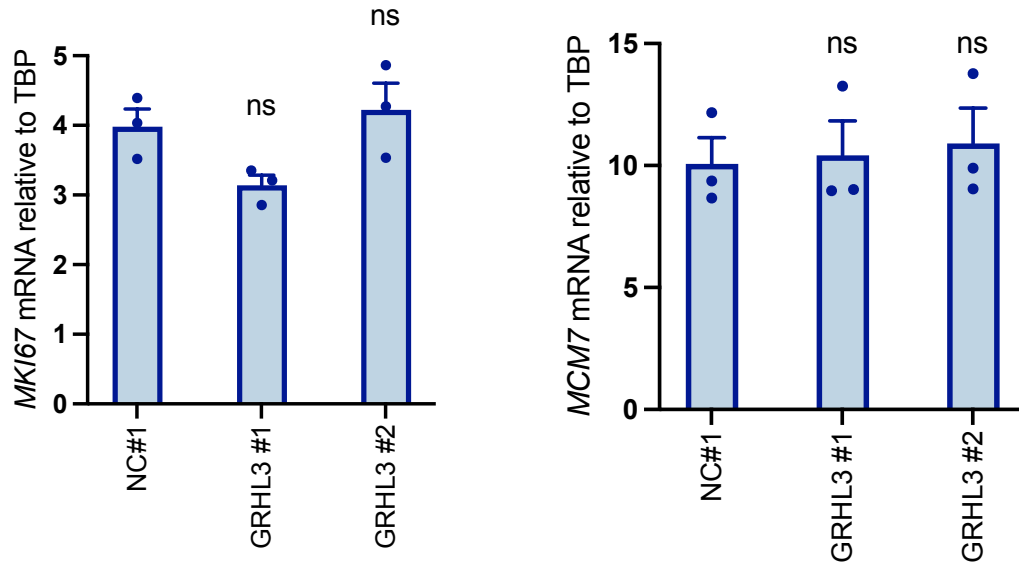
Supplementary Figure 2. *APOBEC3A* is induced upon growth factor withdrawal in NIKS and primary keratinocytes. (a) qRT-PCR measurements of *APOBEC3A*, *KRT10*, and *IVL* mRNA levels in asynchronous growing NIKS ('Async') and following 48 hours of growth factor withdrawal, in the presence of vehicle control (DMSO) or the PKC inhibitor GO6983 at either 100 nM or 500 nM as shown. (b) qRT-PCR-based gene expression measurements of *APOBEC3A*, *KRT10*, *IVL*, *APOBEC3B*, *MKI67* and *MCM7* in NIKS following 24 hours of treatment with vehicle control (DMSO) or inhibitors of MEK1/2 (trametinib), ERK1/2 (ravoxertinib), PI3K (pictilisib), AKT (MK2206) or mTORC1 (everolimus). N = 3, error bars represent SEM. * = p-value < 0.05; ** = p-value < 0.01; *** = p-value < 0.001; **** = p-value < 0.0001. Comparisons to proliferating ('async') cells in (a) and to DMSO-treated cells in (b) were performed using one-way ANOVA with Dunnett's multiple comparisons test.



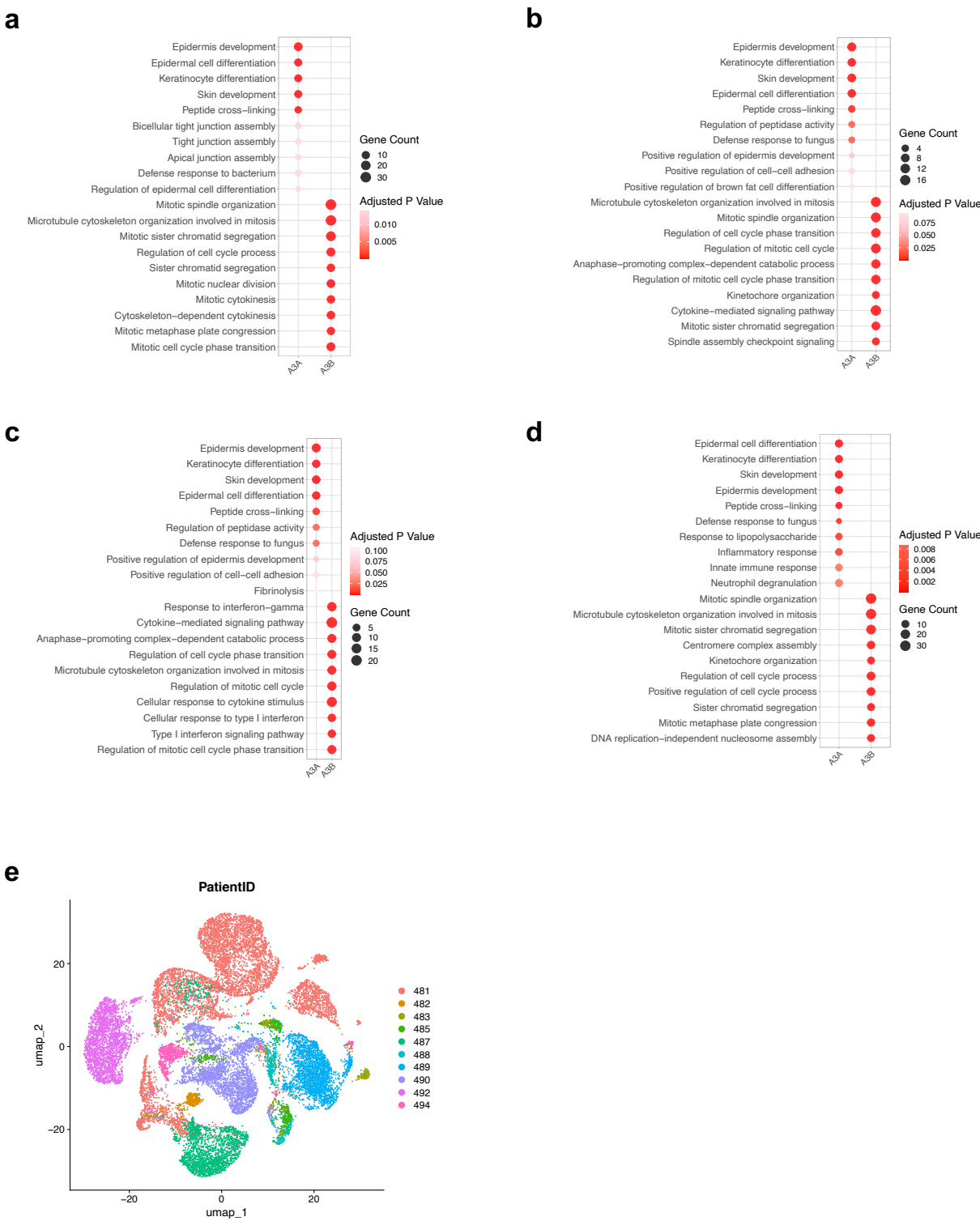
Supplementary Figure 3. SCENIC binary GRHL3 activity (on/off) shows activity in proliferating, differentiating and terminally differentiated clusters in healthy tonsil epithelium. **a)** distribution of GRHL3 activity scores across the dataset. Vertical red line represents the on/off threshold. **b)** UMAP showing binary activity calls for GRHL3 (blue = 'on', grey = 'off'). Accompanies Figure 4.



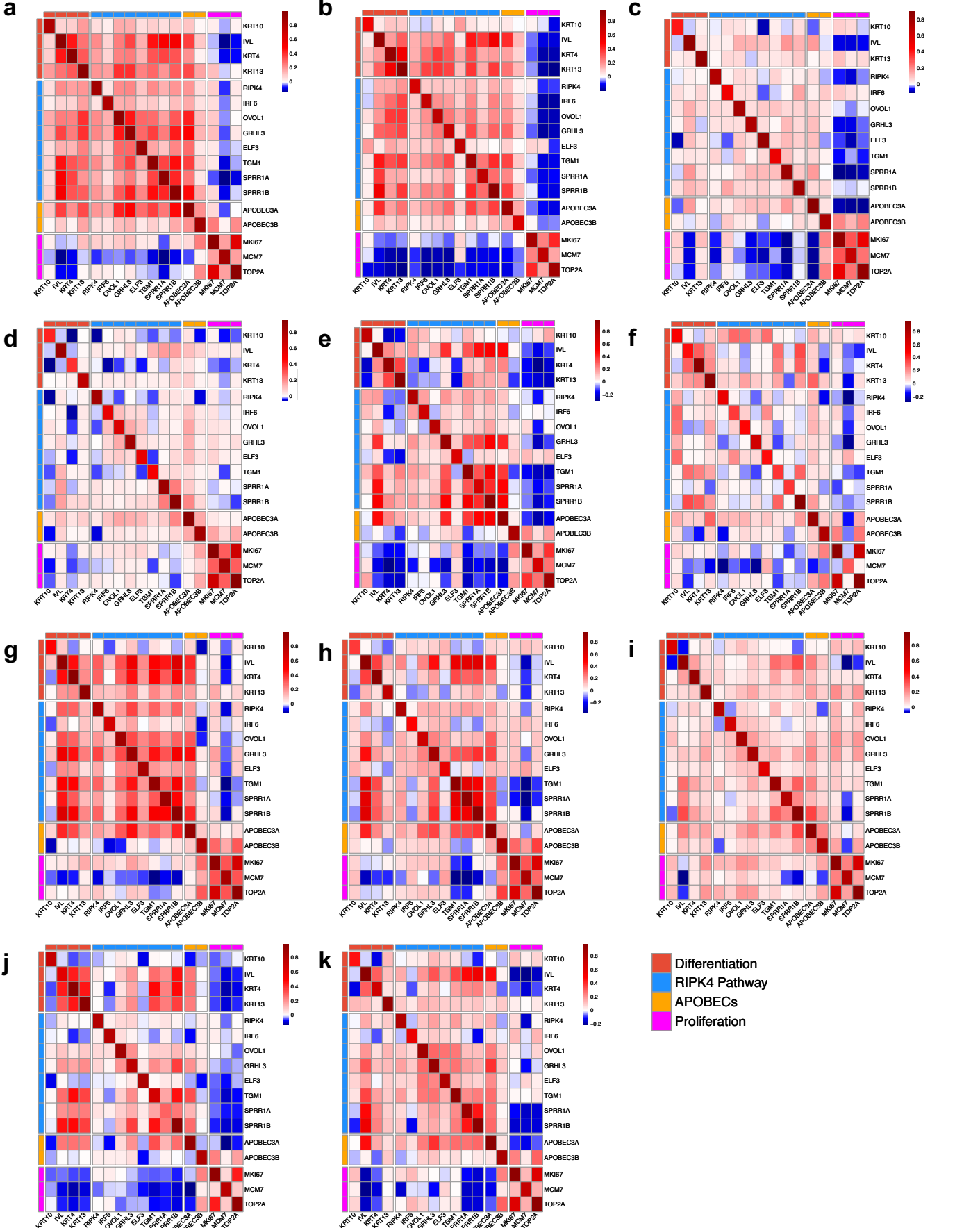
Supplementary Figure 4. *ELF3* is induced under the same conditions as *APOBEC3A* in cultured keratinocytes. qRT-PCR for *ELF3* in proliferating NIKS (async) or following 48 hours of growth factor and serum starvation (**a**), 24 hours of 100 nM afatinib treatment (**b**), growth to high cell density (**c**) or following 24 hours of treatment with inhibitors of MEK1/2 (trametinib), ERK1/2 (ravoxertinib), PI3K (pictilosib), AKT (MK2206) or mTORC1 (everolimus) (**d**). Async cells in panels b and d were treated with vehicle control (DMSO) for 24 hours. N = 3, error bars represent SEM. ** = p-value < 0.01; **** = p-value < 0.0001. Relates to Figure 3 a-c and f. Accompanies Figure 4. Pairwise comparisons in (a) and (b): two-tailed, unpaired t-tests; pairwise comparisons to Day 3 value in (c) and DMSO control treatment in (d): one-way ANOVA with Dunnett's multiple comparisons test.



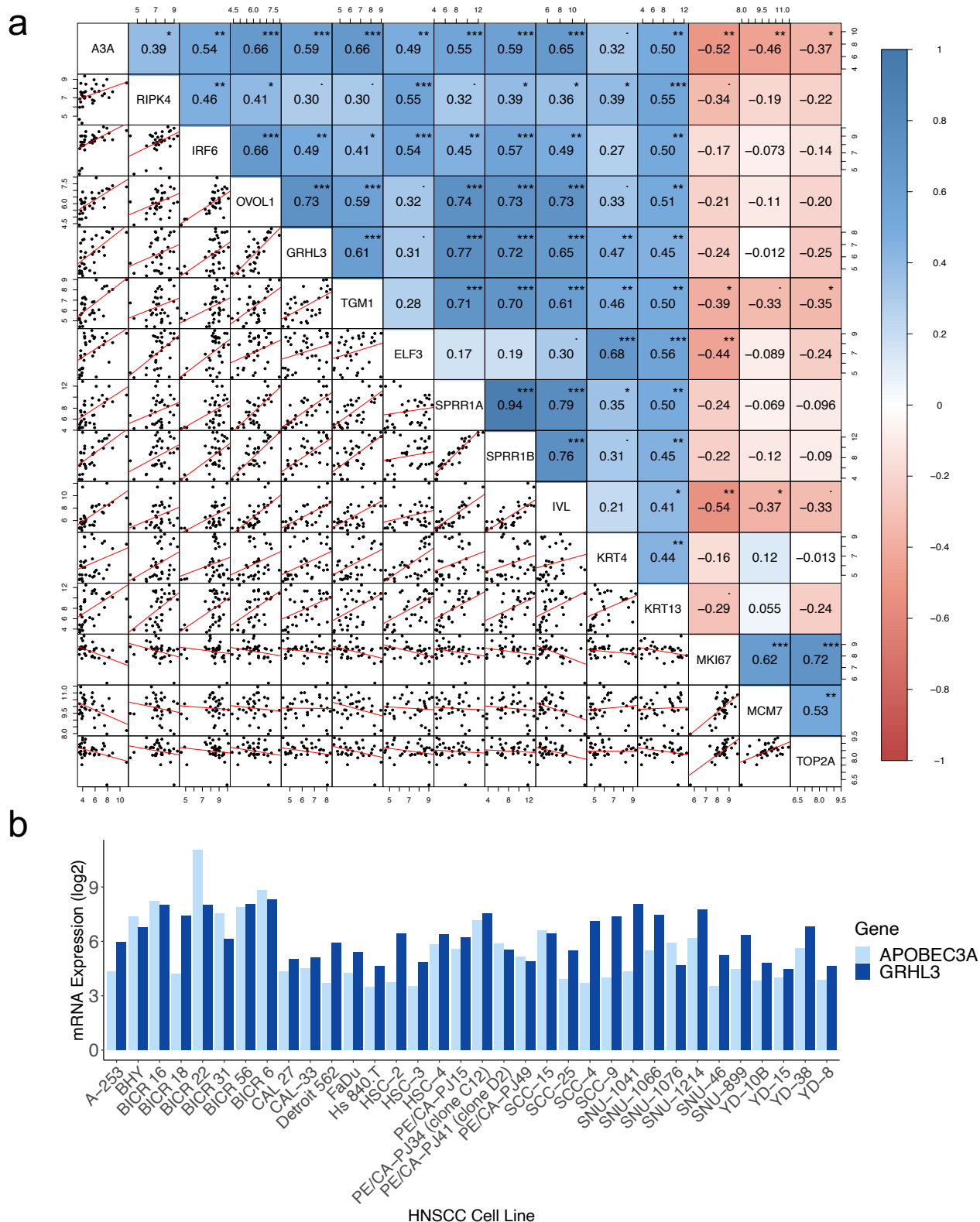
Supplementary Figure 5. *GRHL3* knockdown does not affect levels of the proliferation markers, *MKI67* and *MCM7* in afatinib-treated NIKS. Histograms showing expression of *MKI67* and *MCM7*, in NIKS transfected with control siRNA (NC#1), or two different siRNAs targeting *GRHL3* as indicated. Cells were treated with 100 nM afatinib for 24 hours prior to harvesting to induce differentiation. N = 3, error bars represent SEM. Gene expression in NIKS transfected with siRNAs targeting *GRHL3* was compared with control siRNA-transfected NIKS by one-way ANOVA with Dunnett's multiple comparisons test.



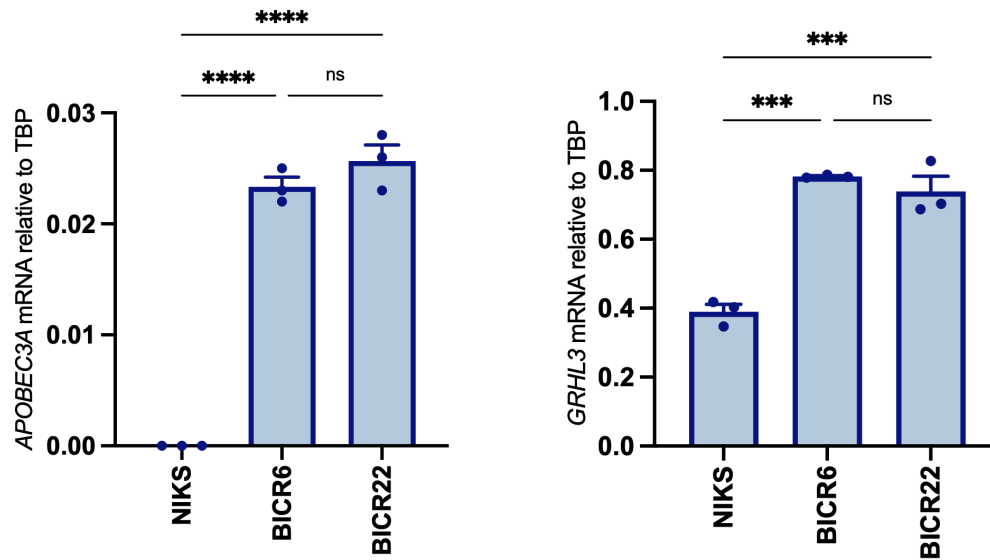
Supplementary Figure 6. HNSCC and ESCC display similar expression patterns of *APOBEC3A* and *APOBEC3B* to those observed in the healthy tonsil epithelium. Top 10 GO biological pathways related to the top 100 genes co-expressed with *APOBEC3A* or *APOBEC3B* for datasets **a)** HNSCC (Southampton), **b)** HNSCC validation dataset 1 (Kürten *et al.* 2021), **c)** HNSCC validation dataset 2 (Puram *et al.* 2023), and **d)** ESCC validation dataset. **e)** UMAPs showing clustering of the Southampton HNSCC dataset by patient.



Supplementary Figure 7. APOBEC3A/B co-expression with differentiation, proliferation, and RIPK4 pathway marker genes in the Southampton HNSCC scRNA-seq dataset. Heatmaps showing COTAN co-expression coefficients (red = positively co-expressed, blue = negatively expressed) for **a)** all 10 cases combined and from individual HPV negative tumours **b)** HN481 (7,032 cells), **c)** HN487 (2,129), **d)** HN492 (2,407), and HPV positive tumour patients **e)** HN482 (354), **f)** HN483 (158), **g)** HN485 (592), **h)** HN488 (621), **i)** HN489 (2,411), **j)** HN490 (2,932), and **k)** HN494 (678). Accompanies Figure 5.



Supplementary Figure 8. *APOBEC3A* is positively correlated with differentiation marker genes and negatively correlated with proliferation marker genes in HNSCC cell lines a) Correlation of *APOBEC3A* expression with the expression of differentiation marker genes, proliferation marker genes, and genes of the RIPK4 pathway in 34 HNSCC cell lines from the CCLE. b) Log2 expression levels of *APOBEC3A* and *GRHL3* in the 34 individual HNSCC cell lines from the CCLE. Accompanies Figure 5.

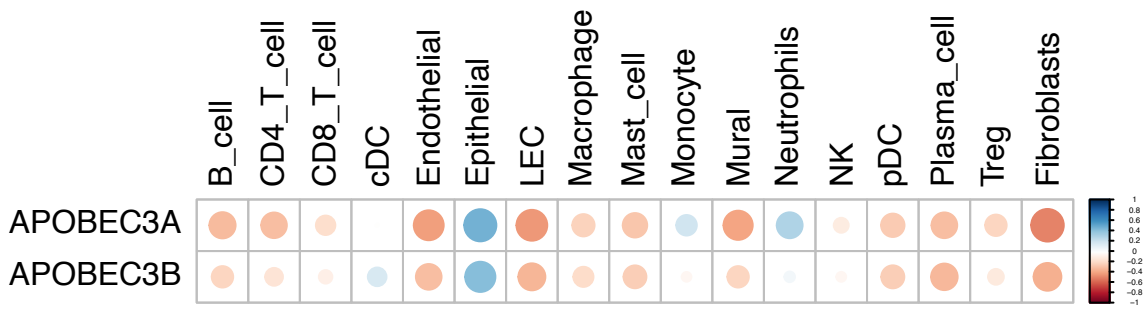


Supplementary Figure 9. *APOBEC3A* and *GRHL3* expression in sub-confluent, asynchronous NIKS, BICR6 and BICR22 cell lines. a) Histograms showing expression of *APOBEC3A* and *GRHL3*, in NIKS, BICR6 or BICR22 cells. Cells were grown to 60-70% confluency and media was replenished 18 hours prior to harvesting. N – 3, * = p-value < 0.05; ** = p-value < 0.01; *** = p-value < 0.001; **** = p-value < 0.0001. One-way ANOVA with Dunnett’s multiple comparisons test. Accompanies Figure 5.

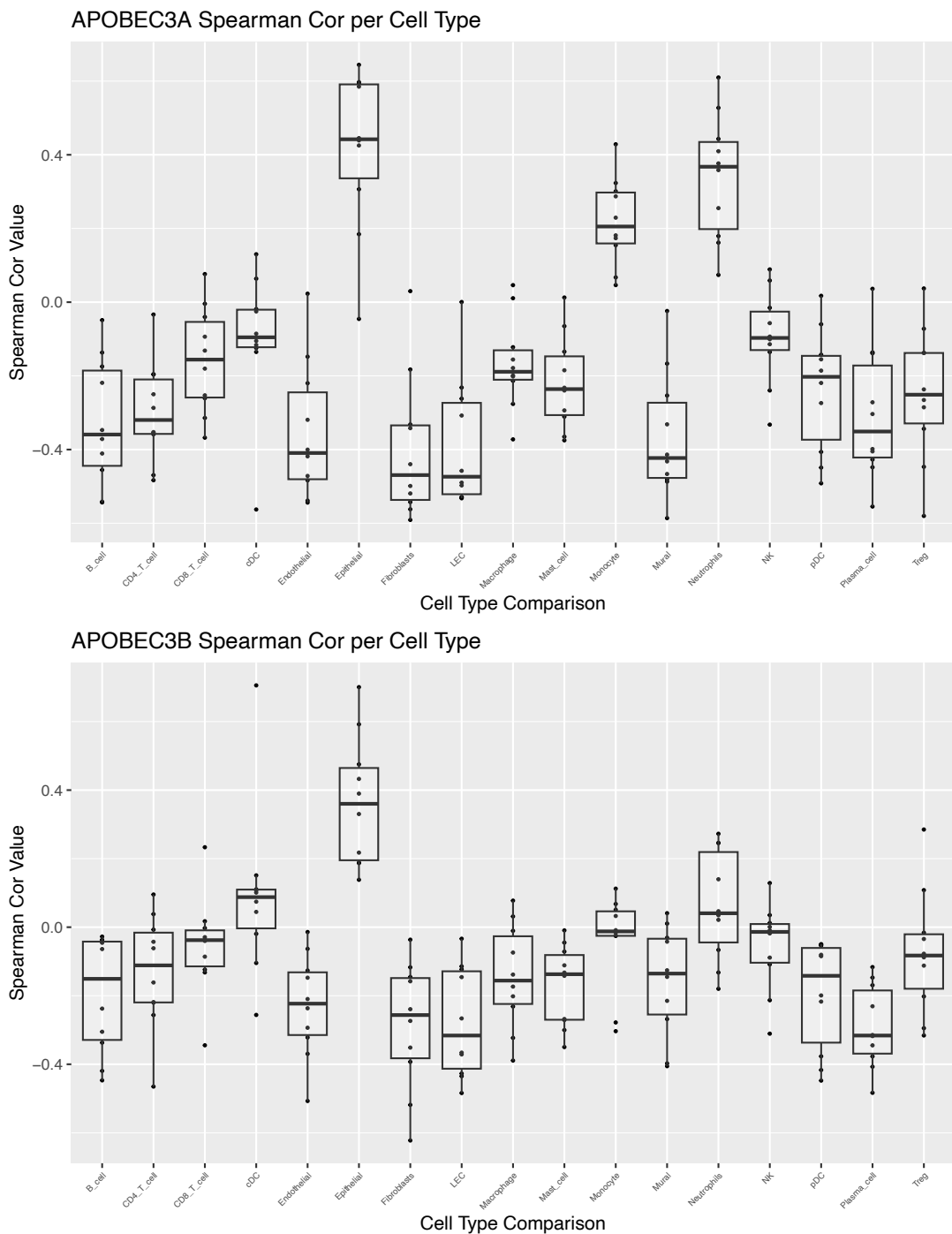


Supplementary Figure 10. Pathway analysis using top 100 genes co-expressed with *APOBEC3A* and *APOBEC3B* in Visium spatial data. Top 10 GO biological processes pathways for *APOBEC3A* and *APOBEC3B* considering the top 100 genes with the highest correlated expression for each gene.

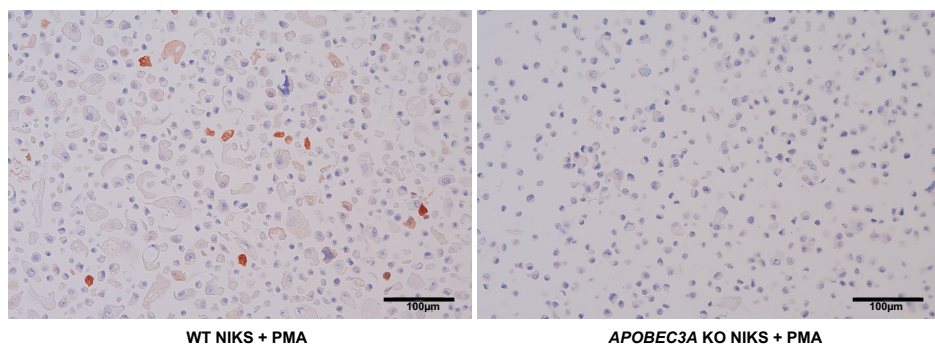
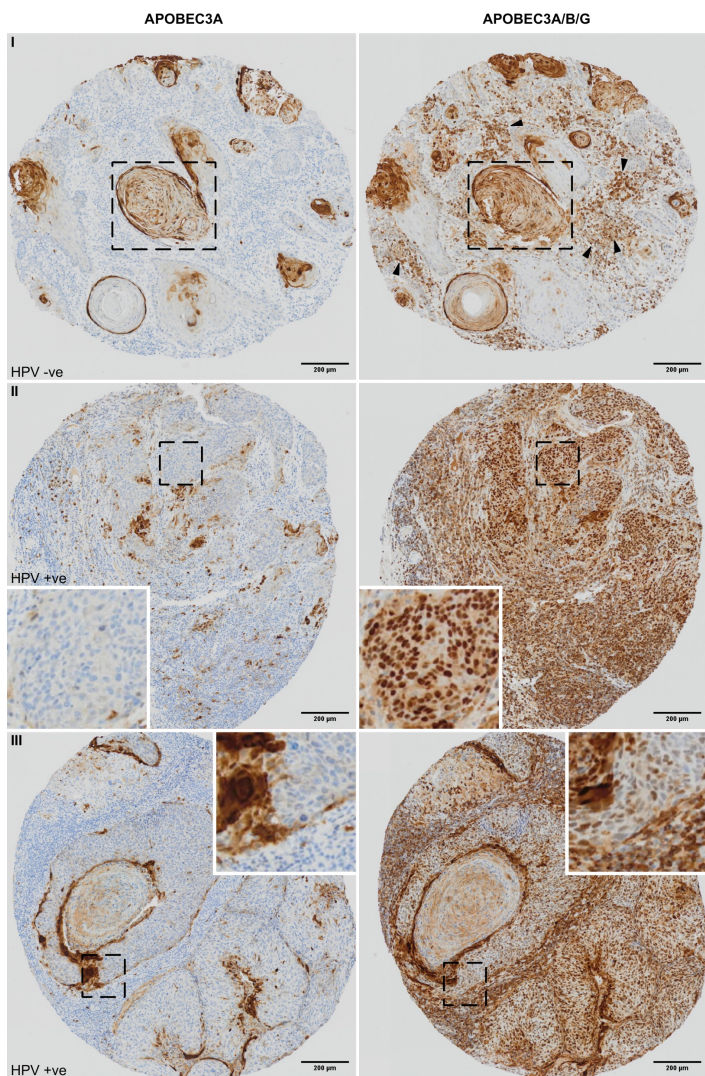
a



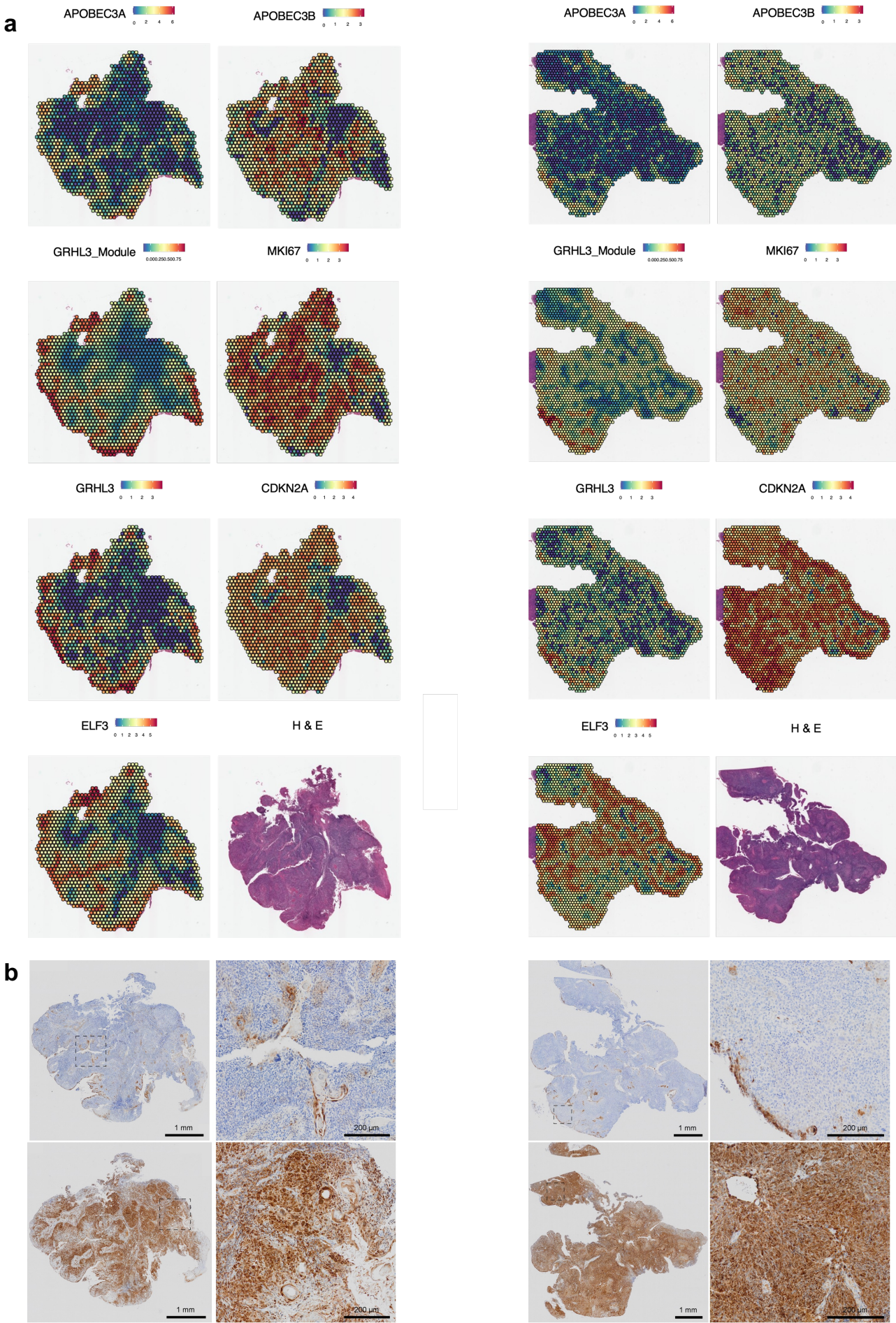
b



Supplementary Figure 11. Deconvolution of spatial transcriptomics data showing the cell types in which APOBEC3A and APOBEC3B are expressed within each spot. Accompanies Figure 5. a) Spearman's correlation for APOBEC3A/B expression and cell-type deconvolution score for every spot in every patient (all correlation coefficients were significant, $p < 0.0001$). b) Boxplot showing Spearman's correlation for APOBEC3A/B expression and cell-type deconvolution score. Each point represents the correlation coefficient value for an individual patient.

a**b**

Supplementary Figure 12. Detection of APOBEC3A and B proteins by immunohistochemistry in formalin-fixed, paraffin-embedded cells and tissues. (a) Wild-type (WT) or *APOBEC3A* knockout (KO) NIKS were treated with phorbol 12-myristate 13-acetate (PMA) for 24 hours prior to cell harvesting. Cell pellets were paraffin-embedded and formalin-fixed prior to sectioning and staining as described in Materials and Methods. (b) Representative images from an HNSCC tissue microarray stained with APOBEC3A-specific (left panels) or APOBEC3A/B/G-specific (right panels) antibodies. Accompanies Figure 5.



Supplementary Figure 13. Expression patterns of selected transcripts from spatial transcriptomics (a) and APOBEC3 protein levels assessed by IHC (b) in HNSCC cases HN482 (left) and HN494 (right). Accompanies Figure 5.