

# Protein Translation Dysregulation and Immune Cell Evasion Define Metastatic Clones in HPV-related Cancer of the Oropharynx

5

Venessa T Chin<sup>1,2,3</sup>, Walter Muskovic<sup>1</sup>, Rachael A McCloy<sup>1</sup>, Drew R Neavin<sup>1</sup>, Jose Alquicira-Hernandez<sup>1</sup>, Himanshi Arora<sup>1</sup>, Anne Senabouth<sup>1</sup>, Patricia Keith<sup>1</sup>, Ellie Spenceley<sup>1</sup>, Angela  
10 Murphy<sup>1</sup>, Dominik Kaczorowski<sup>1</sup>, Peter Floros<sup>2</sup>, Peter Earls<sup>2</sup>, Brett Leavers<sup>2</sup>, Julia Crawford<sup>2,3</sup>,  
Richard Gallagher<sup>2,4</sup>, Joseph E. Powell<sup>1,3,5,\*</sup>.

1. Translational Genomics, Garvan Institute of Medical Research, Darlinghurst, Sydney,  
NSW, Australia
2. The Kinghorn Cancer Centre, St Vincent's Hospital Sydney, Darlinghurst, Sydney,  
15 NSW, Australia
3. Faculty of Medicine, University of New South Wales, Kensington, NSW, Australia
4. Faculty of Medicine, University of Notre Dame, Sydney, NSW, Australia
5. Cellular Genomics Futures Institute, University of New South Wales, Kensington,  
Sydney, NSW, Australia

20 \* Correspondence: j.powell@garvan.org.au

**Keywords:** Oropharyngeal cancer, HPV, clonal evolution, metastasis, tumour microenvironment, translational relief.

# Summary

Head and neck cancers, representing the seventh most common malignancy globally, have seen a shift in causative factors from traditional smoking and alcohol use to human papillomavirus (HPV) infection, now accounting for up to 80% of oropharyngeal cancers. We identify the cellular and clonal mechanisms underlying immune avoidance and metastasis by analysing single-cell and spatial genomic data from primary and metastatic cancers. We first map the clonal evolution of malignant cells based on the accumulation of mutations. We identify metastasising clones based on mutational similarity scores between cells in the primary and lymph node metastasis. Genomic analysis of metastasising and non-metastasising clones identified virally mediated protein translation relief ( $P=4.24 \times 10^{-24}$ ) pathway underlying metastatic expansion. We show that in metastatic clones, this process is driven through upregulation of translation-initiating factors, *EIF4E* ( $P=1.5 \times 10^{-13}$ ) and *EIFG1* ( $P<2.22 \times 10^{-16}$ ), and suppression of regulatory kinases *EIF4EBP1* ( $P=2.1 \times 10^{-10}$ ), *EIF2AK2* ( $P<2.22 \times 10^{-16}$ ), and *EIF2S1* ( $P<2.22 \times 10^{-16}$ ). We subsequently identify that metastatic clones have a corresponding downregulation of the *JAK/STAT* pathway and immunoproteasome genes *PSMB8* ( $P<2.22 \times 10^{-16}$ ) and *PSMB9* ( $P<2.22 \times 10^{-16}$ ), suggesting these clones escape immune surveillance through decreased *INF* inflammatory response and antigen presentation. We validate these results using spatial RNA-seq data, where metastatic cancer clones show decreased cell-to-cell interactions with CD4 T-effector memory cells (CD4<sub>TEM</sub>) ( $P=0.0077$ ), CD8 T-exhausted cells (CD8Ex) ( $P=0.0191$ ), and innate lymphoid cells (ILC) ( $P=0.04$ ). Finally, we demonstrate that the upregulation of cap-independent translational drives cell proliferation in metastatic clones through the expression of translation initiation factors (*EIF4G1*:  $P<2.22 \times 10^{-16}$ ). Our results provide evidence of the mechanisms by which virally induced cancer clones lead to advanced disease and poor prognosis in patients.

# Introduction

Cancers of the head and neck comprise a diverse range of cancers, which represent the 7<sup>th</sup> most common malignancy worldwide, with approximately 900,000 new cases and 450,000 deaths per year<sup>1</sup>. Historically, cancers in the oropharynx region were commonly associated with smoking and alcohol use. However, infection with human papillomavirus (HPV) is becoming an increasingly common cause and is now responsible for up to 80% of oropharyngeal cancers<sup>2</sup>. The HPV-positive oropharyngeal cancer (HPV+) patient population differs from HPV-negative (HPV-), with earlier age of onset, smaller primary tumours, and larger lymph node burden but improved survival outcomes<sup>3-5</sup>. The cellular mechanisms underlying these clinical differences in HPV+ and HPV- patients are poorly understood<sup>6,7</sup>, resulting in missed opportunities to understand the defining differences in cancer cell behaviour, metastasis and treatment response. Despite aggressive therapies, up to 25% of patients with potentially curable HPV+ oropharyngeal cancer experience disease relapse, after which most patients die within two years<sup>8</sup>. To reduce the risk of relapse, adjuvant radiation given with or without chemotherapy can be offered. However, identifying patients who will benefit the most from these toxic therapies with high-risk disease remains difficult without accurate prognostic biomarkers<sup>9</sup>.

Cancer cells are genetically heterogeneous, leading to cancer phenotypes that are transcriptionally and functionally distinct from one another<sup>10</sup>. Primary tumours typically exhibit greater clonal diversity than metastases<sup>11</sup>. This observation is an emergent property because metastatic processes are partly driven by cells acquiring genetic mutations that impact their phenotypes through clonal selection within their microenvironmental context<sup>12-15</sup>. Identifying the phenotypic characteristics and molecular mechanisms of clones prone to

metastasis not only offers a chance to identify biomarkers indicating a poor prognosis but also holds the potential for uncovering novel drug targets.

75 In HPV- oropharyngeal cancer, the epithelial-to-mesenchymal transition differentiates between primary and metastatic cells<sup>16</sup>. Moreover, *AXL*, known for its role in cell migration, EMT, invasion, and proliferation<sup>17</sup>, along with *AURKB*, a pivotal regulator of mitosis<sup>18</sup>, have both been recognised as significant contributors to the development of invasive metastatic cell phenotypes in HPV- oropharyngeal cancers.

80

In HPV+ oropharyngeal cancers, there is limited understanding of the cellular pathways underlying metastasis, which are expected to differ from those in HPV- cancers given the unique clinical and aetiological factors. The inability to clear HPV infections stems from complex mechanisms involving the suppression of the host cell's antiviral response (involving  
85 *NFKB*, *STING*, and *IFN*), the sustained viral replication facilitated by the retinoblastoma tumour suppressor (*RB*), and evasion of immune surveillance by inhibiting apoptosis via *TP53*<sup>19</sup>.

To sustain viral replication, HPV employs various mechanisms to manipulate the host cell. Specifically, viral proteins E6 and E7 incapacitate key tumour suppressor regulators like  
90 *TP53*<sup>20</sup> and *RB*<sup>21</sup>, disrupting apoptosis and cell cycle regulation. Furthermore, E6 directly interferes with DNA damage repair processes<sup>22</sup>. While these strategies effectively support viral replication, they render the host cell incapable of managing defective DNA damage repair, leading to a gradual increase in genomic instability over time<sup>22</sup>. Ultimately, this increasing genomic instability culminates in malignant transformation<sup>23</sup>. Simultaneously, the virus  
95 employs multiple mechanisms to evade the host's immune responses. This immune evasion enables infected cells to remain undetected and allows for the tolerance of emerging pre-

malignant and malignant lesions. Therefore, many mechanisms that maintain viral infection may also contribute to cancer cell evasion of anti-tumour responses, an area that has not been comprehensively explored.

100

Here, we present an analysis of primary and regionally metastatic cancers in patients with HPV+ oropharyngeal cancer who have undergone curative surgical resection. Using single-cell techniques, we use copy number variation (CNV) to define cancer clones and study clonal population phenotypes with regard to metastasis and microenvironmental dynamics. We identify virally mediated protein translation relief and dysregulation of immunoproteasomal and IFN pathways, accompanied by a shift towards cap-independent translation underlying metastatic clonal selection. Furthermore, our study validates immune cell evasion by sequencing matched tumour samples *in situ*.

110

# RESULTS

## Evidence of clonal selection favouring phenotypic predisposition to metastasis

It has long been recognised that the genomic profiles of cancer clones influence cellular phenotypes, impacting clinical outcomes<sup>24</sup>. Using haplotypes and single-cell RNA-sequence (scRNA) data, we called CNVs from matched primary and regional lymph node metastases from four HPV+ patients (**Figure 1A, Figure 1B, Supplementary Table 1**)<sup>25</sup>. An overview of the CNV distribution across all individual cells highlights various genomic alterations indicative of the genomic instability typically observed in cancers (**Figure 2**). We did not observe a consistent enrichment of oropharyngeal cancer oncogenes identified by the cancer genome atlas (TCGA)<sup>26</sup> in the CNV regions (*PIK3CA*, *TRAF3*, *E2F1*) (**Figure 2**). A possible explanation is that TCGA data comprises HPV+ and HPV- cancer samples. Nevertheless, the data illustrates the heterogeneous nature of HPV+ cancer across patients. To understand clonal dynamics, subclonal phylogenies were generated by *Numbat*<sup>27</sup> by aggregating data from single cells into a pseudobulk profile, capturing the evolutionary relationships between cells. From 14,157 single-cell CNV profiles, three to seven distinct clusters per patient were identified (**Figure 3A**). Each cluster, suggesting a unique cancer clone, exhibited a characteristic CNV signature. By constructing a phylogenetic tree based on shared CNV events<sup>25</sup>, we identified the ancestral clones and genomic evolution of clones by acquiring new somatic events. As expected, the clonal evolution of cancers was unique for each patient and is summarised in detail in **Figure 3**. Using patient four as an example, we observe the development of clone two arising from the original clone (ancestral) through deletions at 3p and 15p. Subsequent deletions at 10q and 11q and a loss of heterozygosity (LOH) at 20q led to clone three, followed by a divergent event into clones four and six through a new LOH at four (clone six) and amplifications at 3q and 5q plus a deletion at 7q (clone four). We observe a final clonal

evolution event, with clone five arising from clone four via somatic mutations at 4p<sup>+</sup>, 8q<sup>-</sup>, 19<sup>+</sup> and 20p<sup>+</sup> (**Figure 3**). Similar patterns of clonal evolutionary events are observed across patients.

We next sought to understand the clonal diversity in primary and lymph node sites and identify cancer clones that show significant expansions into the metastatic lymph node sites. As shown in multiple cancers<sup>11</sup>, the clonal diversity was greater in the primary site compared to the lymph metastasis (**Figure 3C**). This difference in diversity can be due to several factors, such as selection pressures on cancer clones, where only a subset of cells from the primary tumour possess the necessary genomic and phenotypic characteristics required to metastasise; tumour microenvironment and adaptation, where differences in immune evasion in the primary and metastatic site favour growth and expansion of specific clones. We observed significant differences in the clonal proportions between the primary and lymph node sites, demonstrating clonal expansion into the metastatic site. This enabled us to identify genomic factors underlying metastatic behaviour. The results revealed a clear pattern, with one or two clones displaying substantial expansion from rare populations in the primary to dominant clones in the metastatic lymph node (**Figure 3**). We refer to these clones as expanding or metastatic clones. Correspondingly, clones common in the primary site are frequently absent or observed in very low frequencies in the lymph node site, suggesting an absence of metastatic phenotypes. We subsequently analysed the genomic differences between these clones to identify pathways and microenvironmental contexts that contribute to the metastatic phenotype.

**Virally-induced dysregulation of protein translation is a defining characteristic in metastasising cancer clones**

In most cases, cancer mortality is primarily attributed to metastatic disease<sup>28</sup>, due to the emergence of cancer cell clones with aggressive phenotypes through a selective process. To identify the genomic phenotypes of expanding clones, we overlapped the clonal data with cell clustering to delineate the expanding and non-expanding cancer cells (**Figure 4A**) and identify genes differentially expressed. We identified significantly differentially expressed between the expanding and non-expanding clones (**Supplementary Table 2**). We performed pathway enrichment on this set of genes (**Figure 4B**). Our results identify three broad cellular pathway groups represented by genes whose expression levels significantly differ between metastatic expanding clones and non-expanding clones: cell cycle, immune regulation, and protein translation (**Figure 4B** and **Supplementary Table 3**). In the top ten enriched pathways for each patient comparison, protein translation pathways show consistent downregulation in the expanding compared to non-expanding clones (**Figure 4, Supplementary Figures 1-6**). For example, the pathway “Formation of a pool of free 40S subunits” is enriched in comparison A for patient 1 with 59 differentially expressed genes from this pathway ( $P$ -value for enrichment =  $4.24 \times 10^{-23}$ ). **Supplementary Figure 7** shows that the genes found in this pathway in comparison 1A consist of protein translation initiating factors and ribosomal subunit protein genes. Further analysis (**Supplementary Figure 1**) demonstrates that many genes identified in comparison 1A within this pathway are downregulated, like *RPL35A* ( $\text{Log}_2(\text{FC}) = -0.89$ ,  $P = 9 \times 10^{-251}$ ) and *RPS26* ( $\text{Log}_2(\text{FC}) = -0.59$ ,  $P = 1.45 \times 10^{-107}$ ), indicating that expanding clones are downregulating protein translation.

The consistent downregulation of protein translational machinery in expanding clones is accompanied by a small but significant decrease in the expression levels of ribosomal protein-encoding genes in the expanding clones ( $P=1.4 \times 10^{-6}$ ) (**Figure 4C**). The presence of 18 enriched protein pathways, as revealed in **Figure 4B**, all characterised by an abundance of

protein translation initiation factors and ribosomal protein genes, collectively underscores the pivotal role of protein translation downregulation as a defining phenotypic factor in expanding clones.

To understand the relationship between viral infection and dysregulation of protein translation in metastatic expanding clones, we tested for the relative expression levels of translation-initiating factors and regulatory kinases associated with virally activated translational relief, a process by which virally infected cells escape normal anti-viral mechanisms<sup>29,30</sup>. Under normal circumstances, a cell will downregulate protein translation and upregulate apoptosis once a virus is detected. We identified significant upregulation of two translation-initiating factors, *EIF4E* ( $P=1.5 \times 10^{-13}$ ) and *EIFG1* ( $P<2.22 \times 10^{-16}$ ), in expanded cancer clones of the lymph node compared with non-expanding clones and corresponding significant downregulation of regulatory kinases *EIF4EBP1* ( $P=2.1 \times 10^{-10}$ ), *EIF2AK2* ( $P<2.22 \times 10^{-16}$ ), and *EIF2S1* ( $P<2.22 \times 10^{-16}$ ) (**Figure 5, Supplementary Table 4**). Further evidence of translational relief is seen with upregulation of the stress kinases *PPP1R15A* ( $P<2.22 \times 10^{-16}$ ), *PPP1CA* ( $P=7.9 \times 10^{-6}$ ) and *PPP1CB* ( $P<2.22 \times 10^{-16}$ ) and downregulation of apoptosis-associated genes *TP53* ( $P<2.22 \times 10^{-16}$ ), *FAS* ( $P=2.9 \times 10^{-8}$ ) and *BAX* ( $P=1.3 \times 10^{-9}$ ) (**Figure 5, Supplementary Table 4**). Dysregulation of protein translation has previously been described in both viral infection<sup>29,31,32</sup> and cancer<sup>33</sup>. After an acute viral infection, the host cell attempts to reduce virus replication through a process called “host shut-off”<sup>34,35</sup>, where ribosomal proteins involved in mRNA translation are downregulated and apoptosis is triggered. In normal circumstances, this process limits viral replication and induces apoptosis in infected cells, thereby containing the infection. However, in persistent viral infection like HPV, these mechanisms fail to eradicate the virus. Our results provide strong evidence of virally induced translational relief in expanding clones, where host cells’ shut-off mechanisms are normalised using stress kinases to keep EIF2AK in

its dephosphorylated state. This maintains protein translational machinery and viral replication and turns off apoptotic pathways.

# **Virally-induced immunosuppression through interferon and immunoproteasome pathways**

In HPV+ cervical cancer, viral infection triggers significant alterations in the microenvironment<sup>36</sup>. During persistent infection, there is a notable decline in the presence of cytotoxic CD8+ and helper CD4+ T-cells within the epithelial tissue. In the precancerous stage, cytotoxic T-cell populations maintain an immunosuppressed profile, characterised by increased regulatory T-cells. Invasive cancers demonstrate a high immunogenic state marked by elevated levels of cytotoxic and helper T-cells and occasionally include a surge in regulatory T-cell numbers. Notably, abundant regulatory T-cells are associated with a more unfavourable prognosis<sup>36</sup>. Viral infection generally results in immune evasion via suppression of IFN-stimulated gene transcription<sup>32</sup> and, in HPV-infection, downregulation of the immunoproteasome<sup>30</sup>. Silencing the IFN regulatory pathway has also been demonstrated to be a feature of metastatic cancer cells<sup>37</sup>. We next sought to understand the consequences of HPV infection on immune evasion and its relationship to clonal expansion. Comparing the expression levels of IFN-inducible genes, we observed a downregulation of the IFN regulatory pathway in expanding metastatic cancer clones compared to non-expanding clones (**Figure 5, Supplementary Table 4**). The effect is observed in genes involved in the *JAK/STAT* pathway, with *JAK1* ( $P=4.5 \times 10^{-7}$ ), *JAK2* ( $P=0.0072$ ), *STAT1* ( $P=1.5 \times 10^{-7}$ ), and *STAT2* ( $P=0.031$ ) all downregulated amongst the expanding clones relative to non-expanding ones. In support of this, we see consistent downregulation of regulators of antiviral immune response *MXI* ( $P=5.2 \times 10^{-12}$ ), *ISG15* ( $P<2.22 \times 10^{-16}$ ), and *IFIT3* ( $P=2.5 \times 10^{-10}$ ) amongst expanded cancer clones. Furthermore, the immunoproteasome genes *PSMB8* ( $P<2.22 \times 10^{-16}$ ) and *PSMB9*

( $P < 2.22 \times 10^{-16}$ ) are downregulated, along with the associated regulatory protein *TAP1*. These results suggest that through mechanisms to dampen the IFN inflammatory response and reduce antigen presentation, expanding clones can avoid the host immune response, metastasise and expand.

## Up-regulation of cap-independent translational pathways is related to metastatic clonal behaviour.

MRNA production in healthy cells involves cap-dependent translation mediated by the protein complex EIF4F. However, in cancer cells, cap-dependent translation is suppressed in favour of cap-independent translation, mediated by EIF4G1. This transition enables the sustained gene expression that promotes cell survival and proliferation<sup>38</sup>. Compared with healthy cells, we observe a higher ratio of *EIF4G1:EIF4F* as expected. However, we hypothesise that cap-independent translation pathway variation would impact metastatic cell phenotypes. To evaluate this, we compared the expression levels of cap-dependent translation genes between the expanded metastatic clones and non-expanding clones. The results display a clear pattern of upregulation of *EIF4G1* ( $P < 2.22 \times 10^{-16}$ ), coupled with the downregulation of genes *RPS2* ( $P < 2.22 \times 10^{-16}$ ), *RPS15* ( $P = 1.1 \times 10^{-8}$ ), *RPL23A* ( $P < 2.22 \times 10^{-16}$ ), and *RPL35A* ( $P < 2.22 \times 10^{-16}$ ) controlled by cap-dependent translation (**Figure 5**). Additionally, we observed the upregulation of genes involved in cancer cell proliferation and angiogenesis, with *HIF1A* ( $P = 0.00012$ ), *VEGFA* ( $P < 2.22 \times 10^{-16}$ ), and *PCBP1* ( $P = 4.4 \times 10^{-7}$ ) all upregulated in expanded versus non-expanded clones. Interestingly, whereas in other cancers, acquisition of somatic mutations has been linked to dysregulation of initiating factor genes<sup>39</sup>, we do not observe any evidence for that here (*EIF4A1*, *EIF4E3*, *EIF4G1*, *EIF4G2*, *EIF4EBP1*; **Figure 3**).

Our results identify the relationship between genomic phenotypes and clonal expansion into lymph node metastases, with expanding clones displaying a combination of dysregulated pathways favouring their survival. These can be summarised as the downregulation of cap-dependent translation and apoptosis, alongside the upregulation of cap-independent translation, co-expression with genes associated with cancer cell proliferation and metastatic processes, and evasion of immune response through the proteasome and IFN pathways.

### **Evasion of immune cells within the primary tumour is a distinct cellular phenotype of metastatic clones.**

Our analysis of expanding clones that result in clonal dominance in the lymph node metastasis identified immune evasion through virally induced interferon and immunoproteasome pathways. We generated in situ spatial RNA-seq data from matched primary and lymph node biopsies to provide supportive evidence for this observation. We interrogated the relationship between cancer clonal subtypes and immune cell interactions within the cancer microenvironmental context. Numerous studies have highlighted the importance of the immune microenvironment within the primary site in influencing the metastatic potential of cancer cells<sup>37,40</sup>, and we aimed to test this in our cohort.

After classifying cancer clones and immune cell subtypes (**Methods**), we tested for the correlation in co-occurrence of pairwise combination of cells. We identified distinct differences between the T-cell correlation with expanding and non-expanding cancer clones (**Figure 6**). Expanding, metastatic clones predominately exhibit negative or neutral correlations ( $-0.361 < R < 0.01$ ) with T-cell subsets compared to more positive correlations in the non-expanding clones ( $-0.334 < R < 0.678$ ). This observation is more pronounced within primary tumours. For example, the mean correlation in the primary between CD8 Exhausted cells

(CD8Ex) and expanding and non-expanding clones (non-expanding = 0.599; expanding = -  
290 0.12) is more disparate than in the lymph node (non-expanding = 0.295, expanding = 0.01). To  
evaluate this further, we tested for a significant difference in the median correlation between  
the expanding and non-expanding cancer clones across all patients. Adjusting for multiple  
testing, we show a significant difference in the co-occurrence of three T-cell subtypes between  
expanding and non-expanding cancer clones (**Figure 7A**). In the primary site, CD4 T-effector  
295 memory cells (CD4TEM) ( $P=0.0077$ ), CD8 T-exhausted cells (CD8Ex) ( $P=0.0191$ ), and innate  
lymphoid cells (ILC) ( $P=0.04$ ) show a significant decrease in their correlation with expanding  
verses non-expanding cancer clones. CD4TEM cells also demonstrate a significant difference  
in correlation ( $P=0.021$ ) with non-expanding clones compared to expanding clones in the  
lymph node site (**Figure 7A**). We observe a similar pattern of relationships in B-cell subsets  
300 (**Supplementary Figure 8**). The correlation analysis identified a distinct trend where  
expanding clones exhibit negative or neutral correlations with B-cells, in contrast to the  
behaviour observed in non-expanding clones, which have consistently strong co-occurrences.  
Summary correlations using mean correlation scores (**Figure 7B**) confirm this pattern. Notably,  
within the primary tumour, a substantial disparity exists in the correlations between expanding  
305 and non-expanding clones with B-memory cells and plasmablasts. Similarly, significant  
differences emerge in the correlations with B-memory ( $P=0.00048$ ) and B-naïve ( $P=0.0065$ )  
cells within the lymph node. In the monocyte populations CD14 monocytes (CD14Mono),  
classical dendritic cells (cDC), plasmacytoid dendritic cells (pDC), a similar relationship to  
that observed in T and B-cell subsets (**Supplementary Figure 9**). The mean correlation scores  
310 (**Figure 7C**) identify a tendency to have positive correlations with monocytes and non-  
expanding clones, in contrast to the negative correlations noted with expanding clones. This  
trend is particularly pronounced in the lymph nodes, where significant differences between

expanding and non-expanding clones exist within the CD14 monocyte ( $P=0.0091$ ) and cDC subsets ( $P=0.0052$ ).

315

The *in situ* data results demonstrate clone-specific immune evasion within the tumour microenvironment in metastasising cells. The observed dysregulation in the interferon and immunoproteasome pathways identified in differential gene expression analysis offers a credible explanation for this observed phenotype.

320

## DISCUSSION

It is widely acknowledged that cancer cells exhibit a range of diverse phenotypic properties<sup>41</sup>. Moreover, distinct cancer clones within tumours significantly influence the responses to anti-cancer treatments, subsequently impacting overall survival outcomes<sup>42</sup>. In locally advanced  
325 cancers, surgical resection aims to remove all the macroscopically identified cancer, with the addition of adjuvant therapies like radiation and chemotherapy to reduce recurrence rates by eradicating microscopic residual disease<sup>43</sup>. It has been shown that clones present in recurrent cancer are often sparsely represented in the primary tumour and that clonal complexity decreases in metastases<sup>12,39</sup>. While the mechanisms are not always resolved, it is clear that  
330 clonal selection occurs for cell phenotypes that enhance survival and metastatic processes. In the case of HPV+ oropharyngeal cancer, persistent HPV infection leads to alterations in the host immune response<sup>30</sup>, which likely results in significant selection pressure on cells, which shapes the tumour microenvironment<sup>44</sup>. This is evident when observing immune cell microenvironment changes through acute and chronic viral infection phases and pre-cancer and  
335 invasive cancer development with HPV infection<sup>36</sup>, highlighting the importance of examining cancer clones in their spatial context. Considering that HPV infection serves as the primary driver of carcinogenesis, our findings delineating metastatic clones are understandably distinct

from those observed in the HPV- population<sup>16</sup>. Our initial results confirm observations presented in other cancers<sup>45</sup>, namely that there is reduced clonal complexity in metastases and that metastasising clones are sparsely represented in the primary tumour, confirming that clonal selection occurs. Through genomic analysis of cancer clones, we identified that the dysregulation of protein translation phenotypically defines expanding metastatic clones.

Given that protein translation dysregulation is seen in both acute<sup>32</sup> and chronic<sup>29,31</sup> viral infections and cancer<sup>38</sup>, we took a deeper look into the regulation of genes. We identified several important mechanisms that define expanding cancer clones in HPV+ oropharyngeal cancer. Our results suggest that virally induced translational relief is a mechanism that defines metastasising cancer clones in HPV+ oropharyngeal cancer. This observation aligns with prior work demonstrating HPV E6 protein is linked to translational control in infected fibrosarcoma cancer cell lines<sup>29</sup>, with E6 integration inducing the production of *EIF2AK2*, which in turn leads to phosphorylation of EIF2S1 and a decrease in protein translation and E6 production rapidly decreases. However, despite significant downregulation, the small amount of remaining E6 can initiate translational relief or rescue mechanisms. E6 can promote dephosphorylation of EIF2AK2 through stress kinase PPP1R15A, allowing protein translation to normalise and apoptosis to be inhibited<sup>29</sup>.

Furthermore, *PPP1R15A* is associated with tight regulation of the IFN-mediated anti-viral response<sup>46</sup>. In an acute viral infection, PPP1R15A is required to relieve protein translation inhibition to allow *IFN* and stress granule production in a pulsatile fashion. This allows the cell to restrict protein translation while releasing inflammatory mediators in periodic bursts. The dynamics of this process are different in chronic infection. When cells are exposed to stress, such as chronic viral infection<sup>47</sup>, PPP1R15A levels are continually elevated with fewer bursts, reducing stress granule release over time and creating an environment conducive to viral

maintenance. Our results show that metastasising clones have increased expression of PPP1R15A and decreased levels of IFN-inducible genes, suggesting that these compensatory mechanisms favour viral replication and are critical in defining the metastatic phenotype. In addition to its role in viral response, *PPP1R15A* has previously been implicated in carcinogenesis and cancer growth. In mouse models of colorectal cancer, PPP1R15A potentiates carcinogenesis by enhancing *IL-6* production and STAT4 activation<sup>48</sup>. It has also been shown to promote autophagy in liver cancer<sup>49</sup> and upregulate the RAC1-GTPase pathway in breast cancer<sup>50</sup>. This highlights a potential treatment target for the prevention of carcinogenesis in patients with chronic viral infection and established cancer.

Further perturbations of protein translation were identified in the metastatic cell phenotype, with the upregulation of cap-independent protein translation. Cap-independent mechanisms do not rely on the normal ribosomal machinery and, therefore, are not as susceptible to changes in functional subunit availability<sup>51</sup>. Thus, cancer clones that can utilise these cap-independent pathways to maintain protein translation may have an evolutionary advantage. Our data suggests that cap-independent translation maintains cancer-related proteins in the expanding metastatic clones. This observation is supported by cell-specific co-expression with cancer survival genes *VEGFA* and *HIF1A*. These results align with previous research in breast cancer models that have demonstrated how cap-independent mechanisms selectively translate mRNAs that encode proteins critical for cancer cell survival like VEGFC, HIF1A and BCL2<sup>52,53</sup>. Previous work shows that lung cancers rely on cap-independent mechanisms to maintain cell replication and that EIF4G1 is an adverse prognostic feature<sup>38</sup>. Given that cap-independent translation appears to be a mechanism underlying the selection of HPV+ oropharyngeal cancer metastatic clones, we propose this may be a useful biomarker to identify patients with poor prognostic disease.

Immune cell evasion was identified as a defining feature of metastatic cancer clones both as a cellular phenotype and through analysis of the tumour microenvironment. For HPV infection to persist, viral proteins E5, E6 and E7 mediate anti-inflammatory and immune-evading functions. For example, in human keratinocytes, E5 can actively suppress *IFN* and IFN-stimulated gene expression, key components of the innate anti-viral response<sup>54</sup>. Here, we show that the downregulation of immunoproteasomes is an important component of immune evasion in the expansion and metastasis of clones in HPV+ oropharyngeal cancer. The immunoproteasome is involved in antigen presentation and is a critical step in the inflammatory response to viral infection and cancer. E5 HPV protein is a powerful negative regulator of the anti-viral immune response by limiting the MHC I antigen repertoire via suppression of the immunoproteasome<sup>30</sup>, allowing infected cells to escape host immune surveillance<sup>55,56</sup>. The importance of the immunoproteasome in antigen recognition is further demonstrated, with high expression being a positive predictive biomarker of immune checkpoint inhibitor blockade response<sup>57</sup>. Our data shows that immune evasion is a second phenotypic feature of metastatic clones. Although HPV+ oropharyngeal cancer is considered an immunotherapy-sensitive cancer, the response is not universal. In the metastatic setting, survival is less than two years despite checkpoint inhibitor use. Ineffective antigen presentation via the immunoproteasome may be a desirable therapeutic target in HPV+ oropharyngeal cancer, given that HPV infection and carcinogenesis are dysregulators of this process. Restoration of the immunoproteasome by IFN- $\gamma$  has been shown to improve immune system recognition of renal cancer cells<sup>58</sup> and cervical cancer<sup>59</sup>. Interestingly, when this approach was used in head and neck cancer<sup>60</sup>, highly heterogeneous responses were seen *ex vivo*. This heterogeneity could be explained by the mixture of HPV+ and HPV- cancers analysed.

Challenges for managing HPV+ oropharyngeal cancer include the lack of reliable prognostic markers and paucity of druggable targets. This limits the ability to personalise treatment decisions in curable disease and offer effective treatment options in the metastatic setting. We have identified mechanisms defining the metastatic clonal phenotype unique to HPV+ oropharyngeal cancer. Given the aetiology of HPV-related cancer, it is expected that many of the cellular mechanisms are shared with those observed in chronic viral infection. *EIF4G1* expression may be helpful as a prognostic marker, which could be used to select patients more likely to metastasise and benefit from more aggressive curative therapies. Finally, the immunoproteasome likely plays a significant role in immune evasion in metastasising clones. Further investigation of immunoproteasome proteins as predictive biomarkers for checkpoint inhibitor response and direct targeting as a therapeutic strategy should be undertaken.

425

## Acknowledgements

We acknowledge the patients and their families.

This research was funded by Cure Cancer Australia, St Vincent's Clinic Research Foundation and The Garvan Institute of Medical Research Foundation.

430

## Author Contributions

V.C. designed the biobanking study and research project, conducted the experiments, analysed data, and wrote the manuscript. W.M. optimised *Numbat* code and assisted with data analysis and graphics. R.M. conducted biobanking, ethics management and conducted experiments.

435 J.A. conducted immune cell subtyping. D.N. and H.A. conducted de-multiplexing and pre-processing of single cell data. A.S. assisted with data analysis, storage and pre-processing.

P.K. conducted biobanking. E.S. conducted biobanking and experiments. A.M. conducted biobanking and experiments. D.K. conducted experiments. P.F, B.L and J.C. consented patients and retrieved specimens. P.E. specimen preparation, cut up and histopathological

440 analysis. R.G. managed the clinical aspects, consented patients, retrieval of specimens. J.P. designed experiments, assisted with data analysis, and wrote the manuscript.

## Declaration of Interests

Nil

445

## Figure Titles and Legends

**Figure 1: Study design and cell annotation.** A. Description of the patient samples. The primary sites of the oropharyngeal of either the tonsil or base of the tongue, nodal stations affected by cancer as per the pathological staging. B. Study schematic. In parallel, the samples were processed for scRNA-seq and spatial RNA sequencing using the Visium platform. Annotated single cells were used to deconvolute the spatial data. Clone delineation was performed based on single-cell CNV calls. C. Canonical markers demonstrate cancer cells, b-cells, t-cells, NK cells and monocytes.

**Figure 2. Karyotype plots.** For each patient, the karyotype plots, including the primary and the LN, are shown—amplification in red, deletion in blue and LOH in green. The location of key genes are shown at the top. Chromosome 8 was shown to be altered in all patients. Locations of genes of interest are listed across the top. HPV infection and cancer (*RBI*, *TP53*, *MYC*), genes identified in the ICGC cohort (*E2F1*, *TRAF3*, *PIK3CA*) and translation initiation (*EIF4E3*, *EIF4G1*, *EIF4EBP1*, *EIF4G2*, *EIF4A1*, *MKNK1*)

**Figure 3: Selection leading to clonal expansion in metastatic lymph nodes.** A. Clonal phylogenies are shown for each patient. B. CNV events underlying the formation of expanding clones. Each row displays the CNV profile for a clone. Deletion, amplification and loss of heterozygosity (LOH) events are shown in blue, red, and green, respectively. C. Percentage representation of each clone in the primary (dark colours) versus the LN (light colours). Significant changes in the percentage representation of a clone between the primary and LN were estimated using a t-test

**Figure 4. Downregulation of protein translation pathways in expanding metastatic**

**clones.** A. The primary and LN cancer cells were clustered into transcriptionally similar groups, with the clones overlayed in their colours shown in the bar graphs. In most cases, the clones formed transcriptionally distinct clusters, which formed the comparisons for pathway enrichment analyses. Multiple comparisons have been done in patients 1 and 3 using these distinct clusters. B. Reactome pathways, which were significantly (study-wide  $p < 0.05$ ) enriched, are shown. The colour relates to the p-value of enrichment. These are grouped into three distinct pathways of cell cycle, immune regulation and protein translation. C. Percentage of ribosomal protein gene expression in the expanding versus non-expanding clones. There is less expression in the expanding clones. A t-test shows that this difference is highly statistically significant. P-cluster = primary cluster. LN-cluster = lymph node cluster.

**Figure 5. Gene expression differences between metastasising expanding clones and non-**

**expanding clones.** Genes are grouped into eight key pathways: Protein translation initiating factors, protein translation regulatory kinases, protein translation stress kinases, apoptosis, cap-dependent translation, cap-independent translation, Proteasome, IFN-inducible genes. The non-expanding cancer clone's gene expression is in the dark blue and expanding clone's expression in the light blue.

**Figure 6. Relationship between expanding and non-expanding cancer clones and t-cell**

**sub types.** Spatial data was deconvoluted into cell type composition, and the colouration represents the estimated percentage of a cell's contribution to a probe spot. Correlative analysis and pie chart plots are shown for the primary (left) and LN (right). CD8TEM = CD8 T effector memory cells, CD4Prol = CD4 proliferating cells, T-reg = T regulatory cells, MAIT =

Mucosal-associated invariant T cells, dnT = Double negative t cells, CD8Ex = CD8 exhausted

495 cells, gdT = gammadelta t cells, ILC = innate lymphoid cells.

**Figure 7. Relationship between immune cell and cancer clone co-occurrence across primary and lymph node sites.** Co-occurrence correlations between expanding and non-

expanding clones and immune cell subtype determined from spatial RNA in situ analysis.

500 Significant differences were calculated and compared using t-tests, and study-wide significant relationships are displayed with a \*.

## Methods

### RESOURCE AVAILABILITY

#### 505 Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, A/Prof Venessa Chin (v.chin@garvan.org.au)

**Materials availability** This study did not generate any new unique reagents

**Data and code availability** All single-cell RNA-sequencing data generated by this study have  
510 been deposited in the [REDACTED] ([https://\[REDACTED\]](https://[REDACTED])). The data can be accessed under the  
accession number [REDACTED]. All other data supporting the findings of this study and Code  
used for all processing and analysis are available from the corresponding author upon  
reasonable request.

#### 515 EXPERIMENTAL MODEL AND STUDY PARTICIPANTS DETAILS

##### Ethics statement

This project received ethics approval from the St Vincent's Hospital Human Research Ethics Committee (HREC/16/SVH/329) in accordance with the ethical standards laid down in the Declaration of Helinski. Written informed consent was obtained per institutional guidelines.

#### 520 Patient cohort

Patients were an average age of 55 and compromised 75% men (n=3) and 25% women (n=1). Newly diagnosed, previously untreated patients with HPV+ Oropharyngeal squamous cell carcinoma were selected if they presented with lymph node metastases at St Vincent's Hospital Sydney, and proceeded to surgery for their tumour.

525

### METHOD DETAILS

#### Tissue sample processing

Within 30 minutes of surgical resection, tumour samples were tumour banked. For Visium, half of the tumor sample was cut off and snap frozen in O.C.T (Tissue-Tek, CAT# 4583) on dry ice. After freezing, samples were moved to -80°C for long-term storage. The remaining sample was cut into 2mm x 2mm chunks with a scalpel blade. For single cell analysis, 2-3 tumour chunks were placed into a cryovial with 90% Fetal Calf Serum (FCS) + 10% DMSO. Cryovials were placed into a CoolCell in a -80°C freezer overnight, then transferred to Vapour Phase for long-term storage.

### **Blood processing**

Blood was collected in 10mL BD Vacutainer EDTA tubes (Becton Dickinson, CAT# 367525). The sample was processed for plasma and buffy coat as previously described<sup>61</sup>. Samples were stored at -80 °C.

### **SNP Genotyping**

DNA was extracted from Buffy Coat samples using the QIAamp DNA Blood Mini Kit (Qiagen, CAT# 51106). Genotyping was performed using the UK Biobank Axiom array (ThermoFisher Scientific, CAT# 902502) at the Ramaciotti Centre for Genomics. Imputation was performed using the Michigan Imputation Server, with Minimac4 and the Haplotype Reference Consortium (HRC) panel.

### **Preparation of Single Cell Suspensions**

The presence of cancer tissue was confirmed on haematoxylin and eosin staining before sequencing. Tissue samples were defrosted in a 37°C water bath. Contents were transferred to a 6 cm culture dish and washed with RPMI-1640 (Gibco, CAT# 11875093) + 10% FCS. Samples were moved to a 2mL Lo-Bind Eppendorf containing 100uL of RPMI-1640 and

mechanically dissociated using scissors and snips. Samples were chemically dissociated with the Miltenyi Tumour Dissociation Kit (Human) (Miltenyi Biotech, CAT# 130-095-929) according to the manufacturer instructions. Samples were incubated for 20-30 minutes on a shaking incubator at 37°C until all tumour chunks were dissociated. Sample solution was passed through a 100um cell strainer into a FACS tube with RMPI-1640 + 10% FCS. Cells were counted and viability assessed using Trypan Blue. If samples had <80% cell viability then dead cells were removed with the EasySep Dead Cell Removal (Annexin V) Kit (StemCell Technologies, CAT# 17899) according to manufacturers instructions. Samples were pooled equally (primaries pooled together and lymph nodes pooled together) to a final concentration of 1500 cells per uL in PBS + 10% FCS. Samples were passed through a 70uM cell strainer.

### **Single-Cell Capture and cDNA Library Preparation**

The 10X Genomics Chromium instrument (10X Genomics) was used to partition viable cells with barcoded beads and cDNA from each cell was prepared using the 10X Genomics Single Cell 3' Library, Gel Bead and Multiplex Kit (v3) (CAT# 10X-1000075) as per the manufacturer's instructions. Patient samples were pooled and multiplexed. Cell numbers were optimised to capture approximately 5000 cells per individual sample, and 20000 cells in the total pool. Samples were loaded onto a 10X Genomics Chromium Single Cell Chip Kit B (CAT# 10X-1000153). Libraries were generated with the 10X Genomics Chromium Single Cell 3' Library construction kit (v3) (CAT# 10X-1000078)

### **scRNA-seq Data Processing**

**Library sequencing** The libraries were sequenced at the Ramaciotti Centre for Genomics on an Illumina NovaSeq 6000 (NovaSeq Control Software v 1.7.0 / Real-Time Analysis v3.4.4) using a NovaSeq S4 200 cycles kit (Illumina, 20482067 as follows: 28bp (Read 1), 91bp (Read

2) and 8bp (Index). A median sequencing depth of 30,000 reads/cell was targeted for each sample. The sequencer generated raw data files in binary base call (BCL) format. The BCL files were demultiplexed and converted to the FASTQ file formats using Illumina Conversion Software (bcl2fastq v2.19.0.316).

### **Bioinformatics pipeline:**

The cellranger -v (3.1.0) count pipeline was used for alignment, filtering, barcode, and UMI counting from FASTQ files. The pipeline was executed on a high-performance cluster with a 3.10.0-1127.el7.x86\_64 operating system. RNA sequencing reads were mapped to the human GRCh38 transcriptome and HPV 16 and 18 transcriptome<sup>62,63</sup> using the STAR aligner<sup>64</sup>.

**DropletQC** – To distinguish between true cells and cell-free ambient RNA, DropletQC.

Empty drops were excluded from the analysis<sup>65</sup>.

**Demultiplexing** – SNP data was generated for each individual as described above. Multiplexed patient pools were demultiplexed using Demuxafy<sup>66</sup>. Doublets were identified and removed.

**Single Cell Data** – The aggregated single-cell gene expression data generated by CellRanger was used as the input for the Seurat analysis software 4.0<sup>67</sup>. Expression levels for each transcript were determined using the number of unique molecular identifiers assigned to the transcript. Quality control and filtering steps were performed to remove outlier genes and cells. Cells with an expression greater than 25% of mitochondrial genes were removed. Cell-cell normalisation was performed using a regularised negative binomial regression (sctransform)<sup>68</sup>. Principal component analysis was performed on the filtered and normalised gene expression matrix, and the first 15 principal components that explained the majority of the variance in the

data were retained. Uniform manifold approximation and projection for dimension reduction (UMAP)<sup>[88]</sup> plots were generated to visualise the gene expression patterns in each cell.

605

**Integration of data** – Seurat was used to identify common anchors between datasets to integrate primary and lymph node data for each individual into the same object<sup>70</sup>.

### **CNV Estimation and identification of malignant cells**

610 Clonal architecture was inferred using the Numbat package<sup>25</sup>. Cancer cells from each individual were examined separately, with cells from both the primary and lymph node metastasis grouped, using the individual's immune cells as a normal reference. The default running parameters were used for all samples except patient 1. In this sample, the clones did not become stable until the LLR reached 400. This was thought to be due to the high number  
615 of malignant cells in this sample. Cancer cells with similar CNV profiles were grouped in a clonal group.

**Immune cell classification** – Cancer cells were identified using *EPCAM* expression. The immune cells were subtyped using automated methods using a reference dataset for immune  
620 cells<sup>71</sup>. This reference dataset does not include exhausted CD8<sup>+</sup> t-cells or macrophages, so these were added manually (Figure 1C). CD8<sup>+</sup> T-central memory cells (CD8 TCM) were identified using the reference dataset. Markers of cytotoxicity (*GNLY*, *GZMK*, *TNF*, *FASLG*, *PRF1*, *IL2*, *IFNG*) and exhaustion (*LAG3*, *BTLA*, *CTLA4*, *PDCD1*, *HAVCR2*) were used to classify the CD8 TCMs further. One sub-population of CD8 TCMs expressed high cytotoxicity  
625 and exhaustion markers and were labeled as CD8TCM-exhausted cells. For macrophages, the monocyte population was identified using the reference dataset. Expression counts within the monocytes of *FCGR2A* were calculated. Any cells expressing *FCGR2A* >2x the mean

expression was assigned as a macrophage. After cell filtering, normalisation, and cell classification we had 14157 cancer cells 36324 immune cells and 308 stromal/endothelial cells for analysis.

**Differential Gene Expression Analysis** – To examine the phenotype of the expanding clones, the cancer clones were grouped into transcriptionally similar clusters with the assigned clones overlaid. The non-expanding clones in the primary were compared with the expanding clones in the lymph node for differential gene expression analysis. Differential gene expression analysis was performed with the “FindAllMarkers” function in Seurat<sup>67</sup>. Calculations were corrected for the patient pool. Where there were transcriptionally and clonally distinct clusters, multiple comparisons were made per patient. Patient 1 and patient 3 had two comparisons done each. Patient 1 had two transcriptionally distinct clusters in the primary - “P-cluster A”, made up of clone 4, and “P-cluster B”, made up of clone 6. These were each separately compared to the LN cluster, denoted 1A and 1B in Figure 4B. Similarly, in patient 3, the primary cluster was compared to each of the LN clusters, “LN-cluster A”, made up of primarily clone 3 and “LN-cluster B”, primarily made up of clones 4 and 5 (denoted comparison 3A and 3B in Figure 4B).

DEG analysis was performed for each patient and analysed with ReactomePA<sup>72</sup> for enriched pathways. Custom code was developed for plots in supplementary figures 2-6.

### Visium Spatial Gene Expression

Following the manufacturer's instruction, O.C.T embedded tissue samples were processed using the Visium Spatial Gene Expression slide and reagent kit (10X Genomics, PN-1000186). Briefly, 10 µm cryosections were placed into the capture areas of the Visium slide. Tissue morphology was assessed with H&E staining and imaging using a Leica DM6000 Power

Mosaic microscope equipped with a 20x lens (Leica). The imaged sections were then permeabilized for 18 minutes using the supplied reagents. The permeabilization condition was previously optimised using the Visium Spatial Tissue Optimization reagent kit (10X Genomics, PN-1000192), and Visium Spatial Tissue Optimization slide kit (10X Genomics, PN-1000191). After permeabilization, cDNA libraries were prepared using the Library Construction Kit (10X Genomics, PN-1000190), checked for quality and sequenced on a NovaSeq 6000 platform (Illumina, US). Sequencing targeted 50,000 reads per occupied spot on the Visium slide. Around 300 million pair-ended reads were obtained for each tissue section. Read 1, i7 index and Read 2 were sequenced with 28, 8 and 91 cycles respectively.

### **Visium spatial transcriptomics data processing**

Three out of four patients had suitable tissue for spatial sequencing for both the primary and lymph node; however, in patient one, only the lymph node sample had sufficient tissue. Reads were demultiplexed and mapped to the reference genome GRCh38 using the Space Ranger Software v1.0.0 (10X Genomics). Count matrices were loaded into the Seurat v3.2 (<https://github.com/satijalab/seurat/tree/spatial>) and STutility (<https://github.com/jbergenstrahle/STUtility>) R packages for all subsequent data filtering, normalisation, filtering, dimensional reduction and visualisation. All spatial spots determined to be over tissue regions by Space Ranger were kept for subsequent analysis. Poor quality tissue locations were then filtered out based on a cut off of 500 unique genes. Genes detected in more than 10 locations were also kept for analysis. Data normalisation was performed on independent tissue sections using the variance stabilizing transformation method implemented in the SCTransform function in Seurat. We applied non-negative matrix factorization (NMF) to the normalised expression matrix using the STutility package (nfactors = 20). NMF reduction was then used for clustering using Seurat with all 20 factors as input (RunUMAP, FindNeighbors and FindClusters functions).

**Deconvolution of Visium Spots** – for each patient, the single cell dataset was used as a reference for deconvolution of the Visium spots. The cell subtype signatures from annotated single cell data is used to determine the cell composition of each Visium spot using a seeded NMF regression<sup>74</sup>. A weighted combination of cell types which reside in each spot are calculated and visualised as a pie chart. With these weighted combinations, correlation analyses were performed, examining the likelihood of a certain cell-types co-locating with one another.

## Correlation plots

Mean correlation scores between cancer clones and cell types of interest were calculated for all individual primaries and LN separately. Bar graphs were generated using ggbarplot<sup>73</sup>, and statistical significance was determined using a t-test.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Statistical analysis performed and software used

Figure 2 – statistical significance was determined using Fisher exact test in rstatix package<sup>75</sup>

Figure 5 – statistical significance was determined using a t-test using the stat\_compare\_means function in ggplot2.

P-values are reported with asterix: \* <0.05, \*\* <0.01, \*\*\*<0.001

### Key Resources Table

| REAGENT or RESOURCE                   | SOURCE  | IDENTIFIER               |
|---------------------------------------|---|--------------------------|
| Biological samples                    |   |                          |
| Fresh surgical tumor resections       | The St Vincent's Head and Neck Cancer Biobank | Ethics # HREC/16/SVH/329 |
| O.C.T embedded tumor tissue blocks    | The St Vincent's Head and Neck Cancer Biobank | Ethics # HREC/16/SVH/329 |
| Buffy Coat, prepared from whole blood | The St Vincent's Head and Neck Cancer Biobank | Ethics # HREC/16/SVH/329 |
|                                       |   |                          |

| Critical commercial assays  |                            |   |
|---|----------------------------|---|
| Chromium Next GEM Single Cell 3' Library, Gel Bead and Multiplex Kit v3 | 10X genomics               | 10X-1000075   |
| Chromium Next GEM Single Cell Chip Kit B                                | 10X genomics               | 10X-1000153   |
| Chromium Single Cell 3' Library construction kit v3                     | 10X genomics               | 10X-1000078   |
| Visium Spatial Gene Expression Slide and Reagent Kit                    | 10X genomics               | PN-1000186  |
| Library Construction Kit  | 10X genomics               | PN-1000190  |
| Miltenyi tumor dissociation kit, human                                  | Miltenyi                   | 130-095-929   |
| EasySep Dead Cell Removal (Annexin V) Kit                               | StemCell Technologies      | 17899   |
| QIAamp DNA Blood Mini Kit   | Qiagen                     | 51106   |
| Deposited data  |                            |   |
| Single cell RNA-seq data?   | This paper                 |   |
| Visium data?  | This paper                 |   |
| Software and algorithms   |                            |   |
| Cell Ranger v 3.1.0   | 10X Genomics               | <a href="https://software.10xgenomics.com/">https://software.10xgenomics.com/</a>   |
| STAR aligner  | Dobin et al., 2013         | <a href="https://github.com/alexdobin/STAR">https://github.com/alexdobin/STAR</a>   |
| DropletQC   | Muskovic and Powell, 2021  | <a href="https://powellgenomicslab.github.io/DropletQC/">https://powellgenomicslab.github.io/DropletQC/</a>                                       |
| Demuxafy  | Neavin et al., 2022        | <a href="https://github.com/drneavin/Demultiplexing_Doublet_Detecting_Docs">https://github.com/drneavin/Demultiplexing_Doublet_Detecting_Docs</a> |
| Seurat 4.2.2  | Satija et al., 2015        | <a href="https://github.com/satijalab/seurat">https://github.com/satijalab/seurat</a>   |
| scPred  | Alquicira-Hernandez., 2019 | <a href="https://github.com/powellgenomicslab/scPred">https://github.com/powellgenomicslab/scPred</a>   |
| Numbat  | Gao et al., 2022           | <a href="https://github.com/kharchenkolab/numbat">https://github.com/kharchenkolab/numbat</a>   |
| karyoploteR   | Gel et al., 2017           | <a href="https://bernatgel.github.io/karyoploteR_tutorial/">https://bernatgel.github.io/karyoploteR_tutorial/</a>                                 |
| ReactomePA  | Yu et al., 2016            | <a href="https://github.com/YuLab-SMU/ReactomePA">https://github.com/YuLab-SMU/ReactomePA</a>   |
| ggplot2   | Wickham. 2016              | <a href="https://ggplot2.tidyverse.org/">https://ggplot2.tidyverse.org/</a>   |
| SPOTlight   | Elosua-Bayes et al., 2021  | <a href="https://marcelosua.github.io/SPOTlight/">https://marcelosua.github.io/SPOTlight/</a>   |

|         |                            |   |
|---------|----------------------------|---|
| rstatix | Kassambara et al.,<br>2023 | <a href="https://rpkgs.datanovia.com/rstatix/">https://rpkgs.datanovia.com/rstatix/</a> |
|---------|----------------------------|---|

## Supplemental information titles and legends

**Supplementary Table 1.** Patient characteristics and survival outcomes

700 **Supplementary Table 2.** Differential gene expression analysis for the comparisons in Figure 4. Each comparison is listed separately.

**Supplementary Table 3.** Gene enrichment analysis results for each comparison in Figure 4. Each comparison is listed separately.

**Supplementary Table 4.** Mean expression levels of genes shown in Figure 5

705 **Supplementary Figure 1.** Patient 1A. The top ten enriched pathways, revealing the direction and extent of gene expression changes

**Supplementary Figure 2.** Patient 1B. The top ten enriched pathways, revealing the direction and extent of gene expression changes

**Supplementary Figure 3.** Patient 2. The top ten enriched pathways, revealing the direction  
710 and extent of gene expression changes

**Supplementary Figure 4.** Patient 3A. The top ten enriched pathways, revealing the direction and extent of gene expression changes

**Supplementary Figure 5.** Patient 3B. The top ten enriched pathways, revealing the direction and extent of gene expression changes

715 **Supplementary Figure 6.** Patient 4. The top ten enriched pathways, revealing the direction and extent of gene expression changes

**Supplementary Figure 7.** Functional protein association network from patient comparison 1A

**Supplemental Figure 8. Spatial data and correlative analysis for cancer clones and b-cell  
720 subsets.** Correlative analysis and pie chart plots are shown for the primary (left) and LN (right). B-intermed = b intermediate cells.

## Supplemental Figure 9. Spatial data and correlative analysis for cancer clones and

**monocyte subsets.** Correlative and pie chart plots are shown for the primary (left) and the LN

(right). CD14 Mono = CD14 positive monocytes, cDC = classical dendritic cells, pDC =

plasmacytoid dendritic cells

## References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018; **68**(6): 394-424.
- 750 2. O'Sullivan B, Huang SH, Su J, Garden AS, Sturgis EM, Dahlstrom K, Lee N, Riaz N, Pei X, Koyfman SA, et al. Development and validation of a staging system for HPV-related oropharyngeal cancer by the International Collaboration on Oropharyngeal cancer Network for Staging (ICON-S): a multicentre cohort study. *Lancet Oncol* 2016; **17**(4): 440-51.
3. Chow LQM. Head and Neck Cancer. *N Engl J Med* 2020; **382**(1): 60-72.
- 755 4. Goldenberg D, Begum S, Westra WH, Khan Z, Sciubba J, Pai SI, Califano JA, Tufano RP, Koch WM. Cystic lymph node metastasis in patients with head and neck cancer: An HPV-associated phenomenon. *Head Neck* 2008; **30**(7): 898-903.
5. Floros P, Rao A, McCloy RA, Sim HW, Chin VT, Leavers BC, Crawford JA, Gallagher RM. Altered presentation of oropharyngeal cancer, a 6-year review. *ANZ J Surg* 2021; **91**(6): 1240-5.
- 760 6. Burtneess B, Harrington KJ, Greil R, Soulieres D, Tahara M, de Castro G, Jr., Psyrri A, Baste N, Neupane P, Bratland A, et al. Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study. *Lancet* 2019; **394**(10212): 1915-28.
- 765 7. Nindra U, Hurwitz J, Forstner D, Chin V, Gallagher R, Liu J. A systematic review of neoadjuvant and definitive immunotherapy in locally advanced head and neck squamous cell carcinoma. *Cancer Med* 2023.
8. Fakhry C, Zhang Q, Nguyen-Tan PF, Rosenthal D, El-Naggar A, Garden AS, Soulieres D, Trotti A, Avizonis V, Ridge JA, et al. Human Papillomavirus and Overall Survival After Progression of Oropharyngeal Squamous Cell Carcinoma. *Journal of Clinical Oncology* 2014; **32**(30): 3365-73.
- 770 9. Ferris RL, Flamand Y, Weinstein GS, Li S, Quon H, Mehra R, Garcia JJ, Chung CH, Gillison ML, Duvvuri U, et al. Phase II Randomized Trial of Transoral Surgery and Low-Dose Intensity Modulated Radiation Therapy in Resectable p16+ Locally Advanced Oropharynx Cancer: An ECOG-ACRIN Cancer Research Group Trial (E3311). *J Clin Oncol* 2022; **40**(2): 138-49.
- 775 10. Lenz G, Onzi GR, Lenz LS, Buss JH, Dos Santos JA, Begnini KR. The Origins of Phenotypic Heterogeneity in Cancer. *Cancer Res* 2022; **82**(1): 3-11.
- 780 11. Hu Z, Li Z, Ma Z, Curtis C. Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nat Genet* 2020; **52**(7): 701-8.
12. Echeverria GV, Powell E, Seth S, Ge Z, Carugo A, Bristow C, Peoples M, Robinson F, Qiu H, Shao J, et al. High-resolution clonal mapping of multi-organ metastasis in triple negative breast cancer. *Nat Commun* 2018; **9**(1): 5079.
- 785 13. Baldwin LA, Bartonicek N, Yang J, Wu SZ, Deng N, Roden DL, Chan CL, Al-Eryani G, Zanker DJ, Parker BS, et al. DNA barcoding reveals ongoing immunoediting of clonal cancer populations during metastatic progression and immunotherapy response. *Nat Commun* 2022; **13**(1): 6539.
14. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* 2017; **168**(4): 613-28.
- 790 15. Greaves M, Maley CC. Clonal evolution in cancer. *Nature* 2012; **481**(7381): 306-13.
16. Quah HS, Cao EY, Suteja L, Li CH, Leong HS, Chong FT, Gupta S, Arcinas C, Ouyang JF, Ang V, et al. Single cell analysis in head and neck cancer reveals potential immune evasion mechanisms during early metastasis. *Nat Commun* 2023; **14**(1): 1680.
- 795 17. Tang Y, Zang H, Wen Q, Fan S. AXL in cancer: a modulator of drug resistance and therapeutic target. *J Exp Clin Cancer Res* 2023; **42**(1): 148.

18. Tang A, Gao K, Chu L, Zhang R, Yang J, Zheng J. Aurora kinases: novel therapy targets in cancers. *Oncotarget* 2017; **8**(14): 23937-54.
19. Della Fera AN, Warburton A, Coursey TL, Khurana S, McBride AA. Persistent Human Papillomavirus Infection. *Viruses* 2021; **13**(2).
20. Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* 1993; **75**(3): 495-505.
21. Rampias T, Sasaki C, Weinberger P, Psyrris A. E6 and e7 gene silencing and transformed phenotype of human papillomavirus 16-positive oropharyngeal cancer cells. *J Natl Cancer Inst* 2009; **101**(6): 412-23.
22. Kesis TD, Slebos RJ, Nelson WG, Kastan MB, Plunkett BS, Han SM, Lorincz AT, Hedrick L, Cho KR. Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. *Proc Natl Acad Sci U S A* 1993; **90**(9): 3988-92.
23. Alazawi W, Pett M, Strauss S, Moseley R, Gray J, Stanley M, Coleman N. Genomic imbalances in 70 snap-frozen cervical squamous intraepithelial lesions: associations with lesion grade, state of the HPV16 E2 gene and clinical outcome. *Br J Cancer* 2004; **91**(12): 2063-70.
24. Paiva B, Corchete LA, Vidriales MB, Puig N, Maiso P, Rodriguez I, Alignani D, Burgos L, Sanchez ML, Barcena P, et al. Phenotypic and genomic analysis of multiple myeloma minimal residual disease tumor cells: a new model to understand chemoresistance. *Blood* 2016; **127**(15): 1896-906.
25. Gao T, Soldatov R, Sarkar H, Kurkiewicz A, Biederstedt E, Loh P-R, Kharchenko PV. Haplotype-aware analysis of somatic copy number variations from single-cell transcriptomes. *Nature Biotechnology* 2022.
26. Cancer Genome Atlas N. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* 2015; **517**(7536): 576-82.
27. Gao T, Soldatov R, Sarkar H, Kurkiewicz A, Biederstedt E, Loh PR, Kharchenko PV. Haplotype-aware analysis of somatic copy number variations from single-cell transcriptomes. *Nat Biotechnol* 2023; **41**(3): 417-26.
28. Dillekas H, Rogers MS, Straume O. Are 90% of deaths from cancer caused by metastases? *Cancer Med* 2019; **8**(12): 5574-6.
29. Kazemi S, Papadopoulou S, Li S, Su Q, Wang S, Yoshimura A, Matlashewski G, Dever TE, Koromilas AE. Control of alpha subunit of eukaryotic translation initiation factor 2 (eIF2 alpha) phosphorylation by the human papillomavirus type 18 E6 oncoprotein: implications for eIF2 alpha-dependent gene expression and cell death. *Mol Cell Biol* 2004; **24**(8): 3415-29.
30. Miyauchi S, Kim SS, Jones RN, Zhang L, Guram K, Sharma S, Schoenberger SP, Cohen EEW, Califano JA, Sharabi AB. Human papillomavirus E5 suppresses immunity via inhibition of the immunoproteasome and STING pathway. *Cell Rep* 2023; **42**(5): 112508.
31. Billingsley CL, Chintala S, Katzenellenbogen RA. Post-Transcriptional Gene Regulation by HPV 16E6 and Its Host Protein Partners. *Viruses* 2022; **14**(7).
32. Gaucherand L, Gaglia MM. The Role of Viral RNA Degrading Factors in Shutoff of Host Gene Expression. *Annu Rev Virol* 2022; **9**(1): 213-38.
33. Xu Y, Ruggero D. The Role of Translation Control in Tumorigenesis and Its Therapeutic Implications. *Annual Review of Cancer Biology* 2020; **4**(1): 437-57.
34. Bercovich-Kinori A, Tai J, Gelbart IA, Shitrit A, Ben-Moshe S, Drori Y, Itzkovitz S, Mandelboim M, Stern-Ginossar N. A systematic view on influenza induced host shutoff. *Elife* 2016; **5**.
35. Stern-Ginossar N, Thompson SR, Mathews MB, Mohr I. Translational Control in Virus-Infected Cells. *Cold Spring Harb Perspect Biol* 2019; **11**(3).

36. Litwin TR, Irvin SR, Chornock RL, Sahasrabudhe VV, Stanley M, Wentzensen N. Infiltrating T-cell markers in cervical carcinogenesis: a systematic review and meta-analysis. *Br J Cancer* 2021; **124**(4): 831-41.
37. Bidwell BN, Slaney CY, Withana NP, Forster S, Cao Y, Loi S, Andrews D, Mikeska  
850 T, Mangan NE, Samarajiwa SA, et al. Silencing of Irf7 pathways in breast cancer cells promotes bone metastasis through immune escape. *Nat Med* 2012; **18**(8): 1224-31.
38. Wu S, Wagner G. Deep computational analysis details dysregulation of eukaryotic translation initiation complex eIF4F in human cancers. *Cell Syst* 2021; **12**(9): 907-23 e6.
39. Leong TL, Gayevskiy V, Steinfort DP, De Massy MR, Gonzalez-Rajal A, Marini KD,  
855 Stone E, Chin V, Havryk A, Plit M, et al. Deep multi-region whole-genome sequencing reveals heterogeneity and gene-by-environment interactions in treatment-naïve, metastatic lung cancer. *Oncogene* 2019; **38**(10): 1661-75.
40. Mlecnik B, Bindea G, Kirilovsky A, Angell HK, Obenauf AC, Tosolini M, Church SE, Maby P, Vasaturo A, Angelova M, et al. The tumor microenvironment and Immunoscore are  
860 critical determinants of dissemination to distant metastasis. *Sci Transl Med* 2016; **8**(327): 327ra26.
41. Capp JP, DeGregori J, Nedelcu AM, Dujon AM, Boutry J, Pujol P, Alix-Panabieres C, Hamede R, Roche B, Ujvari B, et al. Group phenotypic composition in cancer. *Elife* 2021; **10**.
42. Marusyk A, Janiszewska M, Polyak K. Intratumor Heterogeneity: The Rosetta Stone  
865 of Therapy Resistance. *Cancer Cell* 2020; **37**(4): 471-84.
43. Blanchard P, Baujat B, Holostenco V, Bourredjem A, Baey C, Bourhis J, Pignon JP, group M-CC. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): a comprehensive analysis by tumour site. *Radiother Oncol* 2011; **100**(1): 33-40.
44. O'Leary B. Understanding Head and Neck Cancer Evolution to Guide Therapeutic  
870 Approaches. In: Vermorken JB, Budach V, Leemans CR, Machiels J-P, Nicolai P, O'Sullivan B, editors. Critical Issues in Head and Neck Oncology; 2023 2023//; Cham: Springer International Publishing; 2023. p. 63-81.
45. Roh V, Abramowski P, Hiou-Feige A, Cornils K, Rivals JP, Zougman A, Aranyossy T, Thielecke L, Truan Z, Mermoud M, et al. Cellular Barcoding Identifies Clonal Substitution as a  
875 Hallmark of Local Recurrence in a Surgical Model of Head and Neck Squamous Cell Carcinoma. *Cell Rep* 2018; **25**(8): 2208-22 e7.
46. Dalet A, Arguello RJ, Combes A, Spinelli L, Jaeger S, Fallet M, Vu Manh TP, Mendes A, Perego J, Reverendo M, et al. Protein synthesis inhibition and GADD34 control IFN-beta heterogeneous expression in response to dsRNA. *EMBO J* 2017; **36**(6): 761-82.
- 880 47. Klein P, Kallenberger SM, Roth H, Roth K, Ly-Hartig TBN, Magg V, Ales J, Talemi SR, Qiang Y, Wolf S, et al. Temporal control of the integrated stress response by a stochastic molecular switch. *Sci Adv* 2022; **8**(12): eabk2022.
48. Tanaka Y, Ito S, Oshino R, Chen N, Nishio N, Isobe K. Effects of growth arrest and DNA damage-inducible protein 34 (GADD34) on inflammation-induced colon cancer in mice.  
885 *Br J Cancer* 2015; **113**(4): 669-79.
49. Holczer M, Banhegyi G, Kapuy O. GADD34 Keeps the mTOR Pathway Inactivated in Endoplasmic Reticulum Stress Related Autophagy. *PLoS One* 2016; **11**(12): e0168359.
50. Hamamura K, Minami K, Tanjung N, Wan Q, Koizumi M, Matsuura N, Na S, Yokota H. Attenuation of malignant phenotypes of breast cancer cells through eIF2alpha-mediated  
890 downregulation of Rac1 signaling. *Int J Oncol* 2014; **44**(6): 1980-8.
51. Kaiser C, Dobrikova EY, Bradrick SS, Shveygert M, Herbert JT, Gromeier M. Activation of cap-independent translation by variant eukaryotic initiation factor 4G in vivo. *RNA* 2008; **14**(10): 2170-82.
52. Morfioisse F, Kuchnio A, Frainay C, Gomez-Brouchet A, Delisle MB, Marzi S, Helfer  
895 AC, Hantelys F, Pujol F, Guillermet-Guibert J, et al. Hypoxia induces VEGF-C expression in

metastatic tumor cells via a HIF-1 $\alpha$ -independent translation-mediated mechanism. *Cell Rep* 2014; **6**(1): 155-67.

53. Braunstein S. A hypoxia-controlled cap-dependent to cap-independent translation switch in breast cancer. *Mol Cell* 2007; **28**.

54. Scott ML, Woodby BL, Ulicny J, Raikhy G, Orr AW, Songock WK, Bodily JM. Human Papillomavirus 16 E5 Inhibits Interferon Signaling and Supports Episomal Viral Maintenance. *J Virol* 2020; **94**(2).

55. Tripathi SC, Peters HL, Taguchi A, Katayama H, Wang H, Momin A, Jolly MK, Celiktaş M, Rodriguez-Canales J, Liu H, et al. Immunoproteasome deficiency is a feature of non-small cell lung cancer with a mesenchymal phenotype and is associated with a poor outcome. *Proc Natl Acad Sci U S A* 2016; **113**(11): E1555-64.

56. Gobbi G, Mirandola P, Micheloni C, Solenghi E, Sponzilli I, Artico M, Soda G, Zanelli G, Pelusi G, Fiorini T, et al. Expression of HLA class I antigen and proteasome subunits LMP-2 and LMP-10 in primary vs. metastatic breast carcinoma lesions. *Int J Oncol* 2004; **25**(6): 1625-9.

57. Thompson JC, Davis C, Deshpande C, Hwang WT, Jeffries S, Huang A, Mitchell TC, Langer CJ, Albelda SM. Gene signature of antigen processing and presentation machinery predicts response to checkpoint blockade in non-small cell lung cancer (NSCLC) and melanoma. *J Immunother Cancer* 2020; **8**(2).

58. Seliger B, Hammers S, Hohne A, Zeidler R, Knuth A, Gerharz CD, Huber C. IFN- $\gamma$ -mediated coordinated transcriptional regulation of the human TAP-1 and LMP-2 genes in human renal cell carcinoma. *Clin Cancer Res* 1997; **3**(4): 573-8.

59. Street D, Kaufmann AM, Vaughan A, Fisher SG, Hunter M, Schreckenberger C, Potkul RK, Gissmann L, Qiao L. Interferon- $\gamma$  enhances susceptibility of cervical cancer cells to lysis by tumor-specific cytotoxic T cells. *Gynecol Oncol* 1997; **65**(2): 265-72.

60. Berszin M, Michaelides I, Siemert J, Rohl L, Wellhausen J, Wald T, Bohr C, Kunzel J, Gradistanac T, Dietz A, et al. Cytokine Profiles of Head and Neck Squamous Cell Carcinoma Undergoing Dual Immunotherapy With Cetuximab and Pembrolizumab Identify Interferon Gamma-Induced Protein 10 as Novel Biomarker. *Front Oncol* 2022; **12**: 795277.

61. Yazar S, Alquicira-Hernandez J, Wing K, Senabouth A, Gordon MG, Andersen S, Lu Q, Rowson A, Taylor TRP, Clarke L, et al. Single-cell eQTL mapping identifies cell type-specific genetic control of autoimmune disease. *Science* 2022; **376**(6589): eabf3041.

62. Van Doorslaer K, Li Z, Xirasagar S, Maes P, Kaminsky D, Liou D, Sun Q, Kaur R, Huyen Y, McBride AA. The Papillomavirus Episteme: a major update to the papillomavirus sequence database. *Nucleic Acids Res* 2017; **45**(D1): D499-D506.

63. Van Doorslaer K, Tan Q, Xirasagar S, Bandaru S, Gopalan V, Mohamoud Y, Huyen Y, McBride AA. The Papillomavirus Episteme: a central resource for papillomavirus sequence data and analysis. *Nucleic Acids Res* 2013; **41**(Database issue): D571-8.

64. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; **29**(1): 15-21.

65. Muskovic W, Powell JE. DropletQC: improved identification of empty droplets and damaged cells in single-cell RNA-seq data. *Genome Biology* 2021; **22**(1): 329.

66. Neavin D, Senabouth A, Hang Lee JT, Ripoll A, Franke L, Prabhakar S, Ye CJ, McCarthy DJ, Melé M, Hemberg M, et al. Demuxafy. Improvement in droplet assignment by integrating multiple single-cell demultiplexing and doublet detection methods. *bioRxiv* 2022: 2022.03.07.483367.

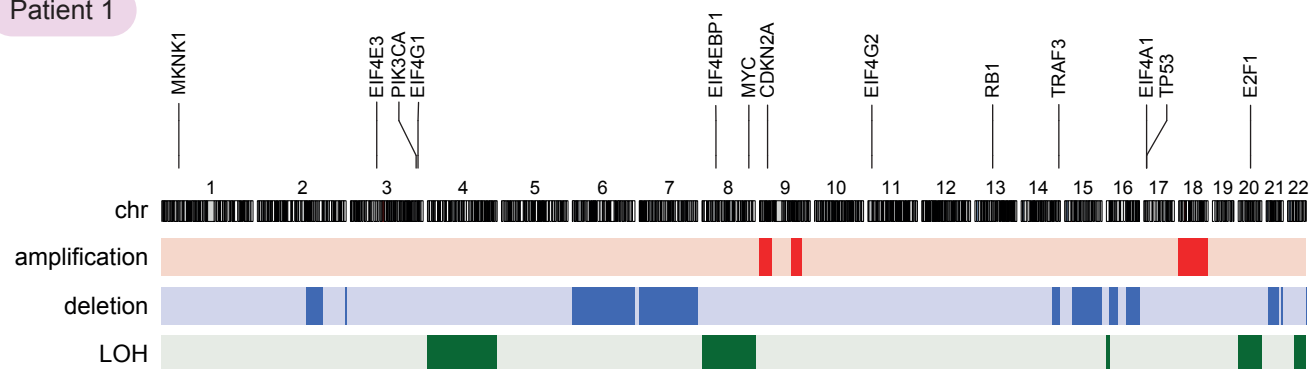
67. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol* 2015; **33**(5): 495-502.

68. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 2019; **20**(1): 296.

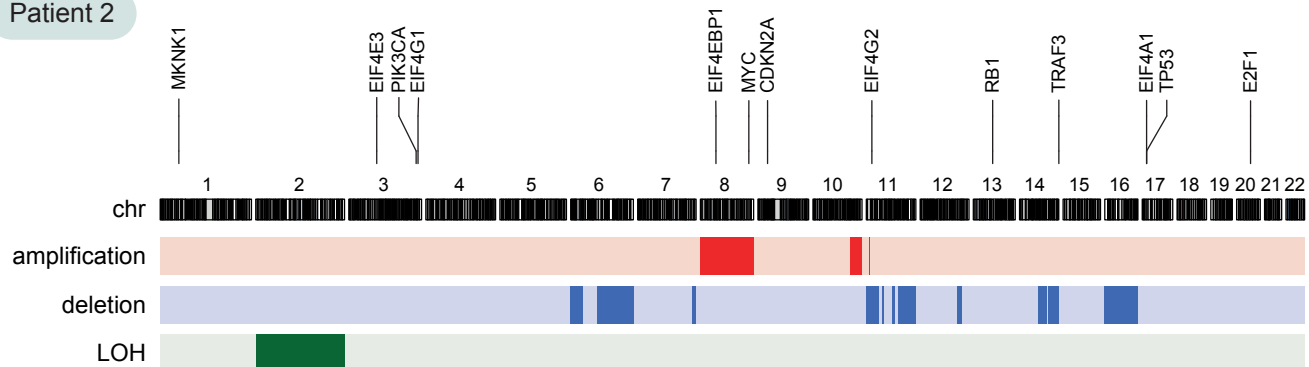
69. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, Ginhoux F, Newell EW. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol* 2018.
- 950 70. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive Integration of Single-Cell Data. *Cell* 2019; **177**(7): 1888-902 e21.
71. Alquicira-Hernandez J, Sathe A, Ji HP, Nguyen Q, Powell JE. scPred: accurate supervised method for cell-type classification from single-cell RNA-seq data. *Genome Biol* 2019; **20**(1): 264.
- 955 72. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst* 2016; **12**(2): 477-9.
73. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York: Springer-Verlag 2016.
- 960 74. Nieto P, Elosua-Bayes M, Trincado JL, Marchese D, Massoni-Badosa R, Salvany M, Henriques A, Nieto J, Aguilar-Fernandez S, Mereu E, et al. A single-cell tumor immune atlas for precision oncology. *Genome Res* 2021; **31**(10): 1913-26.
75. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. R package version 0.7.2. 20232023).

## Figure 2

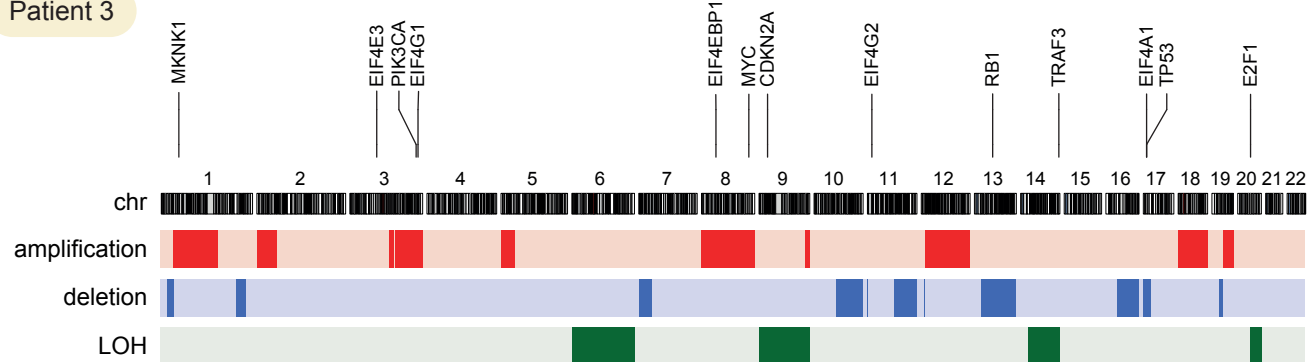
### Patient 1



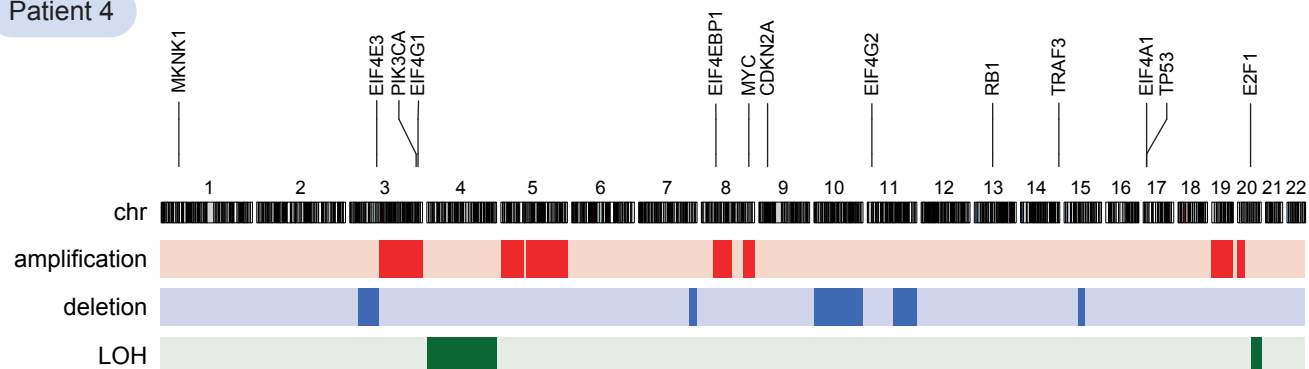
### Patient 2



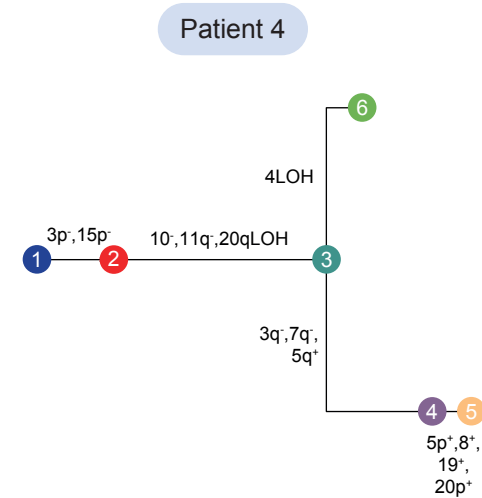
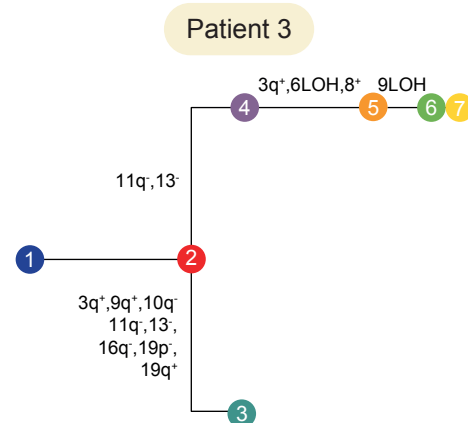
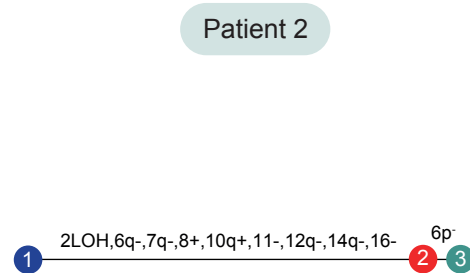
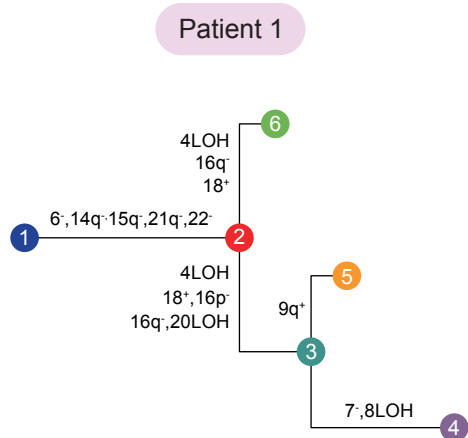
### Patient 3



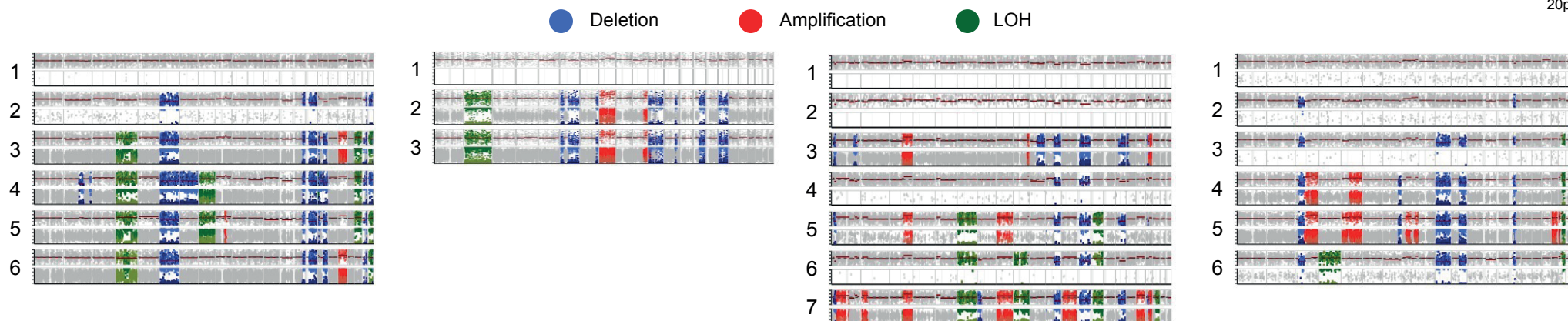
### Patient 4



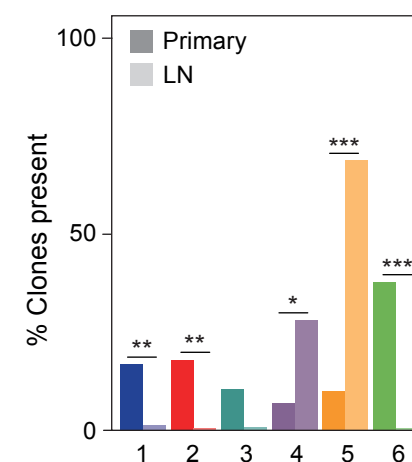
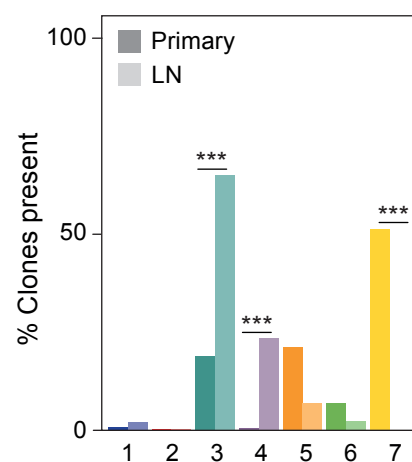
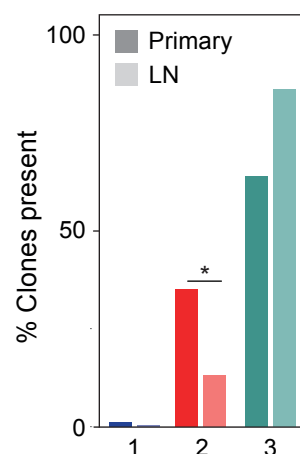
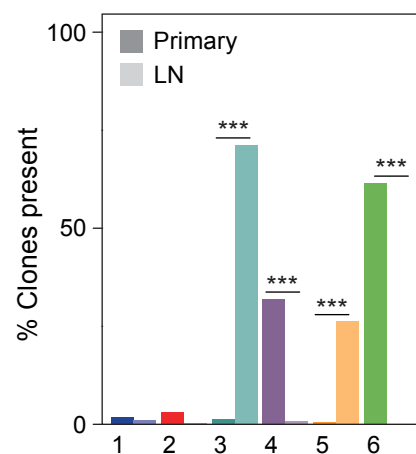
**A**



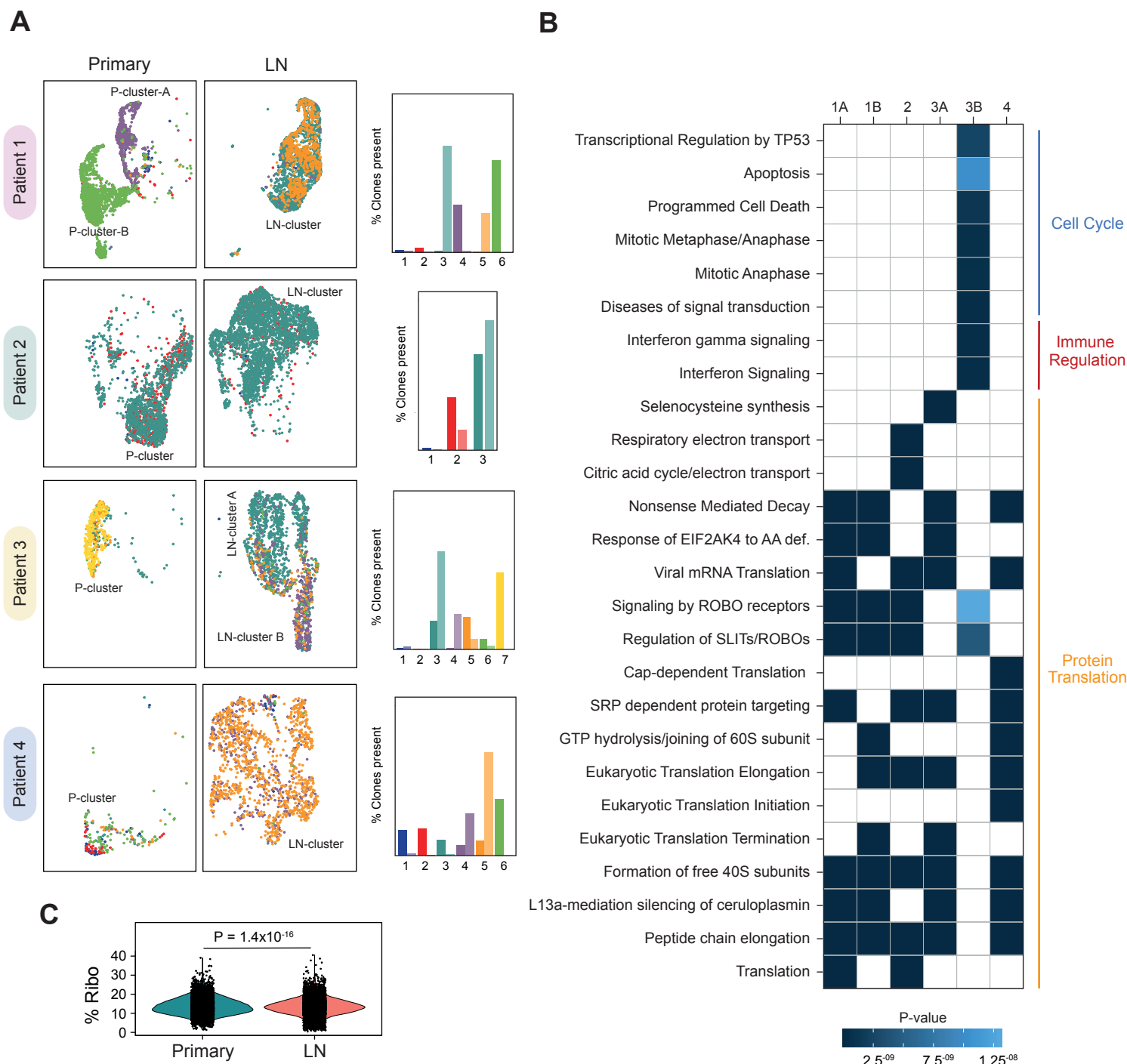
## B



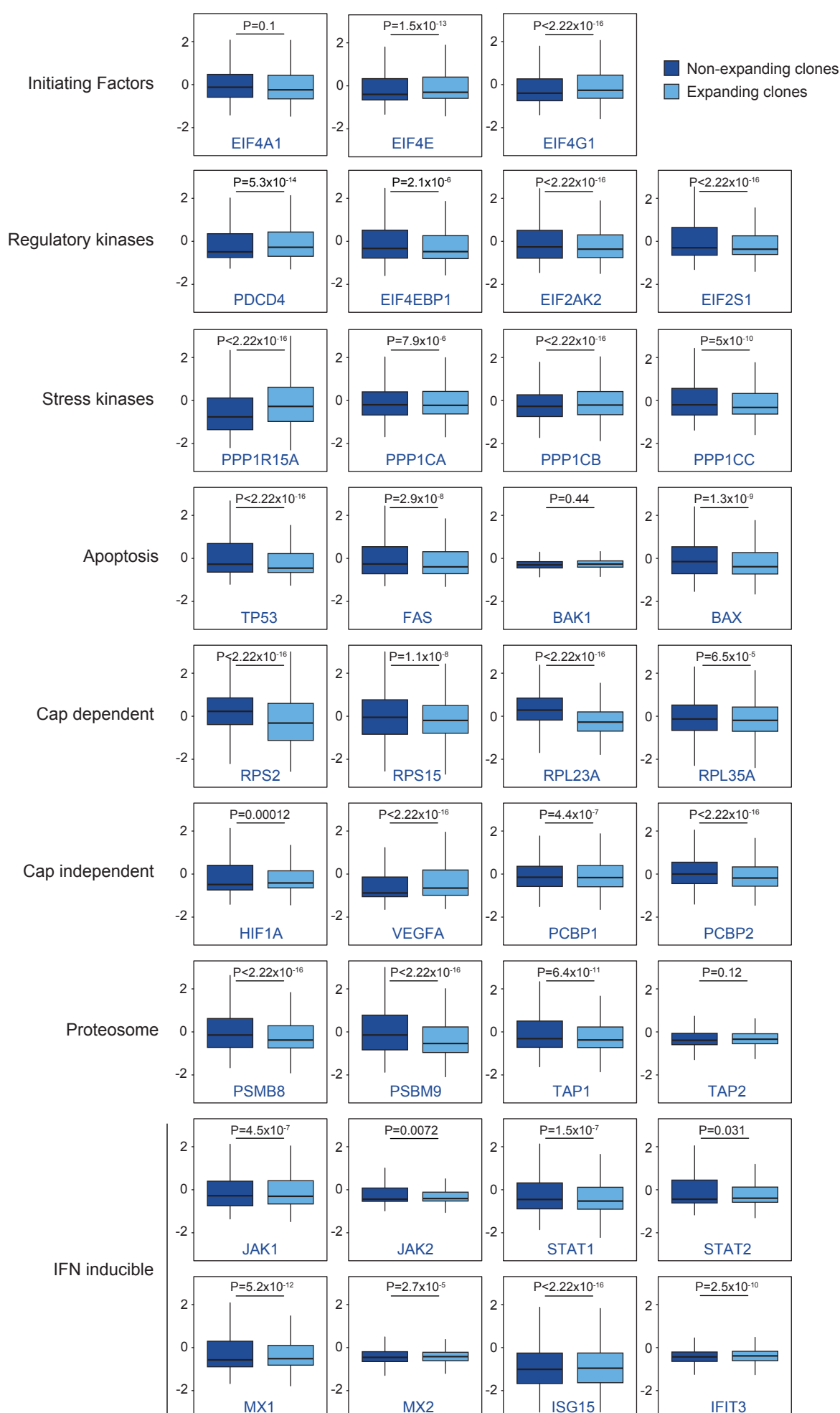
**C**



## Figure 4



**Figure 5**



## Figure 7

Expanding Non-Expanding

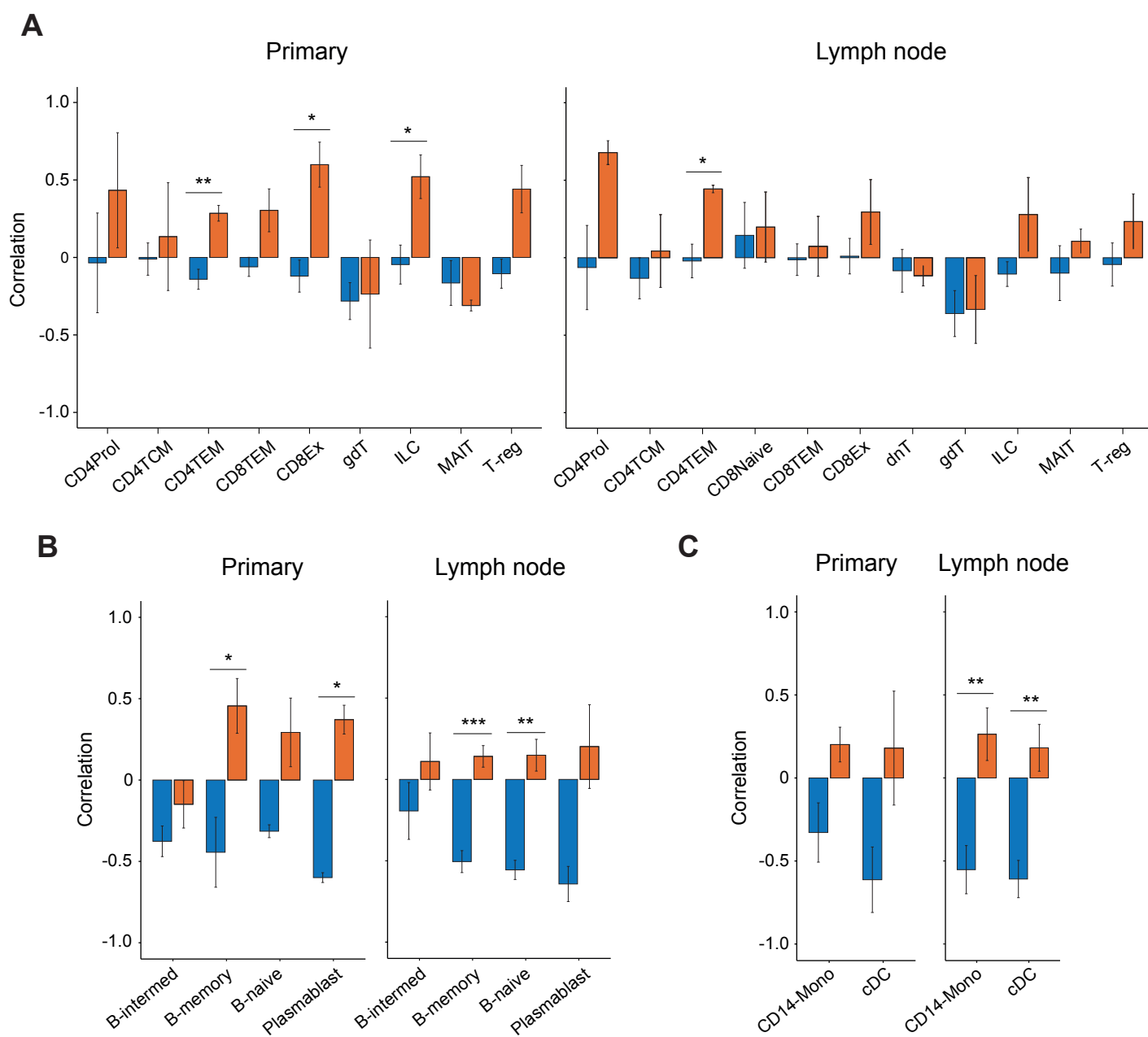
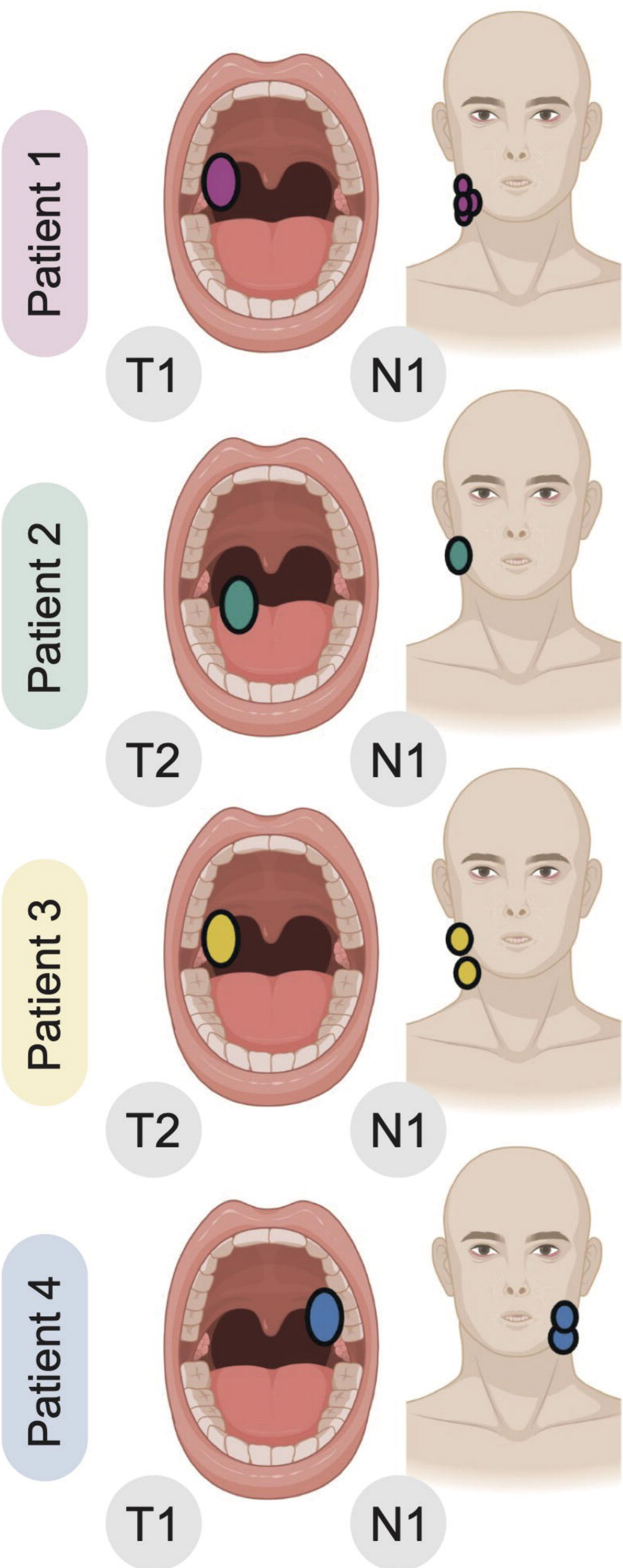
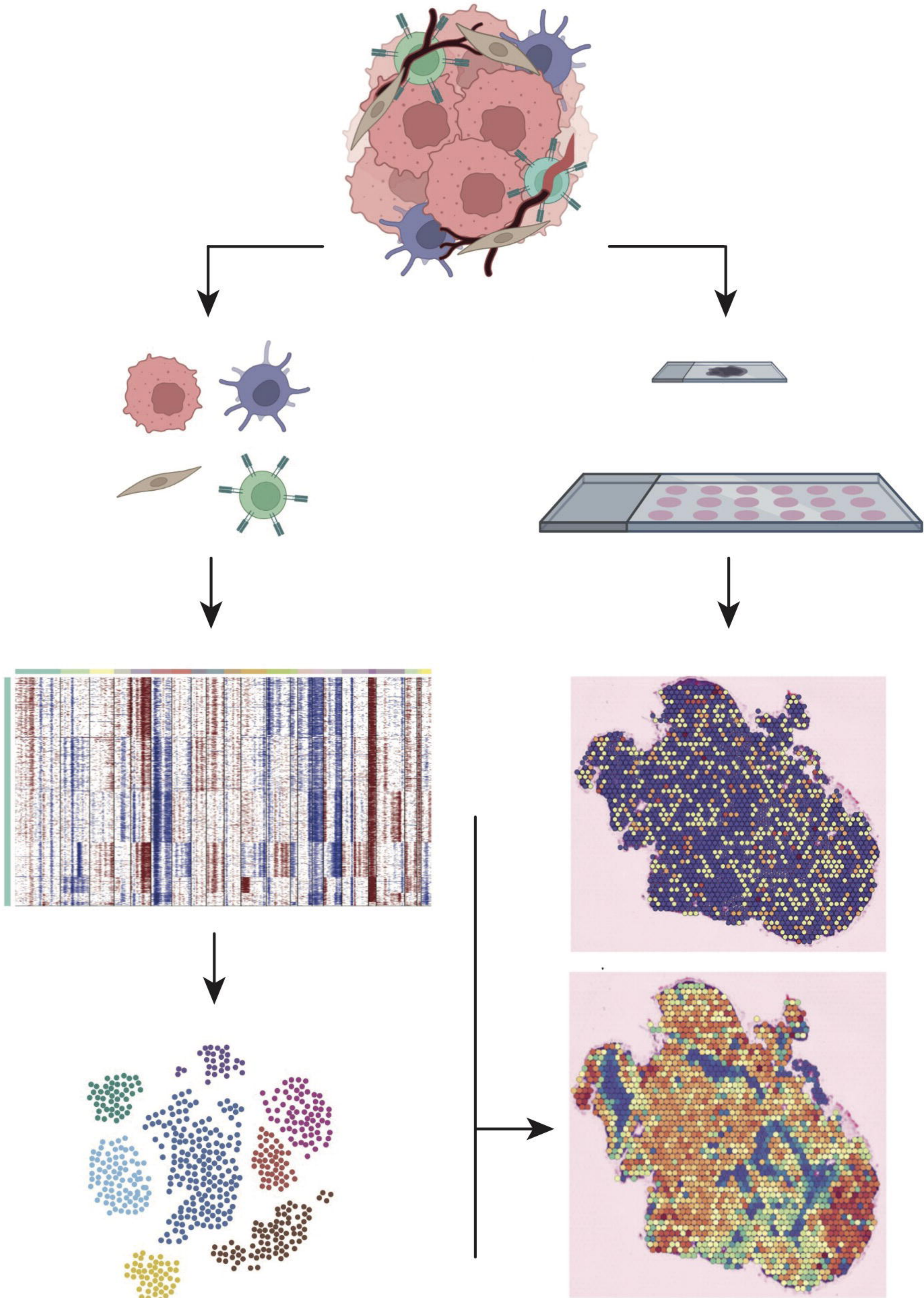


Figure 1

A



B



C

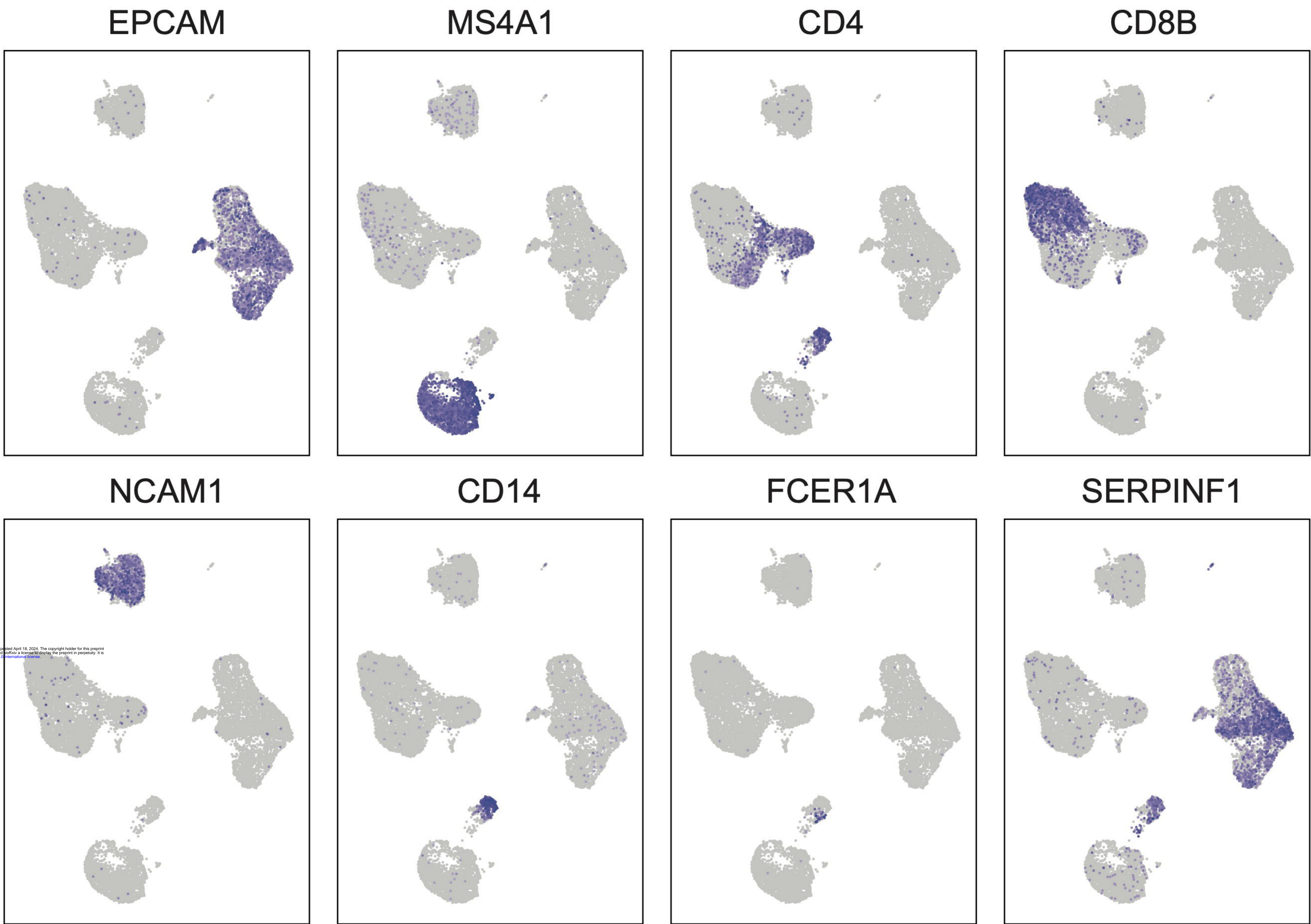
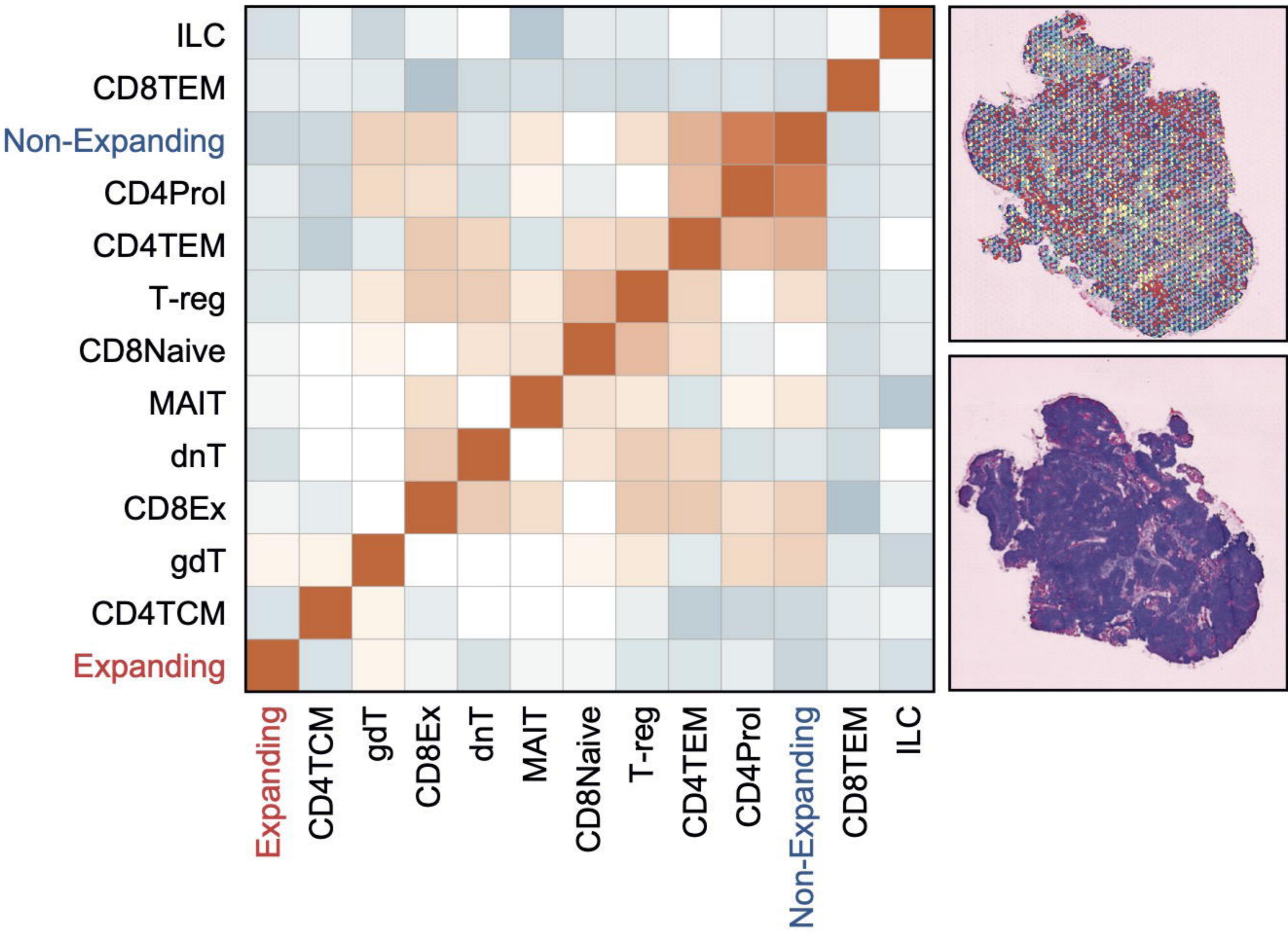
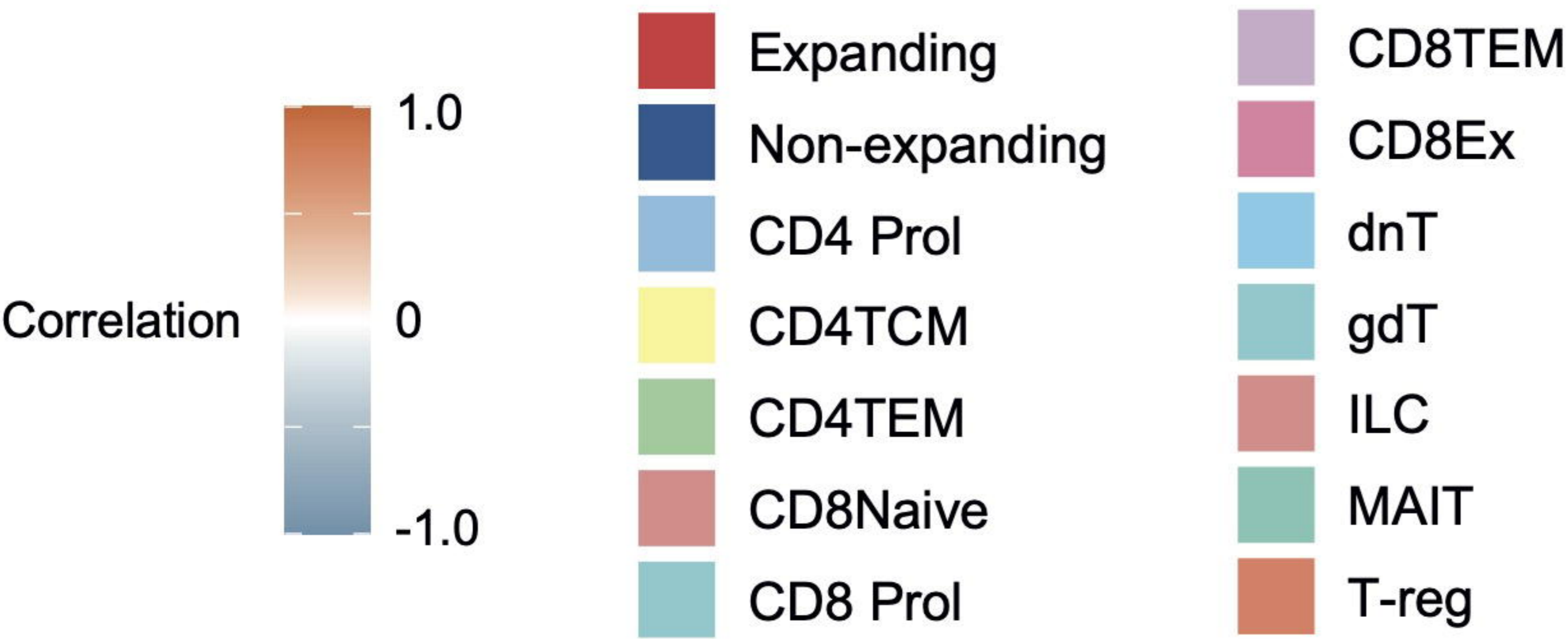


Figure 6

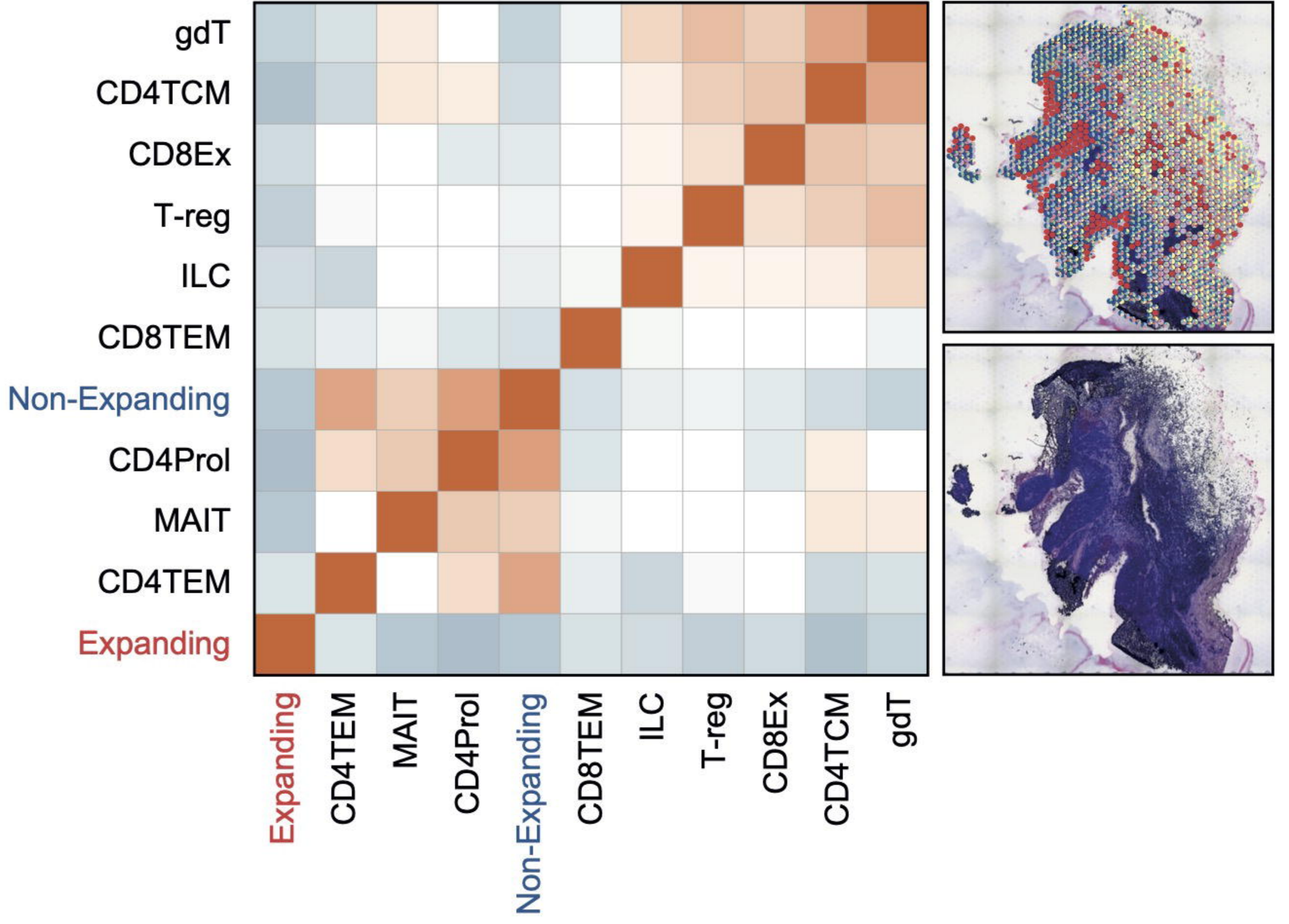
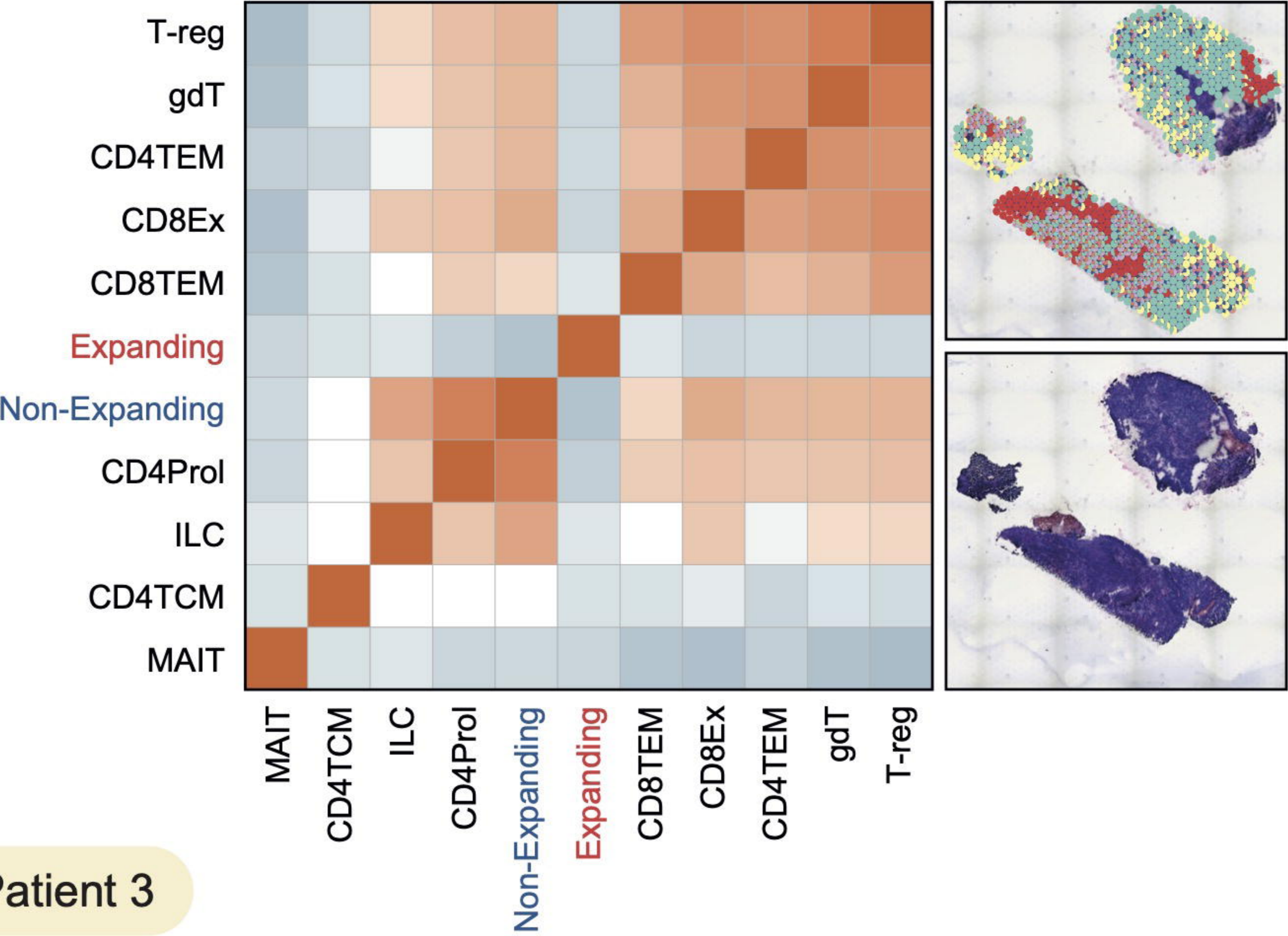
Primary

Lymph node

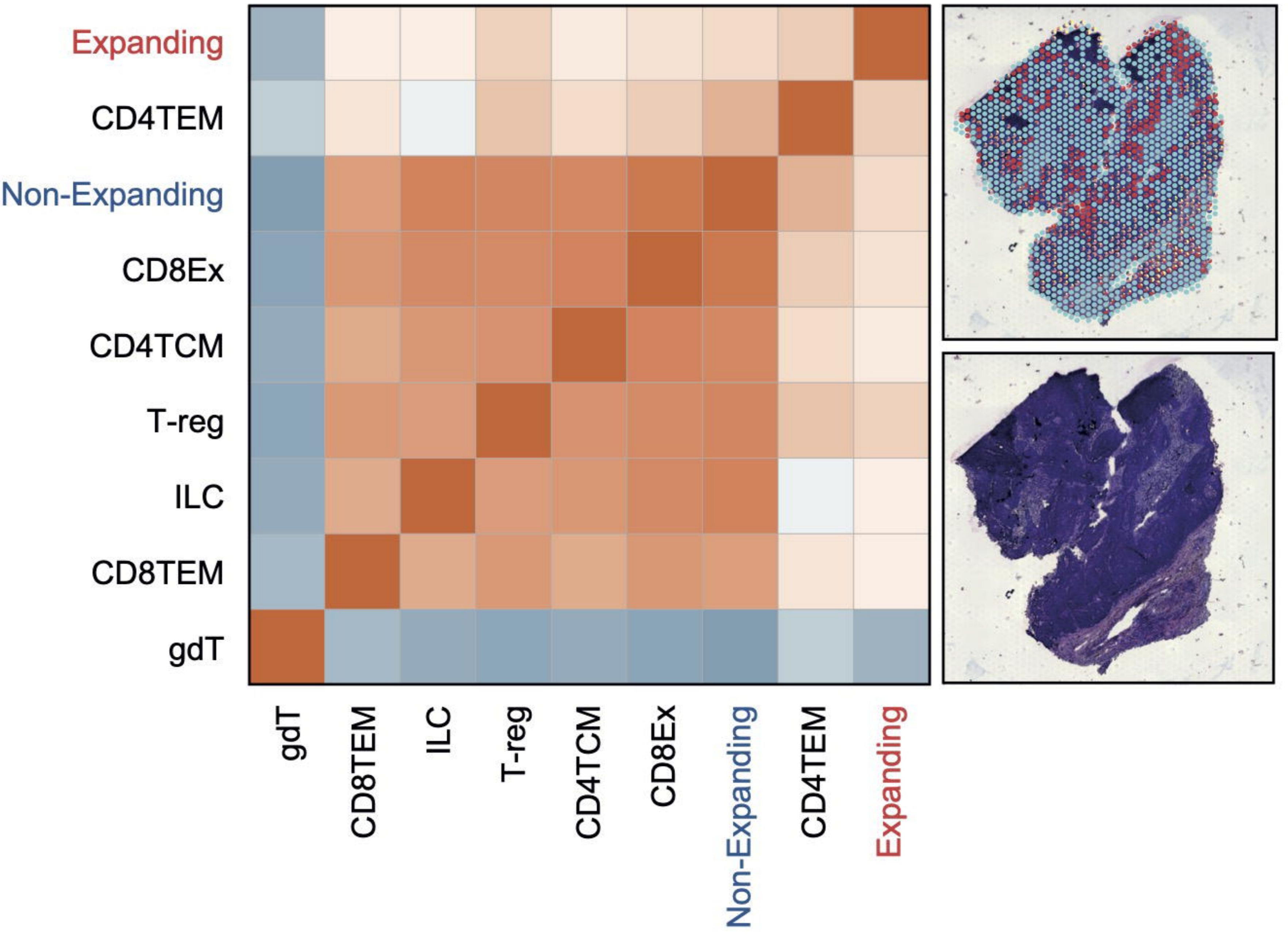
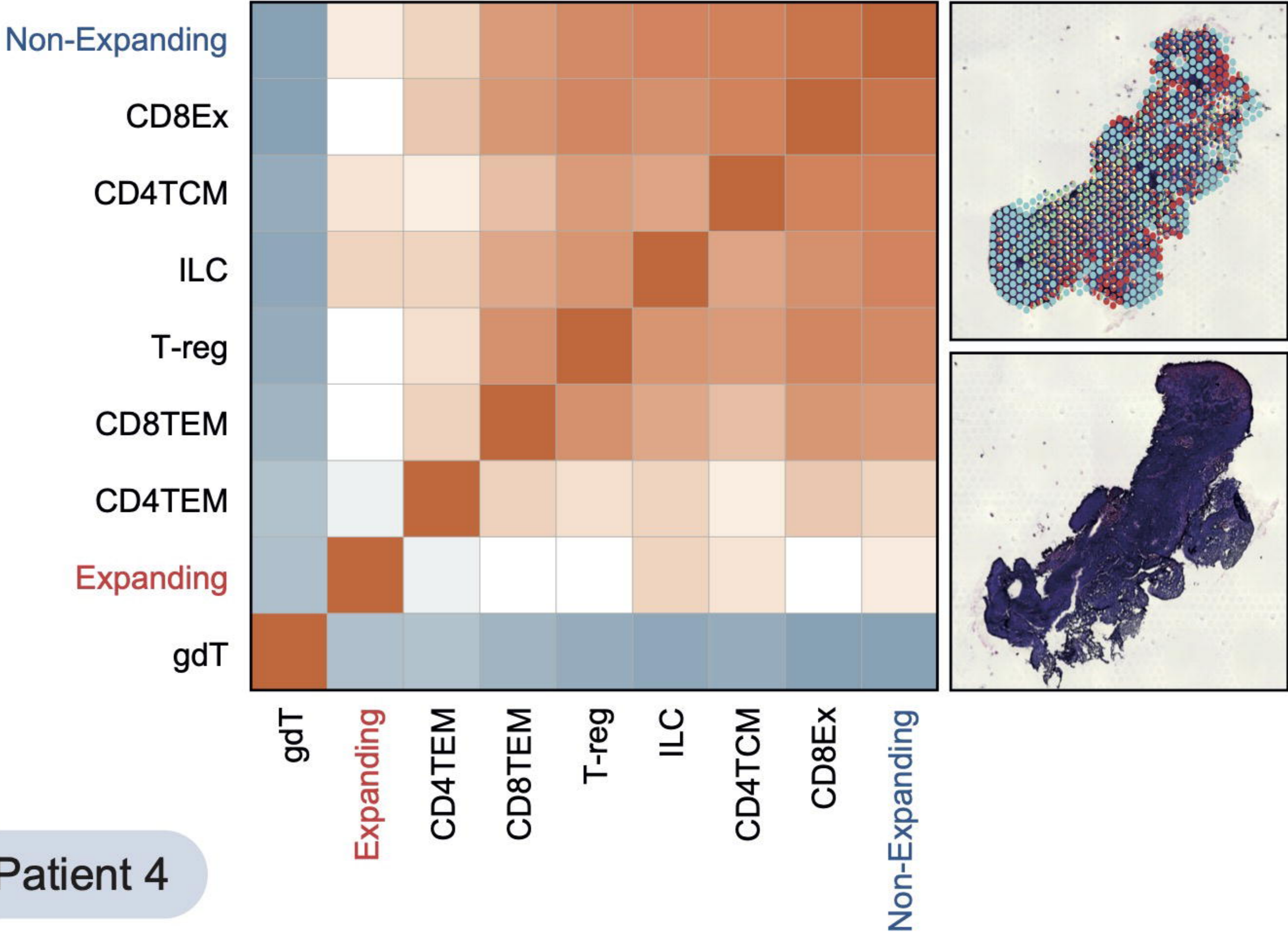
Patient 1



Patient 2



Patient 3



Patient 4

