Spatial transcriptomics reveals the architecture of the tumor/microenvironment interface

Miranda V. Hunter[1]*, Reuben Moncada[2]*, Joshua M. Weiss[1,3], Itai Yanai[2†], Richard M. White[1,4†]
[1] Cancer Biology and Genetics, Memorial Sloan Kettering Cancer Center, New York, NY.
[2] Institute for Computational Medicine, NYU Langone Health, New York, NY.
[3] Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program, Memorial Sloan Kettering Cancer Center, New York, NY.
[4] Lead Contact
*Equal contribution
[†] Correspondence: itai.yanai@nyulangone.org (I.Y); whiter@mskcc.org (R.M.W)

SUMMARY
During tumor progression, cancer cells come into contact with new cell types in the microenvironment, but it is unclear how tumor cells adapt to new environments. Here, we integrate spatial transcriptomics and scRNA-seq to characterize tumor/microenvironment interactions during the initial steps of invasion. Using a zebrafish model of melanoma, we identify a unique “interface” cell state at the tumor/microenvironment boundary. This interface is composed of specialized tumor and microenvironment cells that upregulate a common set of cilia genes, and cilia proteins are enriched only where the tumor contacts the microenvironment. Cilia gene expression is regulated by ETS-family transcription factors, which normally act to suppress cilia genes outside of the interface. An ETS-driven interface is conserved across ten patient samples, suggesting it is a conserved feature of human melanoma. Our results demonstrate the power of spatial transcriptomic approaches in uncovering mechanisms that allow tumors to invade into the microenvironment.

KEYWORDS
melanoma, zebrafish, tumor microenvironment, spatial transcriptomics, cilia, cell-cell interaction

INTRODUCTION
The tumor microenvironment is increasingly appreciated to play a major role in cancer phenotypes, including proliferation, invasion, metastasis and drug resistance (Quail and Joyce, 2013; Ungefroren et al., 2011). As tumors grow and invade into new tissues, they come into contact with a diverse array of microenvironment cell types, but it is poorly understood how these cell-cell interactions allow for successful invasion. In melanoma, these interactions between the tumor and TME include a diverse number of cell types including immune cells (Passarelli et al., 2017), adipocytes (Zhang et al., 2018), fibroblasts (Kaur et al., 2016), and keratinocytes (Kim et al., 2017), all of which have been shown to play roles in tumor growth and metastasis.

Melanomas exhibit a high degree of transcriptomic heterogeneity, which has now been extensively characterized using bulk and single-cell RNA-sequencing approaches (Cancer

Genome Atlas Network, 2015; Tirosh et al., 2016; Winnepenninckx et al., 2006). However, by their nature studies of tumor heterogeneity largely omit the TME from their analyses, making it difficult to investigate tumor invasion using existing datasets. Furthermore, bulk and single-cell RNA-seq require dissociation of the tissue of interest, resulting in a loss of spatial information. Thus, a comprehensive understanding of the transcriptional landscape of tumor invasion *in vivo* is lacking, at least in part due to the limitations of current RNA-sequencing technologies.

Spatial transcriptomics (ST) has recently emerged as a way to address the limitations of both bulk and single-cell RNA-seq, by using spatially-barcoded microarray probes to capture and sequence mRNA from frozen tissue sections (Ståhl et al., 2016). Although the resolution of ST is limited by the diameter of each ST spot (55 µm), we recently developed a technique to integrate ST and scRNA-seq to map the transcriptomic architecture of tumors at true single-cell resolution (Moncada et al., 2020). This provides a unique opportunity to understand the mechanisms that are driving the unique cell-cell interactions that occur at the tumor/TME border.

Here, we integrate ST and scRNA-seq to characterize the transcriptional landscape of melanoma invasion into the TME *in vivo* at cellular resolution. Using a zebrafish model of melanoma, we construct a spatially-resolved gene expression atlas of transcriptomic heterogeneity within tumors and their microenvironment, and discover a histologically invisible but transcriptomically unique “interface” region where tumors contact neighbouring tissues, composed of specialized tumor and muscle cells. We uncover enrichment of cilia genes and proteins at the interface, and find that ETS-family transcription factors regulate cilia gene expression specifically at the interface. We further demonstrate that this unique “interface” transcriptional state is conserved in human melanoma, providing a conserved mechanism that provides new opportunities for halting melanoma invasion and progression.

RESULTS

Spatial transcriptomics reveals the architecture of the melanoma/TME interface

To investigate the transcriptional landscape of melanoma invasion into the TME *in vivo*, we processed frozen sections from adult zebrafish with large, invasive *BRAF*^{V600E}-driven melanomas (Patton et al., 2005) for spatial transcriptomics, using the 10X Genomics Visium platform (**Fig. 1a-c**). Although the size of the tissue section used for ST is limited by the size of the ST array (6.5 mm^2), zebrafish allows us the unique advantage that a transverse section through an adult fish (~5 mm diameter) fits in its entirety on a ST array (**Fig. 1c**). Zebrafish are thus one of the only vertebrate animals that can be used to study both the tumor and all surrounding tissues in their intact forms, without any need for dissection. Our ST dataset contained transcriptomes for 7,281 barcoded array spots across three samples, encompassing 17,317 unique genes (**Fig. 1d-e** and **S1**). We detected approximately 1,000 - 15,000 transcripts (unique molecular identifiers, UMIs) and 500 - 3,000 unique genes per spot (**Fig. S1**).

We first combined our expression matrices using an anchoring framework to identify common cell states across different datasets (Stuart et al., 2019) (**Fig. S2a**). After community-detection based clustering on our integrated dataset, we inferred the identities of 13 distinct clusters.

When we projected the cluster assignments back onto the tissue coordinates (**Fig. 1d**) and onto the UMAP embeddings for each spot (**Fig. 1e**), we found complex spatial patterns in the data that strongly recapitulated tissue histology. Our ST data captured multiple microenvironment cell types (muscle, liver, brain, skin, pancreas, heart, intestine, gills) in addition to the *BRAF*^{V600E}-driven melanomas (**Fig. 1d-f**). We validated our cluster assignments by plotting onto the ST array the expression of marker genes that should be expressed exclusively in the tumor (*BRAF*^{V600E}), muscle (*pvalb4*), heart (*kcn6a*) and nervous system (*mbpa*), and observed that expression of these marker genes was restricted to the expected regions of the tissue (**Fig. 1f**).

To further characterize the transcriptional architecture of the microenvironment, we asked whether we could leverage publicly available, annotated gene sets to uncover spatially-organized patterns of biological activity across the tissue. To this end, we computed the mean expression of genes associated with all zebrafish Gene Ontology (GO) terms, and measured the distance between spots that highly express these genes, reasoning that shorter distances between spots may represent underlying spatial organization of these biological pathways across the tissue. We then compared this distribution to that of a null distribution of distances between random spots, allowing us to identify GO terms with spatially coherent, non-random expression patterns. Applying this to the tumor region of our samples, we identified several GO terms displaying interesting spatial expression patterns related to tissue structure (GO: extracellular structure organization; $p = 2.3\text{x}10^{-8}$) and the immune system (GO: macrophage migration, $p = 7.1\text{x}10^{-4}$), among others (**Figs. 1g** and **S3a**; **Video S1**). We performed the same analysis on the microenvironment, and found several notable spatially-organized pathways that function in tumor growth and invasion (GO: lipid import into cell, $p = 1.2\text{x}10^{-96}$; GO: IMP biosynthetic process, $p = 2.0\text{x}10^{-40}$) (**Figs. 1h** and **S3b**; **Video S2**). Together, these data validate our spatial transcriptomics workflow and demonstrate the existence of discrete tumor and microenvironment regions within our ST dataset.

The tumor-microenvironment interface is transcriptionally unique from the surrounding tissues

We noticed in all of the samples a transcriptionally distinct cluster of array spots that localized to the border between the tumor and the adjacent microenvironment (**Fig. 1c-e**), in which specific biological pathways were upregulated (**Fig. 1h**). This “interface” cluster was present in all three samples (**Fig. S2b-c**). Interestingly, the tissue in this interface region appeared largely indistinguishable from the surrounding microenvironment (**Fig. 2a**), despite it being transcriptionally unique (**Fig. 1e**). We thus hypothesized that this interface cluster represented the region in which the tumor was invading into the surrounding microenvironment. To get a better sense of the transcriptional profile of the interface cluster, we computed the correlation between the averaged transcriptomes of each ST cluster across all 3 samples. We found that the transcriptional profile of the interface cluster was more correlated with the tumor ($R = 0.33$) than with muscle ($R = 0.06$) (**Fig. 2b**), despite the fact that the tissue in this region histologically resembles muscle with few tumor cells visible (**Fig. 2a**).

We next sought to identify genes that may differentiate the interface from muscle (to which it is

most similar histologically) and from tumor (to which it is most similar transcriptionally). We found a number of genes that were upregulated specifically in the interface cluster, including, interestingly, a number of uncharacterized genes (**Fig. 2c**). To identify gene expression programs that are enriched specifically at the interface and provide further evidence for the interface as a transcriptionally distinct tissue region, we performed non-negative matrix factorization (NMF) on all microenvironment spots (including both interface and muscle clusters) across all samples (see Methods). When we projected the NMF factor scores onto each spot, we found several discrete NMF factors specifically enriched at the interface between the tumor and the microenvironment (factors 8, 10 and 15; **Figs. 2d** and **S4**). To investigate the biology underlying the genes contributing to these factors, we looked for significantly enriched GO terms among the top 150 genes contributing to each factor (**Figs. 2e** and **S4**). This revealed that the interface factors were enriched in genes functioning in actomyosin contractility (factor 15; **Fig. 2e**), nuclear biology (factor 8; **Fig. S4**), and membrane trafficking (factor 10; **Fig. S4**). Together, these data uncover a distinct "interface" region bordering the tumor, which histologically resembles the microenvironment, transcriptionally resembles tumor, but expresses unique gene modules that likely promote tumor invasion.

The tumor-microenvironment interface is composed of specialized tumor and muscle cells

Our ST results so far detail a transcriptionally distinct "interface" region where tumors contact the microenvironment. However, spatial transcriptomics data is limited in resolution by the diameter of each spot on the ST array (55 µm). Thus, each ST spot probably captures transcripts from multiple cells. As the interface region is, by nature, likely a mixture of tumor and microenvironment cells, we performed single-cell RNA-seq (scRNA-seq) on tumor and non-tumor cells from adult zebrafish with large melanomas (**Fig. 3a**) in order to better define the cell types present in the interface. As expected, our scRNA-seq data contained tumor cells as well as various microenvironment cell types, including erythrocytes, keratinocytes, and several types of immune cells (**Fig. 3b**). Consistent with our ST results, clustering of our scRNA-seq data revealed a distinct "interface" cell cluster (**Fig. 3b-d**), which we identified based on the fact that cells in this cluster significantly upregulated the same genes that were upregulated in our ST interface cluster ($p = 1.83 \times 10^{-26}$; **Fig. 3c**). Strikingly, principal component analysis of the interface cluster revealed two distinct cell populations, one expressing tumor markers such as *BRAF*^{V600E} and the other expressing muscle genes such as *ckba*, with other genes such as the centromere gene *stra13* upregulated in both populations (**Fig. 3e**). This result confirms that the transcriptionally distinct "interface" region we identified in our ST data is actually composed of two similar, but distinct cell types: a "tumor interface" and a "muscle interface". Based on this, we separated the interface cluster into two subclusters, and confirmed that the two subclusters express anti-correlated levels of tumor markers such as *BRAF*^{V600E}, *mitfa*, and *pmela*, and muscle markers such as *ckba*, *neb* and *ak1* (**Fig. 3f**). The presence of putative "muscle" cells in the interface is particularly notable, in light of the fact that adult skeletal muscle is composed of multinucleated muscle fibers that we were unable to isolate in our scRNA-seq workflow due to their size, evidenced by the lack of a muscle cell cluster in our dataset (**Fig. 3b**). Together, these

data reveal that the interface is composed of specialized tumor and microenvironment cells, which together upregulate a common gene program to mediate tumor invasion.

Cilia pathways and genes are enriched in the interface

To gain further insight into the biological processes underlying the specialized “interface” region identified in our ST and scRNA-seq data, we performed pre-ranked gene set enrichment analysis (GSEA), using differentially expressed genes in the scRNA-seq interface cluster, to identify conserved pathways that may promote melanoma invasion. Strikingly, all of the top 10 enriched GO biological process pathways in the interface were related to cilia (**Fig. 4a**), and there was also a clear enrichment of cilia-related GO cellular component terms in the interface (**Fig. 4b-c**). Cilia-related GO terms were also enriched in the ST interface (**Fig. 4d**). Several recent studies have implicated cilia in an important role in melanoma initiation and progression, although the mechanism by which cilia mediate melanoma progression is unknown (Choudhury et al., 2020; Kim et al., 2011; Lang et al., 2020; Snedecor et al., 2015; Zingg et al., 2018). To further investigate a role for cilia in melanoma invasion, we scored each cell (scRNA-seq) for relative enrichment of cilia genes, using the “gold standard” SYSCILIA gene list (van Dam et al., 2013), and quantified a significant upregulation of cilia genes in the interface cells in our scRNA-seq data (**Fig. 4e**). Furthermore, of the top 100 genes enriched in the scRNA-seq interface, 38 were related to cilia, all of which were upregulated by more than 7-fold in the interface (**Table S1**). Together, these results suggest an important role for cilia in mediating melanoma invasion.

Cilia proteins are polarized at the tumor-microenvironment interface

Interestingly, previous studies have shown that human and mouse melanomas are not ciliated, although they express cilia genes (Choudhury et al., 2020; Kim et al., 2011; Lang et al., 2020; Snedecor et al., 2015; Zingg et al., 2018). Similarly, when we stained zebrafish tissue sections for acetylated tubulin, a common cilia marker (Kim et al., 2011), we saw a significant reduction in acetylated tubulin signal within the tumor region, compared to the surrounding tissue (p = 0.0482, **Fig. 4f-g**). Immunofluorescence staining of a zebrafish melanoma cell line (ZMEL1) (Heilmann et al., 2015; Kim et al., 2017) for the cilia markers ARL13B (Caspary et al., 2007; Cevik et al., 2010) and acetylated tubulin further confirmed that zebrafish melanoma cells are not ciliated *in vitro* (**Fig. S5a**), although they also express cilia genes at high levels both *in vitro* and when transplanted into adult zebrafish (**Fig. S5b**). To reconcile these confounding models, we examined localization of cilia markers at the invasive front between tumors and muscle. We found a specific enrichment of both ciliary and basal body markers specifically at the interface between tumors and the microenvironment (**Fig. 4h-i**). In concordance with our scRNA-seq data, this accumulation of cilia markers seems to occur in both tumor and microenvironment cells (**Fig. 4h-i,** insets). These data suggest that although the bulk of primary melanomas is not ciliated, cilia are important in mediating communication and invasion of tumors into the surrounding microenvironment.

ETS-family transcription factors regulate cilia gene expression at the interface

Next, to identify potential regulators of gene expression within the interface, we performed HOMER motif analysis (Heinz et al., 2010) to identify conserved transcription factor (TF) binding motifs enriched in genes differentially expressed in the interface. When we performed motif analysis on genes differentially expressed in the ST interface compared to normal muscle, the top-ranked motif was the highly conserved ETS DNA-binding domain, containing a core GGAA/T sequence ($p = 1x10^{-22}$; **Fig. 5a**). The ETS domain was also the top-ranked motif enriched in genes differentially expressed in the ST interface compared to all other ST spots ($p = 1x10^{-15}$), and was the second-ranked motif enriched in genes differentially expressed in the interface cluster identified in our scRNA-seq dataset ($p = 1x10^{-13}$; **Fig. 5a**). Although ETS-family transcription factors have not been widely studied in melanoma, they have been reported to function in melanoma invasion (Rothhammer et al., 2004) and phenotype switching (Wouters et al., 2020), and are aberrantly upregulated in many types of solid tumors (Sizemore et al., 2017). Interestingly, zebrafish ETS-family transcription factors were downregulated in the interface in both our scRNA-seq and ST datasets (**Fig. 5b-c**).

To identify potential biological processes that could be regulated by ETS transcription factors at the tumor-microenvironment interface, we investigated putative target genes containing an ETS motif in their promoter. We queried the zebrafish genome for genes with an ETS motif within 500 kb of the transcription start site, filtered these genes to include only those differentially expressed in the tissue/cell type of interest, and performed GSEA on the resulting target gene lists. Surprisingly, within the ETS-target genes in both the ST and scRNA-seq interface clusters, we again found an enrichment of pathways related to cilia (**Fig. 5d-e**). Of the top 10 GO cellular component terms enriched in ETS-target genes in the ST interface, 5 were directly related to cilia or ciliary components (microtubules, centrioles) (**Fig. 5d**). In the interface cells identified in our scRNA-seq dataset, 6 of the top 10 GO biological process terms enriched in ETS-target genes were related to cilia or ciliary components **(Fig. 5e**). As ETS TFs are downregulated specifically in the interface (**Fig. 5b-c**), this suggests that ETS-family TFs may act as a transcriptional repressor of cilia genes. In support of this model, when we scored each cell in the interface for relative expression of both ETS genes and ETS-target genes, the two were strongly anti-correlated ($R = -0.625$, $p = 9.02x10^{-27}$, **Fig. 5f**). Collectively, these data indicate that ETS-family transcription factors act as transcriptional repressors of cilia genes in cells at the interface between tumors and the microenvironment, where upregulation of cilia likely mediates tumor invasion.

Interface cells are conserved in human melanoma

Finally, we investigated whether “interface” cells, distinguished by upregulation of cilia genes and downregulation of ETS transcription factors, are also found in human melanoma. We analyzed a recently published human melanoma scRNA-seq dataset, composed of 4,323 cells from nine patient-derived cultures (MM lines) and a well-characterized human melanoma cell line (A375) (Wouters et al., 2020) (**Fig. 5g**). To identify putative “interface” cells in this human melanoma dataset, we scored each cell for expression of the human orthologs of the genes upregulated by more than 1.5-fold in our zebrafish scRNA-seq interface cluster. Strikingly, in all ten human samples there was a distinct subset of cells that upregulated these interface marker

genes (**Fig. 5h**). Based on this, we separated our human dataset into “interface” cells (cells upregulating interface markers), and the remaining tumor cells. When we quantified both cilia and ETS gene expression across the tumor and interface cells, we found the same anti-correlated relationship between ETS and cilia genes in human melanoma interface cells, in which cilia genes were upregulated and ETS genes were downregulated in the interface (R = -0.076, $p = 8.6 \times 10^{-5}$; **Fig. 5i-j**). This demonstrates that the presence of a transcriptionally unique tumor-microenvironment interface is conserved in human melanoma. Together, our results indicate that melanoma invasion is accomplished by a subset of specialized tumor and muscle cells, which together upregulate a common gene program to enable tumor growth.

DISCUSSION

Here, we combined spatial and single-cell transcriptomics approaches to characterize how tumor cells interface with new tissues in the TME, revealing key regulators of how this interface is formed. We analyzed a total of 14,493 transcriptomes encompassing expression of 20,589 unique genes from 7,281 ST array spots and 7,212 single cells from both zebrafish and human samples. Our results identified a series of spatially-patterned gene modules, some of which specifically localize to the interface between tumors and surrounding tissue. We showed that the interface is composed of specialized tumor and muscle cells, which are distinguished by upregulation of cilia genes and proteins. We further show that ETS transcription factors regulate expression of cilia genes at the interface, and that a unique “interface” cell population is conserved across ten human melanoma samples. Together, our results reveal a novel “interface” transcriptional state that mediates melanoma invasion.

Our results identify a novel role for ETS-family transcription factors in mediating cilia gene expression at the interface to promote melanoma invasion. Interestingly, in recent years cilia have been implicated in multiple facets of melanoma biology, but their role in melanoma progression is still unclear. Melanomas are not ciliated (Kim et al., 2011; Snedecor et al., 2015) (**Fig. 4f-g**), and in fact the “ciliation index” is gaining in prominence as a diagnostic tool to distinguish melanomas from benign nevi (Choudhury et al., 2020; Lang et al., 2020). Furthermore, cilia disassembly has recently been implicated in melanoma metastasis (Zingg et al., 2018), in which deconstruction of cilia, regulated by EZH2, drives metastasis. The paradox is that while most melanoma cells are not ciliated, many melanomas still express cilia genes (**Figs. 4a-e** and **S5b**). Our data adds a new layer onto this complexity, in that we find that not only are cilia genes upregulated specifically at the interface between tumor and TME, but more importantly that only cells at that interface express high levels of cilia proteins. This raises the still not fully answered question of what role cilia play in various steps in melanoma progression. In primary melanoma growth, it is clear that most cells are unciliated, and that EZH2 acts to suppress those genes. Loss of cilia via EZH2 clearly increases metastasis in these models via enhanced Wnt/β-catenin signaling (Zingg et al., 2018). Our finding that most melanoma cells do not have cilia is consistent with this finding, but yet we find a specific subset of cells at the interface that upregulate cilia genes and protein, and these cells appear to be present in human melanoma as well.

How to reconcile these seemingly conflicting pieces of data? Our data would suggest that intact cilia may be most important when they are first encountering new, heterotypic cell types in the TME. We can envision several different possibilities to why cilia are upregulated specifically at this interface. First, this upregulation of cilia genes and proteins at the interface may be transient, induced by heterotypic cell-cell interactions between tumor and muscle. Primary cilia are critical signalling hubs for the cell, and regulate signaling pathways such as Hedgehog and TGF-β/BMP (Anvarian et al., 2019), all of which are important in cancer progression (Hassounah et al., 2012) and cell-cell communication. A second possibility is that the primary cilia are acting as mechanotransducers, and play a role in directional migration of the cells as they invade into new tissues. For example, seminal work on primary cilia demonstrated that cilia can orient in the direction of migration in 3T3 cells (Albrecht-Buehler, 1977), which has been also seen in the context of wound healing (Christensen et al., 2008; Schneider et al., 2010). Finally, it is possible that the emergence of cilia at the interface is actually acting as a barrier to systemic metastatic dissemination, and that heterotypic interactions between melanoma and muscle might be restraining progression. It is notable that our zebrafish melanomas metastasize at a low rate, and in fact skeletal muscle (where we most easily see the interface) is a rare site of metastasis in humans, consistent with this possibility. A major endeavor for future studies will be to delineate how cilia act at each step of tumorigenesis, what signaling nodes are most critical, and whether they act as a barrier or enabler of metastasis.

Our results uncovered a novel role for ETS-family transcription factors in melanoma, as transcriptional repressors of cilia genes. Although most ETS TFs can function as transcriptional activators, at least four ETS TFs also have repressor activity (Mavrothalassitis and Ghysdael, 2000). Despite the fact that ETS TFs have a well-characterized role in several types of solid tumors, their role in melanoma has not been studied in depth, although a recent study found that ETS TFs induce a UV damage signature that correlates with increased mutational burden in human melanoma (Mao et al., 2018). ETS TFs broadly function in various facets of tumorigenesis, including DNA damage, metabolism, self-renewal, and remodelling of the TME (Sizemore et al., 2017). However, most if not all of these situations have been found to be induced by aberrant upregulation of ETS genes. Conversely, we found a role for downregulation of ETS TFs specifically where tumors contact the TME. It is still unclear what triggers this downregulation of ETS genes in such a spatially-restricted region. Despite their role as transcription factors, ETS proteins also participate in a wide range of protein-protein interactions, and their activity is regulated through phosphorylation as a result of signaling cascades (Li et al., 2000). MAPK signaling has been reported to regulate ETS (Wasylyk et al., 1998), and the MAPK pathway is frequently activated in melanoma (Fecher et al., 2008). It is unclear if MAPK or other signaling pathways display spatially-restricted patterns of activation within tumors and/or the microenvironment, but the advent of ST techniques will help to address these questions.

Although it was not a focus of our study, our ST dataset also uncovered spatially-organized transcriptomic heterogeneity within the tumor itself (**Figs. 1f-g**). In recent years, the advent of single-cell transcriptomics approaches has identified a substantial degree of transcriptomic

heterogeneity in most if not all types of cancer (González-Silva et al., 2020). Tumor heterogeneity often increases as tumors progress, and may be a predictor of poor clinical outcomes as it is believed to be a major contributor to drug resistance (Dagogo-Jack and Shaw, 2018). Investigating the underlying cause and complex clonal relationships within different tumor cell subtypes has proven to be challenging for many reasons, one of which being a lack of information regarding the spatial patterning of this heterogeneity. Our dataset acts as a proof-of-principle for the use of spatial transcriptomics in identifying spatially-organized tumor heterogeneity, and lays the groundwork for future studies using our dataset or others to explore the basis of this spatial patterning.

Our study is, to our knowledge, the first spatially-resolved gene expression atlas of the interface between the tumor and its microenvironment. Although we uncovered many genes, pathways and gene modules that exhibit novel spatial patterns within the tumor and/or TME, there are likely many more interesting biological phenomena in our dataset that we have yet to identify. Recently, deep learning methods have been applied to histopathology images to uncover spatially-resolved predictions of molecular alterations, mutations, and prognosis (Fu et al., 2020; Kather et al., 2020). A logical next step would be extension of these approaches to integrate deep learning and pattern recognition algorithms with ST data, to identify interesting spatial patterns of gene expression and also predict transcriptomes based on histopathology. Ultimately, integration of transcriptomics, histopathology, and deep learning techniques will allow us to expand the utility of both ST and histological datasets and broaden our understanding of cancer cell interactions *in vivo*.

ACKNOWLEDGEMENTS

We thank S. Selvaraj and B. Dabovic from the NYU Experimental Pathology Core for technical assistance with imaging for spatial transcriptomics experiments, R. Luther and M. Hogan from the Maurano lab at NYU for assistance with sequencing, and the members of the Yanai and White labs for useful discussions and technical support. M.V.H was funded by a postdoctoral fellowship from the Canadian Institutes of Health Research. J.M.W. was supported by a NIH Kirschstein-NRSA predoctoral fellowship (F30CA236442), a NIH predoctoral fellowship (T32GM008539) from the Cell and Developmental Biology Program at Weill Cornell Graduate School, and a NIH Medical Scientist Training Program grant (T32GM007739). R.M.W was funded by grants from the Melanoma Research Alliance, The Debra and Leon Black Family Foundation, NIH Research Program grants R01CA229215 and R01CA238317, NIH Director's New Innovator Award DP2CA186572, The Pershing Square Sohn Foundation, The Mark Foundation, The Alan and Sandra Gerry Metastasis Research Initiative at Memorial Sloan Kettering Cancer Center, The Harry J. Lloyd Foundation, Consano, and the Starr Cancer Consortium.

AUTHOR CONTRIBUTIONS

M.V.H, R.M, I.Y and R.M.W conceived the study. M.V.H and R.M performed all experiments and data analysis. J.M.W generated transgenic fish. M.V.H and R.M wrote the manuscript and all authors provided feedback before submission.

DECLARATION OF INTERESTS

R.M.W is a paid consultant to N-of-One Therapeutics, a subsidiary of Qiagen. R.M.W is on the scientific advisory board of Consano, but receives no income for this. R.M.W receives royalty payments for the use of the *casper* zebrafish line from Carolina Biologicals.

FIGURE LEGENDS

Figure 1. Spatial transcriptomics reveals the transcriptional architecture of the melanoma microenvironment.

a. Schematic showing the spatial transcriptomics (ST) experiment workflow. **b.** Images of zebrafish with *BRAF*^{V600E}-driven melanomas used for ST. The region where the fish were sectioned is highlighted. **c.** H&E staining of cryosections used for ST. **d.** ST spots colored by clustering assignments of the integrated dataset (see Methods). **e.** UMAP embedding of ST spots from all three samples colored by cluster assignments of the integrated dataset (see Methods). **f.** The expression of select marker genes (*BRAF*^{V600E}, tumor; *pvalb4*, muscle; *kcn6a*, heart; *mbpa*, nervous system) from the ST data projected over tissue space (left), with images of the corresponding histology from the indicated region of the ST array (right). **g-h.** Average, standardized expression of annotated genes for gene ontology (GO) terms displaying spatially-coherent expression patterns in the tumor (g) and microenvironment (h) regions in each sample. *p*-values represent the comparison between the distance between spots expressing that GO term genes and a null-distribution of distances between random spots (Wilcoxon's Rank Sum test, see Methods).

Figure 2. The tumor-microenvironment interface is transcriptionally distinct from the surrounding microenvironment.

a. Interface and muscle-annotated cluster spots projected onto tissue image. Insets show the tissue underlying the interface spots (1) and muscle spots (2). **b.** Correlation matrix between average expression profile of ST clusters across all three datasets. Clusters are ordered by hierarchical clustering of the Pearson's correlation coefficients (see Methods) and bubble sizes correspond to *p*-value (-log10) of correlation, with *p*-values < 10^{-3} omitted. Clustering of tumor and interface together is highlighted in the dendrogram (red). **c.** Volcano plot of differentially expressed genes between the interface cluster versus the muscle and tumor clusters. *p*-values were obtained from the Wilcoxon's rank sum test. **d.** Non-negative matrix factorization (NMF) of the microenvironment spots (muscle and interface clusters). Shown are the standardized factor scores for interface-specific NMF factor 15, projected onto microenvironment spots. **e.** Enriched GO terms for the top 150 scoring genes in NMF factor 15.

Figure 3. The tumor-microenvironment interface is composed of specialized tumor and muscle cells.

a. Schematic showing scRNA-seq experiment workflow. **b.** UMAP dimensionality reduction plot for 2,889 cells sequenced as in (a). Cluster/cell type assignments are labelled and colored. **c.** Expression score per cell (scRNA-seq) for average expression of interface marker genes from the ST interface cluster. **d.** Heatmap showing expression of the top 50 genes upregulated in the

tumor cell cluster (top, orange) and interface cell cluster (bottom, yellow). Selected genes are labelled. **e.** Principal component analysis of cells in the interface cluster, scored for expression of the tumor marker *BRAF*^{V600E}, the muscle marker *ckba*, and the centromere gene *stra13*. Scores for principal components 1 and 2 are plotted. Cells are labelled by standardized expression of the indicated genes. **f.** Dot plot showing expression of tumor and muscle markers. The size of each dot corresponds to the percentage of cells in that cluster expressing the indicated gene, and the colour of each dot indicates the expression level.

Figure 4. Cilia genes and proteins are enriched at the tumor-microenvironment interface.
a. Bar graph showing the top 10 GO biological process terms enriched in differentially expressed genes in the scRNA-seq interface cluster. GO terms are colored by p-value. **b.** Waterfall plot showing the top and bottom 250 GO cellular component terms by NES in the scRNA-seq interface cluster, with cilia GO terms labelled in red. **c.** GSEA barcode plot showing enrichment of genes in the GO: CILIUM pathway in the scRNA-seq interface cluster. NES and FDR are indicated. **d.** Bar graph comparing normalized enrichment scores for cilia-related GO terms in the scRNA-seq (purple) and ST (green) interface clusters. **e.** Violin plots showing expression of fish SYSCILIA genes in the scRNA-seq interface cluster. **f.** Violin plot showing normalized cilia fluorescence intensity within and outside the tumor region. *, $p < 0.05$, paired t-test ($n = 6$ fish). **g.** A section through an adult zebrafish with an invasive melanoma immunostained for GFP (tumor cells), acetylated tubulin (cilia) and Hoescht (nuclei). Arrows denote cilia in the microenvironment. Scale bar, 200 μm. **h-i.** Sections from adult zebrafish with invasive melanomas stained with antibodies against the cilia markers acetylated tubulin and ARL13B, and the basal body marker γ-tubulin, imaged at 10X (**h**) and 63X (**i**) magnification. Scale bars, 200 μm (10X) and 50 μm (63X).

Figure 5. ETS transcription factors regulate cilia gene expression at the interface in zebrafish and human melanoma.
a. Results from HOMER motif analysis of differentially expressed genes in the ST and scRNA-seq interface clusters. **b-c.** Violin plots showing expression of zebrafish ETS-family genes in the tumor and interface clusters from the scRNA-seq (b) and ST (c) datasets. **d-e.** Top 10 enriched GO cellular component terms by normalized expression score (NES) from ETS-target genes in the ST interface cluster. Cilia-related GO terms are labelled in red. **f.** Scatter plot comparing ETS gene expression scores per cell and ETS-target gene expression scores per cell in the scRNA-seq interface cluster. Pearson's correlation coefficient (R) is indicated. **g-h.** UMAP projection of human melanoma scRNA-seq data from Wouters et al. (2020). Cell lines (g) and interface marker gene expression score (h) are indicated. Scoring of interface marker genes was done with the human orthologs of the zebrafish genes upregulated by >1.5-fold in our zebrafish scRNA-seq interface cluster. **i-j.** Expression of human cilia (i) and ETS (j) genes in tumor and interface cells in human melanoma scRNA-seq. **b-c, f, i-j.** The expression score represents the standardized mean expression of all indicated genes per each cell or ST array spot. **, $p < 0.01$, ***, $p < 0.001$ (Wilcoxon rank-sum test).

STAR METHODS

438 **Zebrafish husbandry**
439 Zebrafish lines were maintained at 28.5°C in a dedicated aquatics facility with a 14 hours on/10
440 hours off light cycle. *casper* (White et al., 2008) fish were used for all experiments. Fish were
441 anesthetized with Tricaine (MS-222) at a stock concentration of 4 g/L (pH 7.0). All zebrafish
442 experiments and procedures were carried out in compliance with institutional protocols for
443 vertebrate animals, and were approved by the Memorial Sloan Kettering Cancer Center IACUC
444 (protocol #12-05-008).

445 **Generation of transgenic fish**
446 Transgenic tumor-bearing fish were generated using the *miniCoopR* system as previously
447 described (Ceol et al., 2011; Iyengar et al., 2012). Briefly, *casper* fish with the genotype
448 *mitfa-BRAF^{V600E}; p53$^{-/-}$; mitfa$^{-/-}$* were incrossed, and the resulting 1-cell stage embryos were
449 injected with plasmids containing *mitfa-MITF* and *mitfa-GFP*. Fish were raised to adulthood (4-6
450 months) and screened for the presence of pigmented, GFP-positive tumors.

451 **Cell culture**
452 ZMEL1 cells (Heilmann et al., 2015; Kim et al., 2017) expressing *mitfa*-BRAF^{V600E}/p53$^{-/-}$ and
453 *mitfa*-EGFP were grown in a humidified incubator at 28°C in DMEM supplemented with 10%
454 FBS (Seradigm), 1X penicillin-streptomycin-gluatmine (Gibco), and 1X Glutamax (Gibco).

455 **Adult transplants**
456 Transplantation of ZMEL1 cells into adult fish was performed as previously described (Heilmann
457 et al., 2015; Kim et al., 2017). Briefly, adult *casper* fish were irradiated on two consecutive days
458 on a cesium irradiator at a dose of 15 Gy each day. Fish were then allowed a 3-4 day recovery
459 period. In preparation for transplant, ZMEL1 cells were trypsinized, washed with PBS and
460 resuspended in PBS to a concentration of $5x10^6$ cells/mL. 250,000 cells were injected
461 subcutaneously using a Hamilton syringe in each of the dorsal and ventral fat pads for a total of
462 500,000 transplanted cells/fish. Fish were imaged 2-3 days post-transplant on a Zeiss
463 AxioZoom V16 fluorescence microscope to confirm successful transplant, and tumors were
464 allowed to grow for a total of 10-14 days until a pigmented tumor mass was visible under the
465 skin.

466 **Spatial transcriptomics**
467 ***Sample preparation***
468 Adult tumor-bearing fish were euthanized on ice and washed in 1X PBS. After dissection of the
469 head and tail, the remaining tissue was equilibrated in cold OCT for 2 minutes, before transfer to
470 a tissue mold filled with fresh OCT for snap-freezing in liquid nitrogen-chilled isopentane. Tissue
471 blocks were stored at -80°C. For cryosectioning, both the tissue block and the ST slide were
472 equilibrated inside the cryostat for 15-30 minutes at 16°C before sectioning. Transverse sections
473 through the entire fish were cut at a thickness of 10 µm and immediately placed on the ST slide
474 (Visium Spatial Gene Expression Slides, 10X Genomics). ST slides containing sections were
475 stored at -80°C for a maximum of one week before use.

476 ***Fixation, staining, imaging and construction of cDNA libraries***
477 Samples were processed according to the Visium Spatial Gene Expression User Guide (10X
478 Genomics) and all reagents were from the Visium Spatial Gene Expression Kit (10X Genomics).
479 Briefly, sections were fixed in chilled methanol for 30 min at -20°C, stained with hematoxylin and
480 eosin, and mounted in 85% glycerol for imaging. Imaging was performed on a Leica SCN400 F
481 whole-slide scanner at 40X magnification. After imaging, sections were permeabilized at 37°C
482 for 45 minutes. After permeabilization, the on-slide reverse transcription (RT) reaction was
483 performed at 53°C for 2 hours. Permeabilization time and RT reaction length were determined
484 using the Visium Spatial Tissue Optimization Kit (10X Genomics). Second strand synthesis was
485 subsequently performed on-slide for 15 minutes at 65°C. All on-slide reactions were performed
486 in a thermocycler with a metal slide adaptor plate. Following second strand synthesis, samples
487 were transferred to tubes for cDNA amplification and cleanup. Library quality was assayed using
488 a Bioanalyzer High Sensitivity chip (Agilent).

489 ***Sequencing***
490 10X Genomics Visium libraries were pooled, denatured, and diluted to a loading concentration
491 of 1.8 pM with 1% PhiX control, followed by paired-end sequencing on an Illumina NextSeq 500
492 to a depth of approximately 110 - 180 million paired reads per sample. Sequencing parameters:
493 Read1 28 cycles. i7 10 cycles, i5 10, Read2 120 cycles. Sequencing data was processed using
494 the Space Ranger pipeline v.1.0.0 (10X Genomics).

495 ***Dimensionality reduction and clustering***
496 ST data was processed using R version 3.6.3, Seurat version 3.1.4 (Stuart et al., 2019), Python
497 version 3.6, and MATLAB 2019b. Data was normalized using SCTransform (Hafemeister and
498 Satija, 2019). The 3 ST datasets were integrated using the Seurat SCTransform integration
499 workflow, using 3000 integration features and including all common genes between the 3
500 datasets. Principal component analysis (Jolliffe, 1986) and UMAP dimensionality reduction
501 (McInnes et al., 2018) were done using default parameters. Initial clustering was done using the
502 FindClusters function implemented in the Seurat R package with the resolution parameter = 0.8.
503 Tissue types of each cluster were inferred and clusters were further refined by plotting clusters
504 onto the associated histology images and identifying marker genes using the Wilcoxon's Rank
505 Sum test. Expression scores for ETS and cilia gene sets were calculated using the Seurat
506 function AddModuleScores with default parameters. A list of cilia genes was obtained from the
507 SYSCILIA gold standard list (van Dam et al., 2013). A list of ETS genes was obtained from
508 (Dittmer, 2011).

509 ***Non-negative matrix factorization (NMF)***
510 After normalization and integration of ST data (see *Dimensionality reduction and clustering*),
511 negative values in the integrated expression matrix were set to zero. NMF was performed with a
512 rank of 10. Factor scores were first z-scored across factors prior to plotting onto ST spots.

513 ***Analysis of gene ontology (GO) terms with spatially coherent expression patterns***

514 GO term annotations for *Danio rerio* were downloaded from Biomart (Durinck et al., 2009). For
515 each GO term, the average expression of genes annotated for that GO term was computed. We
516 defined spots that highly express this GO term as spots whose expression level for these genes
517 is above the mean plus two standard deviations (we required the number of these spots to be at
518 least 5 to proceed with the analysis). We then computed the Euclidean distance between these
519 spots. Next, we computed the Euclidean distance between the same number of random spots,
520 and repeated this computation 100 times to generate a null distribution of distances. We then
521 compared the GO term spot distances to the null distribution using a Wilcoxon's Rank Sum test
522 to compute a *p*-value.

523 ***Correlation between ST spots and ST clusters***

524 For computing the correlation across ST clusters, we first computed the average expression of
525 each tissue cluster in the integrated expression matrix of our three datasets. We then used the
526 union of the ~1000 variably expressed genes in each individual dataset to obtain a list of ~2300
527 total variably expressed genes. We then used these genes to compute the Pearson's correlation
528 and associated p-values.

529 ***GSEA and pathway analysis***

530 Lists of differentially expressed genes for pathway analysis were created using the Seurat
531 function FindMarkers using the Wilcoxon rank sum test. Ribosomal genes and genes with
532 *p*-values above 0.05 were removed. Zebrafish genes were converted to their human orthologs
533 using DIOPT (Hu et al., 2011), keeping only human orthologs with a DIOPT score > 6. In cases
534 where there were multiple zebrafish orthologs for one human gene, the gene with the highest
535 log fold change in expression was used. Pathway analysis and GSEA (Subramanian et al.,
536 2005) was done using the fgsea R package (Sergushichev, 2016), using the MSigDB (Liberzon
537 et al., 2011) Gene Ontology (GO) (Ashburner et al., 2000; The Gene Ontology Consortium,
538 2019) biological processes and GO cellular component human genesets.

539 ***HOMER motif analysis***

540 Motif analysis was performed using HOMER (Heinz et al., 2010), using the function
541 findMotifs.pl. Motifs of lengths 8, 10, and 16 were queried within +/- 500 bp of the TSS of
542 differentially expressed genes. Target genes containing the motif of interest were found by
543 filtering the list of differentially expressed genes to contain only those with the desired motif.
544 JASPAR (Khan et al., 2018) was used to annotate motifs.

545 **Single-cell RNA-seq**

546 ***Sample preparation***

547 Adult tumor-bearing fish were dissected to obtain only the tumor and surrounding tissues (i.e
548 head and tail were removed). Tissue was minced with a fresh scalpel and incubated in 0.16
549 mg/mL liberase (Sigma-Aldrich #5401020001) in 0.9X PBS for 15 minutes at RT. Tissue was
550 then further dissociated by repeated pipetting with a wide-bore P1000, followed by incubation
551 for an additional 15 minutes at RT. After adding 500 uL FBS to stop the dissociation reaction,
552 samples were filtered through a 70 µm filter and centrifuged at 500*g* for 5 minutes at RT. The

553 resulting pellet was resuspended in DMEM supplemented with 2% FBS, and cells were sorted
554 at room temperature to remove debris and doublets, using a BD FACSAria III cell sorter (BD
555 Biosciences). Equal numbers of GFP+ (tumor) and GFP- (microenvironment) cells were
556 collected.

557 ***Cell encapsulation and library preparation***
558 Equal numbers of sorted GFP+ (tumor) and GFP- (microenvironment) cells were centrifuged at
559 300*g* for 5 minutes at RT, and resuspended in DMEM + 10% FBS. Droplet-based scRNA-seq
560 was performed using the Chromium Single Cell 3' Library & Gel Bead Kit v3 (10X Genomics)
561 and Chromium Single Cell 3' Chip G (10X Genomics). Approximately 10,000 cells from two fish
562 were split encapsulated in a single v3 reaction. GEM generation and library preparation were
563 performed according to manufacturer's instructions.

564 ***Sequencing***
565 10X scRNA-Seq libraries were pooled, denatured, and diluted to a concentration of 1.8 pM with
566 1% PhiX prior to paired-end sequencing on a NextSeq 500. Each library (corresponding to
567 approximately 5000 cells) were sequenced to a depth of 550M paired-end reads. Sequencing
568 parameters: Read1 28 cycles, index read 8 cycles, Read2 132 cycles.

569 ***Analysis***
570 Data was processed using R version 3.6.3 and Seurat version 3.1.4 (Stuart et al., 2019). Cells
571 with fewer than 200 unique genes or >20% mitochondrial reads were filtered out. Expression
572 data was normalized using SCTransform (Hafemeister and Satija, 2019). Datasets were
573 integrated using the Seurat SCTransform integration workflow, with 3000 integration anchors
574 and including all genes expressed in both datasets (15,154 genes). Principal component
575 analysis (Jolliffe, 1986), UMAP dimensionality reduction (McInnes et al., 2018), HOMER
576 analysis, GSEA, and pathway analysis were performed as described above. Cluster annotations
577 were performed using the Seurat function FindAllMarkers, in conjunction with marker genes
578 used in previous analyses (Baron et al., 2020).

579 **Immunofluorescence and imaging**
580 ***Staining of sections from adult fish***
581 Adult *casper* zebrafish with large pigmented tumors visible under the skin were euthanized on
582 ice and fixed in 4% paraformaldehyde in PBS for 72 hours at 4°C. Fish were then stored in 70%
583 ethanol before embedding in paraffin and sectioning by Histowiz, Inc. FFPE slides were
584 deparaffinized in xylene before several rounds of washing in 100-50% ethanol. Antigen retrieval
585 by heating slides to sub-boiling temperature for 10 minutes in 10 mM sodium citrate pH 6.2.
586 After cooling, slides were blocked in a solution of 5% donkey serum, 1% BSA, and 0.4%
587 Triton-X100 in PBS for 1 hour at room temperature, before overnight incubation with primary
588 antibodies in blocking buffer. Primary antibodies used were: goat anti-GFP (abcam #ab5450,
589 1:200), rabbit anti-ARL13B (Proteintech #17711-1-AP, 1:100), mouse anti-gamma tubulin
590 (Sigma-Aldrich #T6557, 1:100), and mouse anti-acetylated tubulin (Sigma-Aldrich #6793,
591 1:100). Following overnight incubation with primary antibodies, slides were washed in PBS

before incubation with secondary antibodies for 2 hours at room temperature. Secondary antibodies used were: donkey anti-goat IgG conjugated to Alexa 488 (Thermo Fisher Scientific #A11055, 1:250), goat anti-rabbit IgG conjugated to Alexa 555 (Cell Signaling Technology #4413S, 1:250), and goat anti-mouse IgG conjugated to Alexa 647 (Cell Signaling Technology #4410S, 1:250). Hoechst 33342 (Thermo Fisher Scientific #H3570) was added to the secondary antibody solution at 1:1000. Slides were mounted in ProLong Glass (Thermo Fisher Scientific #P36980) and cured overnight at room temperature.

Staining of cultured zebrafish melanoma cells

ZMEL1 cells expressing cytoplasmic GFP were plated on Lab-Tek II CC2 4-well Chamber Slides (Thermo Fisher Scientific #154917PK) at a concentration of $1x10^5$ cells/well. Cells were left to adhere overnight and then fixed in 4% paraformaldehyde in PBS for 15 mins at RT. Cells were then washed with PBS and permeabilized in 0.1% Triton-X100 in PBS for 30 mins at RT, before blocking in 10% goat serum (Thermo Fisher Scientific #50-062Z) for 1 hour at RT. Primary antibody incubation was done overnight at 4°C in 10% goat serum. Primary antibodies used were the same as above but at a concentration of 1:100. After primary antibody incubation, cells were washed with PBS and incubated with secondary antibodies for 2 hours at RT. Secondary antibodies used were the same as above. Cells were then incubated with Hoechst (1:1000 in 10% goat serum) for 15 mins at RT before mounting in ProLong Glass.

Confocal microscopy and image analysis

Slides were imaged on a Leica SP5 upright line-scanning confocal microscope using 10X and 63X (oil) objectives. 12-bit Z-stacks were acquired at 0.5-μm steps using 3X line averaging. Maximum intensity projections were created of the Z-stacks in ImageJ. Quantification of microscopy images was done using the MATLAB (Mathworks) Image Processing Toolbox. For quantification of cilia intensity in and outside tumor regions, the tumor region was segmented using a threshold of the image mean plus one standard deviation. The tumor segmentation was then applied as a mask to the corresponding image of acetylated tubulin (cilia) staining. Cilia images were background subtracted using the image mean as the background value, before calculating the mean pixel value of the acetylated tubulin intensity within and outside the tumor region. Cilia fluorescence within and outside the tumor region was compared using a paired t-test.

Re-analysis of Wouters et al. human melanoma scRNA-seq dataset

The counts matrix containing expression data for 10 human melanoma cell lines (4,323 total cells) at baseline without perturbations was obtained from GEO (GSE134432, sample GSM3946506). All analysis was done in R using Seurat v3.6.3 (Butler et al., 2018). Seurat objects for each of the 10 samples were created and then merged to create one dataset. Dimensionality reduction and clustering was performed with default parameters, using 15 principal components. Expression scores for interface, cilia, and ETS genes were calculated using the Seurat function AddModuleScore with default parameters. Interface marker genes were defined as the human orthologs of all genes with a log fold change > 1.5 in our zebrafish scRNA-seq interface cluster.

633 **Statistical analysis**
634 Statistical analysis and figure generation were performed in MATLAB (Mathworks, R2019a) and
635 R (R Foundation for Statistical Computing, 3.6.3). Image processing and analysis was
636 performed in MATLAB and ImageJ (NIH). Unless otherwise noted, *p*-values were calculated
637 using the Wilcoxon rank-sum test with Bonferroni's correction for multiple groups (R functions
638 wilcox.test and pairwise.wilcox.test). Pearson correlation coefficients and corresponding
639 *p*-values were calculated using the R function cor.test. *p*-value abbreviations are as follows: *, p
640 < 0.05, **, $p < 0.01$, ***, $p < 0.001$, n.s, not significant.

641 **Data availability**
642 The scRNA-seq and ST data reported in this manuscript have been deposited to the Gene
643 Expression Omnibus under accession number GSE159709.

644 **SUPPLEMENTAL INFORMATION TITLES AND LEGENDS**
645 **Figure S1. Spatial transcriptomics data statistics.**
646 **a-b.** Violin plots showing number of transcripts (UMIs) detected per ST spot (a) and number of
647 unique genes detected per ST spot (b). **c-d.** Histograms showing the distribution of the number
648 of UMIs and genes detected per ST spot. **e-f.** Numbers of UMIs and genes detected per spot,
649 plotted onto the tissue array.

650 **Figure S2. A unique interface cluster is found in each sample.**
651 **a.** Cluster assignments (left) and original sample identity (right) labelled in UMAP space for the
652 integrated ST dataset. **b.** Cluster assignments plotted onto the ST array for individual samples.
653 **c.** Cluster assignments labelled in UMAP space for individual samples.

654 **Figure S3. Spatial patterning of biological pathways in the tumor and microenvironment.**
655 **a-b.** Average, standardized expression of annotated genes for gene ontology (GO) terms
656 displaying spatially-coherent expression patterns in the tumor (a) and microenvironment (b)
657 regions in each sample. P-values represent the comparison between the distance between
658 spots expressing that GO term genes and a null-distribution of distances between random spots
659 (Wilcoxon's Rank Sum test, see Methods).

660 **Figure S4. Non-negative matrix factorization (NMF) analysis on ST spots, and enriched**
661 **GO terms for top genes in each NMF factor.**
662 Non-negative matrix factorization (NMF) was performed on the integrated expression matrix of
663 three ST datasets, with k=15 factors. Factor scores were then projected onto spots, and the top
664 150 scoring genes per factor were used for GO enrichment analysis.

665 **Figure S5. A zebrafish melanoma cell line is not ciliated.**
666 **a.** ZMEL1 cells stained with antibodies against the cilia markers ARL13B and acetylated tubulin.
667 Scale bars, 25 μm. **b.** Violin plot showing average normalized mRNA expression from bulk
668 RNA-seq for SYSCILIA genes expressed in ZMEL1 cells *in vitro* (n = 2 replicates) and upon

transplant into zebrafish (*n* = 4 replicates). Each point represents one SYSCILIA gene (*n* = 320 genes). Bulk RNA-seq data from Heilmann et al., 2015. n.s., not significant (*p* = 0.429, Wilcoxon rank-sum test).

Table S1. Top 100 genes upregulated in the scRNA-seq interface cluster.
Cilia genes are highlighted in pink.

Video S1. Spatial patterning of biological pathways in melanoma. 100 pathways are shown within the tumor region of sample A. The average expression of the genes annotated for the indicated pathway is shown. Blue: low expression; red: high expression.

Video S2. Spatial patterning of biological pathways in the melanoma macroenvironment. 100 pathways are shown within the interface and muscle regions of sample C. The average expression of the genes annotated for the indicated pathway is shown. Blue: low expression; red: high expression.

REFERENCES

Albrecht-Buehler, G. (1977). Phagokinetic tracks of 3T3 cells: parallels between the orientation of track segments and of cellular structures which contain actin or tubulin. Cell *12*, 333–339.

Anvarian, Z., Mykytyn, K., Mukhopadhyay, S., Pedersen, L.B., and Christensen, S.T. (2019). Cellular signalling by primary cilia in development, organ function and disease. Nat. Rev. Nephrol. *15*, 199–219.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene Ontology: tool for the unification of biology. Nat. Genet. *25*, 25–29.

Baron, M., Tagore, M., Hunter, M.V., Kim, I.S., Moncada, R., Yan, Y., Campbell, N.R., White, R.M., and Yanai, I. (2020). The Stress-Like Cancer Cell State Is a Consistent Component of Tumorigenesis. Cell Syst.

Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420.

Cancer Genome Atlas Network (2015). Genomic classification of cutaneous melanoma. Cell *161*, 1681–1696.

Caspary, T., Larkins, C.E., and Anderson, K.V. (2007). The graded response to Sonic Hedgehog depends on cilia architecture. Dev. Cell *12*, 767–778.

Ceol, C.J., Houvras, Y., Jane-Valbuena, J., Bilodeau, S., Orlando, D.A., Battisti, V., Fritsch, L., Lin, W.M., Hollmann, T.J., Ferré, F., et al. (2011). The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset. Nature *471*, 513–517.

Cevik, S., Hori, Y., Kaplan, O.I., Kida, K., Toivenon, T., Foley-Fisher, C., Cottell, D., Katada, T.,

Kontani, K., and Blacque, O.E. (2010). Joubert syndrome Arl13b functions at ciliary membranes and stabilizes protein transport in Caenorhabditis elegans. J. Cell Biol. *188*, 953–969.

Choudhury, A., Neumann, N.M., Raleigh, D.R., and Lang, U.E. (2020). Clinical implications of primary cilia in skin cancer. Dermatol Ther (Heidelb) *10*, 233–248.

Christensen, S.T., Pedersen, S.F., Satir, P., Veland, I.R., and Schneider, L. (2008). The primary cilium coordinates signaling pathways in cell cycle control and migration during development and tissue repair. Curr. Top. Dev. Biol. *85*, 261–301.

Dagogo-Jack, I., and Shaw, A.T. (2018). Tumour heterogeneity and resistance to cancer therapies. Nat. Rev. Clin. Oncol. *15*, 81–94.

van Dam, T.J., Wheway, G., Slaats, G.G., SYSCILIA Study Group, Huynen, M.A., and Giles, R.H. (2013). The SYSCILIA gold standard (SCGSv1) of known ciliary components and its applications within a systems biology consortium. Cilia *2*, 7.

Dittmer, J. (2011). Ets Transcription Factors. In Encyclopedia of Cancer, M. Schwab, ed. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 1339–1343.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat. Protoc. *4*, 1184–1191.

Fecher, L.A., Amaravadi, R.K., and Flaherty, K.T. (2008). The MAPK pathway in melanoma. Curr. Opin. Oncol. *20*, 183–189.

Fu, Y., Jung, A.W., Torne, R.V., Gonzalez, S., Vöhringer, H., Shmatko, A., Yates, L.R., Jimenez-Linan, M., Moore, L., and Gerstung, M. (2020). Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. Nat. Cancer.

González-Silva, L., Quevedo, L., and Varela, I. (2020). Tumor Functional Heterogeneity Unraveled by scRNA-seq Technologies. Trends Cancer *6*, 13–19.

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

Hassounah, N.B., Bunch, T.A., and McDermott, K.M. (2012). Molecular pathways: the role of primary cilia in cancer progression and therapeutics with a focus on Hedgehog signaling. Clin. Cancer Res. *18*, 2429–2435.

Heilmann, S., Ratnakumar, K., Langdon, E., Kansler, E., Kim, I., Campbell, N.R., Perry, E., McMahon, A., Kaufman, C., van Rooijen, E., et al. (2015). A quantitative system for studying metastasis using transparent zebrafish. Cancer Res. *75*, 4272–4282.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., and Mohr, S.E.

740 (2011). An integrative approach to ortholog prediction for disease-focused and other functional
741 studies. BMC Bioinformatics *12*, 357.

742 Iyengar, S., Houvras, Y., and Ceol, C.J. (2012). Screening for melanoma modifiers using a
743 zebrafish autochthonous tumor model. J. Vis. Exp. e50086.

744 Jolliffe, I.T. (1986). Principal component analysis and factor analysis. In Principal Component
745 Analysis, (New York, NY: Springer New York), pp. 115–128.

746 Kather, J.N., Heij, L.R., Grabsch, H.I., Loeffler, C., Echle, A., Muti, H.S., Krause, J., Niehues,
747 J.M., Sommer, K.A.J., Bankhead, P., et al. (2020). Pan-cancer image-based detection of
748 clinically actionable genetic alterations. Nat. Cancer.

749 Kaur, A., Webster, M.R., Marchbank, K., Behera, R., Ndoye, A., Kugel, C.H., Dang, V.M.,
750 Appleton, J., O'Connell, M.P., Cheng, P., et al. (2016). sFRP2 in the aged microenvironment
751 drives melanoma metastasis and therapy resistance. Nature *532*, 250–254.

752 Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J.A., van der Lee, R.,
753 Bessy, A., Chèneby, J., Kulkarni, S.R., Tan, G., et al. (2018). JASPAR 2018: update of the
754 open-access database of transcription factor binding profiles and its web framework. Nucleic
755 Acids Res. *46*, D1284.

756 Kim, I.S., Heilmann, S., Kansler, E.R., Zhang, Y., Zimmer, M., Ratnakumar, K., Bowman, R.L.,
757 Simon-Vermot, T., Fennell, M., Garippa, R., et al. (2017). Microenvironment-derived factors
758 driving metastatic plasticity in melanoma. Nat. Commun. *8*, 14343.

759 Kim, J., Dabiri, S., and Seeley, E.S. (2011). Primary cilium depletion typifies cutaneous
760 melanoma in situ and malignant melanoma. PLoS ONE *6*, e27410.

761 Lang, U.E., Love, N.R., Cheung, C., McCalmont, T.H., and Kim, J. (2020). Use of the ciliation
762 index to distinguish invasive melanoma from associated conventional melanocytic nevi. Am. J.
763 Dermatopathol. *42*, 11–15.

764 Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P.
765 (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740.

766 Li, R., Pei, H., and Watson, D.K. (2000). Regulation of Ets function by protein–protein
767 interactions. Oncogene *19*, 6514–6523.

768 Mao, P., Brown, A.J., Esaki, S., Lockwood, S., Poon, G.M.K., Smerdon, M.J., Roberts, S.A., and
769 Wyrick, J.J. (2018). ETS transcription factors induce a unique UV damage signature that drives
770 recurrent mutagenesis in melanoma. Nat. Commun. *9*, 2626.

771 Mavrothalassitis, G., and Ghysdael, J. (2000). Proteins of the ETS family with transcriptional
772 repressor activity. Oncogene *19*, 6524–6532.

773 McInnes, L., Healy, J., Saul, N., and Großberger, L. (2018). UMAP: uniform manifold
774 approximation and projection. JOSS *3*, 861.

775 Moncada, R., Barkley, D., Wagner, F., Chiodin, M., Devlin, J.C., Baron, M., Hajdu, C.H.,
776 Simeone, D.M., and Yanai, I. (2020). Integrating microarray-based spatial transcriptomics and

single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. Nat. Biotechnol. *38*, 333–342.

Passarelli, A., Mannavola, F., Stucci, L.S., Tucci, M., and Silvestris, F. (2017). Immune system and melanoma biology: a balance between immunosurveillance and immune escape. Oncotarget *8*, 106132–106142.

Patton, E.E., Widlund, H.R., Kutok, J.L., Kopani, K.R., Amatruda, J.F., Murphey, R.D., Berghmans, S., Mayhall, E.A., Traver, D., Fletcher, C.D.M., et al. (2005). BRAF mutations are sufficient to promote nevi formation and cooperate with p53 in the genesis of melanoma. Curr. Biol. *15*, 249–254.

Quail, D.F., and Joyce, J.A. (2013). Microenvironmental regulation of tumor progression and metastasis. Nat. Med. *19*, 1423–1437.

Rothhammer, T., Hahne, J.C., Florin, A., Poser, I., Soncin, F., Wernert, N., and Bosserhoff, A.K. (2004). The Ets-1 transcription factor is involved in the development and invasion of malignant melanoma. Cell. Mol. Life Sci. *61*, 118–128.

Schneider, L., Cammer, M., Lehman, J., Nielsen, S.K., Guerra, C.F., Veland, I.R., Stock, C., Hoffmann, E.K., Yoder, B.K., Schwab, A., et al. (2010). Directional cell migration and chemotaxis in wound healing response to PDGF-AA are coordinated by the primary cilium in fibroblasts. Cell. Physiol. Biochem. *25*, 279–292.

Sergushichev, A. (2016). An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. BioRxiv.

Sizemore, G.M., Pitarresi, J.R., Balakrishnan, S., and Ostrowski, M.C. (2017). The ETS family of oncogenic transcription factors in solid tumours. Nat. Rev. Cancer *17*, 337–351.

Snedecor, E.R., Sung, C.C., Moncayo, A., Rothstein, B.E., Mockler, D.C., Tonnesen, M.G., Jones, E.C., Fujita, M., Clark, R.A., Shroyer, K.R., et al. (2015). Loss of primary cilia in melanoma cells is likely independent of proliferation and cell cycle progression. J. Invest. Dermatol. *135*, 1456–1458.

Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M., et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. Science *353*, 78–82.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell *177*, 1888-1902.e21.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA *102*, 15545–15550.

The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. *47*, D330–D338.

Tirosh, I., Izar, B., Prakadan, S.M., Wadsworth, M.H., Treacy, D., Trombetta, J.J., Rotem, A., Rodman, C., Lian, C., Murphy, G., et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science *352*, 189–196.

Ungefroren, H., Sebens, S., Seidl, D., Lehnert, H., and Hass, R. (2011). Interaction of tumor cells with the microenvironment. Cell Commun. Signal. *9*, 18.

Wasylyk, B., Hagman, J., and Gutierrez-Hartmann, A. (1998). Ets transcription factors: nuclear effectors of the Ras-MAP-kinase signaling pathway. Trends Biochem. Sci. *23*, 213–216.

White, R.M., Sessa, A., Burke, C., Bowman, T., LeBlanc, J., Ceol, C., Bourque, C., Dovey, M., Goessling, W., Burns, C.E., et al. (2008). Transparent adult zebrafish as a tool for in vivo transplantation analysis. Cell Stem Cell *2*, 183–189.

Winnepenninckx, V., Lazar, V., Michiels, S., Dessen, P., Stas, M., Alonso, S.R., Avril, M.-F., Ortiz Romero, P.L., Robert, T., Balacescu, O., et al. (2006). Gene expression profiling of primary cutaneous melanoma and clinical outcome. J Natl Cancer Inst *98*, 472–482.

Wouters, J., Kalender-Atak, Z., Minnoye, L., Spanier, K.I., De Waegeneer, M., Bravo González-Blas, C., Mauduit, D., Davie, K., Hulselmans, G., Najem, A., et al. (2020). Robust gene expression programs underlie recurrent cell states and phenotype switching in melanoma. Nat. Cell Biol. *22*, 986–998.

Zhang, M., Di Martino, J.S., Bowman, R.L., Campbell, N.R., Baksh, S.C., Simon-Vermot, T., Kim, I.S., Haldeman, P., Mondal, C., Yong-Gonzales, V., et al. (2018). Adipocyte-Derived Lipids Mediate Melanoma Progression via FATP Proteins. Cancer Discov. *8*, 1006–1025.

Zingg, D., Debbache, J., Peña-Hernández, R., Antunes, A.T., Schaefer, S.M., Cheng, P.F., Zimmerli, D., Haeusel, J., Calçada, R.R., Tuncer, E., et al. (2018). EZH2-Mediated Primary Cilium Deconstruction Drives Metastatic Melanoma Formation. Cancer Cell *34*, 69-84.e14.

Figure 1. Spatial transcriptomics reveals the transcriptional architecture of the melanoma microenvironment.

a. Schematic showing the spatial transcriptomics (ST) experiment workflow.

b. Images of zebrafish with *BRAF*^{V600E}-driven melanomas used for ST. The region where the fish were sectioned is highlighted.

c. H&E staining of cryosections used for ST.

d. ST spots colored by clustering assignments of the integrated dataset (see Methods).

e. UMAP embedding of ST spots from all three samples colored by cluster assignments of the integrated dataset (see Methods).

f. The expression of select marker genes (*BRAF*^{V600E}, tumor; *pvalb4*, muscle; *kcn6a*, heart; *mbpa*, nervous system) from the ST data projected over tissue space (left), with images of the corresponding histology from the indicated region of the ST array (right).

g-h. Average, standardized expression of annotated genes for gene ontology (GO) terms displaying spatially-coherent expression patterns in the tumor (g) and microenvironment (h) regions in each sample. P-values represent the comparison between the distance between spots expressing that GO term genes and a null-distribution of distances between random spots (Wilcoxon's Rank Sum test, see Methods).

Figure 1

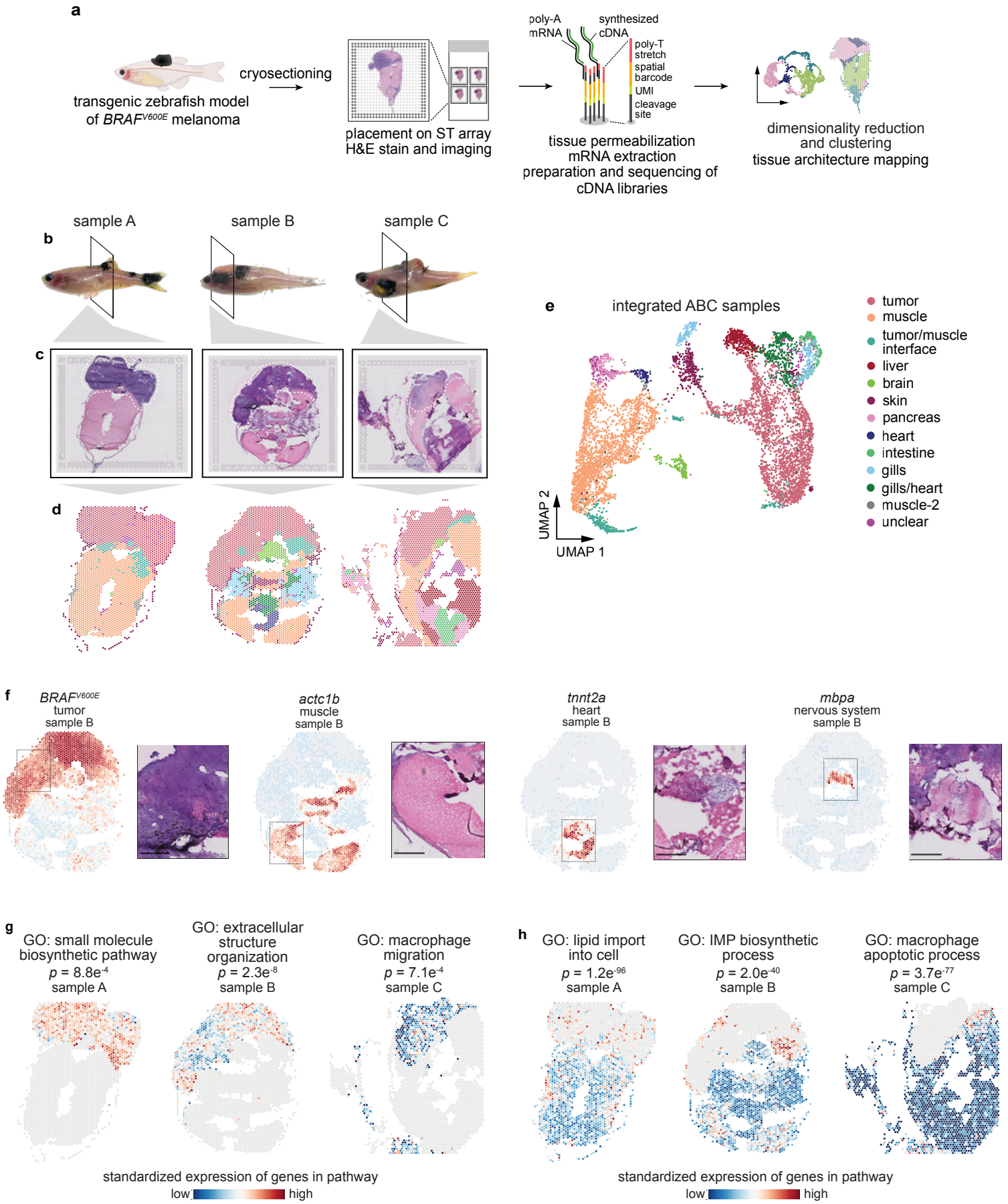

a
transgenic zebrafish model of BRAFV600E melanoma
cryosectioning
placement on ST array
H&E stain and imaging
poly-A mRNA
synthesized cDNA
poly-T stretch
spatial barcode
UMI
cleavage site
tissue permeabilization
mRNA extraction
preparation and sequencing of cDNA libraries
dimensionality reduction and clustering
tissue architecture mapping
sample A
sample B
sample C
b
c
d
e
integrated ABC samples
UMAP 2
UMAP 1
tumor
muscle
tumor/muscle interface
liver
brain
skin
pancreas
heart
intestine
gills
gills/heart
muscle-2
unclear
f
BRAFV600E tumor sample B
actc1b muscle sample B
tnnt2a heart sample B
mbpa nervous system sample B
g
GO: small molecule biosynthetic pathway
p = 8.8e-4
sample A
GO: extracellular structure organization
p = 2.3e-8
sample B
GO: macrophage migration
p = 7.1e-4
sample C
h
GO: lipid import into cell
p = 1.2e-96
sample A
GO: IMP biosynthetic process
p = 2.0e-40
sample B
GO: macrophage apoptotic process
p = 3.7e-77
sample C
standardized expression of genes in pathway
low
high
standardized expression of genes in pathway
low
high

Figure 2. The tumor-microenvironment interface is transcriptionally distinct from the surrounding microenvironment.

a. Interface and muscle-annotated cluster spots projected onto tissue image. Insets show the tissue underlying the interface spots (1) and muscle spots (2).

b. Correlation matrix between average expression profile of ST clusters across all three datasets. Clusters are ordered by hierarchical clustering of the Pearson's correlation coefficients (see Methods) and bubble sizes correspond to P value (-log10) of correlation, with P-values < 10-3 omitted. Clustering of tumor and interface together is highlighted in the dendrogram (red).

c. Volcano plot of differentially expressed genes between the interface cluster versus the muscle and tumor clusters. P-values were obtained from the Wilcoxon's rank sum test.

d. Non-negative matrix factorization (NMF) of the microenvironment spots (muscle and interface clusters). Shown are the standardized factor scores for interface-specific NMF factor 15, projected onto microenvironment spots.

e. Enriched GO terms for the top 150 scoring genes in NMF factor 15.

Figure 2

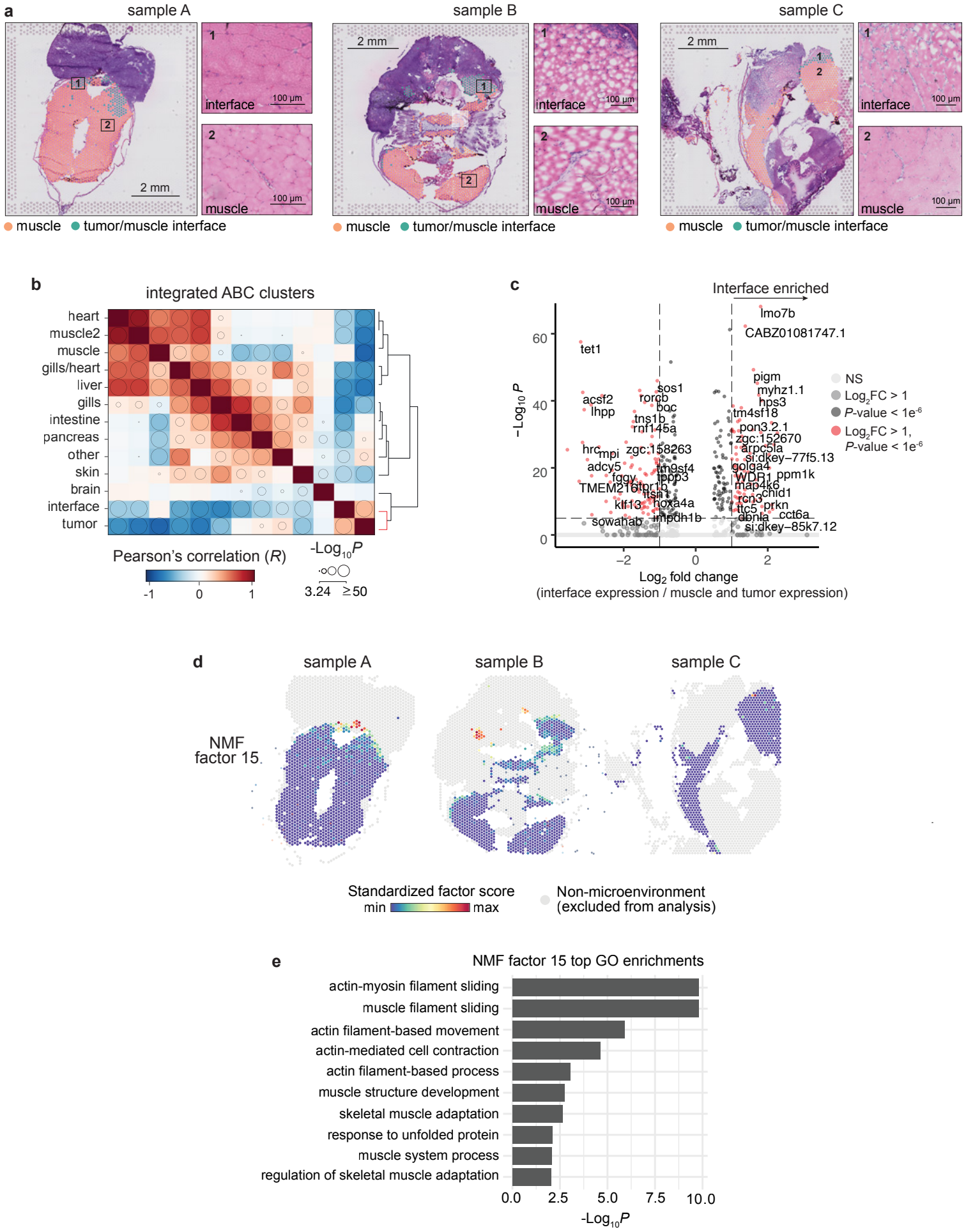

a
sample A
sample B
sample C
2 mm
interface
muscle
100 µm
muscle
tumor/muscle interface
b
integrated ABC clusters
heart
muscle2
muscle
gills/heart
liver
gills
intestine
pancreas
other
skin
brain
interface
tumor
Pearson's correlation (R)
-1
0
1
-Log10P
3.24
≥50
c
Interface enriched
lmo7b
CABZ01081747.1
tet1
pigm
myhz1.1
hps3
tm4sf18
pon3.2.1
zgc:152670
arpc5la
si:dkey-77f5.13
golga4
WDR1
ppm1k
map4k6
chid1
rcn3
prkn
ttc5
dbnla
cct6a
si:dkey-85k7.12
sos1
rorcb
boc
tns1b
rnf145a
acsf2
lhpp
hrc
mpi
zgc:158263
adcy5
tm9sf4
fggy
ippp3
TMEM216
itpr1b
itsn1
klf13
hoxa4a
sowahab
impdh1b
NS
Log2FC > 1
P-value < 1e-6
Log2FC > 1, P-value < 1e-6
−Log10P
Log2 fold change
(interface expression / muscle and tumor expression)
d
sample A
sample B
sample C
NMF factor 15
Standardized factor score
min
max
Non-microenvironment (excluded from analysis)
e
NMF factor 15 top GO enrichments
actin-myosin filament sliding
muscle filament sliding
actin filament-based movement
actin-mediated cell contraction
actin filament-based process
muscle structure development
skeletal muscle adaptation
response to unfolded protein
muscle system process
regulation of skeletal muscle adaptation
0.0
2.5
5.0
7.5
10.0
-Log10P

Figure 3. The tumor-microenvironment interface is composed of specialized tumor and muscle cells.
a. Schematic showing scRNA-seq experiment workflow.
b. UMAP dimensionality reduction plot for 2,889 cells sequenced as in (a). Cluster/cell type assignments are labelled and colored.
c. Expression score per cell (scRNA-seq) for average expression of interface marker genes from the ST interface cluster.
d. Heatmap showing expression of the top 50 genes upregulated in the tumor cell cluster (top, orange) and interface cell cluster (bottom, yellow). Selected genes are labelled.
e. Principal component analysis of cells in the interface cluster, scored for expression of the tumor marker *BRAF*^{V600E}, the muscle marker *ckba*, and the centromere gene *stra13*. Scores for principal components 1 and 2 are plotted. Cells are labelled by standardized expression of the indicated genes.
f. Dot plot showing expression of tumor and muscle markers. The size of each dot corresponds to the percentage of cells in that cluster expressing the indicated gene, and the colour of each dot indicates the expression level.

Figure 3

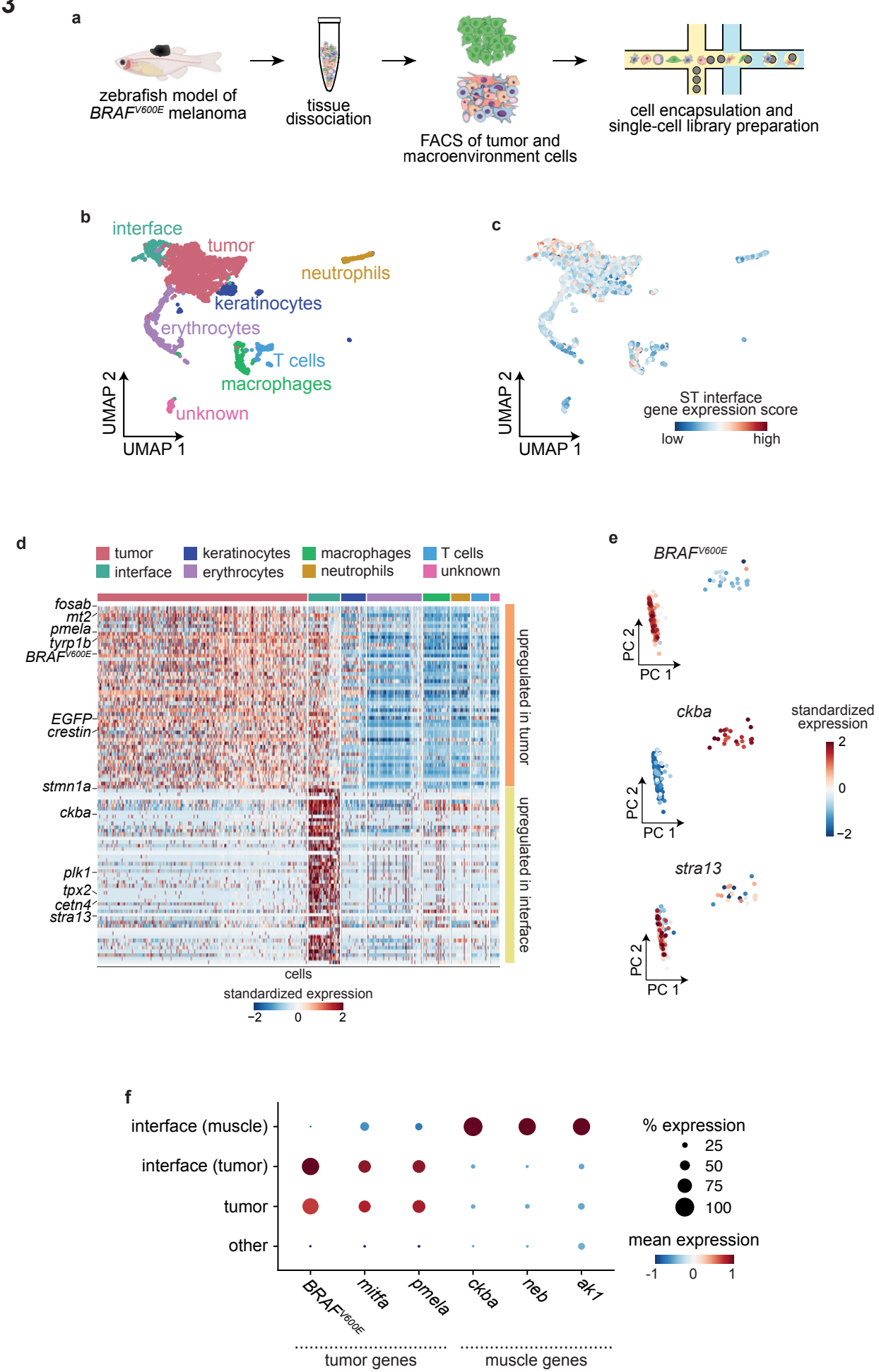

a
zebrafish model of
BRAFV600E melanoma
tissue
dissociation
FACS of tumor and
macroenvironment cells
cell encapsulation and
single-cell library preparation
b
interface
tumor
neutrophils
keratinocytes
erythrocytes
T cells
macrophages
unknown
UMAP 2
UMAP 1
c
UMAP 2
UMAP 1
ST interface
gene expression score
low
high
d
tumor
interface
keratinocytes
erythrocytes
macrophages
neutrophils
T cells
unknown
fosab
mt2
pmela
tyrp1b
BRAFV600E
EGFP
crestin
stmn1a
ckba
plk1
tpx2
cetn4
stra13
upregulated in tumor
upregulated in interface
cells
standardized expression
−2
0
2
e
BRAFV600E
ckba
stra13
PC 2
PC 1
standardized
expression
2
0
−2
f
interface (muscle)
interface (tumor)
tumor
other
BRAFV600E
mitfa
pmela
ckba
neb
ak1
tumor genes
muscle genes
% expression
25
50
75
100
mean expression
-1
0
1

Figure 4. Cilia genes and proteins are enriched at the tumor-microenvironment interface.
a. Bar graph showing the top 10 GO biological process terms enriched in differentially expressed genes in the scRNA-seq interface cluster. GO terms are colored by p-value.
b. Waterfall plot showing the top and bottom 250 GO cellular component terms by NES in the scRNA-seq interface cluster, with cilia GO terms labelled in red.
c. GSEA barcode plot showing enrichment of genes in the GO: CILIUM pathway in the scRNA-seq interface cluster. NES and FDR are indicated.
d. Bar graph comparing normalized enrichment scores for cilia-related GO terms in the scRNA-seq (purple) and ST (green) interface clusters.
e. Violin plots showing expression of fish SYSCILIA genes in the scRNA-seq interface cluster.
f. Violin plot showing normalized cilia fluorescence intensity within and outside the tumor region. *, $p < 0.05$ ($n = 6$ fish).
g. A section through an adult zebrafish with an invasive melanoma immunostained for GFP (tumor cells), acetylated tubulin (cilia) and Hoescht (nuclei). Arrows denote cilia in the microenvironment. Scale bar, 200 μm.
h-i. Sections from adult zebrafish with invasive melanomas stained with antibodies against the cilia markers acetylated tubulin and ARL13B, and the basal body marker γ-tubulin, imaged at 10X (h) and 63X (i) magnification. Scale bars, 200 μm (10X) and 50 μm (63X).

Figure 4

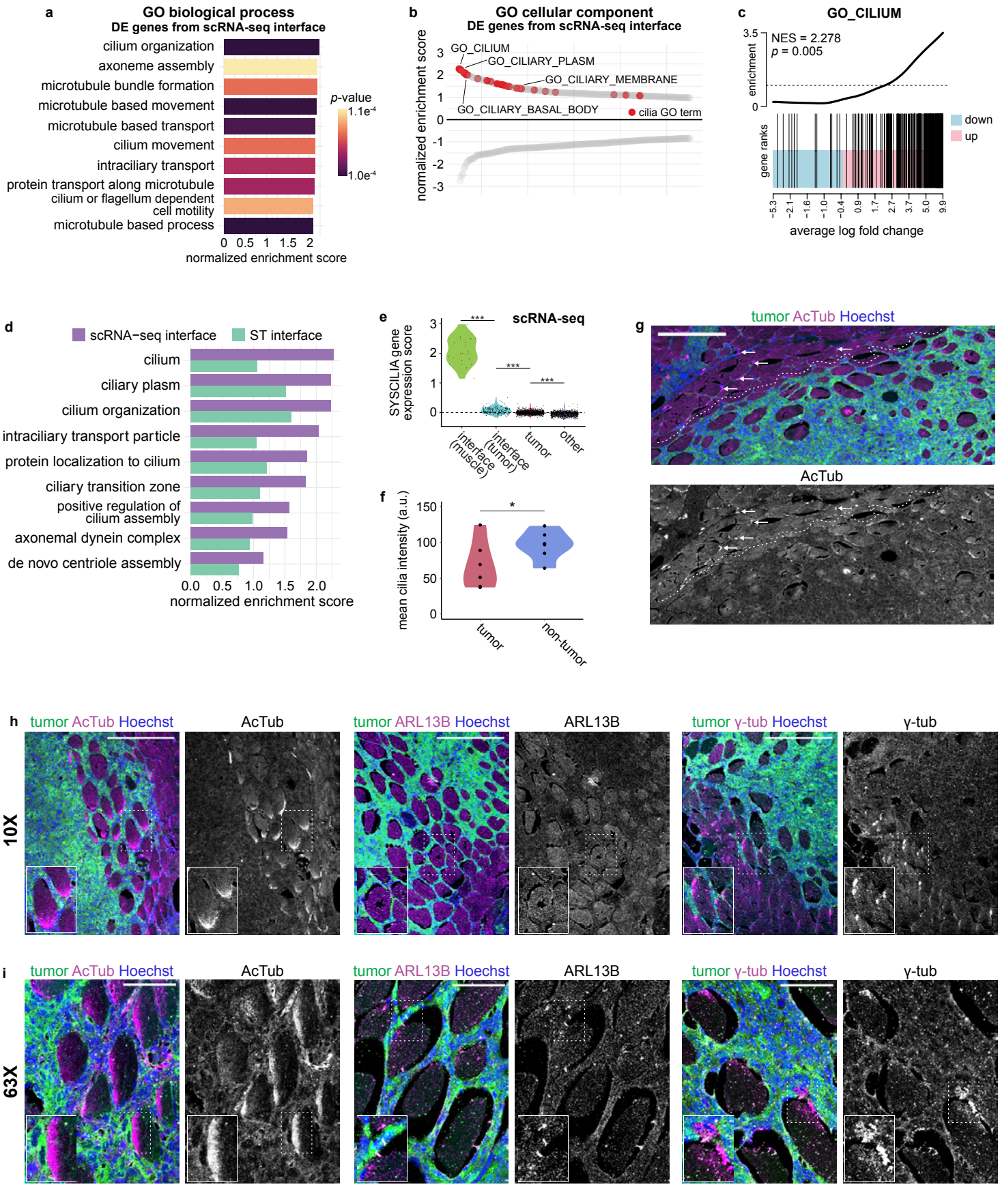

a
GO biological process
DE genes from scRNA-seq interface
cilium organization
axoneme assembly
microtubule bundle formation
microtubule based movement
microtubule based transport
cilium movement
intraciliary transport
protein transport along microtubule
cilium or flagellum dependent cell motility
microtubule based process
p-value
1.1e-4
1.0e-4
0 0.5 1 1.5 2
normalized enrichment score
b
GO cellular component
DE genes from scRNA-seq interface
normalized enrichment score
GO_CILIUM
GO_CILIARY_PLASM
GO_CILIARY_MEMBRANE
GO_CILIARY_BASAL_BODY
cilia GO term
c
GO_CILIUM
NES = 2.278
p = 0.005
enrichment
gene ranks
down
up
average log fold change
d
scRNA−seq interface
ST interface
cilium
ciliary plasm
cilium organization
intraciliary transport particle
protein localization to cilium
ciliary transition zone
positive regulation of cilium assembly
axonemal dynein complex
de novo centriole assembly
0.0 0.5 1.0 1.5 2.0
normalized enrichment score
e
scRNA-seq
SYSCILIA gene expression score
interface (muscle)
interface (tumor)
tumor
other
f
mean cilia intensity (a.u.)
tumor
non-tumor
g
tumor AcTub Hoechst
AcTub
h
tumor AcTub Hoechst
AcTub
tumor ARL13B Hoechst
ARL13B
tumor γ-tub Hoechst
γ-tub
10X
i
tumor AcTub Hoechst
AcTub
tumor ARL13B Hoechst
ARL13B
tumor γ-tub Hoechst
γ-tub
63X

Figure 5. ETS transcription factors regulate cilia gene expression at the interface in zebrafish and human melanoma.

a. Results from HOMER motif analysis of differentially expressed genes in the ST and scRNA-seq interface clusters.

b-c. Violin plots showing expression of zebrafish ETS-family genes in the tumor and interface clusters from the scRNA-seq (b) and ST (c) datasets.

d-e. Top 10 enriched GO cellular component terms by normalized expression score (NES) from ETS-target genes in the ST interface cluster. Cilia-related GO terms are labelled in red.

f. Scatter plot comparing ETS gene expression scores per cell and ETS-target gene expression scores per cell in the scRNA-seq interface cluster. Pearson's correlation coefficient (R) is indicated.

g-h. UMAP projection of human melanoma scRNA-seq data from Wouters et al. (2020). Cell lines (g) and interface marker gene expression score (h) are indicated. Scoring of interface marker genes was done with the human orthologs of the zebrafish genes upregulated by >1.5-fold in our zebrafish scRNA-seq interface cluster.

i-j. Expression of human cilia (i) and ETS (j) genes in tumor and interface cells in human melanoma scRNA-seq.

b-c, f, i-j. The expression score represents the standardized mean expression of all indicated genes per each cell or ST array spot. **, $p < 0.01$, ***, $p < 0.001$ (Wilcoxon rank-sum test).

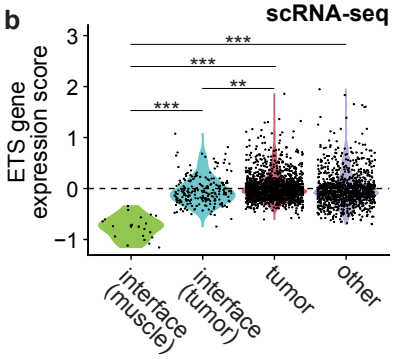
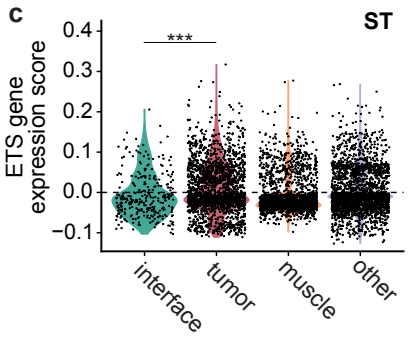
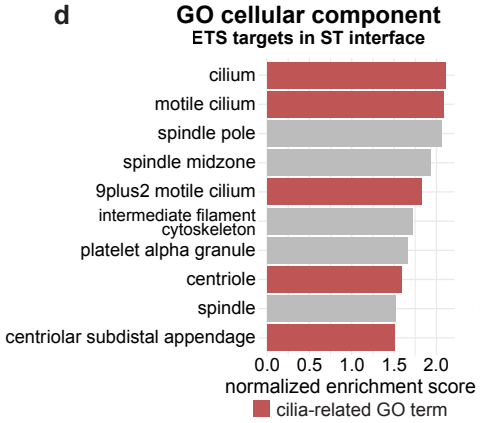
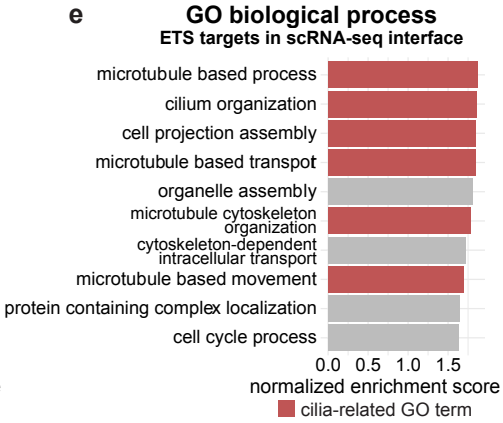
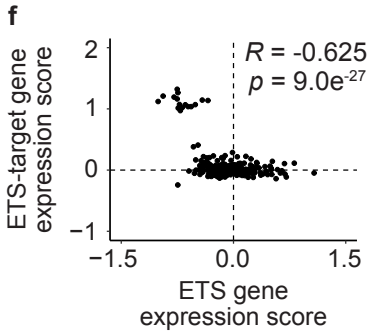
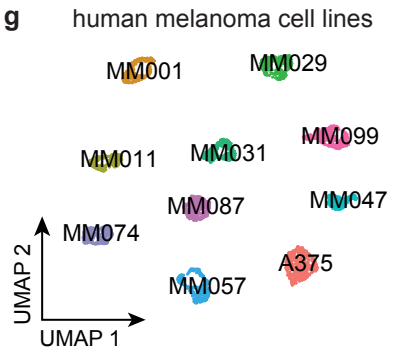
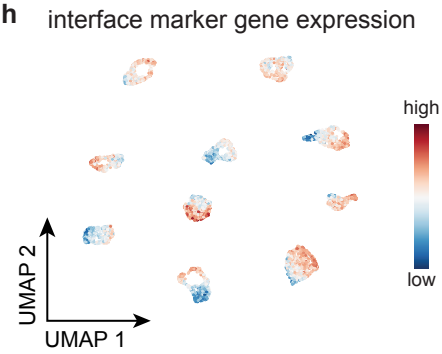
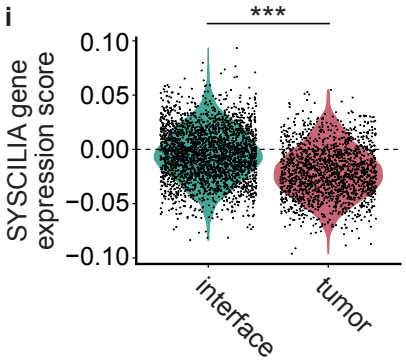
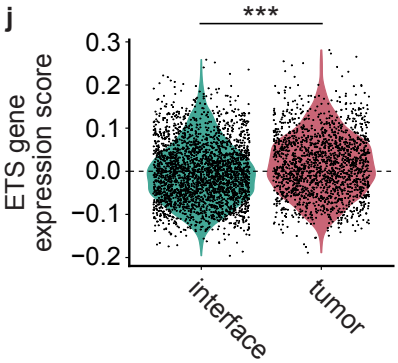
Figure 5

a **HOMER: interface vs. muscle** **ST**

motif	best match	*p*-value	rank
	ETS	$1e^{-22}$	1

HOMER: interface vs. all spots **ST**

motif	best match	*p*-value	rank
	ELF5 (ETS)	$1e^{-15}$	1

HOMER: interface vs. all other cells **scRNA-seq**

motif	best match	*p*-value	rank
	NKX2-3,..., ETV6 (ETS)	$1e^{-13}$	2