RESEARCH ARTICLE

The Spatial Landscape of Progression and Immunoediting in Primary Melanoma at Single-Cell Resolution

Ajit J. Nirmal1,2,3, Zoltan Maliga1,2, Tuulia Vallius1,2, Brian Quattrochi4, Alyce A. Chen1,2, Connor A. Jacobson1,2, Roxanne J. Pelletier1,2, Clarence Yap1,2, Raquel Arias-Camison1,2,4, Yu-An Chen1,2, Christine G. Lian4, George F. Murphy4, Sandro Santagata1,2,4, and Peter K. Sorger1,2,5
ABSTRACT

Cutaneous melanoma is a highly immunogenic malignancy that is surgically curable at early stages but life-threatening when metastatic. Here we integrate high-plex imaging, 3D high-resolution microscopy, and spatially resolved microregion transcriptomics to study immune evasion and immunoediting in primary melanoma. We find that recurrent cellular neighborhoods involving tumor, immune, and stromal cells change significantly along a progression axis involving precursor states, melanoma in situ, and invasive tumor. Hallmarks of immunosuppression are already detectable in precursor regions. When tumors become locally invasive, a consolidated and spatially restricted suppressive environment forms along the tumor–stromal boundary. This environment is established by cytokine gradients that promote expression of MHC-II and IDO1, and by PD-1–PDL1-mediated cell contacts involving macrophages, dendritic cells, and T cells. A few millimeters away, cytotoxic T cells synapse with melanoma cells in fields of tumor regression. Thus, invasion and immunoediting can coexist within a few millimeters of each other in a single specimen.

SIGNIFICANCE: The reorganization of the tumor ecosystem in primary melanoma is an excellent setting in which to study immunoediting and immune evasion. Guided by classic histopathology, spatial profiling of proteins and mRNA reveals recurrent morphologic and molecular features of tumor evolution that involve localized paracrine cytokine signaling and direct cell–cell contact.

INTRODUCTION

Tumorigenesis commonly involves a progressive failure of immune cells, particularly T cells, to detect cancer cells as they accumulate mutations promoting growth, invasion, and metastasis (1). The competition between editing by immune cells and escape by cancer cells generates a complex ecosystem whose molecular features and physical organization determine disease outcomes and responsiveness to therapy (2, 3). In the case of primary cutaneous melanoma, DNA sequencing has identified recurrent mutations in drivers such as \(\text{BRAF}\), \(\text{NRAS}\), \(\text{PTEN}\), and \(\text{TP53}\) (4–6), and dissociative single-cell RNA sequencing (scRNA-seq) has revealed progression-associated changes in immune cell states (7). However, oncogenic transformation and immune escape remain only partly understood due in part to a high mutational burden in morphologically normal skin, estimated in Caucasians to be >100 driver mutations per cm\(^2\) by late middle age (8). Although treatment of metastatic melanoma has benefited from modern targeted therapies guided by genetic biomarkers (BRAF and MEK inhibitors) and by immune-checkpoint inhibitors, primary melanoma is treated surgically. It is diagnosed and staged using classic methods such as histopathologic assessment of hematoxylin and eosin (H&E)–stained formaldehyde-fixed, paraffin-embedded (FFPE) skin biopsies, complemented in some cases by IHC (9).

Normal skin is characterized by evenly spaced melanocytes, which are neural crest–derived melanin-producing cells (10) located between cuboidal basal keratinocytes on the apical face of the dermal–epidermal junction (11). Fields of melanocytic atypia, the earliest signs of oncogenic transformation, involve increases in melanocyte number and density, enlargement, and irregularity of melanocyte nuclei, movement of melanocytes away from the dermal–epidermal junction (12), and loss of 5-hydroxymethylcytosine (5hmC) epigenetic marks (5, 13). These precursor fields can develop into melanoma in situ (MIS), a proliferative and confluence of malignant melanocytes within the epidermis but without invasion into the underlying dermis. MIS can spread within the epidermis and focally invade the superficial dermis without expansile growth, giving rise to radial growth phase melanoma, which has an excellent prognosis upon complete excision. However, invasive growth into the dermis is both expansile and highly mitotic, giving rise to vertical growth phase melanoma with a high potential for metastasis (15). Vertical growth phase melanomas can be endophytic or exophytic, corresponding to vertical growth down into the dermis or upward above the skin, at times resulting in polypoid lesions that erupt from the surrounding skin (16).

The study of recurrent mutations found in cutaneous melanoma has yielded models of sequential tumor evolution starting with the formation of dysplastic nevi (4). However, while the removal of dysplastic nevi with higher grades of atypia is standard clinical practice (17), it is now thought that the majority of primary cutaneous melanomas are not derived from nevi, but rather arise de novo from fields of melanocytic atypia, particularly in sun-damaged skin (18, 19). The key features of these precursor fields and the sequence of genetic events and
immunosuppressive features that promote their progression to invasive melanoma remain poorly understood, as do the extent and impact of interpatient and patient-to-patient variability. From a prognostic perspective, the depth of tumor invasion into the dermis (Breslow thickness) is a particularly important parameter (20) and is used in conjunction with the standard tumor–node–metastasis (TNM) system used for melanoma staging (21). The number and locations of tumor-infiltrating lymphocytes (TIL) also have prognostic value (22). Finally, the Clark scoring system recognizes three distinct patterns for TILs: absent, nonbrisk, and brisk (23). Absent describes both the absence of TILs and their failure to infiltrate tumor, nonbrisk describes the restriction of TILs to scattered foci in the vicinity of the tumor, and brisk describes infiltration throughout vertical growth phase tumors or widely distributed along the invasive tumor front (24). In general, the greater the number of infiltrating TILs, the brisker the response and the more favorable the prognosis (25, 26). In some tumors, regions of inflammatory regression are also observed. In these regions, T cells are observed to eradicate malignant melanocytes, leading to fields of fibrosis, vascular proliferation, and pigment incontinence, which are indicative of terminal regression (27). Inflammatory regression represents an example of successful and ongoing immunoeediting but is currently incidental to diagnosis and of uncertain prognostic significance (28).

The great majority of studies on immune surveillance in primary and metastatic melanoma have involved either histologic analysis of H&E or IHC images, which are restricted to one to three markers per section, or sequencing of genomic mutations or mRNA profiling. However, several recent studies have demonstrated the potential for multiplexed imaging to provide greater insight into the spatially restricted tumor and immune programs in melanomas at different stages (29, 30).

Here, we focus on the molecular and morphologic analysis of histologic features commonly found in primary melanoma. We focus on features used for diagnosis and treatment decisions in specimens containing multiple distinct stages of disease. These include precursor fields, MIS, radial growth phase melanoma, and/or invasive vertical growth phase melanoma as well as regions of inflammatory regression. Specimens were acquired from the Brigham and Women’s Hospital dermatopathology tissue bank and, like virtually all primary melanomas, were available only in fixed form (FFPE) as a diagnostic necessity; only a few were subjected to or consented for DNA sequencing. The spatial organization of the tumor microenvironment (TME) was analyzed using 20- to 30-plex fluorescence microscopy (CyCIF) and either conventional wide-field microscopy or 3D optical sectioning followed by deconvolution (31). We also performed transcriptional profiling of selected microregions using two different methods for microregion transcriptomics [microregion sequencing (mrSEQ); GeoMx and PicSeq; refs. 32, 33]. The resulting molecular and morphologic data were then correlated with local histopathology as determined from H&E images by board-certified dermatopathologists. To preserve the spatial relationships of different histologies and to provide sufficient statistical power (34), CyCIF and H&E imaging were performed on whole slides, not tissue microarrays (TMA) or small fields of view (FOV; ref. 35).

Using differential expression (DE) analysis and unsupervised clustering of mrSEQ data and spatial statistics on CyCIF data, we identified molecular programs associated with histopathologic progression. In many cases, immunoeediting by activated T cells was observed within a few millimeters of near-complete immune exclusion from invasive melanoma. Immunosuppressive niches were highly localized—in some cases only a few cells thick—and high-resolution imaging showed that they contained PD-L1 expressing myeloid cells in direct contact with PD1-expressing T cells.

RESULTS

Multimodal Profiling of Spatially Distinct Regions within Cutaneous Melanoma

A total of 70 tissue regions [histologic regions of interest (ROI)] with precancer or cancer histologies were identified in 11 FFPE specimens of primary cutaneous melanoma, one locoregional metastasis, and one distant skin metastasis (specimens MEL1 to MEL13; Supplementary Tables S1 and S2; histologic features and annotations are described in Supplementary Table S3). Analysis of H&E-stained specimens by board-certified dermatopathologists confirmed the presence of one to five histologic ROIs (average 2.4 per specimen) corresponding to precursor fields, MIS, invasive melanoma (IM), exophytic melanoma (EM), and inflammatory regression (IR) ~5 to 20 mm apart from one another (summarized in Supplementary Fig. S1A). Serial FFPE sections (5 μm thick) were subjected to whole-slide, subcellular-resolution, 20- to 30-plex CyCIF imaging with different combinations of antibodies to generate complementary sets of image data (Fig. 1A–C; Supplementary Fig. S1A; antibody panels described in Supplementary Tables S3 and S4). Antibodies included pan-cytokeratin (pan-CK) to stain keratinocytes in the epidermis, SOX10 and MITF to stain normal and atypical melanocytes and tumor cells (Supplementary Fig. S1B), smooth muscle actin (αSMA) to stain stromal cells, and CD31 to stain endothelial cells lining vessels. Immune cells were stained with lineage-specific cell-surface proteins and functional markers (e.g., PD1), as described in Supplementary Fig. S1C and Supplementary Tables S5 and S6. Image analysis and data processing were performed using algorithms integrated into the open-source MCMICRO pipeline (36); staining intensities for lineage markers such as CD4, CD8 (CD8A), and CD163 were then binarized to distinguish among 13 immune cell types (Fig. 1D–F).

More extensive molecular analysis was performed of specimen MEL1, which had the greatest number of distinct histologies (and spanned three tissue blocks MEL1-1, MEL1-2, and MEL1-3). MEL1 was an NFI-mutant, BRAFwt tumor, which is one of four recurrent cutaneous subtypes identifiable in The Cancer Genome Atlas (TCGA) data (37). It was a large primary tumor involving both inward-projecting vertical growth phase (nodular melanoma) and outward-growing EM. A region of MIS was coextensive with regions of inflammatory and terminal regression in which immune editing had reduced or eliminated tumor cells; these regions contained dense infiltrates of CTLs, the majority of which were PD1+ (and thus activated) as well as regulatory T cells (Treg). IM was located ~10 mm away from the MIS, and the invasive boundary (IB) of the nodular component had reached a depth of 4 to 5 mm and was surrounded by a domain of immune-rich stroma that was scored as a brisk TIL (bTIL) response. The patient from whom MEL1...
was obtained developed locoregional recurrence and distant metastases but was alive at the time of the last follow-up. MEL1 was characterized with a total of 80 different antibodies on five serial sections subjected to microregion transcriptomics (GeoMx and PickSeq [mrSEQ]). Abbreviations for annotated histologies are shown below with color-coding used in subsequent figure panels. A 3D-plex CyCIF image of a section of specimen MEL1-1 showing selected markers for epidermis (panCK, cyan) and tumor cells (SOX10, red), highlighting annotated histologies and microregions (mROIs) that were subjected to mrSEQ (white “+” symbols). This specimen was likely torn during slide processing, and, thus, spatial arrangements in the region marked with a blue dashed boundary are not considered reliable. Other mrSEQ sites are shown in Supplementary Fig. S2A. Multimodal profiling of cutaneous melanoma. A, Conceptual framework of sample processing for cyclic immunofluorescence (CyCIF), high-resolution CyCIF, and microregion transcriptomics [GeoMx and PickSeq (mrSEQ)]. B, A 3D-plex CyCIF image of a section of specimen MEL1-1 showing selected markers for epidermis (panCK, cyan) and tumor cells (SOX10, red), highlighting annotated histologies and microregions (mROIs) that were subjected to mrSEQ (white “+” symbols). This specimen was likely torn during slide processing, and, thus, spatial arrangements in the region marked with a blue dashed boundary are not considered reliable. Other mrSEQ sites are shown in Supplementary Fig. S2A. 

The ability of immune cells to make functional contacts with one another and with tumor cells is a fundamental feature of cancer immunoediting commonly quantified using spatial statistics (proximity analysis; ref. 38). In images collected at standard resolution (∼450 nm laterally), it is not possible to visualize the distinctive morphologies of immune synapses or PD1 binding to PD1 (39). We, therefore, used 3D high-resolution 3D deconvolution microscopy was performed. The region highlighted with orange is magnified in G. (continued on next page)
point of contact between the tumor cells and the Treg also revealed a contact (Fig. 1; Supplementary Fig. S1I) that may be associated with the programming of tolerogenic activity. Comprehensive characterization and quantification of cell–cell contacts detected by high-resolution tissue imaging await the development of better image recognition tools, but our data provide clear and hitherto unavailable evidence that immune and tumor cells in close proximity to one another have structures characteristic of functional cell-to-cell contacts. Proximity analysis likely overestimates the frequency of these contacts, whereas visual inspection of thin sections almost certainly results in an undercount because long processes perpendicular to the image plane are lost.

**Recurrent Cellular Neighborhoods Associated with Melanoma Progression**

To identify patterns of immune and tumor cell interaction that recur across patients and correlate with tumor progression, we used latent Dirichlet allocation (LDA; ref. 42; Supplementary Fig. S2A). LDA is a probabilistic modeling method that reduces complex assemblies of intermixed entities into distinct component communities [recurrent cellular neighborhoods (RCN)]. LDA is widely used in biodiversity studies because it can detect both gradual and abrupt changes in the composition and arrangements of natural elements (cells in a tissue or trees in a forest) while effectively accounting for uncertainty and missing data (43, 44). To identify RCNs, ~1.7 × 10^6 single cells from MEL1 to MEL13 were assigned to one of 12 basic classes based on the expression of cell type and state markers (e.g., proliferating, regulatory, and exhausted) in 22-plex CyCIF data (Fig. 2A; Supplementary Fig. S2B). The data exhibited good signal to noise across critical markers, and cell-type assignment was robust to variation in gating (Fig. 2B). Across 70 histologic ROIs that were annotated (regions of disease progression), we observed significant increases in percentage of S100A4.
Figure 2. Recurrent cellular neighborhoods associated with melanoma progression. A, Uniform manifold approximation and projection (UMAP) of single-cell data from 70 ROIs in 12 patients. The plot was generated using 50,000 single cells that were randomly sampled from the full data set of $1.5 \times 10^6$ cells. The UMAP is colored based on the phenotype (left), disease progression stage (center), and patient ID (right). B, UMAPs (shown also in A) representing feature plots of expression of selected protein markers. C, The percentage of SOX10+ melanocytes or tumor cells expressing S100A within each stage of progression. **, $P < 0.01$; ***, $P < 0.001$. D, Heat map showing the abundance of cell types within the 30 LDA-based cellular neighborhood clusters (numbers to the right of the plot); these were then reduced to the 10 meta-clusters (RCNs) shown to the left of the plot. The bar chart to the right of the heat map depicts the distribution of progression stages within each cluster, and the bar chart to the left of the heat map represents the distribution of patients within each cluster. E, Bar plot depicting the detailed breakdown of cell-type proportions within each RCN (RCN1–10; x-axis). Pie charts depicting a simplified breakdown of cell types in each RCN: myeloid (green; dendritic cells, CD11C+ macrophages, macrophages, and Langerhans cells), lymphoid (light orange; CTL, Treg, and T helper), immune-suppressive (dark orange; PDL1+ Macs, PDL1+ DCs, and PD1+ CTL), melanocytes (dark blue), and keratinocytes (yellow). APC, antigen-presenting cell; DC, dendritic cell; LC, Langerhans cell; Macs, macrophages.
SOX10+ cells were densely packed together with few infiltrating cells (Fig. 3A and B). Isolated pockets of RCN9 and RCN10 were also found in normal skin and in regions with adjacent melanocytic atypia and regions where SOX10+ cells clustered together (Fig. 3C; Supplementary Fig. S3A). The most abundant immune cells in RCN9 were CD11C+ macrophages and dendritic cells (80%), and the prevalence of this neighborhood increased significantly from precursor to MIS to invasive tumor, highlighting the formation of a myeloid-enriched tumor boundary (Fig. 3D; Supplementary Fig. S3B). When we quantified the proximity of tumor and CD11C+ myeloid cells using a 10-μm cutoff, we found that proximity volume scores increased from precursor to MIS to IM stages, independently confirming the observed increase in RCN9 frequency with progression (Supplementary Fig. S3B and S3C).

RCNs that were primarily made of immune cells could be subdivided into three classes: enriched for myeloid cells (RCN2–4), enriched for T cells (RCN6–7), and immune-suppressed (RCN5 and 8). RCN2–4 contained overlapping sets of cells, with tissue-resident macrophages predominating in RCN2 and CD11C+ cells in RCN3 and 4 (Fig. 2D). RCN2 was found throughout the dermis (and had a distribution similar to that of tissue-resident macrophages), while RCN3 and 4 were found close to the invasive tumor (Supplementary Fig. S3D). RCN6 was rich in CD4+ helper T cells (T helper) and Tregs, and RCN7 was enriched for CTLs. RCN5 and 8 had high proportions of activated PD1+ CTLs as well as Tregs and PDL1+ myeloid cells, which are immunosuppressive (47). Five of the seven immune-enriched niches (RCN3–7) significantly (P < 0.05) increased in frequency between precursor and MIS, while only one (RCN4) significantly increased between normal and precursor fields, reflecting recruitment of myeloid cells. Two significant changes were observed between MIS and IM, and this involved RCN9, which increased in abundance due to the formation of a PDL1+ sheath of cells at the IB, and RCN1, which fell in abundance due to the displacement of keratinocytes and proliferation of tumor cells in IM (Supplementary Fig. S3E).

When we quantified the proximity of immune-rich RCNs (RCN2–8) to SOX10+ cells in RCN10 (i.e., melanocytes or tumor cells), we found that myeloid-enriched (RCN2, 4) and PDL1-enriched (RCN5) communities were significantly closer to RCN10 in precursor ROIs than adjacent uninvolved skin or later disease stages. In contrast, a cytotoxic community (RCN7) appeared closer to RCN10 in precursor samples than in MIS or IM (Fig. 3E; Supplementary Table S7). To confirm this finding, we measured the distance between melanocytic cells and the nearest PDL1+ myeloid cell or CTLs. We observed a significant decrease in distances for both cell types between normal and precursor stages. Tregs also showed a significant decrease in proximity to melanocytic cells in precursor fields (Fig. 3F). Thus, at the precursor stage, the recruitment of cytotoxic T cells was accompanied not only by immune resolution but also by the first signs of immunosuppression by myeloid cells. When RCNs were mapped back to the landscape of MEL1–1 (see Fig. 1), we found that the community of tumor cells near CD11C+ myeloid cells (RCN9) that were sparsely present in association with MIS had become a nearly

---

**Figure 3.** A, Scatter plot (top) showing an FOV of the IM region (specimen MEL1-1). The cells are colored based on recurrent cellular neighborhoods (RCN1–10) to which they belong. The red and blue boxes represent regions that are magnified in the bottom panel (left and right, respectively) depicted as Voronoi diagrams. B, Example CyCIF highlighting RCNs in the invasive front of specimen MEL1-1. Top, an overall view of the invasive front stained for tumor cells (S100B: blue), macrophages (CD163: cyan), T cells (CD3: red), and dendritic cells (CD11C: green). The inset squares correspond to magnified panels at the bottom. H&E staining of a serial section of the same region is represented in the top right corner. B, Bar plot depicting the proportional distribution of RCNs (RCN1–10) to which it belongs (as in A). Examples of corresponding CyCIF images from one patient in each case are provided at the bottom. A magnified view is available in Supplementary Fig. S3A. C, Bar plots depicting the proportional distribution of RCNs (RCN1–10) among the disease progression stages (N, P, MIS, IM, and EM). E, Box plots of the distribution of the shortest distance between cells in RCN2 to RCN7 and RCN10 grouped based on progression stages. The t-test (*, P < 0.05) was used to depict significant changes in mean distances between the compared stages. The comparison made is described in the upper right corner of each plot (e.g., N vs. P). F, Shift plot shows the distance between melanocytes and CTLs, PDL1+ myeloid cells, and Tregs in normal (top) and later stages (bottom) grouped for each percentile (10, 20, 30, 40, 50, 60, 70, 80, and 90) using the robust Harrell-Davis quantile estimator. Red indicates a significant difference (P < 0.05) and gray represents nonsignificance for each percentile.
continuous sheath at the IB of IM (Fig. 3A–C). Immediately adjacent to the sheath of RCN9 cells we observed RCN3 and 4 myeloid niches in a mosaic pattern with RCN6 (T helper and Treg) and RCN5 (PDL1+ immunosuppressive) neighborhoods. The density of immunosuppressive niches was also highly variable even between nearby locations (Fig. 3A and B). RCNs containing cytotoxic T cells (RCN7) and PD1+ CTLs (RCN8) were also intermingled, consistent with local activation of T cells. Moreover, whereas intermixing of tumor cells (RCN10) and multiple immune-rich RCNs was evident in MIS, the myeloid and immunosuppressive RCN (RCN5) was largely confined to areas immediately surrounding the CD31+ vasculature in EM and IM. Individual tumors differed in the specific arrangements of RCNs, and LDA generates statistical models subject to instance to instance variation, but it was consistently true that progression was associated with not only greater levels of invasion but also the formation of increasingly complex immune environments.

### PDL1-Mediated Immune Suppression Primarily Involves Myeloid Cells, Not Tumor Cells

The importance of PD1–PDL1 interaction in melanoma is demonstrated by the success of anti-PD1 therapy. Across all 70 ROIs from 13 specimens, ~70% of CTLs expressed the activation marker PD1, but we detected very few tumor cells expressing significant levels of PD1, even in regions of tumor in which IFNG was highly expressed based on transcript profiling (see below; IFNG is a known inducer of PDL1). 3D deconvolution imaging proved to be more sensitive than conventional imaging in detecting PD1, but even in MIS, in which immune and tumor cells were intermixed, only five of 106 tumor cells imaged at high resolution in 12 FOVs were judged to be PD1+ positive. In these cases, imaging showed that the PD1+ ligand on tumor cells and the PD1 receptor on CTLs was colocalized, consistent with ligand–receptor binding (Fig. 4A; Supplementary Fig. S4A). In contrast to the paucity of PD1+ tumor cells across all patients, significant co-occurrence (P < 0.05) was observed between PD1+ CTLs and PDL1+ macrophages and dendritic cells in 44 of 70 annotated histologic domains; the frequency of this co-occurrence also increased with disease stage (Fig. 4B). To confirm that co-occurrence involved cell-to-cell interactions at least some of the time, we performed high-resolution 3D imaging of FOVs spanning the invasive front in MEL1-1 and observed frequent contact between PD1+ CTLs and either PDL1+ macrophages or dendritic cells with a concentration of PD1 and PDL1 at the site of cell-to-cell interaction (Supplementary Fig. S4B and S4C). In some cases, macrophages formed presumed inhibitory synapses with CTLs via cellular processes that extended at least one cell diameter (10 μm) from the macrophage (Fig. 4C; Supplementary Video), showing that nonadjacent cells can make functional contacts with one another. A substantial subset of PD1+ myeloid cells also expressed TIM3, which is associated with immune suppression (Fig. 4D and E).

We were surprised to find so few PD1+ tumor cells in our specimens and therefore sought confirmation via analysis of an additional set of 25 primary melanomas. These specimens were annotated by dermatopathology as containing for radial (6/25) and vertical growth phase (16/25) histologies based on H&E images (as before) and subjected to low-plex immunofluorescence imaging for PD1, PDL1, SOX10, and CD11C, followed by visual inspection of staining patterns by trained tissue biologists and pathologists. In these specimens, PDL1+ SOX10+ tumor cells were abundant (estimated to be >20% positive) in only one specimen, present (5%–20% positive) in two specimens, and infrequent (0%–5% positive) in 22 of 25 specimens (Fig. 4F and G; note that 5% threshold for PDL1 has previously been used to score tumor “PDL1 positivity” in melanoma; refs. 48–50). In contrast, >25% of CD11C+ myeloid cells scored as PDL1+ in 19 of 25 specimens. Moreover, using the same reagents, we have routinely found that nearly all metastatic melanomas contain abundant PDL1+ SOX10+ tumor cells (51). When the three melanomas containing relatively abundant PDL1+ tumor cells were examined further, we found that PDL1 positivity was strongly enriched at the IB of vertical growth phase melanoma, which is enriched in cytokines secreted by immune cells (e.g., IFNG) and the probable site of metastasis formation. We conclude that in the primary melanomas we imaged (n = 33 of 36 in total), the cells most expressing PDL1 were dendritic cells and/or macrophages, not tumor cells.

Recent data from the MC38 murine syngeneic model of colorectal cancer suggest that dendritic cells, not macrophages, may also be the relevant myeloid cell type for PDL1-mediated immunosuppression of activated CTLs in colonic adenocarcinomas (52). However, whereas the murine tumors analyzed by Oh and colleagues (52) contained many more PDL1+ macrophages than PD1+ dendritic cells, we found that these two types of myeloid cells were similar in abundance in primary melanoma (1.2%–1.4% of all cells). By high-resolution imaging of the invasive front, we also found multiple fields in which tumor cells, CTLs, dendritic cells, and other immune cell types (a subset of which expressed PD1 or PDL1) were all in direct contact with one another as part of extended networks (Fig. 4E). The presence of multidentate cell interactions and extended cellular processes containing immune-regulatory molecules suggests that multiple different immune cell types might communicate with one another via cell–cell contacts as well as autocrine or paracrine signaling. A more complete understanding of these interactions awaits high-resolution 3D reconstructions of the TME.

### Single-Cell Analysis of Invasive Tumor Reveals Large-Scale Gradients in Lineage, Immune, and Proliferation Markers

Because LDA detects discrete differences between cells (most commonly in immune differentiation markers), it is insensitive to qualitative differences between cells of a single type. To quantify such differences—specifically in tumor cells—we used principal component analysis (PCA) and shift-lag analysis, focusing on cells in the invasive tumor (~5 × 10^5 malignant single cells; Supplementary Fig. S1A). Principal components 1 and 2 (PC1 and PC2) explained 40% of the variance in these data, which represents good performance for a PCA model. The top loadings in score plots were K167, the S100A and S100B proteins, and the MITF transcription factor (Supplementary Fig. S5A). K167 is widely used to measure proliferation (53), and MITF is a master regulator of melanocyte differentiation (54). MITF is both a melanoma oncogene...
Figure 4. A, Field of MIS from a whole-slide CyCIF image of MEL1-1. A PDL1+ melanocyte (SOX10: white, PDL1: green) and CTLs (CD8: red) are highlighted with an orange box (left). Right, the polarization of PD1 (red) and PDL1 (green) to the point of contact between the interacting cells. B, Line plot showing the percentage of ROIs that displayed significant ($P < 0.05$) co-occurrence based on a proximity analysis performed between PDL1+ CD11C+ CD163− dendritic cells and PD1+ CTLs. C, Field of IM from a whole-slide CyCIF image of MEL1-1 stained for tumor (SOX10: red), macrophages (CD163: green), and CTLs (CD8: white), with three fields of macrophage–CTL contacts (yellow boxes). Maximum-intensity projections imaged at high-resolution in fields 1 and 2 are stained for DNA (blue), PDL1 (red), and PD1 (green), with cells labeled as myeloid cells (M) and engaged T cells (T). Field 3 shows tumor cells (SOX10: red), CTLs (CD8: white), and a macrophage (CD163: green). Inset white boxes in the bottom right show concentration of PD1 (green) and PDL1 (red) to the point of contact, and the long connection between a macrophage (CD163: white) and a CTL is shown in a 3D reconstruction of field 3. D, Left shows the same CyCIF FOV as in C, stained for DNA (blue), TIM3 (red), and CD8A (green). The white inset box illustrates the staining of one CD163+ CD11C+ TIM3+ myeloid cell next to a CTL (right). E, Maximum-intensity projection from the bTIL region (top left) stained for DNA (blue), macrophages (CD163: green), and T cells (CD3D: white). The white inset is magnified and stained for T-cell polarity (CD4: green, CD8A: red), the PD1–PDL1 axis (PD1: green, PDL1: red), and exhaustion markers (TIM3: red, LAG3: green). A Treg in this field is indicated with the label “Tr.”. F, PDL1 positivity in SOX10+ tumor cells (top) and CD11C+ myeloid cells (bottom). The proportions of PDL1+ tumor cells to all tumor cells (0%–5%, 5%–20%, and >20%) and PDL1+ myeloid cells to all myeloid cells (<1%, 1%–25%, and >25%) are presented in both primary melanoma cohorts (cohort 1: MEL1–13 and cohort 2: 25 primary melanomas). G, Fields of a primary melanoma and a melanoma metastasis from CyCIF images stained for DNA (blue), SOX10 (green), PDL1 (red), and CD11C (white). Top, an example of PDL1+ SOX10+ tumor cells at the deepest invasive region. PDL1+ metastasis is shown in the bottom panel. The tumor-stroma interface is indicated with a white dashed line.
(55) and a determinant of drug resistance (56); a MITF<sup>hi</sup> state has been associated with dedifferentiation and resistance to RAF/MEK therapy (57). Across the whole tumor, S100A, S100B, and MITF exhibited striking gradients in expression levels on short- and long-length scales (∼100–3,000 μm), with the highest protein levels at the invasive margin and the lowest in the middle of the EM (Fig. 5A and B; Supplementary Fig. S5B). Thus, whereas clustering of sequencing data (7) emphasizes the presence of dichotomous MITF or S100-high and -low states, imaging reveals continuous changes in protein levels through space. Spatial gradients involving morphogens have been widely studied in tissue development (58) but infrequently in cancer (59).

Spatial lag is a common spatial statistic in geography and ecology (60), and we used it to identify recurrent tumor cell communities (TCC) based on continuous differences in protein levels. Clustering of spatial lag vectors revealed the presence of 10 TCCs (see Methods for details of clustering; Fig. 5C and D) that differed from one another in hyperdimensional features (combinations of markers), although in a few cases, single markers were dominant: MHC-II positivity for TCC3 and a MITF<sup>hi</sup> Ki67<sup>hi</sup> state for TCC8 (Fig. 5C and D). TCC1 corresponded to an S100A hi state that was primarily found in EM, while MITF<sup>hi</sup> cells in TCC2 were primarily found in IM. TCC3 and TCC4 were either MHC-II hi or CCND1 hi and had distinctive spatial localization (Fig. 5C and D). The component of the IM facing the dermis had seven distinct TCCs, each of which was 2 to 5 cell diameters thick. For example, TCC3 and TCC4 were found at the IB and significantly colocalized (P < 0.05 by co-occurrence analysis) with immune cells (Supplementary Fig. S5C). TCC8 was found internal to TCC3 and TCC4, and TCC1 and TCC2 were primarily found internal to this, at the trailing edge (Fig. 5E). H&E imaging has previously suggested that vertical growth phase melanoma might have the layered arrangement of tumor cell states revealed by spatial lag analysis of CyCIF data (61).

The invasive state of cutaneous melanoma cells is thought to involve a MITF<sup>hi</sup> slowly cycling state (56). However, we found that the TCC2 community at the invasive front was comprised of 70% to 85% MITF<sup>hi</sup> Ki67<sup>hi</sup> cells (Fig. 5F; Supplementary Fig. S5D and S5E). Further evidence of proliferation was provided by positive staining of many cells in this TCC with antibodies against cyclin A2, cyclin B1, phospho-Rb (pRb, which is highest in S-phase), and phospho-histone H3 (pHH3 a marker of mitosis; Supplementary Fig. S6A), with the highest rates of proliferation in IM (∼3-fold fewer proliferating cells were present in EM; Supplementary Fig. S6A and S6B). A second previously described feature of IM is MITF<sup>hi</sup> nor slowly proliferating (relative to the rest of the tumor). NGFR (CD271) and the AXL receptor tyrosine kinase are two other proteins widely studied for their roles in state switching and drug resistance in metastatic disease (64). However, we detected only sporadic NGFR expression in MEL1 tumor cells by either mrSEQ or imaging. AXL was detected only on the plasma membranes of keratinocytes and immune cells, not tumor cells (65). Thus, the primary melanomas we imaged differed in MITF, NGFR, and AXL expression from the melanoma cell lines used in most laboratory studies, most of which were derived from metastatic disease (66).

**Microregional Transcript Profiling Identifies Spatially Distinct Immune, Mitogenic, and Survival Programs**

To study the transcriptional programs associated with different immune neighborhoods and TCCs that we identified using LDA and spatial lag analyses, we performed microregion transcriptomics (mrSEQ) on a total of 292 microregions of interest (mROI) of specimen MEL1 using PickSeq (32), which recovers 5 to 20 cells per 40-μm diameter mROI, and GeoMx (a commercial technology), which recovers ∼200 to 400 cells per ∼200-μm diameter mROI; Fig. 1B and Supplementary Fig. S6D; ref. 67). PCA of mrSEQ data revealed three primary clusters corresponding to (i) MIS, (ii) malignant tumor (EM plus IM), and (iii) regions of active immune response (IR, which were adjacent to the MIS and a bTIL region adjacent to the IB; Fig. 6A). We found that markers commonly used to detect and subtype malignant melanoma (PMEL, MLANA, TYR, MITF, and CSPG4) were all strongly and consistently expressed at the gene level in mROIs from tumor domains (EM and IM), sporadically in MIS and not in immune-rich regions (IR, bTIL), confirming the annotation of these regions and the selectivity of the method (Fig. 6B; gene names are listed in Supplementary Table S6). Single-sample gene set enrichment analysis (ssGSEA) confirmed high enrichment of melanocyte signatures in tumor but not in immune mROIs; conversely, immune signatures in IR and bTIL regions. Keratinocyte signatures were enriched in skin adjacent to MIS and IR (Fig. 6C), as expected. Moreover, results were consistent between PickSeq and GeoMx.

To investigate molecular determinants of the spatial heterogeneity within vertical growth phase melanoma revealed by spatial lag analysis, we performed DE analysis on the IM and EM domains of MEL1; this uncovered 81 significantly upregulated genes in IM and 69 upregulated genes in EM (FDR < 0.05; Fig. 6D; Supplementary Table S6). In IM, GSEA revealed significant enrichment of KRAS signaling and the downstream NF-κB and MYC programs (Fig. 6E and F). Uregulation of the KRAS pathway is expected in a tumor such as MEL1 that is mutant in NF1, which functions as a Ras GTPase-activating protein (GAP; ref. 37). BCL2A1 (68), an antiapoptotic prosurvival member of the BCL2 gene family, was expressed in IM but not in EM (Fig. 6F). EMT-associated genes were also differently expressed: the S100A4 metalloproteases, β-catenin, and vimentin (DMKN, MMP2, CTNNB1, and VIM genes) were upregulated in IM, and GSEA confirmed enrichment of an EMT-associated signature within this region (Fig. 6G; Supplementary Fig. S6C). EMT-related genes are known to promote invasion and metastasis in many human neoplasms (69), consistent with the observed invasion of this melanoma into the underlying dermis. In contrast, an RNA-sensing protein DDX58/RIG-I implicated...
Figure 5. Single-cell analysis of invasive tumor. A, CyCIF images of MEL1-1 stained for S100A (top), MITF (middle), and S100B (bottom). Boxes represent regions highlighted in B. B, Insets from A of the tumor region (IM) showing gradient expression patterns for MITF (top) and S100B (bottom). Contours describe averaged quantified marker expression. C, Heat map showing median expression of protein markers identified within TCC1-10 TCCs. The bar plot on top of the heat map shows the proportional estimate of the TCCs within histologic annotations (EM, IM, or IB). The heat map at the bottom shows the properties related to the shape of the cells (area, solidity, extent, and eccentricity) derived from the segmentation masks. D, Scatter plot mapping the physical location of the derived tumor cell clusters (TCC1-10: dark blue) in MEL1-1. Each subplot represents the location of cells within a TCC and other cells in gray. E, Scatter plot (left) showing an FOV of the IM region. Cells are colored based on their TCC (TCC1-10). The yellow circle highlights the region in B. Right, a CyCIF image of the same FOV (from specimen MEL1-1) stained for CD163 (green), MITF (yellow), Ki67 (red), and MHC-II (HLADPB1: blue). Voronoi diagram (right) generated from an FOV at the apex of the invasive front (inset highlighted in yellow). Cells are colored based on the TCC (TCC1-10) to which they belong. F, Bar plots showing the percentage of S100B-, S100A-, MITF-, Ki67-, and MHC-II (HLADPB1)-positive cells within each TCC (TCC1-10).
**Figure 6.** Microregional transcript profiling. A, PCA plot of melanoma mrSEQ transcriptomes (GeoMx). Colors indicate regional histopathology: bTIL (pink), IR (brown), MIS (green), invasive front (IB: light green), EM (gray), and center of invasive melanoma tumor (IM: yellow). EM and IM are enriched for tumor cells in this analysis, and IB contains mostly tumor cells with marginal immune infiltration. B, Expression of selected melanoma-related marker genes in mrSEQ data (PickSeq) split into three broad groups based on the PCA of GeoMx data (A). Data are mean ± SEM, **P < 0.01; ns, not significant. C, ssGSEA on mrSEQ data (PickSeq). ssGSEA scores highlight enrichment of melanoma-related gene signatures in tumor mROIs (primarily IB, IM, and EM) and immune-related signatures in the immune-rich mROIs (IR, bTIL). D, Fold difference (log2) and significance (log2 FDR) for expression of 19,500 genes between EM (n = 34) and IM (n = 16) mROIs (PickSeq). Differentially expressed genes above (brown) and below (blue/gray) a significance threshold (P < 0.05) and above a fold change threshold (log2 fold change = 10) are indicated. E, GSEA for upregulation of the KRAS pathway in IM (n = 16) compared with EM (n = 34) mROIs (PickSeq). FDR < 0.05. E, Expression (log2) of MYC, NFKB1, IGFBP2, IGF1R, and BCL2A1 in IM and EM mROIs (PickSeq). G, Heat map showing expression of genes (listed on the y-axis) known to play a role in EMT transition (PickSeq). All genes showed a significant difference between their mean expression in IM versus EM mROIs (P < 0.05). H, CyCIF image showing an FDI in MIS (top) and EM (bottom) regions. The tissue is stained for melanocytes (S0X10: yellow), endothelial cells (CD31: green), keratinocytes (pan-CK: pan-Ck white), and tumor cells (S100B: magenta). Arrows mark examples of melanocytes and tumor cells. I, Correlation network subgraph of genes associated with S100B expression in mrSEQ data (PickSeq). Nodes represent genes, and the edges correspond to the correlation between them. Brown nodes represent the genes that belong to the S100B module. Selected genes are annotated. J, Mean expression of 35 genes identified within the S100B module in mrSEQ data (PickSeq). The x-axis represents the mROIs grouped into the histopathologic annotation category from which they were isolated. K, Density plots illustrating the log-scaled protein expression of PMEL and CD63 in MIS and tumor (EM and IM) regions imaged with CyCIF.
in the suppression of cancer migration (70) was upregulated in EM (P < 0.05; Supplementary Fig. S6E). The insulin-like growth factor receptor IGF1R and the IGF binding protein IGFBP2, which is a mitogenic factor (71), were significantly upregulated in EM relative to IM (Fig. 6F). Thus, even though IM and EM are contiguous and both in the vertical growth phase, they exhibited significant differences in mitogenic, survival, and EMT-associated programs.

To identify genes differentially expressed with tumor progression, we compared mrSEQ data of tumor in aggregate (EM plus IM) with MIS; this yielded 1,327 differentially expressed genes (FDR < 0.05; Supplementary Table S6). However, differences in cellular composition were a complicating factor in this analysis: EM and IM contained mostly tumor cells with very few immune cells, but MIS was rich in immune cells and keratinocytes in addition to tumor cells. To correct for this effect, we searched for a gene shown by imaging and mrSEQ to be expressed in SOX10+ tumor cells from EM and IM but not in MIS and then constructed a correlation-based gene network to identify genes coexpressed with that gene (see Methods); S100B was found to be an ideal candidate for this purpose (epidermal Langerhans cells also stain positive for S100B, but they were too infrequent to affect the analysis; Fig. 6H). The resulting S100B correlation module comprised 35 genes (at r = 0.6), all of which exhibited statistically significant upregulation in EM-IM (FDR < 0.05; Fig. 6I and J; Supplementary Table S6). Among these genes, we validated by CyCIF the upregulation of CD63 and PMEL at the protein level (Fig. 6K). The S100B module included (i) genes implicated in metastasis or invasion in diverse cancers such as SERPINE2 (72), CTSL (73, 74), TBC1D7 (75), and NR2P2 (76); (ii) MITF-regulated genes such as SCD (77) and CDK2 (78); and (iii) oncogenes, such as the ETV5 transcription factor (79, 80). When we examined TCGA melanoma data we found that multiple genes in the S100B module (BRI3BP, CDK2, MT-ND2, PMEL, SOX10, TBC1D7, TSPAN10, and TLR) were associated with lower survival (P < 0.05; Supplementary Fig. S6F). Thus, half of the genes differentially expressed between MIS and EM-IM have established roles in oncogenesis, invasion, or progression in one or more cancers, and ~25% are associated with lower survival in melanoma. This gene set may warrant further analysis as a means to refine current approaches to determining melanoma risk by gene set analysis (e.g., using methods such as DecisionDx-Melanoma; ref. 81).

**Tumor–Immune Interaction Induces Multiple Immune Suppression Programs at the Invasive Boundary**

To better understand invasive properties of primary melanoma (Fig. 7A), we combined mrSEQ, conventional CyCIF (with a total of 80 antibodies in five separate panels), and 3D high-resolution deconvolution microscopy from tumor MEL1. In the IB region, mrSEQ data revealed significant and localized upregulation of IFNγ and JAK–STAT signaling as well as the IFN-inducible cytokines CXCL10 and CXCL11 (Fig. 7B; Supplementary Fig. S7A–S7C). CXCL10 and CXCL11 (along with CXCL9 and the CXCR3 receptor) have diverse roles in regulating immune cell migration, differentiation, and activation and play a role in response to immune-checkpoint inhibitor therapy (82). IFNγ-mediated JAK–STAT signaling can promote upregulation of the metabolic enzyme IDO1 (83), which has previously been reported to inhibit CTL activation (84, 85), and also recruitment of Tregs and myeloid-derived suppressor cells (MDSC; ref. 86). Consistent with this observation, we observed spatially restricted expression of IDO1 at the IB (Fig. 7B). Additionally, both mrSEQ and CyCIF of tumor cells revealed spatially restricted expression of MHC-II (HLA-DPB1; Fig. 7C; Supplementary Fig. S7D), which is known to be IFNγ-inducible (87). MHC-II binds to LAG3 on CD4+/CD8+ T cells, promoting melanoma persistence by upregulating MAPK/P38 signaling, and can also facilitate immune escape by suppressing FAS-mediated apoptosis (88). Thus, a tightly restricted microenvironment exists at the IB involving multiple cytokines that induce, and are induced by, the JAK–STAT–IDO1 pathway (89), leading to the formation of a highly localized immune-suppressive environment.

Expression of other interferon-stimulated genes (ISG) such as IRF1 and IRF5 was also evident at the IB: Imaging revealed nuclear staining of IRF1 in tumor cells and strong IRF5 staining in CD11c+ myeloid cells directly adjacent to the tumor boundary (Fig. 7D and E; Supplementary Fig. S7E). By integrating protein intensities across this boundary, we found that the half-maximal width for IRF1 staining was ~40 μm (Fig. 7E) and that of MHC-II expression roughly twice as wide (i.e., ~100 μm or 4 cell diameters; Fig. 7E; Supplementary Fig. S7D). Thus, mrSEQ and imaging are consistent with a paracrine signaling mechanism in which IFNγ arising in the peritumoral stroma (including the bTIL region) diffuses into the tumor, inducing ISGs at the invasive front (90).

A reciprocal mechanism involved the macrophage migration inhibitory factor (MIF), an inflammatory cytokine over-expressed in a variety of cancers (91). MIF was more abundant in tumors (MIS, IM, and EM) than in immune-rich regions (bTIL, IR; DE with P < 0.05; Fig. 7B) and was confirmed by imaging (Supplementary Fig. S7F). MrSEQ data showed that the MIF receptor CD74 (which is induced by IFNγ; ref. 92) was expressed in immune-rich (bTIL) regions adjacent to the IB (Fig. 7B), and CyCIF confirmed this at the protein level (Supplementary Fig. S7G). CD74 was also found to be expressed in melanoma cells (where it can promote PI3K/AKT activation and cell survival) but was spatially restricted to cells at the IB (Supplementary Fig. S7G). We also detected elevated expression of a second MIF receptor, CXCR4, and another cognate ligand, CXCL12, in the bTIL region; CXCR4 activation leads to expansion of immunosuppressive Tregs (93). CXCR4 is the chemokine receptor most commonly found on cancer cells, and binding to CXCL12 is thought to promote invasive and migratory phenotypes leading to metastasis (94). However, mrSEQ showed that CXCR4 levels were low in IM and EM (Supplementary Fig. S7H). Thus, mrSEQ data are most consistent with MIF expression in tumor cells that acts in a paracrine manner on immune cells with overlapping CXCL12–CXCR4 signaling, also in the immune compartment.

Overall, these data reveal the pattern of immune cell activation and immunosuppression involving highly localized cytokine signaling and direct cell-to-cell contact all within a few cell diameters of the IB.

However, successful immune editing and clearance of SOX10+ tumor cells were also observed at regions of inflammatory and terminal regression in MIS, only a few
**A** Invasive tumor boundary

**B**

- **CXCL10**
  - bTIL: 0, IB: 3.00 ± 0.30, IM: 4.50 ± 0.30
  - bTIL: 3.00 ± 0.30, IB: 0, IM: 4.50 ± 0.30

- **CXCL11**
  - bTIL: 0, IB: 2.25 ± 0.25, IM: 3.75 ± 0.25
  - bTIL: 2.25 ± 0.25, IB: 0, IM: 3.75 ± 0.25

**C** MHC-II–expressing tumor cells

**D**

- SOX10
  - bTIL: 0, IB: 3.00 ± 0.30, IM: 4.50 ± 0.30
  - bTIL: 3.00 ± 0.30, IB: 0, IM: 4.50 ± 0.30

- CD11C
  - bTIL: 0, IB: 2.25 ± 0.25, IM: 3.75 ± 0.25
  - bTIL: 2.25 ± 0.25, IB: 0, IM: 3.75 ± 0.25

**E**

- RIF1 from tumor
  - MIF1 autocrine loop
  - CXCL12–CXCR4

- MHC-II binds LAG3
  - T cells

- Rare PDL1+ tumor

**F**

- Fraction of cells
  - PD1+ DCs
  - PD1+ Macs
  - Treg
  - Exhaused CTL

**G**

- Epidermis
  - MHC-II–PD1

- Dermis

- Z score

**H**

- MS
  - CTLA4
  - PDCD1
  - GNLY
  - HAVCR2
  - TIGIT
  - LAG3
  - PRF1
  - GZMB
  - GZMA

**I**

- Normal
  - Precursor
  - Melanoma in situ
  - Invasive melanoma

- Remodeling of the TME during progression

- Immune response
  - Cytolytic T cells are farther away

- Cytotoxic T cells are farther away

- Immune-suppressive environment

**J**

- Tumor
  - Interface
  - Stroma

- PD1+ tumor
  - PDL1+ DCs/Macrophages
  - DCs/Macs
DISCUSSION

In this article, we exploit histologic features used clinically to stage primary cutaneous melanoma as a framework for analyzing multiplexed imaging and mRNASeq data along an axis of tumor progression from precursor fields to IM. We also examine regions near the dermal–epidermal junction in which immunooediting is ongoing or had reached a resolution with few or no tumor cells remaining. Molecular evidence of progression was obtained using protein markers (by CyCIF) and oncogenic programs (by mRNA expression) both within specimens, each of which comprised several distinct histologies, and also across a patient cohort. Disease-relevant morphologic features ranged in length scale from 0.5 μm (organelles) to 20 mm (invasive fronts), and we found that imaging the entirety of individual specimens up to 1 cm in length—not a TMA or a small ROI—was essential for retaining information on tissue context and for the success of our approach (54). Accompanying high-resolution 3D imaging revealed the presence of immune synapses and PD1–PDL1 colocalization to the plasma membranes of neighboring cells; we interpret these as evidence of functional cell-to-cell interaction. Juxtaocular receptor-ligand interactions of this type appear to be relatively common among cells lying along the immune-rich invasive tumor boundary (up to 20% of adjacent cells making contact in an exemplary field). At the current state of the art, however, only a relatively small number of whole-slide images could be analyzed for spatial patterns in their entirety (n = 13 patients and 70 histologic ROIs). Thus, the progression-associated changes described in this article should be regarded as representative rather than comprehensive; in contrast, discovery of new (progression) biomarkers by traditional IHC typically involves analysis of at least 100 specimens followed by clinical trials (95).

The use of LDA on high-plex data made it possible to identify recurrent combinations and arrangements of cell types across ROIs. The frequency of recurrent cellular neighborhoods (RCN), and their proximity to one another, changed with disease stage (Fig. 7J). Relative to adjacent normal skin, changes in the immune environment were detectable in fields of melanocytic atypia (precursor fields), but the largest differences along the progression axis appeared to involve precursor fields and MIS. This involved the recruitment to the tumor domain of CTLs, many of which were PD1-positive and presumably activated, as well as increases in suppressive Tregs and PDL1-expressing myeloid cells. The resulting immunosuppressive environment became more consolidated between MIS and invasive stages. For example, in sample MEL1, a community of cells involving tumor and PDL1+ myeloid cells (macrophages and dendritic cells in roughly equal proportion) formed a thin and continuous sheath along the invasive front. TILs were largely excluded from the tumor at this stage except in the immediate proximity of small vascular structures that were found throughout the EM.

Whereas LDA was effective at identifying neighborhoods involving different types of cells, spatial lag modeling on CyCIF data identified recurrent patterns involving continuous differences in protein levels on a scale of 10 to 100 cell diameters. Spatial gradients on similar scales were also observed for several protein markers—MITF or S100B, for example. Thus, whereas LDA and clustering of transcriptional data highlight discrete differences in cell states, imaging demonstrates the presence of gradients reminiscent of those found in developing tissues (58, 59). In general, significant differences among communities of cancer cells identified by shift-lag modeling involved hyperdimensional features (combinations of markers instead of single proteins) consistent with the current understanding of molecular determinants.

Figure 7. A, CyCIF image of specimen MEL1-1 showing a protruding edge of the invasive tumor (SOX10: violet, S100B: pink) into the dermis; outside the tumor boundary marked by a white line (dashed) is the brisk TIL region, which contains activated/exhausted T cells (CD3: green, PD1: red) and myeloid cells (CD11c: blue). B, Expression of CXCL10, CXCL11, IDO1, MIF, and CD74 among histologic sites (PickSeq data). Values represent mean ± SEM, *P < 0.05, **P < 0.01, ***P < 0.001 ns, not significant. C, CyCIF FOV of MEL1-1 highlighting the spatial arrangement of MHC-II+ tumor cells at the invasive front. Tumor cells were stained with SOX10 (cyan), MHC-I (HLA-A: green), and MHC-II (HLA-DRP1: red). Magnified regions outlined in magenta and yellow squares illustrate MHC-II and MHC-I staining of tumor cell membranes. D, CyCIF of specimen MEL1-1. Left, zoomed-out view of invasive front stained for melanocytes (SOX10: blue), myeloid cells (CD11c: red), and interferon signaling (IFNγ: green). Top right, zoomed-in view of invasive front stained for melanocytes (SOX10: blue), myeloid cells (CD11c: red), and interferon signaling (IFNγ: green). Bottom right, zoomed-in view of invasive front stained for melanocytes (MART1: green), myeloid cells (CD11c: blue), and interferon signaling (IFNγ: red). E, Line plot showing scaled fluorescence intensity of SOX10 (blue) and IFNγ (pink) within tumor (left of the dashed blue line) and outside (stroma, right of the dashed blue line) the invasive tumor front seen in D. F, Stacked bar graph showing the proportions of lymphoid and myeloid cells between the histologic regions (IR, MIS, and b/IR) in specimen MEL1-1. APC, antigen-presenting cell; Macs, macrophages. G, CyCIF maximum-intensity projection images of MEL1-1 of the region of inflammatory regression (shown in C). H, Heat map showing expression of genes related to immune checkpoints and T-cell activation between histologic mROIs in patient MEL1-1 (GeoMx). Significant upregulation in comparison with the EM region (P < 0.05) is highlighted in red and nonsignificant in gray. I, Schematics of remodeling of the TME with disease progression; see text for details. J, Summary of mechanisms of immune suppression detected in sample MEL1-1.
of cellular morphology (96). Moreover, gradients in MITF or S100B are likely to be indications that tumor cell phenotypes are graded in space but not causes of this variation. One unexpected finding involved the “invasive” state of melanoma cells, which is often described as being MITFhi with slow proliferation. Spatial lag modeling showed that MITFhi K6C2hi cells were common in MEL1 in the immediate proximity of the invasive front, and mrSEQ showed that these cells were significantly enriched in EMT-associated programs, which are common along the invasive boundaries of many other types of tumors. Future studies on paired primary and metastatic tumors will be required to understand how these data relate to previous analysis of MITF-high and -low states, which has largely been performed in cell lines.

CyCIF and mrSEQ revealed a pattern of cytokine production and receptor expression at the IB of MEL1 consistent with paracrine regulation of both tumor and immune cells (Fig. 7J). The dermis in this region was rich in TILs (corresponding to a brisk TIL response in the Clark grading system) and was the site of highest IFNγ production. A band of cells ~2 cell diameters wide in the adjacent IM stained positive for nuclear-localized IRF1, the master regulator of interferon response (yellow cells in Fig. 7J); mrSEQ showed that JAK–STAT signaling was active in this region, and IDO1 was differentially expressed. IDO1 converts tryptophan into kynurenine, which activates Tregs and MDSCs, and is known to be immunosuppressive in melanoma (97). MHC-II was also expressed in both immune and tumor cells at the IB, in a band roughly twice as wide as IRF1, and is known to function in this context by binding to LAG3 on TILs, leading to inhibition of TCR signaling and T-cell activation (98, 99). MIF1 was another inflammatory cytokine found at the invasive front and was expressed primarily in the invasive tumor region; responsive CXCR4-expressing immune cells were found in the stroma. MIF1 may also have an autocrine activity because expression of the MIF1 receptor CD74 was detected in the tumor itself. CXCR4-expressing immune cells are also responsive to CXCL12, which was expressed in the TIL-rich stroma. CXCR4 is the cytokine receptor most commonly found on melanoma and other types of cancer cells, and CXCR4-CXCL12 signaling is thought to promote metastasis (94), but we did not observe CXCR4 expression in MEL1 by mrSEQ. We conclude that the immunosuppressive activity of IFNγ manifests itself in MEL1 in a spatially restricted manner involving a sheath of tumor and myeloid cells surrounding the invasive tumor. Undoubtedly, the magnitude of these effects will vary among primary tumors, but our analysis illustrates how reciprocal cytokine signaling between tumor and immune cells can shape the local TME.

Performing spatial proximity analysis on imaging data (with a 10–20 μm cutoff) made it possible to identify cells that are sufficiently close to one another that physical contact is probable. We were able to visualize these contacts and infer function using high-resolution 3D imaging of ~5 × 104 cells. The most informative images were those involving cytotoxic T and melanoma cells that resulted in the polarization of CD8 (a coreceptor for the TCR) at the point of contact, suggesting the formation of a synapse. PD1+ CTL cells were also observed in contact with PDL1-expressing macrophages and dendritic cells, resulting in receptor-ligand colocalization. In some cases, these contacts involved surprisingly extended processes (~10 μm) in which macrophages appeared to stretch toward T cells. In other cases, multiple CTLs, T helper, and myeloid cells were found to be in physical contact with one another and with tumor cells with evidence of receptor or ligand polarization. The functional significance of these clusters awaits further analysis using a greater diversity of immune markers, but they are presumably a physical manifestation of the competing activating and inhibitory effects of other immune cells on CTLs.

Overall, we found evidence of at least six immunosuppressive mechanisms operating near the invasive front. Particularly striking was the overlap in the binding of PD1+ CTLs to PDL1+ macrophages and dendritic cells and tumor cell–intrinsic phenotypes such as MHC-II and IDO1 expression. Unexpectedly, we only rarely detected high expression of PDL1 on tumor cells by either whole-slide imaging or high-resolution microscopy (even when IFNγ expression was detected). This finding was validated using a separate cohort of 25 primary cutaneous melanomas and contrasts with data collected in parallel from metastases, in which strongly PDL1+ tumor cells are common. We conclude that myeloid cells are likely to provide the predominant source of PDL1 bound to PD1+ T cells in the tumors in our cohort. Data obtained by Oh and colleagues (52) in mouse models suggest that the functionally significant cell type is likely to be PDL1+ dendritic cells, but in our specimens, dendritic cells and macrophages appear to be similar in abundance at tumor-invasive boundaries.

CTLs were found to engage tumor cells in a region of inflammatory regression adjacent to MIS in MEL1. The additional presence of an adjacent region of complete regression, which was rich in immune cells but free of tumor cells, suggests that immune editing was ongoing and successful. However, these regions also had a preponderance of terminally exhausted CTLs, showing that the characteristics of a successful and self-limiting antitumor–immune response in data such as those presented here can resemble those of immunosuppression in IM. The primary difference we observed between regions of regression and invasion with immunosuppression was a substantially lower level of PDL1+ myeloid cells, but further research will be required to determine if this is generally true.

Limitations of This Study

One challenge encountered in molecular analysis of primary melanoma is that, as a diagnostic necessity, specimens are available only in FFPE form, generally precluding scRNA-seq for research purposes. Sequencing of carefully selected microregions by mrSEQ provides meaningful information on activated pathways and differential gene expression linked to histologic features but is not yet single-cell resolution. A second challenge is that meaningful outcome analysis requires long follow-up: All patients whose tumors were analyzed in this study were diagnosed between 2017 and 2019 and were alive at the time of the last follow-up; ~75% were disease-free. Thus, we used histologic progression, not outcome, to organize the data in a biologically meaningful fashion. A final limitation in any molecular study of
patient-derived specimens is that only one time point can be evaluated per patient. Our analysis of tumor samples exhibiting progression within the same specimen helps to mitigate this issue.

Despite the scope of the current data collection effort, 13 specimens are too few to be representative of the diversity of cutaneous melanoma. We estimate that data collection will need to be scaled up 5- to 10-fold to determine whether many of the features observed in MEL1 are significantly associated with progression in other specimens. Moreover, spatially resolved mRNA expression and high-plex imaging data support each other in many cases, but this was not always true. This is not unexpected because mRNA and protein expression are known to be uncorrelated in many cases (100), and cell morphology represents a hyperdimensional feature in the gene expression space (96). 3D image data have provided valuable insight into cell-to-cell interactions, but automated segmentation of these data remains difficult, and most conclusions were derived from a human inspection of images. More generally, the greatest limitation in the current work is related to the underdevelopment of software tools for characterizing large high-plex tissue images. Much therefore remains to be discovered from the images we have collected. Full-resolution level 3 images (101) and associated single-cell data are therefore being released in their entirety, without restriction, for follow-on analysis.

METHODS

Contact for Reagent and Resource Sharing

This manuscript does not contain any unique resources and reagents; all data are provided for download without restrictions. Any questions should be directed to the lead contact Peter Sorger (peter_sorger@hms.harvard.edu).

Clinical Samples

Using medical records and pathologic review of H&E-stained diagnostic specimens, we retrospectively identified 13 patients with tissue samples containing various stages of melanoma progression (Supplementary Tables S1 and S2). The samples were retrieved from the archives of the Department of Pathology at Brigham and Women’s Hospital and collected under Institutional Review Board approval (FWA00007071, Protocol IRB18-1363) under a waiver of consent. Fresh FFPE tissue sections were cut from each tumor block. The first section of each block was H&E-stained and used to annotate ROIs (Supplementary Table S3). The remaining subsequent FFPE slides were used for cyclic multiplex immunofluorescence imaging (CyCIF) experiments to characterize markers of melanoma progression and the features of the immune microenvironment within various stages of melanoma. A specimen from a single patient, MEL1 (samples MEL1-1, MEL1-2, and MEL1-3), was selected for deeper profiling with CyCIF and high-resolution imaging, in addition to microregion transcriptomics (PickSeq and GeoMus). The clinical, biospecimen, and imaging-level metadata were all collected following the Minimum Information About Highly Multiplexed Tissue Imaging (MITI) standards (101).

Based on the melanoma diagnostic criteria, the histopathologic annotations included normal skin (N), melanoma precursor lesions (P: melanocytic atypia, dysplasia, and hyperplasia), M1S, vertical growth phase of melanoma, radial growth phase of melanoma, invasive (IM), and nodular melanoma; the exophytic component of the polypoid melanoma was labeled as EM. These ROIs were further classified and subdivided based on the presence of immune infiltrate (bTIL, IR, none) and various histologically distinct structures [epidermis, dermis, invasive front (IB)]. The bTIL region was defined as a dense lymphocytic infiltrate in the stroma adjacent to the invasive tumor. IB was defined as the tumor region extending ∼20 μm from the tumor–stroma interface. The most representative regions of each histologic category from each specimen were selected in order to avoid interobserver variability. In the case that a single specimen contained more than one histologic region in each category (e.g., precursor regions on both sides of vertical growth phase melanoma), we performed neighborhood analyses separately, as these regions were not physically adjacent.

Imaging (H&E and Tissue-Based CyCIF)

H&E-stained FFPE slides were digitized using an Olympus VS-120 automated microscope using a 20x objective (0.75 NA) at the Neurobiology Imaging Core at Harvard Medical School. CyCIF was performed as described in (31) and at protocols.io (https://dx.doi.org/10.17504/protocols.io.bjukkew). In brief, the BOND RX Automated IHC Stainer was used to bake FFPE slides at 60°C for 30 minutes, dewax using Bond Dewax solution at 72°C, and perform antigen retrieval using Epitope Retrieval 1 (Leica) solution at 100°C for 20 minutes. Slides underwent multiple cycles of antibody incubation, imaging, and fluorophore inactivation. Antibodies were incubated overnight at 4°C in the dark; in contrast to the protocol.io method, this was performed using a solution that also included Hoechst 33342 for DNA staining. Before imaging, glass coverslips were wet-mounted using 100 μL of 70% glycerol in 1× PBS. Images were acquired using a CyteFinder slide scanning fluorescence microscope (RareCyte Inc.) with a 20x/0.75 NA objective. Slides were soaked in 42°C PBS to facilitate coverslip removal, and then fluorophores were inactivated by incubating slides in a solution of 4.5% H2O2 and 24 mmol/L NaOH in PBS and placing them under an LED light source for 1 hour. The list of all antibody panels used in the experiments is presented in Supplementary Table S4. All the used antibodies were validated with a multistep process, including comparing multiple antibodies with one another and with clinical standards and by visual inspection on individual stained FFPE tissue sections. Antibodies that passed these criteria and followed the expected staining pattern were included in only downstream analysis.

One FFPE section from sample MEL1-1 was imaged with CyCIF at high-resolution using a DeltaVision ELITE microscope (Cytova; formerly GE Sciences) equipped with an Edge 42 (PCO) sCMOS camera. For accurate deconvolution, an oil refractive index of 1.524 was selected through optimizing multiple acquired point-spread functions, as it provided the highest image quality. The slide was wet-mounted with a high-resolution 1.5-grade coverslip (ThorLabs CG15KH) using 105 μL of 90% glycerol. The fields for image acquisition were selected by evaluating SOX10 staining to locate and identify melanocytes and tumor cells, yielding a total of 42 fields across the annotated regions (Fig. 1C, Supplementary Fig. S1D). Images were acquired in 5-μm Z-stacks at 200-nm step size to create a 3D representation of the sample. Excitation wavelengths were 632/22 nm, 542/27 nm, 475/28 nm, and 390/18 nm for four-channel imaging.

PDL1 expression was also quantified in an additional cohort of 25 additional primary melanomas (cohort 2; Fig. 4F and G) selected from the Brigham and Women’s Hospital tissue bank using the same criteria as the 13 specimens described above. These specimens were subjected to lower-plex CyCIF imaging using antibodies listed in Supplementary Table S4. The frequency of PDL1 positivity on tumor and myeloid cells was then visually quantified as the percentage of PDL1+ SOX10+ tumor cells (binned as follows: 0%-5%, 5%-20%, >20%) or CD11c+ myeloid cells (binned as follows: <1%, 1%-25%, >25%). Broad bins were chosen to make the results robust to counting errors in regions of tissue where cells were densely packed.
Microregion Transcriptomics

For the microregion transcriptomic profiling (mrSEQ) using PickSeq and GeoMx, we identified mROIs of MIS, EM, IM, IB, and bTIL from samples MEL1-1, MEL1-2, and MEL1-3 based on the corresponding H&E-stained sections. Freshly cut serial sections from the corresponding tissue blocks were used for the mrSEQ experiments.

PickSeq Processing and Library Preparation

PickSeq is a method by which 40-μm mROIs are physically extracted using a robotic arm followed by mRNA extraction and RNA-seq (32). Two hundred twenty-two ROIs representing five morphologically distinct sites (MIS, IM, IB, bTIL, and EM; Supplementary Fig. S6D) were selected for collection and library preparation. The FFPE sections were deparaffinized and rehydrated using the Histogene Refill Kit (Arcturus). Slides were immersed in xylene for 5 minutes, immersed in a second jar of xylene for 5 minutes, and then incubated in a series of ice-cold solutions with 0.0025% RNase Plus (Promega): 100% ethanol for 1 minute, 95% ethanol for 1 minute, 75% ethanol for 1 minute, 1× PBS for 1 minute, and another tube of 1× PBS for 1 minute. Slides were stained with 50 μmol/L DRAQ5, a Far-Red DNA Dye (Thermo Fisher), in PBS, with 0.1% RNasin Plus. A volume of 20 μL of the lysis buffer [1:16 mix of Proteinase K solution (Qiagen) in PKD buffer × cold 1×] was added to each slide. The slides were then incubated at 80°C until ready for downstream RNA-seq workflow.

PCR tubes containing tissue microregions in the lysis buffer were removed from the freezer, allowed to thaw at room temperature for 5 minutes, and incubated at 56°C for 1 hour. The slides were briefly vortexed, spun down, and placed on ice. Dynabeads Oligo(dT)25 beads (Thermo Fisher) were washed three times with ice-cold 1× hybridization buffer [NorthernMax buffer (Thermo Fisher) with 0.05% Tween 20 and 0.0025% RNase Plus] and resuspended in original bead volume with ice-cold 2× hybridization buffer (NorthernMax buffer with 0.1% Tween 20 and 0.005% RNase Plus). A volume of 20 μL of washed beads was added to each lysis sample, mixed by pipette, and incubated at 56°C for 1 minute, followed by room temperature incubation for 10 minutes. Samples were placed on a magnet and washed twice with an ice-cold 1× hybridization buffer and then once with ice-cold 1× PBS with 0.0025% RNase Plus. The supernatant was removed, and the pellet was resuspended in 10.5 μL nuclease-free water. Samples were incubated at 80°C for 2 minutes and immediately placed on a magnet. The supernatant was transferred to new PCR tubes or plates and placed on ice for subsequent whole-transcriptome amplification or stored at −80°C.

Reverse transcription and cDNA amplification were performed using the SMART-Seq v4 Ultra Low Input RNA Kit for Sequencing (Takara Bio). The resulting amplified cDNA libraries were assessed for RNA concentration using the Qubit dsDNA HS Assay Kit (Thermo Fisher) and for fragment size distribution using the BioAnalyzer 2100 High Sensitivity DNA Kit. The sequencing libraries were prepared with the TruSeq DNA-seq Kit (Taike Bio). The resulting libraries were characterized using the Qubit dsDNA HS Assay Kit and BioAnalyzer 2100 High Sensitivity DNA Kit, pooled at equimolar ratios, and sequenced using a MiSeq (Illumina) or NextSeq (Illumina) sequencer.

GeoMx Processing and Data Collection

NanoString GeoMx gene expression analysis utilizing the cancer transcriptome array (CTA) probe set was performed by the Technology Access Program at NanoString using previously described methods (33). Briefly, a 5-μm section of FFPE melanoma was dewaxed and stained overnight with antibodies targeting melanocytes (PMEL), epithelial (pan-CK), cells, and immune cells (CD45), defining cell morphology and highlighting ROIs. The section was hybridized with the CTA probes before being loaded into the instrument. Seventy ROIs representing five morphologically distinct sites (MIS, IM, IB, bTIL, and EM; Supplementary Fig. S6D) were selected for collection and library preparation. Sample processing and sequencing were performed by the Technology Access Program at NanoString. Probe measurements and quality control data were provided by NanoString.

3D Image Processing, Alignment, and Visualization

Acquired images were deconvolved using constrained iterative in SoftWorx to reassign photons to the focal plane and increase image contrast. Maximum-intensity-projection images were also generated. Subsequently, cycles were aligned using a custom script written in MATLAB (Mathworks). Briefly, 2D image registration was first carried out using the Hoechst channel maximum-intensity projections. This was followed by registration along the z-axis. The registered 3D data sets were visualized in Imaris (Bitplane) and surface-rendered for visualization.

PickSeq Data Alignment and Expression Matrix Generation

The raw FASTQ files were examined for quality issues using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to ensure library generation and sequencing were suitable for further analysis. The reads were processed using the bcbio pipeline v1.2.1 software (102). Briefly, reads were mapped to the GRCh38 human reference genome using HISAT2 and Salmon. Length-scaled transcripts per million derived from Salmon were used for the downstream analysis.

Differential Gene Expression and Pathway Analysis

DESeq2 R package was used to generate the normalized read count table based on their estimateSizeFactor() function with default parameters by calculating a pseudoreference sample of the geometric means for each gene across all samples and then using the “median ratio” of each sample to the pseudoreference as the sizeFactor for that sample. The sizeFactor was then applied to each gene’s raw count to get the normalized count for that gene. DESeq2 (103) was used for differential gene expression analysis. A corrected P value cutoff of 0.05 was used to assess significant genes that were upregulated or downregulated using the Benjamini-Hochberg (BH) method. PCA was performed using the prcomp R package. A compendium of biological and immunologic signatures was identified from publicly available databases or published articles for performing enrichment analysis. To perform GSEA, two previously published methods (GSEA and ssGSEA) were primarily used. The R package clusterProfiler was used to perform GSEA and the R package GSVA was used to perform ssGSEA, which calculates the degree to which the genes in a particular gene set are coordinate upregulated or downregulated within a sample. The KEGG and JAK–STAT signatures were curated from MSigDB (104), and immune cell-related and melanoma-related signatures were curated from published studies (7, 105, 106).

Network Analysis to Identify Genes within the S100B Module

The normalized expression matrix (PickSeq data) was loaded into the network analysis tool BioLayout (107). Within the tool, a Pearson correlation matrix was generated, that is, an all versus all comparison matrix.
of expression profiles across all samples. A gene correlation network (GCN) was then generated using a correlation threshold value of 0.6. In the context of a GCN, nodes represent genes and edges represent the correlations between them. A single-step neighbor walk was performed within the tool from S100B to determine the S100B module.

**CyCIF Image Preprocessing and Quality Control**

The complete preanalytical CyCIF image processing (stitching, registration, illumination correction, segmentation, and single-cell feature extraction) was performed using the MCIMICRO pipeline (36), an open-source multiple-choice microscopy pipeline, versions 60929d5b82 and 7547e0:42a (full codes available on GitHub at https://github.com/labsyspharm/mcimicro). For the generation of probability maps and the nuclei segmentation, a trained U-Net model, UnMicsv1, was used followed by a marker-controlled watershed used for single-cell segmentation (108). A diameter range of 3 to 5 pixels was used for nuclei detection. The cytoplasmic area was captured by expanding the nuclei mask by 3 pixels. After generating the segmentation masks, the mean fluorescence intensities of each marker for each cell were computed, resulting in a single-cell data table for each acquired whole-slide CyCIF image. The X/Y coordinates of annotated histologic regions on the whole-slide image were used to extract the quantified single-cell data of cells that lie within the ROI range.

Multiple approaches were taken to ensure the quality of the single-cell data. At the image level, the cross-cycle image registration and tissue integrity were reviewed; regions that were poorly registered or contained severely deformed tissues and artifacts were identified, and cells inside those regions were excluded. Antibodies that gave low confidence staining patterns by visual evaluation were excluded from the analyses. The quality of the segmentation was assessed, and the segmentation parameters were iteratively modified to improve the accuracy of the segmentation masks. On the single-cell data level, correlations of DNA staining intensities in different cycles were used to filter out cells that were lost in the cyclic process with a threshold of correlation coefficient less than 0.8.

**Single-Cell Phenotyping**

We developed a gating-based phenotyping approach to classify cells (109). First, an open-source OpenSeadragon-based visual gating tool (https://github.com/labsyspharm/cycif_viewer) was used to derive gates (the cutoff value that distinguishes cells that express and do not express a particular marker). The identified gates for each marker were subsequently used to rescale (similar to batch correction) the single-cell data between 0 and 1, such that the values above 0.5 identify cells that express the marker and vice versa (rescale function within scimap). We repeated this process on every image independently and merged them into a single large single-cell data set. The scaled single-cell data were used for cell-type calls.

We built an algorithm (phenotype_cells function within the scimap python package) that assigns phenotype labels to individual cells based on a sequential probability classification approach. The underlying assumption is that the probability of a real signal would be higher than the blood-through/artifact signals that arise due to chromatic or segmentation artifacts. For example, if a B cell (CD20+ and T cell (CD3+ are physically next to each other with some blood-through of CD3 signal into the B cell (CD20+ cell), the algorithm compares the scaled intensity of CD3 and CD20 and assigns the cell as a B cell due to higher levels of CD20 expression. It is sequential as we follow a tree structure, whereby the cells are initially classified into large groups such as tumor (e.g., based on SOX10/S100B) and immune (CD45 expression), and as a second step, the immune cells are further divided into cell types such as T cells, B cells, etc., which are further divided into finer subtypes in a sequential step. An input to this algorithm is a relationship chart (phenotyping workflow; Supplementary Figs. S1C and S2B; Supplementary Table S5) between markers and cell types (phenotypes). Each cell is binned into a phenotype class based on the highest expression of a given marker. If a cell does not express any of the markers (i.e., <0.5) in the phenotyping workflow sheet, it is assigned to an unknown class. On average, we found that ~25% of cells (15%–39% across all 13 patients) were classified as unknown. Based on inspection of H&E images, these cells are likely to include fibroblasts, adipocytes, muscle, and other stromal cells. By using “AND, OR, ANY, ALL” as parameters, in combination with “POS or NEG” expression patterns, we were able to define the desired cell types identified via unsupervised clustering and manual inspection of the images. The assigned cell types were then verified by overlaying the phenotypes onto the image using Napari (image_viewer function within scimap). In total, we assigned phenotype labels to 1.7 × 10^6 single cells from 70 CyCIF ROIs corresponding to all progression stages (specimens MEL2–MEL13) and a whole-slide data set from specimen MEL1-1.

**Phenotype Co-occurrence Analysis**

For each cell in the CyCIF data set, its local neighborhood was captured by querying a radius of 20 μm from the cell centroid as measured by the Euclidean distance between X/Y coordinates. The phenotypes of those cellular neighbors were mapped to generate a neighborhood matrix containing the neighbor phenotype for every cell. We then randomly permuted (1,000 times) the neighborhood phenotypes without changing the number of neighbors (to maintain the tissue structure) and generated 1,000 random cell–cell neighborhood matrices. The frequency of all cell-to-cell pairwise proximity from the real neighborhood matrix was compared with the 1,000 randomly generated neighborhood matrices to identify significant proximity or avoidance between pairs of cell types. The P values were derived for every pairwise proximity according to the following formulas:

\[ p_{ij} = \frac{1}{2} \left( 1 + \frac{c_{ij}}{g_{ij}} \right) \]

where \( c_{ij} \) is the number of times the ith cell type was found proximal to the jth cell type. Its associated P value \( p_{ij} \) was calculated by

\[ z = \frac{c_{ij} - \mu_{ij}}{\sigma_{ij}} \]

\[ p_{ij} = \text{erfc} \left( \frac{z}{\sqrt{2}} \right) \]

where \text{erfc} is the complementary error function calculated using the python function “scipy.stats.norm.sf.” The method is implemented under the spatial_interaction function in the scimap python package.

**Spatial Lag Analysis to Define TCCs**

For each tumor cell in the CyCIF data set (MEL1-1), its local neighborhood was captured by querying a radius of 20 μm from the center cell as measured by the Euclidean distance between X/Y coordinates. A spatial lag vector was derived for each neighborhood by taking the product of the expression matrix and a weighted proximity matrix. The weights were assigned such that the closest cell within a neighborhood received the highest weight (weight = 1) and the farthest received the lowest weight (weight = 0). The weights were then normalized to account for the number of cells within each neighborhood. The spatial lag matrix was then clustered using Python’s sklearn implementation of K-means with \( k = 20 \) and manually grouped (hierarchical clustering assisted) into meta-clusters (10 clusters) based on similar expression patterns visualized using a heat map. The method is implemented under the spatial_expression function in scimap python package.

**Proximity Volume Scoring**

To quantify the abundance of cell-to-cell proximity between cell types of interest (COI) observed in CyCIF images, we developed a scoring system that weights user-defined proximity patterns. The proximity volume score is defined as the proportion of COI found in proximity to one another (10 μm) compared with the total number of cells within...
that image. We calculated the spatial volume score between COIs (tumor and CD11C+ myeloid cells) for each image and averaged them across images belonging to the same stage. The scoring is implemented under the `spatial_volume` function in the `scimap` python package.

**RCN Analysis to Identify Microenvironmental Communities**

For every single cell from specimens MEL1 to MEL13, its local neighborhood was captured by querying a radius of 20 μm from the center cell as measured by the Euclidean distance between X/Y coordinates. The cells within each neighborhood were mapped to the cell-type assignment made, and their frequency within each neighborhood was computed. The frequency matrix was then used for microenvironment modeling using a method called LDA, which is commonly used in the natural language processing and information retrieval community. Python’s `gensim` ([https://pypi.org/project/gensim/](https://pypi.org/project/gensim/)) implementation of LDA model estimation was used to train the algorithm. The number of latent motifs to be extracted from the training corpus was determined empirically (motifs = 10). The latent vectors (weights) were recovered from the model and clustered using `scikit-learn` implementation of K-means with k = 30. The optimal number of K-means clustering was determined by looking for the elbow point in the computed cluster heterogeneity during convergence (Supplementary Fig. S2D). A fairly lenient elbow point (k = 30) was used to capture the maximal variance in our data set and to account for smaller communities. The clusters were then manually grouped (hierarchical clustering-assisted) into meta-clusters (11 clusters) based on similar microenvironmental community patterns. To validate the RCN assignment, these meta-clusters were overlaid on the original tissue H&E-stained and fluorescence images. For example, RCN1 generally mapped to the epidermis and the dermis. To validate the RCN assignment, these meta-clusters were overlaid on the original tissue H&E-stained and fluorescence images. In parallel, we derived RCNs using an alternative approach, whereby we directly cluster the cell-type frequency table generated before feeding into the LDA model. We were able to identify similar communities (Supplementary Fig. S2E), thereby validating the communities that we describe using an alternative approach. However, we believe the LDA model was more robust to noise compared with directly clustering the cell-type frequency table. The method is implemented under the `spatial_volume` function, and the LDA approach is implemented under the `spatial_volume` function in `scimap` python package.

**Statistical Tests**

All statistical tests to infer P value for significant differences (P < 0.05) in mean were performed using Python’s `scipy` implementation of the t test.

**Data and Software Availability**

mrSEQ data are available via the Gene Expression Omnibus (GEO; GSE171888). All full-resolution images derived from image data (e.g., segmentation masks) and all cell count tables are available via the NCI Human Tumor Atlas Network data portal ([https://data.humanumoratlas.org/](https://data.humanumoratlas.org/)). These data are also available via the Harvard Tissue Atlas Portal ([https://www.tissue-atlas.org/atlas-datasets/nirmal-maliga-vallius-2021/](https://www.tissue-atlas.org/atlas-datasets/nirmal-maliga-vallius-2021/)). Note that individual files are ~100 GB in size, so an AWS S3-compatible download tool should be used. Several of the figure panels in this paper are available with text and audio narration for anonymous online browsing using MINERVA software ([https://pypi.org/project/minerva.software/](https://pypi.org/project/minerva.software/)), which supports zoom, pan, and selection actions without requiring the installation of software.

**Authors’ Disclosures**

A.J. Nirmal reports grants from the NIH during the conduct of the study. T. Vallius reports grants from the Finnish Medical Foundation and the Relander Foundation during the conduct of the study. S. Santagata reports grants from the NIH/NCI and the Ludwig Center at Harvard during the conduct of the study, as well as personal fees from RareCyte, Inc. outside the submitted work. P.K. Sorger reports personal fees, nonfinancial support, and other support from RareCyte Inc., personal fees from NanoString, and other support from Glencoe Software during the conduct of the study, as well as personal fees from Montal Health, Merck, and Applied Biomath outside the submitted work. No disclosures were reported by the other authors.

**Authors’ Contributions**

A.J. Nirmal: Conceptualization, data curation, software, formal analysis, validation, investigation, visualization, methodology, writing—original draft, writing—review and editing. Z. Maliga: Conceptualization, data curation, validation, investigation, visualization, writing—original draft, writing—review and editing. T. Vallius: Conceptualization, validation, investigation, visualization, writing—original draft, writing—review and editing. A.A. Chen: Project administration, investigation, writing—review and editing. C.A. Jacobson: Resources, data curation, writing—review and editing. C. Yapp: Data curation, software, investigation, visualization, writing—review and editing. A.A. Chen: Resources, data curation, software, investigation, visualization, writing—review and editing. R.J. Pelletier: Resources, data curation, investigation, writing—review and editing. C. Yapp: Data curation, software, investigation, visualization, writing—review and editing. A.A. Chen: Resources, data curation, software, investigation, visualization, writing—review and editing. G.F. Murphy: Conceptualization, resources, supervision, funding acquisition, writing—review and editing. S. Santagata: Conceptualization, resources, supervision, funding acquisition, writing—review and editing. P.K. Sorger: Conceptualization, resources, supervision, funding acquisition, writing—original draft, writing—review and editing.

**Acknowledgments**

The authors are investigators in the NCI-funded Human Tumor Atlas Network. They thank David Liu, Genevieve Boland, Jeremy Muhlich, David Weinstock, Robert Krueger, Jared Jessup, and Simon Warchol for their help in multiple stages of this project; they are deeply grateful to Keith Ligon for hosting their clinical research coordinator. This work was supported by NIH grants U2C-CA233262 (P.K. Sorger and S. Santagata), K99-CA256497 (A.J. Nirmal), R50-CA252138 (Z. Maliga), and to the Ludwig Center at Harvard (P.K. Sorger and S. Santagata), and by grants from the Finnish Medical Foundation and the Relander Foundation (T. Vallius). Access to the GeoMx mrSEQ platform was kindly provided by NanoString as part of their Technology Access Program. All Human Tumor Atlas Network consortium members are listed at humantumoratlas.org. The authors thank Dana-Farber/Harvard Cancer Center for the use of the Specialized Histopathology Core, which provided histopathology services supported by P30-CA06516. Imaging at the Harvard Medical School Neurobiology Imaging Facility (of H&E specimens) was supported by the National Institute on Neurological Disorders and Stroke (NINDS) Core Center Grant P30-NS072030.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 13, 2021; revised February 5, 2022; accepted April 1, 2022; published first April 11, 2022.

**REFERENCES**

Atlas of Primary Melanoma


