

Spatially Resolved Transcriptomics Enables Dissection of Genetic Heterogeneity in Stage III Cutaneous Malignant Melanoma

Kim Thrane¹, Hanna Eriksson^{2,3}, Jonas Maaskola¹, Johan Hansson^{2,3}, and Joakim Lundeberg¹



Abstract

Cutaneous malignant melanoma (melanoma) is characterized by a high mutational load, extensive intertumoral and intratumoral genetic heterogeneity, and complex tumor microenvironment (TME) interactions. Further insights into the mechanisms underlying melanoma are crucial for understanding tumor progression and responses to treatment. Here we adapted the technology of spatial transcriptomics (ST) to melanoma lymph node biopsies and successfully sequenced the transcriptomes of over 2,200 tissue domains. Deconvolution combined with traditional approaches for dimensional reduction of transcriptome-wide data enabled us to both visualize the transcriptional landscape within the tissue and identify gene expression profiles linked to specific histologic entities. Our unsupervised analysis revealed a complex spatial intratumoral composition of melanoma metastases that was not evident through morphologic annotation. Each biopsy showed distinct gene expression profiles and included examples of the coexistence of multiple melanoma signatures

within a single tumor region as well as shared profiles for lymphoid tissue characterized according to their spatial location and gene expression profiles. The lymphoid area in close proximity to the tumor region displayed a specific expression pattern, which may reflect the TME, a key component to fully understanding tumor progression. In conclusion, using the ST technology to generate gene expression profiles reveals a detailed landscape of melanoma metastases. This should inspire researchers to integrate spatial information into analyses aiming to identify the factors underlying tumor progression and therapy outcome.

Significance: Applying ST technology to gene expression profiling in melanoma lymph node metastases reveals a complex transcriptional landscape in a spatial context, which is essential for understanding the multiple components of tumor progression and therapy outcome. *Cancer Res*; 78(20); 5970–9. ©2018 AACR.

Introduction

Molecular heterogeneity both within and between tumors in the same patient has been recognized in most cancers, and is especially prevalent in cutaneous malignant melanoma (melanoma; refs. 1–5). The combination of an expansion of tumor subclones with varying genetic and genomic alterations and the interaction between the tumor cells and the tumor microenvironment (TME), for example, malignant, stromal, and immune cells, during cancer progression presents an intricate situation that influences disease development (6–8). Previous large-scale sequencing studies have concluded that the extensive inter- and intratumoral genetic heterogeneity observed in melanoma reflects

one of the highest mutational loads in all cancers (4, 5, 8–13). Furthermore, melanoma is a highly immunogenic cancer that induces immune responses associated with immune cell infiltration both in localized and metastatic disease. Clinically, the genetic diversity of the tumor in relation to that of the TME may be an important factor in explaining tumor progression, response to treatments, emergence of therapy resistance, and outcome differences in melanoma. The heterogeneity of the disease is clearly demonstrated by patients diagnosed with stage III melanoma, which is a group characterized by a high risk of relapse and 5-year relative survival rates between 40% and 70% (14).

Previous research has shown that chemotherapeutic approaches for treating disseminated melanoma are inefficient and have no noticeable impact on overall survival (15). Therapies that both target *BRAFV600* mutated melanomas and utilize immune checkpoint inhibitors have changed the outlook for patients with melanoma and also offer potential therapy options in the adjuvant setting of stage III melanoma (16). Still, remissions do not frequently last long.

Thus, further research is needed to elucidate the molecular mechanisms underlying melanoma pathogenesis, to improve diagnostics and to identify prognostic as well as treatment-predictive biomarkers.

The inter- and intratumoral molecular heterogeneity in stage III melanoma has already been analyzed with an array of analytical approaches, including gene expression analyses, IHC, and proteomic approaches (10, 11, 13, 17–21). Tirosch and colleagues (13)

¹Department of Gene Technology, KTH Royal Institute of Technology, SciLifeLab, Stockholm, Sweden. ²Department of Oncology-Pathology, Karolinska Institutet, SE-17176 Stockholm, Sweden. ³Department of Oncology, Karolinska University Hospital, SE-17176 Stockholm, Sweden.

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

K. Thrane and H. Eriksson contributed equally to this article.

Corresponding Author: Joakim Lundeberg, Royal Institute of Technology, Science for Life Laboratory, Stockholm 171 76, Sweden. Phone: 46704582346; E-mail: joakim.lundeberg@scilifelab.se

doi: 10.1158/0008-5472.CAN-18-0747

©2018 American Association for Cancer Research.

analyzed the transcriptional heterogeneity in melanoma, including lymph node metastases, with single-cell RNA sequencing (scRNA-seq) and concluded that the intra- and interindividual spatial, functional, and genomic heterogeneity in melanoma and the associated tumor components shapes the TME. Although scRNA-seq is a powerful tool for addressing transcriptional heterogeneity, knowledge about the spatial origins of single cells is vital when studying the TME.

We aimed to optimize and apply spatial transcriptomics (ST) technology (22) for the *in situ* and quantitative detection of gene expression in stage III melanoma lymph node metastases. For this purpose, we analyzed the transcriptomes of ~2,200 tissue domains in fresh frozen tumor sections. By retaining the positional information of the tissue domains, we could superimpose the transcriptomic data onto histological tissue images and visualize gene expression throughout the tumor.

Materials and Methods

Melanoma samples and RNA quality control

Lymph node metastases from four patients diagnosed with stage III melanoma were included in this study. The metastases have previously been described as well as molecularly annotated regarding *BRAF*, *NRAS* (23), and *CDKN2A* (24) status. The fresh frozen tumor tissue was stored at -70°C until analysis. The patients were previously categorized as short-term survivors (≤ 13 months) or long-term survivors (≥ 60 months) following lymphadenectomy (24, 25). The staging was performed according to the American Joint Committee on Cancer system 7 (26) at the time of diagnosis. This study was conducted in accordance with the Declaration of Helsinki with written informed consent from all patients. The project was approved by the Regional Ethics Committee at the Karolinska Institutet, Stockholm, Sweden.

The tumor samples were embedded in OCT (#4532, Sakura, Alphen aan den Rijn, the Netherlands) before cryosectioning for RNA extraction. A total of four sections per metastasis, each with a width of $10\ \mu\text{m}$, were placed in Lysing Matrix D (#116913050; MP Biomedicals) and lysed using FastPrep (MP Biomedicals). Total RNA was extracted using the RNeasy Plus Mini Kit (#74134; Qiagen) according to the manufacturer's protocol. RIN values were determined with a 2100 Bioanalyzer, RNA 6000 Pico Kit (Agilent) according to the manufacturer's protocol.

Hematoxylin and eosin-stained (H&E-stained) images of each of the four tumor samples were manually annotated by a trained pathologist to identify melanoma, stromal and lymphoid tissue. The lymphoid tissue was a separate region in two of the samples. The annotation was performed on one of the two consecutive sections used for the spatial arrays.

Preparation of arrays

Two types of arrays, both prepared on CodeLink-activated microscope glass slides (SurModics) as described by Ståhl and colleagues (22), were used in this study. For the optimization assay, reverse transcription (RT)-primers containing a poly-T20VN capture region were immobilized uniformly over the array surface, with six identical surfaces generated per glass slide according to the manufacturer's instructions.

Surface RT-primer for optimization assays:

[AmC6]UUUUUUGACTCGTAATACGACTCACTATAGGGACACGACGCTCTCCGATCTNNNNNNNNNTTTTTTTTTTTTTTTTTTTVN

The arrays for spatial experiments (referred to as spatial arrays) were printed as 1,007 spots of RT-primers containing a poly-T20VN capture region, a spatial barcode and a unique molecular identifier (UMI). The spots, with $200\ \mu\text{m}$ center-to-center spacing, a diameter of $100\ \mu\text{m}$, and approximately 200 million RT-primers per spot, were printed according to the manufacturer's instructions (ArrayJet Ltd.). Six identical array surfaces, with dimensions of $6,200 \times 6,600\ \mu\text{m}$, were generated per glass slide.

Surface RT-primer for spatial arrays:

[AmC6]UUUUUUGACTCGTAATACGACTCACTATAGGGACACGACGCTCTCCGATCT]18mer_Spatial_Barcode_1to1007]WSNNWSNNVTTTTTTTTTTTTTTTTTTTTVVN

Surface frame oligonucleotide:

[AmC6]AAATTCGCTCTGCTATCGCGCTTCTGTACC

Preparation of tumor samples for the ST analysis

Fresh frozen tumor biopsies were embedded in OCT, cryosectioned at 8 and $14\ \mu\text{m}$, and placed onto the array surfaces, followed by incubation at 37°C for 1 min. A new glass slide was used for each of the four tumor samples, however only two array surfaces out of six were used to generate sequencing libraries for the spatial experiments. A fixation solution, which was prepared by diluting 36.5%–38.0% formaldehyde (#F8775; Sigma-Aldrich) 1:10 in 1xPBS (#09-9400; Medicago), was added on top of the tissue sections, incubated for 10 minutes, and washed off by dipping the glass slides in $1 \times$ PBS. The fixated tissue sections were covered with propan-2-ol (#A461-1; Fisher Scientific). Following evaporation for 45 seconds, Mayer's hematoxylin (#S3309; Dako) was added and, after a seven-minute incubation, the glass slides were washed in RNase and DNase free MQ water. The glass slides were briefly dried and bluing buffer (#CS702; Dako) was added and washed off with RNase and DNase free MQ water after 2 minutes. For eosin staining, the glass slides were immersed in a solution of Eosin (#HT110216) diluted 1:20 in Tris-base [0.45M Tris, 0.5M acetic acid, pH 6.0]. The glass slides were then washed in RNase and DNase free MQ water and incubated at 37°C until dry.

Before imaging, the glass slides were mounted with 85% glycerol (#104094; Merck Millipore) and covered with coverslips (#BB014060A1; Menzel-Gläser). Bright field imaging was performed at 20x resolution with the Metafer Slide Scanning platform (MetaSystems). VSlide and VSViewer software (MetaSystems) was used to stitch the raw images and extract .jpeg files.

The coverslips and glycerol were removed by submerging the glass slides in RNase and DNase free MQ water until the coverslips came off, after which the slides were dipped in 70% ethanol.

After drying at 37°C , the glass slides were placed in ArrayIT masks to enable separation of reagents for the six arrays per glass slide.

The glass slides were kept in the masks for the following steps, and removed for the wash preceding tissue removal.

The mRNA capture procedure

For each array surface, $70\ \mu\text{L}$ containing $1 \times$ Exonuclease I Reaction buffer [#B0293S; New England Biolabs (NEB); Ipswich] and $0.19\ \mu\text{g}/\mu\text{L}$ BSA (#B9000S; NEB) was added, followed by an incubation time of 30 minutes at 37°C . After incubation, each array was washed by pipetting $100\ \mu\text{L}$ $0.1 \times$ SSC (#S6639; Sigma-Aldrich), diluted in MQ RNase and DNase free water. For the mRNA optimization assay, $70\ \mu\text{L}$ of 0.1% pepsin (#P7000-25G;

Sigma-Aldrich) dissolved in 0.1M HCl (#318965- 1000ML; Sigma-Aldrich) was added to six optimization surfaces per sample. Incubations were performed at 37°C in duplicates for 10, 12, and 14 minutes. The same reagents and volumes were used in the spatial arrays, but incubation times were based on results from the optimization assay. After incubation, reagents were removed by pipetting and the array surface was washed with 100 μ L 0.1 \times SSC as described above.

Reverse transcription, tissue removal, and material cleavage

Two different RT mixtures for were used for the optimization- (in which fluorescent cDNA was to be generated) and spatial arrays (intended for library preparation and sequencing). Both incubations were conducted with a total volume of 70 μ L per array surface, overnight (ON) at 42°C for approximately 15 hours. The reaction mixture for the optimization assay contained 1 \times first strand buffer (#18080-044; Invitrogen), 5 mmol/L DTT (#18080-044; Invitrogen), 0.19 μ g/ μ L BSA, 50 ng/ μ L Actinomycin D (#A1410-2MG; Sigma-Aldrich), 1% DMSO (#472301-500ML; Sigma-Aldrich), 20 U/ μ L Superscript III (#18080-04; Invitrogen), 2 U/ μ L RNaseOUT (#10777-019; Invitrogen), 500 μ mol/L each of dATP, dGTP, and dTTP (12.5 μ mol/L of dCTP, and 25 μ mol/L of Cyanine 3-dCTP (#NEL576001EA; PerkinElmer). The same reagents were used for the spatial arrays, except that each dNTP (#R0192; Fisher Scientific) was present at the same concentration of 500 μ mol/L and that Cyanine 3-dCTP was excluded.

Residues of melanoma tissue were degraded and removed in two steps. Each array surface was first incubated with 70 μ L of β -mercaptoethanol (#444203; Calbiochem) in RLT buffer (#79216; Qiagen) at a 1:100 ratio, at 56°C for 1.5 hours, with continuous shaking, which was followed by washing with 100 μ L 0.1 \times SSC as described above. In the second step, each array surface was incubated with 70 μ L of Proteinase K (#19131; Qiagen) in PKD buffer (#1034963; Qiagen, pH 7.5) at a 1:7 ratio, at 56°C for 1 hour, with interval shaking. The glass slides were released from the masks and washed in three steps by submerging them in beakers containing: 2 \times SSC with 0.1% SDS (#71736-100ML; Sigma-Aldrich), 0.2 \times SSC, and 0.1 \times SSC, respectively. All washing steps were performed in a thermomixer with continuous shaking; the first at 50°C for 10 minutes, and the remaining two at room temperature for 1 minute each.

All of the glass slides were spin-dried, after which, the optimization arrays were subjected to fluorescent scanning whereas the spatial arrays were put back into the masks for cleavage of the mRNA:cDNA material.

Spatial library preparation and sequencing

The barcoded mRNA/cDNA material was cleaved off from the arrays and sequencing libraries were prepared in solution. Briefly, second-strand cDNA was synthesized, followed by *in vitro* transcription and adapter ligation. Sequencing handles and indexes were added in a PCR, and the finished libraries were sequenced on the Illumina NextSeq platform with paired-end sequencing, 31 bases from read 1 and 121 from read 2. The protocol is described in detail in Ståhl and colleagues (22)

Image alignment

Spot visualization and image alignment were both performed as previously described by Ståhl and colleagues (22).

Read alignment and generation of count data

Sequencing data were processed using the ST pipeline, with standard settings, as described by Fernández and colleagues (27). In short, read 2 was mapped against the human genome (GRCh38) and read 1 was used for UMI filtering and spatial information. The ST pipeline generated one matrix (.tsv-file) per sample containing gene counts per spatial barcode (i.e., tissue domain) as the output. Data points for complexity curves were generated using a built-in function of the ST pipeline (27). Annotation reads were subsampled before proceeding to UMI-filtering for each subsample, resulting in a computed number of unique molecules for a stepwise increasing number of annotated reads. The resulting data points were plotted in RStudio, version 0.99.903, The R Project for Statistical Computing, Vienna, Austria.

Correlation of biological replicates

The count data were normalized by log2-transformation of counts per million (CPM) +1 pseudocount, and scatterplots were generated for each pair of consecutive section using the ggplot2 package (28) in RStudio (version 1.1.453). The built-in stats package was used to compute Pearson correlations.

Principal component analysis and hierarchical clustering of "bulk" data

Matrix data from the ST pipeline were imported into RStudio (version 0.99.903) and count data from all spatial IDs (tissue domains) within a dataset were merged. Read counts were normalized with DESeq2 (29), a principal component analysis (PCA) plot was generated with ggplot2 (28), and hierarchical clustering was performed with pheatmap (30).

Factor analysis

Count data generated from the ST pipeline were analyzed with a factor analysis method previously described by Berglund, Maaskola, Schultz and colleagues (31). Four out of 20 generated factors were named and classified according the manual histopathological annotation.

PCA of individual tissue sections

The matrix with count data generated from the ST pipeline was imported into RStudio (version 0.99.903) and filtered by removing all tissue domains with less than 500 genes. The data were normalized using the scran package (32). Variance calculations and PCA were conducted using built-in functions.

Results

Overview of subjects

This study included samples from lymph node metastases from four patients with stage IIIc melanoma (Table 1). All but one sample were *BRAF*v600 positive and one patient was a long-term survivor according to the outcome classification by Johansson and colleagues (>60 months; Table 1; ref. 25). The RNA quality control prior to the ST analysis showed RNA integrity numbers (RIN) between 6.2 and 7.4 (Table 1). A manual histologic annotation, performed by a trained pathologist, identified areas of tumor cells (i.e., melanoma), stroma, and lymphoid within the metastases (Supplementary Fig. S1). The annotations were used for comparison with results from the ST gene expression analyses.

Table 1. Clinical data^a

Tumor sample	Gender	BRAF status	NRAS status	CDKN2A status	Survival ^b	RIN value
1	Male	mut	wt	hd	10+ years	6.2
2	Female	wt	wt	hd	401 days	7.1
3	Male	mut	wt	wt	208 days	7.4
4	Male	mut	wt	wt	305 days	6.9

NOTE: Overview of clinical and genetic data, including RNA quality (RIN value), for each of the tumor samples from patients with stage IIIc melanoma.

Abbreviations: hd, hemizygous deletion; mut, mutated; wt, wild type.

^aAll data except for the RIN value was generated prior to this study.

^bOverall survival from the diagnosis of stage III melanoma.

Procedure for spatial analysis of lymph node metastases

The two-dimensional gene expression analysis provided by the ST procedure was based on fresh frozen tissue sections from the four tumor biopsies (denoted 1–4; Fig. 1). Briefly, the sections were placed on glass slides containing RT-primers arrayed as spots that corresponded to tissue domains comprising between 5 and 40 cells. The histologic sections were stained with H&E and imaged before the cells were permeabilized and the mRNA was captured. The RT-primers at each spot had a unique spatial ID barcode, which was sequenced along with the transcript to enable trace-back to a specific tissue domain. After the release of cDNA: mRNA, library construction, and sequencing, the transcriptome-wide data were decomposed into factors (31) and the gene

expression profiles were visualized using t-distributed stochastic neighbor embedding (t-SNE; Fig. 1; ref. 33).

An optimization assay employing arrays coated with polyA-capturing RT-primers followed by RT with fluorescently labeled nucleotides was performed to determine the optimal time for melanoma tissue permeabilization (Supplementary Fig. S2A and S2B).

Generating spatially resolved transcriptomic data from melanoma lymph node metastases

Barcoded arrays were used to generate spatially resolved transcriptomic sequence libraries. An average of 286 tissue domains per section were analyzed, and a total of over 2,200 domains were

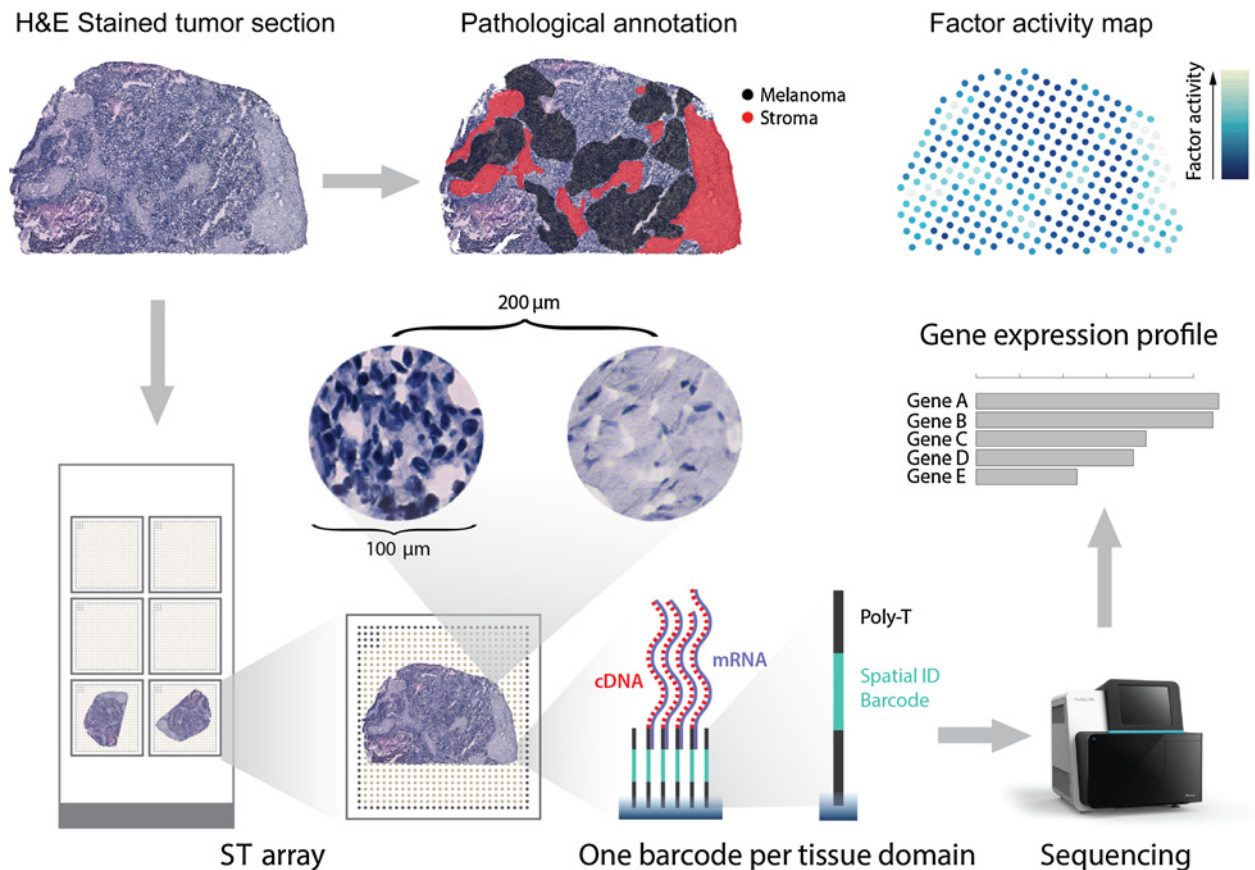


Figure 1. Spatial transcriptomics overview. Schematic overview of the STs technology, along with the downstream analysis. The barcoded microarrays contained 1,007 printed spots of RT-primers with unique barcode sequences. Each spot had a diameter of 100 μm, thus corresponding to a tissue domain. The center-to-center distance was 200 μm.

Downloaded from <http://aacrjournals.org/cancerres/article-pdf/78/20/5970/272986/5970.pdf> by guest on 27 August 2022

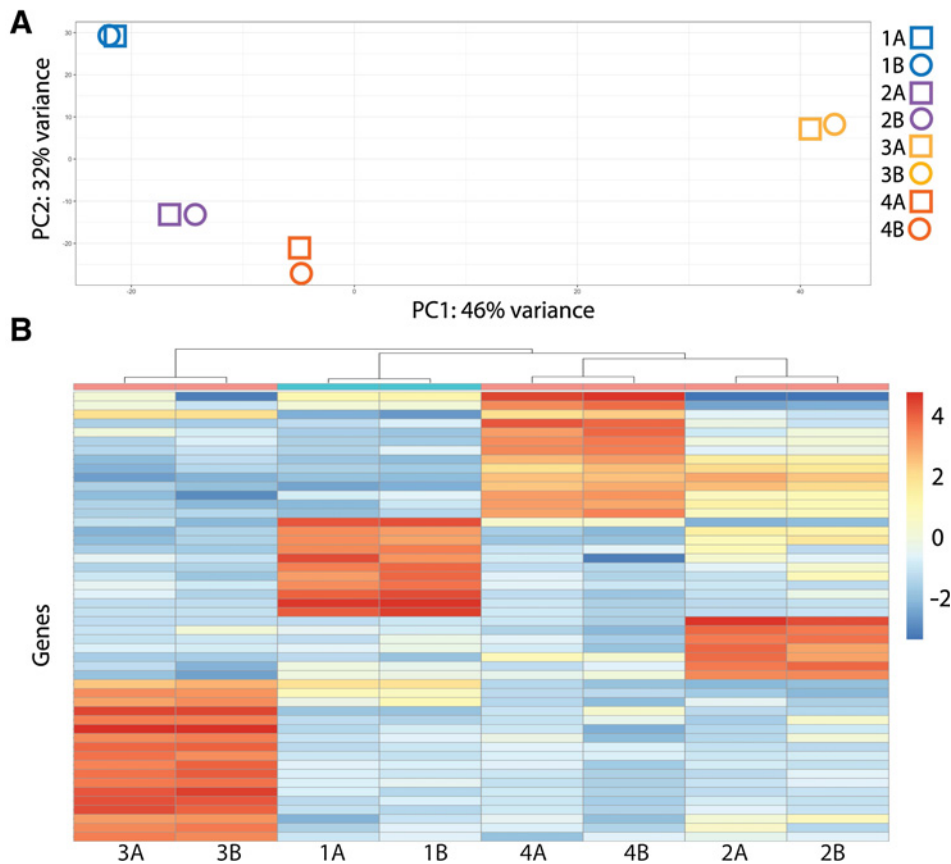


Figure 2. Analysis of *in silico* bulk data. Data from all spots (tissue domains) within a section was merged to mimic bulk RNA-seq data. **A**, PCA of all sections (variance per component is shown in Supplementary Fig. S4). **B**, Hierarchical clustering of the 50 most variable genes.

investigated in this study. After data processing and UMI filtering, each tissue section was represented by between 1.4 and 4.2 million unique molecules (transcripts), or an average of nearly 3,000 transcripts per tissue domain (Supplementary Table S3A). All libraries were sequenced at a depth close to saturation (Supplementary Fig. S3B). Scatterplots for each pair of biological replicates and calculated Pearson correlations are included in Supplementary Fig. S3C.

We first explored the transcriptomic data as an *in silico* bulk sample and thereby mimicking bulk RNA-seq data, by merging gene expression data from all of the tissue domains within a tissue section. To achieve an overview of the metastases, the eight datasets (two sections from each metastasis) from the four different biopsies were analyzed with PCA (Fig. 2A; Supplementary Fig. S4). The four metastases were clearly separated in the PCA plot whereas each pair of tissue sections from the same tumor showed low variance. To further investigate the merged datasets, hierarchical clustering was performed for the 50 most variable genes (Fig. 2B).

Visualizing gene expression in 2,200 tissue domains by factor analysis and t-SNE

The spatial information provided by the barcoded arrays was then included in the analysis to visualize the tissue context of gene expression and increase the granularity of the analysis. The gene expression data of all 2,200 tissue domains from the four pairs of melanoma sections were subjected to a factor analysis method that had earlier been used to explore tumor heterogeneity in prostate cancer (31). In short, unsupervised dimensionality

reduction was used to decompose the data into 20 factors describing gene expression profiles, which were then organized as spatial factor activity maps for the tissue domains. The principles of factor analysis as well as factor contributions are illustrated in Supplementary Fig. S5A to S5C. In several cases, a factor's activity map showed a clear link to the histologic patterns noted in the metastases. By using the manual histopathologic annotation as a reference, these factors were classified and named accordingly. Gene expression profiles showing the relative contribution of genes to a specific factor are listed in Supplementary Fig. S6.

To achieve an overview of all gene expression profiles across the tissue sections, the 20 factors were further decomposed into a color scale, and each tissue domain assigned a color based on factor activity (Supplementary Fig. S7; ref. 31). Areas of a similar color thus share similar gene expression profiles, and each pair of tumor sections shared certain regions having the same color.

Most of the areas (colors) were shared between tumor section pairs across the four studied metastases.

In general, this analysis demonstrated extensive heterogeneity between the four different metastases, which confirmed the initial PCA and hierarchical clustering analyses described above.

The tumor sections were examined in further detail to explore gene expression profiles in relation to the manually annotated areas of melanoma, stroma, and lymphoid tissue. All of these regions, especially those associated with melanoma, showed highly diverse gene expression across biopsies. For example, the t-SNE results for biopsy 1 demonstrated a distinct delineation between the melanoma, stromal and lymphoid tissue regions, with the tumor region represented by a single color (Fig. 3A and

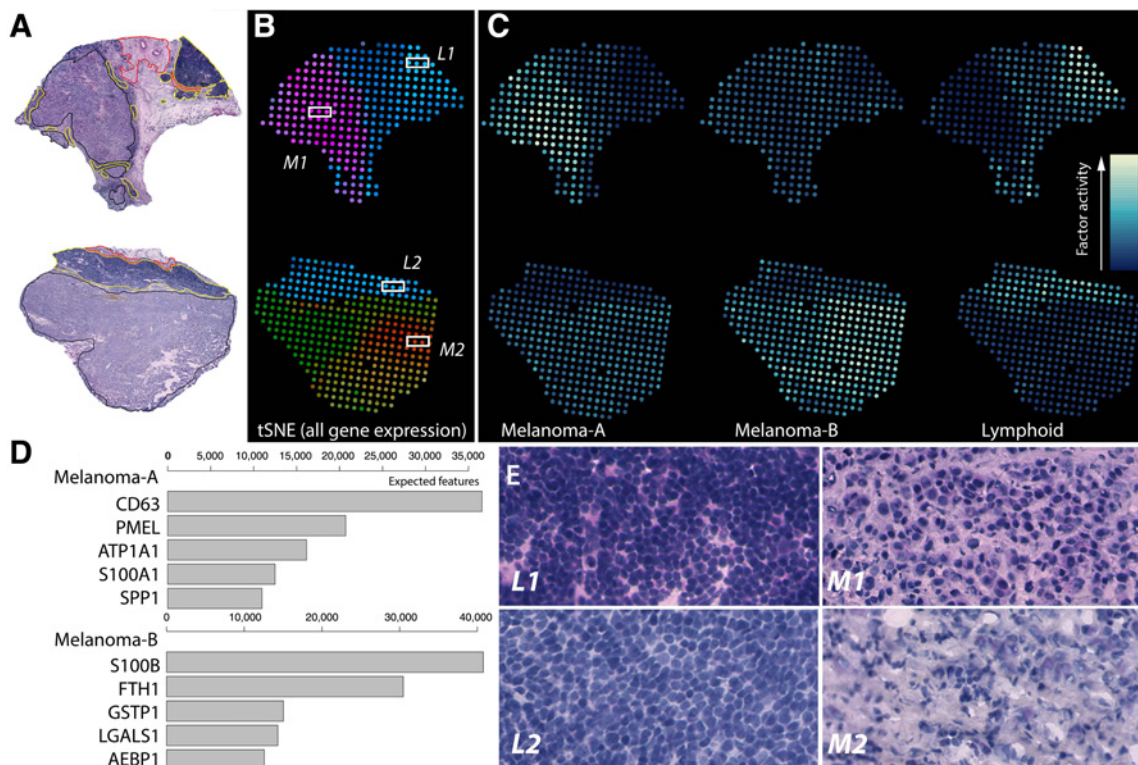


Figure 3.

Tumor morphology and results from factor analysis. **A**, H&E-stained tissue images of biopsy 1 (top) and 2 (bottom) with marked melanoma (black), stromal (red), and lymphoid (yellow) tissue regions. **B**, t-SNE color visualization of gene expression, summarizing gene expression profiles from the factor analysis. **C**, Spatial activity maps of factors Melanoma-A, Melanoma-B, and Lymphoid. **D**, A representation of the most abundant genes in factors Melanoma-A and Melanoma-B. **E**, Magnified images of areas manually annotated as lymphoid tissue and melanoma, respectively.

B). In contrast, tumor cells in biopsy 2 were characterized by extensive heterogeneity within the melanoma area, which indicates the presence of several distinct expression profiles (Fig. 3A and B). Notably, the "monoclonal" biopsy 1 was from a patient classified as a long-term survivor whereas the more heterogeneous biopsy 2 was from a patient facing short-term survival.

Two factors showed recurrent activity across biopsies in areas manually annotated as melanoma (Fig. 3C). One of the factors, referred to as Melanoma-A (A representing one of two melanoma tumor profiles identified) mainly included *CD63* and *PMEL* overexpression whereas the other, Melanoma-B, showed heightened expression of *S100B* and *FTH1*, among others. The top five genes of these factors, according to detected mRNA transcript amounts, are listed in Fig. 3D, and the complete gene expression profiles are shown in Supplementary Fig. S6. Melanoma-A was prevalent in regions annotated as melanoma in biopsy 1, but not as noticeable in biopsies 2 and 4 (Supplementary Figs. S1 and S8). Melanoma-A and Melanoma-B had an overlapping pattern in biopsy 2 (Supplementary Fig. S8). This overlapping profile across biopsies was not visible in the color scale generated by t-SNE of all factors, which emphasizes the increased granularity that examining the factors individually can provide (Supplementary Fig. S8). Melanoma-B was clearly present in biopsy 4, but was not identified for biopsy 1. The two factors (Melanoma-A and -B) coexisted in the tumor areas of biopsies 2 and 4 (Supplementary Fig. S8).

The lymphoid factor, corresponding to lymphoid tissue (Fig. 3C and E), was evident in biopsies 1 and 2 (Supplementary

Fig. S8), and was also revealed in the t-SNE analysis (Fig. 3B; Supplementary Fig. S7). The other two biopsies (3 and 4) did not contain any lymphoid tissue based on the manual and computational annotations.

A visualization of gene expression in the TME

According to the manual annotation, biopsy 1 had regions of lymphoid tissue that were both in close proximity to, and at a distance from, the tumor area; therefore, this biopsy was chosen for further studies concerning how the TME affects gene expression. The initial factor analysis identified a transition area (Supplementary Figs. S6 and S8), which showed high expression of *FTL*, *B2M*, *APOE*, and HLA-associated genes (*HLA A-C*).

We performed a PCA using the gene expression data obtained from all tissue domains (284 spots) of biopsy 1 to provide a more detailed analysis of spatial gene expression (Fig. 4A and B). The result showed four clusters that strongly reflected the histological areas annotated by the pathologist. The tissue domains within the melanoma area shared common gene expression profiles that differed from the expression profiles dominating in distant lymphoid regions (Fig. 4B and C). This result confirms the previous factor analysis.

The tissue domains corresponding to lymphoid regions in close proximity to the melanoma area were grouped in a separate PCA cluster than lymphoid regions further away from the tumor (Fig. 4B and C). This clearly demonstrates that multiple gene

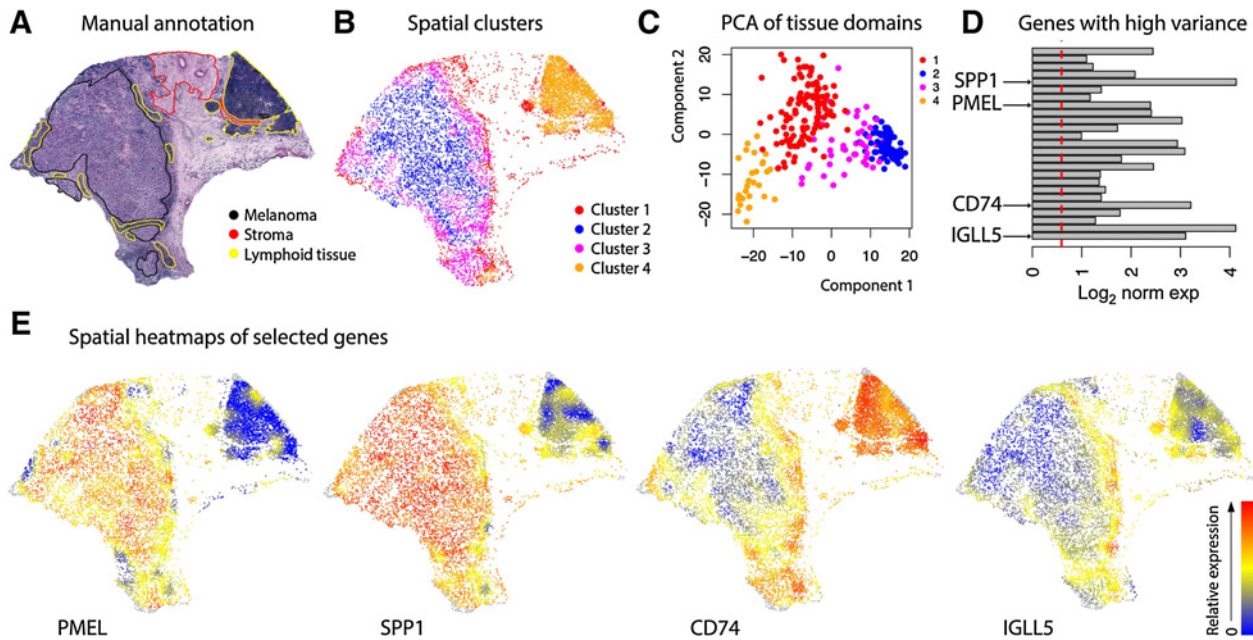


Figure 4. PCA on an individual section and spatial heatmaps. **A**, H&E-stained tissue image with pathologic annotation. **B** and **C**, PCA conducted on tissue domains resulted in four clusters. Using the pathologic annotation as reference, cluster 1 corresponds to stromal or lymphoid tissue surrounding the melanoma areas, whereas cluster 2 clearly represents the melanoma region. Cluster 3 includes tissue domains in the border area between the lymphoid and tumor tissues, likely containing a mixture of cells from these two tissue types. Cluster 4 represents lymphoid tissue physically separated from the cancer region—a separation also revealed by the PCA plot. **D**, Mean expression of the 25 genes with highest variance across tissue domains. Dashed red line shows the average value across all features. **E**, Four genes that showed both high variance and mean expression were selected for visualization as spatial heatmaps.

expression programs, which are possibly influenced by distance from tumor cells, exist within lymphoid tissue. Four of the major genes contributing to the cluster separations (Fig. 4D) are displayed in Fig. 4E as spatial heat maps. *PMEL* is clearly exclusive to melanoma areas, whereas *SPP1* was detected in a slightly larger area that also covers the lymphoid tissue in close proximity to the tumor cells. The factor analysis described above included both of these genes in Melanoma-A, a factor that was clearly prevalent in the tumor area of biopsy 1. In areas manually annotated as lymphoid tissue, *CD74* was primarily detected at a distance from the melanoma area whereas *IGLL5* was more prevalent closer to the tumor cells.

Discussion

In this exploratory study of the applications of ST technology to melanoma research, we demonstrate intra- and intertumoral gene expression heterogeneity in lymph node metastases using *in situ* transcriptome-wide methodology. The analysis of spatial RNA sequencing information from over 2,200 tissue domains, when compared with single bulk analysis of tumors, provides numerous avenues for analysis using unsupervised machine learning approaches and is particularly robust in revealing spatial trends in gene expression. Overall, we could confirm matches between histological entities and gene expression profiles, but the tested approach revealed a significantly more complex transcriptional landscape than what can be observed from a histopathologic analysis.

The routine investigation of melanoma and metastasis biopsies relies on a pathologist’s manual evaluation in combination with

molecular analyses and IHC, which are especially important before starting targeted therapies or administering checkpoint inhibitors. To date, there are no tools in clinical use that allow healthcare professionals to analyze or follow genetically heterogeneous tumor clones in relation to melanoma progression or treatment outcome. By using unsupervised factor analysis, we not only confirmed both malignant and nonmalignant tissue areas, but also identified various expression “clones” within the metastases and delineated location-dependent expression patterns in lymphoid tissue.

A number of technologies have been introduced in the last years to spatially detect RNA and proteins in histologic sections (34–37), and multiplexed immunofluorescence (IF) has improved IHC by enabling simultaneous detection of multiple antigen markers (34). The CO-Detection by indexing (CODEX) approach utilizes antibodies tagged with oligonucleotide indexes that are decoded by iterative fluorescence labeling and scanning. In a preprint by Goltsev and colleagues 66 antigens were detected in fresh frozen tissue using the CODEX technology (35). Imaging Mass Cytometry allows for detection of up to 40 target proteins at a resolution of 1 μm, which can be increased at the expense of reduced sensitivity (36). Compared with ST, IF approaches typically provide a higher degree of spatial resolution, but with a limited number of antigen markers, and thereby enable a powerful complement for in-depth studies. The digital spatial profiling (DSP) by Nanostring was recently used in a study by Ziai and colleagues to spatially detect 29 proteins and 39 RNA species in regions of interest (ROI) from lung cancer tissue (37). In addition to spatial approaches, RNA-seq is comparable to ST by also providing

unbiased, transcriptome-wide data. Ståhl and colleagues (22) evaluated the performance of ST by comparing *in silico* RNA-seq libraries from spatial data with libraries prepared in solution from extracted and fragmented total RNA from mouse olfactory bulb. They found 0.89% of the genes to be shared, whereas 0.053%, and 0.052% were exclusive to the RNA-seq, and the spatial library, respectively. In total nearly 18,000 genes were detected and the two libraries showed a Pearson correlation of $r = 0.94$.

Although the PCA performed on *in silico* bulk samples indicated low variance between the two consecutive sections of each tumor sample, the spatial t-SNE (but not PCA) visualization suggested heterogeneity, particularly in sample 4. This fact did not appear to affect the results of the factor analysis, as we still found the same factors being active in the two replicates, indicating quite subtle differences between consecutive sections. The scatterplot of the biological replicates also showed an even representation of highly and lowly expressed genes, with a Pearson correlation of 0.86. The reproducibility of ST has previously been evaluated on more homogenous nontumor tissue (mouse olfactory bulb), showing a Pearson correlation $r = 0.97$ between two consecutive sections (22).

The signatures of factors Melanoma-A and -B included both cancer- and immune-associated genes that are most likely active in the tumor area. For example, *CD63*, which showed increased expression in Melanoma-A, may act either as a tumor suppressor or participate in cell signaling and anoikis resistance through the PI3K signaling pathway independently of Akt phosphorylation (38). Interestingly, Biopsy 1 from a long-term survivor demonstrated a homogenous expression pattern in the melanoma cell area as compared to the corresponding region in Biopsy 2 from a short-term survivor. This could represent a picture of tumor clonal heterogeneity but it may also be related to phenotypic switching, which occurs in melanoma through epithelial–mesenchymal transition (EMT; ref. 39). Many pathways play a role in EMT, including RAS/RAF/MEK/ERK, PI3K/AKT/mTOR, Wnt/ β -catenin as well as downstream effectors of these pathways, which induce EMT expression transcription factors. For example, among the gene expression profiles in our report showing a large contribution of genes related to a specific factor, RTK-genes, such as *ERBB3*, were represented. Phenotypic switching is critical in melanoma through melanocyte lineage differentiation during tumor progression and in response to treatment, for example, BRAF- and MEK-inhibiting therapy (40, 41). The ST technology may be applied to guide microdissection of defined sections to address this in future.

The gene expression profile of the lymphoid factor involved HLA genes and *CD74*, the latter of which has been shown to be associated with favorable recurrence-free survival and overall survival in stage III melanoma (42). The expression of HLA genes is essential for immune-mediated regression of metastases, with certain HLA variants causing resistance to immune therapy (43–44). There may thus be a link between HLA genotype and the immunoeediting of oncogenes and tumor suppressors (45). IFN γ or TIS signatures described by Ayers and colleagues (46) were not found in the factor analysis. As indicated by the manual pathologic annotation, lymphoid areas were only present in two of the metastases. Therefore, a possible explanation is that T cell activation is present in too few of the 2,200 tissue domains for the signature to represent one of the 20 factors generated in our

analyses. However, when examining the signature genes in the spatial data of one of the tumors with lymphoid tissue, the IFN γ or TIS signatures were expressed particularly within the transition area described below (count matrixes for exploring genes / gene signatures of interest are available at <http://www.spatialtranscriptomicsresearch.org/>).

The expression of HLA genes was also found in the factor we classified as the transition area, which corresponds to regions between the melanoma and lymphoid areas that could not be manually defined. Further examination of this transition area could provide more information regarding tumor heterogeneity and resistance mechanisms in melanoma. Tirosh and colleagues (13) analyzed the transcriptional heterogeneity in melanoma, including lymph node metastases, by scRNA-seq, and concluded that the TME is shaped by spatial, functional, and genomic heterogeneity of melanoma as well as the associated tumor components. They also found that genes expressed by cancer-associated fibroblasts seem to affect the proportion of T-cells in the melanoma lymph node metastases. We can only speculate about whether the genetically defined transition areas in the TME identified through our approach host cancer-derived factors that interact with immune cells similarly as what has been described for stromal-derived cells in melanoma and situations of immune cell abundance. In the present report, the most abundantly expressed genes in the transition area were *FTL*, *B2M*, *APOE* and HLA-associated genes (*HLA A-C*). These genes have been linked with tumor growth regulation through the *GADD45/JNK* pathway in other cancers such as in glioblastomas (*FTL*; ref. 47), along with loss of immunogenicity and acquired resistance to PD-1 blockade in patients with melanoma through defects in the pathways involved in interferon-receptor signaling, such as *B2M*, and in antigen presentation (43, 44, 48).

For validation, we performed a joint analysis of our data and a public available dataset from Tirosh and colleagues (13). In summary, the single-cell RNA-seq data confirmed the extensive heterogeneity across patients that was also found in the spatial data. Interestingly however, patient-specific signatures from the single-cell data could be detected in specific regions of the metastases.

Our findings that cell–cell interactions in the TME may support tumor progression and influence therapy response warranted further analyses. We therefore performed individual analyses of tissue sections to visualize the TME using principle component analysis. This approach identified multiple gene expression programs in lymphoid tissue that might depend on distance from the tumor cell clusters. Notably, genes such as *PMEL* and *SPP1* were overexpressed in the tumor cell cluster whereas lymphoid tissue regions far away from, and in close proximity to, the tumor cell areas were characterized by expression of the immune-related genes *CD74* and *IGLL5*, respectively. The transition area could only be clearly identified in one out of the four tumors analyzed in this study and more samples would be required for a deeper understanding of the TME in this region.

In conclusion, the spatial information added to transcriptomic analyses has revealed a more detailed portrait of the heterogeneity among melanoma metastases. The presented results should motivate researchers to integrate the spatial component into analyses of the multiple factors underlying tumor progression and therapy outcome.

Disclosure of Potential Conflicts of Interest

J. Lundeberg is a member of the board for Spatial Transcriptomics AB; has ownership interest (including stock, patents, etc.) in Spatial Transcriptomics AB. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: K. Thrane, H. Eriksson, J. Lundeberg
Development of methodology: K. Thrane
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): H. Eriksson, J. Hansson
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): K. Thrane, H. Eriksson, J. Maaskola, J. Hansson
Writing, review, and/or revision of the manuscript: K. Thrane, H. Eriksson, J. Hansson, J. Lundeberg
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): K. Thrane, H. Eriksson
Study supervision: J. Hansson, J. Lundeberg

References

- Betensky RA, Louis DN, Cairncross JG. Influence of unrecognized molecular heterogeneity on randomized clinical trials. *J Clin Oncol* 2002;20:2495–9.
- Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. *Nature* 2013;501:328–37.
- Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. *Science* 2013;339:1546–58.
- Hodis E, Watson IR, Kryukov GV, Arola ST, Imielinski M, Theurillat JP, et al. A landscape of driver mutations in melanoma. *Cell* 2012;150:251–63.
- Curtin JA, Fridlyand J, Kageshita T, Patel HN, Busam KJ, Kutzner H, et al. Distinct sets of genetic alterations in melanoma. *N Engl J Med* 2005;353:2135–47.
- Merlo LMF, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. *Nat Rev Cancer* 2006;6:924–35.
- Klein CA. Selection and adaptation during metastatic cancer progression. *Nature* 2013;501:365–72.
- Shain AH, Yeh I, Kovalyshyn I, Sriharan A, Talevich E, Gagnon A, et al. The genetic evolution of melanoma from precursor lesions. *N Engl J Med* 2015;373:1926–36.
- Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multi-region sequencing. *Nat Rev Cancer* 2012;366:883–92.
- Katona TM, Jones TD, Wang M, Eble JN, Billings SD, Cheng L. Genetically heterogeneous and clonally unrelated metastases may arise in patients with cutaneous melanoma. *Am J Surg Pathol* 2007;31:1029–37.
- Harbst K, Staaf J, Mäsback A, Olsson H, Ingvar C, Vallon-Christersson J, et al. Multiple metastases from cutaneous malignant melanoma patients may display heterogeneous genomic and epigenomic patterns. *Melanoma Res* 2010;20:381–91.
- Yancovitz M, Litterman A, Yoon J, Ng E, Shapiro RL, Berman RS, et al. Intra- and inter-tumor heterogeneity of BRAF(V600E) mutations in primary and metastatic melanoma. *PLoS One* 2012;7:e29336.
- Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 2016;352:189–96.
- Balch CM, Gershenwald JE, Soong SJ, Thompson JF, Atkins MB, Byrd DR, et al. Final version of 2009 AJCC melanoma staging and classification. *J Clin Oncol* 2009;27:6199–206.
- Middleton MR, Grob JJ, Aaronson N, Fierlbeck G, Tilgen W, Seiter S, et al. Randomized phase III study of temozolomide versus dacarbazine in the treatment of patients with advanced metastatic malignant melanoma. *J Clin Oncol* 2000;18:158–66.
- Long GV, Stroyakovskiy D, Gogas H, Levchenko E, De Braud F, Larkin J, et al. Dabrafenib and trametinib versus dabrafenib and placebo for Val600 BRAF-mutant melanoma: a multicentre, double-blind, phase 3 randomised controlled trial. *Lancet* 2015;386:444–51.
- Zand S, Buzney E, Duncan LM, Dadras SS. Heterogeneity of metastatic melanoma. *Am J Clin Pathol* 2016;146:353–60.
- Harbst K, Lauss M, Cirenajwis H, Isaksson K, Rosengren F, Törnqvist T, et al. Multiregion whole-exome sequencing uncovers the genetic evolution and mutational heterogeneity of early-stage metastatic melanoma. *Cancer Res* 2016;76:4765–74.
- Sabatino M, Zhao Y, Voiculescu S, Monaco A, Robbins P, Karai L, et al. Conservation of genetic alterations in recurrent melanoma supports the melanoma stem cell hypothesis. *Cancer Res* 2008;68:122–31.
- Takata M, Morita R, Takehara K. Clonal heterogeneity in sporadic melanomas as revealed by loss-of-heterozygosity analysis. *Int J Cancer* 2000;85:492–7.
- Eriksson H, Zebary A, Vassilaki I, Omholt K, Ghaderi M, Hansson J. BRAFV600E protein expression in primary cutaneous malignant melanomas and paired metastases. *JAMA Dermatol* 2015;151:410–6.
- Stahl PL, Salmen F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 2016;353:78–82.
- Edlundh-Rose E, Egyházi S, Omholt K, Månsson-Brahme E, Platz A, Hansson J, et al. NRAS and BRAF mutations in melanoma tumours in relation to clinical characteristics: a study based on mutation screening by pyrosequencing. *Melanoma Res* 2006;16:471–8.
- Grafström E, Egyházi S, Ringborg U, Hansson J, Platz A. Biallelic deletions in INK4 in cutaneous melanoma are common and associated with decreased survival. *Clin Cancer Res* 2005;11:2991–7.
- Johansson CC, Egyházi S, Masucci G, Harlin H, Mouggiakakos D, Poschke I, et al. Prognostic significance of tumor iNOS and COX-2 in stage III malignant cutaneous melanoma. *Cancer Immunol Immunother* 2009;58:1085–94.
- Balch CM, Buzaid AC, Soong SJ, Atkins MB, Cascinelli N, Coit DG, et al. Final version of the American Joint Committee on Cancer staging system for cutaneous melanoma. *J Clin Oncol* 2001;19:3635–48.
- Navarro JF, Sjöstrand J, Salmén F, Lundeberg J, Ståhl PL. ST Pipeline: an automated pipeline for spatial mapping of unique transcripts. *Bioinformatics* 2017;33:2591–3.
- Wickham H. *ggplot: Elegant graphics for data analysis*. New York, NY: Springer Verlag; 2009.
- Love MI, Anders S, Huber W. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550. doi: 10.1186/s13059-014-0550-8.
- Kolde R. Package 'pheatmap'. *Bioconductor* 2012;1–6.
- Berglund E, Maaskola J, Schultz N, Marklund M, Friedrich S, Bergenstråhle J, et al. Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nat Commun* 2018;9:2419. doi: 10.1038/s41467-018-04724-5.
- Lun ATL, McCarthy DJ, Marioni JC. A step-by-step workflow for low-level analysis of single-cell RNA-seq data with Bioconductor. *F1000Research* 2016;5:2122. doi: 10.12688/f1000research.9501.2.

Acknowledgments

We acknowledge help from the National Genomics Infrastructure (NGI) Stockholm and SciLifeLab in providing infrastructural support and from Ludvig Larsson and Konstantin Carlberg in bioinformatics support. We also thank all patients for participating.

This work was supported by Stockholm County Council - Health, Medicine and Technology (grant nos. 20130751, 20160852), Knut and Alice Wallenberg Foundation (grant no. 2015.0296), the Swedish Cancer Society (grant no. CAN2014-424), the Swedish Research Council (grant no. 621-2014-5629), the Swedish Society for Medical Research (grant no. 160502), Cancer Research Foundations of Radiumhemmet (grant no. 161052), Swedish Society of Medicine (grant no. 6844391), Alex and Eva Wallström Foundation (grant no. 2018-00222), and the KI funds (grant no. 2016fobi50190).

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 March 16, 2018; revised July 3, 2018; accepted August 23, 2018; published first August 28, 2018.

33. Van Der Maaten L, Hinton G. Visualizing data using t-SNE. *J Mach Learn Res* 2008;620:267–84.
34. Parra ER. Novel platforms of multiplexed immunofluorescence for study of paraffin tumor tissues. *J Cancer Treat Diagnosis* 2018;2:42–53.
35. Goltsev Y, Samusik N, Kennedy-Darling J, Bhate S, Hale M, Vasquez G, et al. Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. *bioRxiv* 2018; doi: <https://doi.org/10.1101/203166>.
36. Chang Q, Ornatsky OI, Siddiqui I, Loboda A, Baranov VI, Hedley DW. Imaging mass cytometry. *Cytometry A* 2017;91:160–9.
37. Ziai J, Caplazi P, Decalf J, Liang Y, de Almeida P, Zollinger D, et al. Highly multiplexed analysis of immune cell subsets in non-small cell lung cancer: validation of protein and RNA analysis by the Nanostring Digital Spatial Profiling (DSP) platform. *Proc Am Assoc Ca Res* 2018;59poster #2089.
38. Radford KJ, Mallesch J, Mersey P. Suppression of human melanoma cell growth and metastasis by the melanoma-associated antigen CD63 (ME491). *Int J Cancer* 1995;62:631–5.
39. Li FZ, Dhillon AS, Anderson RL, McArthur G, Ferrao PT. Phenotype switching in melanoma: implications for progression and therapy. *Front Oncol* 2015;5:31. doi: 10.3389/fonc.201500031.
40. Kemper K, De Goeje PL, Peeper DS, Van Amerongen R. Phenotype switching: Tumor cell plasticity as a resistance mechanism and target for therapy. *Cancer Res* 2014;74:5937–41.
41. Konieczkowski DJ, Johannessen CM, Abudayyeh O, Kim JW, Cooper ZA, Piris A, et al. A melanoma cell state distinction influences sensitivity to MAPK pathway inhibitors. *Cancer Discov* 2014;4:816–27.
42. Ekmekcioglu S, Davies MA, Tanese K, Roszik J, Shin-Sim M, Bassett RL, et al. Inflammatory marker testing identifies CD74 expression in melanoma tumor cells, and its expression associates with favorable survival for stage III melanoma. *Clin Cancer Res* 2016;22:3016–24.
43. Budczies J, Bockmayr M, Klauschen F, Endris V, Fröhling S, Schirmacher P, et al. Mutation patterns in genes encoding interferon signaling and antigen presentation: a pan-cancer survey with implications for the use of immune checkpoint inhibitors. *Genes Chromosom Cancer* 2017;56:651–9.
44. Zaretsky JM, Garcia-Diaz A, Shin DS, Escuin-Ordinas H, Hugo W, Hu-Lieskovan S, et al. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N Engl J Med* 2016;375:819–29.
45. Marty R, Kaabinejadian S, Rossell D, Slifker MJ, van de Haar J, Engin HB, et al. MHC-I genotype restricts the oncogenic mutational landscape. *Cell* 2017;171:1272–83.
46. Ayers M, Luceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, et al. IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest* 2017;127:2930–40.
47. Wu T, Li Y, Liu B, Zhang S, Wu L, Zhu X, et al. Expression of ferritin light chain (FTL) is elevated in glioblastoma, and FTL silencing inhibits glioblastoma cell proliferation via the GADD45/JNK pathway. *PLoS One* 2016;11:e0149361.
48. Sucker A, Zhao F, Real B, Heeke C, Bielefeld N, Maßen S, et al. Genetic evolution of T-cell resistance in the course of melanoma progression. *Clin Cancer Res* 2014;20:6593–604.