



Published in final edited form as:

Cell Syst. 2023 May 17; 14(5): 404–417.e4. doi:10.1016/j.cels.2023.03.008.

## Deciphering tumor ecosystems at super resolution from spatial transcriptomics with TESLA

Jian Hu<sup>1,7,\*</sup>, Kyle Coleman<sup>2</sup>, Daiwei Zhang<sup>2</sup>, Edward B. Lee<sup>3</sup>, Humam Kadara<sup>4</sup>, Linghua Wang<sup>5,6,\*</sup>, Mingyao Li<sup>2,\*</sup>

<sup>1</sup>Department of Human Genetics, School of Medicine, Emory University, Atlanta, GA 30322, USA

<sup>2</sup>Department of Biostatistics, Epidemiology and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>3</sup>Translational Neuropathology Research Laboratory, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>4</sup>Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>5</sup>Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

<sup>6</sup>The University of Texas MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences (GSBS), Houston, TX 77030, USA

<sup>7</sup>Lead contact

### SUMMARY

Cell populations in the tumor microenvironment (TME), including their abundance, composition, and spatial location, are critical determinants of patient response to therapy. Recent advances in spatial transcriptomics (ST) have enabled the comprehensive characterization of gene expression in the TME. However, popular ST platforms, such as Visium, only measure expression in low-resolution spots and have large tissue areas that are not covered by any spots, which limits their usefulness in studying the detailed structure of TME. Here, we present TESLA, a machine learning framework for tissue annotation with pixel-level resolution in ST. TESLA integrates histological information with gene expression to annotate heterogeneous immune and tumor cells directly on the histology image. TESLA further detects unique TME features such as tertiary lymphoid structures, which represents a promising avenue for understanding the spatial

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\*Correspondence: jian.hu@emory.edu (J.H.), lwang22@mdanderson.org (L.W.), mingyao@penncmedicine.upenn.edu (M.L.).

#### AUTHOR CONTRIBUTIONS

This study was conceived and led by M.L. J.H. designed the model and algorithm. J.H. implemented the TESLA software and led the data analysis with input from M.L., L.W., E.B.L., H.K., K.C., and D.Z. K.C. performed cell-type deconvolution analysis. E.B.L. examined the histology images. L.W. and H.K. provided marker genes for cancer and interpreted the results. L.W. provided guidance for the tumor TME analysis. J.H., M.L., and L.W. wrote the paper with feedback from all other coauthors.

#### DECLARATION OF INTERESTS

M.L. received research funding from Biogen Inc. unrelated to the current manuscript.

#### SUPPLEMENTAL INFORMATION

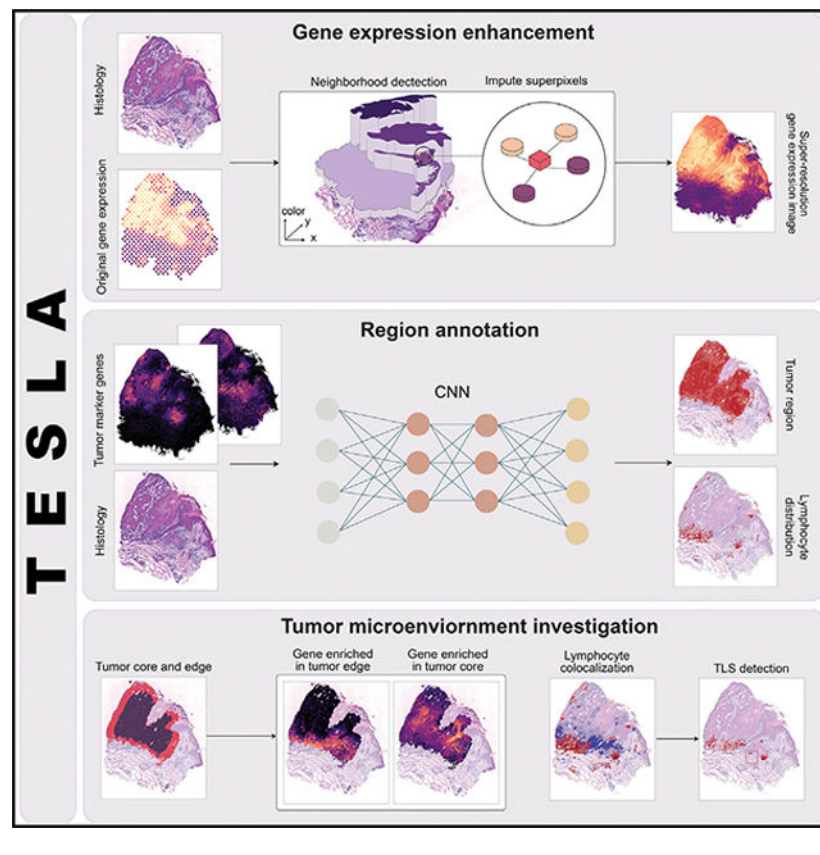
Supplemental information can be found online at <https://doi.org/10.1016/j.cels.2023.03.008>.

architecture of the TME. Although we mainly illustrated the applications in cancer, TESLA can also be applied to other diseases.

## In brief

We present TESLA, a machine learning framework for tissue annotation with pixel-level resolution, which uses spatial transcriptomics. TESLA is able to enhance gene expression to a higher resolution using histology images, and by integrating morphology and molecular profiles, TESLA provides a comprehensive characterization of the tumor microenvironment in various tumor types.

## Graphical Abstract



## INTRODUCTION

The tumor microenvironment (TME) contains networks of cells and structures that surround tumor cells.<sup>1</sup> Tumor-TME crosstalk regulates the initiation, progression, and metastasis of tumors.<sup>2</sup> A comprehensive analysis of the multiple exchanges between tumor cells and their TME is essential for understanding the underlying mechanisms of tumor growth and response to therapy. The TME has a great impact on the efficacy of anti-cancer therapies and clinical outcomes.<sup>3,4</sup> For example, hypoxic avascular regions deeply embedded inside the tumors hinder the delivery of therapeutic agents,<sup>5</sup> and hypoxic tolerant tumor cells are resistant to most anti-cancer treatments.<sup>6,7</sup>

Tumor-infiltrating lymphocytes (TILs) consist of lymphocytic cell populations that have invaded the tumor tissue and have emerged as an important biomarker in predicting the efficacy and outcome of anti-cancer treatment.<sup>8,9</sup> Recent studies suggest that not only the density of TILs, but also their spatial organization, such as the presence of tertiary lymphoid structures (TLSs), play key roles in determining tumor immune phenotypes and response to immunotherapy.<sup>10</sup> TLSs are broadly found in the TME for various solid cancers. Similar to secondary lymphoid organs, mature TLSs are believed to be the site of immune response activation against tumors by recruiting and activating TILs, and they represent the key to understanding antitumor immune responses.

Spatial transcriptomics (ST) enables gene expression profiling while preserving location information in tissues, which innovates a promising avenue to understand the spatial context and the nature of cellular heterogeneity of the TME.<sup>11–13</sup> In an ST experiment using spatial barcoding followed by next-generation sequencing-based technologies, such as ST<sup>14</sup> and 10x Genomics Visium, the expression levels for thousands of genes in a tissue section are simultaneously measured, complemented by a high-resolution histology image obtained from the same tissue section. With the power of ST, we aim to provide a detailed annotation of tumor structure and different lymphocytes by integrating gene expression and histology image information.

A major challenge that hinders gene expression and histology integration in spatial barcoding-based ST is the relatively low resolution of gene expression data compared with histology images, which prevents deciphering detailed TME structures, such as the TLS. Although methods such as BayesSpace<sup>15</sup> can enhance the gene expression resolution, their enhancement is only for regions captured by spots, but they still leave a large portion of the tissue unmeasured for gene expression. For example, for 10x Visium, each spot is 55  $\mu\text{m}$  in diameter with a 100- $\mu\text{m}$  center-to-center distance between spots. The lack of full tissue coverage by spots leaves approximately 54%–80% of the tissue unmeasured for gene expression (Note S1). A recently developed method, XFuse,<sup>16</sup> uses a deep generative model to predict super-resolution gene expression at each pixel by leveraging the information in hematoxylin and eosin (H&E)-stained histology images. However, XFuse is extremely slow—it often takes more than 2 weeks to analyze a tissue section generated from 10x Visium. Additionally, it only shows high accuracy for ~100 highly expressed genes. For genes that have moderate to low expression, it is unclear whether they can faithfully recover the original pixel-level gene expression.

To overcome the above-mentioned challenge, we present tumor edge structure and lymphocyte multi-level annotation (TESLA), a machine learning framework that integrates gene expression and histology image information in ST to investigate the TME. By generating super-resolution gene expression images, TESLA can leverage histology information to annotate different tumor/TME cell types on the histology image with pixel-level resolution, detect TIL structure, such as TLS, by colocalization analysis of different lymphocytes and dendritic cells, and characterize high-resolution cellular and molecular spatial structures of tumors by separating the tumor edge and core and elucidating differential transcriptome programs between the edge and core. The detailed multi-level annotations performed by TESLA provide a comprehensive understanding of the TME.

## RESULTS

### Overview of TESLA and evaluation

We describe the TESLA workflow using ST data from tumor as an example, though the method can be easily modified to analyze ST data from other tissue types. TESLA takes spatial gene expression and histology image as input. The first step is to enhance the relatively low-resolution gene expression to the same resolution as the histology image. Next, the gene expression image and histology image are stacked together and used as input for a neural network to annotate tissue regions. Both steps are evaluated for their accuracy.

**Enhancement of gene expression resolution**—The most important step in TESLA is to enhance gene expression resolution by leveraging the information provided by the companion high-resolution histology image generated from the same tissue section (Figure 1A). This is based on the intuition that tissue regions that share similar histological features are also likely to share the similar gene expression, implying that histology similarity can help impute gene expression. To impute super-resolution gene expression and fill in the tissue gaps, TESLA will first generate superpixels from the histology image. At each superpixel, TESLA will identify its nearest neighboring spots based on physical location and histology similarity. TESLA will then impute the superpixel's gene expression through weighted aggregation of spot-level gene expression values from these neighboring spots. By the end of this imputation, TESLA will return a super-resolution gene expression image for each gene.

**Multi-level tissue annotation**—With the super-resolution gene expression images generated for tumor, immune, and stromal cells, TESLA can annotate the tumor region and tumor-immune interface at multiple levels. To infer tumor cell distribution, TESLA first combines tumor-related marker genes into a meta gene, which captures the shared expression patterns of the markers (Figure 1B). After stacking the meta gene images and the histology image as input, TESLA then employs a convolutional neural network (CNN) approach to segment the tissue region in an unsupervised manner, which enables the annotation of the tumor region directly on the histology image with pixel-level resolution. Because the tumor cells and microenvironment at the invasive edge of the tumor are known to be different from that of the tumor core,<sup>2</sup> TESLA also draws the leading edge along the tumor boundary and outlines the tumor edge and tumor core (Figure 1C).

To characterize spatial cellular and molecular heterogeneity between the tumor edge and core and better understand tumor-TME dynamics, TESLA first performs differential gene expression analysis between the edge and core, which enables the identification of differentially expressed genes (DEGs) that are specific to the tumor edge or core regions. With the region-specific DEGs, TESLA can also compare cell compositions and transcriptome programs between the edge and the core and further profile spatial heterogeneity of the TME. In addition to tumor region annotation, TESLA can also annotate a target cell type's distribution on the histology image by using cell-type-specific marker genes, e.g., lymphoid cell lineage markers, as input (Figure 1B). Based on the colocalization

of specific cell populations, TESLA can further detect multicellular structures such as TLS (Figure 1D).

**Evaluation strategy**—To showcase the strength and properties of TESLA, we applied it to eight publicly available ST datasets<sup>15,17–22</sup> (Table S1). First, we systematically compared our method with BayesSpace for gene expression enhancement. We next show that the super-resolution gene expression from TESLA correlates well with protein expression obtained from immunofluorescence staining. We further show that the cell-type distribution predicted by TESLA agrees well with independent cell-type deconvolution results by robust cell-type decomposition (RCTD),<sup>23</sup> and the predicted TLSs are in agreement with independent annotation by a pathologist.

### Imputation of super-resolution gene expression by integrating histology image information

#### Evaluation of gene expression imputation accuracy by hold-off experiments

—TESLA's multi-level cellular and molecular annotation of ST data relies on the super-resolution gene expression images. We conducted hold-off experiments on five public datasets<sup>15,17,18,21,22</sup> to show that TESLA can faithfully recover the missing expression patterns for genes with spatial variability. For each dataset, we first selected spatially variable genes (SVGs) using our previously developed method, SpaGCN.<sup>24</sup> Next, we masked each measured spot and used the remaining spots to impute the SVGs' expression for the masked spot. We repeated this masking for all of the spots and examined whether TESLA can faithfully impute the missing gene expression at the masked spots. Figure 2A shows the correlation between the observed and imputed SVGs' expression of the masked spots. The median correlation among the SVGs is 0.60 for the invasive ductal carcinoma (IDC) dataset,<sup>15</sup> 0.71 for the cutaneous squamous cell carcinoma (CSCC) dataset,<sup>17</sup> 0.69 for the cutaneous malignant melanoma dataset,<sup>18</sup> 0.70 for the mouse posterior brain tissue dataset, and 0.54 for the mouse kidney dataset. These high correlations indicate that TESLA can recover the underlying expression pattern for unmeasured spots. Because the distances between superpixels and their nearest neighboring spots in real ST data are smaller than the distances between the masked spots and their neighboring spots in the hold-off experiments, we expect TESLA to perform better in real data analysis.

**Benchmark evaluation and comparison with BayesSpace**—To showcase the strength of TESLA's imputed super-resolution gene expression, we compared the performance of TESLA with BayesSpace,<sup>15</sup> a recently developed tool that can also enhance the gene expression resolution for ST data. TESLA features two benefits relative to BayesSpace: (1) TESLA can incorporate high-resolution histology image information and offers more flexibility for resolution enhancement by adjusting the superpixel size, whereas BayesSpace can only enhance the gene expression resolution by splitting a spot into fixed number of subspots, and (2) TESLA can fill in gene expression for unmeasured tissue regions between spots, whereas BayesSpace leaves unmeasured tissue areas blank. For 10x Visium, each spot is 55  $\mu\text{m}$  in diameter with a 100- $\mu\text{m}$  center-to-center distance between spots, and this leaves a large portion of the tissue unmeasured for gene expression. We compare the super-resolution gene expression from TESLA and BayesSpace in the following three aspects using real data.

First, a good gene expression enhancement method should increase the expression resolution while retaining the original expression pattern at the spot level, as this will ensure no artificial patterns are introduced in the enhanced gene expression. To show that TESLA outperforms BayesSpace in retaining the original gene expression patterns, we considered the top 2,000 highly variable genes selected by BayesSpace and obtained the spot-level gene expression from the enhanced expression generated from both TESLA and BayesSpace. For TESLA, we obtained the spot-level gene expression from the super-resolution gene expression image by extracting expression from circles that exactly overlap measured spots. For BayesSpace, we summed up the expression from all sub-spots within a spot to get the spot-level expression. As shown in Figure 2B, TESLA's super-resolution gene-expression-derived spot-level expression yields significantly higher correlations with the original spot-level gene expression than BayesSpace across all five datasets. Although BayesSpace requires that the sum of principal components of the subspots equals the principal component of the original spot, it cannot guarantee that the expression pattern is not distorted in the original gene expression space (Figure S1).

Second, an accurate gene expression enhancement method should be able to recover gene expression pattern at a higher resolution compared with the original. To show that TESLA has better ability to recover high-resolution gene expression than BayesSpace, we benchmarked these two methods using a 10x Xenium dataset from human breast cancer tissue,<sup>25</sup> which has a single-cell resolution. We made pseudo-Visium data from the Xenium data by extracting square regions with sizes equal  $40 \mu\text{m} \times 40 \mu\text{m}$ , which have the same area as the  $55\text{-}\mu\text{m}$  circle spots in Visium. As shown in Figure 2C, the center-to-center distances between pseudo-spots are  $100 \mu\text{m}$ , leaving the same tissue gaps as Visium. The number of cells in these pseudo-spots is between 5 and 10, which is also similar to Visium (Figure S2). After the pseudo-Visium data were generated, we then performed gene expression enhancement using TESLA and BayesSpace. To evaluate gene expression imputation accuracy, we calculated Pearson correlations between the enhanced gene expression and the ground truth gene expression for all 313 genes measured in Xenium at different resolutions, i.e., 50, 32, and  $16 \mu\text{m}$ . As shown in the Figure 2D, TESLA has higher correlations with the ground truth than BayesSpace across all resolutions. We also show that TESLA's gene expression enhancement will not detect the tumor boundary (Note S2).

Third, an ideal super-resolution gene expression imputation approach should yield gene expression that is similar to the corresponding protein expression obtained from immunofluorescence staining.<sup>15</sup> To check this ability, we evaluated the super-resolution gene expression from TESLA and BayesSpace with the immunofluorescence staining for the same gene/protein pair obtained from the same tissue section in the IDC dataset, which includes immunofluorescence staining for CD3e, a T cell co-receptor, which is involved in activating both the cytotoxic T cell and T helper cells. Because protein and RNA expression measure different modalities, rather than directly calculating a correlation coefficient, we compared the similarity of these two modalities using the structural similarity index (SSIM),<sup>26</sup> a commonly used metric for measuring the similarity between two images, for evaluation. SSIM ranges from  $-1$  to  $1$  with a value of  $1$  indicating two identical images,  $0$  indicating no structural similarity, and  $-1$  indicating strong dissimilarity. Compared with the CD3e immunofluorescence staining image (Figure 2E), the super-resolution *CD3E* gene

image from TESLA is visually similar (SSIM 0.33; Figure 2F). In contrast, the SSIM is only 0.08 for BayesSpace (Figure 2G), presumably because of its inability to fill in gene expression in the unmeasured tissue area.

In addition to BayesSpace, we also compared TESLA with XFuse (Note S3). In summary, compared with other existing methods, the enhanced gene expression from TESLA can better recover the high-resolution expression pattern while retaining the original expression pattern untwisted. The super-resolution gene expression image from TESLA also shows a high similarity with the correspondent protein immunofluorescence staining.

### Characterization of the high-resolution cellular and molecular spatial structure of the tumor

The tumor edge is considered the tumor invasion front because of its strong invasive ability, and it has been reported that the tumor edge has a unique microenvironment and distinct morphological, structural, and molecular features from that of the tumor core, such as the level of immune cell infiltration and compositions, vascular density, hypoxia status, metabolic rate, etc.<sup>27,28</sup> The super-resolution gene expression images generated by TESLA enables the detection and precise annotation of the tumor edge, which enables further exploration of the structural, cellular, and molecular differences in tumor and TME cells between the primary tumor core and edge.

**Application to CSCC**—To show that such analyses can advance our understanding of intra-tumor heterogeneity, we first analyzed the CSCC dataset of the skin.<sup>17</sup> As reported in the original study, the tumors include basal, cycling, and differentiating keratinocyte cell populations that are similar to normal skin, and a tumor-specific keratinocyte (TSK) population. Therefore, we identified the tumor region using nine TSK marker genes suggested in the original study (*NT5E*, *PTHLH*, *INHBA*, *LAMC2*, *MMP10*, *FEZ1*, *CD151*, *IL24*, and *SLITRK6*), plus three commonly used marker genes (*TOP2A*, *KIF1C*, and *BUB1B*). Utilizing these genes (Figure S3), TESLA derived a meta gene, which captured all their patterns and further detected the tumor region by joint segmentation of the meta gene and histology images. Figure 3A shows that the tumor region detected by TESLA agrees well with the dark region in the histology image except for the upper right and bottom right regions. Further examination of the histology image of the upper right region by a pathologist (E.B.L.) revealed that this region consists of dead cells with hypereosinophilic cytoplasm and pyknotic nuclei (yellow box). This is further confirmed by the relatively low total unique molecular identifier (UMI) counts compared with other regions (Figure S4). The region on the right represents non-neoplastic skin adjacent to the squamous cell carcinoma (blue box). This example shows that by combining gene expression and histology images together, TESLA can accurately annotate live tumor region from non-viable or non-neoplastic tissue without the reliance on a pathologist's manual annotation.

The leading edge drawn by TESLA also agrees well with the manual annotation provided by a pathologist (Figure 3B). After detecting the tumor boundary, TESLA further separated the tumor into core and edge (Figure 3C). The tumor edge is defined as the area that is within 200  $\mu\text{m}$  interface of the tumor and adjacent non-malignant tissue, and the tumor core

is defined as the proximal tumor area within the tumor edge.<sup>27</sup> DEG analysis between the tumor core and edge identified genes that are highly enriched in the core (3,665 genes) or edge (106 genes). For example, in Figures 3D and S5, oncogenes such as *IGFBP2*, *CCNB1*, *KRT15*, and *HRAS* are highly expressed in the core region, whereas tumor-edge-enriched genes included *CXCL12*, *IL2RB*, *LAIR1*, and *LILRB2*, which are immune related, implying a distinct microenvironment from that of the tumor core. We further performed gene set enrichment analysis for the 106 tumor-edge-enriched genes and the top 300 tumor-core-enriched genes with the highest fold changes, by computing overlaps with curated gene sets in the Molecular Signature Database<sup>29</sup> (Table S2). Our analysis suggested significant enrichment of the metabolism of lipids under hypoxic conditions and mitotic phase in cell cycle in the tumor core. The high proliferation rate of cancer cells in the tumor core leads to hypoxia and nutrient deprivation, which affect the lipid metabolism of cells.<sup>30,31</sup> Pathways that are highly activated in the tumor edge included innate immune system, adaptive immune system, interferon- $\gamma$  response, inflammatory response, and cytokine signaling. These highly activated immune systems in the tumor edge are possibly due to a higher level of immune cell density and tumor-immune interactions in the edge. Together, TESLA's unique feature to characterize and profile tumor edge and core may provide great insights into better understanding of the TME and the mechanisms of immune suppression and evasion.

**Application to melanoma**—To show the generalizability of TESLA in characterizing the cellular and molecular spatial structure of tumors, we next analyzed the cutaneous malignant melanoma dataset.<sup>18</sup> Using a list of genes for clinical melanoma diagnosis<sup>32</sup> (*MITF*, *CSPG4*, *MAGEA1*, *MLANA*, *TYR*, and *SOX10*; Figure S6), TESLA identified the melanoma region, which strongly recapitulate the pathologist's manual annotation in Figure 3E. Following similar steps as described previously, we further separated the tumor region into tumor core and edge (Figure 3F) and detected 3,510 genes enriched in the tumor core and 155 genes enriched in the tumor edge. Some hypoxia-related genes<sup>33,34</sup> (e.g., *HSPB1*, *MIF*, and *TPI1*) were detected to be highly expressed in the tumor core, whereas most genes highly expressed in the tumor edge are immune related (Figures 3G and S7). Next, we performed gene set enrichment analysis for the 155 edge enriched genes and the top 300 core-enriched genes with the highest fold changes (Table S3). Interestingly, for the top 300 core-enriched genes, we observed a significant upregulation of the signaling by Rho GTPases, receptor tyrosine kinase, and Wnt and that MET activates RAS signaling, which are known to be involved in melanoma progression through regulating cell proliferation and invasion.<sup>35,36</sup> In addition, we observed increased signaling by MET, which has emerged as a paradigm of tumor resistance to modern targeted therapies, and the assessment of its expression in patients' samples may be a valuable biomarker of tumor progression and response to targeted therapy.<sup>37</sup> The enriched genes in the tumor edge are largely related to inflammatory response, such as *PDCD1*, *IL10RA*, *PAX5*, *CCL19*, *TNFSF4*, *IFI27*, *IL32*, and *IL4R*. Consistently, pathway enrichment analysis revealed increased expression of interferon gamma response, cytokine signaling and cytokine-cytokine receptor interaction, and increased signaling by Interleukins, suggesting activated anti-tumor immune response. Interestingly, we also observed up-regulated expression of genes defining epithelial-mesenchymal transition (EMT). EMT is a process through which epithelial tumor cells acquire mesenchymal phenotypic properties, and it contributes to



both metastatic dissemination and therapy resistance in cancer.<sup>38</sup> The cell-type distributions inferred by cell-type deconvolution analysis with RCTD also showed significant difference between the tumor edge and core (Figure S8), with the tumor edge enriched for T cells, whereas the core enriched for malignant cells.

We also demonstrate that all these findings, based on the super-resolution annotation, cannot be achieved with the original spot-level data (Note S4). The failure of detecting tumor-edge-enriched genes at the spot level is possibly due to two reasons. First, the number of observations is much smaller when considering a spot as the analysis unit, and the reduced sample size in DEG analysis will lead to less power. Second, the spot-level data do not have a single-cell resolution. The number of cells within a spot depends on their type and state and local morphology in the corresponding tissue. Indeed, the diameter size of each spot is 100  $\mu\text{m}$  in ST, which is much larger than a single cell. Based on a recently published study that utilized the ST platform, it is estimated that the number of cells per spot is  $\sim 40$ .<sup>39</sup> Because each spot may contain many cells, the mixture of cells from different cell types will dilute the differential expression signal, especially when the immune cells are rare in the tumor edge. Numerous studies have shown that the tumor-immune interface plays an important role in understanding the TME with primary relevance to the success of immunotherapy.<sup>40</sup> Therefore, we believe that until the sequencing-based ST technologies reach a single-cell resolution, gene expression resolution enhancement will be needed when the goal is to detect gene expression changes that occur only in a small region of the tissue.

### Identification of lymphoid cell types and TLS

**Application to CSCC**—Next, we show that TESLA can profile the spatial distribution of different immune cell subsets based on super-resolution expression images of lineage-specific genes. We first analyzed the CSCC data.<sup>17</sup> Using canonical cell-type marker genes (Figure S9), we detected the locations for B cells, CD4<sup>+</sup> T cells, dendritic cells, and CXCL13, a marker of TLS that is known to be also constitutively expressed in secondary lymphoid tissue. As shown in Figure 4A, TESLA's cell-type distribution has much higher resolution than results obtained from spot-level gene expression by SpaGCN,<sup>24</sup> a clustering method developed for ST data (Figure 4B). For ST and Visium data, each spot may contain multiple cells. By leveraging high-resolution histology image data, TESLA's cell-type annotation achieves pixel-level resolution, which better describes the underlying cell type distributions. By colocalizing B cells, CD4<sup>+</sup> T cells, dendritic cells, and CLCX13 near the tumor region, TESLA further detected TLSs (Figure 4C). Examination of the histology image by a pathologist (E.B.L.) indicates that the blue-boxed region contains aggregates of lymphocytes consistent with a TLS on the edge of the tumor. We also observed the colocalization of lymphocytes in the yellow-boxed region, which are not forming tight, aggregated groups compared with the blue-boxed region. We suspect that these lymphocytes are collecting in this region but are not forming a TLS because of geospace's constraint of the tissue or perhaps because this region represents the edge of a TLS, which is not present due to the relative thinness of the histology image. TLS is an ectopic lymphoid formation within nonlymphoid tissue. TLS can additionally foster tumor antigen presentation and T cell activation and the presence of TLS has been associated with improved response to cancer immunotherapy and prolonged patient survival.<sup>41–43</sup> Similar analysis was performed

on the cutaneous malignant melanoma data<sup>18</sup> using canonical cell-type marker genes (Figure S10).

We further validated the cell-type distribution by TESLA in Figures 4A and 4D by performing cell-type deconvolution analysis using RCTD with annotated single-cell RNA-seq data from Ji et al.<sup>17</sup> and Tirosh et al.<sup>19</sup> as the references for the CSCC data and the melanoma data, respectively. The deconvolution results revealed similar distribution and density for B and T cells as predicted by TESLA (Figures S11 and S12). These results also demonstrate that by utilizing known marker information, TESLA can infer cell subtypes without the reliance on single-cell reference, which is advantageous compared with reference-based deconvolution methods.<sup>23</sup>

Figure 4E shows TLS density predicted by TESLA. Aggregations of lymphocytes are observed at the bottom and the right side of the melanoma region, indicating strong immune activity in the yellow and blue boxes, where we observed colocalization of B cells, CD4<sup>+</sup> T cells, and dendritic cells with a high expression of CXCL13 (Figure 4D). Because the top-region in the green box is not adjacent to the tumor region but to aggregated lymphocytes, it is annotated as a lymph node instead of a TLS. These results highlight the importance of maintaining the spatial relationships of gene expression in heterogeneous tissue as the gene signature of TLS versus normal reactive lymph nodes are highly similar, reflecting the formation of lymph-node-like architecture of TLS's associated with tumors. By identifying the neoplastic tissue core and edge (Figure 3F), together with the TLS signature (Figure 4E), TESLA can segment tumor-associated TLS versus adjacent non-neoplastic lymph node, which would have been lost in dissociated or homogenized tissue analyses, such as single-cell transcriptomics. The presence of TLS was further validated by the co-location of follicular helper T cells and B cells (Figure S13).

**Application to renal cell carcinoma**—In addition to the CSCC and cutaneous malignant melanoma data, we also analyzed a dataset<sup>44</sup> that has manual annotation and immunostaining-validated TLS information. This dataset was obtained from human clear cell renal cell carcinoma (ccRCC) primary tumors, where the TLSs were manually annotated by pathologists and further validated by triple immunofluorescence labeling (CD20/CD3/PD-1) on an immediately adjacent slide for frozen samples and CD3/CD20 labeling by immunohistochemistry for formalin-fixed paraffin-embedded (FFPE) samples. We performed TLS detection using TESLA on a tissue section in this dataset. As shown in Figure 4F, the TLS scores agree well with the two validated TLS regions. We then treated the TLS score, ranging from 0 to 1, as the probability of being a TLS and calculated the pixelwise area under the ROC curve (AUC) score using pathologist's annotation provided in the original study as ground truth. TESLA achieved an AUC of 0.905, which further demonstrates TESLA's ability to correctly locate the TLSs.

### Inferring cancer subtypes in breast cancer

Next, we show that the integration of gene expression and histology information can help infer cancer subtypes. We first analyzed a HER2+, ER+, and PR- breast cancer dataset generated by 10x Genomics.<sup>20</sup> Figure 5A summarized the cancer subtypes distribution,

which is concordant with the diagnosis. We next analyzed an annotated HER2+ human breast cancer dataset from Andersson et al.<sup>45</sup> We selected three patients H, G, and B and analyzed one tissue section from each patient. All three patients are HER2 positive and ER negative, and only patient B is PgR positive, whereas the other two are PgR negative. Based on the expression of basal markers, such as *KRT5*, *KRT7*, *KRT14*, and *KRT18*, and breast cancer markers previously reported, e.g., *ERBB2*, *GATA3*, *PIP*, and *SCGB2A2* (Figure S14), TESLA detected the tumor region, demonstrating concordance with manual annotation from the original study (Figure 5B).

Based on specific marker gene expression, we detected the HER2+, ER+, and PgR+ regions within the tumor for each tissue section. Figures 5B–5D summarized the cancer subtypes distribution and the percentages of area for HER2, ER, and PgR for the three tissue sections. We observed high expression of *ERBB2*, which encodes HER2, in the tumors of all 3 patients, presented in 76%, 94%, and 76% of tumor cells, respectively, for patients H, G, and B, confirming the diagnosis of HER2+ cancers. In addition, we detected the expression of ER and PgR in patient B (Figure 5D), and low-level expression (0%–3%) of ER and PgR in the tumors of patients H and G, indicating the high degree of intratumoral heterogeneity in these tumors, owing to the super-resolution gene expression images generated by TESLA. For patient B, although it is diagnosed to be ER– and PR+, we detected 33.6% of the tumor cells to be ER+ and only 20.0% of the tumor cells to be PR+. We contacted the authors of the original paper and they claimed that the cancer subtype classifications were based on immunohistochemistry (IHC) (and PAM50) on a separate piece of the tumor. The low proportion of PR+ tumor cells and relatively high proportion of ER+ tumor cells may be due to the difference in hormone receptors between the piece used for the Visium experiment and the piece used for IHC. These results demonstrate the power of TESLA in quantitative evaluation of tumor subtypes and the additional information it provides that is beyond the simple positive or negative annotation of a tumor.

## DISCUSSION

In this paper, we presented TESLA, a machine learning framework for multi-level tissue annotation on the histology image with pixel-level resolution in ST. By integrating information from high-resolution histology images, TESLA can impute gene expression at superpixels and fill in the missing gene expression in tissue gaps. The increased gene expression resolution makes it possible to treat gene expression data as images, which enables the integration with histological features for joint tissue segmentation and annotation of different cell types directly on the histology image with pixel-level resolution. Additionally, TESLA can detect unique structures of tumor-immune microenvironments, such as TLSs, and separate a tumor into core and edge to examine their cellular compositions, expression features, and molecular processes. Principally, TESLA can distinguish lymphoid aggregates, immature and mature TLSs, if appropriate marker genes are provided by the users, but this will also depend on the quality of the ST data, such as sequence coverage, the average number of genes detected in each spot, and dropout rate. TESLA has been evaluated on seven cancer datasets, including CSCC of the skin, cutaneous malignant melanoma, four breast cancer datasets, and ccRCC. We further conducted a benchmark evaluation based on data generated from 10x Xenium and showed that TESLA

has higher correlations with the ground truth than BayesSpace. Although our results consistently showed that TESLA can generate high-quality, super-resolution gene expression images, neither TESLA nor BayesSpace can infer single-cell-level gene expression with high accuracy.

H&E-stained histology images are routinely used in clinics for disease diagnosis. Utilizing manual annotation of histology images provided by pathologists, supervised methods have been developed to computationally annotate tissues; for example, Saltz et al.<sup>46</sup> developed a deep learning model focusing on detecting TIL using histology images in The Cancer Genome Atlas Program. However, these supervised methods require large, well-labeled training data in which expert pathologists are demanded to review a large set of images and mark detailed regions of lymphocytes and necrosis. Although pathologists' annotation is highly effective for disease diagnosis, it is qualitative and prone to both inter- and intra-observer variability, particularly when quantifying or characterizing feature-rich phenomena such as tumor-associated lymphocytic infiltrates. Additionally, existing supervised methods only utilize histology information, which limit their usefulness in studying detailed structures in TME, as histology alone cannot reveal the subtypes of lymphocytes, but they can be distinguished using gene expression.

The cell-type annotation in TESLA shares similarity with traditional clustering-based cell-type annotation, but TESLA has several advantages. First, TESLA's annotation has much higher resolution than annotation obtained from spot-level gene expression clustering. Second, in traditional clustering, each spot is assigned to a single cell type. However, for spatial barcoding-based ST data, e.g., ST and 10x Visium, each spot may contain multiple cells. Although deconvolution-based methods, such as RCTD,<sup>23</sup> can infer cell-type proportions within each spot, they cannot tell the cell identity for each cell within the spot. By contrast, through leveraging high-resolution histology image information, TESLA's annotation can provide cell-type enrichment information with pixel-level resolution. Additionally, deconvolution-based methods require an annotated single-cell reference,<sup>23,47-49</sup> which would limit their applications when single-cell references are not available.

In this paper, we only showed examples of combining histology and gene expression for cell-type annotation, but it is also possible to incorporate immunofluorescence images for cell-type-specific proteins to infer targeted cell types using TESLA. Although immunofluorescence images can be used directly to infer cell identity for each cell, artifacts in the images may render the results less reliable. When analyzing the CD3e immunofluorescence data, we noticed that the light in the original CD3e immunofluorescence image is not flat across the tissue area, which leads to brighter staining in the upper part and darker staining in the lower part of the tissue (Figure 2E). This uneven illumination results in an inaccurate estimation of CD3E abundance. By combining the gene image for *CD3E* and the CD3E immunofluorescence image into a 2-channel image as input, TESLA inferred the distribution of T cells. As shown in Figure S15, the upper edge now has a lower degree of T cell enrichment, suggesting that combining gene expression and immunofluorescence images can correct immunofluorescence image artifacts.

Although we illustrated the application of TESLA in cancer, the framework is generic and can be applied to other medical conditions as long as high-resolution histology images are available. For example, when applied to mouse posterior brain, TESLA is able to separate the “granular cell layer of the cerebellum” and the “molecular layer of the cerebellum” (Note S5), a task that is difficult to achieve in the absence of super-resolution gene expression data. As shown in Figure 2B, TESLA achieved high accuracy when imputing super-resolution gene expression for non-cancer tissues (e.g., mouse posterior brain and mouse kidney). We also demonstrate that TESLA is computationally fast and memory efficient compared with BayesSpace (Note S6). With the increasing popularity of ST in biomedical research, we expect TESLA will be an attractive tool for ST data analysis. Results from TESLA will enable researchers to increase gene expression resolution and annotate their tissues of interest with high confidence.

We are aware of the new ST technologies, e.g., 10x Xenium,<sup>25</sup> NanoString CosMx,<sup>50</sup> MERSCOPE, and Stereo-seq<sup>51</sup> etc. Although these platforms provide sub-cellular resolution, none of them are perfect. For example, Xenium, CosMx, and MERSCOPE are *in situ* hybridization based, and these platforms can only measure a few hundred genes. The lack of full transcriptome coverage will limit their applicability when the goal is to do biological discoveries. Stereo-seq, on the other hand, is next-generation sequencing based; thus, it measures gene expression for the entire transcriptome. However, Stereo-seq is not easily accessible by a typical research lab due to the lack of commercialization. In addition, the sequencing cost will be extremely high because the number of spots to be sequenced is much larger than Visium. Given these constraints, we think Stereo-seq is not a practical choice for most of the research labs. We believe that TESLA would offer a practical solution to existing Visium users, because Visium covers the entire transcriptome, has a reasonable cost, and is the most popular ST platform based on a recent survey.<sup>52</sup>

There are some specific issues in ST data that TESLA does not account for. Although it has shown promising performance, it might be overly simplistic for complex tissue types. We are currently exploring a more sophisticated approach to further improve the performance of TESLA. The main idea of this extension is to extract histology image features that account for both local and long-range dependencies of pixel intensities and use a weakly supervised machine learning method to impute super-resolution gene expression by imposing constraints provided by spot-level gene expression. Fully implementing and testing this idea would require a substantial amount of work. Given the promising performance of our current version of TESLA over existing methods, we believe that implementing and testing more sophisticated approaches is beyond the scope of the current study. Nevertheless, it is an exciting future direction that we will pursue.

## STAR★METHODS

### RESOURCE AVAILABILITY

**Lead contact**—Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jian Hu (jian.hu@emory.edu).

**Materials availability**—This study did not generate new materials.

**Data and code availability**—This paper analyzes existing, publicly available data. These accession numbers for the datasets are listed in the key resources table.

All original code has been deposited at GitHub (<https://github.com/jianhuupenn/TESLA>) and is publicly available as of the date of publication. This repository has been archived at Zenodo (<https://doi.org/10.5281/zenodo.7703260>).

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## METHOD DETAILS

**Data preprocessing**—TESLA takes spatial gene expression and histology image data as input. The spatial gene expression data contain an  $N \times G$  matrix of unique molecular identifier (UMI) counts with  $N$  spots and  $G$  genes, along with the  $(x, y)$  2-dimensional spatial coordinates of each spot. The gene expression values in each spot are normalized such that the UMI count for each gene in a given spot is divided by the total UMI count across all genes in that spot, multiplied by 10,000, and then transformed to a natural log scale.

**Super-resolution gene expression image generation**—After preprocessing, TESLA extracts the measured tissue region from the histology image and generates a super-resolution gene expression image for each gene in the tissue using histological, spatial location, and the original spot-level gene expression data. The super-resolution gene expression generation step in TESLA is based on two assumptions: 1) spots that are physically close have similar gene expression, and 2) the expression patterns for spatially variable genes are correlated with histology image features. Therefore, this enhancement step works better for genes whose expression patterns have some spatial variability as compared to non-spatially variable genes. Below we describe each step in super-resolution gene expression generation in detail.

**Superpixel generation:** TESLA first detects the contour of the captured tissue area in the histology image. When the tissue region can be easily separated from the histology image background, TESLA uses the Canny-edge-detection algorithm implemented in the python package “opencv” to draw the contour. When the histology image contains contaminated stains or tissue fragments, the Canny-edge-detection algorithm may result in an inaccurate contour. In this case, TESLA detects the boundary of the measured spots and uses the enlarged boundary as the tissue contour. After contour detection, TESLA separates the tissue region inside the contour into equal-sized squares, referred to as superpixels. The size of the superpixels can be adjusted depending on the dataset and is usually much smaller than spots captured by next generation sequencing-based ST technologies. The default size of a superpixel is  $50 \times 50$  pixels, but can be adjusted by user depending on the diameter size of each spot in the ST data. As a comparison, the diameter is  $\sim 350$  pixels for a Spatial transcriptomics spot and  $\sim 200$  pixels for a 10x Visium spot (Table S1).

**Superpixel gene expression imputation:** We aim to impute the gene expression at each superpixel using observed spot-level gene expression. The gene expression of a superpixel is

expected to be similar to that of its neighboring spots. Thus, to impute the gene expression at a given superpixel  $v$ , TESLA detects its top 10 nearest neighboring measured spots based on the Euclidean distance metric described in SpaGCN,<sup>24</sup> which considers the similarity between the superpixel  $v$  and a measured spot with respect to both physical location and histological features. We first extract the approximate region of each superpixel on the histology image and calculate the mean color value for the RGB channels,  $(r_v, g_v, b_v)$  of all pixels that fall in that region. Next, a weighted sum of the RGB value is calculated to represent the histology image features:

$$z_v = \frac{r_v \times V_r + g_v \times V_g + b_v \times V_b}{V_r + V_g + V_b},$$

where  $V_r = \text{Variance}(r_v)$ ,  $V_g = \text{Variance}(g_v)$ , and  $V_b = \text{Variance}(b_v)$  for all  $v \in V$ . Then,  $z_v$  is rescaled to the same scale as  $x_v$  and  $y_v$  as

$$z_v^* = \frac{z_v - \mu_z}{\sigma_z} \times \max(\sigma_x, \sigma_y),$$

where  $\mu_z$  is the mean of  $z_v$ ,  $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$  are the standard deviations of  $x_v$ ,  $y_v$  and  $z_v$ ; respectively. The 3D coordinates for all the measured spots are derived using the same approach, and the Euclidean distance between superpixel  $v$  and a measured spot  $u$  is calculated as

$$d(u, v) = \sqrt{(x_u - x_v)^2 + (y_u - y_v)^2 + (z_u^* - z_v^*)^2}.$$

For each superpixel  $v$ , TESLA detects its top 10 nearest neighboring measured spots, denoted as set  $N_v$ . Then, the gene expression value of the superpixel  $v$  is imputed using a weighted sum of the neighboring spots' gene expression values. The weight for a given neighboring spot  $u_i \in N_v$  is negatively associated with the distance from that spot to the superpixel and is defined as

$$w(u_i, v) = \frac{d(u_i, v)^{-2}}{\sum_{u_j \in N_v} d(u_j, v)^{-2}},$$

where  $d(u_i, v)$  is the Euclidean distance calculated using the approach described in SpaGCN. For a given gene  $g$ , its imputed expression value for superpixel  $v$  is calculated as

$$e_{g,v} = \sum_{u_j \in N_v} w(u_j, v) e_{g,u_j},$$

where  $e_{g,u_j}$  is the measured spot-level gene expression for gene  $g$  at spot  $u_j$ . For each pixel within the superpixel, its expression is imputed as the gene expression at the corresponding superpixel.

After imputation, TESLA generates a super-resolution gene expression image for each gene in the tissue with tissue gaps imputed as well. We note that the center-to-center distance between two adjacent spots in the current 10x Visium platform is 100  $\mu\text{m}$ , leaving a large portion of the tissue unmeasured for gene expression. Although BayesSpace can enhance the gene expression resolution within measured spots, regions that are not covered by spots are still left blank.

**Multi-level tissue annotation**—TESLA can annotate a tissue section at multiple levels by integrating different super-resolution gene expression images together with a histology image for joint segmentation using a convolutional neural network approach. For example, using integrated cell-type-specific marker gene images and the histology image as input, TESLA can reveal cell type distributions at the same pixel-level resolution as the histology. Additionally, TESLA can identify tumor regions if the tumor marker gene images are used as input for joint segmentation. The joint consideration of tumor marker genes and histology images enables the detection of intra tumor heterogeneity that might be missed by traditional pathology-based tumor diagnosis. For simplicity, we use cell type annotation to demonstrate the annotation method.

**Generation of meta gene image:** The joint segmentation takes both gene expression and histology images as input. A histology image has 3 color channels representing red (R), green (G), and blue (B) color intensities, while the gene expression are represented by  $K$  channels for a cell type with  $K$  marker genes. If the channels are weighted equally, the gene expression would dominate the joint segmentation results when  $K$  is larger than 3, whereas the histology image would dominate when  $K$  is less than 3. Therefore, it is important to find a balance between the histology and gene expression. Since the number of marker genes varies across cell types, we combine the  $K$  markers into one meta marker gene, which can be represented by one color channel, to make our method robust across different cell types.

Since not all marker genes are expressed, an ideal meta gene should preserve the expression pattern for at least a subset of the marker genes. Given  $K$  marker genes and a pre-specified number  $k$ , for any superpixel  $i$ , we first rank all markers' relative expression in descending order as  $\{e_{1,i}, e_{2,i}, e_{3,i}, \dots, e_{K,i}\}$ . Next, we select the top  $k$  expression values and calculate the meta gene's relative expression at superpixel  $i$  as:

$$e_{meta,i} = \min(e_{1,i}, e_{2,i}, \dots, e_{k,i}).$$

At a given superpixel, if only the top  $k - 1$  or fewer marker genes are expressed and the  $k^{\text{th}}$  marker has zero expression, the meta gene will have an estimated expression value of 0 at that superpixel. This step ensures that expression patterns that are present in less than  $k$  genes are not included in the meta gene, thus avoiding generating patterns that are less representative. Defining the meta gene's expression in this way guarantees that its expression level is high at a given superpixel if and only if all top  $k$  markers are highly expressed at that superpixel. In our analyses, we set  $k$  equal to 1 to capture the expression patterns of all marker genes.



**Tissue segmentation using convolutional neural network:** Next, we convert the RGB channels from the histology image into one gray channel. For the meta gene image, we use the gene expression at each superpixel to represent the expression for each pixel that resides within the superpixel. Since the gray and meta gene channels are the same size, we can simply stack them to create a 2-channel image, where each pixel value is normalized to  $[0, 1]$  and fed into a convolutional neural network for unsupervised segmentation<sup>53</sup> (Figure 6). Let  $M$  be the number of pixels in the 2-channel image denoted by  $I = \{v_m \in \mathbb{R}^2\}_{m=1}^M$ . A  $p$ -dimensional feature map  $\{x_m \in \mathbb{R}^p\}_{m=1}^M$  is computed from  $\{v_m\}_{m=1}^M$  through two convolutional components, each of which consists of a 2-dimensional convolution, a ReLU activation function, and a batch normalization function. Subsequently, a response map  $\{r_m = \Lambda x_m\}_{m=1}^M$  is obtained by applying a linear classifier on  $\{x_m \in \mathbb{R}^p\}_{m=1}^M$ , where  $\Lambda \in \mathbb{R}^{p \times p}$  represents the weights in the convolutional layer. The response map is then normalized to  $\{r'_m\}_{m=1}^M$ , which has zero mean and unit variance. Finally, the initial cluster label  $c_m$  for each pixel is obtained by selecting the cluster that has the maximum value in  $r'_m$ . To initialize the neural network, we start with a large initial number of clusters (default  $p = 100$ ).

The loss function  $L$  consists of a constraint on spatial continuity and a constraint on gene expression and image similarity, defined as follows:

$$L = \mu L_{con}(\{r'_m\}) + L_{feature}(\{r'_m, c_m\}),$$

where  $\mu$  represents the weight for balancing the two constraints ( $\mu = 5$  as default). Following Kim et al.,<sup>27</sup> we utilize the L1-norm of horizontal and vertical differences of the response  $\{r'_m\}_{m=1}^M$  as a spatial constraint, which is defined as

$$L_{con}(\{r'_m\}) = \sum_{w=1}^{W-1} \sum_{h=1}^{H-1} (\|r'_{w+1,h} - r'_{w,h}\|_1 + \|r'_{w,h+1} - r'_{w,h}\|_1),$$

where  $W$  and  $H$  represent the width and height of the 2-channel image, and  $r'_{w,h}$  represents the pixel value at  $(w, h)$  in the response map  $\{r'_m\}_{m=1}^M$ . By applying this spatial continuity loss, an excessive number of clusters due to complicated patterns can be suppressed. The following cross entropy loss between  $\{r'_m\}_{m=1}^M$  and  $\{c_m\}_{m=1}^M$  is calculated as the constraint on gene expression and image similarity:

$$L_{feature}(\{r'_m, c_m\}) = \sum_{m=1}^M \sum_{i=1}^p -\delta(i - c_m) \log r'_{m,i},$$

where  $\delta(\cdot)$  is an indicator function. This loss ensures that pixels assigned to the same cluster have similar meta gene expression and H&E image features. Since the input gray-scaled H&E image and meta gene image are normalized to the same scale, they contribute equally to the final segmentation.

In the iterative update procedure, cluster memberships for spatially close pixels or pixels with similar meta gene expression values are merged by considering spatial continuity and gene expression similarity. This process leads to the reduction in the number of unique clusters  $p'$ . The normalized response map vectors and cluster assignments are updated iteratively as follows. Since some clusters may have no pixels assigned to them, the number of unique clusters can be a number  $p'$  ranging from 1 to  $p$ . During each training epoch, TESLA has a cluster refinement step to merge minor clusters to its neighboring major clusters. First, clusters containing less than 1000 (default) pixels will be selected as minor clusters. Next, for each pixel in a given minor cluster, its neighboring pixels, defined as pixels located within a circle of radius 3 centered at that pixel, are collected into a set. Based on the cluster identities of all neighboring pixels of the pixels in that minor cluster, TESLA will detect the major cluster which has the most pixels in the neighboring set, referred to as the neighboring major cluster. TESLA then merges the minor cluster with its neighboring major cluster, which increases the integrity of the segmentation. Due to the cluster refinement step, the number of clusters decreases during the training process. The model will stop training when the number of clusters reaches a pre-defined threshold (default is 30). We tested the value of this threshold from 10 to 50 and found that it does not affect the target cell type detection much. We use 30 as default to allow the algorithm to assign some tissue edge and small contaminated regions as small clusters.

**Target cell type detection**—After obtaining the cluster assignments from segmentation, TESLA identifies which cluster(s) are enriched for a specified target cell type. Since the segmentation output has the same resolution as the meta gene image, TESLA can calculate the average meta gene relative expression across all pixels within each cluster and sort the clusters in descending order of this value as  $\{\bar{e}_{meta, cluster1}, \bar{e}_{meta, cluster2}, \dots, \bar{e}_{meta, clusterp}\}$ . Next, clusters whose average meta gene relative expression is greater than  $t \times \bar{e}_{meta, cluster1}$  are considered as clusters that are enriched for the target cell type. We find that  $t = 1/2$  generally generates the best performance in practice. Finally, each cluster is annotated on the histology image with color corresponding to the average meta gene expression level. This framework is flexible, where users can customize their own gene list for a cell type of interest; for example, by changing the cell type marker genes to tumor marker genes (Table S4), one can detect the tumor region in the tissue.

**Intra-tumor heterogeneity analysis**—With tumor region annotation, we can further draw the leading edge and separate the tumor into edge and core, which allows us to study intra-tumor heterogeneity. TESLA first detects the boundary of the tumor region using the Canny-edge-detection algorithm implemented in the python package “opencv”. Tumor edge is defined as the area that is within 200  $\mu\text{m}$  interface of tumor and adjacent to non-malignant tissue and the tumor core is defined as the proximal tumor area within the tumor edge.<sup>27</sup> After the tumor edge and core are identified, we can perform differential expression analysis to detect genes enriched in the edge and the core, which will help in understanding the heterogeneity inside the tumor. We performed DE analysis between pixels in the tumor core and edge using Wilcoxon rank-sum test. Genes with a false discovery rate (FDR) adjusted p-value  $< 0.05$  are first selected. To ensure only genes with enriched expression patterns in the core/edge region are selected, we further require a gene to meet the following three criteria:

1) mean log (gene expression) >0.1; 2) in/out group expression fraction ratio  $\geq 1$ , where “in group” refers to spots in the target region (e.g., tumor edge or tumor core), and “out group” refers to the remaining spots, “in group expression fraction” is the percentage of spots in the “in group” where the gene is expressed, and “out group expression fraction” is the percentage of spots in the “out group” where the gene is expressed; 3) in group expression fraction >0.5; 4) fold change of in group gene expression/out group gene expression >1.2.

**Detection of tertiary lymphoid structure**—Previous studies have suggested that TLS formation results from a complex interplay between B cells, CD4+ T cells, and dendritic cells, with reciprocal signaling between these cells mediated by chemokine CXCL13.<sup>10</sup> TESLA detects the potential location of TLS based on the colocalization of B cells, CD4+ T cells, dendritic cells, and chemokine CXCL13 (marker genes in Table S4). For any pixel  $i$ , TESLA first calculates the molecular abundance of each cell type as the mean expression across all pixels for the corresponding meta gene of that cell type. TESLA then ranks the molecular abundances of these cell types and CXCL13 in descending order as  $\{b_{1,i}, b_{2,i}, b_{3,i}, b_{4,i}\}$ . Next, the TLS score is calculated as the minimum value of the top three abundances:

$$TLS\ score_i = \min(b_{1,i}, b_{2,i}, b_{3,i}).$$

This minimum pooling from the top three values ensures that a pixel has a high TLS score if and only if at least three out of the four abundances are high. We do not perform minimum pooling using all four values because we expect the enrichment of B cells, CD4+ T cells, and dendritic cells, and CXCL13 may be closely related to TLS but not perfectly overlap with the other three cell types. Next, we annotate the whole tissue with color corresponding to the TLS score, where dense regions embedded inside the tumor with high TLS scores are identified as TLSs.

**Comparing image similarity using SSIM**—The Structural SIMilarity (SSIM) index is a method for measuring the similarity between two images. The SSIM between two images  $x$  and  $y$  of the same size is calculated as

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1) \times (2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1) \times (\sigma_x^2 + \sigma_y^2 + c_2)},$$

where  $(\mu_x, \mu_y)$  and  $(\sigma_x, \sigma_y)$  are the means and standard deviations of  $x$  and  $y$ ,  $\sigma_{xy}$  is the covariance of  $x$  and  $y$ ,  $c_1 = (0.01 \times L)^2$  and  $c_2 = (0.03 \times L)^2$  are two variables used to stabilize the division, and  $L$  is the dynamic range of the pixel-values. When calculating the SSIM between protein's immunofluorescence staining image and the super-resolution gene image, the protein staining intensity and the super-resolution gene expression are scaled to the range (0,255).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

This work was supported by the following grants: National Institutes of Health grants R01GM125301 (to M.L.), P01AG066597 (to E.B.L. and M.L.), U01CA264583 (to H.K., L.W., and M.L.), and Cancer Prevention and Research Institute of Texas (CPRIT) grant RP220101 (to H.K. and L.W.). L.W. was also supported in part by the Start-up Research fund and the Institutional Research Grant (IRG) awards provided by U.T. MD Anderson Cancer Center, the Andrew Sabin Family Fellowship provided by the Andrew Sabin Family Foundation, and the RP200385 award provided by CPRIT. We thank Dr. Kim Thrane for sharing the melanoma histology image data.

## REFERENCES

- Whiteside TL (2008). The tumor microenvironment and its role in promoting tumor growth. *Oncogene* 27, 5904–5912. 10.1038/onc.2008.271. [PubMed: 18836471]
- Quail DF, and Joyce JA (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* 19, 1423–1437. 10.1038/nm.3394. [PubMed: 24202395]
- Klemm F, and Joyce JA (2015). Microenvironmental regulation of therapeutic response in cancer. *Trends Cell Biol.* 25, 198–213. 10.1016/j.tcb.2014.11.006. [PubMed: 25540894]
- Roma-Rodrigues C, Mendes R, Baptista PV, and Fernandes AR (2019). Targeting tumor microenvironment for cancer therapy. *Int. J. Mol. Sci.* 20, 840. 10.3390/ijms20040840. [PubMed: 30781344]
- Jahanban-Esfahlan R, de la Guardia M, Ahmadi D, and Yousefi B (2018). Modulating tumor hypoxia by nanomedicine for effective cancer therapy. *J. Cell. Physiol.* 233, 2019–2031. 10.1002/jcp.25859. [PubMed: 28198007]
- Oliveira G, Stromhaug K, Klaeger S, Kula T, Frederick DT, Le PM, Forman J, Huang T, Li S, Zhang W, et al. (2021). Phenotype, specificity and avidity of antitumor CD8(+) T cells in melanoma. *Nature* 596, 119–125. 10.1038/s41586-021-03704-y. [PubMed: 34290406]
- Jing X, Yang F, Shao C, Wei K, Xie M, Shen H, and Shu Y (2019). Role of hypoxia in cancer therapy by regulating the tumor microenvironment. *Mol. Cancer* 18, 157. 10.1186/s12943-019-1089-9. [PubMed: 31711497]
- Rosenberg SA, Packard BS, Aebersold PM, Solomon D, Topalian SL, Toy ST, Simon P, Lotze MT, Yang JC, Seipp CA, et al. (1988). Use of tumor-infiltrating lymphocytes and interleukin-2 in the immunotherapy of patients with metastatic melanoma. A preliminary report. *N. Engl. J. Med.* 319, 1676–1680. 10.1056/NEJM198812223192527. [PubMed: 3264384]
- Rosenberg SA, Spiess P, and Lafreniere R (1986). A new approach to the adoptive immunotherapy of cancer with tumor-infiltrating lymphocytes. *Science* 233, 1318–1321. 10.1126/science.3489291. [PubMed: 3489291]
- Sautès-Fridman C, Petitprez F, Calderaro J, and Fridman WH (2019). Tertiary lymphoid structures in the era of cancer immunotherapy. *Nat. Rev. Cancer* 19, 307–325. 10.1038/s41568-019-0144-6. [PubMed: 31092904]
- Lewis SM, Asselin-Labat ML, Nguyen Q, Berthelet J, Tan X, Wimmer VC, Merino D, Rogers KL, and Naik SH (2021). Spatial omics and multiplexed imaging to explore cancer biology. *Nat. Methods* 18, 997–1012. 10.1038/s41592-021-01203-6. [PubMed: 34341583]
- Rao A, Barkley D, França GS, and Yanai I. (2021). Exploring tissue architecture using spatial transcriptomics. *Nature* 596, 211–220. 10.1038/s41586-021-03634-9. [PubMed: 34381231]
- Longo SK, Guo MG, Ji AL, and Khavari PA (2021). Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat. Rev. Genet.* 22, 627–644. 10.1038/s41576-021-00370-8. [PubMed: 34145435]
- Ståhl PL, Salmén F, Vickovic S, Lundmark A, Navarro JF, Magnusson J, Giacomello S, Asp M, Westholm JO, Huss M, et al. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* 353, 78–82. 10.1126/science.aaf2403. [PubMed: 27365449]
- Zhao E, Stone MR, Ren X, Guenthoer J, Smythe KS, Pulliam T, Williams SR, Uyttingco CR, Taylor SEB, Nghiem P, et al. (2021). Spatial transcriptomics at subspot resolution with BayesSpace. *Nat. Biotechnol.* 39, 1375–1384. 10.1038/s41587-021-00935-2. [PubMed: 34083791]

16. Bergenstr hle L, He B, Bergenstr hle J, Abalo X, Mirzazadeh R, Thrane K, Ji AL, Andersson A, Larsson L, Stakenborg N, et al. (2022). Super-resolved spatial transcriptomics by deep data fusion. *Nat. Biotechnol.* 40, 476–479. 10.1038/s41587-021-01075-3. [PubMed: 34845373]
17. Ji AL, Rubin AJ, Thrane K, Jiang S, Reynolds DL, Meyers RM, Guo MG, George BM, Mollbrink A, Bergenstr hle J, et al. (2020). Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. *Cell* 182, 497–514.e22. 10.1016/j.cell.2020.05.039. [PubMed: 32579974]
18. Thrane K, Eriksson H, Maaskola J, Hansson J, and Lundeberg J (2018). Spatially resolved transcriptomics enables dissection of genetic heterogeneity in Stage III cutaneous malignant melanoma. *Cancer Res.* 78, 5970–5979. 10.1158/0008-5472.CAN-18-0747. [PubMed: 30154148]
19. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, et al. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352, 189–196. 10.1126/science.aad0501. [PubMed: 27124452]
20. Dataset (2021). Human Breast Cancer (Block A Section 1). [https://support.10xgenomics.com/spatial-gene-expression/datasets/1.1.0/V1\\_Breast\\_Cancer\\_Block\\_A\\_Section\\_1](https://support.10xgenomics.com/spatial-gene-expression/datasets/1.1.0/V1_Breast_Cancer_Block_A_Section_1).
21. Dataset (2021). Mouse Brain Sagittal Anterior. [https://support.10xgenomics.com/spatial-gene-expression/datasets/1.1.0/V1\\_Mouse\\_Brain\\_Sagittal\\_Anterior](https://support.10xgenomics.com/spatial-gene-expression/datasets/1.1.0/V1_Mouse_Brain_Sagittal_Anterior).
22. Dataset (2020). Mouse Kidney Section (Coronal). <https://www.10xgenomics.com/resources/datasets/mouse-kidney-section-coronal-1-standard-1-1-0>.
23. Cable DM, Murray E, Zou LS, Goeva A, Macosko EZ, Chen F, and Irizarry RA (2022). Robust decomposition of cell type mixtures in spatial transcriptomics. *Nat. Biotechnol.* 40, 517–526. 10.1038/s41587-021-00830-w. [PubMed: 33603203]
24. Hu J, Li X, Coleman K, Schroeder A, Ma N, Irwin DJ, Lee EB, Shinohara RT, and Li M (2021). SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods* 18, 1342–1351. 10.1038/s41592-021-01255-8. [PubMed: 34711970]
25. Janesick A, Shelansky R, Gottscho AD, Wagner F, Rouault M, Beliakoff G, de Oliveira MF, Kohlway A, Abousoud J, Morrison CA, et al. (2022). High resolution mapping of the breast cancer tumor microenvironment using integrated single cell, spatial and in situ analysis of FFPE tissue. 10.1101/2022.10.06.510405.
26. Wang Z, Bovik AC, Sheikh HR, and Simoncelli EP (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13, 600–612. 10.1109/tip.2003.819861. [PubMed: 15376593]
27. Lin CM, Yu CF, Huang HY, Chen FH, Hong JH, and Chiang CS (2019). Distinct tumor microenvironment at tumor edge as a result of astrocyte activation is associated with therapeutic resistance for brain tumor. *Front. Oncol.* 9, 307. 10.3389/fonc.2019.00307. [PubMed: 31106146]
28. Jim nez-S nchez J, Bosque JJ, Jim nez Londo o GA, Molina-Garc a D, Mart nez  , P rez-Beteta J, Ortega-Sabater C, Honguero Mart nez AF, Garc a Vicente AM, Calvo GF, et al. (2021). Evolutionary dynamics at the tumor edge reveal metabolic imaging biomarkers. *Proc. Natl. Acad. Sci. USA* 118, e2018110118. 10.1073/pnas.2018110118. [PubMed: 33536339]
29. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, and Mesirov JP (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 102, 15545–15550. 10.1073/pnas.0506580102. [PubMed: 16199517]
30. Mylonis I, Simos G, and Paraskeva E (2019). Hypoxia-inducible factors and the regulation of lipid metabolism. *Cells* 8, 214. 10.3390/cells8030214. [PubMed: 30832409]
31. Munir R, Lisec J, Swinnen JV, and Zaidi N (2019). Lipid metabolism in cancer cells under metabolic stress. *Br. J. Cancer* 120, 1090–1098. 10.1038/s41416-019-0451-4. [PubMed: 31092908]
32. Garraway LA, Widlund HR, Rubin MA, Getz G, Berger AJ, Ramaswamy S, Beroukhir R, Milner DA, Granter SR, Du J, et al. (2005). Integrative genomic analyses identify MITF as a lineage survival oncogene amplified in malignant melanoma. *Nature* 436, 117–122. 10.1038/nature03664. [PubMed: 16001072]

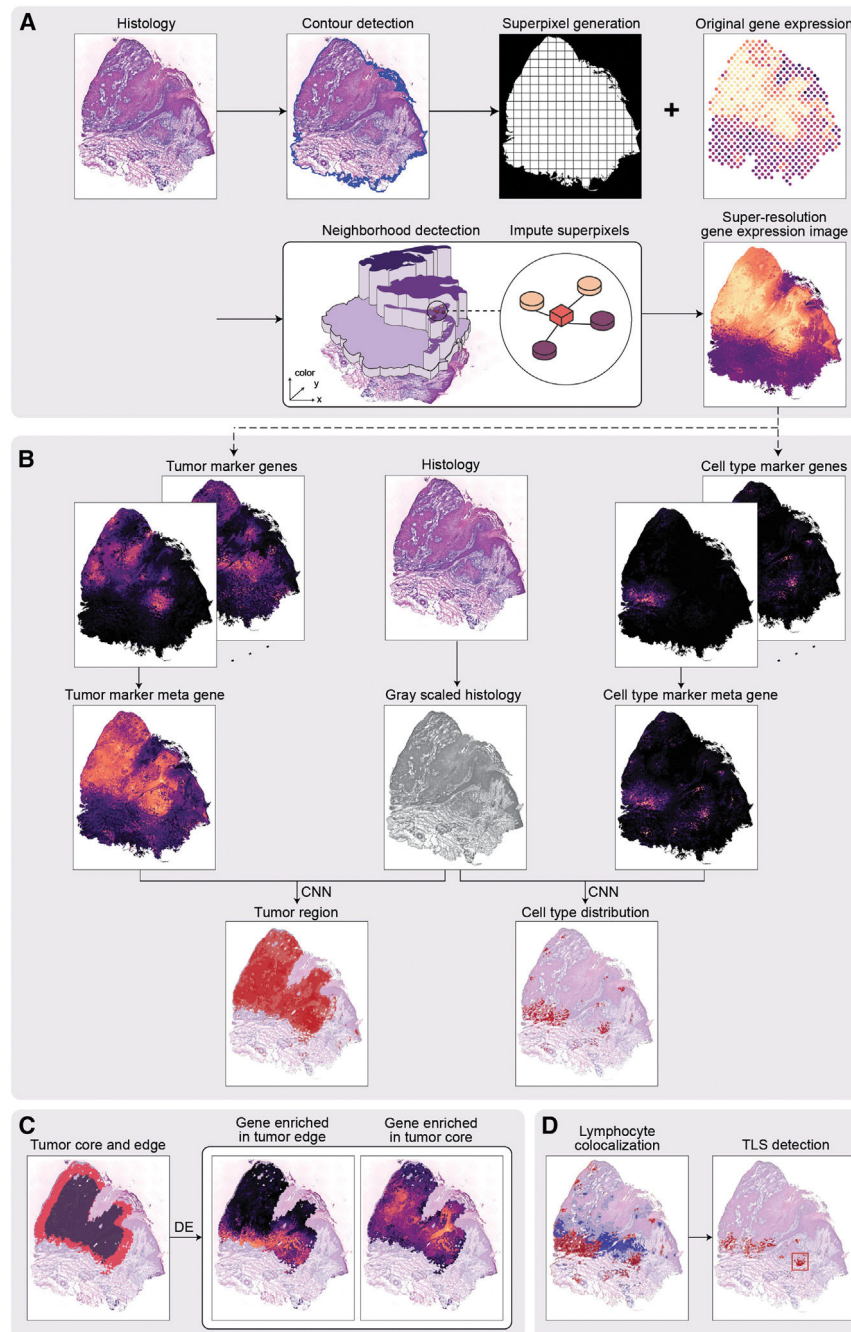
33. Buffa FM, Harris AL, West CM, and Miller CJ (2010). Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br. J. Cancer* 102, 428–435. 10.1038/sj.bjc.6605450. [PubMed: 20087356]
34. Ye Y, Hu Q, Chen H, Liang K, Yuan Y, Xiang Y, Ruan H, Zhang Z, Song A, Zhang H, et al. (2019). Characterization of hypoxia-associated molecular features to aid hypoxia-targeted therapy. *Nat. Metab.* 1, 431–444. 10.1038/s42255-019-0045-8. [PubMed: 31984309]
35. Lopez-Bergami P, Fitchman B, and Ronai Z (2008). Understanding signaling cascades in melanoma. *Photochem. Photobiol.* 84, 289–306. 10.1111/j.1751-1097.2007.00254.x. [PubMed: 18086245]
36. Paluncic J, Kovacevic Z, Jansson PJ, Kalinowski D, Merlot AM, Huang ML, Lok HC, Sahni S, Lane DJ, and Richardson DR (2016). Roads to melanoma: key pathways and emerging players in melanoma progression and oncogenic signaling. *Biochim. Biophys. Acta* 1863, 770–784. 10.1016/j.bbamcr.2016.01.025. [PubMed: 26844774]
37. Zhou Y, Song KY, and Giubellino A (2019). The role of MET in melanoma and melanocytic lesions. *Am. J. Pathol.* 189, 2138–2148. 10.1016/j.ajpath.2019.08.002. [PubMed: 31476283]
38. Pedri D, Karras P, Landeloos E, Marine JC, and Rambow F (2022). Epithelial-to-mesenchymal-like transition events in melanoma. *FEBS Journal* 289, 1352–1368. 10.1111/febs.16021. [PubMed: 33999497]
39. Moncada R, Barkley D, Wagner F, Chiodin M, Devlin JC, Baron M, Hajdu CH, Simeone DM, and Yanai I (2020). Integrating microarray--based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat. Biotechnol.* 38, 333–342. 10.1038/s41587-019-0392-8. [PubMed: 31932730]
40. Zanetti M, Xian S, Dosset M, and Carter H (2022). The unfolded protein response at the tumor-immune interface. *Front. Immunol.* 13, 823157. 10.3389/fimmu.2022.823157. [PubMed: 35237269]
41. Helmink BA, Reddy SM, Gao J, Zhang S, Basar R, Thakur R, Yizhak K, Sade-Feldman M, Blando J, Han G, et al. (2020). B cells and tertiary lymphoid structures promote immunotherapy response. *Nature* 577, 549–555. 10.1038/s41586-019-1922-8. [PubMed: 31942075]
42. Cabrita R, Lauss M, Sanna A, Donia M, Skaarup Larsen M, Mitra S, Johansson I, Phung B, Harbst K, Vallon-Christersson J, et al. (2020). Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature* 577, 561–565. 10.1038/s41586-019-1914-8. [PubMed: 31942071]
43. Petitprez F, de Reyniès A, Keung EZ, Chen TW, Sun CM, Calderaro J, Jeng YM, Hsiao LP, Lacroix L, Bougouïn A, et al. (2020). B cells are associated with survival and immunotherapy response in sarcoma. *Nature* 577, 556–560. 10.1038/s41586-019-1906-8. [PubMed: 31942077]
44. Meylan M, Petitprez F, Becht E, Bougouïn A, Pupier G, Calvez A, Giglioli I, Verkarre V, Lacroix G, Verneau J, et al. (2022). Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity* 55, 527–541.e5. 10.1016/j.immuni.2022.02.001. [PubMed: 35231421]
45. Andersson A, Larsson L, Stenbeck L, Salmén F, Ehinger A, Wu SZ, Al-Eryani G, Roden D, Swarbrick A, Borg Å, et al. (2021). Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions. *Nat. Commun.* 12, 6012. 10.1038/s41467-021-26271-2. [PubMed: 34650042]
46. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, Samaras D, Shroyer KR, Zhao T, Batiste R, et al. (2018). Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. *Cell Rep.* 23, 181–193.e7. 10.1016/j.celrep.2018.03.086. [PubMed: 29617659]
47. Andersson A, Bergensträhle J, Asp M, Bergensträhle L, Jurek A, Fernández Navarro J, and Lundeberg J (2020). Single-cell and spatial transcriptomics enables probabilistic inference of cell type topography. *Commun. Biol.* 3, 565. 10.1038/s42003-020-01247-y. [PubMed: 33037292]
48. Elosua-Bayes M, Nieto P, Mereu E, Gut I, and Heyn H (2021). SPOTlight: seeded NMF regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes. *Nucleic Acids Res.* 49, e50. 10.1093/nar/gkab043. [PubMed: 33544846]
49. Dong R, and Yuan GC (2021). SpatialDWLS: accurate deconvolution of spatial transcriptomic data. *Genome Biol.* 22, 145. 10.1186/s13059-021-02362-7. [PubMed: 33971932]

50. He S, Bhatt R, Brown C, Brown EA, Buhr DL, Chantranuvatana K, Danaher P, Dunaway D, Garrison RG, Geiss G, et al. (2022). Highplex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat. Biotechnol.* 40, 1794–1806. 10.1038/s41587-022-01483-z. [PubMed: 36203011]
51. Chen A, Liao S, Cheng M, Ma K, Wu L, Lai Y, Qiu X, Yang J, Xu J, Hao S, et al. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* 185, 1777–1792.e21. 10.1016/j.cell.2022.04.003. [PubMed: 35512705]
52. Moses L, and Pachter L (2022). Museum of spatial transcriptomics. *Nat. Methods* 19, 534–546. 10.1038/s41592-022-01409-2. [PubMed: 35273392]
53. Kim W, Kanazaki A, and Tanaka M (2020). Unsupervised learning of image segmentation based on differentiable feature clustering. *IEEE Trans. Image Process.* 29, 8055–8068.

### Highlights

- TESLA enhances gene expression to super resolution using histology images
- TESLA integrates gene expression and histology to annotate tumor and lymphoid cells
- TESLA characterizes intra-tumor heterogeneity by separating tumor into core and edge
- Tertiary lymphoid structures can be identified by colocalization analysis in TESLA





**Figure 1. Workflow of TESLA**

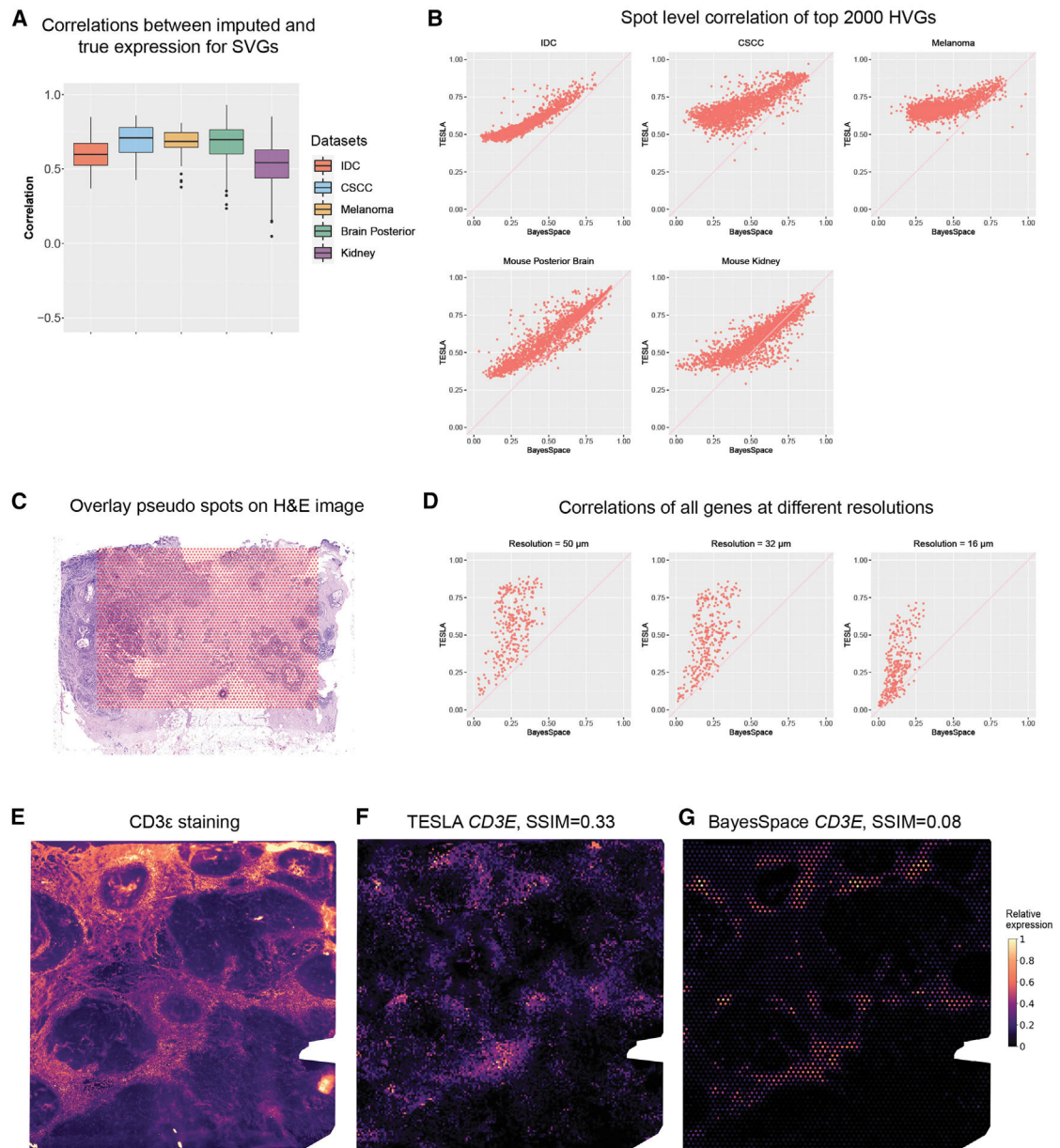
(A) TESLA starts from generating superpixels to cover the entire tissue area. The gene expression of each superpixel is imputed by the weighted sum of its nearest measured spots in a 3D space, defined by the spatial coordinates  $(x, y)$  and the 3rd dimensional coordinate  $z$ , obtained from the RGB values in the histology image.

(B) TESLA combines a set of tumor/cell-type marker genes into a meta gene and stacks the meta gene image with the gray-scaled histology image into a two-channel image, which is fed into a convolutional neural network (details shown in Figure 6) for segmentation.

Tumor region/cell-type distribution can be displayed on the histology image based on the segmentation results.

(C) TESLA can further separate the tumor region into edge and core. Downstream differential expression analysis is performed to identify core or edge enriched genes.

(D) Based on the colocalization of specific lymphoid cell types detected in (B), TESLA can identify TLS on the histology image.



**Figure 2. Evaluation of super-resolution gene images from TESLA**

(A) Boxplot of Pearson correlations between the original gene expression and TESLA's imputed gene expression for masked spots for SVGs detected by SpaGCN. Each point in the boxplot represents a gene. SpaGCN detected 68 SVGs ( $n = 68$ ) from the IDC dataset, 85 SVGs ( $n = 85$ ) from the CSCC dataset and 30 SVGs ( $n = 30$ ) SVGs from the cutaneous malignant melanoma dataset. The lower and upper hinges correspond to the first and third quartiles, and the center refers to the median value. The upper (lower) whiskers extend from the hinge to the largest (smallest) value no further (at most) than  $1.5\times$  interquartile range from the hinge. Data beyond the end of the whiskers are plotted individually.

(B) Scatter plots of Pearson correlations between the original spot-level gene expression and “spot-level” gene expression obtained from the enhanced expression generated by TESLA

and BayesSpace for the top 2,000 highly variable genes selected by BayesSpace for the five datasets. Each dot in the plot represents a gene.

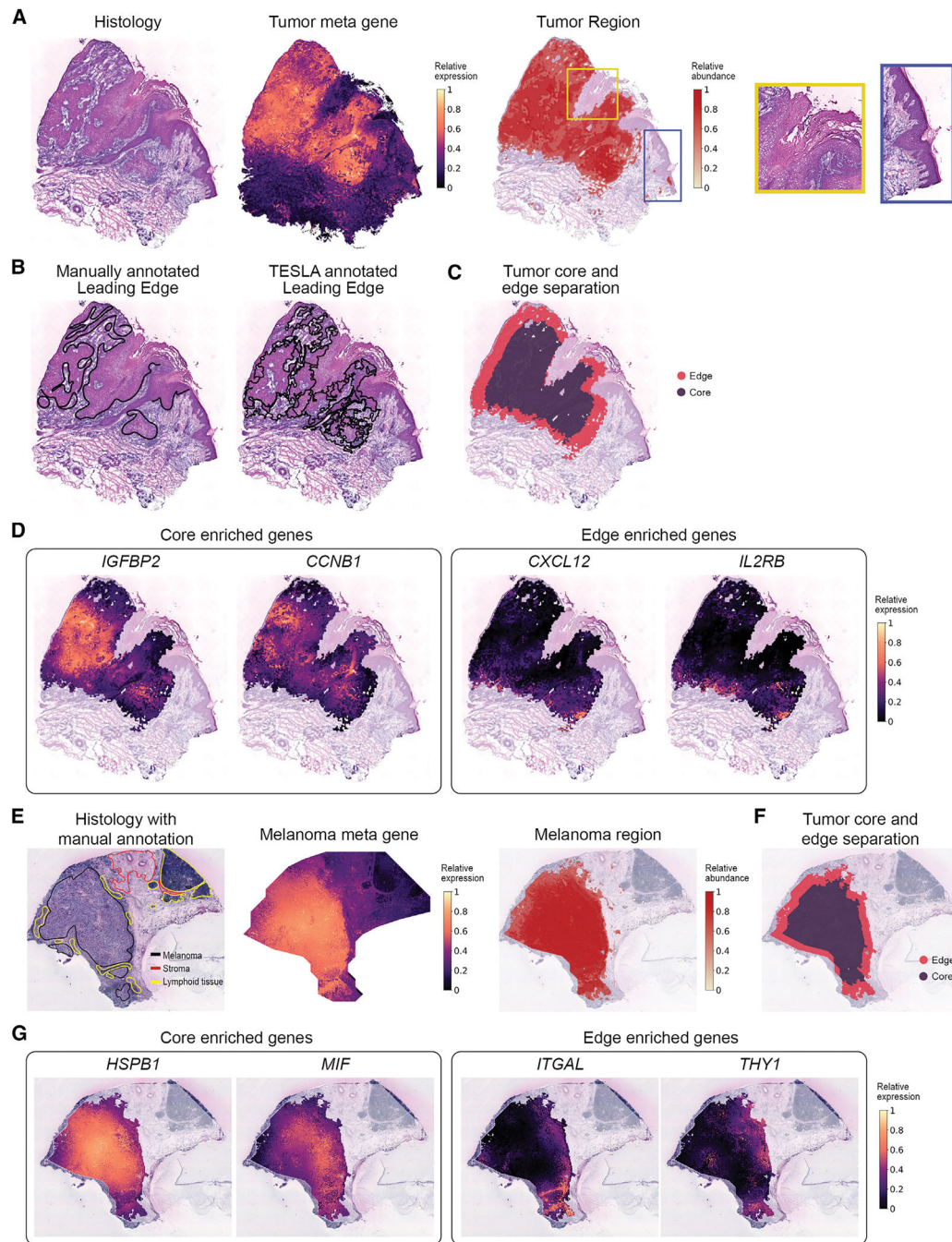
(C) Spot layout of the pseudo-Visium data derived from 10× Xenium breast cancer dataset, in which the pseudo spots are overlaid on the H&E image.

(D) Scatter plots of Pearson correlations between observed gene expression in Xenium and enhanced gene expression generated by TESLA and BayesSpace at different resolutions for all 313 genes in Xenium. Each dot in the plot represents a gene.

(E) Intensity of the CD3ε immunofluorescence staining. Intensity was scaled to (0, 1) for visualization.

(F) Super-resolution gene expression image of *CD3E* from TESLA.

(G) Super-resolution gene expression image of *CD3E* from BayesSpace.



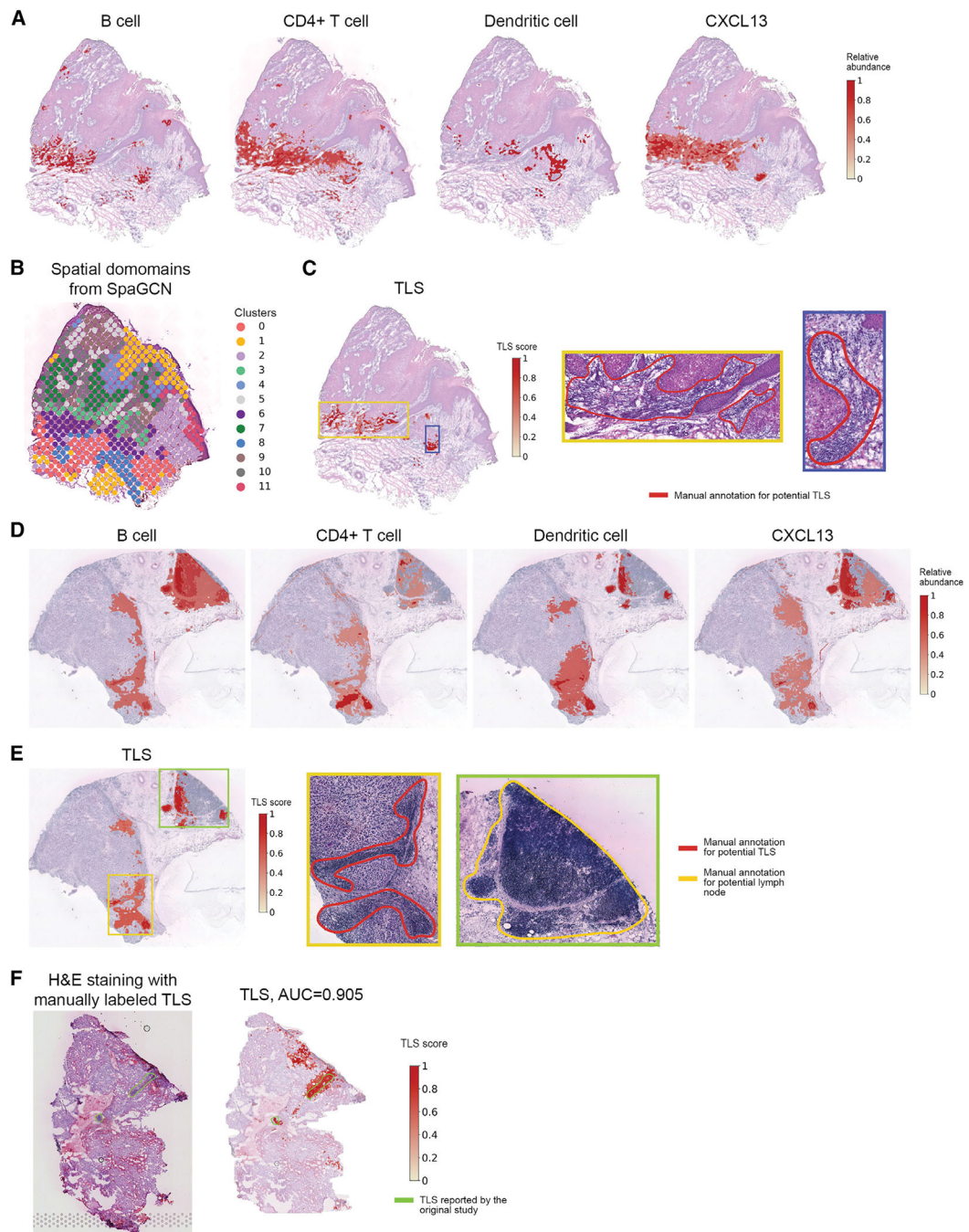
**Figure 3. TESLA characterizes intra-tumor heterogeneity in a CSCC sample and a melanoma sample**

(A) Histology image (left), tumor meta marker gene (middle) and the tumor region detected by TESLA (right) from the CSCC tissue section. The yellow and blue boxes mark two regions with dark color on histology image but not detected as tumor region.

(B) Pathologist's manually annotated leading edge from the original study and the leading edge drawn by TESLA for the CSCC tissue section.

(C) Tumor core and edge separation obtained by TESLA for the CSCC tissue section.

- (D) Examples of genes enriched in the tumor core (*IGFBP2*, *CCNB1*) and tumor edge (*CXCL12*, *IL2RB*) detected by TESLA for the CSCC tissue section.
- (E) Pathologist annotated histology image from original study, melanoma meta marker gene expression, and the melanoma region detected by TESLA.
- (F) Tumor core and edge separation of the melanoma region by TESLA.
- (G) Examples of the tumor core (*HSPB1*, *MIF*) and edge (*ITGAL*, *THY1*) enriched genes detected by TESLA for the melanoma tissue section.



**Figure 4. TESLA identifies lymphoid cell types and TLS with pixel-level resolution in the CSCC and melanoma samples**

(A) The distribution of B cells, CD4<sup>+</sup> T cells, dendritic cells, and chemokine CXCL13 at pixel-level resolution characterized by TESLA in the CSCC tissue section.

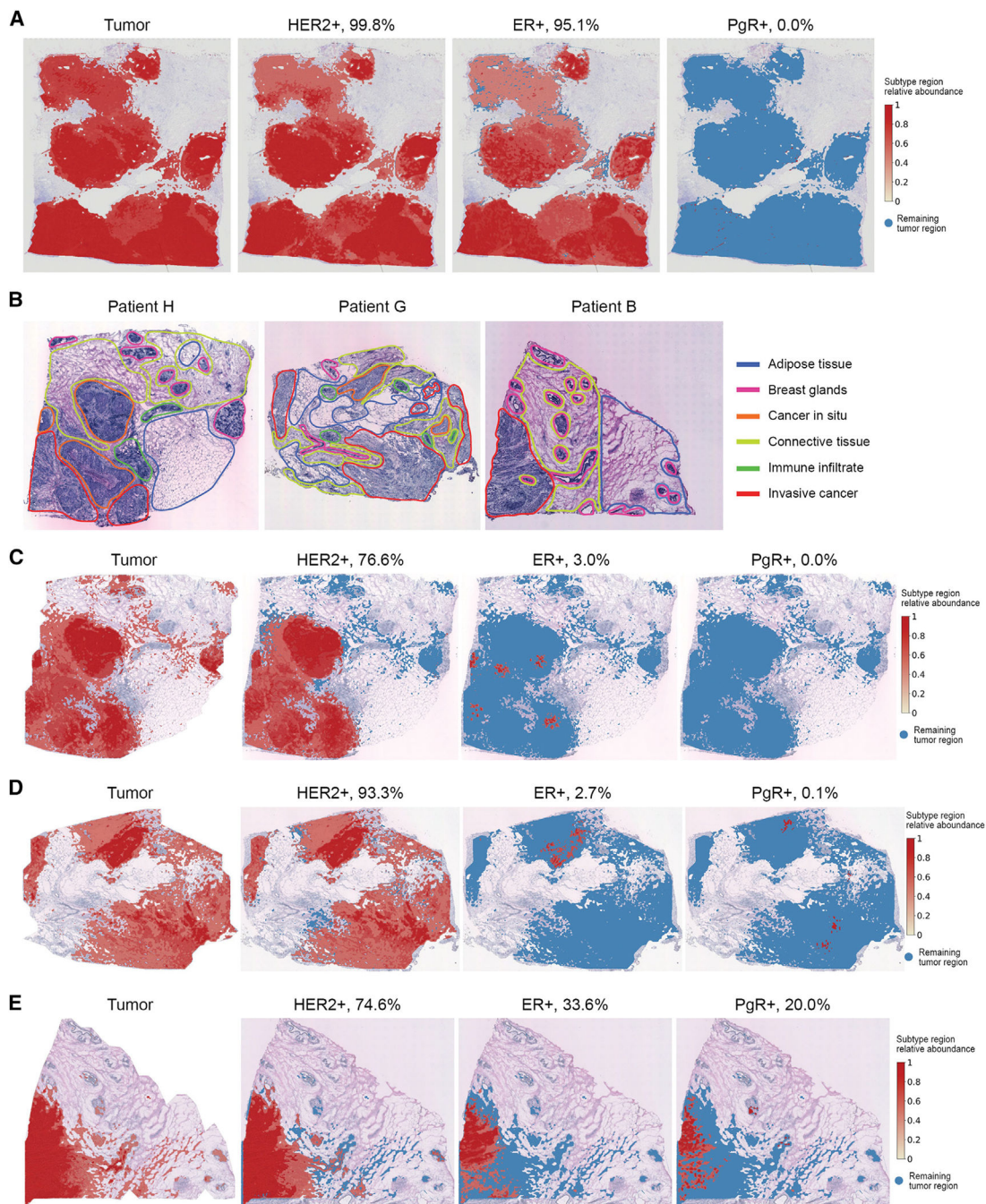
(B) Clustering of spots of the CSCC tissue section using SpaGCN.

(C) TLS scores calculated by TESLA for the cutaneous CSCC tissue section. The blue box marks a highly possible TLS, whereas the green box marks a region with aggregated lymphocytes but not forming a TLS.

(D) The distribution of Bcells, CD4<sup>+</sup> Tcells, dendritic cells, and chemokine CXCL 13 at pixel-level resolution characterized by TESLA in the melanoma tissue section. (E) TLS scores calculated by TESLA for the melanoma tissue section. The yellow and green boxes mark two potential TLSs, and the blue box marks a lymphoid node apart from the tumor region.

(F) H&E staining and TLS scores calculated by TESLA for the clear cell renal cell carcinoma (ccRCC) primary tumor data. The yellow line outlines the true TLS annotated by pathologists, which was further validated by triple immunofluorescence labeling (CD20/CD3/PD-1) on an immediately adjacent slide for frozen sample and CD3/CD20 labeling by immunohistochemistry for FFPE sample.



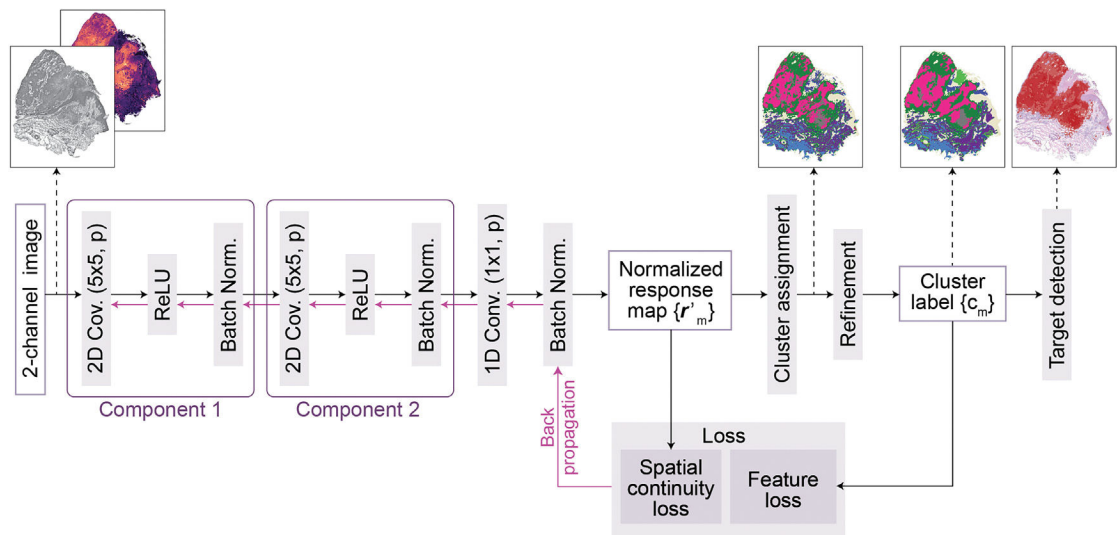


**Figure 5. TESLA detects subtypes of breast cancer samples**

(A) Tumor region and tumor subtype regions (HER2+, ER+, and PgR+) detected by TESLA in the HER2+, ER+, and PR<sup>-</sup> breast cancer dataset from 10x Genomics.

(B) Histology images with pathologist's manual annotation from the original HER2+ breast cancer study.

(B–D) Tumor region and tumor subtype regions (HER2+, ER+, and PgR+) detected by TESLA in the three breast cancer tissue sections from patients H, G, and B in the HER2+ breast cancer study.



**Figure 6. Architecture of the CNN for image segmentation in TESLA**

A 2-channel image is fed into the CNN to extract deep features through two convolutional components, each of which consists of a two-dimensional convolution ( $5 \times 5, p$ ), a ReLU activation function, and a batch normalization function. Subsequently, the response vectors of the features in  $p$ -dimensional cluster space are calculated through a one-dimensional convolutional layer ( $1 \times 1, p$ ) and normalized to  $\{r'_m\}_{m=1}^M$  across the axes of the cluster space using a batch normalization function. The normalized response map  $\{r'_m\}_{m=1}^M$  is used to calculate the spatial continuity loss. Further, raw cluster labels are determined by assigning the cluster to  $\{r'_m\}_{m=1}^M$  using an argmax function. Next, the raw cluster labels are refined to  $\{c_m\}_{m=1}^M$  and used as pseudo labels to compute the gene expression loss. Finally, the spatial continuity loss, as well as the gene expression loss, are combined and backpropagated. After training, this CNN can segment input image into different clusters, and TESLA next identifies the clusters corresponding to the target region.

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Deposited data</b>		
Human Invasive ductal carcinoma spatial transcriptomics data	10x Genomics	10x Data: <a href="https://support.10xgenomics.com/spatialgene-expression/datasets/1.2.0/V1_Human_Invasive_Ductal_Carcinoma">https://support.10xgenomics.com/spatialgene-expression/datasets/1.2.0/V1_Human_Invasive_Ductal_Carcinoma</a>
Human Cutaneous squamous cell carcinoma spatial transcriptomics and single-cell RNA sequencing data	Ji et al. <sup>17</sup>	GEO: GSE144240
Human cutaneous malignant melanoma spatial transcriptomics data	Thrane et al. <sup>18</sup>	Spatial Research Lab Data: <a href="https://www.spatialresearch.org/resources-published-datasets/doi-10-1158-0008-5472-can-18-0747/">https://www.spatialresearch.org/resources-published-datasets/doi-10-1158-0008-5472-can-18-0747/</a>
Human melanoma tumor single-cell RNA sequencing data	Tirosh et al. <sup>19</sup>	GEO: GSE72056
Human HER2+, ER+ and PR- breast cancer spatial transcriptomics data	10x Genomics	10x Data: <a href="https://support.10xgenomics.com/spatialgene-expression/datasets/1.1.0/V1_Breast_Cancer_Block_A_Section_1">https://support.10xgenomics.com/spatialgene-expression/datasets/1.1.0/V1_Breast_Cancer_Block_A_Section_1</a>
Human HER2+ breast tumor spatial transcriptomics data	Andersson et al. <sup>45</sup>	GitHub Data: <a href="https://github.com/almaan/her2st">https://github.com/almaan/her2st</a>
Mouse Posterior brain (sagittal) spatial transcriptomics data	10x Genomics	10x Data: <a href="https://support.10xgenomics.com/spatialgene-expression/datasets/">https://support.10xgenomics.com/spatialgene-expression/datasets/</a>
Mouse kidney spatial transcriptomics data	10x Genomics	10x Data: <a href="https://support.10xgenomics.com/spatialgene-expression/datasets/">https://support.10xgenomics.com/spatialgene-expression/datasets/</a>
Human clear cell renal cell carcinoma primary tumors spatial transcriptomics data	Meylan et al. <sup>44</sup>	GEO: GSE175540
Human breast cancer spatial transcriptomics data	10x Genomics	10x Data: <a href="https://www.10xgenomics.com/products/xenium-in-situ/preview-dataset-human-breast">https://www.10xgenomics.com/products/xenium-in-situ/preview-dataset-human-breast</a>
<b>Software</b>		
TESLA	This paper	<a href="https://doi.org/10.5281/zenodo.7703260">https://doi.org/10.5281/zenodo.7703260</a>
BayesSpace	Zhao et al. <sup>15</sup>	<a href="https://github.com/edward130603/BayesSpace">https://github.com/edward130603/BayesSpace</a>
SpaGCN	Hu et al. <sup>24</sup>	<a href="https://github.com/jianhuupenn/SpaGCN">https://github.com/jianhuupenn/SpaGCN</a>
RCTD	Cable et al. <sup>23</sup>	<a href="https://github.com/dmcable/RCTD">https://github.com/dmcable/RCTD</a>