Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____   04/10/2025
Xueqi Shen   Date

Systematic Evaluation of Spatial Transcriptomics Alignment Methods

By

Xueqi Shen
Master of Science in Public Health

Department of Biostatistics and Bioinformatics

_____

Jian Hu, Ph.D
(Thesis Advisor)

_____

Steve Qin, Ph.D
(Thesis Reader)

Systematic Evaluation of Spatial Transcriptomics Alignment Methods

By

Xueqi Shen

B.S.

University of California, Santa Barbara

2023

Thesis Advisor: Jian Hu, Ph.D

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Department of Biostatistics and Bioinformatics

2025

Abstract

Systematic Evaluation of Spatial Transcriptomics Alignment Methods

By Xueqi Shen

Introduction: Accurate alignment of tissue sections is crucial for integrating spatial transcriptomics with histological analyses. Current alignment methods often prioritize a single modality—either gene expression or morphological features—which can be insufficient, especially for distantly spaced sections. This study evaluates two alignment strategies, PASTE and VALIS, to determine their effectiveness in aligning HER2+ breast cancer sections characterized by heterogeneous spatial structures.

Methods: We applied PASTE and VALIS to align serial sections from HER2+ breast cancer tissues. To assess alignment effectiveness, we developed a per-spot alignment cost function incorporating bidirectional nearest-neighbor relationships and penalties for misalignments due to reflection, scaling and rotation. This function quantitatively evaluates alignment performance across both gene expression and morphological modalities.

Results: Our analysis revealed that PASTE frequently produced reflections, likely due to its gene expression minimization process, leading to misalignments. Conversely, VALIS exhibited scaling and rotation inconsistencies, particularly in regions where spatial landmarks were less distinct. The per-spot alignment cost function effectively quantified these discrepancies, highlighting the limitations of both methods in accurately aligning sections with weak spatial landmarks.

Conclusion: The study highlights key limitations in current alignment strategies when applied to tissue sections with high gene expression heterogeneity or lacking clear spatial landmarks. The findings emphasize the necessity for advanced alignment approaches capable of integrating multiple modalities and addressing spatial inconsistencies to enhance the accuracy and reliability of spatial transcriptomics analyses.

Systematic Evaluation of Spatial Transcriptomics Alignment Methods

By

Xueqi Shen

B.S.

University of California, Santa Barbara

2023

Thesis Advisor: Jian Hu, Ph.D

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Department of Biostatistics and Bioinformatics

2025

Acknowledgements

I would like to thank the Department of Biostatistics and Bioinformatics for all the support. In particular, I would like to thank all the professors and faculty members who have taught me over the past two years—you are incredibly knowledgeable and always responsive to my questions.

I am especially grateful to my advisor, Dr. Jian Hu, who guided me through the entire process. Thank you for encouraging me to think more deeply and for mentoring me in how to work like a researcher.

Finally, I would like to thank my family for their unwavering support. I could not have come this far without them.

Table of Contents

Introduction

Advancements in spatial transcriptomics (ST) have revolutionized the characterization of spatially defined tissue regions by enabling spatially resolved gene expression analysis[1]. These data enable mapping the intricate micro-landscapes of tissues and deeper understanding of the cellular dynamics underlying health and disease. In recent years, several biotechnology companies released platforms like 10X Visium[2], 10X Xenium[3], MERFISH[4] and STARmap[5], to accommodate diverse research needs and budgets for ST. With further reductions in the cost per tissue section, current studies have increasingly generated multiple ST datasets from consecutive tissue sections of a single sample, expanding the availability of data for analysis.

Despite increased availability and spatial continuity, applications remain limited, and tasks such as aligning a series of consecutive sections of the same tissue sample remain complex due to variations in tissue placement, local deformation, and sampling heterogeneity across tissue sections. This challenge necessitates the development of robust computational alignment methods. Resolving this issue would pave the way for broader applications of ST datasets, such as three-dimensional tissue reconstruction from limited sections and the transfer of pathological labels from one section to another to reduce the cost of annotation.

To address this challenge, researchers have developed novel algorithms that are founded upon different modalities of available ST data. For instance, Probabilistic Alignment of Spatial Transcriptomics Experiments (PASTE)[6], first published in 2022, incorporates both gene expression count data and spatial coordinates data. This method performs pairwise alignments followed by 3D representation reconstruction to generate a consensus tissue representation. By integrating spatial information with expression profiles, PASTE allows for high-resolution tissue

reconstruction and facilitates downstream analyses, including cell-type identification and spatial expression pattern derivation. Ideally, PASTE should be able to effectively align adjacent tissue sections, capturing both molecular and spatial continuity.

Complementing ST data-based methods, Virtual Alignment of Pathology Image Series (VALIS)[7] focuses on morphology-based alignment using whole-slide image[8] (WSI) registration. Hematoxylin and eosin (H&E) staining[9] is the common technique for tissue section staining in ST. For this reason, the H&E-stained image of the tissue section is almost always available alongside the generation of ST data, making this method both feasible and powerful for our purpose. VALIS employs scalable, fully automated pipelines to align tissue sections based solely on morphological features, addressing challenges related to deformation and large-scale data processing. After aligning histological images, VALIS enables coordinate warping, allowing the conversion of alignments to spot-levels, which can then be aligned with the original ST data. In this way, VALIS serves as a perfect example for validating alignment strategies in our evaluation, without relying on other modalities. Having introduced PASTE and VALIS as key methods for ST alignment, this study systematically evaluates two alignment strategies: whole-section alignment based on ST using PASTE and morphology-based WSI alignment using VALIS.

To facilitate the evaluation process with PASTE, Integrative and Reference-Informed Segmentation (IRIS)[10] is incorporated. IRIS segments tissues based on cell type composition, leveraging single-cell RNA sequencing data to achieve high accuracy in detecting spatial domains. This method assists the validation of alignment results for biologically relevant spatial domains. Through this strategy, we can gain deeper insights into the alignment results of PASTE in addition to statistical metrics. For a fair and comprehensive evaluation, we propose a unified workflow that integrates multiple modalities, enabling a more thorough assessment of alignment. This approach

incorporates various aspects of tissue sections (**Figure 1**), including gene expression data and image feature data, ensuring a robust and balanced evaluation. For instance, in the case of PASTE, although the algorithm may produce an unlikely alignment—visually observed—it still outputs this result because it achieves the lowest gene expression cost according to its model. Such cases emphasize the importance of addressing extreme alignments. The proposed workflow is particularly valuable for tackling issues like these, where minimizing gene expression cost could result in inaccurate mapping, such as large rotation angles or reflection. Ultimately, this method provides a more accurate and interpretable measure of alignment quality.

By validating gene expression- and morphology-based alignment strategies with our proposed method, this study highlights their strengths and limitations under varying conditions. The findings provide a roadmap for developing hybrid alignment frameworks that integrate data from multiple modalities, aiming to improve tissue section alignment and facilitate 3D tissue reconstruction.
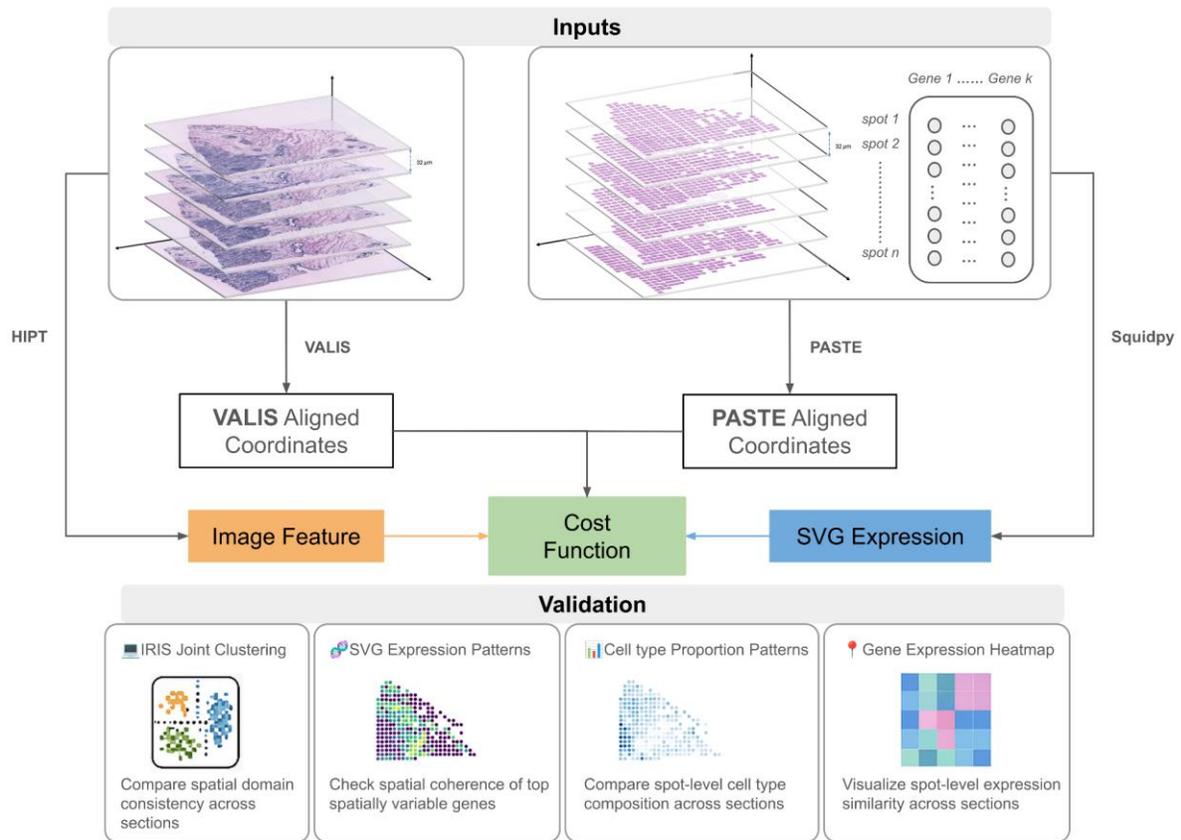
Figure 1. Overview of the alignment evaluation workflow. H&E-stained images and gene expression data are aligned using VALIS and PASTE, respectively. Spot-level features are extracted and used to compute alignment costs. Validation is performed through spatial domain detection, SVG patterns, cell type composition, and gene expression heatmaps.

Methods

Dataset used in this study

In this study, we utilized a publicly available HER2-positive breast cancer dataset established by Anderson et al. in 2021[11]. This dataset comprises eight samples, each derived from a different individual. Within each sample, three to six consecutive cryosections were generated, with a 32μm gap between adjacent sections. The dataset has been preprocessed using a standard ST pipeline and includes corresponding H&E-stained images for each section, which provide additional morphological context for our analysis.

To systematically evaluate the performance of different alignment strategies, we selected two representative sets of cryosections. ST data from two sets of tissue sections were analyzed, referred to as series B (B1–B6) and series D (D1–D6). Each section contained between 270 and 315 spatial spots, with series B ranging from 270 to 298 spots and series D from 301 to 315 spots. The number of detected genes per section ranged from 15,109 to 15,387 in series B, and from 15,396 to 15,666 in series D (**Table S1**). Overall, slightly higher spot counts and gene detection levels were observed in series D compared to series B.

The first set, consisting of six cryosections from sample B (B1 to B6), was characterized by a triangular shape and well-defined tissue regions in the H&E images. This set was used for whole-series alignment analysis. Additionally, B4 and B6 were aligned using B1 as the reference section to investigate the effect of increasing section distance on pairwise alignment. To introduce a more challenging scenario, the second set of cryosections, D1 to D6, was included. These sections were rectangular in shape and lacked distinct spatial patterns in their H&E images. The alignment

approach was based on the assumption that adjacent tissue sections share the most similar gene expression profiles and morphological features.

By analyzing these three scenarios—(i) full-sequence pairwise alignments (B1 to B6), (ii) distant section alignments (B1 and B4, B1 and B6), and (iii) pairwise alignments of sections with much less distinct morphology (D1 to D6)—we aimed to assess the robustness and limitations of each alignment approach under varying tissue structures and spatial complexities.

PASTE for Pairwise ST Section Alignment

In the sequential scenarios, we followed the default pairwise alignment workflow of PASTE. Specifically, the pairwise alignment was performed between adjacent sections in sequence, meaning section 1 and section 2 were aligned first, followed by alignment of section 2 and section 3, continuing until the last section in the sequence was reached.

For each section, two inputs were provided: the spot-level spatial coordinates, with $n \times 2$ dimensions, and the gene expression count matrix, with $n \times m$ dimensions, where n represents the number of spots and m is the number of genes. These data were integrated into an Anndata object, which then underwent pre-processing. The count data were normalized to a total count of 10,000 per spot, followed by the addition of a pseudo-count of 1 to each cell in the matrix to prevent infinite results, and then log-transformation. These preprocessing steps were necessary to correct for differences in expression levels across spots and to ameliorate the heavy tail of the count distribution, thereby making the data more comparable across spots. Finally, the gene expression patterns were scaled to have a mean of zero and a standard deviation of one with respect to each gene. Thus, it guaranteed that datasets were brought to the same scale, making them consistent across tissue sections.

In the pairwise alignment, the first section in each pair was chosen as the reference. It underwent translation only, with its spatial coordinates centered relative to the section's center after alignment. The second section was then aligned to the reference, and its spot-level spatial coordinates were transformed and projected onto the reference coordinate. After alignment, the spot-level spatial coordinates of each section were obtained and subsequently used to compute the cost.

VALIS for Pairwise WSI Registration

For a fair comparison, we registered the H&E-stained images following the same principle as with PASTE. Specifically, for a given set of tissue sections with a known order, we registered the images pairwise—aligning section 1 with section 2, followed by section 2 with section 3, and so on. Both images first underwent pre-processing using a specific function for H&E images implemented in VALIS. During the pre-processing, either the hematoxylin-stained or eosin-stained patterns which helped to better extract landmarks.

In each pairwise WSI registration, the first section in each pairwise WSI registration was chosen as the reference, meaning it underwent no transformation, and its spot-level spatial coordinates were preserved after the registration. The second image in each pair was registered to the reference through rotation, transformation, and additional local alignment. As a result, the corresponding spot-level spatial coordinates of the second image were warped according to the registration outcome. Similarly, the resulting spatial coordinates were then applied to compute the cost.

Evaluation metric: Per-Spot Cost

We proposed a novel cost function to assess the alignment between ST sections using both spatial and gene expression data. This methodology was designed to evaluate the quality of pairwise tissue section alignment while accounting for potential geometric transformations, such as rotation and scaling, as well as feature-wise variations in gene expression. The cost function incorporated several critical components, each contributing to a comprehensive assessment of alignment accuracy.

Since VALIS does not provide image features as output, we employed HIPT to extract raw image features from the H&E images used in our analysis. For each section, only the corresponding H&E image was required for the initial feature extraction step in HIPT. The extracted image features were obtained at the pixel level. To derive spot-level image features, we utilized the spot-level spatial coordinates and computed the mean image feature values within each spot area, which measured 280×280 pixels.

After assessing reflection in PASTE and obtaining the necessary inputs for the evaluation metric, we proceeded to the first step in our approach: assessing the global alignment of the query section (the second section in the pair) with the reference section (the first section in the pair). This assessment was based on the nearest Euclidean distance between spots in the query section and their closest neighbors in the reference section. We applied a threshold distance r to determine whether the aligned spot pairs met a predefined spatial proximity criterion. This threshold accounted for potential misalignments caused by image transformations such as rotation and scaling (i.e., spatial changes due to resizing). By establishing this threshold, we ensured that only

spatially close spot pairs were considered aligned, while outlier pairs—those that could distort cost calculations—were excluded and later penalized.

$$C = \begin{cases} \frac{1}{N_1+N_2} \left(\sum_{j=1}^{J} \|X_{1j} - X_{2j}\|^2 + \sum_{k=1}^{K} \|X_{2k} - X_{1k}\|^2 \right), & \text{if 90\% spots have a neighbor in r} \\ \frac{1}{N_1+N_2} \left(2\sum_{i=1}^{n} \|X_{1i} - X_{2i}\|^2 + \sum_{j'=1}^{J'} \|X_{1j'}\|^2 + \sum_{k'=1}^{K'} \|X_{2k'}\|^2 \right), & \text{otherwise} \end{cases}$$

The per-spot cost C quantifies alignment performance based on the spatial proximity of corresponding spots across sections. Here, N_1 and N_2 denote the number of spots in section 1 and section 2, respectively, while X represents the input modality, which can be either gene expression or image features. The formulation of C depends on whether 90% of spots have a nearest neighbor within distance r. If this condition is met, the cost is computed using the squared Euclidean distances between spots with their nearest neighbors in both sections (X_{1j} in section 1 and X_{2k} in section 2). Otherwise, the cost is computed using the squared Euclidean distances between spots with bidirectional nearest neighbors (X_{1i} and X_{2i}), with an additional penalty is applied for spots without a nearest neighbor within r, where $X_{1j'}$ and $X_{2k'}$ represent such spots in sections 1 and 2, respectively.

To incorporate gene expression data into the alignment process, we focused on spatially variable genes (SVGs)—genes whose expression patterns vary significantly across the spatial domains of tissue sections. SVGs are often biologically relevant as they reflect tissue-specific processes or regions with distinct molecular characteristics. We selected a set of top SVGs that captured the most spatially relevant gene expression data, thereby reducing dimensionality. This reduction retained only the most informative spatial features, enhancing the robustness and interpretability of our alignment assessment. By focusing on SVGs, we ensured that our evaluation was driven by meaningful spatial patterns rather than background noise.

For a balanced evaluation, all features needed to contribute equally to the cost function. Therefore, we standardized each feature across all spots, scaling them to have a mean of zero and a standard deviation of one. Standardization was crucial as it prevented any single feature from disproportionately influencing the overall cost due to differences in its original scale or range. By standardizing the features, we ensured that each had equal weight in the final cost calculation, allowing for a fair comparison across different alignments.

The final output of our cost function was the per-spot cost based on pairwise alignment, calculated by averaging the total alignment cost over the number of spots involved. It was essential not to use the total cost for comparison, as section sizes varied, and directly comparing total costs across sections with different spot counts could be misleading. By averaging the cost per spot, we ensured that alignment quality was assessed independently of section size, allowing for fair comparisons across different alignments and facilitating the identification of the most effective alignment method, regardless of the underlying data distribution.

Validation of Alignment Costs via IRIS Joint Spatial Domain Partitions

To better evaluate the performance of pairwise alignment using PASTE in cases where reflection occurred, we applied IRIS, a spatial domain detection algorithm designed for the joint analysis of multiple sections. IRIS utilized single-cell RNA sequencing data as a reference during its clustering process, enhancing its ability to identify spatial domains across sections. This approach allowed us to compare gene expression patterns and spatial domain structures within the section pair of interest, helping to detect potential misalignment due to reflection.

For the reference data, the 10X Chromium Human Breast Cancer dataset from Wu et al.[12] was utilized. This dataset was derived from 26 breast cancer patients and included 29,733 genes and

100,064 cells. Additionally, breast cancers from the three major clinical subtypes were represented: 11 estrogen receptor-positive (ER+), 5 human epidermal growth factor receptor 2-positive (HER2+), and 10 triple-negative breast cancer (TNBC) cases.

In addition to the reference input, each section required two inputs: (1) spot-level spatial coordinates with dimensions $n \times 2$, and (2) a gene expression count matrix with dimensions $m \times n$, where m represents the number of genes and n the number of spots. These data were integrated into an object and preprocessed using IRIS's default function. Specifically, genes expressed in fewer than five spots and spots with fewer than 100 expressed genes were filtered out to denoise the input object. The resulting clusters were assigned numerical labels, which were maintained consistently across sections. This consistency allowed for the identification of shared spatial domains and similar structures, ultimately helping to pinpoint potential misalignment issues caused by reflection in the aligned sections.

Validation of Alignment Costs via Shared SVG Expression Patterns

We chose to visualize shared SVG expression patterns across sections to validate the resulting alignment costs. After identifying shared SVGs between paired sections, we compared their expression patterns on a log10 scale before alignment. If the patterns between the original coordinates were similar, it indicated the original orientation should be retained. Therefore, any noticeable deviations, such as reflections or distortion, would suggest errors in the alignment process, helping to identify potential issues in preserving the spatial structure using a single modality.

Validation of Alignment Costs via Heatmap

To validate the computed per-spot alignment costs, we used gene expression heatmaps as a visual confirmation of alignment accuracy. For each section pair, we selected common genes that were detected in both sections. We then generated heatmaps for these genes, with yellow indicating high expression and purple indicating low expression at each spot.

We compared the heatmaps of both sections to assess the consistency of gene expression patterns. Consistent distributions between the sections would imply similar spatial arrangements, while discrepancies in the patterns would suggest a discontinuity in spatial distributions between the sections due to the spatial distance between the obtained sections. This visual validation approach further strengthened the evidence supporting the computed alignment costs, reinforcing the effectiveness of the proposed cost function. This approach further strengthened the evidence, reinforcing the effectiveness of the proposed cost function.

Validation of Alignment Costs via Spot-level Cell-type Proportion

We assessed cell-type composition at the spot level to gain more insight into the nature of alignments. We used cell-type proportion data, which was obtained along with HER2+ ST data, considering eight major cell types: B-cells, cancer-associated fibroblasts (CAFs), endothelial cells, epithelial cells, myeloid cells, plasma cells, perivascular-like cells (PVL), and T-cells. For each section pair, we generated proportion annotation plots, where darker blue shades represented higher proportions of each cell type at individual spots.

We compared the cell-type proportions across aligned sections to evaluate the consistency of spatial distributions. Consistent distributions across both sections would indicate similar spatial

distributions across sections, while discrepancies patterns suggested non-continuity of spatial distributions across sections. This cell-type level validation provided additional support for the alignment accuracy, reinforcing the findings from gene expression and heatmap analysis.

Results
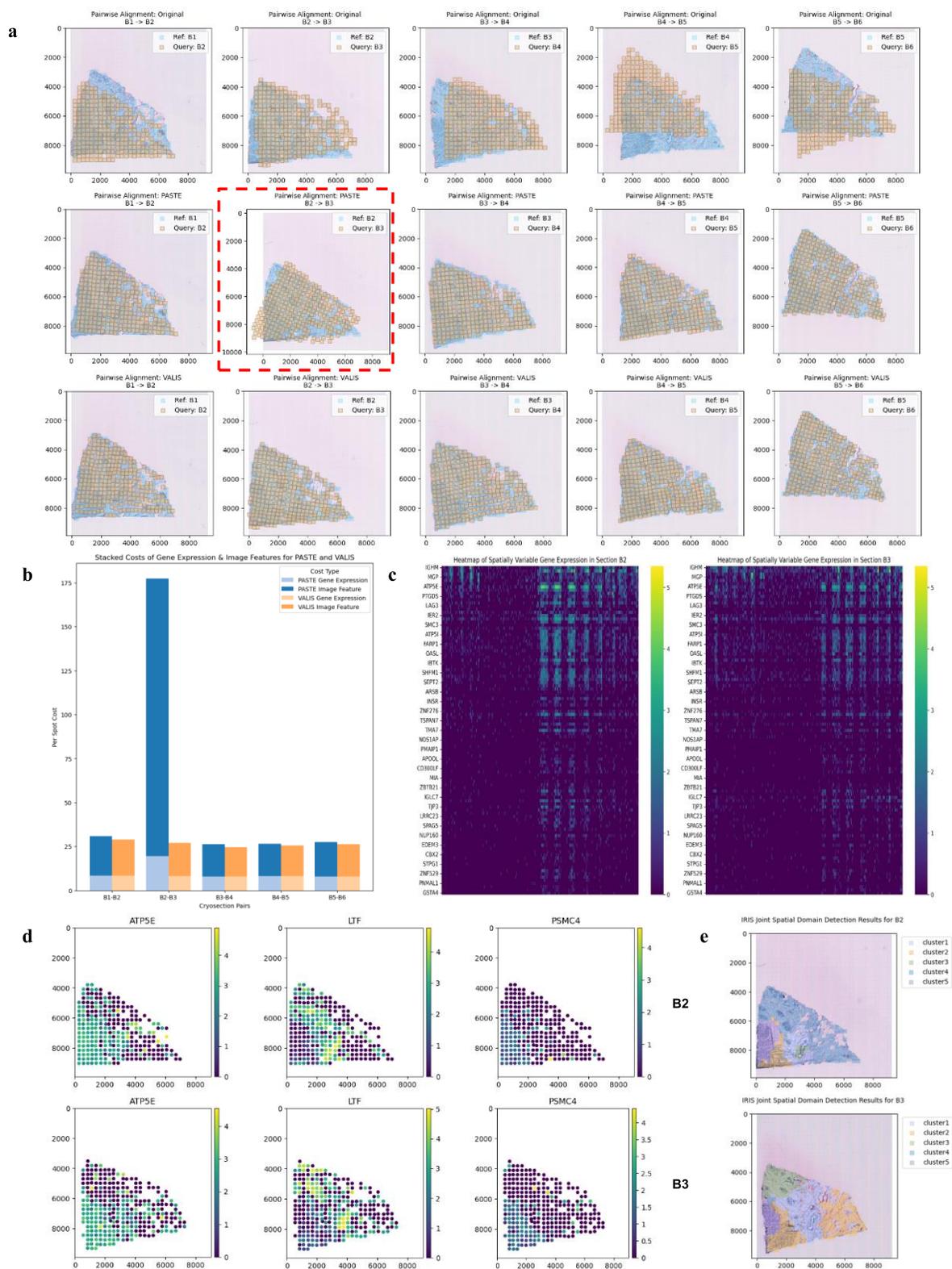
Figure 2. Analyzing the HER2 + Breast Cancer Data: Cryosections of Sample B. **a**, Alignment plots for B1–B6 showing original, PASTE-aligned, and VALIS-aligned coordinates. **b**, The bar plots of the per-spot alignment cost for PASTE and VALIS, with stacked gene expression and image feature components. **c**, Heatmaps (log10) of 3 common genes in B2 and B3. **d**, Expression patterns (log10) of 3 shared SVGs in B2 and B3. **e**, IRIS joint clustering of B2 and B3, highlighting shared spatial domains.

Sequential Pairwise Alignment on High-Quality Cryosections

We performed sequential pairwise alignment on six cryosections (B1 to B6) using two algorithms, PASTE and VALIS. Each consecutive pair (B1–B2, B2–B3, B3–B4, B4–B5, B5–B6) was aligned independently, and the resulting aligned spot coordinates were overlaid on the reference H&E image for visual assessment. PASTE, which relied on gene expression similarity, and VALIS, which employed image-based morphological features, were evaluated based on their alignment performance across the dataset. In **Figure 2a**, three plots are displayed for each pair of cryosections. The first plot shows the original coordinates of both cryosections overlaid on the H&E image of the reference section. The second plot presents the PASTE-aligned coordinates of the query section along with the original coordinates of the reference section, both overlaid on the H&E image of the reference section. The third plot illustrates the VALIS-aligned coordinates of the query section, along with the original coordinates of the reference section, overlaid on the same H&E image of the reference section. These plots are arranged from top to bottom, respectively.

For the B1–B2, B3–B4, B4–B5, and B5–B6 section pairs, both algorithms produced visually consistent alignments, preserving tissue morphology and structural integrity. The per-spot cost, which incorporated both gene expression and image features, remained below 30 for these section

pairs, as shown in **Figure 2b**. However, in the alignment of B2 and B3 using PASTE, we observed an unexpected reflection in the registered section, with the corresponding per-spot cost exceeding 175. To determine whether this reflection indicated a misalignment, we examined the heatmaps of commonly expressed genes in B2 and B3 (**Figure 2c**). The heatmaps revealed consistent expression patterns across spots on the $\log_{10}$ scale. To further validate this observation, we assessed the gene expression patterns of three SVGs in the affected region, as presented in **Figure 2d**. Again, the expression patterns remained consistent across the three selected SVGs, indicating the erroneous nature of the reflection produced by PASTE.

Additionally, we employed IRIS to jointly analyze B2 and B3 and assess whether the detected spatial domain patterns corroborated the observed reflection. As shown in **Figure 2e**, spatial domains 1, 2, and 5 were consistent across B2 and B3, reinforcing that the reflection was most likely an alignment error rather than a biologically meaningful pattern. This finding highlights the potential limitations of gene expression information in guiding pairwise alignment.

To confirm that the reflection was a misalignment between B2 and B3, we further examined cell type composition at the spot level. We considered eight major cell types: B-cells, cancer-associated fibroblasts (CAFs), endothelial cells, epithelial cells, myeloid cells, plasma cells, perivascular-like cells (PVL), and T-cells. We generated proportion annotation plots for B2 and B3, with darker blue shades indicating higher cell type proportions per spot (**Figure S1**). We observed consistent spatial distributions of all eight cell types across both sections further supporting misalignment due to the reflection. This consistency at the cell-type level reinforced the conclusion that PASTE has misaligned B2 with B3.

One plausible explanation for the reflection observed in the alignment between B2 and B3 using PASTE is that the gene expression patterns in the reflected regions were more similar to each other than to the true corresponding regions. This similarity caused the alignment process to minimize the gene expression cost by aligning the reflected regions instead of the true spatially corresponding ones. This misalignment arose because PASTE considers all genes in the dataset, rather than focusing on SVGs, which introduces noise from less relevant genes. Without accounting for additional contextual information, such as morphological features from the H&E images, this noise can lead to distortions in the alignment. Thus, when gene expression data is used in isolation, misalignments like the one observed can occur. In contrast, VALIS, which leverages image features, did not exhibit such distortions, suggesting that image-based alignment is more robust in this context. This superior performance was quantified by the per-spot alignment cost for each pair. Across all section pairs, VALIS consistently achieved lower alignment costs than PASTE, further supporting the reliability of image-based alignment.

Overall, while PASTE demonstrated reasonable alignment performance in most cases, the reflection error in B2–B3 underscores its sensitivity to gene expression variability. These findings suggest that image-based features, as utilized by VALIS, provide a more stable foundation for ST alignment, particularly in cases where gene expression patterns may not be sufficiently robust to ensure accurate registration.
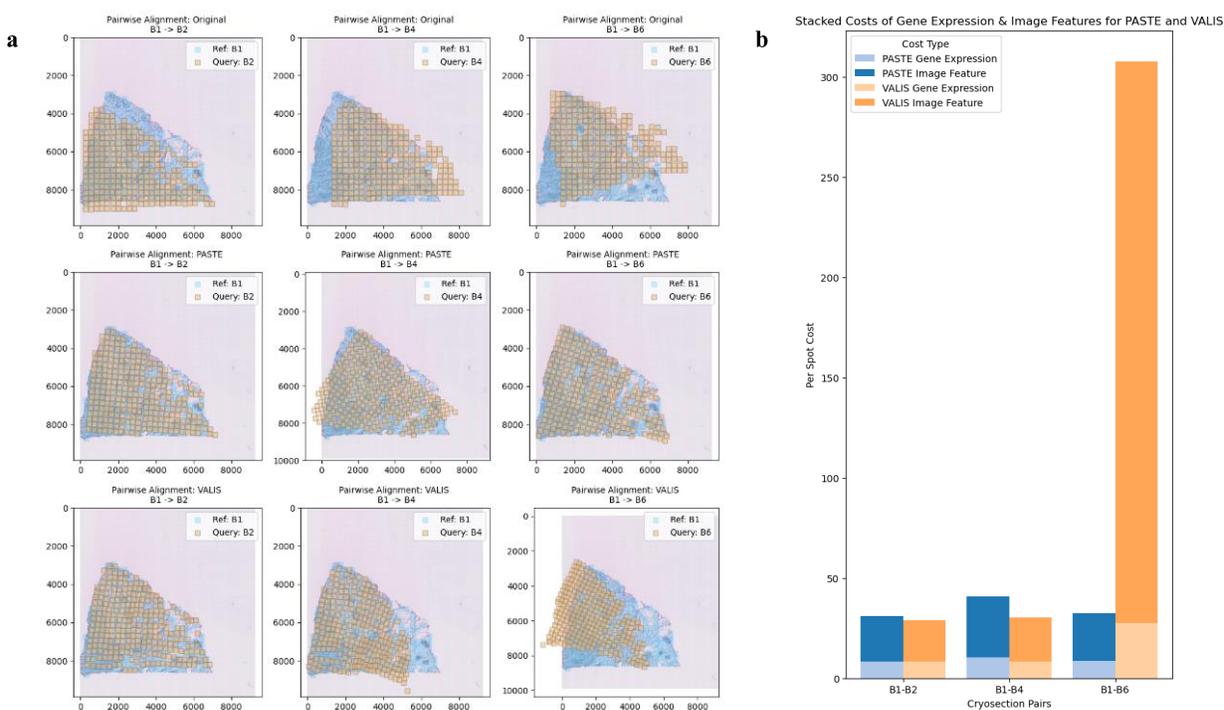
Figure 3. Analyzing Pairwise Alignment of Cryosections with Increased Distance. **a**, Alignment plots for B2, B4, and B6 showing original, PASTE-aligned, and VALIS-aligned coordinates. **b**, The bar plots of the per-spot alignment cost for PASTE and VALIS, with stacked gene expression and image feature components.

Pairwise Alignment of Cryosections with Increased Distance

To further investigate the impact of spatial distance on alignment performance, we conducted pairwise alignments between B1 and B4, B1 and B6, where the z-axis distance was three times and five times greater than that of the original consecutive pairs, respectively. This scenario assesses whether increased section spacing influenced the performance of PASTE and VALIS, particularly in the presence of well-preserved morphological features in the H&E images.

In **Figure 3a**, for the B1–B4 pair, PASTE produced a reflection, whereas VALIS resulted in substantial deformation. The increased alignment costs for both gene expression and image features, compared to the B1–B2 pair, indicate that the distance between sections has a significant impact on pairwise alignment performance. For the B1–B6 pair, PASTE achieved a good alignment, whereas VALIS failed to properly align the images, leading to a high image feature cost. In PASTE, the increased section spacing had noticeable effects on alignment performance; however, it successfully recovered the alignment in the most distant pair in sample B and achieved a low cost, as shown in **Figure 3b**. This suggests that gene expression information remains effective for aligning cryosections even with increased spacing. In contrast, VALIS was significantly impacted by the increased distance between cryosections, as shown in **Figure 3a**, with morphological patterns becoming less preserved as the sections got farther apart. These findings underscore the limitations of each method, particularly the challenge of preserving either gene expression consistency or morphological integrity in distantly spaced sections.
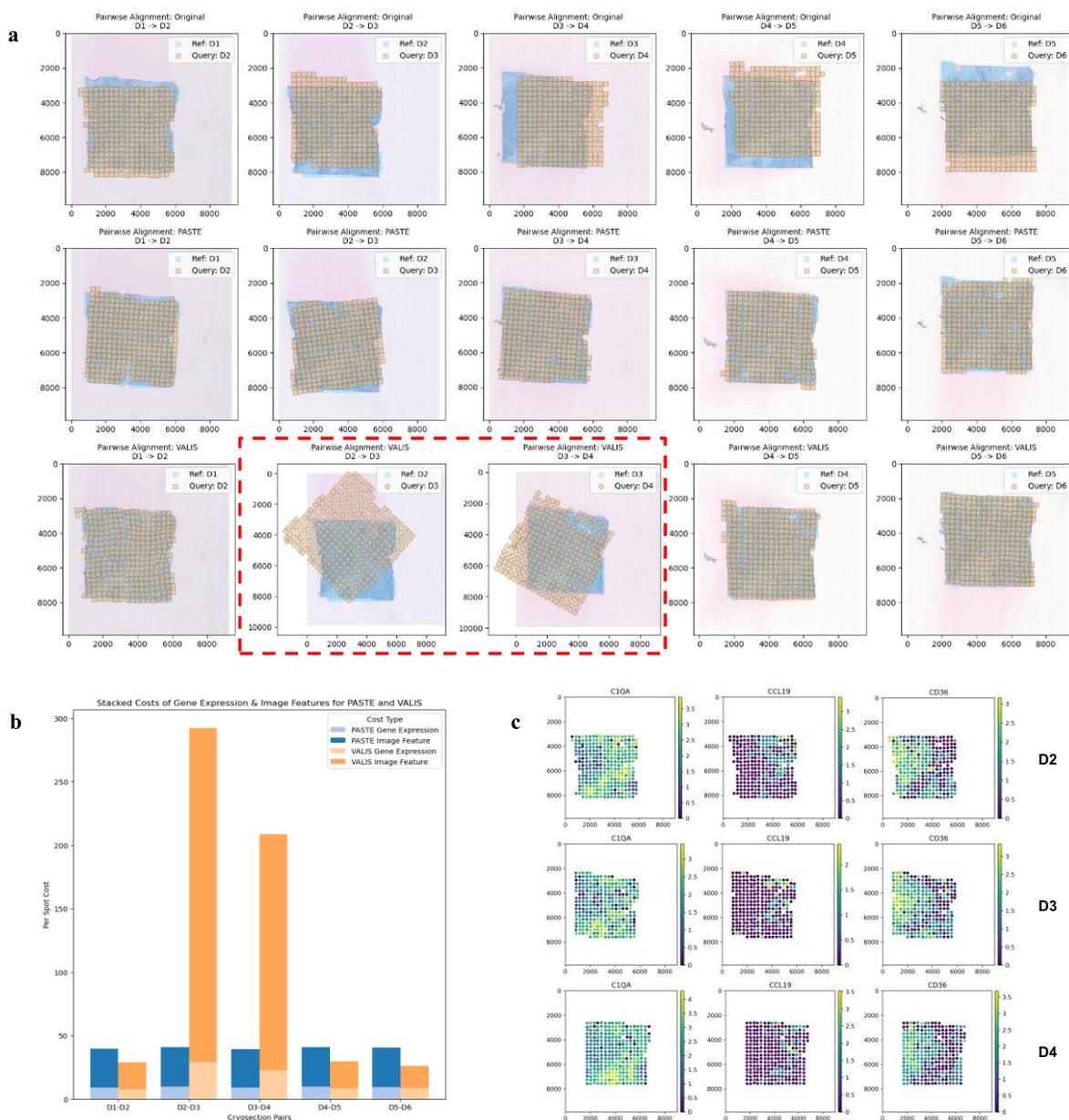
Figure 4. Analyzing the HER2 + Breast Cancer Data: Cryosections of Sample D. **a**, Alignment plots for D1–D6 showing original, PASTE-aligned, and VALIS-aligned coordinates. **b**, The bar plots of the per-spot alignment cost for PASTE and VALIS, with stacked gene expression and image feature components. **c**, Expression patterns (log10) of 3 shared SVGs in D2, D3 and D4.

Sequential Pairwise Alignment on Challenging Cryosections

To further evaluate the performance of PASTE and VALIS under more challenging conditions, we applied both alignment algorithms to a second set of cryosections, D1 to D6. Similar to the previous analysis, we performed sequential pairwise alignment for each consecutive pair (D1–D2, D2–D3, D3–D4, D4–D5, D5–D6) and assessed the resulting alignments through visual inspection and quantitative metric for gene expression and morphological similarities. Unlike the B-series cryosections, this dataset presented additional challenges due to heterogeneous and/or symmetric gene expression patterns and the lack of distinct spatial patterns in the H&E images, increasing the complexity of accurate alignment. In **Figure 4a**, three plots are shown for each pair of consecutive cryosections from sample D (D1 to D6), following the same format as for sample B.

PASTE produced reflections across all section pairs due to its gene expression minimization process. The reliability of these reflections was evaluated by comparing the per-spot alignment costs of PASTE and VALIS for each pair. As shown in **Figure 4b**, the alignment costs for the D1–D2, D4–D5, and D5–D6 pairs were significantly lower in VALIS than in PASTE, suggesting that the reflections in PASTE were likely misalignments caused by noise introduced by less relevant genes. In contrast, the alignment costs for the D2–D3 and D3–D4 pairs were significantly higher in VALIS, primarily due to scaling and rotation issues.

As shown in **Figure 4c**, even when utilizing representative SVGs, the gene expression patterns in sample D remained heterogeneous or sparse, as observed for genes like C1QA and CCL19. Moreover, some patterns lacked clear orientation, such as CD36, where the highly expressed region formed a triangular shape with symmetry along the x-axis. These characteristics of heterogeneity and symmetry further highlight the challenges posed by this dataset.

We generated cell type proportion annotation plots for sections D2, D3, and D4, with darker blue shades indicating higher cell type proportions per spot (**Figure S2**). Consistent spatial distributions of most of the eight major cell types were observed across the three sections, further supporting the conclusion that the reflections were alignment errors. This cell-type level consistency reinforced the inference that PASTE had misaligned D2 with D3 and D3 with D4.

Overall, these findings suggest that neither PASTE nor VALIS is fully sufficient for aligning cryosections that lack clear spatial patterns and landmarks. The observed distortions and elevated alignment costs underscore the need for improved alignment strategies capable of addressing cases with weak spatial structures in both gene expression and image features, and potentially incorporating additional modalities when available.

Discussion

This study evaluated two alignment strategies for ST data: whole-section alignment using PASTE and morphology-based WSI alignment using VALIS. Our findings highlight key differences in their performance under varying conditions, particularly in response to increased section spacing. While PASTE effectively leveraged gene expression information to align distant sections, it was susceptible to alignment distortions, such as reflections, when symmetric gene expression patterns were present. In contrast, VALIS maintained high alignment accuracy for closely spaced sections but struggled as section spacing increased, leading to diminished morphological preservation. These results emphasize the complementary nature of gene expression- and morphology-based alignment strategies and the need for hybrid approaches that integrate both modalities.

A key strength of this study is its systematic evaluation of alignment performance using multiple metrics, incorporating both biological and computational perspectives. By leveraging IRIS for region-specific validation, we were able to assess alignment accuracy beyond simple spatial metrics, considering biologically relevant spatial domains. Additionally, the proposed evaluation workflow provides a structured approach to assessing alignment performance across different methods, ensuring fair comparisons.

However, this study has some limitations. First, while PASTE and VALIS were evaluated across different spacing conditions, the dataset was limited to a specific tissue type, and findings may not generalize to other tissues with different structural and molecular characteristics. Second, the evaluation focused on pairwise alignment, which may not fully capture the challenges of aligning multiple consecutive sections simultaneously. Third, although statistical and visual assessments

were conducted, a more detailed validation using independent spatial ground-truth data would further strengthen the conclusions.

Future work should explore hybrid alignment frameworks that integrate both gene expression and morphological features to improve overall performance. One potential direction is to refine alignment models by incorporating additional spatial constraints to mitigate extreme transformations, such as large rotation angles or reflections observed in PASTE. Additionally, expanding the evaluation to diverse tissue types and sequencing platforms could provide deeper insights into the generalizability of these alignment methods. Lastly, developing automated quality control metrics to flag misalignments could enhance the robustness of alignment strategies for practical applications.

Overall, this study provides a comparative assessment of gene expression- and morphology-based alignment strategies, highlighting their respective strengths and limitations. The results highlight the significance of multi-modal approaches for tissue section alignment and pave the way for future advancements in 3D reconstruction and pathological applications of ST data.

Bibliography

1. Ståhl, P. L., Salmén, F., Vickovic, S., et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*. 2016;353(6294):78–82.

2. Maynard, K. R. et al. Transcriptome-scale spatial gene expression in the human dorsolateral prefrontal cortex. *Nat. Neurosci.* 2021;24(4):612.

3. Janesick A, Shelansky R, Gottscho AD, Wagner F, Williams SR, Rouault M, Beliakoff G, Morrison CA, Oliveira MF, Sicherman JT, et al. High resolution mapping of the tumor microenvironment using integrated single-cell, spatial and in situ analysis. *Nat Commun.* 2023;14(1):8353.

4. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. Y. & Zhuang, X. W. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*. 2015;348(6233).

5. Wang, X. et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science*. 2018;361(5691).

6. Zeira, R., Land, M., Strzalkowski, A. & Raphael, B. J. Alignment and integration of spatial transcriptomics data. *Nat. Methods*. 2022;19(5):567–575.

7. Gatenbee, C. D. et al. Virtual alignment of pathology image series for multi-gigapixel whole slide images. *Nat. Commun.* 2023;14(1): 4502.

8. Pantanowitz L. Digital images and the future of digital pathology. *J Pathol Inform*. 2010;1:15.

9. Fischer AH, Jacobson KA, Rose J, Zeller R. Hematoxylin and eosin staining of tissue and cell sections. *CSH Protoc*. 2008;2008(5).

10. Ma, Y. & Zhou, X. Accurate and efficient integrative reference-informed spatial domain detection for spatial transcriptomics. *Nat. Methods*. 2024;21(7):1231-1244.

11. Andersson, A., Larsson, L., Stenbeck, L. et al. Spatial deconvolution of HER2-positive breast cancer delineates tumor-associated cell type interactions. *Nat. Commun.* 2021;12(1):6012.

12. Wu SZ, Al-Eryani G, Roden DL, et al. A single-cell and spatially resolved atlas of human breast cancers. *Nat Genet*. 2021;53(9):1334-1347.

Appendix

Software and Package Utilization

The study's computational analyses were conducted using Python version 3.10 and R version 4.4.1, leveraging the strengths of each environment for different stages of the workflow. Python was primarily used for image processing, ST analysis, and alignment evaluation, while R supported advanced statistical modeling and validation steps. For stability, we set up a virtual environment using Conda version 25.3.0, ensuring consistent dependencies and an isolated environment for the analysis.

Python Packages:

- os: Used for file and directory management throughout the analysis. It facilitated tasks such as navigating the file system, handling file paths, and organizing output files, ensuring a smooth workflow for data processing and analysis.

- NumPy: This core library for numerical computing supported the manipulation of multidimensional arrays and matrices, which was essential for handling count data, image data and spatial coordinates.

- Pandas: Used for data preparation and pre-processing, particularly in managing ST metadata, image-derived features, and gene expression matrices. Its DataFrame structure offered a flexible and efficient way to store and join information across modalities.

- Scanpy: A specialized toolkit for analyzing single-cell and ST data. It was utilized for data storage and data loading. Scanpy's AnnData structure provided an efficient format for storing count matrices along with spatial coordinates and other metadata.

- OpenCV (cv2): This powerful computer vision library was employed for image preprocessing tasks such as grayscale conversion, resizing, and format conversion between OpenCV and PIL formats.

- Pillow (PIL): Essential for working with high-resolution PNG images and constructing composite visualizations, such as 3D-stacked images for section alignment illustrations.

- Matplotlib and Seaborn: These libraries facilitated the visualization of alignment outcomes, cost function comparisons, and spatial landmarks, allowing clear interpretation of method performance.

- SciPy: Used for spatial computations, nearest-neighbor analysis, and clustering evaluations. It also supported metric calculations for image similarity assessment.

- Scikit-learn: Provided tools for dimensionality reduction (e.g., PCA), standardization, clustering algorithms.

- Squidpy: utilized to identify SVGs in each section, with the package's advanced spatial statistics tools, such as Moran's I, playing a key role in this process.

- VALIS: Employed as a histology-based alignment tool using non-rigid image registration techniques. VALIS performed well in aligning H&E images based on morphological patterns alone and served as a benchmark for image-only alignment performance.

- PASTE: Used as a gene expression-based alignment tool that minimizes expression dissimilarity across sections using optimal transport. PASTE was evaluated for its performance in capturing biological continuity across sections using ST data.

- HIPT: Applied to extract multi-scale, high-dimensional morphological features from H&E-stained images.

R Packages:

- IRIS: Used for joint spatial domain partitioning by integrating gene expression and morphological features. IRIS facilitated region-specific alignment and performance evaluation through unsupervised clustering of tissue sections.

- ggplot2: Used for visualization of IRIS clustering results.

Supplementary Figures
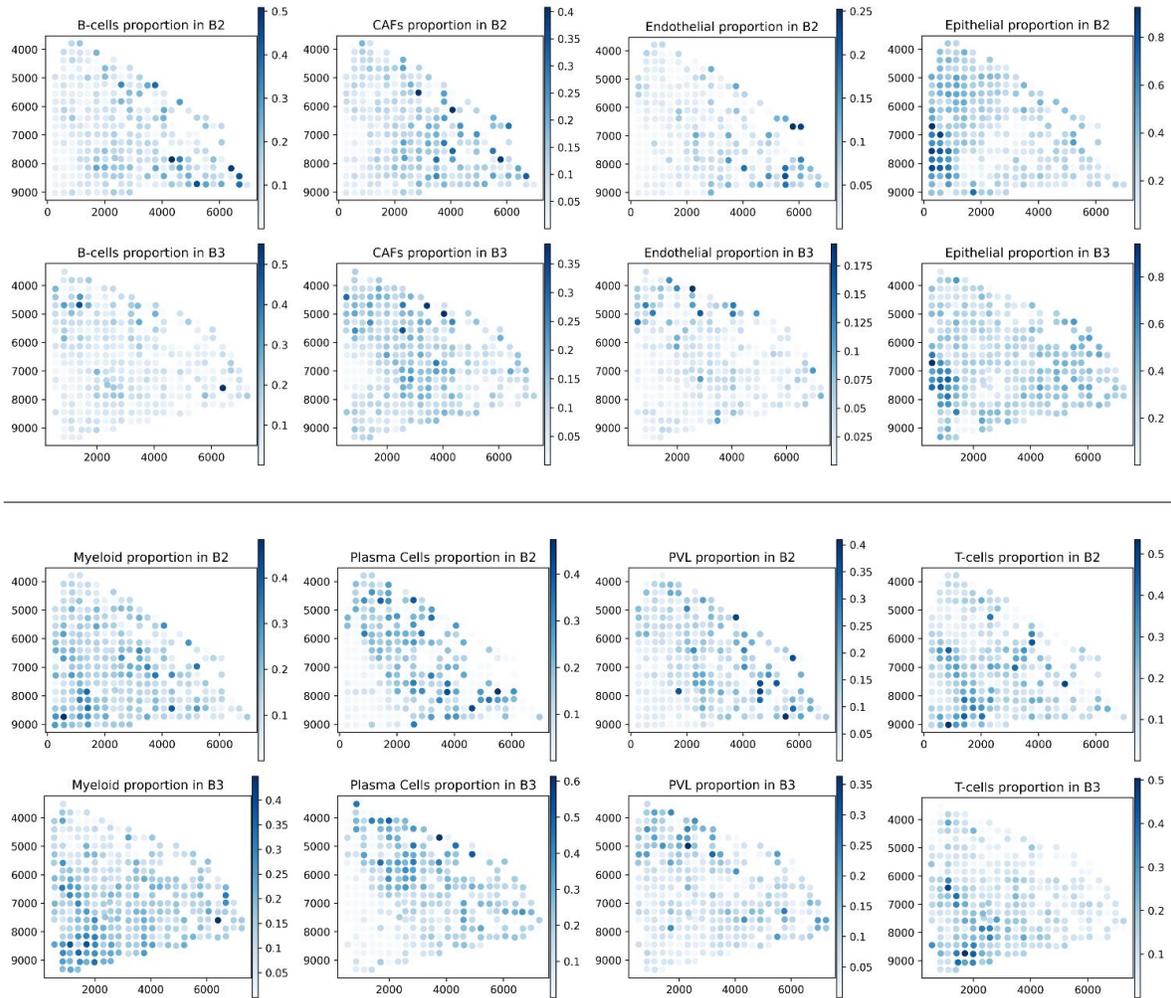
Figure S1. Spot-level cell type composition of B2 and B3
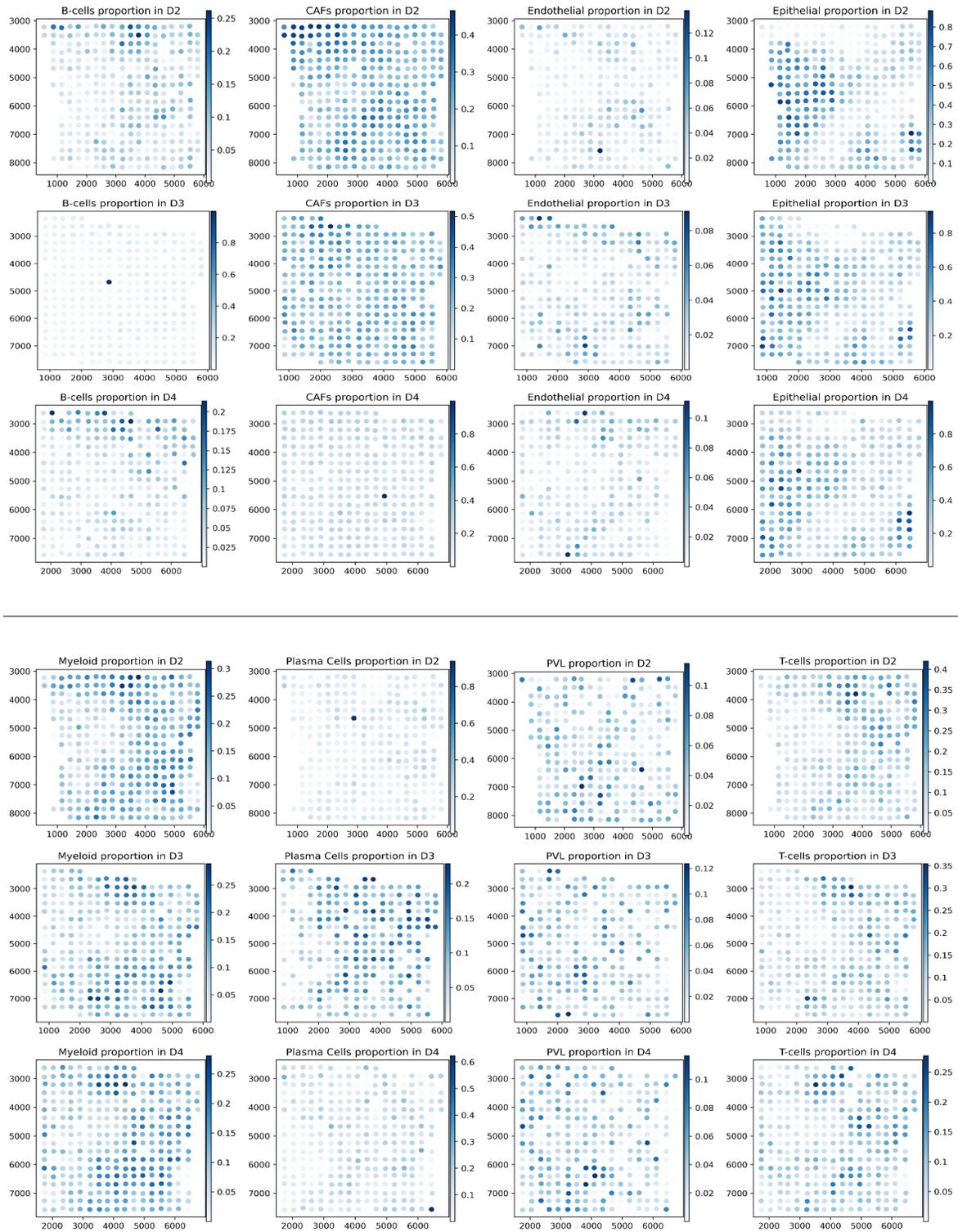
Figure S2. Spot-level cell type composition of D2, D3 and D4

Supplementary Table

Section	Number of Spots	Number of Genes	Number of Image Features
B1	295	15109	579
B2	270	15290	579
B3	298	15215	579
B4	283	15289	579
B5	289	15273	579
B6	277	15387	579
D1	306	15661	579
D2	303	15396	579
D3	301	15529	579
D4	302	15503	579
D5	306	15666	579
D6	315	15409	579

Table S1. Data Description