

---

1

## 2 Inferring spatially transient gene expression pattern from spatial 3 transcriptomic studies

---

4

5 Jan Kueckelhaus<sup>1,2,3#</sup>, Jasmin von Ehr<sup>1,2,3#</sup>, Vidhya M. Ravi<sup>1,2,3</sup>, Paulina Will<sup>1,2,3</sup>, Kevin Joseph<sup>1,2,3</sup>, Juergen Beck<sup>2,3</sup>, Ulrich G.  
6 Hofmann<sup>3,4</sup>, Daniel Deley<sup>5,6</sup>, Oliver Schnell<sup>2,3,7</sup>, Dieter Henrik Heiland<sup>1,2,3</sup>

7

8 <sup>1</sup>Microenvironment and Immunology Research Laboratory, Medical Center, University of Freiburg, Germany

9 <sup>2</sup>Department of Neurosurgery, Medical Center, University of Freiburg, Germany

10 <sup>3</sup>Faculty of Medicine, Freiburg University, Germany

11 <sup>4</sup>Neuroelectronic Systems, Medical Center, University of Freiburg, Germany

12 <sup>5</sup>Department of Neurosurgery, RWTH University of Aachen, Aachen, Germany

13 <sup>6</sup>Neurosurgical Artificial Intelligence Laboratory Aachen (NAILA), Department of Neurosurgery, RWTH University of Aachen,  
14 Aachen, Germany

15 <sup>7</sup>Translational NeuroOncology Research Group, Medical Center, University of Freiburg, Germany

16

17 **# Equal contributed first authorship**

18 **DISCLOSURE OF CONFLICTS OF INTEREST:** No potential conflicts of interest were disclosed by the authors.

19 **Corresponding author:**

20 Dieter Henrik Heiland

21 Department of Neurosurgery

22 Medical Center University of Freiburg

23 Breisacher Straße 64

24 79106 Freiburg

25 -Germany-

26 Tel: +49 (0) 761 270 50010

27 Fax: +49 (0) 761 270 51020

28 E-mail: [dieter.henrik.heiland@uniklinik-freiburg.de](mailto:dieter.henrik.heiland@uniklinik-freiburg.de)

29

30

## 31 **Abstract**

32 Spatial transcriptomic is a technology to provide deep transcriptomic profiling by preserving the spatial  
33 organization. Here, we present a framework for SPAtial Transcriptomic Analysis (SPATA,  
34 <https://themilolab.github.io/SPATA>), to provide a comprehensive characterization of spatially resolved  
35 gene expression, regional adaptation of transcriptional programs and transient dynamics along spatial  
36 trajectories.

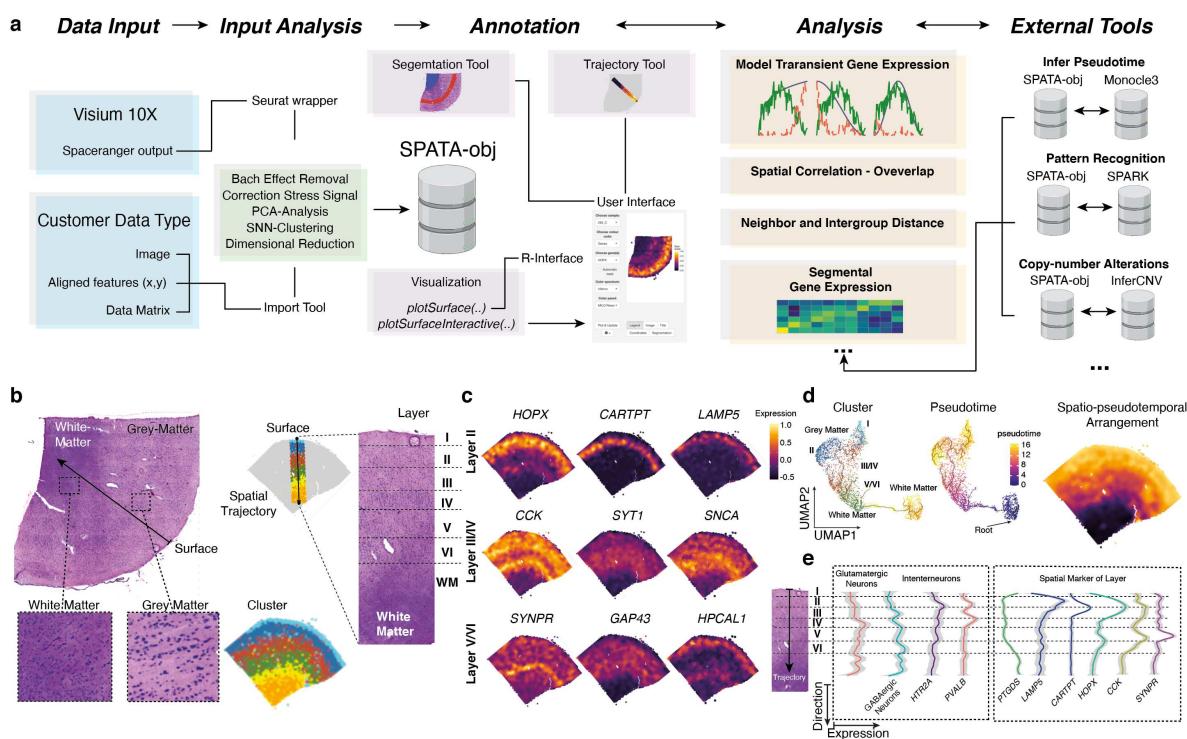
37

38

## 39 **Brief Communication**

40 Deep transcriptional profiling of single cells by RNA-sequencing maps the cellular composition of tissue  
41 specimens regarding cellular origin, developmental trajectories and transcriptional programs<sup>1-3</sup>.  
42 However, information determining the spatial arrangement of specific cell types or transcriptional  
43 programs are lacking and thus can only be predicted indirectly<sup>4</sup>, which is a considerable drawback of  
44 this method. Spatial tissue organization was traditionally investigated by imaging technologies which  
45 provide information at high resolution but are strongly limited by the number of genes or proteins to be  
46 mapped. Several novel technologies such as MERFISH<sup>5</sup>, FISH-seq<sup>6</sup>, Slide-seq<sup>7</sup> or spatial  
47 transcriptomics<sup>8,9</sup> are able to preserve the spatial context of transcriptional data, however all these  
48 technologies are limited by either the spatial resolution or depth of transcriptional profiling. Further, data  
49 integration, visualization and analysis of transcriptomic and spatial information remains challenging.  
50 Here, we present a software tool to provide a framework for integration of high-dimensional  
51 transcriptional data within a spatial context. By combining user-friendly interfaces for visualization,  
52 segmentation or trajectory analysis and command-based pipe-friendly functions for data manipulation  
53 and modeling, we provide a broad range of applications for different analytical demands. In addition, we  
54 implemented interfaces to provide easy exchange of numerous external tools. Previously published  
55 tools focus mainly on the visualization of gene expression using known tools from scRNA-seq analysis  
56 rather than addressing gene expression within its spatial context<sup>10-12</sup>. In particular, we focus on transient  
57 changes of gene expression and aim to infer transcriptional programs that are dynamically regulated as  
58 a function of spatial organization.

59 In order to present an overview of possible analytic capabilities of the SPATA workflow, Figure 1a, we  
60 generated spatial transcriptomic datasets from human cortex and human glioblastoma samples using  
61 the Visium technology (10X Genomics). The human cortex is separated into defined layers containing  
62 different types of neurons and cellular architecture. In a first step, we combine shared-nearest neighbor  
63 clustering and spatial pattern recognition by an external tool (spatial pattern recognition via kernels,  
64 "SPARK"<sup>13</sup>) in order to determine genes with a defined spatially resolved expression pattern. We found  
65 that the cortical layering is accurately reflected by our clustering approach. In order to gain insights into  
66 the spatial organization we provided a tool to compute the spatial distance within the defined layers or  
67 correspondent clusters. An increasing distance within individual clusters allows to differentiate between  
68 narrowly related or a widespread dispersion of spots within the cluster.

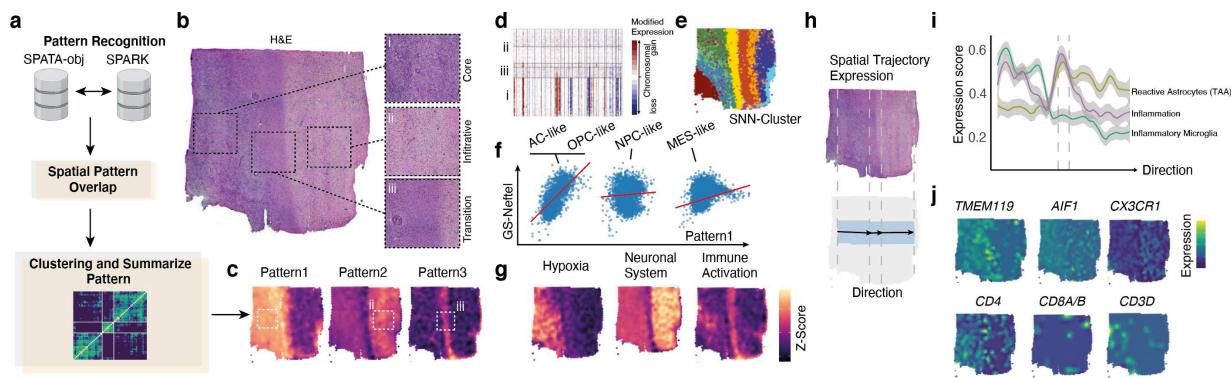


69

70 *Figure 1: a) Illustration of the SPAtial Transcriptomic Analysis (SPATA) workflow containing Data Input and a predefined set of*  
 71 *analysis which will be saved in a SPATA object. In the following step of the workflow, annotation of segment trajectories can be*  
 72 *performed using a user-friendly interface. Additionally, multiple tools for visualization are available within the user interface. This*  
 73 *also includes tools for geneset enrichment analysis (GSEA) and gene set variation (GSVA). After region of interests are defined,*  
 74 *a list of analytic tools is provided which includes wrapper for external tools. b) H&E staining of a human cortex sample with the*  
 75 *corresponding SNN clustering (bottom right). The annotation tool is used to draw a trajectory along all cortical layer (right side) c)*  
 76 *Infer genes with defined peaks along the trajectory revealed genes with layer specific gene expression. d) Integrating an external*  
 77 *package (monocle3) the pseudotime within the sample was computed and visualized by a 2D UMAP representation (left) and*  
 78 *within its spatial context (right). e) Comparison of traditional markers (left) and markers given by our model (inferring transient*  
 79 *gene expression) along the cortical layering.*

80

81 Next, the spatial overlap of transcriptional programs or gene expression was analyzed using a Bayesian  
 82 approach, resulting in an estimated correlation which quantifies the identical arrangement of expression  
 83 in space. In a further step, we aimed to analyze dynamic changes, which were annotated using  
 84 pseudotime estimation or RNA-velocity. We directly implemented the pseudotime inference from the  
 85 monocle3<sup>14</sup> package, but also allow the integration of any other tool such as “latent-time” extracted from  
 86 RNA-velocity (scVelo<sup>15</sup>). Another option for dynamic gene expression analysis is the detection of defined  
 87 transcriptional programs along a defined trajectory. In our example, we mapped different activation  
 88 states of astrocytes and microglia within the cortical layering.



89

90 *Figure 2: a) Illustration of the SPATA workflow integrating SPARK for pattern recognition into the analysis of human glioblastoma.*  
91 SPARK will estimate to what extent a gene is present in a spatial pattern. The output is piped into a spatial overlap analysis and  
92 clustered to extract set of genes which belong to the same pattern. b) H&E staining of glioblastoma with 3 histological distinct  
93 regions. c) Predicted pattern visualized by the z-scored gene expression of all genes aligned into a pattern. d-e) Copy-number  
94 analysis of the sample and cluster annotation (e). f) Comparison of recognized pattern with known gene expression classification,  
95 here the Neftel classification. g) Expression of significantly expressed pathways within a pattern. h) Spatial trajectory analysis  
96 along the tumor infiltration region. i-j) Change of z-scored geneset expression along the trajectory (i) and marker genes of  
97 microenvironmental alterations and inflammation (j).

98

99 Moreover, we provide the opportunity to screen for gene expression or transcriptional programs which  
100 transiently change along predefined trajectories by modelling gene expression changes in accordance  
101 to various biologically relevant behaviors. All genes or transcriptional programs which significantly  
102 followed one or multiple predefined models were ranked and visualized. The detection of dynamic  
103 spatially defined gene expression patterns is also of great interest in malignant specimens. In another  
104 example, we profiled tissue of a human glioblastoma, the most malignant tumor of the central nervous  
105 system (CNS) as SPATA provides numerous tools to analyze datasets with malignant origin. In a first  
106 step, integrating inferred copy-number alterations (CNV)<sup>2,16</sup>, spatial pattern recognition and shared-  
107 nearest neighbor clustering provides a broad overview of spatially defined transcriptional programs  
108 within the subclonal architecture of tumor samples, *Figure 2a-g*. Using this information, specific  
109 segments can be specified and analyzed to gain insights into their spatially differentially expressed  
110 genes. We showed that segments of higher cellular density also contained increased signaling of the  
111 hypoxic pathway including expression of *VEGFA*, *HIF1A* and *GAPDH*. Additionally, mapping the  
112 subclonal architecture based on a CNV clustering allowed to screen for gene expression differences  
113 within regions of exclusive genetic context *Figure 2d*. Inferring spatially transient gene expression along  
114 trajectories connecting particular tumor regions, i.e. between tumor core and infiltration zone, provided  
115 the opportunity to map transcriptional programs executed during tumor infiltration and tumor-induced

116 microenvironment changes of the surrounding areas. Thus, we were able to show that immune related  
117 genes from myeloid cells and reactive astrocytes were localized in a “glial-scars” resembling structure,  
118 sharply separating normal brain from tumor regions *Figure 2h*. We observed a transient increase of  
119 macrophage and microglia activation directed towards the tumor boarder. Mapping transcripts that mark  
120 for lymphoid cells, we found more T cells abundance within the normal brain compared to tumor regions  
121 which is in line with the reported immunosuppressive environment within glioblastoma. Inferring  
122 pseudotime, we were able to confirm a dynamic adaptation of myeloid cells along our defined trajectory  
123 *Figure 2i-j*. Recently, Neftel and colleagues established a classification of 4 transcriptional states using  
124 single-cell RNA-sequencing, *Figure 2f,i*. Using these signatures, we were able to map the spatial  
125 distribution of assumed tumor heterogeneity. We implemented a 2D representation of all 4 states which  
126 could be used to map the distribution of all transcriptional states within defined segments or along spatial  
127 trajectories. Of utmost importance, our tool enables the usage of a variety of different biological data  
128 containing a spatial context such as spatially resolved mass-spectroscopy or imaging mass cytometry  
129 (IMC). SPATA is a resource developed from scientists for scientists incorporating the FAIR principles of  
130 providing findable, accessible, interoperable, and reusable data<sup>17</sup>.

131

132 **Methods:**

133 **SPATA software and functions**

134 A detailed overview of all included functions and the structure of the package is given at the package  
135 website (<https://themilolab.github.io/SPATA/index.html>). We implemented tutorials for all described  
136 analytic approaches to provide a simple-as-possible solution to trace the individual analytic steps.

137

138 **Data preparation, per-analysis and SPATA object implementation**

139 We offer two possible input options. On one side, we implemented the direct input from spaceranger by  
140 using the Seurat wrapper for spatial transcriptomics. On the other hand, we used the Seurat v3.0  
141 package to normalize gene expression values by dividing each estimated cell by the total number of  
142 transcripts and multiplied by 10,000, followed by natural-log transformation. As described for single cell-  
143 RNA sequencing, we removed batch effects and scaled data using a regression model including sample  
144 batch and percentage of ribosomal and mitochondrial gene expression. For further analysis we used  
145 the 2000 most variable expressed genes and decomposed eigenvalue frequencies of the first 100  
146 principal components and determined the number of non-trivial components by comparison to  
147 randomized expression values. The obtained non-trivial components were used for SNN clustering  
148 followed by dimensional reduction using the UMAP and TSNE algorithm. After analysis all data will be  
149 saved in a SPATA object, detailed information of the S4 object structure is given at the package  
150 information. Another option is to provide 3 files that will be used to create a SPATA object, one file  
151 containing barcode information or other identifier of each spot with the given x and y coordinates  
152 determining the spatial position of each spot within the H&E image. The second file contains an  
153 expression or intensity matrix with identifier as colnames and genes or other features as rownames. The  
154 last file is an image with x and y coordinates corresponding to the identifier of file1. If the inputs are gene  
155 expression counts we run the standard pipeline (Seurat wrapper), otherwise (IMC, MALDI or MERFISH)  
156 we provide a data analysis pipeline which is designed for non-integer inputs and normal distributed data.

157

158 **Modeling of transient gene expression along spatial trajectories**

159 A given trajectory includes multiple spots summarized into predefined bins of the directed trajectory. In  
160 order to model the gene expression of single genes or genesets we created a set of mathematical  
161 models which represent defined biological behaviors, including linear, logarithmic or gradient  
162 ascending/descending expression pattern, one-, or multiple peak expression, detailed information in the

163 package description. The analysis is implemented into the function `assessTrajectoryTrends()`. Further,  
164 if a defined pattern is requested, we open the possibility to add a vector containing the requested model  
165 for which the algorithm will screen. Next, we fitted the summarized expression values of each bin using  
166 a non-parametric kernel estimation (Gaussian or Cauchy-Kernel), input vectors were normalized and z-  
167 scored: (1)  $n_{exp\ i} = \frac{A_{exp\ i} - min(A_{exp})}{max(A_{exp}) - min(A_{exp})}$  (2)  $\hat{f}_h(n_{exp\ i}) = \frac{1}{n} \sum_{i=1}^n K_h(n_{exp} - n_{exp\ i})$  K is the kernel and  $0.7 >$   
168  $h > 0.3$  is used to adjust the estimator. Next, we computed residuals for each input vector (gene  
169 expression) and estimated area under the curve (AUC) using the trapezoidal numerical integration.  
170 (3)  $\int_b^a f(res) dx \approx \sum_{k=1}^n \frac{f(res_{k-1}) + f(res_k)}{2} \Delta res_k$  The distance and direction is defined by [a,b]  $a=x_0 < x_1 <$ ,  
171  $\dots < x_{n-1} < x_n = b$ . We use the AUC to rank the estimated models and predict genes that follow our  
172 predefined behavior. The implemented function `plotTrajectoryFit()` shows the model fit with respect  
173 to the given residuals.

174

## 175 **Enrichment analysis for SPATA**

176 Gene sets were obtained from the database MSigDB v7 and internally created gene sets are available  
177 at within the package. For enrichment analysis we provide multiple methods listed in the description of  
178 the `plotSurface()` function. Per default, we use a probability distribution fitting of the input values which  
179 could be genes or summarized gene sets and transformed the distribution to representative colors.  
180 Further adaptation of the applied color scale can be performed by using the  
181 `confuns::scale_color_add_on()`. Further additions for geneset enrichment analysis or gene set variation  
182 analysis are implemented by using the GSVA package. As input for a GSEA the normalized and  
183 centered expression data are used and further transformed to z-scores ranging from 1 to 0. Genes were  
184 ranked in accordance to the obtained differential expression values and used as the input for GSEA.  
185

## 186 **Two-dimensional representation of cellular states**

187 Within the SPATA toolbox, we allow to plot a recently popular 2D presentation of multiple cellular states.  
188 As usual in SPATA, we provide two versions to acquire the data, on one side plotting from inside a  
189 SPATA-object is possible (`plotFourStates()`) and on the other hand, data can be used from outside (for  
190 example an expression matrix containing a single-cell dataset) by applying version 2 of the  
191 function(`plotFourStates2()`). Therefore, we aligned spots to variable states based on defined gene sets:  
192  $GS_{(1,2,\dots,n)}$ . We separated cells into  $GS_{(1+2)}$  versus  $GS_{(2+4)}$ , using the following equation:  $A_1 =$

193  $\| GS_{(1)}, GS_{(2)} \|_{\infty} - \| GS_{(3)}, GS_{(4)} \|_{\infty}$  A1 defines the y-axis of the two-dimensional representation. In a next  
194 step, we calculated the x-axis separately for spots  $A1 < 0$  and  $A1 > 0$ :  $A1 > 0: A_2 = \log 2 (\overline{GS_{(1)}} -$   
195  $[\overline{GS_{(2)}} + 1])$   $A1 < 0: A_2 = \log 2 (\overline{GS_{(3)}} - [\overline{GS_{(4)}}])$  For further visualization of the enrichment of subsets  
196 of cells according to gene set enrichment across the two-dimensional representation, using a probability  
197 distribution fitting - we transformed the distribution to representative colors. This representation is an  
198 adapted method published by Neftel and colleges recently<sup>2,3</sup>.

199

## 200 **Spatial distance measurement**

201 In order to measure the spatial distance, we use either a defined factorized input or a continuous vector.  
202 We first measure the spatial distance from each spot to all other spots and compute a distance matrix  
203 with spots as rows and columns ( $n_r = n_c$ ). If factorized input was applied, we factorize the matrix and  
204 calculate the mean distance per factor ( $f_1 \rightarrow f_i$ ): (1)  $dist_{fi} = \frac{\sum_{k=1}^n a}{2n}$  (2)  $a = \begin{bmatrix} s_1 & s_1 \\ \dots & \dots \\ s_{nr} & s_{nc} \end{bmatrix}$ . If a distance is  
205 numeric, we create bins of spots with common gene expression of gene set enrichment resulting in  
206 factorized values. Using the distance computation, we estimate to what extent a gene is expressed in  
207 exclusive spots (lower distance) or diffuse without spatial enrichment.

208

## 209 **Spatial overlap and correlation analysis**

210 Spatial overlap of spatial correlation was designed to estimate the similarity of gene expression pattern  
211 within the spatial organization. In order to map spatial correlated gene expression or gene set  
212 enrichments, we used z-scored ranked normalized expression values. We used a Bayesian approach  
213 to compute the correlation distribution within two different genes or gene sets (~5-20 minutes runtime,  
214 MacOS 2019). The spatial reference is given by the x and y coordinates of each spot. In order to provide  
215 an alternative method which is computationally less-intensive (~1-3 minutes runtime, MacOS 2019) we  
216 construct a trajectory of spots from lowest ranked to highest ranked spot (based on z-scored input  
217 vectors). The genes of interest (which were correlated with the spatial trajectory) were fitted by loess-fit  
218 from the stats-package (R-software) and aligned to the ranked spots and fitted by a probability  
219 distribution. Correlation analysis was performed by Pearson's product moment correlation coefficient.  
220 For heatmap illustration the gene order was computed by ordering the maximal peak of the loess fitted  
221 expression along the predefined spatial trajectory.

222

223 **Implantation of external tools: SPARK**

224 For pattern recognition of spatially distinct expressed genes we integrated the R package SPARK<sup>13</sup>,  
225 which was shown to perform beneficial compared to other tools such as SpatialDE<sup>18</sup>. We transformed  
226 the required data into a SPARK object which is externally analyzed and reimported to SPATA. We add  
227 the possibility to group genes with a significant spatial pattern by overlap estimation and SNN clustering  
228 of the given correlation matrix.

229

230 **Implantation of external tools: InferCNV**

231 Copy-number Variations (CNVs) were estimated by aligning genes to their chromosomal location and  
232 applying a moving average to the relative expression values, with a sliding window of 100 genes within  
233 each chromosome, as described recently<sup>16</sup>. First, we arranged genes in accordance to their respective  
234 genomic localization using the CONICSmat package (R-software). As a reference set of non-malignant  
235 spots, we used cortex from epilepsy patients. To avoid the considerable impact of any particular gene  
236 on the moving average we limited the relative expression values [-2.6,2.6] by replacing all values  
237 above/below  $exp(\vartheta)=|2.6|$ , by using the infercnv package (R-software). This was performed only in the  
238 context of CNV estimation as previously reported<sup>19</sup>.

239

240 **Implantation of external tools: Monocle3 or RNA-velocity**

241 We implemented a wrapper to easily switch between cds-objects (monocle3) and SPATA objects. First,  
242 we compute minimum spanning tree (MST) to estimate the most separate paths and order these cells  
243 to annotate pseudotime. By using the createPseudotime() function, a shiny-interface from  
244 monocle3 will give the possibility to select a root for pseudotime annotation. Further, we provide the  
245 possibility to implement each vector, for example “latent time” extracted from RNA-velocity using scvelo,  
246 to integrate into our SPATA object.

247

248 **Data acquisition of spatial transcriptomics**

249 All Visium Gene Expression experiments were performed according to 10X Genomics user guide  
250 ‘Visium Spatial Gene Expression Reagent Kits’. In brief, 10µm thick, cryosectioned slices of fresh frozen  
251 brain tissue were applied onto capture areas of Visium Spatial Gene Expression Slide, hematoxylin and  
252 eosin stained and imaged for subsequent alignment with spatial RNA data. During permeabilization,  
253 mRNA was liberated from cells and captured by primers on the slide’s surface which enable downstream

254 reassignment of barcoded mRNA sequences to their former, spatial location. Permeabilization times  
255 had been determined in advance (Cortex: 18 min; Tumor: 12 min; Cerebellum: 12 min) according to  
256 manufacturer's instructions (10X Genomics, Spatial Tissue Optimization Reagent Kit). After reverse  
257 transcription, second strand synthesis and denaturation of cDNA, second strands were amplified by  
258 PCR and desired cDNA fragments were selected via SPRIselect reagent. Successful amplification was  
259 confirmed by QC via Agilent Fragment Analyzer system. During the following fragmentation and double-  
260 sided size selection via SPRIselect reagent, length of cDNA fragments was optimized for analysis via  
261 Illumina NextSeq Sequencing System. Each fragment was provided with unique, dual indexes as well  
262 as adapters binding to oligonucleotides on Illumina flow cell. Post Library Construction QC via Agilent  
263 Fragment Analyzer system and Invitrogen Qubit Fluorometer was performed before normalization of  
264 libraries. For more information consult Illumina 'Denature and Dilute Libraries Guide - Protocol A:  
265 Standard Normalization Method'. Phix control at a concentration of 1.8pM was added to each library in  
266 a dilution of 1:100. Sequencing was performed using the NextSeq 500/550 High Output Kit (150 Cycles).

267

268

## 269 **Data and code Availability**

270 Further information and requests for resources, raw data and reagents should be directed and will be  
271 fulfilled by the Contact: D. H. Heiland, [dieter.henrik.heiland@uniklinik-freiburg.de](mailto:dieter.henrik.heiland@uniklinik-freiburg.de). The source code of  
272 SPATA is available at <https://github.com/theMILab/SPATA>, additional functions are at  
273 [https://github.com/heilanddd/SPATA\\_Developer](https://github.com/heilanddd/SPATA_Developer) and <https://github.com/kueckelj/confuns>. Spatial  
274 Transcriptomic data will be provided at GEO (in preparation) and SPATAobjects at [www.themilolab.com](http://www.themilolab.com)  
275 (in preparation).

276

## 277 **Acknowledgement**

278 This study is funded by the Else Kröner-Fresenius-Stiftung (2020\_EKSMS.24) (DHH), BMBF  
279 (MEPHISTO) (DHH and DD) and the Krebshilfe (Mildred-Scheel-Program) (JK and JE)

280

## 281 **Conflict of interests**

282 No potential conflicts of interest were disclosed by the authors.

283

284 Bibliography

285

- 286 1. Rosenberg, A. B. *et al.* Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**, 176–182 (2018).
- 287 2. Tirosh, I. *et al.* Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma. *Nature* **539**, 309–313 (2016).
- 288 3. Neftel, C. *et al.* An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell* **178**, 835–849.e21 (2019).
- 289 4. Cang, Z. & Nie, Q. Inferring spatial and signaling relationships between cells from single cell transcriptomic data. *Nat. Commun.* **11**, 2084 (2020).
- 290 5. Chen, K. H., Boettiger, A. N., Moffitt, J. R. & Wang, S. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* (2015).
- 291 6. Eng, C.-H. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH. *Nature* **568**, 235–239 (2019).
- 292 7. Rodrigues, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).
- 293 8. Berglund, E. *et al.* Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nat. Commun.* **9**, 2419 (2018).
- 294 9. Vickovic, S. *et al.* High-density spatial transcriptomics arrays for in situ tissue profiling. *BioRxiv* (2019). doi:10.1101/563338
- 295 10. Fernández Navarro, J., Lundeberg, J. & Ståhl, P. L. ST viewer: a tool for analysis and visualization of spatial transcriptomics datasets. *Bioinformatics* **35**, 1058–1060 (2019).
- 296 11. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
- 297 12. Dries, R. *et al.* Giotto, a pipeline for integrative analysis and visualization of single-cell spatial transcriptomic data. *BioRxiv* (2019). doi:10.1101/701680
- 298 13. Sun, S., Zhu, J. & Zhou, X. Statistical analysis of spatial expression patterns for spatially resolved transcriptomic studies. *Nat. Methods* **17**, 193–200 (2020).
- 299 14. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
- 300 15. Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* (2020). doi:10.1038/s41587-020-0591-3
- 301 16. Patel, A. P. *et al.* Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **344**, 1396–1401 (2014).
- 302 17. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
- 303 18. Svensson, V., Teichmann, S. A. & Stegle, O. SpatialDE: identification of spatially variable genes. *Nat. Methods* **15**, 343–346 (2018).
- 304 19. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).

324