Reference-free cell-type deconvolution of multi-cellular pixel-resolution spatially resolved transcriptomics data

Brendan F. Miller[1,2], Feiyang Huang[1,2], Lyla Atta[1,2], Arpan Sahoo[1,3], Jean Fan[1,2,3,*]

[1]Center for Computational Biology, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD 21211, USA

[2]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21218, USA

[3]Department of Computer Science, Johns Hopkins University, Baltimore MD 21218, USA

[*]To whom correspondence should be addressed

Correspondence should be addressed to:

Jean Fan (jeanfan@jhu.edu)

Key words:

Bioinformatics, Computational biology, Gene expression, Single Cell, Deconvolution, Spatial Transcriptomics, Topic Modeling

Abstract

Recent technological advancements have enabled spatially resolved transcriptomic profiling but at multi-cellular pixel resolution, thereby hindering the identification of cell-type-specific spatial patterns and gene expression variation. To address this challenge, we developed STdeconvolve as a reference-free approach to deconvolve underlying cell-types comprising such multi-cellular pixel resolution spatial transcriptomics (ST) datasets. Using simulated as well as real ST datasets from diverse spatial transcriptomics technologies comprising a variety of spatial resolutions such as Spatial Transcriptomics, 10X Visium, DBiT-seq, and Slide-seq, we show that STdeconvolve can effectively recover cell-type transcriptional profiles and their proportional representation within pixels without reliance on external single-cell transcriptomics references. STdeconvolve provides comparable performance to existing reference-based methods when suitable single-cell references are available, as well as potentially superior performance when suitable single-cell references are not available. STdeconvolve is available as an open-source R software package with the source code available at https://github.com/JEFworks-Lab/STdeconvolve.

35 **Introduction**

36 Delineating the spatial organization of transcriptionally distinct cell-types within tissues
37 is critical for understanding the cellular basis of tissue function[1]. Recent technologies have
38 enabled spatial transcriptomic (ST) profiling within tissues at multi-cellular pixel-resolution[2]. As
39 such, these ST measurements represent cell mixtures that may comprise multiple cell-types. This
40 lack of single-cell resolution hinders the characterization of cell-type specific spatial
41 organization and gene expression variation.

42 To address this challenge, several reference-based, supervised and semi-supervised
43 deconvolution approaches have recently been developed to predict the proportion of cell-types
44 within ST pixels. Of these, SPOTlight[3] uses cell-type marker genes derived from a single-cell
45 RNA-sequencing (scRNA-seq) reference to seed a non-negative matrix factorization. RCTD[4]
46 uses the cell-type specific mean expression of marker genes derived from a scRNA-seq reference
47 to build a probabilistic model of the contribution of each cell-type to the observed gene counts in
48 each pixel. SpatialDWLS[5] uses cell-type signature genes derived from a scRNA-seq reference to
49 first enrich for cell-types likely to be in each pixel, then applies a dampened weighted least
50 squares approach to infer the cell-type composition. DSTG[6] uses synthetic pseudo-mixtures of
51 scRNA-seq references to train a semi-supervised graph-based convolutional network. As such,
52 these approaches rely on the availability of a suitable single-cell reference, which may present
53 limitations if such a reference does not exist due to budgetary, technical[7], or biological
54 limitations[8]. While the rise of scRNA-seq references through atlasing efforts such as the BRAIN
55 Initiative Cell Census Network (BICCN)[9], the Human BioMolecular Atlas Program
56 (HuBMAP)[10], and Human Cell Atlas[11] may help alleviate such limitations particularly for
57 healthy tissues, processing independent tissue samples or different sections of the same tissue

58 may still result in systematically different gene expression quantifications due to batch effects as
59 well as inter- and intra-sample heterogeneity. Additionally, difficulties dissociating and capturing
60 certain cell-types via single-cell sequencing may result in missing or inconsistent cell-types
61 between scRNA-seq references and ST datasets[12,13]. Further, scRNA-seq references and ST
62 datasets may be affected by different perturbations manifesting as distinct transcriptional
63 differences affecting reference-based deconvolution accuracy and subsequent biological
64 interpretations. As such, a reference-free deconvolution approach provides an alternative strategy
65 for deconvolving cell-types when an appropriate reference is not available.

66 Here, we developed STdeconvolve (available at https://github.com/JEFworks-
67 Lab/STdeconvolve and as Supplementary Software) as a reference-free, unsupervised approach
68 for deconvolving multi-cellular pixel resolution ST data (Figure 1). STdeconvolve builds on
69 latent Dirichlet allocation (LDA), a generative statistical model commonly used in natural
70 language processing for discovering latent topics in collections of documents. In the context of
71 natural language processing, given a count matrix of words in documents, LDA infers the
72 distribution of words for each topic and the distribution of topics in each document. In the
73 context of ST data, given a count matrix of gene expression in multi-cellular ST pixels,
74 STdeconvolve applies LDA to infer the putative transcriptional profile for each cell-type and the
75 proportional representation of each cell-type in each multi-cellular ST pixel (Methods). While
76 LDA has previously been applied in the context of deconvolving cell-types in bulk RNA-seq
77 data[14], STdeconvolve leverages several unique features of ST data that make this application of
78 LDA particularly amenable (Supplementary Note 1). Briefly, these include i) the limited number
79 of cells and cell-types represented in each ST pixel, ii) the limited impact of batch effects on the
80 measured gene expression across pixels, iii) the large number of pixels compared to cell-types,

81 and iv) the likely heterogeneity of cell-type proportional distribution across pixels in tissues. To
82 improve the application of LDA, where latent cell-types are characterized by co-expressed and
83 ideally non-overlapping groups of genes, STdeconvolve feature selects for such genes likely to
84 be informative of latent cell-types. Specifically, STdeconvolve selects for significantly
85 overdispersed genes, or genes with higher-than-expected expression variance across pixels[15]
86 (Methods). In addition, as the application of LDA requires the number of transcriptionally
87 distinct cell-types, *K*, to be set *a priori*, STdeconvolve provides several data-driven metrics to
88 guide the estimation of an appropriate *K* (Methods, Supplementary Note 2).
89

90 **Results**

91 **STdeconvolve accurately recovers cell-type proportions and transcriptional profiles in**
92 **simulated ST data**

93 As a proof of concept, we first evaluated the performance of STdeconvolve in recovering
94 the proportional representations of cell-types and their transcriptional profiles using simulated
95 ST data. We simulated ST data by aggregating the gene expression of cells from single-cell
96 resolution multiplex error-robust fluorescence *in situ* hybridization (MERFISH) data of the
97 mouse medial pre-optic area (MPOA)[16] within spatially contiguous pixels. Previously,
98 MERFISH was previously applied to map the spatial distribution of 135 select genes within
99 MPOA brain tissue. These select 135 genes were chosen to distinguish between major non-
100 neuronal cell-types as well as neuronal subtypes. Imaging-based cell segmentation was
101 performed and the counts of genes per cell were quantified to achieve single-cell resolution
102 spatially resolved transcriptomic profiling. Subsequent transcriptional clustering analysis on the
103 single-cell resolution gene expression measurements identified 9 major cell-types, including

104 excitatory and inhibitory neurons. Further clustering found that these excitatory and inhibitory
105 neurons could be subdivided into 69 finer neuronal sub-types.
106 To simulate multi-cellular pixel-resolution ST data, we aggregated the single-cell
107 resolution MERFISH data into 100 μm^2 pixels (Figure 2A, Supplementary Figure S1A-B,
108 Methods). Given the already limited 135 gene panel chosen to distinguish between cell-types,
109 additional feature selection for this dataset was not necessary (Supplementary Note 3). Applying
110 STdeconvolve, we identified K=9 cell-types and deconvolved their proportional representation
111 and transcriptional profiles in each simulated pixel (Figure 2B, Supplementary Figure S1C, S2A,
112 Supplementary Methods). To infer the identities of the deconvolved cell-types for benchmarking
113 purposes, we matched their deconvolved transcriptional profiles with the transcriptional profiles
114 of ground truth cell-types by testing for enrichment of ground truth cell-type specific marker
115 genes (Methods, Supplementary Figure S2B). We observed strong correlations between the
116 transcriptional profiles of each deconvolved cell-type and matched ground truth cell-type across
117 genes (Figure 2C). Likewise, we observed strong correlations between the proportions of each
118 deconvolved cell-type and matched ground truth cell-type across simulated pixels (Figure 2D).
119 We further quantified this performance using the root-mean-square-error (RMSE) of the
120 deconvolved cell-type proportions compared to ground truth across simulated pixels (Methods,
121 Figure 2E). In this manner, STdeconvolve can accurately recover the proportional representation
122 and transcriptional profiles of major cell-types.
123
124 **STdeconvolve achieves competitive performance to reference-based, supervised**
125 **deconvolution approaches**

126 We next sought to compare the performance of STdeconvolve to existing reference-
127 based, supervised and semi-supervised deconvolution approaches SPOTlight, RCTD,
128 spatialDWLS, and DTSG using our simulated 100 μm^2 resolution ST data of the MPOA. As
129 described previously, these approaches require a single-cell transcriptomics reference for
130 deconvolution. As an ideal single-cell transcriptomics reference, we used the original single-cell
131 MERFISH data that was used to construct the simulated ST data (Supplementary Figure S3A,
132 Supplementary Methods). We again quantified the performance of each approach using the
133 RMSE of the deconvolved cell-type proportions compared to ground truth across simulated
134 pixels. DSTG was unable to deconvolve distinct cell-types in the data and was omitted from
135 further comparison (Supplementary Figure S3A). In general, we find the performance of
136 STdeconvolve to be comparable to these reference-based deconvolution approaches when such
137 an ideal single-cell transcriptomics reference is used (Figure 2E-F).
138 One potential limitation of such existing reference-based deconvolution approaches is
139 their reliance on a suitable single-cell transcriptomics reference. We thus sought to evaluate the
140 performance of these reference-based deconvolution approaches when a suitable single-cell
141 reference is not available. To this end, we removed excitatory and inhibitory neuronal cell-types
142 to simulate a less suitable single-cell transcriptomics reference (Supplementary Methods). We
143 then deconvolved the simulated ST data of the MPOA using each reference-based deconvolution
144 approach with this new reference and computed the RMSE across pixels. Because STdeconvolve
145 does not use a reference, its performance does not change. However, the performance for all
146 reference-based deconvolution approaches resulted in a significantly higher RMSE (Diebold-
147 Mariano p-value $< 2.2 \times 10^{-16}$) than STdeconvolve (Figure 2G). Likewise, pixels previously
148 comprised of neurons were now erroneously predicted by reference-based deconvolution

149 approaches to be comprised primarily of immature oligodendrocytes (Supplementary Figure
150 S3B). In addition, we evaluated the performance of each reference-based deconvolution
151 approach after removing rarer ependymal cells from the single-cell transcriptomics reference.
152 Again, given this less suitable single-cell transcriptomics reference, pixels previously comprised
153 of ependymal cells were now erroneously predicted by reference-based deconvolution
154 approaches to be comprised primarily of astrocytes (Supplementary Figure S3C). Thus, the
155 performance of reference-based deconvolution approaches is sensitive to differences in cell-type
156 composition between the ST data and the single-cell transcriptomics reference used.
157 Likewise, such an ideal single-cell transcriptomics reference that optimally matches the
158 cell-type composition and measurement sensitivities of the ST data to be deconvolved may not
159 be available. Therefore, this ideal MERFISH MPOA single-cell transcriptomics reference likely
160 provides an upper bound on performance for reference-based deconvolution approaches. To
161 provide a more realistic evaluation of performance for reference-based deconvolution
162 approaches, we sought to deconvolve our simulated ST data of the MPOA using a scRNA-seq
163 reference from a mouse brain atlasing effort[17]. Again, as a reference-free deconvolution
164 approach, the performance of STdeconvolve does not change. However, again, the performance
165 for all reference-dependent methods resulted in a significantly higher RMSE (Diebold-Mariano
166 p-value $< 2.2 \times 10^{-16}$) than STdeconvolve (Figure 2H, Supplementary Methods). Thus,
167 STdeconvolve achieves comparable performance to reference-based, supervised deconvolution
168 approaches when an ideal single-cell transcriptomics reference is used, and potentially better
169 performance when an ideal single-cell transcriptomics reference is not used.
170
171 **STdeconvolve recovers perturbation specific gene expression profiles**

172 Though reference-based deconvolution approaches may accurately recover cell-type
173 proportions in ST data, they currently do not deconvolve cell-type specific gene expression
174 profiles. Nonetheless, perturbations may induce cell-type-specific transcriptional changes in ST
175 data that would not be identifiable by current reference-based deconvolution approaches unless
176 perturbation-matched single-cell transcriptomics references are used. While the availability of
177 scRNA-seq references grows due to single-cell atlasing initiatives, these datasets primarily
178 represent collections of healthy tissues[9-11,18,19]. As such, there is a particular scarcity of suitable
179 scRNA-seq references available for reference-based deconvolution of ST data in the context of
180 disease and other perturbations.

181 In contrast to current reference-based deconvolution approaches, STdeconvolve can
182 estimate cell-type transcriptional profiles in a manner that is not constrained by the expression
183 profiles of specific cell-types defined in single-cell transcriptomics references. We therefore
184 sought to explore the potential of STdeconvolve in detecting these perturbation-driven cell-type-
185 specific gene expression changes using simulated ST data from mixtures of single cells assayed
186 by scRNA-seq (Figure 3A, Supplementary Methods). Briefly, we took advantage of scRNA-seq
187 data previously collected from mammary tissues of aged and young mice[20]. Previous
188 transcriptional clustering analysis revealed a subpopulation of macrophages with age-associated
189 gene expression changes. Specifically, aged macrophages upregulated *Cd274* and *Clec4d*, and
190 downregulated *Coro1a* compared to young macrophages. Therefore, we simulated ST data of
191 aged tissue using mixtures of aged macrophages and other luminal cells and ST data of young
192 tissue using mixtures of young macrophages and other luminal cells (Figure 3B). We then sought
193 to evaluate the ability of STdeconvolve to recover these age-associated gene expression changes
194 in macrophages (Supplementary Methods). Applying STdeconvolve using K=2 cell-types to the

195 simulated ST data of both aged and young tissue, we found that the deconvolved transcriptional
196 profiles were highly correlated with the matched ground truth gene expression profiles from
197 scRNA-seq in all cases (Supplementary Figure S4A-B). Further, when we compared the
198 deconvolved transcriptional profiles of aged versus young macrophages, we were able to identify
199 upregulated genes included *Cd274* and *Clec4d*, and downregulated genes included *Coro1a*,
200 consistent with the original study (Figure 3C). Thus, STdeconvolve can potentially recover
201 perturbation-driven cell-type-specific gene expression changes in ST data.
202
203 **Deconvolution provides distinct insights compared to clustering analysis**
204 Generally, we note that deconvolution of multi-cellular pixel resolution ST data can
205 provide distinct insights from clustering analysis. To demonstrate this, we again simulated ST
206 data using mixtures of single cells assayed by scRNA-seq (Supplementary Methods).
207 Specifically, we simulated ST pixels comprised of mixtures of either luminal cells and pericytes
208 or pericytes and macrophages (Figure 3D). Applying clustering analysis to these ST pixels, we
209 identified 2 clusters corresponding to either mixtures of luminal cells and pericytes or mixtures
210 of pericytes and macrophages (Figure 3E). In contrast, applying STdeconvolve with K=3, we
211 were able to recover the proportional representations of luminal cells, pericytes, and
212 macrophages as well as their original cell-type specific transcriptional profiles (Figure 3F).
213 Such differences between deconvolution and clustering analysis extends to resolution
214 enhancing clustering approaches such as BayesSpace[21]. Briefly, BayesSpace utilizes a spatial
215 prior that encourages spatially neighboring pixels to cluster into the same transcriptional cluster.
216 Enhanced resolution clustering is obtained after subdividing each pixel and modeling the
217 expression profiles of the subpixels as additional latent parameters estimated in the Bayesian

218 model. Applying BayesSpace with 3 clusters to our simulated ST data, we obtained 3 spatially
219 discrete clusters corresponding to different mixtures of luminal cells and pericytes and mixtures
220 of pericytes and macrophages (Figure 3G, Supplementary Methods). Compared to
221 STdeconvolve, both regular transcriptional clustering and resolution enhanced clustering with
222 BayesSpace exhibited significantly higher RMSE (Diebold-Mariano p-value $< 2.2 \times 10^{-16}$)
223 (Figure 3H).

224 To further demonstrate the difference between deconvolution and clustering analysis for
225 ST data, we again simulated ST data using a single-cell resolution MERFISH dataset of a
226 coronal section of the mouse brain[22]. We analyzed the single-cell resolution transcriptional
227 profiles to identify 20 transcriptionally distinct cell-types and again simulated multi-cellular
228 pixel-resolution ST data by aggregating the single cells into 100 μm^2 pixels (Figure 3I,
229 Supplementary Figure S5A-B, Supplementary Methods). The organization of cell-types within
230 the mouse brain is highly complex with many regions including the thalamus at the central
231 region of this coronal section being composed of mixtures of multiple transcriptionally distinct
232 cell-types. We thus sought to evaluate whether STdeconvolve could better recover the
233 proportional representation of cell-types compared to resolution enhanced clustering with
234 BayesSpace. Applying both STdeconvolve and BayesSpace, we generally recover the cell-type
235 pixel proportions and visually recapitulate the spatial organization of cell-types within various
236 brain structures (Figure 3J-K, Supplementary Methods). However, focusing in on the central
237 region of the coronal section encompassing the thalamus, we indeed saw a visual difference
238 between the spatial organization of cell-types recovered by deconvolution via STdeconvolve
239 compared to resolution enhanced clustering via BayesSpace (Figure 3J-K inset). Quantifying
240 performance, BayesSpace exhibited significantly higher RMSE compared to STdeconvolve

241 (Diebold-Mariano p-value < 2.2×10^{-16}) as a whole (Supplementary Fig S5C, though more
242 discernably in the thalamus region (Figure 3L). Taken together, deconvolution approaches such
243 as STdeconvolve can provide distinct results from clustering and resolution enhanced clustering
244 approaches when applied to multi-cellular pixel resolution data.
245
246 **STdeconvolve characterizes the spatial organization of transcriptionally distinct cell-types**
247 **in real ST data**
248 Having demonstrated the capacity of STdeconvolve to recover cell-type proportions and
249 transcriptional profiles in simulated ST data, we next sought to evaluate the performance of
250 STdeconvolve by analyzing real 100 μm^2 resolution ST data of the mouse main olfactory bulb
251 (MOB)[23]. The MOB consists of multiple bilaterally symmetric and transcriptionally distinct cell
252 layers due to topographically organized sensory inputs[24]. While previous clustering analysis of
253 ST data of the MOB revealed coarse spatial organization of coarse cell layers, finer structures
254 such as the rostral migratory stream (RMS) could not be readily observed (Supplementary Figure
255 S6A-B). We applied STdeconvolve to identify K=12 cell-types (Figure 4A, Supplementary
256 Figure S6C, Supplementary Methods) that either overlapped with or further split coarse cell
257 layers previously identified from clustering analysis (Supplementary Figure S6D). In particular,
258 deconvolved cell-type X7 overlapped with the granule cell layer previously identified from
259 clustering analysis and was spatially placed where the RMS is expected[25] (Figure 4B).
260 Upregulated genes in its deconvolved transcriptional profile, including *Nrep*, *Sox11,* and *Dcx*, are
261 known to be associated with neuronal differentiation and upregulated in neuronal precursor cells
262 within the RMS[26] (Figure 4C, Supplementary Figure S6E). Higher resolution ISH staining of
263 these genes further demarcates a region within the granule cell layer where the RMS is

264 expected[19] (Figure 4D). This suggests that deconvolved cell-type X7 may correspond to the
265 neuronal precursor cell-type within the RMS unidentified from clustering analysis.
266 To further evaluate the biological reproducibility of deconvolved cell-types, we applied
267 STdeconvolve independently to 3 additional biological replicates of ST data of the MOB
268 (Supplementary Methods). In each biological replicate, STdeconvolve consistently identified
269 approximately 12 cell-types (Supplementary Figure S7A). Transcriptional profiles between
270 deconvolved cell-types were also highly correlated across biological replicates (Supplementary
271 Figure S7B-D). This suggests that STdeconvolve can reliably deconvolve consistent cell-types,
272 even across biological replicates.
273 As noted previously using simulated ST data, the performance of reference-based
274 deconvolution approaches is sensitive to differences in cell-type composition between the single-
275 cell transcriptomics reference and the ST data to be deconvolved. To demonstrate this with real
276 ST data, we first compared STdeconvolve and reference-based deconvolution approaches using
277 an appropriate MOB scRNA-seq reference[27] (Supplementary Methods). We found strong
278 correlations between cell-type proportions estimated by STdeconvolve and other reference-based
279 deconvolution approaches with a high degree of correspondence among all evaluated methods
280 (Supplementary Figure S8A-B). Notably, the proportion and transcriptional profile of
281 deconvolved cell-type X8 identified by STdeconvolve to be enriched in the olfactory nerve layer
282 correlated strongly with the proportion of olfactory ensheathing cells (OECs) identified by the
283 reference-based deconvolution approaches.
284 Next, to simulate a less suitable scRNA-seq reference, we removed OECs from the MOB
285 scRNA-seq reference and again evaluated the performance of reference-based deconvolution
286 approaches given this new scRNA-seq reference without OECs (Supplementary Methods).

287 Again, as a reference-free deconvolution approach, the results of STdeconvolve do not change.
288 However, for some reference-based deconvolution approaches, given this new reference without
289 OECs, pixels in the olfactory nerve layer previously comprised of OECs were now predicted to
290 be comprised of N2 cells (Supplementary Figure S9A). Although we do not know the ground
291 truth cell-type composition of this olfactory nerve layer, we have reasons to believe that this
292 placement of N2 cells is erroneous. First, when a scRNA-seq reference with OECs was used,
293 reference-based deconvolution approaches generally estimated N2 cells to be relatively rare.
294 However, when a scRNA-seq reference without OECs was used, reference-based deconvolution
295 approaches substantially increased their estimated abundance of N2 cells (Supplementary Figure
296 S9B). Second, while the transcriptional profiles of OECs and N2 cells are highly correlated
297 (Supplementary Figure S9C), the two cell-types exhibit significant transcriptionally differences.
298 For example, top differentially upregulated genes in OECs are highly expressed in the olfactory
299 nerve layer (Supplementary Figure S8C) whereas top differentially upregulated genes in N2 cells
300 are not well detected in the olfactory nerve layer (Supplementary Figure S9D). This lack of
301 detection of N2 cell marker genes within the olfactory nerve layer coupled with the rarity of N2
302 cells in the original reference-based deconvolution with OECs suggests that the placement of N2
303 cells in the olfactory nerve layer by reference-based deconvolution approaches when using a
304 reference without OECs is erroneous.

305 Further, a single-cell transcriptomics reference may not always exist for the same tissue
306 from which ST data was generated, prompting the use of a reference from a related but
307 inherently different tissue source. To evaluate the potential effect of using a single-cell
308 transcriptomics reference from a different tissue source on reference-based deconvolution
309 approaches, we sought to deconvolve the MOB ST data using the scRNA-seq reference from the

310 mouse brain described previously. Given this mouse brain reference, pixels in the olfactory nerve
311 layer previously comprised of OECs were now predicted to be comprised of vascular
312 leptomeningeal cells (VLMC) (Supplementary Figure S10A). Again, although we do not know
313 the ground truth cell-type composition of this olfactory nerve layer, top differentially upregulated
314 genes in VLMCs are not well detected in the MOB (Supplementary Figure S10B) and are
315 therefore likely not truly present. Taken together, all this suggests that reference-based
316 deconvolution approaches are sensitive to the cell-types represented in the single-cell
317 transcriptomics reference that is used, which may lead to inaccurate results and spurious cell-
318 type assignments when a suitable reference is not available.
319
320 **STdeconvolve is applicable across diverse ST dataset resolutions and technologies**
321 We anticipate that continual technological improvements will enhance the resolution of
322 ST data. Already, ST technologies such as Visium (10X Genomics), Slide-seq[28], and DBiT-seq[29]
323 have achieved resolution that can range from 50 μm^2 to 10 μm^2. Therefore, we sought to
324 evaluate the performance of STdeconvolve on higher resolution ST data using both simulated as
325 well as real ST data from higher resolution ST technologies including Visium, Slide-seq, and
326 DBiT-seq.
327 First, to simulate higher resolution ST data, we again aggregated single-cell resolution
328 MERFISH data of the MPOA into 50, 20, and 10 μm^2 resolution pixels. Applying
329 STdeconvolve, we observed similarly strong correlations between the deconvolved cell-type
330 transcriptional profiles and proportions with the ground truth (Supplementary Figure S11A-D).
331 Although the number of cells in each multi-cellular pixel did decrease as the resolution of the
332 pixel increased as expected, we note that even higher resolution pixels may still contain multiple

333 cells representing multiple cell-types (Supplementary Figure S11E-F). Thus, deconvolution may
334 still be applicable to higher resolution ST data and STdeconvolve can still accurately deconvolve
335 cell-types within these higher resolution multi-cellular pixels.

336 Encouraged by STdeconvolve's ability to recover cell-types in simulated high-resolution
337 ST data, we then applied STdeconvolve to real high-resolution multi-cellular ST data from
338 several different technologies. First, we applied STdeconvolve to 50 μm^2 resolution ST data of a
339 coronal section of the mouse brain from 10X Visium[30]. Briefly, for 10XVisium, mRNAs from
340 tissue sections are captured onto an array of DNA barcoded spots, resulting in RNA-sequencing
341 measurements with gridded 2D spatial positional information. We applied STdeconvolve to
342 identify K=20 cell-types that exhibit spatially distinct patterns that demarcate known brain
343 structures such as the isocortex and fiber tracts (Figure 4E, Supplementary Figure S12,
344 Supplementary Methods).

345 We next applied STdeconvolve to 25 μm^2 resolution ST data of the lower body of the
346 E11 mouse embryo from DBiT-seq. Briefly, for DBiT-seq, parallel microfluidic channels are
347 used to deliver DNA barcodes to the surface of a tissue to enable direct barcoding of mRNAs *in*
348 *situ*, resulting in RNA-sequencing measurements in a 2D mosaic of spatial pixels. Previously, the
349 authors identified 13 transcriptionally and spatially distinct features in the E11 mouse embryo
350 including the atrium, ventricle, liver, and blood vessels containing erythrocyte coagulation.
351 Applying STdeconvolve with K=13, we identify deconvolved cell-types that corresponded with
352 similar spatially distinct features in agreement with the original findings (Figure 4F,
353 Supplementary Figure S13A, Supplementary Methods). Moreover, the top genes in the
354 deconvolved cell-type specific transcriptional profiles contained the expected marker genes of
355 the matching features, such as *Myh6* for the atrium, *Myh7* for the ventricle, *Apoa2* for the liver,

356 and *Hba.a2* for the blood vessels containing erythrocyte coagulation in agreement with the
357 original findings (Supplementary Figure S13B).

358 Finally, we applied STdeconvolve to 10 μm^2 resolution ST data of the mouse cerebellum
359 from Slide-seq. Briefly, for Slide-seq, mRNAs from tissue sections are captured onto densely
360 packed barcoded beads, resulting in RNA-sequencing measurements with 2D spatial positional
361 information. Previously, RCTD was also applied to this Slide-seq dataset with a matched Drop-
362 seq scRNA-seq reference of the mouse cerebellum[31] to identify beads representing a distinct
363 layers of Purkinje neurons and Bergmann glia. Applying STdeconvolve, we identified K=14 cell-
364 types (Figure 4G, Supplementary Methods) whose transcriptional profiles correlated strongly
365 with cell-types from the scRNA-seq dataset of the mouse cerebellum (Supplementary Figure
366 S14A). In particular, we found that the deconvolved transcriptional profiles of cell-type X4 and
367 cell-type X2 correlated strongly with the transcriptional profiles of Purkinje neurons and
368 Bergmann glia. Likewise, the deconvolved proportional representation of cell-type X4 and cell-
369 type X2 also agreed significantly (Fisher's Exact p-value $< 2.2 \times 10^{-16}$) with the predicted
370 proportions of Purkinje neuron and Bergmann glia from RCTD (Supplementary Figure S14B-C).
371 Taken together, these results indicate that STdeconvolve can be applicable to a range of multi-
372 cellular resolution ST technologies.

373 As the resolution of ST data improves, the number of spatially resolved pixels and cell-
374 types represented in the data will presumably also increase. We therefore sought to evaluate the
375 scalability of STdeconvolve in anticipation of these increasingly larger datasets. To this end, we
376 benchmarked the runtime and total memory usage by STdeconvolve when deconvolving varying
377 numbers of cell-types using varying numbers of genes across varying numbers of pixels
378 (Methods). We found that both the runtime and memory usage by STdeconvolve increased

379 linearly with the number of pixels and genes in the input dataset (Supplementary Figure S15A)
380 and is comparable to existing reference-based deconvolution methods when applied to current
381 ST datasets[5]. Likewise, runtime scales with the number of deconvolved cell-types *K* in the input
382 dataset though memory usage remains stable (Supplementary Figure S15B). To enhance runtime
383 efficiency, STdeconvolve has built in parallelization. In this manner, we anticipate that
384 STdeconvolve will be amenable to larger ST data.
385
386 **STdeconvolve identifies immune infiltrates in breast cancer**
387 Finally, to demonstrate the potential of an unsupervised, reference-free deconvolution
388 approach, we applied STdeconvolve to 100 μm^2 resolution ST data of 4 breast cancer sections[32].
389 Here, a matched scRNA-seq reference was not available and using a scRNA-seq reference from
390 another breast cancer sample may be inappropriate due to potential inter-tumoral heterogeneity[33].
391 Transcriptional clustering of the ST pixels previously identified 3 transcriptionally distinct
392 clusters that corresponded to 3 histological regions of the tissue: ductal carcinoma *in situ* (DCIS),
393 invasive ductal carcinoma (IDC), and non-malignant[32] (Figure 5A, Supplementary Figure 16A-
394 B). However, the tumor microenvironment is a complex milieu of many additional cell-types[34].
395 We thus applied STdeconvolve to identify potential additional cell-types and interrogate their
396 spatial organization, resulting in *K*=15 identified cell-types (Figure 5B, Supplementary Figure
397 S16C, Online Methods). Of these, deconvolved cell-types X3 and X13 pixel proportions
398 corresponded spatially with pixels annotated as the non-malignant and DCIS regions,
399 respectively (Supplementary Figure S16D). Likewise, the deconvolved expression profiles for
400 X3 and X13 included *KRT1*, a keratin gene specifically expressed in mammary myoepithelial
401 cells[35], and *PRSS23*, a serine protease associated with proliferation of breast cancer cells[36],

402 respectively, consistent with the non-malignant and DCIS annotations (Supplementary Figure
403 S17). Interestingly, the deconvolved expression profile for cell-type X15 included immune genes
404 such as *CD74* and *CXCL10* (Figure 5C-E, Supplementary Figure S18). Gene set enrichment
405 analysis also suggested that genes in the deconvolved expression profile for cell-type X15 was
406 significantly enriched in immune processes such as T cell activation (Supplementary Table S1,
407 Online Methods). This suggests that deconvolved cell-type X15 may correspond to immune
408 infiltrates. Further, we find a significant the number of pixels with a high proportion of
409 deconvolved cell-type X15 corresponding to IDC regions (Fisher's exact p-value = 0.001257)
410 based on previous clustering and pathology annotations. In contrast, we do not see a significant
411 number of pixels with a high proportion of deconvolved cell-type X15 corresponding to DCIS
412 regions (Fisher's exact p-value = 0.5625). This is consistent with previous observations that
413 when comparing pure DCIS and IDC, infiltration of immune cells was significantly higher in
414 IDC to pure DCIS[37,38].

415 The spatial organization of immune cells within tumors has been previously implicated to
416 be relevant in breast cancer prognosis[39]. In particular, whether immune cells are infiltrated or
417 excluded from the tumor is associated with tumor microenvironments that stratify patient
418 outcomes[40]. To evaluate whether STdeconvolve may be able to distinguish infiltrated versus
419 excluded spatial organization of immune cells in tumors, we simulated ST data representing
420 infiltrated and excluded spatial organizations using mixtures of single cells assayed by scRNA-
421 seq (Figure 5F, Online Methods). In both the simulated infiltrated and excluded cases, we find
422 that STdeconvolve can effectively recover the cell-type transcriptional profiles (Figure 5G) and
423 enable the quantification of immune infiltration to help distinguish between infiltrated versus
424 excluded spatial organization of immune cells (Figure 5H). Therefore, we anticipate that

425 STdeconvolve may be able to assist in deconvolving cell-types in heterogeneous cancer tissues
426 to recover potentially clinically interesting spatial organizational patterns.
427
428 **Discussion**
429 Multi-cellular pixel-resolution ST technologies have enabled high-throughput
430 transcriptomic profiling of small mixtures of cells within tissues but accurate identification of the
431 underlying cell-types within each pixel is critical for elucidating cell-type specific spatial
432 organizational patterns and gene expression variation. Although several deconvolution methods
433 have already been developed to address this challenge, they currently rely on suitable single-cell
434 transcriptomics references. As we have shown, this reliance on single-cell transcriptomics
435 references constrains the spatial mapping of cell-types to those in the reference, which may
436 present limitations if there are missing cell-types, mismatched cell-types, perturbations, and
437 batch effect differences between the single-cell transcriptomics reference and ST data to be
438 deconvolved. Here, we have presented STdeconvolve, a reference-free computational approach
439 to deconvolve cell-type proportions and their transcriptional profiles in multi-cellular pixel
440 resolution ST data. We have demonstrated that STdeconvolve can accurately recover underlying
441 cell-type proportions and their transcriptional profiles across a range of different ST technologies
442 and resolutions. STdeconvolve further provides competitive performance to reference-based
443 deconvolution approaches when an ideal single-cell transcriptomics reference is available and
444 potentially better performance in more realistic circumstances where such an ideal reference is
445 not available. Additionally, we showed the advantage of deconvolution over clustering-based
446 analysis methods to interrogate heterogeneous mixtures of cell-types. Likewise, using simulated
447 ST data of aged-perturbed tissues, we showed that STdeconvolve can recover perturbation-

448 driven cell-type-specific gene expression changes. Finally, we applied STdeconvolve to identify
449 putative immune infiltration in real and simulated breast cancer ST data.

450 Though we have shown that STdeconvolve can effectively recover cell-type proportions
451 and transcriptional profiles in simulated and real ST data, its use of LDA modeling relies on
452 several underlying assumptions, which may present limitations when these assumptions are not
453 satisfied. Notably, the performance of LDA in accurately deconvolving cell-types depends on the
454 size of the dataset with respect to the number of pixels and the number of genes[41]. As such,
455 deconvolution accuracy generally decreases for ST data containing fewer than 10 pixels
456 (Supplementary Figure S19). While we have generally found the number of pixels in most ST
457 datasets to be well beyond 10 pixels after quality control filtering, the application of ST to profile
458 tissue slivers or other thin structures covering only a few pixels may present challenges to
459 deconvolution by STdeconvolve. Further, LDA modeling attempts to identify tightly occurring,
460 and ideally non-overlapping groups of genes in the pixels as cell-types. In this manner, if genes
461 do not exhibit variability across pixels due to a homogeneous or uniform proportional
462 representation of cell-types across pixels (Supplementary Figure S20), STdeconvolve may fail to
463 deconvolve distinct cell-types. Likewise, if the gene expression in the ST data is too sparse with
464 high rates of stochastic drop-outs[42], then the LDA model may struggle to identify distinct groups
465 of co-expressed genes and as such STdeconvolve may also struggle to deconvolve distinct cell-
466 types as well. Still, when such failures happen, STdeconvolve will indicate to users when distinct
467 cell-types are not detected.

468 Although we have demonstrated the applicability of STdeconvolve to high resolution
469 multi-cellular pixel resolution ST data, as the resolution of ST data continues to increase, sub-
470 cellular pixel-resolution ST technologies will also become more accessible. Already, a number

471 sub-cellular pixel-resolution ST technologies have emerged[43-47]. As the capture efficiency at this
472 resolution and likewise the biological questions of interest may differ substantially from multi-
473 cellular ST data, we anticipate that new methods specifically suited for sub-cellular resolution ST
474 data will be needed. Thus, STdeconvolve may not be best suited to analysis of such sub-cellular
475 resolution ST data. Still, as we have noted previously, even as the resolution of ST data
476 increases, some pixels may still contain multiple cells representing multiple cell-types,
477 suggesting that deconvolution may still be necessary. Likewise, the number of cells present in a
478 pixel ultimately will depend on cell size, which can vary depending on the organism, tissue,
479 and/or disease state being profiled. Ultimately, we believe that there will be a need to balance
480 between resolution and throughput of ST technologies depending on the biological question of
481 interest. The potentially larger tissue regions able to be covered by multi-cellular pixel resolution
482 ST data may still be of interest and thus still require deconvolution. We anticipate that
483 STdeconvolve will be applicable to data from a variety of current and future ST technologies as
484 well as potentially inferred ST data[48] to reveal cell-type specific spatial organizational patterns
485 and transcriptional changes. In general, we foresee that reference-free deconvolution approaches
486 such as STdeconvolve will contribute to the interrogation of the spatial relationships between
487 transcriptionally distinct cell-types in heterogeneous tissues.
488

489 **Methods**

490

491 **STdeconvolve Overview**

492 STdeconvolve uses latent Dirichlet allocation (LDA)[49], a generative probabilistic model, to
493 deconvolve the latent cell-types contained within multi-cellular pixels of spatially resolved
494 transcriptome (ST) measurements. In this context, each pixel is defined as a mixture of K cell-
495 types represented as a multinomial distribution of cell-type probabilities (θ), and each cell-type
496 is defined as a probability distribution over the genes (β) present in the ST dataset.

497

498 *LDA Modeling*

499 The ST dataset is represented as a $D \times N$ matrix of discrete gene expression counts for each
500 pixel d and gene n. The total number of unique molecules, or total gene expression, in a given
501 pixel d is M_d.

502

503 As a generative probabilistic model, the LDA model generates a set of new pixels as follows:

504

505 For each pixel d:

506 a. draw a cell-type distribution $\theta_d \sim Dir(\alpha)$, where θ_d is a multinomial distribution of
507 length K drawn from a uniform Dirichlet distribution with scaling parameter α.

508 b. for each observed molecule m in M_d:

509 i. draw cell-type assignment $z_{d,m} \sim mult(\theta_d)$

510 ii. draw a gene $w_{d,m} \sim mult(\beta_{z_{d,m}})$

511

512 The central goal is to identify the posterior distribution of the latent parameters given the input

513 data, where for each pixel d:

514 $$p(\theta_d, \boldsymbol{z} \mid \boldsymbol{w}, \alpha, \beta) = \frac{p(\theta_d, \boldsymbol{z}, \boldsymbol{w} \mid \alpha, \beta)}{p(\boldsymbol{w} \mid \alpha, \beta)}$$

515 where $\boldsymbol{z}$ is a vector of M_d cell-types assigned to each unique molecule in pixel d, and $\boldsymbol{w}$ is the

516 vector of M_d genes assigned to each unique molecule for pixel d. A variational expectation-

517 maximization approach is used to estimate the values of the latent parameters[49,50]. By default, β

518 is initialized with 0 for all cell-types and genes, and α as 50/K.

519

520 The resulting estimated θ and β matrices represent the deconvolved proportions of cell-types in

521 each pixel and the gene expression profiles for each cell-type, scaled to a library size of 1. β

522 represents a $K \times N$ gene-probability (*i.e.*, expression) matrix for each cell-type k and each gene

523 n with each row summing to 1. The β matrix can be multiplied by a scaling factor of one million

524 to be more like conventional counts-per-million expression values for interpretability. θ

525 represents a $D \times K$ pixel-cell-type proportion matrix for each pixel d and each cell-type k. LDA

526 modeling in STdeconvolve is implemented through the `topicmodels` R package[50].

527

528 Of note, LDA assumes for each cell-type that there is a group of genes highly co-expressed with

529 high probability. Therefore, STdeconvolve uses feature selection for genes more likely to be

530 highly co-expressed within cell-types, which can improve cell-type deconvolution.

531

532 *Selection of genes for LDA model*

533 Latent cell-types are best discovered by LDA modeling if cell-type specific marker genes

534 are included in the input ST data while genes whose expression is shared across cell-types are

535 excluded. Therefore, to filter for genes that are more likely to be specifically expressed in
536 particular cell-types to improve cell-type deconvolution by LDA, STdeconvolve first removes
537 genes that are not detected in a sufficient number of pixels. By default, genes detected in less
538 than 5% of pixels are removed. Likewise, STdeconvolve also removes genes that are expressed
539 in all pixels. By default, genes detected in 100% of pixels are removed. STdeconvolve then
540 selects for significantly overdispersed genes, or genes with higher-than-expected expression
541 variance across pixels, as a means to detect transcriptionally distinct cell-types[15]. We assume that
542 the proportion of cell-types will vary across pixels and thus differences in their cell-type-specific
543 transcriptional profiles manifest as overdispersed genes across pixels in the dataset.
544 If there are too many genes included in the input ST data, LDA may also struggle to
545 identify non-overlapping clusters composed of distinct combinations of co-expressed genes. In
546 these circumstances, users may modulate the number of informative genes included in the input
547 matrix to ensure LDA convergence. By default, only the top 1000 most overdispersed genes are
548 retained in the input ST data.
549 Additional gene filtering or cell-type specific marker genes to include in the input ST
550 data may also be augmented by the user.
551
552 *Selection of LDA model with optimal number of cell-types*
553 The number of cell-types K in the LDA model must be chosen *a-priori*. To determine the
554 optimal number of cell-types K to choose for a given dataset, we fit a set of LDA models using
555 different values for K over a user defined range of positive integers greater than 1. We then
556 compute the perplexity of each fitted model:
557

$$Perplexity(D) = \exp\left\{-\frac{\log\big(p(D)\big)}{\sum_{d=1}^{D}\sum_{n=1}^{N} c_{d,n}}\right\}$$

558 Where *p(D)* is the likelihood of the dataset and $c_{d,n}$ is the gene count, or expression level, of
559 gene *n* in pixel *d*. We can interpret *p(D)* as the posterior likelihood of the dataset conditional on
560 the cell-type assignments using the final estimated θ and β. The lower the perplexity, the better
561 the model represents the real dataset. Thus, the trend between choice of *K* and the respective
562 model perplexity can then serve as a guide. By default, the perplexity is computed by comparing
563 *p(D)* to the entire input dataset used to estimate θ and β.

564 In addition, STdeconvolve also reports the trend between *K* and the number of
565 deconvolved cell-types with mean pixel proportions < 5% (as default). We chose this default
566 threshold based on the difficulty of STdeconvolve and reference-based deconvolution
567 approaches to deconvolve cell-types at low proportions, (i.e., "rare" cell-types) (Supplementary
568 Note 2). We note that as *K* is increased for fitted LDA models, the number of such "rare" cell-
569 types generally increases. Such rare deconvolved cell-types are often distinguished by fewer
570 distinct transcriptional patterns in the data and may represent non-relevant or spurious
571 subdivisions of primary cell-types. We can use this metric to help set an upper bound on *K*.

572 Generally, perplexity decreases and the number of "rare" deconvolved cell-types
573 increases as *K* increases. Given these model perplexities and number of "rare" deconvolved cell-
574 types for each tested *K,* the optimal *K* can then be determined by choosing the maximum *K* with
575 the lowest perplexity while minimizing number of "rare" deconvolved cell-types. To further
576 guide the choice of *K*, an inflection point ("knee") is derived from the maximum second
577 derivative of the plotted *K* versus perplexity plot and *K* versus number of "rare" deconvolved
578 cell-types.

579 Still, for a given K, the fitted LDA model may fail to identify distinct cell-types e.g., the
580 distribution of cell-type proportions in each pixel is uniform. In such a situation, the Dirichlet

distribution shape parameter α of the LDA model will be >= 1 and STdeconvolve will indicate to the user that the fitted LDA model for a particular K has an α above this threshold by greying out these *Ks* in the trend plot.

Ultimately, the choice of *K* is left up to the user and can be chosen taking into consideration prior knowledge of the biological system.

Simulating ST data from single-cell resolution spatially resolved MERFISH data

MERFISH data of the mouse medial preoptic area (MPOA) was obtained from the original publication[16]. Normalized gene expression values were converted back to counts by dividing by 1000 and multiplying by each cell's absolute volume. Datasets for an untreated female animal (FN7, datasets 171021_FN7_2_M22_M26 and 171023_FN7_1_M22_M26) containing counts for 135 genes assayed by MERFISH were used. Genes with non-count expression intensities assayed by sequential FISH were omitted. Counts of blank control measurements were also removed. Cells were previously annotated as being one of 9 major cell-types (astrocyte, endothelial, microglia, immature or mature oligodendrocyte, ependymal, pericyte, inhibitory neuron, excitatory neuron). Cells originally annotated as "ambiguous" were removed from the dataset to ensure the ground truth was composed of cells with distinguishable cell-types. Because certain cell-types may be enriched in specific regions of the MPOA, we combined 12 tissue sections across the anterior and posterior regions to ensure that all expected cell-types would be well represented in the final simulated ST dataset. After filtering, the final dataset contained 59651 cells representing 9 total cell-types and counts for the 135 genes.

To simulate a multi-cellular pixel resolution ST dataset from such single-cell resolution spatially resolved MERFISH data, we generated a grid of squares, each square with an area of

604 100 μm^2. Each square was considered a simulated pixel and the gene counts of cells whose x-y
605 centroid was located within the coordinates of a square pixel were summed together. A grid of
606 square pixels was generated for each of the 12 tissue sections separately and the simulated pixels
607 for all 12 tissue sections were subsequently combined into a single ST dataset. For a given tissue
608 section, the bottom edge of the grid was the lowest y-coordinate of the cell centroids and the left
609 edge of the grid was the lowest x-coordinate. Square boundaries were then drawn from each of
610 these edges in 100 μm^2 increments until the position of the farthest increment from the origin
611 was greater than the highest respective cell centroid coordinate. After generating the grid, square
612 pixels whose edges formed one of the outside edges of the grid were discarded in order to
613 remove simulated pixels, which by virtue of their placement, encompassed space outside of the
614 actual tissue sample. The retained pixels covered 49142 out of the original 59651 cells in the 12
615 tissue sections. This resulted in a simulated ST dataset with 3072 pixels by 135 genes. We used
616 the original cell-type labels of each cell to compute the ground truth proportions in each
617 simulated pixel. Likewise, to generate the ground truth transcriptional profiles of each cell-type,
618 we averaged the gene counts for cells of the same cell-type from the original 59651 cells and
619 normalized the resulting gene count matrix to sum to 1 for each cell-type. To simulate pixels of
620 50, 20, and 10 μm^2, an identical approach was taken using the same cells except those square
621 boundaries were drawn from each edge in 50, 20, or 10 μm^2 increments.
622
623 **Annotation and matching of deconvolved and ground truth cell-types**
624 Each deconvolved cell-type was first matched with the ground truth cell-type that had the highest
625 Pearson's correlation between their transcriptional profiles. This was done by computing the

626 Pearson's correlation between every combination of deconvolved and ground truth cell-type
627 transcriptional profiles.
628 The assignment of deconvolved cell-types to ground truth cell-types was confirmed by
629 testing for enrichment of differentially upregulated genes of the ground truth cell-types in the
630 deconvolved cell-type transcriptional profiles. To determine the differentially upregulated genes
631 of the ground truth cell-types, ground truth transcriptional profiles were converted to counts per
632 thousand and low expressed genes, defined as those with average expression values less than 5,
633 were removed. For each ground truth cell-type, the $\log_2$ fold-change of each remaining gene with
634 respect to the average expression across the other ground-truth cell-types was computed.
635 Differentially upregulated genes were those with $\log_2$ fold-change > 1. We performed rank-based
636 gene set enrichment analysis of the ground truth upregulated gene sets in each deconvolved cell-
637 type transcriptional profile using the `liger` R package[51]. A match to a ground truth cell-type was
638 confirmed and assigned if the ground truth gene set had the lowest gene set enrichment adjusted
639 *p*-value that was at least < 0.05, followed by the highest positive edge score[52], and then highest
640 positive enrichment score to break ties.
641
642 **Deconvolution of additional simulated and real ST data**
643 Deconvolution of simulated and real ST data using STdeconvolve in addition to deconvolution
644 of simulated and real ST data using supervised and semi-supervised reference-based
645 deconvolution approaches with various single-cell transcriptomics references is further detailed
646 in Supplementary Methods.
647
648 **Comparison of deconvolution approaches**

649 How each supervised and semi-supervised deconvolution approach was run is further detailed in
650 the Supplementary Methods. To compare the performance between deconvolution methods, the
651 root mean squared error (RMSE) was computed for each pixel between the deconvolved and
652 matched ground truth cell-type proportions for each pixel in the ST dataset:

653 $$RMSE = \sqrt{\frac{\sum_{k=1}^{K}(\hat{y}_k - y_k)^2}{K}}$$

654 where K is the number of cell-types, $\hat{y}_k$ is the predicted cell-type proportion for the cell-type k,
655 and y_k is the ground truth cell-type proportion for the cell-type k. To assess whether the
656 distribution of pixel RMSEs was significantly lower for STdeconvolve compared to other
657 methods, a one-sided Diebold-Mariano Test[53] was used.
658

659 **Runtime and memory evaluation**

660 Using the Visium dataset described in Supplementary Methods `*Deconvolution of 10X Visium*
661 *data with STdeconvolve*`, we generated an input ST dataset of 2702 pixels and feature selected
662 for the top 1000 most significant overdispersed genes. Runtime of STdeconvolve was measured
663 on randomly drawn subsets of input data. Five subsets are drawn with 2702 pixels and 50, 100,
664 200, 400, and 1000 genes, respectively. Another five subsets are drawn with the 1000 top
665 overdispersed genes and 50, 100, 200, 400, and 1000 pixels, respectively. All subsets are
666 deconvolved with cell type number (K) between 4 and 20 as input parameters. Runtime was
667 measured using the R package `microbenchmark`[54] (v1.4-7). Memory usage of STdeconvolve
668 was measured using a similar sub-setting procedure and the R package `profmem`[55] (v0.6.0).
669 Total memory allocation was measured, which provides an upper bound for peak memory usage.

Runtime and memory analyses were run on a machine with i7-6600U 2.60GHz CPU with 8GM of RAM.

Availability of Code

STdeconvolve is available as an open-source R software package[56] with the source code available in the Supplemental Material and on GitHub at https://github.com/JEFworks-Lab/STdeconvolve. Additional documentation and tutorials are available at https://jef.works/STdeconvolve/

Acknowledgements

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35-GM142889.

683 **References**

684

685 1 Zhuang, X. Spatially resolved single-cell genomics and transcriptomics by imaging. *Nat*
686 *Methods* **18**, 18-22, doi:10.1038/s41592-020-01037-8 (2021).

687 2 Larsson, L., Frisen, J. & Lundeberg, J. Spatially resolved transcriptomics adds a new
688 dimension to genomics. *Nat Methods* **18**, 15-18, doi:10.1038/s41592-020-01038-7
689 (2021).

690 3 Elosua-Bayes, M., Nieto, P., Mereu, E., Gut, I. & Heyn, H. SPOTlight: seeded NMF
691 regression to deconvolute spatial transcriptomics spots with single-cell transcriptomes.
692 *Nucleic Acids Res*, doi:10.1093/nar/gkab043 (2021).

693 4 Cable, D. M. *et al.* Robust decomposition of cell type mixtures in spatial transcriptomics.
694 *Nat Biotechnol*, doi:10.1038/s41587-021-00830-w (2021).

695 5 Dong, R. & Yuan, G. C. SpatialDWLS: accurate deconvolution of spatial transcriptomic
696 data. *Genome Biol* **22**, 145, doi:10.1186/s13059-021-02362-7 (2021).

697 6 Song, Q. & Su, J. DSTG: deconvoluting spatial transcriptomics data through graph-based
698 artificial intelligence. *Brief Bioinform* **22**, doi:10.1093/bib/bbaa414 (2021).

699 7 Kiemen, A. *et al.* In situ characterization of the 3D microanatomy of the pancreas and
700 pancreatic cancer at single cell resolution. *bioRxiv*, 2020.2012.2008.416909,
701 doi:10.1101/2020.12.08.416909 (2020).

702 8 Nguyen, Q. H., Pervolarakis, N., Nee, K. & Kessenbrock, K. Experimental
703 Considerations for Single-Cell RNA Sequencing Approaches. *Front Cell Dev Biol* **6**, 108,
704 doi:10.3389/fcell.2018.00108 (2018).

705 9 Network, B. I. C. C. A multimodal cell census and atlas of the mammalian primary motor
706 cortex. *Nature* **598**, 86-102, doi:10.1038/s41586-021-03950-0 (2021).
707 10 Hu, B. C. The human body at cellular resolution: the NIH Human Biomolecular Atlas
708 Program. *Nature* **574**, 187-192, doi:10.1038/s41586-019-1629-x (2019).
709 11 Regev, A. *et al.* The Human Cell Atlas. *bioRxiv*, 121202, doi:10.1101/121202 (2017).
710 12 Deutsch, A., Feng, D., Pessin, J. E. & Shinoda, K. The Impact of Single-Cell Genomics
711 on Adipose Tissue Research. *Int J Mol Sci* **21**, doi:10.3390/ijms21134773 (2020).
712 13 Haque, A., Engel, J., Teichmann, S. A. & Lonnberg, T. A practical guide to single-cell
713 RNA-sequencing for biomedical research and clinical applications. *Genome Med* **9**, 75,
714 doi:10.1186/s13073-017-0467-4 (2017).
715 14 Dey, K. K., Hsiao, C. J. & Stephens, M. Visualizing the structure of RNA-seq expression
716 data using grade of membership models. *PLoS Genet* **13**, e1006599,
717 doi:10.1371/journal.pgen.1006599 (2017).
718 15 Fan, J. *et al.* Characterizing transcriptional heterogeneity through pathway and gene set
719 overdispersion analysis. *Nat Methods* **13**, 241-244, doi:10.1038/nmeth.3734 (2016).
720 16 Moffitt, J. R. *et al.* Molecular, spatial, and functional single-cell profiling of the
721 hypothalamic preoptic region. *Science* **362**, doi:10.1126/science.aau5324 (2018).
722 17 Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell
723 transcriptomics. *Nat Neurosci* **19**, 335-346, doi:10.1038/nn.4216 (2016).
724 18 Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain
725 transcriptome. *Nature* **489**, 391-399, doi:10.1038/nature11405 (2012).
726 19 Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature*
727 **445**, 168-176, doi:10.1038/nature05453 (2007).

728 20 Li, C. M. *et al.* Aging-Associated Alterations in Mammary Epithelia and Stroma
729 Revealed by Single-Cell RNA Sequencing. *Cell Rep* **33**, 108566,
730 doi:10.1016/j.celrep.2020.108566 (2020).
731 21 Zhao, E. *et al.* Spatial transcriptomics at subspot resolution with BayesSpace. *Nat*
732 *Biotechnol*, doi:10.1038/s41587-021-00935-2 (2021).
733 22 Vizgen Data Release V1.0. May 2021.
734 23 Stahl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by
735 spatial transcriptomics. *Science* **353**, 78-82, doi:10.1126/science.aaf2403 (2016).
736 24 Nagayama, S., Homma, R. & Imamura, F. Neuronal organization of olfactory bulb
737 circuits. *Front Neural Circuits* **8**, 98, doi:10.3389/fncir.2014.00098 (2014).
738 25 Hintiryan, H. *et al.* Comprehensive connectivity of the mouse main olfactory bulb:
739 analysis and online digital atlas. *Front Neuroanat* **6**, 30, doi:10.3389/fnana.2012.00030
740 (2012).
741 26 Wang, C. *et al.* Identification and characterization of neuroblasts in the subventricular
742 zone and rostral migratory stream of the adult human brain. *Cell Res* **21**, 1534-1550,
743 doi:10.1038/cr.2011.83 (2011).
744 27 Tepe, B. *et al.* Single-Cell RNA-Seq of Mouse Olfactory Bulb Reveals Cellular
745 Heterogeneity and Activity-Dependent Molecular Census of Adult-Born Neurons. *Cell*
746 *Rep* **25**, 2689-2703 e2683, doi:10.1016/j.celrep.2018.11.034 (2018).
747 28 Rodriques, S. G. *et al.* Slide-seq: A scalable technology for measuring genome-wide
748 expression at high spatial resolution. *Science* **363**, 1463-1467,
749 doi:10.1126/science.aaw1219 (2019).

750 29 Liu, Y. *et al.* High-Spatial-Resolution Multi-Omics Sequencing via Deterministic
751 Barcoding in Tissue. *Cell* **183**, 1665-1681 e1618, doi:10.1016/j.cell.2020.10.026 (2020).
752 30 https://www.10xgenomics.com/resources/datasets.
753 31 Saunders, A. *et al.* Molecular Diversity and Specializations among the Cells of the Adult
754 Mouse Brain. *Cell* **174**, 1015-1030 e1016, doi:10.1016/j.cell.2018.07.028 (2018).
755 32 Yoosuf, N., Navarro, J. F., Salmen, F., Stahl, P. L. & Daub, C. O. Identification and
756 transfer of spatial transcriptomics signatures for cancer diagnosis. *Breast Cancer Res* **22**,
757 6, doi:10.1186/s13058-019-1242-9 (2020).
758 33 Karaayvaz, M. *et al.* Unravelling subclonal heterogeneity and aggressive disease states in
759 TNBC through single-cell RNA-seq. *Nat Commun* **9**, 3588, doi:10.1038/s41467-018-
760 06052-0 (2018).
761 34 Wei, R., Liu, S., Zhang, S., Min, L. & Zhu, S. Cellular and Extracellular Components in
762 Tumor Microenvironment and Their Application in Early Diagnosis of Cancers. *Anal*
763 *Cell Pathol (Amst)* **2020**, 6283796, doi:10.1155/2020/6283796 (2020).
764 35 Karantza, V. Keratins in health and cancer: more than mere epithelial cell markers.
765 *Oncogene* **30**, 127-138, doi:10.1038/onc.2010.456 (2011).
766 36 Chan, H. S. *et al.* Serine protease PRSS23 is upregulated by estrogen receptor alpha and
767 associated with proliferation of breast cancer cells. *PLoS One* **7**, e30397,
768 doi:10.1371/journal.pone.0030397 (2012).
769 37 Kim, M. *et al.* Immune microenvironment in ductal carcinoma in situ: a comparison with
770 invasive carcinoma of the breast. *Breast Cancer Res* **22**, 32, doi:10.1186/s13058-020-
771 01267-w (2020).

772 38 Beguinot, M. *et al.* Analysis of tumour-infiltrating lymphocytes reveals two new
773 biologically different subgroups of breast ductal carcinoma in situ. *BMC Cancer* **18**, 129,
774 doi:10.1186/s12885-018-4013-6 (2018).
775 39 Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative
776 Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373-1387 e1319,
777 doi:10.1016/j.cell.2018.08.039 (2018).
778 40 Gruosso, T. *et al.* Spatially distinct tumor immune microenvironments stratify triple-
779 negative breast cancers. *J Clin Invest* **129**, 1785-1800, doi:10.1172/JCI96313 (2019).
780 41 Jian, T., Zhaoshi, M., Xuanlong, N., Qiaozhu, M. & Ming, Z. 190-198 (PMLR, 2014).
781 42 Asp, M., Bergenstrahle, J. & Lundeberg, J. Spatially Resolved Transcriptomes-Next
782 Generation Tools for Tissue Exploration. *Bioessays* **42**, e1900221,
783 doi:10.1002/bies.201900221 (2020).
784 43 Eng, C. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA
785 seqFISH. *Nature* **568**, 235-239, doi:10.1038/s41586-019-1049-y (2019).
786 44 Xia, C., Fan, J., Emanuel, G., Hao, J. & Zhuang, X. Spatial transcriptome profiling by
787 MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene
788 expression. *Proc Natl Acad Sci U S A* **116**, 19490-19499, doi:10.1073/pnas.1912459116
789 (2019).
790 45 Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional
791 states. *Science* **361**, doi:10.1126/science.aat5691 (2018).
792 46 Wang, F. *et al.* RNAscope: a novel in situ RNA analysis platform for formalin-fixed,
793 paraffin-embedded tissues. *J Mol Diagn* **14**, 22-29, doi:10.1016/j.jmoldx.2011.08.002
794 (2012).

47 Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat Methods* **15**, 932-935, doi:10.1038/s41592-018-0175-z (2018).

48 Levy-Jurgenson, A., Tekpli, X., Kristensen, V. N. & Yakhini, Z. Spatial transcriptomics inferred from pathology whole-slide images links tumor heterogeneity to survival in breast and lung cancer. *Sci Rep* **10**, 18802, doi:10.1038/s41598-020-75708-z (2020).

49 Blei, D. M. a. N., Andrew Y and Jordan, Michael I. Latent dirichlet allocation. *The Journal of Machine Learning Research* **3**, 993-1022 (2003).

50 Grün, B. & Hornik, K. topicmodels: An R Package for Fitting Topic Models. *2011* **40**, 30, doi:10.18637/jss.v040.i13 (2011).

51 Fan, J. Differential Pathway Analysis. *Methods Mol Biol* **1935**, 97-114, doi:10.1007/978-1-4939-9057-3_7 (2019).

52 Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550, doi:10.1073/pnas.0506580102 (2005).

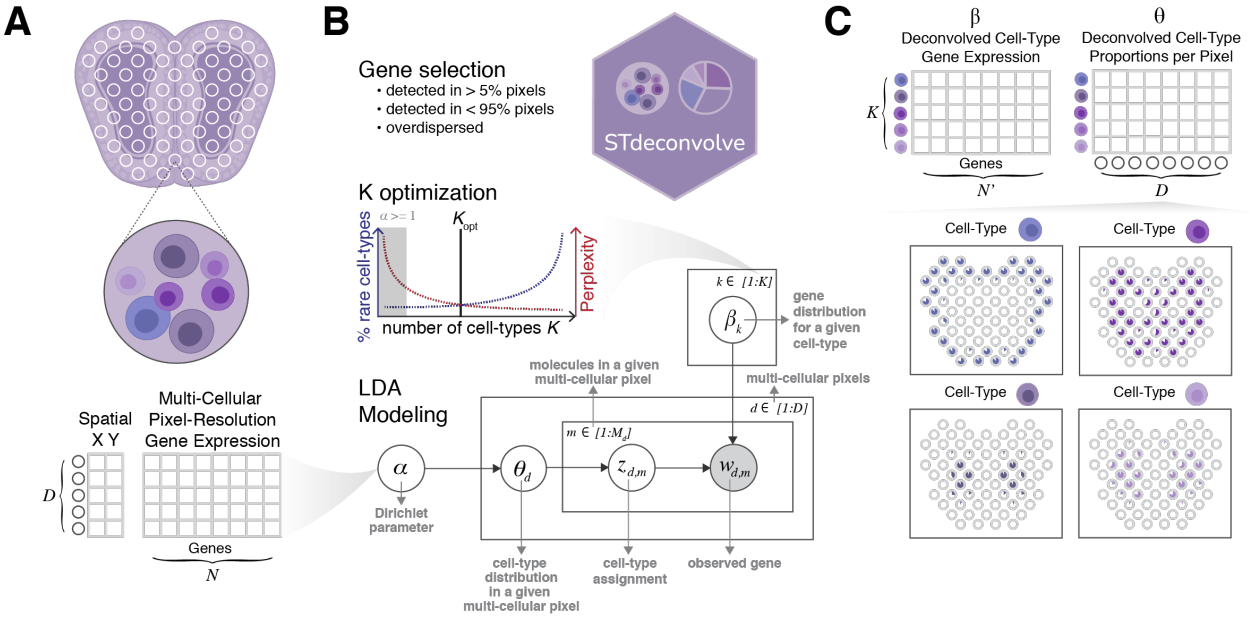
53 Diebold, F. X. & Mariano, R. S. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* **13**, 253-263, doi:10.1080/07350015.1995.10524599 (1995).

54 Accurate Timing Functions. https://github.com/joshuaulrich/microbenchmark/.

55 Simple Memory Profiling for R. https://github.com/HenrikBengtsson/profmem.

56 R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing* (2021).

816 **Figure Legends**



817

818 **Figure 1. Overview of STdeconvolve.** A) STdeconvolve takes as input a spatial transcriptomics
819 (ST) gene counts matrix of *D* pixels (rows) by *N* genes (columns). A matrix of spatial
820 coordinates for each of the *D* pixel can also be used for visualization. B) STdeconvolve first
821 feature selects genes for deconvolution, such as genes with counts in more than 5% and less than
822 95% of the pixels, and overdispersed across the pixels. STdeconvolve then guides the selection
823 of the optimal number of cell-types to be deconvolved, *K*. STdeconvolve finally applies LDA
824 modeling. A graph representation of LDA modeling and the parameters to be learned is shown.
825 Shaded circle indicates observed variables and clear circles indicate latent variables. C)
826 STdeconvolve outputs two matrices: (1) β, the deconvolved transcriptional profile matrix of *K*
827 cell-types over *N'* feature selected genes, and (2) θ, the proportions of *K* cell-types across the *D*
828 pixels. The proportion of deconvolved cell-types can then be visualized across the pixels.

829

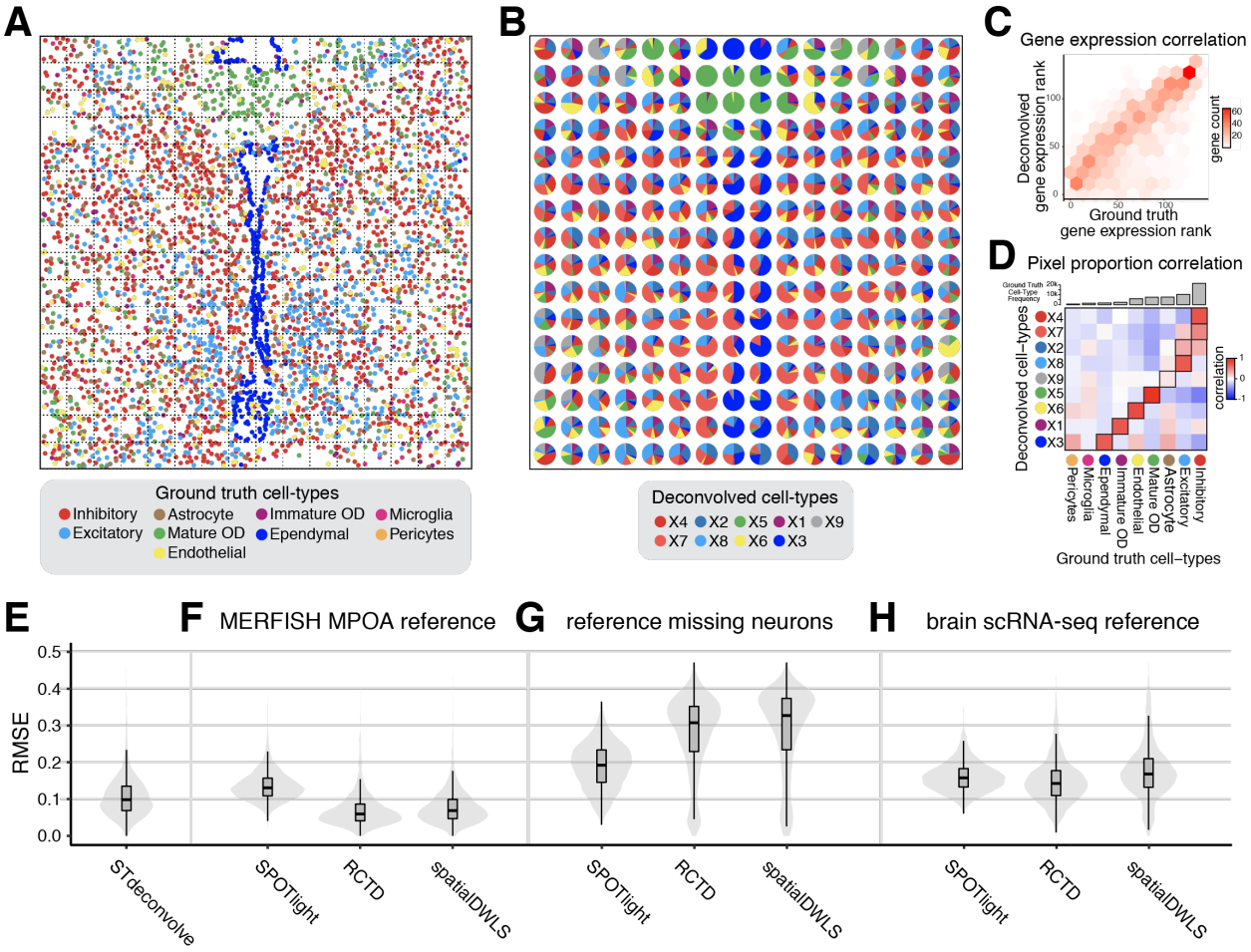


Figure 2. Deconvolution of simulated ST data. A) Ground truth single-cell resolution MERFISH data of one section of the MPOA partitioned into 100 μm^2 pixels (black dashed squares). Each dot is a single cell colored by its ground truth cell-type label. B) Proportions of deconvolved cell-types from STdeconvolve represented as pie charts for each simulated pixel. C) The ranking of each gene based on its expression level in the deconvolved cell-type transcriptional profiles compared to its gene rank in the matched ground truth cell-type transcriptional profiles. C) Heatmap of Pearson's correlations between the deconvolved cell-type proportions and ground truth cell-types proportions across simulated pixels. D) Ground truth cell-types are ordered by their frequencies in the ground truth dataset. Matched deconvolved and ground truth cell-types are boxed. E) Root-mean-square-error (RMSE) of the deconvolved cell-

842 type proportions compared to ground truth for STdeconvolve, F) for supervised deconvolution
843 approaches using the ideal single cell transcriptomics MERFISH MPOA reference, G) for
844 supervised deconvolution approaches using the single cell transcriptomics MERFISH MPOA
845 reference with missing neurons, and H) for supervised deconvolution approaches using a brain
846 single-cell RNA-seq reference.
847

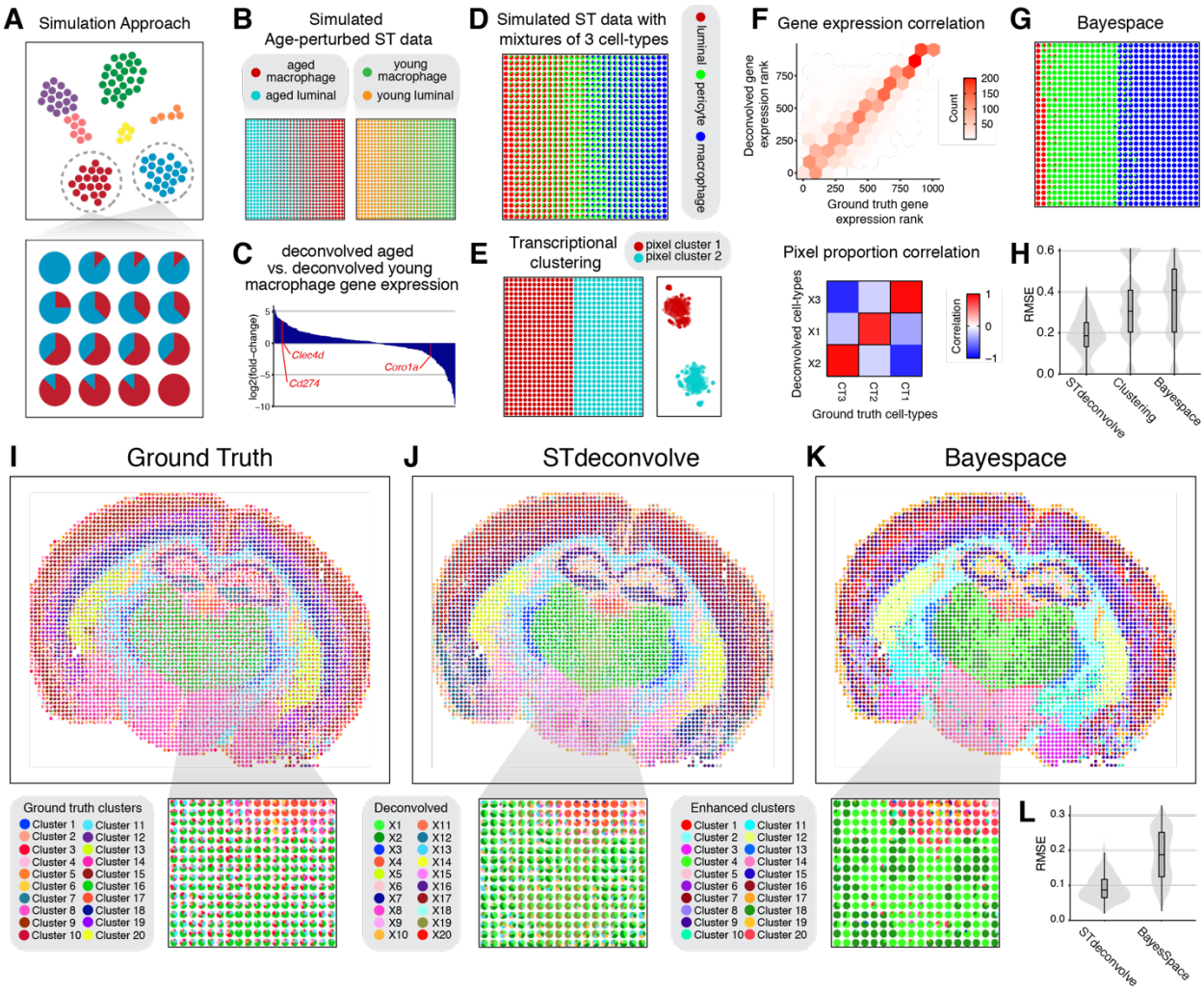


Figure 3. Comparing clustering versus deconvolution analysis for ST data. A) Overview of simulation approach. Starting from a single-cell RNA-seq clustering result visualized as a 2D tSNE embedding with cells colored by cell-type (top), gene expression counts from cells are combined to simulate cell-type mixtures, with the proportions of cell-types are represented as pie charts for each arbitrary spatial pixel (bottom). B) Simulated ST datasets of aged and young tissues using mixtures of aged macrophages with aged luminal cells and young macrophages with young luminal cells respectively represented as pie charts for each simulated ST pixel. C) Bar chart of $\log_2$ fold-change for deconvolved aged macrophage versus deconvolved young macrophage gene expression. Select genes are highlighted in red. D) Simulated ST dataset with 3

858 cell-types represented as pie charts for each simulated ST pixel. E) Clustering analysis results of
859 simulated ST dataset with 3 cell-types. Pie chart proportional representation (left) and tSNE
860 representation (right). F) Deconvolution results for the simulated ST dataset with 3 cell-types by
861 STdeconvolve. The ranking of each gene based on its expression level in the deconvolved-cell-
862 type transcriptional profiles compared to its gene rank in the matched ground truth cell-type
863 transcriptional profiles (top). Heatmap of Pearson's correlations between the deconvolved cell-
864 types proportions and ground truth cell-types proportions across simulated pixels (bottom). G)
865 BayesSpace enhanced resolution clustering results for the simulated ST dataset with 3 cell-types
866 represented as pie charts. H) Root-mean-square-error (RMSE) of the deconvolved cell-type
867 proportions compared to ground truth for the simulated ST dataset with 3 cell-types. I) Ground
868 truth cell-type proportions derived from single-cell resolution MERFISH data of the mouse brain
869 partitioned into 100 μm^2 pixels. J) Deconvolved cell-type proportions for the mouse brain by
870 STdeconvolve. K) Enhanced resolution clustering for the mouse brain by BayesSpace. Inset
871 highlights an interior region corresponding approximately to the thalamus. L) Root-mean-square-
872 error (RMSE) of the deconvolved cell-type proportions compared to single-cell clustering for the
873 MERFISH mouse brain data for the inset interior region corresponding approximately to the
874 thalamus.
875

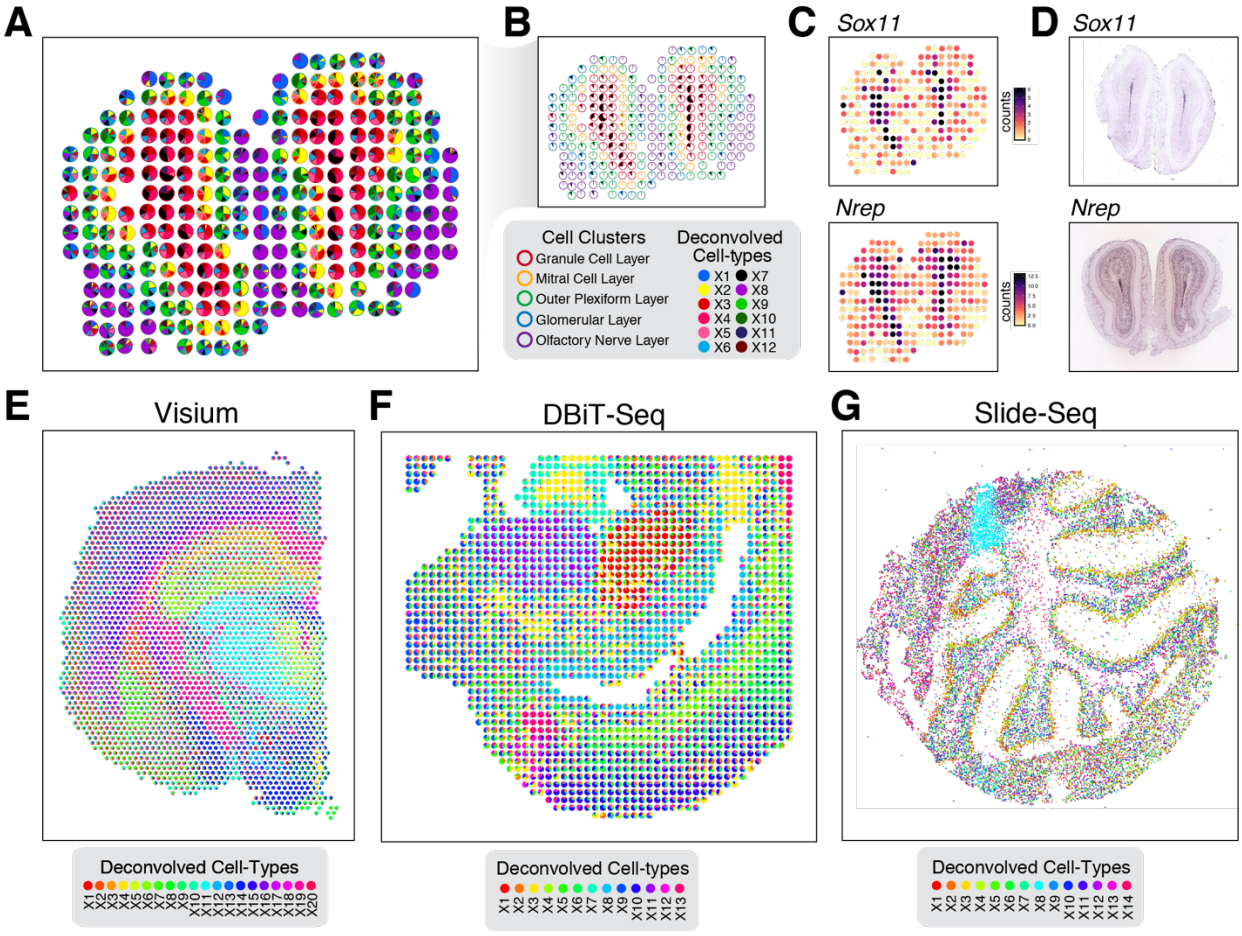


Figure 4. Deconvolution of ST data of varying resolution from multiple technologies by STdeconvolve. A) Deconvolved cell-type proportions for ST data of the MOB, represented as pie charts for each ST pixel. Pixels are outlined with colors based on the pixel transcriptional cluster assignment corresponding to MOB coarse cell layers. B) Highlight of deconvolved cell-type X7. Pixel proportion of deconvolved cell-type X7 are indicated as black slices in pie charts. Pixels are outlined with colors as in A). C) Gene counts in each pixel of the MOB ST dataset for deconvolved cell-type X7's select top marker genes *Sox11* and *Nrep*. D) Corresponding ISH images for *Sox11* and *Nrep* from the Allen Brain Atlas[19]. E) Deconvolved cell-type proportions for Visium data of the mouse brain. F) Deconvolved cell-type proportions for DBiT-seq data of

886 the lower body of an E11 mouse embryo. G) Deconvolved cell-type proportions for Slide-seq

887 data of the mouse cerebellum.

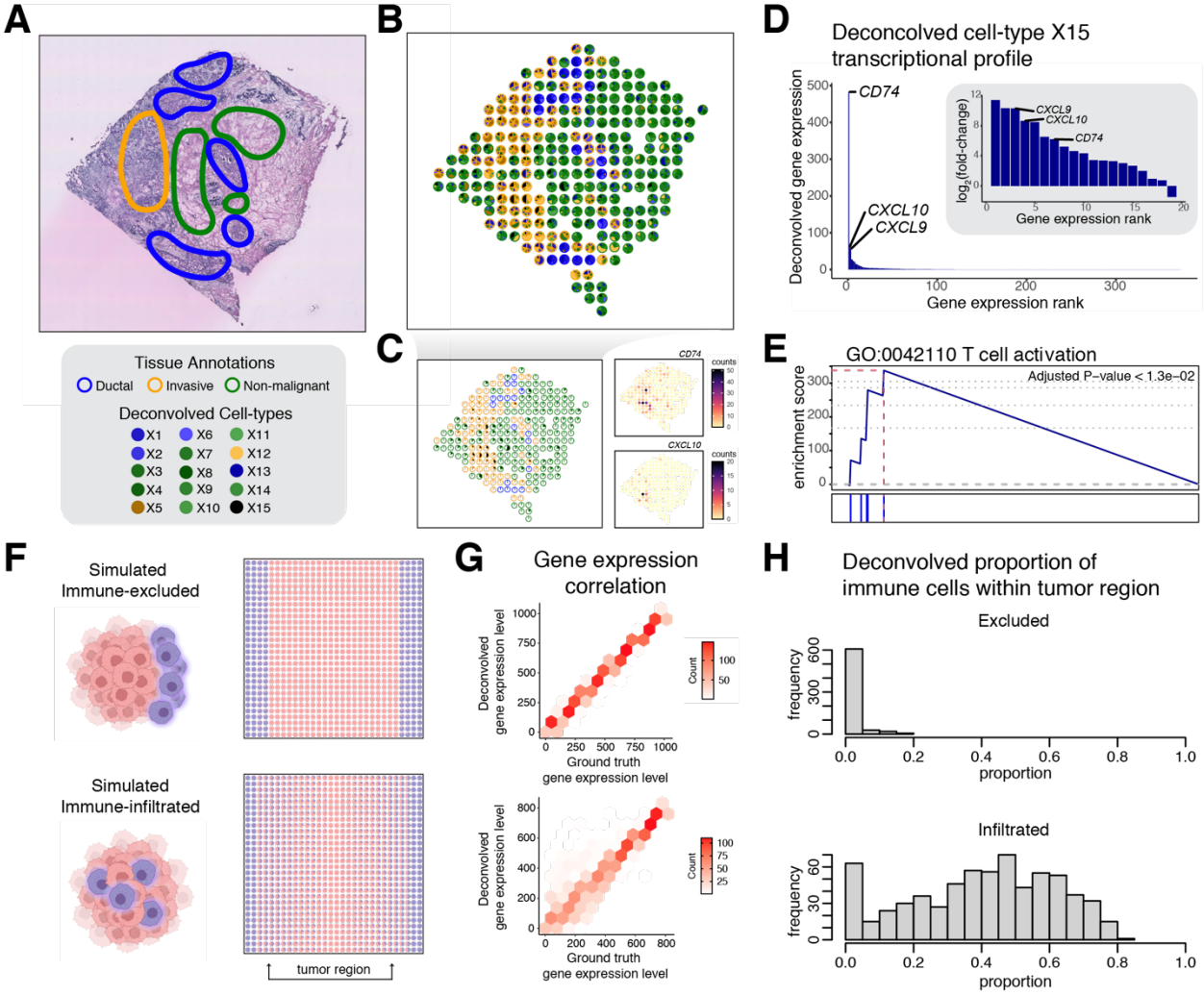


Figure 5. STdeconvolve characterizes the spatial organization of immune cells in real and simulated breast cancer ST data. A) An H&E-stained image of the breast cancer tissue with pathological annotations adapted from Yoosuf *et al.*[32]. B) Deconvolved cell-type pixel proportions for ST data of a breast cancer tissue section, represented as pie charts. Pixels are outlined with colors based on the pixel transcriptional cluster assignment corresponding to 3 pathological annotations. C) Highlight of deconvolved cell-type X15. Pixel proportion of deconvolved cell-type X15 are indicated as black slices in pie charts. Pixels are outlined with colors as in B). Select genes corresponding cell-type X15's select top marker genes are shown.

898 D) Barplot of the deconvolved transcriptional profile of cell-type X15 ordered by magnitude.
899 Inset represents the $\log_2$ fold-change of the deconvolved transcriptional profile of cell-type X15
900 with respect to the mean expression of the other 14 deconvolved cell-type transcriptional
901 profiles. Select highly expressed and high fold-change genes are labeled. E) Gene set enrichment
902 plot for significantly enriched GO term "T cell activation" for deconvolved cell-type X15. F)
903 Simulated ST datasets of an immune-excluded tumor sample (top) and immune-infiltrated tumor
904 sample (bottom) using mixtures of single cells represented as pie charts for each simulated ST
905 pixel. G) Deconvolution results for the simulated ST data by STdeconvolve. The ranking of each
906 gene based on its expression level in the deconvolved-cell-type transcriptional profiles compared
907 to its gene rank in the matched ground truth cell-type transcriptional profiles for the simulated
908 immune-excluded tumor sample (top) and immune-infiltrated tumor sample (bottom). H)
909 Histogram of the deconvolved proportion of immune cells in the tumor region defined in (F) for
910 the simulated immune-excluded tumor sample (top) and immune-infiltrated tumor sample
911 (bottom).
912