Inference of single cell profiles from histology stains with the Single Cell omics from Histology Analysis Framework (SCHAF)

Charles Comiter[1,2,3,#], Xingjian Chen[2,3], Eeshit Dhaval Vaishnav[4], Koseki J. Kobayashi-Kirschvink[1,12,23], Metamia Ciampricotti[6], Ke Zhang[3], Jason Murray[18], Francesco Monticolo[3], Jianhuan Qi[3], Ryota Tanaka[3,19], Sonia E. Brodowska[3], Bo Li[1,13,14,24], Yiming Yang[1,13,24], Scott J. Rodig[7], Angeliki Karatza[22], Alvaro Quintanal Villalonga[6], Madison Turner[8,16], Kathleen L. Pfaff[8], Judit Jané-Valbuena[1], Michal Slyper[1], Julia Waldman[1], Sebastian Vigneau[1], Jingyi Wu[1], Timothy R. Blosser[9], Åsa Segerstolpe[1], Daniel L. Abravanel[14,15], Nikhil Wagle[1,15], Shadmehr Demehri[3,19], Xiaowei Zhuang[9,10], Charles M. Rudin[6], Johanna Klughammer[1,11], Orit Rozenblatt-Rosen[1,13], Collin M. Stultz[2,17,20,21,#], Jian Shu[1,3,20,#], Aviv Regev[1,5,10,24,#]

[1]Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

[2]Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA 02142, USA

[3]Cutaneous Biology Research Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02129, USA

[4]Sequome Inc., 329 Oyster Point Blvd, Suite 300, South San Francisco, CA 94080, USA

[5]Department of Biology, MIT, Cambridge, MA 02140, USA

[6]Druckenmiller Center for Lung Cancer Research, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

[7]Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA

[8]Center for Immuno-Oncology, Dana-Farber Cancer Institute, Boston, MA 02115, USA

[9]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

[10]Howard Hughes Medical Institute

[11]Gene Center and Department of Biochemistry, Ludwig Maximilians Universität München, Munich, Germany

[12]Laser Biomedical Research Center, G. R. Harrison Spectroscopy Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[13]Center for Immunology and Inflammatory Diseases, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA.

[14]Department of Medicine, Harvard Medical School, Boston, MA, USA

[15]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA, USA

[16]Department of Microbiology, Immunology, and Cancer Biology, University of Virginia, Charlottesville, VA 22908, USA

[17]Divison of Cardiology, Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02129, USA

[18]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

[19]Center for Cancer Immunology, Krantz Family Center for Cancer Research, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

[20]Harvard-MIT Program in Health Sciences and Technology, Cambridge, MA 02142, USA

[21]Institute of Medical Engineering and Science, MIT, Cambridge, MA 02142, USA

[22]Perlmutter Cancer Center, New York University Langone Health, New York, NY 10016, USA

[23]Department of Medicine, University of Chicago, Chicago, IL 60637

[24]Current address: Genentech, 1 DNA Way, South San Francisco, CA 94080, USA

#To whom correspondence should be addressed: ccomiter@mit.edu (CC), cmstultz@mit.edu (CMS), jian.shu@mgh.harvard..edu (JS), aviv.regev.sc@gmail.com (AR)

Tissue biology involves an intricate balance between cell-intrinsic processes and interactions between cells organized in specific spatial patterns, which can be respectively captured by single cell profiling methods, such as single cell RNA-seq (scRNA-seq) and spatial transcriptomics, and histology imaging data, such as Hematoxylin-and-Eosin (H&E) stains. While single cell profiles provide rich molecular information, they can be challenging to collect routinely in the clinic and either lack spatial resolution or high gene throughput. Conversely, histological H&E assays have been a cornerstone of tissue pathology for decades, but do not directly report on molecular details, although the observed structure they capture arises from molecules and cells. Here, we leverage vision transformers and adversarial deep learning to develop the Single Cell omics from Histology Analysis Framework (SCHAF), which generates a tissue sample's spatially-resolved whole transcriptome single cell omics dataset from its H&E histology image. We demonstrate SCHAF on a variety of tissues—including lung cancer, metastatic breast cancer, placentae, and whole mouse pups—training with matched samples analyzed by sc/snRNA-seq, H&E staining, and, when available, spatial transcriptomics. SCHAF generated appropriate single cell profiles from histology images in test data, related them spatially, and compared well to ground-truth scRNA-Seq, expert pathologist annotations, or direct spatial transcriptomic measurements, with some limitations. SCHAF opens the way to next-generation H&E analyses and an integrated understanding of cell and tissue biology in health and disease.

INTRODUCTION

Advances in massively parallel, high-resolution molecular profiling now provide cellular and tissue level measurements at a genomic scale[1]. These include methods for massively parallel single cell or single-nucleus (sc/sn) profiling of RNA, chromatin, proteins, and their multi-modal combinations[2]. Simultaneous advancements in data-driven analytics, largely in artificial intelligence / machine learning (AI/ML)[3], have allowed us to derive biological insights from such rich data[4–8], in addition to key modalities such as H&E stains, which are used in both research and clinical practice[9].

However, substantial challenges remain in realizing the promise of these methods in the context of tissue biology, especially for histopathology, a cornerstone of medicine. While the costs and complexity of single cell omics have reduced dramatically[10], they remain relatively expensive and time consuming and are not yet routinely applied in clinical settings. Experiments also remain prone to technical variations, leading to inter-sample discrepancies and batch effects[11,12]. Moreover, single cell genomics does not directly capture spatial information nor is it directly related to the rich legacy of histology. While spatial transcriptomics methods, such as MERFISH[13], seqFISH[+14], Visium[15], Xenium[16], ISS[17], Barista-seq[18], smFISH[19], osmFISH[20], STARmap[21], Slide-tags[22], or Targeted ExSeq[23] measure spatially-resolved expression data, their throughput is still limited, and they involve high per-sample costs and complexities. Computational methods[24–26] have used limited spatial information measured experimentally to project cell profiles to spatial positions, but they have thus far required some shared variables for such mapping (*i.e.*, transcripts measured both in the dissociated cells and spatially in tissue). In sum, they do not spatially resolve single cell omics data based solely on the microscopic morphology in a histology image.

88 Recent advances in applied deep learning may open the way to address the challenge of mapping
89 single cell profiles to histology. In other domains, deep learning methods have successfully related
90 data modalities of the same entities even when they do not have nominally shared variables (such
91 as audio and video[27–29]). Moreover, in recent studies, models were trained to generate a tissue's
92 bulk RNA-seq or its non-single cell spatial transcriptomics measurements (55 micron diameter),
93 from its histology image[30,31]. Given these successes, we hypothesized that deep learning could also
94 be applied to the more difficult translation problem of inferring single cell expression profiles from
95 histology.

96 Here, we present the Single Cell omics from Histology Analysis Framework (SCHAF), a deep
97 learning-based framework for translation from histology images to single cell omics data. SCHAF
98 is based on the assumption that a histology image and a single cell omics dataset from the same
99 tissue sample can be explained by a single underlying latent distribution. Given a corpus of tissue
100 data, where each sample has a histology image, sc/snRNA-seq data, and, when available, spatial
101 transcriptomics data, SCHAF discovers this common latent space from these modalities across
102 different samples. SCHAF then leverages this latent space to construct an inference engine
103 mapping a histology image to its corresponding (model-generated) single cell molecular profiles.
104 This inferred dataset has high genomic coverage, yielding expression information on all the genes
105 in the training corpus, and is, in some cases, spatially resolved, with each predicted gene expression
106 profile mapped at single cell resolution. Given the spatial information inherent in its predictions,
107 SCHAF can further predict a spatial portrait of a tissue's cell types. We demonstrate SCHAF's
108 performance on these tasks with data from a variety of tissue types, through multiple criteria,
109 including a comparison of SCHAF predictions to experimentally measured spatial transcriptomics
110 data and to expert pathologist annotations of cell types. We further demonstrate that SCHAF

outperforms related models for the task of predicting transcriptomic information from H&E images alone.

RESULTS

SCHAF: Single Cell omics from Histology Analysis Framework

SCHAF infers single cell omics data from a corresponding histological image. The model is trained on a corpus of samples, each with a matched histology image and single cell transcriptomics data. SCHAF comes in two flavors depending on the type of single cell transcriptomics data available for training (**Fig. 1**).

If spatial transcriptomics data from the same tissue section as the histology image are available in addition to corresponding sc/snRNA-seq, we train Paired SCHAF, through the following steps (**Fig. 1, right, Methods)**. First, Paired SCHAF tiles the histology image into partially overlapping smaller square tiles, each centered around an individual cell, capturing it and its surroundings. Paired SCHAF is then trained in two stages. In the first stage, the model is trained to predict the spatial transcriptomics profile corresponding to each tile (**Fig. 1a, supervised spatial pretraining**), where the initial weights for the image encoder are obtained from a previously developed whole-slide foundational digital pathology model[32]. In the second stage, additional transcriptome-wide expression data from genes not measured in the spatial transcriptomic experiments are mapped onto the spatial locations using Tangram[7], a computational method that infers spatial locations for scRNA-seq, and Paired SCHAF is updated to predict these additional spatially resolved transcripts (**Fig. 1a, Paired SCHAF**). This second stage improves the model's ability to infer accurate spatial locations for transcripts that were not captured in the spatial transcriptomics training data. At inference, Paired SCHAF only requires a histology image as input

133 to generate single cell profiles at high (whole transcriptome) gene throughput that are spatially
134 resolved at single cell resolution (as the generated profile is associated with the cell at the center
135 of the tile).

136 As experimentally determined spatial transcriptomic information is not always available for
137 training, we also develop Unpaired SCHAF, as a method for inferring single cell transcript profiles
138 from histology images that does not rely on spatial transcriptomic training data (**Fig. 1, left,**
139 **Methods**). Like Paired SCHAF, it first segments the histology image into partially-overlapping
140 tiles, each centered on a cell (**Methods**; cells are segmented by a pretrained AI model for cell
141 segmentation). Unpaired SCHAF then uses encoders for both histopathology and scRNA-seq to
142 build a model that translates from one H&E tile to one scRNA-seq profile. We take the initial
143 weights for each of these encoders from existing foundation models for both histopathology[32] and
144 scRNA-seq[33] and use a discriminative network to impose that the latent representations from both
145 encoders be similar, thereby enabling a model that translates from one H&E tile to one scRNA-
146 seq profile. Like Paired SCHAF, the resulting predictions generated at inference have single cell
147 resolution and high gene throughput, albeit with more limited spatial information. When
148 histopathologist annotations are available for training data, they, with scRNA-seq cell type
149 annotations, can also be incorporated into Unpaired SCHAF's training to encourage the method to
150 infer cell types with some spatial accuracy guarantees. While Paired SCHAF provides spatial
151 transcriptomic information at single cell resolution, Unpaired SCHAF provides information about
152 the distribution of RNA transcripts in a given tissue in addition to spatially resolved cell type
153 information.

154 The final models (Paired SCHAF and Unpaired SCHAF) are fusions of different encoders and
155 decoders, yielding an end-to-end process to generate a single cell expression profile from a

156 histology image tile (**Fig. 1, bottom, Methods**). At inference, we split the segmented histology
157 image into tiles centered on each cell and generate each cell's expression profile using the
158 corresponding SCHAF model to infer single cell resolution profiles. In principle, when spatially
159 resolved transcript information was used during training (Paired SCHAF), we expect better
160 performance than when it was not available (Unpaired SCHAF).

161 **Paired SCHAF infers spatial gene expression at single cell resolution**

162 We developed and tested Paired SCHAF using three tissue corpora, each with H&E, spatial
163 transcriptomics, and sc/snRNA-seq training data available (**Fig. 2a, top, and 2b, left**). **First**, in
164 the one day old **mouse pup**, we found one single tissue slice assayed via H&E and experimentally
165 measured spatial transcriptomics at single cell resolution (ExpSCR) via Xenium, profiling 365
166 genes over more than 1,000,000 cells[16], and a second tissue from a different whole one day old
167 mouse pup by scRNA-seq[34]. We partitioned the mouse pup data into four spatial folds and
168 evaluated Paired SCHAF via cross validation, training, for each fold, a model using H&E, scRNA-
169 seq, and ExpSCR data from the remaining three folds, evaluating on the fourth fold, and averaging
170 results over all four folds (**Fig. 2a, top left**). **Second**, in estrogen receptor positive metastatic breast
171 cancer (ER+ MBC), we relied on data we recently collected from one large tumor, analyzed with
172 three adjacent sections[16]: the first was assayed by H&E and ExpSCR/Xenium, the second by
173 scRNA-seq, and the third by Experimentally Measured spatial transcriptomics at low-resolution
174 (ExpLR) via Visium (profiling 17,742 genes, at 55 micron spot resolution), used for validation.
175 We again split the data into four folds, trained models using H&E, scRNA-seq, and ExpSCR data,
176 and evaluated via cross validation and with the ExpLR/Visium data (**Fig. 2a, top right**). **Third**, in
177 estrogen receptor negative metastatic breast cancer (ER- MBC), we used for training a whole
178 single section assayed by H&E and ExpSCR/Xenium[16]. To further evaluate Paired SCHAF's

179 ability to generalize to new data samples that were not available during training, we trained it using
180 the ER+ MBC data and evaluated its performance on the ER- MBC sample (**Fig. 2b-left**). For each
181 of these three cases, we randomly partitioned the ExpSCR genes into two equally-sized sets:
182 "training genes" (used in supervised spatial pretraining, **Fig. 1a, right**), and "hold-out genes" that
183 are not used during training. We expect performance for training genes to be higher than for hold-
184 out genes, and that the prediction for ER- MBC samples from a model trained only on ER+ MBC
185 to be the most difficult task.

186 Paired-SCHAF performed well in inferring the spatial expression levels of both training and hold-
187 out genes in four-fold spatial cross validation, based on Pearson's correlation *vs*. ground truth
188 ExpSCR data (**Fig. 2a, bottom**, mean 0.445 (training), 0.377 (hold-out) for mouse pup; 0.402
189 (training), 0.373 (hold-out) for ER+ MBC) with 48.7% and 52.1% of genes with strong (>0.4)
190 spatial correlations[7] for mouse pup and ER+ MBC, respectively. These were reduced, but still
191 positive for predicting ER- MBC from an ER+ MBC model (**Fig. 2b**; mean Pearson's correlation
192 0.276 (hold-out) and over 26% genes with high correlations (>0.4)).

193 Moreover, spatial cell type distributions of Paired SCHAF predictions (using a neural network
194 trained to classify cell type from a single cell profile) preserve spatial patterns accurately (**Fig. 2c-**
195 **e**), based on the proportion of SCHAF inferred profiles with classifier assigned annotations that
196 were the same as those inferred from the ground truth ExpSCR profiles. Mouse Pup predictions
197 had the highest accuracy across folds (**Fig. 2c**, 0.811-0.904), followed by the ER+ MBC (**Fig. 2d**,
198 0.584-0.665) and ER- MBC (**Fig. 2e**, 0.621), which still showed strong qualitative agreement.
199 Notably, Paired SCHAF preserves heterogeneity *within* each cell type, such that the expression
200 variation across predicted Paired SCHAF profiles is similar to that across measured ExpSCR
201 (**Supplementary Fig. 1**).

202 While there are no other extant methods addressing the problem solved by Paired SCHAF, we
203 compared Paired SCHAF's inferences of training genes to inferences of training genes by four
204 supervised methods (SPIRIT[35], ST-Net[31], DeepPT[36], HE2RNA[30]) that learn transcriptomics from
205 H&E. Paired SCHAF outperformed the other methods in all corpora both in gene level spatial
206 correlation and in cell type accuracy (**Supplementary Fig. 2**).

207 **A gene-specific quality score identifies genes that Paired SCHAF will accurately predict**

208 While Paired SCHAF infers spatially resolved single cell expression profiles at a higher accuracy
209 than other approaches, not all predicted gene transcripts are spatially accurate. We hypothesized
210 that genes whose predicted expression varies substantially across predicted cells in different spatial
211 positions may be more accurate predictions than those with little spatial variation, which may
212 indicate the models' inability to capture their signal. Indeed, a gene's standard deviation over
213 single cell SCHAF predictions agreed well with its actual spatial correlation value for both training
214 and hold-out genes (**Fig. 3a**; r=0.69 **(**hold-out genes, mouse pup) to 0.99 (training genes, ER+
215 MBC) including in ER- MBC (0.730 (hold-out) to 0.803 (training)). Based on this observation we
216 defined a gene's standard deviation over single cell SCHAF predictions as its predictive quality
217 score (PQS). Because this predictive quality score is solely a function of the model *output*, it
218 indicates *a priori* if the model's prediction is likely to be accurate.

219 The predictive quality score helps identify reliable spatial expression patterns in Paired SCHAF
220 predictions, as seen when comparing ExpSCR measurements with Paired SCHAF inferences for
221 genes with a variety of quality scores (**Fig. 3b-d**, **Extended Data Fig. 1-2**). For example, in the
222 Mouse Pup, for both the epithelial marker gene *Epcam* (PQS=0.337 to 1.458, **Fig. 3b**) and the liver
223 hepatocyte marker *Hmgcs2* (**Extended Data Fig. 1a**, PQS = 0.15 to 1.409), higher PQS is

224 associated with better agreement in complex visual patterns between ExpSCR and Paired SCHAF,
225 and lower PQS with noisier gene expression across the sample. Neuronatin (*Nnat*) (**Extended**
226 **Data Fig. 1b**), an important gene in brain development, has a broader PQS range (0.206, .238,
227 .763, and 2.661), but even with regions with lower PQS and lower spatial correlation between
228 SCHAF prediction and ExpSCR measurement, Paired-SCHAF captured some highly subtle spatial
229 patterns correctly (**Extended Data Fig. 1b, bottom right**). We observed similar patterns in ER+
230 MBC for the Lumican gene (*LUM*, PQS = 0.696 to 1.002, **Fig. 3c**), *CD3E* (PQS = 0.626 to 0.809,
231 **Extended Data Fig. 2a**), and lipoprotein lipase (*LPL*) (PQS = 0.268, 0.453, 0.477, and 0.697,
232 **Extended Data Fig. 2b**). Similarly to Neuronatin, *LPL* has sparse punctate expression, which is
233 nevertheless qualitatively predicted by Paired SCHAF (**Extended Data Fig. 2b, bottom right**).
234 In ER- MBC, Paired SCHAF performed better for *TACSTD2*, with strong patterned expression
235 (**Fig. 3d**), less well for *CCDC80* with a somewhat weaker but patterned expression (**Extended**
236 **Data Fig. 2c**), and worst for *CD68*, with weak and sparse expression (**Extended Data Fig. 2d**).
237 Notably, the PQS is not simply a reflection of expression level *per se*. Thus, a lowly expressed
238 gene that has a distinct spatial pattern, such as *CD3E* (PQS = 0.626 to 0.809, **Extended Data Fig.**
239 **2a**), can be predicted well and have a high PQS.

240 **Paired SCHAF generates high quality whole transcriptome predictions from signature**
241 **spatial training data**

242 Next, we assessed how well Paired SCHAF predicts expression profiles over the entire
243 transcriptome by comparing its predictions to lower resolution Visium spatial transcriptomics
244 (ExpLR) for ER^+ MBC. Because these low-resolution data measure aggregate transcriptomes in
245 55 micron spots, we 'coarsed-up' the single cell resolution predictions from Paired SCHAF to a
246 comparable map, summing each predicted transcript expression within a spot.

247 There was a strong agreement between a gene's predictive quality score and the spatial correlation
248 between the ExpLR maps and low-resolution Paired SCHAF expression profiles, for genes that
249 were not available for the supervised spatial pretraining phase of Paired SCHAF (**Fig. 4a,**
250 **Extended Data Fig. 3a**; Pearson's r= 0.88 over all 17,742 genes measured by ExpLR/Visium).
251 For example, low-resolution expression profiles for Tetraspanin 1 (*TSPAN1)* have a range of
252 predictive quality scores over the four folds (0.62, 0.694, 0.876, 1.038) which correlate with the
253 qualitative agreement between Paired SCHAF low-resolution maps and Visium, with lower spatial
254 correlation in lower PQS folds driven by loss of lower signal by Paired SCHAF predictions (**Fig.**
255 **4b**). We observed similar patterns for additional genes (e.g., *ATP2C2* and *CENPL*, **Extended Data**
256 **Fig. 3b,c**). Notably, in all these cases, the original (pre-coarsing) single cell resolution predictions
257 from Paired SCHAF add substantial definition compared to the low-resolution Visium maps (**Fig.**
258 **4b,d**, **Extended Data Fig. 3b,c**).

259 Because lowly and sparsely expressed individual genes may be captured more poorly by spatial
260 transcriptomics and predicted less robustly by Paired SCHAF, we reasoned that the spatial
261 expression of co-functional or co-regulated gene programs in transcriptome-wide data would be
262 more robust and better predicted. To this end, we analyzed 797 cancer gene programs[37,38], whose
263 genes are well-covered by Paired SCHAF predictions (>90% gene overlap, **Methods**), scoring the
264 overall expression of each gene program (**Methods**) in each spot in the measured and predicted
265 maps. Notably, the spatial correlation values between measured and predicted values increased
266 (compare **Extended Data Fig. 3a and 4a,** P=0.0, Mann-Whitney U test), with the vast majority
267 showing good positive correlation (772 (96.6%) programs *vs*. 6,858 (38.7%) of the constituent
268 program genes with spatial correlation r>0.4). We further defined the program's predictive quality
269 score as the standard deviation of its program score over the cells in the Paired SCHAF generated

270 data. There was a moderately strong relationship between a program's PQS and the spatial
271 correlation (Pearson's r=0.606, **Fig. 4c**). This was also qualitatively reflected in the spatial patterns
272 of multiple programs, including a DNA Repair program (**Fig. 4d**), a Tumor Epithelial gene
273 program (**Extended Data Fig. 4b**) and a Serine Proteases program (**Extended Data Fig. 4c**).

274 ***In situ* hybridization experiments validate Paired SCHAF predictions**

275 To determine whether Paired SCHAF can accurately predict high resolution expression profiles
276 for genes that are not available for model training, we selected eight genes that were neither
277 training nor hold-out genes and that had a wide range of predictive quality scores across the mouse
278 pup. We then obtained a new tissue slice for *in situ* hybridization (ISH), spanning the entirety of
279 another (new) one-day old mouse pup. We developed a new protocol to perform both H&E and
280 ISH from the same slice (**Methods**) and performed this for the eight genes in the same slice. We
281 trained a Paired SCHAF model on all data from the original mouse pup and used it to predict the
282 expression values of these eight genes from the new pup's H&E slice. Because the H&E and ISH
283 assays were performed on the same slice, we could calculate exact, single cell resolution spatial
284 correlations between Paired SCHAF predictions and ExpSCR (ISH) measurements.

285 The spatial correlation was in strong agreement with the predictive quality score for the eight
286 genes, such that genes with higher predictive quality scores indeed had much higher spatial
287 correlation between Paired SCHAF predictions and experimental measurements (**Fig. 4e**). These
288 were also supported by qualitative assessment of images of Paired SCHAF predicted gene
289 expression and the experimental maps from ISH experiments for genes from different tissues (**Fig.
290 4f, Extended Data Fig. 5**). For example, *Trf*, a liver marker with a relatively high quality score
291 (0.261) and spatial correlation (0.692) also showed a stronger qualitative agreement than the

292 cardiomyocyte and skeletal muscle marker *Csrp3* (0.262 and 0.462, respectively) and the sparsely
293 expressed neural marker *Rtn1* (0.124, 0.015) (**Fig. 4f, Extended Data Fig. 5**). Nevertheless, for
294 *Csrp3*, there was substantial qualitative similarity between ISH measurement and Paired SCHAF
295 predicted patterns, suggesting that Paired SCHAF captured the subtler patterns. *Rtn1*, with a low
296 PQS, had a low spatial correlation. Both ExpCSR and SCHAF show sparse, scattered expression
297 but not in the same specific cells (**Extended Data Fig. 5**).

298 **Unpaired SCHAF infers single cell distributions of gene expression**

299 Unlike Paired SCHAF, which requires spatial transcriptomics for training, Unpaired SCHAF is
300 trained only with H&E images and sc/snRNA-seq data (**Fig. 1a**). Unpaired SCHAF training can
301 optionally incorporate cell type annotations when available for both scRNA-seq data (from
302 expression) and H&E data (via pathologist annotations), to encourage correspondence between
303 input tiles and their corresponding prediction (**Supplementary Fig. 3**). Thus, Unpaired SCHAF
304 uses an H&E image to infer the distribution of RNA transcripts in a given tissue sample.

305 To assess Unpaired SCHAF, we relied on three human tissue datasets, where both sc/snRNA-Seq
306 and H&E histology were available for training. Two datasets were from human tumors. As non-
307 canonical tissues, tumors are a challenging use case for prediction on new specimens (new tumor
308 from a new patient). The **small cell lung cancer (SCLC)** dataset consisted of 24 tumors profiled
309 by snRNA-seq and matching H&E stains from an unpublished study from Memorial Sloan
310 Kettering Cancer Center (MSKCC). The snRNA-seq data were collected from a portion of the
311 same slice used in H&E staining. No cell type annotations were performed on these H&E data.
312 The **metastatic breast cancer (MBC)** dataset consisted of nine patient samples from the Human
313 Tumor Atlas Project Pilot (HTAPP)[39], where each pair of H&E and snRNA-seq/scRNA-seq data

314 came from different biopsies of the one metastasis. We also included previously collected spatial
315 profiling by MERFISH[13] (used for validation), on another adjacent section. In addition, we
316 generated coarse cell type annotations by an expert pathologist for the H&E sections, broadly
317 delineating large regions that contain mostly cells of the same type. The third dataset was from
318 four **human placenta** samples, which we previously profiled by scRNA-seq[40] and added here
319 matching whole-slide H&E stains from an adjacent section with detailed cell type annotations at
320 nearly single cell resolution, newly generated by a clinical pathologist with expertise in placental
321 histopathology. We also had previously collected spatial profiling by STARmap (used for
322 validation)[40], using a portion of another adjacent section. For each tissue sample, we evaluated
323 Unpaired SCHAF via leave-one-out cross validation (LOOCV), where we trained on all but one
324 of the samples in the dataset and evaluated on the sample excluded from training. The reported
325 results are the average across all patient samples. We incorporated cell types into training Unpaired
326 SCHAF for the two datasets (MBC/HTAPP and placenta) with expert histopathologist annotations.

327 Unpaired SCHAF generates adequate single cell profiles from H&E based on its ability to correctly
328 estimate (**1**) gene expression distributions across single cells, (**2**) distribution (composition) of cell
329 types and psuedobulk profiles; (**3**) heterogeneity of expression profiles within each cell type; and
330 (**4**) appropriate correlation structure and clustering. First, we quantified the difference between
331 Unpaired SCHAF predicted expression distributions and the ground truth distributions for a given
332 gene using the Earth Mover's Distance (EMD), a widely used quantitative measure of similarity
333 between two continuous, nonparametric distributions (**Methods**; low values indicate high
334 similarity). The average gene EMD between Unpaired-SCHAF predictions and real sc/snRNA-
335 Seq was very low in each of the three datasets (**Fig. 5a**; EMD=0.09 (MBC), 0.054 (Placenta),
336 0.062 (SCLC), with many genes' EMD nearly at zero, showing that most genes had predicted

337 distributions very similar to their measured scRNA-seq distributions. Consistently, the gene-gene
338 correlation structures were relatively preserved, with some variation between datasets (Placenta
339 performing best) and between samples within a dataset (**Supplementary Fig. 4**). Second, the cell
340 type distributions of Unpaired SCHAF predictions were similar to those from measured
341 sc/snRNA-seq, based on the Jensen-Shannon Divergence (JSD, **Methods**), with cell types
342 assigned by a previously trained neural network classifier to the Unpaired SCHAF generated data
343 (**Extended Data Fig. 6a, b**). The pseudobulk profiles across different cell types were typically
344 well correlated between Unpaired SCHAF inferred data and measured scRNA-seq
345 (**Supplementary Fig. 5**), with some notable exceptions (e.g., HTAPP 4531). Third, the extent of
346 gene expression heterogeneity within each cell type using the same heterogeneity score as in Paired
347 SCHAF (**Methods**) was comparable between generated and real sc/snRNA-seq, with more
348 disparity in the Placenta dataset than MBC or SCLC (**Supplementary Fig. 6**), and the distribution
349 of EMD for genes was typically low at a cell type specific level (**Supplementary Fig. 7, 8**). Fourth,
350 overall, the relationship between generated profiles is preserved in terms of the extent of clustering,
351 both locally (coherence of clusters) and globally (cluster separation) and as reflected visually in
352 low dimensional embeddings (**Fig. 5b, Extended Data Fig. 6c, d, e**). Heatmaps summarizing these
353 metrics highlight varying degrees of consistency in different samples (**Supplementary Fig. 9**).

354 While no other methods explicitly solve the problem addressed by Unpaired SCHAF (to the best
355 of our current knowledge), we compared it to several state-of-the-art methods for generalized
356 domain translation. Because none can exploit cell type annotation in their training, we compared
357 them to Unpaired SCHAF trained without cell type information. Across both gene level Earth
358 Mover's Distance and cell type distribution Jensen-Shannon Divergence criteria, Unpaired
359 SCHAF outperformed competing methods in all corpora (**Supplementary Fig. 10a**). On other

360 criteria (pseudobulk profiles, gene-gene correlation structure, etc.), unpaired SCHAF performed
361 on par with or better than benchmarks (**Supplementary Fig. 10b**).

362 **A gene-specific distribution matching score identifies gene distributions that Unpaired**
363 **SCHAF will accurately predict**

364 We next developed a dedicated per-gene prediction quality score to *a priori* anticipate when the
365 generated expression distribution from Unpaired SCHAF is similar to the measured sc/snRNA-
366 seq. Unlike Paired SCHAF, where a low standard deviation was empirically associated with poor
367 predictions (**Fig. 3a**), in Unpaired SCHAF, we reasoned that a high standard deviation (more than
368 mean expression or CV), typically associated with more lowly expressed (and less reliably
369 measured) genes in scRNA-seq[41], is associated with lower prediction accuracy. Accordingly, we
370 defined a Gene Distribution Matching Score (GDMS) as one minus one-half the standard deviation
371 of that gene's Unpaired SCHAF inferred expression, which ensures, given that the maximum
372 standard deviation over inferred values is 2, that the score is always non-negative.

373 Indeed, in all three datasets, the GDMS is highly negatively correlated with EMD distance between
374 Unpaired SCHAF predictions and measured sc/snRNA-seq (**Fig. 5c**; Pearson's r=-0.885 to -
375 0.936), such that a high GDMS is associated with a successful prediction of expression
376 distributions. This is also supported when examining expression distributions of individual genes
377 across all three datasets and multiple GDMS scores (**Fig. 5d**, **Extended Data Fig. 7**). Genes with
378 complex multimodal distributions but high GDMS (>0.49) were well predicted by SCHAF
379 (EMD=0.02-0.12) (**Fig. 5d**, **Extended Data Fig. 7** top row), whereas those with low GDMS often
380 resulted in poor prediction (**Extended Data Fig. 7**, bottom four rows, e.g., *LRMDA* in SCLC).

381 **Unpaired SCHAF inferences spatially agree with pathologist annotations and marker genes**

382 Finally, we assessed the spatial accuracy of Unpaired SCHAF predictions – trained with no spatial
383 molecular data, focusing on its ability to generate spatially accurate cell type assignments. We
384 generated single cell RNA expression data using Unpaired SCHAF for each cell in the H&E
385 section, assigned a cell type to the generated profiles and compared these to expert pathologist
386 annotations.

387 We found strong quantitative agreement in the placenta dataset and a weaker – but still marked –
388 agreement in the less canonical HTAPP tumor metastases (**Fig. 6**, **Supplementary Fig. 11**). In the
389 placenta, there was also strong qualitative agreement between spatial cell type annotations from
390 Unpaired SCHAF inferences and from the expert pathologist (**Fig. 6a**), especially if pathologist
391 cell type annotations were included in the training (**Supplementary Fig. 12**). In all four samples,
392 inferences of major cell types in the placenta, such as extravillous trophoblasts (EVTs),
393 syncytiotrophoblasts (STBs), villous cytotrophoblast (vCTBs), fibroblasts, and decidua, are
394 spatially accurate (**Fig. 6a**), as also reflected in confusion matrices (**Supplementary Fig. 11b**),
395 albeit with some specific challenges for particular cell types in some samples. In particular, in the
396 W9 and W11 samples, idiosyncratic regions populated by mostly decidua, fibrin, and maternal
397 cells were well captured by Unpaired SCHAF. In W7, W8, and W9 samples, small but dense
398 regions of EVTs were also captured well. In the W8 and W11 samples, the sparse spatial patterning
399 of fetal blood and endothelial cells is correctly present. The two main cell groups across all
400 samples, one consisting of fetal fibroblasts and Hofbauer cells, and the other of STBs and vCTBs,
401 are both inferred well by Unpaired SCHAF (**Fig. 6a**). Ablating cell type information from training
402 lowered this spatial agreement (**Supplementary Fig. 12**). In the tumor dataset, where samples are
403 far less canonical, there was still notable but weaker qualitative agreement with pathologist
404 annotations in three complex samples (**Supplementary Fig. 11a,c**), with similar reduced

405 performance when cell type information is ablated (**Supplementary Fig. 13**). Note that elementary
406 image analyses show only a weak, if any, correspondence between the number of cells per tile and
407 Unpaired SCHAF's inferred spatial cell type maps (**Supplementary Fig. 14-17**), suggesting that
408 Unpaired SCHAF is learning high complexity features, rather than merely cell type dependent
409 density.

410 We further compared SCHAF predictions to cell type annotations determined by ExpSCR
411 (MERFISH for HTAPP, STARmap for placenta) and found relatively compelling qualitative and
412 quantitative results in the placenta and HTAPP datasets (**Supplementary Fig. 18, 19**). Since the
413 H&E and ExpSCR data for each sample came from adjacent, but not the same, slices, we could
414 perform cell-to-cell mapping between the two. To address this, we aligned lower-resolution spots
415 (~5 cells on average), handling our ExpSCR data as ExpLR data and comparing to coarsed-up
416 (low-resolution) SCHAF predictions. In the Placenta, qualitative and quantitative comparison
417 show substantial, but imperfect agreement (including due to the loss of signal due to coarsing up)
418 (**Supplementary Fig. 18c**). In the HTAPP data, the agreement for three complex samples with
419 multiple cell types was weaker, especially struggling with immune cells, although still notable
420 (**Supplementary Fig. 19c**).

421 Unpaired SCHAF also generated correct spatial patterns for more specific cell types derived from
422 ExpSCR (STARmap) (**Fig. 6b**) in addition to four marker genes associated with various annotation
423 categories: *QSOX1* (EVT), *COL3A1* (fibroblasts), *EFEMP1* (vCTB) and *CD163* (Hofbauer), by
424 qualitative comparison to STARmap[21] measurements in adjacent portions (**Fig. 6b**; sections could
425 not be fully matched up at exact single cell granularity). Overall, more than 100 genes were well
426 captured in the placenta as assessed by comparison to ExpLR measurements (low resolution
427 STARmap) (**Supplementary Fig. 20**). In HTAPP, results were weaker and very variable across

428 tumors (**Supplementary Fig. 21a-c**). Notably, Unpaired SCHAF trained without cell types
429 inferred few (in HTAPP) to no (in placenta) genes with a correlation over 0.4 (**Supplementary**
430 **Fig. 21d**). Ablating cell type information from training also lowered spatial agreement with
431 pathologist annotations (**Supplementary Fig. 12, 13**).

432 **DISCUSSION**

433 We describe SCHAF, a computational framework to predict spatially resolved single cell omics
434 profiles from H&E images. SCHAF is the first method to infer high gene throughput, single cell
435 resolution, spatially resolved (in the case of Paired SCHAF) transcriptomic data using only
436 histology images, a relatively low-cost technique in wide clinical practice. To address the range of
437 opportunities for training, we present two versions of SCHAF: Paired SCHAF, which requires
438 limited molecular spatial data in training, but has spatial accuracy guarantees for its predictions,
439 and Unpaired SCHAF, which, does not require any spatial transcriptomic data during training, but
440 at the cost of making predictions with weaker spatial guarantees. We demonstrate SCHAF across
441 both mouse and human data, from both healthy, canonical developmental tissues and from non-
442 canonical human tumors. We assess its predictions in terms of genes, gene programs, and cell types
443 in single cell and spatial patterns with experimental validation by sc/snRNA-seq, ISH, Xenium,
444 STARmap, MERFISH, and Visium.

445 Paired SCHAF leverages a state-of-the-art foundation[32] model based on a vision transformer
446 architecture to match the spatial transcriptomic data. Instead of supervised training using a small

447 number of spatial marker transcripts, Paired SCHAF augments supervised training with
448 information obtained by mapping scRNA-seq data to spatial positions in the histology image. This
449 allowed the model to generate highly quality spatial, transcriptome-wide profiles. While many
450 spatial transcriptomics methods only measure a small portion of tissue, H&E can be found in a
451 much larger tissue view. Accordingly, spatial transcriptomics from a small part of tissue can be
452 used to train Paired SCHAF, which can then be used to extrapolate spatial data for the entire tissue
453 during inference. Indeed, Paired SCHAF was able to accurately infer large parts of a tissue sample
454 not seen in training and generalize to entirely new tissue samples. SCHAF produces data that is
455 higher resolution and more realistic than extant methods trained to infer low-resolution Visium
456 data from H&E. Our prediction quality score helps users identify *a priori* genes and gene programs
457 that are likely well-predicted by Paired SCHAF.

458 Unpaired SCHAF leverages adversarial deep learning and foundation models and does not require
459 spatial transcriptomic data to train. It also performed quite well in generating accurate gene
460 distributions and cell type compositions, producing high gene throughput, single cell resolution
461 expression data from an H&E stain. The Gene Distribution Matching Score allows a user to *a*
462 *priori* determine the genes that are predicted well. By incorporating expert histopathologist
463 annotations into our training, when including expert pathologist cell annotations in the training,
464 Unpaired SCHAF was able to spatially replicate both expert pathologist annotations and marker
465 gene expression measured via STARmap on unseen tissues.

466 SCHAF was successful on corpora of matched H&E and transcriptomics data from a wide variety
467 of tissue types. As diseased tissues are highly variable and idiosyncratic, predicting a new sample's
468 profile, even of the same type, is a challenging task. At present, the paucity of organized multi-
469 modal datasets limits the generalizability of models like SCHAF across new tissue types. Future
470 studies are required to assess the ability to train a single model across a single corpus of multiple
471 types of tissues, tumors, or technologies (e.g., snRNA-Seq, scATAC-Seq, and spatial multiomics).
472 Doing so would require additional organized spatial datasets with matched single cell profiling
473 and histology imaging, which are still surprisingly scarce, but are growing thanks to efforts, such
474 as the Human Tumor Atlas Network[42], HubMAP[43], CZI Virtual Cell initiative[44], and the Human
475 Cell Atlas[4,2]. To encourage this, we have released the "SCHAF-1M" data repository with this
476 study, encapsulating all the data analyzed here and used in training, as a resource to the broader
477 community.

478 The main experimental limitation of Paired SCHAF lies in the requirement of expensive spatial
479 transcriptomic data in training and/or to test the spatial accuracy of its predictions. To guarantee
480 spatial accuracy, spatial transcriptomic data must be incorporated in training Paired SCHAF. Even
481 in Unpaired SCHAF, spatial accuracy of cell type level predictions required expert histopathologist
482 annotations in the training. Future work should aim to minimize the number of genes required in
483 training models like Paired SCHAF.

484 SCHAF's success motivates the use of a modality-agnostic, lower-dimensional latent space in
485 computational techniques for a multitude of applications in data integration and modality transfer.
486 The underlying principles defining SCHAF can be extended to other cases in biology, both with
487 similar data modalities (*e.g.*, cell profiles and microscopy images measured in cultured cells[45],
488 single cell and spatial multi-omics) in different biological modalities (such as temporal tracing[46]),
489 leading to models producing multimodal outputs from one readily-accessible input. More
490 fundamentally, interpreting SCHAF's predictions could have implications for our understanding
491 of tissue biology, including how molecular information leads to tissue structures and vice versa.

492 Finally, SCHAF's ability to produce high quality transcriptomic data *in silico* from histological
493 imaging would significantly increase the accessibility of high-quality molecular profiles. Beyond
494 the savings in time, effort, and expense, SCHAF could be used to unlock molecular information
495 for the vast number of clinical samples archived over decades and allow scientists to benefit from
496 the insights provided by high-resolution methods in low resource settings. Furthermore, clinical
497 biopsies, yielding minimal patient measurements, could leverage SCHAF to glean more useful
498 clinical information in minimal time and sample access. SCHAF, being the first method to generate
499 high gene throughput, single cell resolution, spatially resolved (in the case of Paired SCHAF)
500 transcriptomic data at the relatively low cost of a histology stain, is well poised to do such.

501

502

503 **Figure Legends**

504 **Figure 1. SCHAF learns to predict a tissue's single cell omics data from its H&E image**

505 **a.** Overview of training. Training data (input) consist of matched cell segmented and tiled
506 histology (H&E) stains, scRNA-seq, and, for Paired SCHAF, spatial transcriptomics from the
507 same section as the H&E from multiple samples. Paired SCHAF (right) is trained in three steps:
508 supervised spatial pretraining of a vision transformer (initialized with weights from a pretrained
509 H&E foundational model) to infer a cell's spatial transcriptomic profile from its H&E tile;
510 mapping of scRNA-seq onto spatial transcriptomics; and fine-tuning of the encoder of the pre-
511 trained model and a decoder of several new linear layers to infer a cell's scRNA-seq from its H&E
512 tile. Unpaired SCHAF (left) is trained in two steps. First, a pretrained scRNA-seq transformer
513 encoder, with frozen weights, and a decoder of several linear layers are trained to reconstruct
514 scRNA-seq profiles. Then, a two-part encoder (a pretrained H&E foundational model, with frozen
515 weights, followed by several linear layers) and a decoder, are trained to encode, via training against
516 an adversarial model, to a latent space indistinguishable from the scRNA-seq encoder's outputs,
517 and to reconstruct the embeddings produced by the pretrained H&E foundational model. **b.**
518 Overview of inference. An input H&E slide is first cell segmented and tiled. In Paired SCHAF,
519 the model from the last part of training is used to generate spatially resolved, genome wide, single
520 cell resolution profiles (right). In Unpaired SCHAF, the H&E tile encoder and the scRNA-seq
521 decoder are used to generate high gene throughput, single cell resolution transcriptomic data with
522 some spatial accuracy guarantees (left).

523

524 **Figure 2. Paired SCHAF inferred profiles accurately recover spatial transcriptomics**

525 **a.** Evaluation by cross validation. Top: A whole mouse pup (left) or ER+ MBC tumor (right) slide
526 is partitioned into four equal-sized folds. Paired SCHAF is trained by 4-fold cross validation with
527 three parts used for training (red) and one held out for evaluation (blue). The results are averaged.
528 Bottom: Distribution of spatial correlations (Pearson's correlation coefficient of the expression of
529 a gene measured by spatial transcriptomics and predicted by Paired SCHAF) for all spatially-
530 measured genes in training (yellow) and hold-out (blue). Error bars: standard error of folds
531 assessed by cross validation. **b.** Evaluation on an independent tumor. Top: Slidea from an
532 independent ER+ MBC tumor used in training (red) and an ER- MBC tumor used in evaluation
533 (blue). Bottom: Distribution of spatial correlations (Pearson's r) for all spatially-measured genes
534 in training (yellow) and held out (blue) samples. **c-e.** Paired SCHAF inferred profiles preserve cell
535 types. Spatial cell type annotations (called by a classifier trained on ground truth ExpSCR data,
536 color legend) for measured (ExpSCR) and Paired SCHAF-inferred profiles in the Mouse Pup (**c**),
537 ER+ MBC (**d**), and ER- MBC (**e**) datasets. Accuracy: proportion of cells in each fold assigned to
538 the correct cell type label.

539 **Figure 3. Quality scores anticipate which genes Paired SCHAF infers accurately**

540 **a.** Quality scores agree with spatial correlations. Quality scores (x axis) and spatial correlation
541 between measured and SCHAF-inferred expression levels (y axis) for genes (dots) in training (top)
542 and hold-out (bottom) data in the three datasets. Red line: best fit. **b-d.** Enhanced spatial inference
543 for genes with higher qualitative scores. Expression in each spatially-resolved cell (blue color
544 scale) in real and SCHAF-inferred data for training gene *Epcam* in Mouse Pup corpus (**b**), hold-
545 out gene *LUM* in ER+ MBC (**c**) and hold-out gene *TACSTD2* in ER- MBC (**d**).

546

547 **Figure 4. Paired SCHAF infers the whole transcriptome accurately**

548 **a,c.e.** Quality scores agree with the degree of spatial correlation. Quality scores (x axis) and extent
549 of spatial correlation between measured and Paired SCHAF values (y axis) for the ER+ MBC
550 dataset genes (a), the ER+ MBC gene programs (c), or for each gene measured by ExpSCR (ISH)
551 in a new mouse pup sample (e). Low-resolution spatial correlation is Pearson's *r* across ExpLR
552 spots. Dashed red line: best fit. **b,d,f**. Illustrative examples of SCHAF-inferred spatial
553 measurements. **b,d**. ExpLR-measured (left), low-resolution (middle), and single cell resolution
554 (right) SCHAF inferred spatial levels (color bar) for *TSPAN1* (b) and the DNA Repair gene
555 program (d) for each of four folds. Quality scores and low-resolution spatial correlations are noted.
556 **f.** Measured (ISH, left) and SCHAF-inferred (right) spatial expression levels for *Trf*.

557

558 **Figure 5. Unpaired SCHAF inferences are similar to measured scRNA-seq**

559 **a.** Most genes have similar expression distributions across cells in measured and SCHAF-inferred
560 scRNA-Seq. Distribution of Earth Mover's Distance (EMD) between the measured and SCHAF-
561 inferred expression distributions of genes in the SCLC (MSKCC, left), breast cancer (HTAPP,
562 middle) and placenta (right) datasets, in the cross validation (holding out one sample at a time).
563 Error bars: standard error of samples assessed by cross validation. **b.** Representative UMAPs.
564 UMAP embedding of measured (bottom) or Unpaired SCHAF-inferred (top) scRNA-seq profiles
565 (dots) colored by cell types (legends), for three representative samples from each of the placenta
566 (right) or SCLC (MSKCC, left) datasets. The measured and inferred scRNA-seq are projected in
567 a common space. **c,d.** Agreement of expression distributions. **c.** Distribution Matching Score
568 (DMS, x-axis) and Earth Mover's Distance (EMD, y-axis) between the measured and SCHAF-
569 inferred expression distribution of each gene (dot). Red line: best fit. **d.** KDE plots for measured
570 (blue) and SCHAF inferred (orange) scRNA-seq gene expression for selected genes (labeled on
571 top) from the SCLC (MSKCC, top), breast cancer (HTAPP, middle) and placenta (bottom)
572 datasets. EMD and DMS values are denoted.

573

574 **Figure 6. Unpaired SCHAF inferences spatially match expert pathologist annotations and**
575 **STARmap spatial measurements in the human placenta.**

576 **a.** Agreement with pathologist annotations. Four human placenta samples (rows) colored by
577 annotations from a cell type classifier of Unpaired SCHAF inferences (left) or an expert
578 pathologist (right). **b.** Agreement with STARmap measurements. Fragments from four human
579 placenta samples (as boxed in **a**), measured by ExpSCR (STARmap), colored by cell type
580 annotations and marker gene expression, in both Unpaired-SCHAF-inferred (top) and ExpSCR
581 (STARmap, bottom) measured profiles.

582

583 **Extended Data Figure Legends**

584 **Extended Data Figure 1. Paired SCHAF predictions of individual gene expression profiles in**
585 **a healthy mouse pup**

586 Mouse Pup samples, with cells colored by measured (left) or Paired-SCHAF inferred (right)
587 expression of training gene *Hmgcs2* (**a**) or hold-out gene *Nnat* (**b**) in each of four-folds assessed
588 by cross-validation. Quality scores and spatial correlations are noted.

589

590 **Extended Data Figure 2. Paired SCHAF predictions of individual gene expression profiles in**
591 **human metastatic breast cancer**

592 **a,b.** ER+ MBC predictions. ER+ MBC samples with cells colored by measured (ExpSCR, top) or
593 Paired-SCHAF inferred (bottom) expression of training genes *CD3E* (**a**) and *LPL* (**b**) in each of
594 four-folds assessed by cross-validation. Quality scores and spatial correlations are noted. **c-d.** ER-
595 MBC predictions. ER- MBC samples with cells colored by measured (ExpSCR, left) or Paired-
596 SCHAF inferred (right) expression of training genes *CCDC80* (**c**) and *CD68* (**d**). Quality scores
597 and spatial correlations are noted.

598

599 **Extended Data Figure 3. Paired SCHAF whole transcriptome inference in ER+ MBC**

600 **a.** Spatial correlations between measured and SCHAF-inferred expression. Spatial correlations
601 (Pearson's r, x axis) between low-resolution measured (ExpLR) and the SCHAF inferred (coarsed
602 up) expression values for each measured gene across all ExpLR spots. Error bars: standard error
603 of folds assessed by cross validation. **b-c.** Quality scores capture performance. ER+ MBC samples
604 colored by expression in measured (ExpLR, left, low-resolution spots), coarsed-up SCHAF
605 inferences (middle) and cell-resolution SCHAF inferences (right) for illustrative genes with higher
606 (*ATP2C2*, b) and lower (*CENPL*, c) quality scores (denoted), in each of four folds assessed by
607 cross validation.

608

Extended Data Figure 4. Paired SCHAF gene program inference in ER+ MBC

a. Spatial correlations between measured and SCHAF-inferred program scores. Spatial correlations (Pearson's r, x axis) between low-resolution measured (ExpLR) and the SCHAF inferred (coarsed up) expression scores for each gene program (**Methods**). Error bars: standard error of folds assessed by cross validation. **b-c.** Quality scores capture performance. ER+ MBC samples colored by expression in measured (ExpLR, left, low-resolution spots), coarsed-up SCHAF inferences (middle) and cell-resolution SCHAF inferences (right) for illustrative gene programs with higher (Tumor Epithelial gene program, b) and lower (Serine Proteases gene program, c) quality scores (denoted), in each of four folds assessed by cross validation.

619 **Extended Data Figure 5. Paired SCHAF whole transcriptome inference in a new mouse pup**

620 New mouse pup samples colored by expression in measured (ExpSCR (ISH), left) and cell-
621 resolution SCHAF inferences (right) for ISH-measured test genes with higher (*Csrp3*, a) and lower
622 (*Rtn1*, b) quality scores (denoted).

623

624 **Extended Data Figure 6. Clustering of unpaired SCHAF inferred profiles resembles**
625 **measured scRNA-seq**

626 **a.** Jensen-Shannon divergence. Jensen-Shannon Divergence (JSD, x axis) between each sample's
627 cell type distribution in measured and Unpaired SCHAF-inferred scRNA-seq in SCLC (MSKCC,
628 left), MBC (HTAPP, middle), and placenta (right) samples. **b.** Agreement of cell type proportions
629 between measured and SCHAF-inferred scRNA-Seq. Proportion of cells (y axis) of each type (x
630 axis, as called by a classifier trained on scRNA-seq) in measured (blue) and SCHAF-inferred
631 (orange) scRNA-seq in the SCLC (MSKCC, left), breast cancer (HTAPP, middle) and placenta
632 (right) datasets. **c,d.** Correspondence of clustering metrics. Cluster separation (c, y axis) and local
633 coherence (d, y axis) scores in Uniform Manifold Approximation and Projection (UMAPs)
634 embedding of measured (blue) or Unpaired SCHAF-inferred (orange) scRNA-seq profiles for each
635 sample (x axis, assessed by cross-validation) in the placenta (right), MBC (HTAPP, middle) or
636 SCLC (MSKCC, left) datasets. **e.** Representative UMAPs. UMAP embedding of measured
637 (bottom) or Unpaired SCHAF-inferred (top) scRNA-seq profiles (dots) colored by cell types
638 (legends), for three representative samples from the MBC (HTAPP) dataset. The measured and
639 inferred scRNA-seq are projected in a common space.

640 **Extended Data Figure 7. Cell type distributions in Unpaired SCHAF inferences are similar**

641 **to measured scRNA-seq**

642 KDE plots for measured (blue) and SCHAF inferred (orange) scRNA-seq gene expression for

643 selected genes (labeled on top) from MBC (HTAPP, top row) or SCLC (MSKCC, other rows).

644 EMD and DMS values are denoted.

645

646 **Supplementary Figure Legends**

647 **Supplementary Figure 1. Paired SCHAF preserves expression heterogeneity within cell**
648 **types.**

649 Heterogeneity (y axis, **Methods**) of expression profiles within each cell type (*x* axis) in measured
650 (ExpSCR, blue) and SCHAF-inferred (orange) profiles in the mouse pup (left), ER+ MBC
651 (middle) and ER- MBC (right) datasets. Error bars: standard error of folds assessed by cross
652 validation.

653

654 **Supplementary Figure 2. Benchmarking Paired SCHAF inferences with other methods**

655 Average spatial correlation across spatially-measured genes (**a,** y axis) and cell type accuracy (**b**,
656 y axis; defined as proportion of cells in each fold assigned to the correct cell type label) for
657 inferences by Paired SCHAF (red), SPiRiT[35] (dark blue), ST-Net[31] (light blue), DeepPT[36] (purple),
658 and HE2RNA[30] (green) in each dataset (x axis) for the mouse pup (left), ER+ MBC (middle) and
659 ER- MBC (right) datasets. Error bars: standard error of folds assessed by cross validation.

660

661 **Supplementary Figure 3. Unpaired SCHAF learns to predict a tissue's single cell omics data**

662 **from its H&E image**

663 Detailed overview of Unpaired SCHAF training (top and middle) and inference (bottom).

664

665 **Supplementary Figure 4. Characteristics of scRNA-seq profiles inferred by Unpaired**
666 **SCHAF**

667 **a.** Agreement with pseudobulk RNA-Seq. Correlation (Pearson's r, y axis) between pseudobulk
668 profiles from the measured and Unpaired SCHAF inferred scRNA-Seq for each sample (x axis) in
669 the placenta (right), MBC (HTAPP, middle) and SCLC (MSKCC, left) datasets. **b.** Preservation
670 of heatmaps of order-ranked entries of gene-gene correlation matrices in inferred profiles.
671 Heatmaps, as described above (and in **Methods**), were formed from the 100 most highly variable
672 genes in each sample's scRNA-seq data, from SCHAF inferred (top) and scRNA-seq (bottom)
673 from three samples from each of the Placenta (bottom), MBC (HTAPP, top right), and SCLC
674 (MSKCC, top left) datasets. Measured heatmaps are clustered and the same order is imposed on
675 the SCHAF heatmaps. Samples' heatmap meta correlations (Pearson's *r* between corresponding
676 flattened entries of the SCHAF and scRNA-seq heatmaps) are noted, **c.** Agreement of meta
677 correlations. Heatmap meta correlations (y axis, **Methods**, described above) for each sample (x
678 axis) in the placenta (right), MBC (HTAPP, middle) and SCLC (MSKCC, left) datasets.

679

680 **Supplementary Figure 5. Cell type pseudobulk profiles inferred by Unpaired SCHAF**

681 **a.** Cell type pseudobulk agreement across samples. Cell type pseudobulk correlation score (y axis,
682 **Methods**), with dots representing each cell type's pseudobulk correlation and error bars capturing
683 standard error across cell types, between cell type specific pseudobulk profiles from measured and
684 SCHAF inferred scRNA-seq in each sample (x axis) in the Placenta (right), MBC (HTAPP,
685 middle), and SCLC (MSKCC, left) datasets. **b.** Cell type pseudobulk agreement in representative
686 samples. Correlation (Pearson's *r*, y axis) between pseudobulk profiles from measured and
687 SCHAF inferred scRNA-seq in each cell type (x axis) in three representative samples from the
688 Placenta (bottom), MBC (HTAPP, middle), and SCLC (MSKCC, top).

689

690 **Supplementary Figure 6. Unpaired SCHAF preserves the extent of expression heterogeneity**
691 **within cell types**

692 Mean heterogeneity (x axis, **Methods**; across samples from cross validation) of expression profiles
693 within a cell type (*x* axis) in measured (blue) and Unpaired SCHAF-inferred (orange) scRNA-seq
694 for SCLC (MSKCC, left), MBC (HTAPP, middle), and placenta (right). Error bars: standard error
695 of samples assessed by cross validation.

696

697

698

699

700

701

702

703

704

705

706

707 **Supplementary Figure 7. Agreement of gene expression distributions in each cell type in**
708 **Unpaired SCHAF**

709 Distribution of Earth Mover's Distance (EMD, x axis) between the measured and Unpaired
710 SCHAF-inferred expression distributions of each gene in each cell type (label on top) in the breast
711 cancer (HTAPP, left), placenta (middle) and SCLC (MSKCC, right) datasets, in the cross
712 validation (holding out one sample at a time). Error bars: standard error of samples assessed by
713 cross validation.

714

715 **Supplementary Figure 8. Agreement of highly variable gene expression distributions in each**
716 **cell type in Unpaired SCHAF**

717 Distribution of Earth Mover's Distance (EMD, x axis) between the measured and Unpaired
718 SCHAF-inferred expression distributions of each of highly variable genes in each cell type (label
719 on top) in the breast cancer (HTAPP, left), placenta (middle) and SCLC (MSKCC, right) datasets,
720 in the cross validation setting (holding out one sample at a time). Error bars: standard error of
721 samples assessed by cross validation. Highly variable genes: a gene that is found to be one of the
722 500 most highly variable in any sample of a dataset.

723

724 **Supplementary Figure 9. Aggregation of evaluation metrics for Unpaired SCHAF**

725 Heat maps displaying values of metrics used in evaluating Unpaired SCHAF across samples in
726 breast cancer (HTAPP, left), SCLC (MSKCC, middle), and placenta (right) datasets.

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741 **Supplementary Figure 10. Benchmarking Unpaired SCHAF inferences**

742 Mean Earth Mover's Distance from measured scRNA-seq (EMD, y axis, **a**, left), Jensen-Shannon
743 Divergence from measured scRNA-Seq (JSD, **a**, y axis, right), mean correlation to measured
744 pseudobulk RNA-seq (Pearson's r, y axis, **b**, left), mean cell type pseudobulk correlation
745 (Pearson's r, y axis, **b**, middle), and gene-gene correlation heatmap meta correlation (Pearson's r,
746 y axis, **b**, right), for scRNA-seq profiles inferred by Unpaired SCHAF (without cell types in
747 training) (dark blue) and other methods (a-c, color legends), in each data set (x axis). Error bars:
748 standard error of samples assessed by cross validation.

749

750 **Supplementary Figure 11. Comparison of Unpaired SCHAF spatial inferences to expert**
751 **pathologist annotations**

752 **a.** Global spatial accuracy. Proportion of cells (y axis) with cell types assigned to SCHAF
753 inferences that agree with expert pathologist annotations in each sample (x axis) from the Placenta
754 (left) and MBC (HTAPP, right) datasets. **b**. Cell type specific accuracy in the placenta. Number of
755 cells with each pathologist annotation (rows) and SCHAF-inferred annotation (columns) in each
756 placenta sample. **c.** Regional accuracy in relatively complex MBC samples. Left: Sample sections
757 with cells (left) or regions (right) colored by one of five region categories based on SCHAF-
758 inferred (left) or coarse pathologist (right) annotations. Right: Number of cells with each
759 pathologist annotation (by region; rows) and SCHAF-inferred annotation (columns) in the
760 corresponding MBC sample.

761

762 **Supplementary Figure 12. Unpaired SCHAF spatial annotations match expert pathologist**
763 **annotations in the placenta better when trained with cell type information**

764 **a.** Global spatial accuracy. Proportion of cells (y axis) with cell types assigned to SCHAF
765 inferences that agree with expert pathologist annotations in each sample (x axis) from the Placenta,
766 when SCHAF is trained with (blue) and without (orange) cell types in training. **b.** Human placenta
767 samples (rows, labeled on left) colored by annotations from an expert pathologist (right) or based
768 on Unpaired SCHAF data, with models trained with (middle) and without (left) cell type
769 information. **c.** Cell type specific accuracy in the placenta. Number of cells with each pathologist
770 annotation (rows) and SCHAF-inferred (without cell types in training) annotation (columns) in
771 each placenta sample.

772

773 **Supplementary Figure 13. Unpaired SCHAF spatial annotations match expert pathologist**
774 **annotations in MBC (HTAPP) better when trained with cell type information**

775 **a.** Global spatial accuracy. Proportion of cells (y axis) with cell types assigned to SCHAF
776 inferences that agree with expert pathologist annotations in each sample (x axis) from MBC
777 (HTAPP) samples, when SCHAF is trained with (blue) and without (orange) cell types in training.
778 **b.** MBC (HTAPP) samples (rows, labeled on left) colored by annotations from an expert
779 pathologist (right) or based on Unpaired SCHAF data, with models trained with (middle) and
780 without (left) cell type information. **c.** Cell type specific accuracy in MBC (HTAPP). Number of
781 cells with each pathologist annotation (rows) and SCHAF-inferred (without cell types in training)
782 annotation (columns) in each placenta sample.

783

784

785

786

787

788

789

790

791

792 **Supplementary Figure 14. Comparison of spatial cell types from Unpaired SCHAF**
793 **inferences to cell neighborhood density in the placenta**

794 **a**. Placenta sections from each sample (rows, labeled on left) colored by SCHAF inferred cell types
795 (left) or the number of cells in each cell's H&E tile (right, color bar at bottom). **b**. Distribution of
796 number of cells per H&E tile (y axis) in cells annotated by Unpaired SCHAF to each category (x
797 axis, color legend). **c**. Distribution of number of cells per H&E tile (y axis) in cells annotated by a
798 pathologist to each category (x axis, color legend).

799

800 **Supplementary Figure 15. Comparison of spatial cell types from Unpaired SCHAF**
801 **inferences to RGB value intensity in the placenta**

802 **a**. Placenta sections from each sample (rows, labeled on left) colored by SCHAF inferred cell types
803 (left) or by mean RGB value intensity in each cell's H&E tile (right, color bar at bottom). **b**.
804 Distribution of mean RGB value intensity per H&E tile (y axis) in cells annotated by Unpaired
805 SCHAF to each category (x axis, color legend). **c**. Distribution of mean RGB value intensity per
806 H&E tile (y axis) in cells annotated by a pathologist to each category (x axis, color legend).

807

808 **Supplementary Figure 16. Comparison of spatial cell types from Unpaired SCHAF**
809 **inferences to cell neighborhood density in MBC (HTAPP)**

810 **a**. MBC (HTAPP) sections from each sample (rows, labeled on left) colored by SCHAF inferred
811 cell types (left) or the number of cells in each cell's H&E tile (right, color bar at bottom). **b**.
812 Distribution of number of cells per H&E tile (y axis) in cells annotated by Unpaired SCHAF to
813 each category (x axis, color legend). **c**. Distribution of number of cells per H&E tile (y axis) in
814 cells annotated by a pathologist to each category (x axis, color legend).

815

816

817 **Supplementary Figure 17. Comparison of spatial cell types from Unpaired SCHAF**
818 **inferences to RGB value intensity in HTAPP**

819 **a**. MBC (HTAPP) sections from each sample (rows, labeled on left) colored by SCHAF inferred
820 cell types (left) or by mean RGB value intensity in each cell's H&E tile (right, color bar at bottom).
821 **b**. Distribution of mean RGB value intensity per H&E tile (y axis) in cells annotated by Unpaired
822 SCHAF to each category (x axis, color legend). **c**. Distribution of mean RGB value intensity per
823 H&E tile (y axis) in cells annotated by a pathologist to each category (x axis, color legend).

824

825

826

827

828

829

830

831

832

833

834 **Supplementary Figure 18. Comparison of Unpaired SCHAF inferences to spatial STARmap**
835 **cell type annotations in the placenta.**

836 **a.** Global spatial cell type accuracy. Proportion of ExpLR (low-resolution (coarsed up) STARmap)
837 spots (y axis) with cell types that match those on (coarsed up) SCHAF inferences (x axis) for the
838 portion of the placenta data measured by STARmap and H&E. **b,c**. High and low resolution cell
839 type agreement. **b, and c top**, Placenta portion measured by both ExpSCR (STARmap) and H&E
840 (b, and c top) colored by cell type annotations (legend) on SCHAF-inferred (top) and measured
841 (bottom) data at high (ExpSCR) resolution (b) and at coarsed-up low resolution (c, top). **c, bottom**,
842 Number of cells of each cell type in ExpLR (coarsed-up STARmap) measured (rows) and SCHAF-
843 inferred (column) data for each sample.

844

845 **Supplementary Figure 19. Comparison of Unpaired SCHAF inferences to spatially**
846 **MERFISH cell type annotations in MBC (HTAPP)**

847 **a.** Spatial cell type accuracy. Proportion of ExpLR (low-resolution (coarsed up) MERFISH) spots
848 (y axis) with cell types that match those on (coarsed up) SCHAF inferences in each sample (x axis)
849 in the MBC (HTAPP) dataset. **b,c**. Comparison of spatial cell annotations in individual sections.
850 **b, and c top**, Tumor sections from different samples (columns) colored by cell type (color legend)
851 from SCHAF inferred (top) and MERFISH measured (bottom) data at cellular resolution (**b,**
852 ExpSCR) or coarsed-up low resolution (**c top,** ExpLR). **c,bottom,** Number of cells of each type
853 in SCHAF inferred (columns) and coarsed-up MERFISH (ExpLR, rows) spatial data for each of
854 the samples in b,c.

855

856

857

858

859

860

861

862

863

Supplementary Figure 20. Comparison of Unpaired SCHAF spatial expression to STARmap measured genes in the placenta

a. Spatial gene expression accuracy. Average spatial correlation (y axis, left), total number of genes with higher (>0.4) spatial correlation (y axis, middle) and fraction of genes with higher (>0.4) spatial correlation (y axis, right) between ExpLR measurements (coarsed-up, low resolution STARmap) and Unpaired-SCHAF inferred spatial expression in each placenta sample (x axis). **b,c,** Comparison of individual genes. Normalized expression level (color bar) in each placenta sample (major rows, labeled on left) of each of four genes (labeled on top) in SCHAF inferred (top panels) and STARmap measured (bottom panels) data at either single cell resolution (b) or coarsed-up low resolution (c).

875 **Supplementary Figure 21. Comparison of Unpaired SCHAF spatial expression to MERFISH**
876 **measured genes in MBC (HTAPP)**

877 **a.** Spatial gene expression accuracy. Average spatial correlation (y axis, left), total number of genes
878 with higher (>0.4) spatial correlation (y axis, middle) and fraction of genes with higher (>0.4)
879 spatial correlation (y axis, right) between ExpLR measurements (coarsed-up, low resolution
880 MERFISH) and Unpaired-SCHAF inferred spatial expression in each MBC (HTAPP) sample (x
881 axis). **b,c,** Comparison of individual genes. Gene expression level (color bar) in each MBC sample
882 (major rows, labeled on right) of each of four genes (labeled on top) in MERFISH measured (top
883 panels) or SCHAF inferred (bottom panels) data at either single cell resolution (b) or coarsed-up
884 low resolution (c). **d.** Comparison to other inference methods. Mean number (across samples) of
885 coarsed-up, low resolution MERFISH (ExpLR) measured genes with spatial correlation>0.4 (y
886 axis) with each inference method (x axis, color legend). Dots represent individual samples.

887 **Supplementary Table Legends**
888
889 **Supplementary Table 1.** The number of cells and genes measured by each modality in each
890 fold/sample.
891
892 **Supplementary Table 2.** For each gene in the predicted original mouse pup data, in each fold and
893 the cross validation average, the spatial correlation measurement (when the gene was measured by
894 ExpSCR), predictive quality score, and whether or not the gene was training or hold-out (when the
895 gene was measured by ExpSCR).
896
897 **Supplementary Table 3.** For each gene in the predicted new mouse pup data, the spatial
898 correlation measurement (when the gene was measured by ExpSCR) and predictive quality score.
899
900 **Supplementary Table 4.** For each gene in the predicted ER+ MBC data, in each fold and in the
901 cross validation average, the spatial correlation measurement (when the gene was measured by
902 ExpSCR), low resolution spatial correlation measurement (when the gene was measured by
903 ExpLR), predictive quality score, and whether or not the gene was training or hold-out (when the
904 gene was measured by ExpSCR).
905
906 **Supplementary Table 5.** For each program analyzed in the predicted ER+ MBC data, in each fold
907 and the cross validation average, the low resolution spatial correlation measurement, predictive
908 quality score, and, for the genes in the program that were also measured by ExpLR (Visium), the
909 number of said genes, proportion of said genes from the original program, and the list of genes.
910

Supplementary Table 6. For each gene in the predicted ER- MBC data, the spatial correlation measurement (when the gene was measured by ExpSCR) and predictive quality score, and whether or not the gene was training or hold-out (when the gene was measured by ExpSCR).

Supplementary Table 7. For each gene predicted in the HTAPP data, in each sample and in the cross validation average, the Earth Mover's Distance, Gene Distribution Matching Score, and the spot-level spatial correlation measurement (when the gene was measured by ExpSCR).

Supplementary Table 8. For each gene predicted in the Placenta data, in each sample and in the cross validation average, the Earth Mover's Distance, Gene Distribution Matching Score, and the spot-level spatial correlation measurement (when the gene was measured by ExpSCR in the portion of the tissue measured by ExpSCR)

Supplementary Table 9. For each gene predicted in the MSKCC data, in each sample and in the cross validation average, the Earth Mover's Distance and Gene Distribution Matching Score.

932 **METHODS**

933 **Human subjects**

934 All samples used in this study were collected under protocols approved by the relevant Institutional
935 Review Boards. All HTAPP MBC samples used in this study underwent IRB review and approval
936 at Dana Farber Cancer Institute (DFCI) (protocol 05-246), as well as at the Broad Institute
937 (protocol #15-370B). All SCLC samples were from patients who underwent a surgical resection
938 or tissue biopsy at Memorial Sloan Kettering Cancer Center (MSKCC), with patients identified
939 and biospecimens collected prospectively from 2005 to 2013 (MSKCC protocol #06-107 A(13)).
940 All patients provided written informed consent through the corresponding Institutional Review
941 Board-approved biospecimen collection and analysis protocol. Tumor specimens were collected
942 intraoperatively, immediately flash-frozen in liquid nitrogen in optimal cutting temperature (OCT)
943 compound, and stored at –80°C until further processing.

944 **Matched transcriptomic and H&E images data**

945 Several different tissue corpora were used.

946 One tissue slice spanning a whole, one day old mouse pup that was profiled by ExpSCR (Xenium)
947 and H&E stained was obtained from a public repository[16]. scRNA-seq from another one day old
948 mouse pup was obtained from a different public repository[34]. One tissue slice spanning all of
949 another one day old mouse pup was assayed by both ISH and H&E in this study (as described
950 below). Cell types of ExpSCR profiles were manually annotated based on k-means (k=10)
951 clustering performed by curators of the publicly available data[16].

952 One tissue section from a primary ER+ metastatic breast cancer (MBC) tumor profiled by both
953 Xenium and H&E, scRNA-seq from an adjacent section, and Visium (ExpLR) from another
954 adjacent section as well as one primary ER- MBC tumor section profiled by both Xenium and
955 H&E were all obtained from a single public repository[16].

956 Twenty-four (24) small cell lung cancer (SCLC) samples profiled by snRNA-seq and matching
957 H&E stains were obtained from an unpublished study (as described below) from Memorial Sloan
958 Kettering Cancer Center (MSKCC). The snRNA-seq data came from a portion of the same
959 specimen used in H&E staining. Cell types of snRNA-seq profiles were manually annotated by
960 the curators of the data.

961 Nine other MBC tissue samples, with H&E and snRNA-seq/scRNA-seq and MERFISH data, were
962 obtained from the Human Tumor Atlas Project Pilot (HTAPP)[39]. The H&E, sc/sn-RNA-seq and
963 MERFISH data of each patient sample were obtained from different biopsies of the same
964 metastasis. Cell types of scRNA-seq profiles were manually annotated. Coarse annotations were
965 manually performed on H&E images by an expert histopathologist. Each scRNA-seq profile had
966 two annotations: a specific cell type annotation, and a coarser one with the histopathology
967 annotation categories. Each MERFISH profile had a specific cell type annotation.

968 Four human placenta samples profiled by scRNA-seq, STARmap, and matching H&E stains were
969 obtained from a study from the Massachusetts General Hospital[40]. scRNA-seq was collected from
970 a section adjacent to that used in H&E staining. STARmap data came from a portion of another
971 section adjacent to the H&E section. Cell types of scRNA-seq profiles were manually annotated
972 by the curators of the data[40]. The H&E stains were collected for the purpose of this study, and
973 precise annotations were manually performed on H&E images by an expert placental

974 histopathologist (as described below). Each scRNA-seq profile had two annotations: a specific cell
975 type annotation, and a coarser one to match up with one of the histopathology annotations. Each
976 STARmap profile had a specific cell type annotation.

977 Further details on these data sets are in **Supplementary Tables 1-9**.

978 **New mouse pup ISH data for this study**

979 **Sample preparation.** Glasses were coated with poly(d-lysine) solution. An OCT-embedded P1
980 mouse (C57BL/6J) was cut into 10-μm, sections then fixed with 4% paraformaldehyde in PBS for
981 15 min and permeabilized by prechilled methanol at −20 °C for 20 min before hybridization.

982 **Library construction.** STARmap-ISH genes were selected as described in a below section.
983 SNAIL probes were designed and the probe library was constructed according to previous
984 studies[21]. The probes were dissolved in ultrapure RNase-free water and pooled to the final
985 concentration of 10 nM per probe. The probe mixture was heated at 40 °C for 15 min and then
986 equilibrated to 37 °C. Tissue samples were removed from −20 °C tissues, equilibrated to RT and
987 treated with 10 mM Tris, pH 7.5 for 10 min. The samples were then washed by PBS-TR (0.1%
988 Tween- 20, 0.1 U μl−1 of SUPERase•In in PBS) and incubated in hybridization buffer (2× saline–
989 sodium citrate (SSC), 10% formamide, 1% Tween-20, 20 mM ribonucleoside vanadyl complex,
990 0.1 mg ml−1 of yeast transfer RNA, 0.1 U ul−1 of SUPERase•In, SNAIL probes with 5 nM per
991 probe) at 40 °C with gentle shaking for 48 h. The samples were washed by PBS-TR twice for 20
992 min at 37 °C and then by 4× SSC in PBS-TR for 20 min at 37 °C, following a rinse by PBS-TR at
993 RT. The SNAIL padlock probes annealed to the samples were ligated by incubating with T4 DNA
994 ligation mixture (1:10 dilution of T4 DNA ligase, 0.2 mg ml−1 of BSA, 0.5 U ul−1 of
995 SUPERase•In) for 2 h at RT with gentle agitation, followed by a 5-min wash of PBS-TR twice.

996 The samples were then incubated in the RCA mixture (1:50 dilution of Phi29 DNA polymerase,
997 250 μM dNTP, 20 μM 5-(3-aminoallyl)-dUTP, 0.2 mg ml−1 of BSA, 0.2 U μl−1 of SUPERase•In)
998 for 2 h at 30 °C with gentle agitation, followed by a 5-min wash of PBS-TR twice. Subsequently,
999 the samples were treated with 25 mM acrylic acid N-hydroxysuccinimide ester for 2 h at RT with
1000 agitation, rinsed by PBS-T (0.1% Tween-20 in PBS) once and incubated with monomer buffer
1001 (4% acrylamide, 0.2% bis-acrylamide, 2× SSC in H2O) for 15 min for polymerization
1002 pretreatment. The buffer was removed and 30 μl of monomer mixture (0.1% ammonium persulfate,
1003 0.1% tetra- methylethylenediamine in monomer buffer) was directly added to the center of each
1004 sample, which was immediately covered by a coverslip (no. 2 coverslip was coated with Gel-Slick
1005 Solution according to the manufacturer's instructions) and allowed to polymerize in ambiance for
1006 1 h. The tissue–gel hybrid was washed with PBS-T twice and cleared by proteinase K digestion
1007 mixture (50 mM Tris, pH 7.5, 100 mM NaCl, 1% SDS, 0.2 mg ml−1 of proteinase K in H2O) at
1008 37 °C overnight. On the next day, the samples were treated with a dephosphorylation mixture
1009 (1:100 dilution of shrimp alkaline phosphatase, 0.2 mg ml−1 of BSA, 1:10 dilution of CutSmart
1010 buffer in H2O) and rinsed with PBS-T.

1011 **Imaging and sequencing.** The 19-nt fluorescent oligo complementary to DNA amplicon was
1012 diluted at 100 nM in 1× SSC dissolved in PBS-T and samples were incubated at RT for 60 min,
1013 then washed by PBS-T 3× for 5 min each, DAPI staining was performed for cell segmentation.
1014 Images were acquired by Leica Stellaris 5 confocal microscope with a 405 diode, white light laser
1015 and 20x air objective.

1016 **In situ hybridization image processing.** For STARmap-ISH images, the process was initiated by
1017 detecting spots within the raw images and generating multiplexed images, encompassing multiple
1018 genes within a single tissue. The initial step involved aligning images from two rounds of

1019 acquisition using the 2D Fourier transform. Subsequently, a method was employed to identify
1020 regions with local maxima, which serve as signal spots for visualization. Specifically, maximum
1021 and minimum filters were applied from the *scipy.ndimage.filters* python library to locate potential
1022 local maxima and minimal values within neighborhoods sized at 15 pixels. Regions where the
1023 difference between local maxima and minimal values exceeded a defined threshold of 100 were
1024 identified and marked as local maxima. Connected regions were delineated as a result of this
1025 identification process. Subsequently, cell segmentation was performed on DAPI channel images
1026 using the pretrained *2D_versatile_fluo* model from StarDist[47]. After segmentation, the centroids
1027 and areas of the identified cells were extracted. These centroids served as markers for further
1028 refinement using a watershed algorithm. The resulting segmented image was then used to assign
1029 decoded transcripts to individual cells and to generate the count matrix used for downstream
1030 analysis.

1031 **Staining procedure for new mouse pup and placenta H&E data for this study**

1032 For Hematoxylin and Eosin (H&E) staining, slides containing 10 µm sections were loaded onto a
1033 Histo-Tek® SL Slide Stainer (Sakura Finetek, USA) and stained with the appropriate H&E
1034 staining program. H&E-stained slides were mounted with Richard-Allan Scientific™ Cytoseal™
1035 XYL (Thermo Fisher Scientific 8312-4) using glass coverslips (VWR 48393-059). Slide images
1036 were obtained by an Axio Scan Z1 Slide Scanner and analyzed with the ZEN Blue Software (Zeiss,
1037 Germany).

1038 **Small cell lung cancer H&E stains and snRNA-seq data for this study**

1039 **Histology (H&E Staining).** Formalin-fixed, paraffin-embedded (FFPE) tissue sections were
1040 prepared from a portion of each tumor and stained with hematoxylin and eosin (H&E) using

1041 standard protocols at the MSKCC pathology core facility. Stained slides were reviewed by a board-
1042 certified pathologist to confirm tumor content and histological features.

1043 **Sample Processing and snRNA-seq Library Preparation.** Frozen SCLC tumor samples (n =
1044 24) were used for single-nucleus RNA sequencing (snRNA-seq). Nuclei were isolated from each
1045 sample using a Dounce homogenizer in an ice-cold lysis buffer, following standard protocols for
1046 frozen tissue. The resulting nuclear suspensions were filtered through a 40 μm cell strainer to
1047 remove debris. Nuclei were stained with DAPI and sorted by fluorescence-activated cell sorting
1048 (FACS) to enrich for intact, singlet nuclei. snRNA-seq libraries were prepared using the Chromium
1049 Single Cell 3' Reagent Kits (10x Genomics), with 23 samples processed using v3 chemistry and
1050 one sample using v2 chemistry, according to the manufacturer's instructions. Approximately
1051 10,000 nuclei per sample were loaded onto the Chromium instrument. Libraries were sequenced
1052 on an Illumina platform with a minimum target depth of 50,000 reads per nucleus.

1053 **snRNA-seq Data Processing.** The snRNA-seq data processing pipeline used in this study is
1054 similar to another popular approach[48]. Demultiplexed FASTQ files were generated from raw
1055 sequencing reads using Cell Ranger mkfastq (v2.0 and v3.0, 10x Genomics). Reads were aligned
1056 to the human GRCh38 genome and gene counts were quantified as unique molecular identifiers
1057 (UMIs) using Cell Ranger count (v3.1.0, 10x Genomics). For snRNA-seq data, reads mapping to
1058 both introns and exons were counted to increase gene detection sensitivity, following a
1059 conventional approach[49]. A pre-mRNA reference was built using Cell Ranger mkref (v3.0) and a
1060 modified GTF file, as described in the referenced protocol.

1061 Sequencing data were processed using the Cumulus Cellranger workflow on the Terra cloud
1062 platform. CellBender was used to remove ambient RNA contamination from each sample. Count

1063 matrices were aggregated across all 24 SCLC samples using the Cumulus workflow, which wraps
1064 Pegasus for cloud-based computation. Quality control filtering retained nuclei expressing 200–
1065 6,000 genes and with less than 20% of UMIs mapping to mitochondrial genes. Robust genes were
1066 selected at a rate of 0.05% (i.e., expressed in at least 3 out of 6,000 cells). Counts were normalized
1067 to TP100K and log-transformed. The top 2,000 highly variable genes were selected in a batch-
1068 aware manner, and principal component analysis (PCA) was performed using 50 principal
1069 components. Data integration was conducted using *harmony-pytorch*, and a k-nearest neighbor
1070 (kNN) graph (k = 100) was constructed using hnswlib. Clustering was performed using the Leiden
1071 algorithm (resolution = 2.0), and UMAP was used for visualization. Doublets were identified and
1072 removed using Pegasus' doublet detection algorithm. After doublet removal, the analysis steps
1073 from highly variable gene selection to UMAP generation were repeated. Differential expression
1074 analysis was performed using the Mann-Whitney U test (FDR = 5%) to identify cluster-specific
1075 markers, and putative cell types were annotated based on these results and canonical cell type
1076 markers. Annotated epithelial cell clusters were further subclustered (Leiden resolution = 1.0) to
1077 distinguish tumor cells from AT1, AT2, and secretory cells.

1078 **Placenta expert histopathologist annotations for this study**

1079 Annotations were manually performed on H&E images by a physician pathologist with fellowship
1080 training in placental and perinatal pathology. Digital scans of the frozen placenta tissue were
1081 obtained and annotated by manually "drawing" over the tissue using photo editing software.
1082 Annotations were performed by region and groups/clusters of cell types rather than by individual
1083 cells, due to the processing artifact inherent to freezing fresh tissue, with each group of cell types
1084 being assigned a different color. Cell types were grouped based on tissue processing artifact, image
1085 resolution, and correlation with known placental physiologic relevance. Syncytiotrophoblast and

villous cytotrophoblast were marked green, extravillous trophoblast (villous growth columns) were marked blue, decidua and maternal fibrin and blood marked brown. Villous parenchyma (including fibroblasts, stromal cells, Hofbauer cells, and villous vessels including endothelial cells, vascular smooth muscle, and fetal red blood cells) were marked orange. Areas inside the villous parenchyma where fetal vessels were identified with a high degree of likelihood (probable or definitive vessel lumen, with or without intraluminal nucleated red blood cells) were marked red. Foci of intravillous hemorrhage (villi filled with fetal red blood cells), likely a physical artifact caused by the clinical procedure of collecting tissue from the patient, were also marked orange. Software was then used to assign each nucleus from the original H&E image to a cell type group based on the overlying color annotation, so that cell-less empty white space was not annotated with a cell identity.

Transcriptomics data pre-processing

All ExpSCR data and their corresponding H&E images were aligned at single cell resolution with the Xenium Explorer 2 software program. Cell types of ExpSCR profiles in the original datasets were manually annotated based on kmeans (k=10) clustering performed by curators of the publicly available data.

Cell profiles with less than eight genes were removed (except for ISH profiles, where only eight genes were measured), followed by the removal of genes expressed in no cells. Each cell's counts x were transformed by a *log1p* ($f(x) = log(x + 1)$) transformation. Next, within each tissue dataset, all genes shared across all samples in the dataset were identified and retained for further analysis. The above quality-control steps were performed with python's *scanpy* package[50] with default parameters except for the following: *sc.pp.filter_cells*, *sc.pp.filter_genes*, and *sc.pp.log1p*.

Cell segmenting and tiling histology images

Cells in H&E images from sections also profiled by ExpSCR were segmented according to the location of the single cell transcriptomic measurements.

StarDist[47] was used to segment cells in the H&E images in all samples used in Unpaired SCHAF.

Histology images were tiled into smaller, potentially-overlapping tiles, each tile centered around a segmented cell's location, with tile size selected via hyperparameter tuning.

Paired SCHAF training

All machine learning models were built and trained using the python deep learning library *pytorch*[51].

Given a training dataset of tissue samples, each with a corresponding H&E image, spatial transcriptomics from the same section as the H&E, and sc/snRNA-seq data, Paired SCHAF was trained to infer an entire spatially resolved, single cell resolution, transcriptome-scale dataset from a histology image.

First, the Supervised Spatial Pretraining model was trained to infer a cell's training genes from its corresponding H&E tile. The model follows an encoder decoder architecture, with the encoder initialized with weights from UNI[32], a state-of-the-art foundation model for computational pathology, currently trained on the most data, with the largest number of parameters, and best performance[32]. UNI follows a large vision transformer architecture, with an H&E encoder initialized via the python deep learning library *timm*, via the function *timm.create_model("vit_large_patch16_224", img_size=224, patch_size=16, init_values=1e-5,*

num_classes=0, dynamic_img_size=True). A 1024-dimensional embedding of each tile was output by the encoder. The decoder consisted of two fully connected layers: one from 1024 to 512 dimensions, and one from 512 dimensions to the number of training genes. After the first linear layer but before the second was a batch normalization layer followed by a ReLU layer.

An expression profile was predicted by resizing the tile to have both length and width 224 pixels, normalizing it to a RGB mean of (0.485, 0.456, 0.406) and an RGB standard deviation of (0.229, 0.224, 0.225), passing the input tile to the encoder, and then the outputs of the encoder to the decoder. Before training, the expression data for training were z-score normalized so each gene had a mean value of 0 and a standard deviation of 1. The encoder decoder model was then trained on all the spatial transcriptomic profiles of all training samples, with a mean-squared-error (MSE) loss, optimized via stochastic gradient descent, using the Adam optimizer[52], for up to 20 epochs, with a batch size of 64, at a learning rate of .000005.

After training Supervised Spatial Pretraining, and prior to training Paired SCHAF, Tangram was used to map scRNA-seq/snRNA-seq data onto the spatial transcriptomics data with the *tangram* software library from python. To this end, after filtering out cells in both scRNA-seq and spatial transcriptomics data with fewer than 8 genes with non-zero expression (via *scanpy*'s *sc.pp.filter_cells* function with *min_genes*=8), all spatial transcriptomics data and all scRNA-seq data were separately z-scored. Due to memory constraints of *tangram* software, if there were more than 45,000 scRNA-seq profiles in the sample, 45,000 cells were randomly sampled and retained. The intersection of the genes measured by scRNA-seq and the spatial transcriptomics training genes was termed *tangram_genes*. Spatial transcriptomics data were then partitioned into 25,000 cell chunks to accommodate *tangram* memory constraints. The following three functions were run in succession on each chunk: *tg.pp_adatas(sc_rna, st_chunk, genes=tangram_genes)*, *ad_map =*

tg.map_cells_to_space(sc_rna, st_chunk,mode="cells", density_prior='uniform',num_epochs=125,), projected_sc = tg.project_genes(ad_map, sc_rna), where *sc_rna* was the processed scRNA-seq data and *st_chunk* was a chunk of the spatial transcriptomics data. All *projected_sc* were then concatenated together to represent the final Tangram-mapped scRNA-seq data onto the spatial transcriptomics. Mouse scRNA-seq data were mapped to ExpSCR data, and cancer scRNA-seq data were mapped to ER+ ExpSCR data.

Finally, Paired SCHAF was trained. Paired SCHAF is an encoder decoder model. The encoder had the same architecture as Supervised Spatial Pretraining, with a large vision transformer, where an H&E tile encoder initialized with the python deep learning library *timm*, via the function *timm.create_model("vit_large_patch16_224", img_size=224, patch_size=16, init_values=1e-5, num_classes=0, dynamic_img_size=True)*. The encoder was initialized with weights from the encoder at the end of the training of Supervised Spatial Pretraining. The decoder consisted of four fully-connected layers in the following order (1) from 1024 to 2048 dimensions, (2) from 2048 to 4096 dimensions, (3) from 4096 to 8192 dimensions, and (4) from 8192 dimensions to the number of genes in the Tangram-mapped expression data. After each of the first three linear layers (but before the linear layer following each one) was a batch normalization layer followed by a ReLU layer.

An expression profile was predicted by resizing the tile to have both length and width 224 pixels, normalizing it to a RGB mean of (0.485, 0.456, 0.406) and an RGB standard deviation of (0.229, 0.224, 0.225), passing the input tile to the encoder, and then the outputs of the encoder to the decoder. Before training, spatial transcriptomics data for training was z-score normalized for each gene to a mean value of 0 and a standard deviation of 1. The encoder decoder model was then trained on all the spatial transcriptomics profiles of all training samples, with a mean-squared-error

1174 (MSE) loss, optimized via stochastic gradient descent, using the Adam optimizer, for up to 20
1175 epochs, with a batch size of 64, at a learning rate of .000005.

1176 **Unpaired SCHAF training**

1177 All machine learning models were built and trained using the python deep learning library
1178 *pytorch*[51].

1179 For Unpaired SCHAF, a model was trained to learn a single cell's expression profile from a single
1180 histology-image tile, using an adversarial encoder-decoder[53] based framework for domain
1181 translation applied to image-tile and gene-expression domains.

1182 First, histology images were segmented and tiled (as described above) and the foundation model
1183 UNI, as described above, was used to create representations of every histology image tile. The
1184 scRNA-seq foundation model scGPT was used to create representations of every scRNA-
1185 seq/snRNA-seq profile. scGPT was chosen at the time of the study as the largest in the number of
1186 parameters and training data, and a top performer[33]. The resulting H&E embeddings represented
1187 each tile as a 1,024-dimensional vector and each scRNA-seq profile as a 512-dimensional vector.

1188 An encoder-decoder, *G*, was trained to reconstruct an scRNA-seq profile on all the expression
1189 profiles of all training samples, with a mean-squared-error (MSE) loss, optimized via stochastic
1190 gradient descent, for up to 15 epochs, with a batch size of 128, at a learning rate of .0001, with the
1191 Adam optimizer. Specifically, let *L* be the latent space of scRNA-seq encodings. If cell type
1192 annotations were available, a cell type discriminator, *C,* was trained on all encodings by *G* of all
1193 training scRNA-seq profiles. *C* was trained to map an scRNA-seq profile's encoding to its cell
1194 type annotation. Lastly, an adversarial autoencoder *T* was trained on all image tiles of all training

samples' histology tiles, to simultaneously reconstruct image tile, encode to a latent space indistinguishable from *L* via an adversarial training regime of the tile network against an adversarial discriminator, and to encode tiles and scRNA-seq profiles with the same cell type annotation to the same part of the shared latent space.

C was trained according to a binary-cross-entropy loss with-logits (BCE), optimized via stochastic gradient descent for 15 epochs, with a batch size of 32, at a learning rate of .001. The adversarial discriminator was trained according to a binary-cross-entropy loss with-logits (BCE), optimized via stochastic gradient descent for up to 50 epochs, with a batch size of 128, at a learning rate of .004. *T* was trained according to a regularized mean-squared-error loss, given by the loss function $f(i', i, z) = MSE(i', i) + beta*BCE(z, t) + gamma*BCE(y, a)$, where *i* is the input tile's embedding by UNI, *i'* is the reconstruction of *i* by *T, t* are the labels signifying that a latent space comes from the gene-expression latent space, *z* is the output of the adversarial discriminator on the encoding of *i* by *T, y* is the output of *C* on the encoding of *i* by *T,* and *a* is the cell type annotation corresponding to *i*. *L* was trained with *beta*=.001 and *gamma*=5, optimized via stochastic gradient descent, for up to 50 epochs, with a batch size of 128, at a learning rate of .001. The adversarial discriminator and *L* were trained in adversarial fashion with alternating gradient updates.

G and *T* both had an encoder-decoder architecture. *G*'s encoder was of a frozen scGPT model, and the decoder was composed of four sequential [Linear, BatchNorm, ReLU] blocks, followed by one [Linear, ReLU] block. The dimensions of the five linear layers of the decoder were, in order, 512, 2048, 2048, 2048, 2048, and *number_of_genes* neurons, where *number_of_genes* was the number of genes in each dataset being retained for analysis. *T*'s encoder was of a frozen UNI model, followed by four sequential [Linear, BatchNorm, ReLU] blocks, followed by one [Linear, ReLU] block. The dimensions of these five linear layers were, in order, 1024, 1024, 1024, 1024, 512, and

512 neurons. *T*'s decoder was of four sequential [Linear, BatchNorm, ReLU] blocks, followed by one [Linear, ReLU] block. The dimensions of these five linear layers were, in order, 512, 512, 1024, 1024, 1024, and 1024 neurons.

The adversarial discriminator had an architecture consisting of four [Linear, SpectralNorm, ReLU] blocks, followed by one [Linear, SpectralNorm] block. The dimensions of the five linear layers were, in order, 512, 256, 128, 64, 32, and 2 neurons. *C* had an architecture consisting of four [Linear, SpectralNorm, ReLU] blocks, followed by one [Linear, SpectralNorm] block. The dimensions of the five linear layers were, in order, 512, 256, 128, 64, 32, and *num_cell_types* neurons, with *num_cell_types* as the number of cell types common to both H&E and scRNA-seq data. The Linear, BatchNorm, ReLU, and SpectralNorm layers were implemented via pytorch's *nn.Linear, nn.BatchNorm1D, nn.ReLU,* and *nn.utils.spectral_norm* functions, respectively.

To translate from image-tile to a single cell expression profile, after training *G* and *T*, the tile was first encoded with *T's* encoder and then decoded with *G*'s decoder. For a full inference pipeline from H&E-stain to single cell dataset, the input histology image was tiled with the same parameters used for training images, and a single cell expression profile was inferred for each tile.

Train/test dataset splits

In all datasets, for all models, for both Paired and Unpaired SCHAF, training data were partitioned to a training set (90%) and a validation set (10%), and the latter was used for hyperparameter tuning. Hyperparameter tuning was performed in standard fashion: training 10% of the data over a grid search of possible hyperparameters until a metric computed from the validation data was minimized, at which time the corresponding set of hyperparameters would be used to train the final model. For Paired SCHAF, the average gene spatial correlation of the validation data was

1240 minimized. For Unpaired SCHAF, the average gene Earth Mover's Distance was minimized. Ten
1241 percent of the training data, in each case, were used as validation data for hyperparameter tuning
1242 in each training corpus. Hyperparameter values found via tuning included, but were not limited to,
1243 learning rate (swept over a wide range of values from .01 to .0000001), batch size (swept over a
1244 wide range of values from 32 to 256), H&E tile radius (swept over a wide range of values from 16
1245 to 1024 pixels), and the number of hidden layers in each model's decoder (swept from 1 to 5).

1246 **Evaluating SCHAF via cross validation criteria**

1247 To evaluate SCHAF, leave-one-out cross validation was used for each data corpus. For each
1248 fold/sample *s* in the corpus, SCHAF was trained on all other folds/samples and evaluated on *s*.
1249 Each evaluation benchmark was computed for each sample *s*, using the model trained without *s*.
1250 A weighted average was computed across all samples for the final result, with each sample
1251 weighted by the number of cells in its ground truth transcriptomic dataset. Corresponding error
1252 bars represent one standard error measurement above and below the related cross validation result.
1253 In histogram settings, histograms were found by finding histograms for each individual
1254 fold/sample and constructing a new cross validation histogram, with each bar having the height of
1255 the cross validation averaged height of each fold/sample's corresponding bar, with error bars
1256 representing the standard error of the bar height.

1257 **Paired SCHAF evaluation by spatial cell type annotations**

1258 To compare the spatial organization of cell type profiles in measured ExpSCR and Paired SCHAF
1259 inferred data, for each sample in a dataset, a cell type classifier was first trained on labeled (cell
1260 type annotated) profiles and then applied to predict a cell type label for each Paired SCHAF
1261 inferred profile. The classifier had a linear neural network architecture, composed of three

sequential [Linear, BatchNorm, ReLU] blocks, followed by one [Linear, SoftMax] block. The dimensions of the four linear layers were, in order, *num_genes*, 1024, 256, 64, and *num_cell_types* neurons, where *num_genes* is the number of genes predicted by SCHAF, and *num_cell_types* is the number of different cell types in the dataset. The Linear, BatchNorm, ReLU, and SoftMax layers were implemented via pytorch's *nn.Linear, nn.BatchNorm1D, nn.ReLU,* and *nn.Softmax* functions, respectively. The model was trained with a standard cross-entropy loss. Input transcriptomic profiles were z-scored, such that each gene had mean zero and standard deviation 1.

Each ExpSCR and Paired SCHAF inferred profile was plotted spatially and colored by cell type. Accuracy was computed as the proportion SCHAF inferred profiles with a given cell type annotation that matched the annotation of the ExpSCR profile.

Expression profile heterogeneity of cells of each type was computed for cells with each cell type annotation in a sample/fold, in measured (ExpSCR) or Paired SCHAF inferred data. In each case, the standard deviation was computed for the min-max normalized vector of the expression of each gene g, and then averaged over all genes as the heterogeneity measure.

Paired SCHAF evaluation by spatial correlation

For each ExpSCR measured gene, the correlation of measured (ExpSCR) and Paired SCHAF inferred values was computed with the *numpy.corrcoef(a, b)* function in the python *numpy* library, where a and b were vectors of all SCHAF inferred and ExpSCR measured values. For visualization of individual genes, normalized expression/score were computed by z-scoring the expression vector x of gene g to a mean 0 and standard deviation 1, followed by min-max normalization to minimum of 0 and maximum of 1. The line of best fit was computed between normalized ExpSCR

1284 value and normalized SCHAF inferred values with x-coordinates *numpy.unique(ps)* and y-
1285 coordinates *numpy.poly1d(numpy.polyfit(ps, ts, 1))(numpy.unique(ps)),* with python's *numpy*
1286 library, where *ps* and *ts* are SCHAF inferred and ExpSCR measured values, respectively.

1287 **Paired SCHAF evaluation by quality scores**

1288 To find the quality score of a gene *g*, the standard deviation was computed for the expression
1289 vector of *g*, over all inferred profiles. For genes measured by ExpSCR, the correlation between a
1290 gene's quality score and its spatial correlation value was computed with the *numpy* function
1291 *numpy.corrcoef*, as well the line of best fit between these two using *numpy.poly1d*.

1292 Low-resolution SCHAF Inferred data were generated as follows. First, the ExpLR spot containing
1293 each ExpSCR inferred profile was identified (if available), and all ExpSCR values in one ExpLR
1294 spot were aggregated at spot level and normalized. Spatial correlations between low-resolution
1295 SCHAF inferred data and ExpLR measured data were computed for each ExpLR-measured gene
1296 expression (line of best fit, as above) and visualized spatially.

1297 Similar analysis was performed for each of the gene programs with >90% genes overlapping with
1298 genes inferred by SCHAF, comparing program scores at low-resolution between ExpLR and
1299 SCHAF inferences. Following previous studies[54], the expression score of program *p* was the mean
1300 of z-scored expression of program genes. Similar to individual genes, the quality score of a
1301 program *p* was found as the standard deviation, over all inferred cells, of the expression score of
1302 program *p*.

1303 The eight genes measured by ISH in the new pup placenta sample were randomly selected from
1304 the genes that were not measured by ExpSCR and that had quality scores that varied highly across

1305 folds (from <0.2 to nearly 2.0). A Paired SCHAF model was then trained on the entire In-Sample
1306 Mouse data and used to infer the new mouse sample's ISH.

1307 **Evaluation of overall properties of RNA profiles inferred by Unpaired SCHAF**

1308 'Pseudobulk' RNA profiles were calculated for each sample analyzed by Unpaired SCHAF, as the
1309 mean expression profiles of either measured or Unpaired SCHAF-inferred scRNA-seq profiles,
1310 and Pearson's correlation coefficient between them was calculated.

1311 For each sample analyzed by Unpaired SCHAF, the 100 most highly variable genes were found in
1312 the measured scRNA-seq via *scanpy*'s function *scanpy.pp.highly_variable_genes*, and gene-gene
1313 correlation matrices were computed between measured and inferred profiles by converting each to
1314 a pandas dataframe and then using the .corr function. Each entry in the matrix was then assigned
1315 an integral rank by magnitude, and the matrix was converted into a numpy array using the *scipy*
1316 python library's *stats.rankdata* function. Meta correlation was computed as the Pearson's
1317 correlation between corresponding flattened entries of the measured and SCHAF inferred numpy
1318 arrays.

1319 Per-cell-type 'pseudobulk' RNA-Seq profiles and their correlations were computed as for
1320 pseudobulk above, but separately for all cells of each cell type *c*. In addition, a weighted average
1321 of Pearson's r was calculated, weighted by cell type prevalence, as the "cell type pseudobulk
1322 correlation score."

1323 **Unpaired SCHAF evaluation by gene count distributions**

1324 For each gene *g,* the Earth Mover's Distance (Wasserstein Distance) between the distributions of
1325 SCHAF inferred and measured scRNA-seq values (over all profiles) was computed with the

function *scipy.stats.wasserstein_distance(a, b)* from the *scipy.stats* python library, where *a* and *b* are the vectors of all counts of gene *g* from all SCHAF inferred and measured profiles, respectively. *a* was un-normalized after prediction (multiplied by the standard deviation from training and added with the mean in training). Each distribution was visualized using the *pandas.plot.kde* function from the python *pandas* library, with default parameters except for *bw_method*, which was set to *n**(-.2)*, where *n* was the number of profiles in the distribution.

Unpaired SCHAF evaluation by distribution matching scores

The Unpaired SCHAF inferred expression vector of gene *g* was un-normalized (multiplied by the standard deviation from training and added with the mean in training), followed by subtracting half of its standard deviation from 1. This value served as a gene's distribution matching score. For genes measured by scRNA-seq, the correlation between a gene's distribution matching score and its Earth Mover's Distance value was found, plotting the line of best fit (found with the *numpy* function *numpy.poly1d*) between the two, in addition to a scatterplot visualizing the two.

Unpaired SCHAF evaluation by UMAPs

For each dataset *t* analyzed by Unpaired SCHAF, UMAP embeddings were generated from the measured scRNA-Seq by the following series of *scanpy* functions: *scanpy.pp.highly_variable_genes(t, n_top_genes=4096), scanpy.pp.pca(t, use_highly_variable=1), scanpy.pp.neighbors(t), scanpy.tl.umap(t)*. Next, *scanpy.tl.ingest(p, t, embedding_method='umap')* was used to find the embedding of the corresponding Unpaired SCHAF data *p* in the same space as the UMAP of the scRNA-seq measured data *t*. UMAPs were then visualized using the *scanpy.pl.umap* function, coloring by cell type, where Unpaired SCHAF

1347 inferred profiles were assigned to cell types by the same classifier-based method used to evaluate
1348 Paired SCHAF.

1349 UMAP organization was quantified using a local coherence score, defined as the average, over all
1350 cells, of the proportion of each cell's *k*=10 nearest neighbors in UMAP space that are of its type.
1351 Cluster separation was quantized using a cluster separation score, defined as the ratio of an inter-
1352 cluster to intra-cluster distance scores. Inter-cluster distance score was defined as the average, over
1353 all cells, of each cell's average distance to cells of its type among its k=15 nearest neighbors. The
1354 intra-cluster distance is defined as the average, over all cells, of the cell's average distance to cells
1355 of different types among its k=15 nearest neighbors.

1356 **Unpaired SCHAF evaluation by cell type distributions**

1357 Cell types were assigned to Unpaired SCHAF inferred profiles using the same classifier approach
1358 as for Paired SCHAF.

1359 Jensen-Shannon Divergence between Unpaired SCHAF inferred and measured scRNA-seq
1360 distributions for each sample's cell type were computed using the
1361 *scipy.spatial.distance.jennsenshannon(a, b)* function from the *scipy.spatial* python library, where
1362 *a* and *b* are the vectors of cell type proportions in SCHAF inferred and measured scRNA-seq data,
1363 respectively.

1364 Expression heterogeneity in Unpaired SCHAF inferred and measured scRNA-seq were computed
1365 as for Paired SCHAF, but for Unpaired SCHAF inferences, only cells annotated by the classifier
1366 for a given cell type were considered as belonging to that cell type.

1367 **Aggregation of evaluation metrics for Unpaired SCHAF**

For each dataset, heatmaps were created to show the value of many Unpaired SCHAF evaluation metrics for each sample. Within each dataset, for each metric, the normalized metric value was computed by z-scoring the vector of all samples' metric values to a mean 0 and standard deviation 1, followed by min-max normalization to minimum of 0 and maximum of 1. This normalized metric value was then shown for each (sample, metric) pair in the dataset, with samples being noted on the x-axis and metrics on the y-axis.

Comparison of Unpaired SCHAF to expert pathologist annotations, STARmap or MERFISH

To visually compare to expert pathologist annotations, SCHAF profiles were colored by annotations and regions with two annotations were colored such that a random half of the region's points had one color and the other half had the other.

H&E image analysis was performed to compare cell type assignments to imaging features. For each H&E image and cell, both the mean pixel intensity and the number of cells were computed from the cell's H&E tile. The number of cells in each tile was found by placing the coordinates of all cells in each fold/sample in a KDTree (from the python *scipy.spatial* library) and querying the tree for all points that would fall in a tile centered around the cell in question.

The Unpaired SCHAF and STARmap (placenta) or MERFISH (HTAPP MBC) measured spatial expression value of marker genes was also compared for marker genes.

Spatial accuracy of Unpaired SCHAF predictions

Adjacent slices of H&E and STARmap/MERFISH in parts of the Placenta/HTAPP data were manually aligned. When there was no one-to-one correspondence between individual cells (two

separate slices), low-resolution spots (~five cells / spot) were generated for both inferred (Unpaired SCHAF) and measured (STARmap/MERFISH) data, with average expression of the gene profiles of individual cells in the spot. Spots were assigned a cell type as the most common cell type of the individual cells in the spot. When there was one-to-one spot level alignment, the proportion of correctly annotated cells in each sample was quantified with accuracy and confusion matrix measures, and individual gene correlation measurements across all spots were calculated.

Ablating cell type information in Unpaired SCHAF training

To assess the importance of cell type information in Unpaired SCHAF, versions of Unpaired SCHAF that did not include the annotation classifier *C* were trained, and their inferred spatial cell type maps were compared to ground truth pathologist annotations as described above.

SCHAF evaluation against benchmark methods

Paired and Unpaired SCHAF were compared to two sets of benchmark methods.

Paired SCHAF was compared to methods that predict expression profiles from H&E in supervised spatial fashion: SPiRiT[35], ST-Net[31], DeepPT[36], and HE2RNA[30]. Each model was implemented according to the specifications in its study's Methods section, trained in the same cross validation scheme as SCHAF with hyperparameter tuning, and compared by average spatial correlation and the sample/fold cell type accuracy.

Unpaired SCHAF was compared to Cross-modal Autoencoders[55], CycleGAN[56], and DistanceGAN[57]. Cross-modal Autoencoders was implemented according to the image-to-omics setting described in its study's Methods section[55]. The generator architecture for CycleGAN and DistanceGAN mirrored the final inference model used in Cross-modal Autoencoders, with the

same discriminator architecture for all benchmarks as used in the Unpaired SCHAF model. Methods were evaluated by the gene Earth Mover's Distance, fold/sample cell type distribution Jensen-Shannon Divergence, sample bulk RNA-seq correlation, sample meta correlation score, and sample cell type pseudobulk score, all as defined as described above for Unpaired SCHAF.

Code availability

Code for comprehensive analysis as described here, and figure generation as shown here can be found at https://github.com/ccomiter1/SCHAF.

Data availability

All unpublished data, in addition to all other publicly available/published data, have been organized into a single repository named "SCHAF-1M" at the Zenodo link listed on https://github.com/ccomiter1/SCHAF. HTAPP data are part of the HTAN-HTAPP data release on Synapse (syn20834712) as referenced in its own study[39]. MSKCC will be formally released in its own study (MC, leader of that study, can be reached at metamiaciampricotti@gmail.com). Until its publication, its data can be obtained from the "SCHAF-1M" database at the Zenodo link listed on https://github.com/ccomiter1/SCHAF. Placenta transcriptomic data can be found in its original publication[40], and corresponding H&E images and annotations from the "SCHAF-1M" database at the Zenodo link listed on https://github.com/ccomiter1/SCHAF. Publicly available ExpSCR (Xenium), ExpLR (Visium), and H&E data are at https://www.10xgenomics.com/datasets. New mouse pup ExpSCR (ISH) and H&E data can be found in the "SCHAF-1M" databased at the Zenodo link listed on https://github.com/ccomiter1/SCHAF.

Author Contributions

CC, EDV, and AR conceived and designed the study. CC designed and executed the study with guidance from JS, CMS, and AR. MC, BL, YY, AK, AQV, and CMR mined and provided SCLC tissue data. JJ, MS, JWu, SV, JWa, TRB, AS, DLA, NW, XZ, and JK mined and provided the HTAPP/MBC tissue data. SJR, MT, and KLP provided expert pathological annotations of HTAPP/MBC tissue data. KZ, JQ, RT, FM, SEB, and SD mined and provided both new mouse data and placenta H&E data in addition to helping process publicly-available ExpSCR (Xenium) data. JM provided expert pathologist annotations of placenta data. XC assisted with software engineering and figure preparation. XC, EDV, CMS, KJKK, and JS provided methodological suggestions. JS, CMS, and AR provided feedback and guidance. CC, JS, CMS, and AR wrote the manuscript, with input from all authors.

Conflict of interest

AR is a co-founder and was equity holder of Celsius Therapeutics, is an equity holder in Immunitas, and until July 31, 2020 was an SAB member of ThermoFisher Scientific, Syros Pharmaceuticals, Neogene Therapeutics, and Asimov. From August 1, 2020, AR is an employee of Genentech and has equity in Roche. JS is a scientific advisor for Johnson & Johnson. A patent application has been filed related to this work through Broad Institute. CMR has consulted regarding oncology drug development with AbbVie, Amgen, AstraZeneca, Daiichi Sankyo, Genentech/Roche, Jazz, and Novartis. He serves on the scientific advisory boards of Auron, DISCO, and Earli and has received licensing fees and royalties related to DLL3-directed therapeutics. EDV is the founder & CEO of Sequome, Inc. BL, YY, ORR, and NW are employees of Genentech as of 9 August 2021, 7 June 2021,19 October 2020, and Feb 13, 2023, respectively and have equity in Roche. SJR receives research support from Bristol-Myers-Squibb and KITE/Gilead. SJR is a member of the SAB of Immunitas Therapeutics. NW is an equity holder in

Relay Therapeutics and a consultant and equity holder in Flare Therapeutic. Prior to Jan 31 2023 he was an SAB member of Relay Therapeutics, an advisory board member for Eli Lilly, and received research support from Astra-Zeneca and Puma Biotechnologies. MC is an employee of Tempus AI as of 2 October 2023. AQV is an employee of AstraZeneca as of 17 July 2024.

Special Acknowledgement

This work is dedicated to the memory of Minh-Thi Nguyen, the best friend of CC, who, at 24 years of age, on the morning of June 21, 2024, following a horrific bicycle collision, tragically lost her life. This work would not have been possible without her years of unparalleled support, encouragement, camaraderie, care, humor, and loving friendship, for which CC is and will be forever grateful.

Other Acknowledgements

We thank Ania Hupalowska and Leslie Gaffney for help with graphics and members of the Shu, Stultz, and Regev labs for helpful discussions. Work was supported by the Klarman Cell Observatory and Howard Hughes Medical Institute (AR), NIH New Innovator Award (DP2TR004354, JS), NIH Common Fund (UH3CA275687, UG3CA275687, JS), NICHD (R00HD096049), Massachusetts Life Science Center (JS), Burroughs Wellcome Fund (JS), Additional Ventures (JS), and Massachusetts General Hospital (JS). This project was also funded in part by federal funds from NCI, NIH Task Order HHSN261100039 under Contract HHSN261201500003. CC was supported by a Seibel Scholars Fellowship in addition to a MIT Department of Electrical Engineering and Computer Science Fellowship, Teaching Assistantship, and Research Assistantship. CMR is supported by NCI R35 CA263816 and U24 CA213274. JM is supported by NIH training grant 5T32CA193145. JK was supported by an

EMBO Long-Term Fellowship (ALTF 738-2017), an HFSP long-term fellowship (LT000452/2019-L) and an EKFS starting grant (2019_A70).

1479 **References**

1480 1. Tang, X., Huang, Y., Lei, J., Luo, H. & Zhu, X. The single-cell sequencing: new
1481 developments and medical applications. *Cell Biosci.* **9**, 53 (2019).
1482 2. Rood, J. E., Maartens, A., Hupalowska, A., Teichmann, S. A. & Regev, A. Impact of the
1483 Human Cell Atlas on medicine. *Nat. Med.* **28**, 2486–2496 (2022).
1484 3. Yan, L. C., Yoshua, B. & Geoffrey, H. Deep learning. *Nature* **521**, 436–444 (2015).
1485 4. Regev, A. *et al.* The Human Cell Atlas. *Elife* **6**, (2017).
1486 5. Rozenblatt-Rosen, O., Stubbington, M. J. T., Regev, A. & Teichmann, S. A. The Human
1487 Cell Atlas: from vision to reality. *Nature* **550**, 451–453 (2017).
1488 6. Regev, A. *et al.* The Human Cell Atlas White Paper. *arXiv [q-bio.TO]* (2018).
1489 7. Biancalani, T. *et al.* Deep learning and alignment of spatially-resolved whole transcriptomes
1490 of single cells in the mouse brain with Tangram. *bioRxiv* 2020.08.29.272831 (2020)
1491 doi:10.1101/2020.08.29.272831.
1492 8. Ding, J. & Regev, A. Deep generative model embedding of single-cell RNA-Seq profiles on
1493 hyperspheres and hyperbolic spaces. *bioRxiv* (2019) doi:10.1101/853457.
1494 9. van der Laak, J., Litjens, G. & Ciompi, F. Deep learning in histopathology: the path to the
1495 clinic. *Nat. Med.* **27**, 775–784 (2021).
1496 10. Ziegenhain, C. *et al.* Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol.*
1497 *Cell* **65**, 631–643.e4 (2017).
1498 11. Lähnemann, D. *et al.* Eleven grand challenges in single-cell data science. *Genome Biol.* **21**,
1499 31 (2020).
1500 12. Luecken, M. D. *et al.* Benchmarking atlas-level data integration in single-cell genomics.
1501 *bioRxiv* 2020.05.22.111161 (2020) doi:10.1101/2020.05.22.111161.

1502 13. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. RNA imaging.
1503 Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090
1504 (2015).
1505 14. Eng, C.-H. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA
1506 seqFISH. *Nature* **568**, 235–239 (2019).
1507 15. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial
1508 transcriptomics. *Science* **353**, 78–82 (2016).
1509 16. Janesick, A. *et al.* High resolution mapping of the breast cancer tumor microenvironment
1510 using integrated single cell, spatial and in situ analysis of FFPE tissue. *bioRxiv* (2022)
1511 doi:10.1101/2022.10.06.510405.
1512 17. Ke, R. *et al.* In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods*
1513 **10**, 857–860 (2013).
1514 18. Chen, X., Sun, Y.-C., Church, G. M., Lee, J. H. & Zador, A. M. Efficient in situ barcode
1515 sequencing using padlock probe-based BaristaSeq. *Nucleic Acids Res.* **46**, e22 (2018).
1516 19. Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by cyclic
1517 smFISH. *bioRxiv* (2018) doi:10.1101/276097.
1518 20. Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by osmFISH.
1519 *Nat. Methods* **15**, 932–935 (2018).
1520 21. Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional
1521 states. *Science* **361**, (2018).
1522 22. Russell, A. J. C. *et al.* Slide-tags enables single-nucleus barcoding for multimodal spatial
1523 genomics. *Nature* **625**, 101–109 (2024).
1524 23. Alon, S. *et al.* Expansion Sequencing: Spatially Precise In Situ Transcriptomics in Intact

Biological Systems. *bioRxiv* 2020.05.13.094268 (2020) doi:10.1101/2020.05.13.094268.

24. Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).

25. Biancalani, T. *et al.* Deep learning and alignment of spatially resolved single-cell transcriptomes with Tangram. *Nat Methods* **18**, 1352–1362 (2021).

26. Achim, K. *et al.* High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat. Biotechnol.* **33**, 503–509 (2015).

27. Zhu, H., Luo, M.-D., Wang, R., Zheng, A.-H. & He, R. Deep Audio-visual Learning: A Survey. *Int. J. Autom. Comput.* **18**, 351–376 (2021).

28. Kumar, L. A., Renuka, D. K., Rose, S. L., Shunmuga priya, M. C. & Wartana, I. M. Deep learning based assistive technology on audio visual speech recognition for hearing impaired. *International Journal of Cognitive Computing in Engineering* **3**, 24–30 (2022).

29. Ngiam, J. *et al.* Multimodal Deep Learning. (2011).

30. Schmauch, B. *et al.* A deep learning model to predict RNA-Seq expression of tumours from whole slide images. *Nat. Commun.* **11**, 3877 (2020).

31. He, B. *et al.* Integrating spatial gene expression and breast tumour morphology via deep learning. *Nat Biomed Eng* **4**, 827–834 (2020).

32. Chen, R. J. *et al.* Towards a general-purpose foundation model for computational pathology. *Nat. Med.* **30**, 850–862 (2024).

33. Cui, H. *et al.* scGPT: toward building a foundation model for single-cell multi-omics using generative AI. *Nat. Methods* **21**, 1470–1480 (2024).

34. Qiu, C. *et al.* A single-cell time-lapse of mouse prenatal development from gastrula to birth. *Nature* **626**, 1084–1093 (2024).

35. Zhao, Y. *et al.* Inferring single-cell spatial gene expression with tissue morphology via explainable deep learning. *bioRxivorg* (2024) doi:10.1101/2024.06.12.598686.

36. Hoang, D.-T. *et al.* A deep-learning framework to predict cancer treatment response from histopathology images through imputed transcriptomics. *Nat Cancer* (2024) doi:10.1038/s43018-024-00793-2.

37. 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 (2005).

38. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).

39. Klughammer, J. *et al.* A multi-modal single-cell and spatial expression map of metastatic breast cancer biopsies across clinicopathological features. *Nat. Med.* **30**, 3236–3249 (2024).

40. Ounadjela, J. R. *et al.* Spatial multiomic landscape of the human placenta at molecular resolution. *Nat. Med.* **30**, 3495–3508 (2024).

41. Hicks, S. C., Townes, F. W., Teng, M. & Irizarry, R. A. Missing data and technical variability in single-cell RNA-sequencing experiments. *Biostatistics* **19**, 562–578 (2018).

42. Rozenblatt-Rosen, O. *et al.* The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. *Cell* **181**, 236–249 (2020).

43. HuBMAP Consortium. The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature* **574**, 187–192 (2019).

44. Istrate, A.-M., Li, D. & Karaletsos, T. scGenePT: Is language all you need for modeling single-cell perturbations? *bioRxiv* (2024) doi:10.1101/2024.10.23.619972.

45. Kobayashi-Kirschvink, K. J. *et al.* Raman2RNA: Live-cell label-free prediction of single-

cell RNA expression profiles by Raman microscopy. *bioRxiv* 2021.11.30.470655 (2022) doi:10.1101/2021.11.30.470655.

46. Schofield, J. A., Duffy, E. E., Kiefer, L., Sullivan, M. C. & Simon, M. D. TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding. *Nat. Methods* **15**, 221–225 (2018).

47. Weigert, M. & Schmidt, U. Nuclei Instance Segmentation and Classification in Histopathology Images with Stardist. in *2022 IEEE International Symposium on Biomedical Imaging Challenges (ISBIC)* (IEEE, 2022). doi:10.1109/isbic56247.2022.9854534.

48. Slyper, M. *et al.* A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat. Med.* **26**, 792–802 (2020).

49. Creating a Reference Package with cellranger mkref -Software -Single Cell Gene Expression -Official 10x Genomics Support. https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/advanced/references.

50. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

51. Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. in *Advances in Neural Information Processing Systems* (eds. Wallach, H. et al.) vol. 32 (Curran Associates, Inc., 2019).

52. Kingma & Ba. Adam: A method for stochastic gradient descent. *ICLR: International Conference on Learning*.

53. Makhzani, A., Shlens, J., Jaitly, N., Goodfellow, I. & Frey, B. Adversarial Autoencoders. *arXiv [cs.LG]* (2015).

54. Schiebinger, G. *et al.* Optimal-transport analysis of single-cell gene expression identifies

developmental trajectories in reprogramming. *Cell* **176**, 1517 (2019).

55. Yang, K. D. *et al.* Multi-domain translation between single-cell imaging and sequencing data using autoencoders. *Nat. Commun.* **12**, 31 (2021).

56. Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. in *2017 IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2017). doi:10.1109/iccv.2017.244.

57. Tran, N.-T., Bui, T.-A. & Cheung, N.-M. Dist-GAN: An improved GAN using distance constraints. in *Computer Vision – ECCV 2018* 387–401 (Springer International Publishing, Cham, 2018).

Figure 1

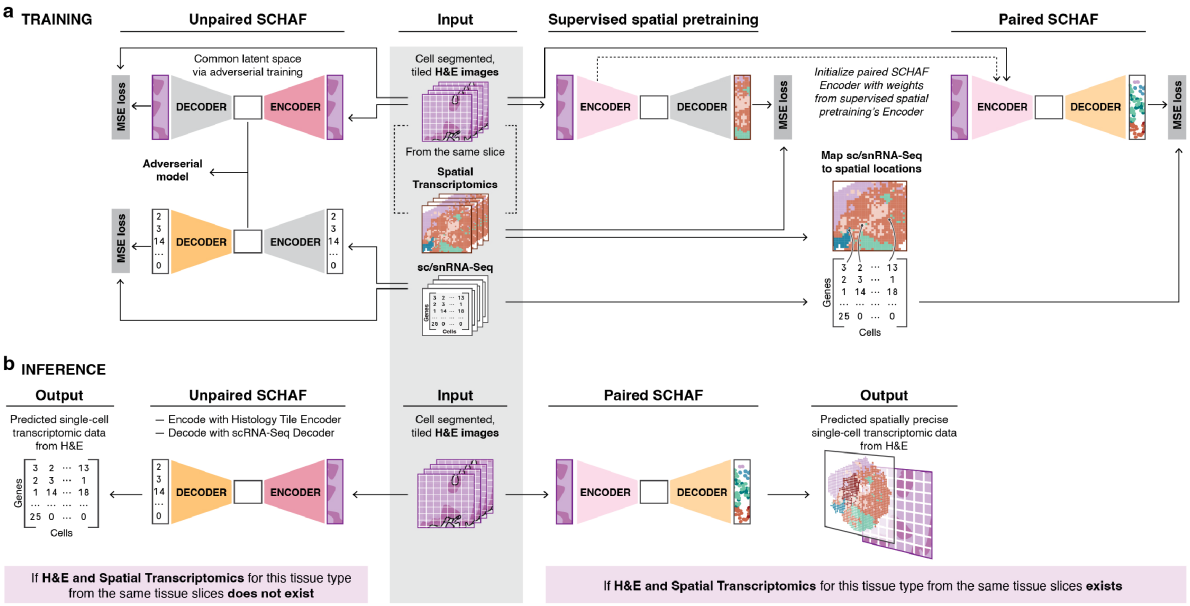

a TRAINING
Unpaired SCHAF
Input
Supervised spatial pretraining
Paired SCHAF
Common latent space via adverserial training
MSE loss
DECODER
ENCODER
Adverserial model
Cell segmented, tiled H&E images
From the same slice
Spatial Transcriptomics
sc/snRNA-Seq
Initialize paired SCHAF Encoder with weights from supervised spatial pretraining's Encoder
Map sc/snRNA-Seq to spatial locations
Genes
Cells
b INFERENCE
Output
Predicted single-cell transcriptomic data from H&E
Unpaired SCHAF
— Encode with Histology Tile Encoder
— Decode with scRNA-Seq Decoder
Input
Cell segmented, tiled H&E images
Paired SCHAF
Output
Predicted spatially precise single-cell transcriptomic data from H&E
If H&E and Spatial Transcriptomics for this tissue type from the same tissue slices does not exist
If H&E and Spatial Transcriptomics for this tissue type from the same tissue slices exists

Figure 2

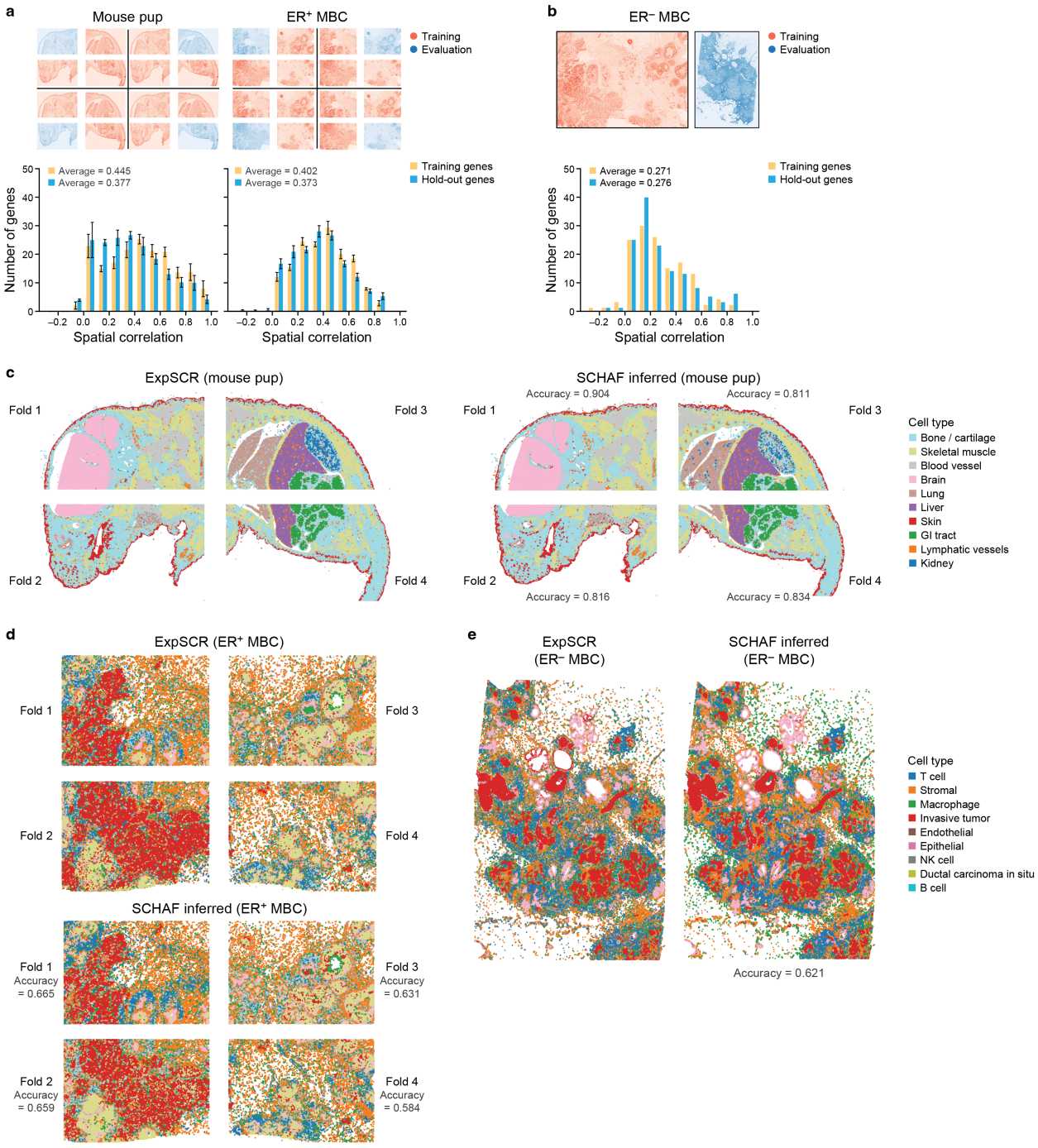

a
Mouse pup
ER+ MBC
Training
Evaluation
Average = 0.445
Average = 0.377
Average = 0.402
Average = 0.373
Training genes
Hold-out genes
Number of genes
Spatial correlation
b
ER− MBC
Average = 0.271
Average = 0.276
c
ExpSCR (mouse pup)
SCHAF inferred (mouse pup)
Fold 1
Fold 2
Fold 3
Fold 4
Accuracy = 0.904
Accuracy = 0.811
Accuracy = 0.816
Accuracy = 0.834
Cell type
Bone / cartilage
Skeletal muscle
Blood vessel
Brain
Lung
Liver
Skin
GI tract
Lymphatic vessels
Kidney
d
ExpSCR (ER+ MBC)
SCHAF inferred (ER+ MBC)
Fold 1 Accuracy = 0.665
Fold 3 Accuracy = 0.631
Fold 2 Accuracy = 0.659
Fold 4 Accuracy = 0.584
e
ExpSCR (ER− MBC)
SCHAF inferred (ER− MBC)
Accuracy = 0.621
Cell type
T cell
Stromal
Macrophage
Invasive tumor
Endothelial
Epithelial
NK cell
Ductal carcinoma in situ
B cell

Figure 3

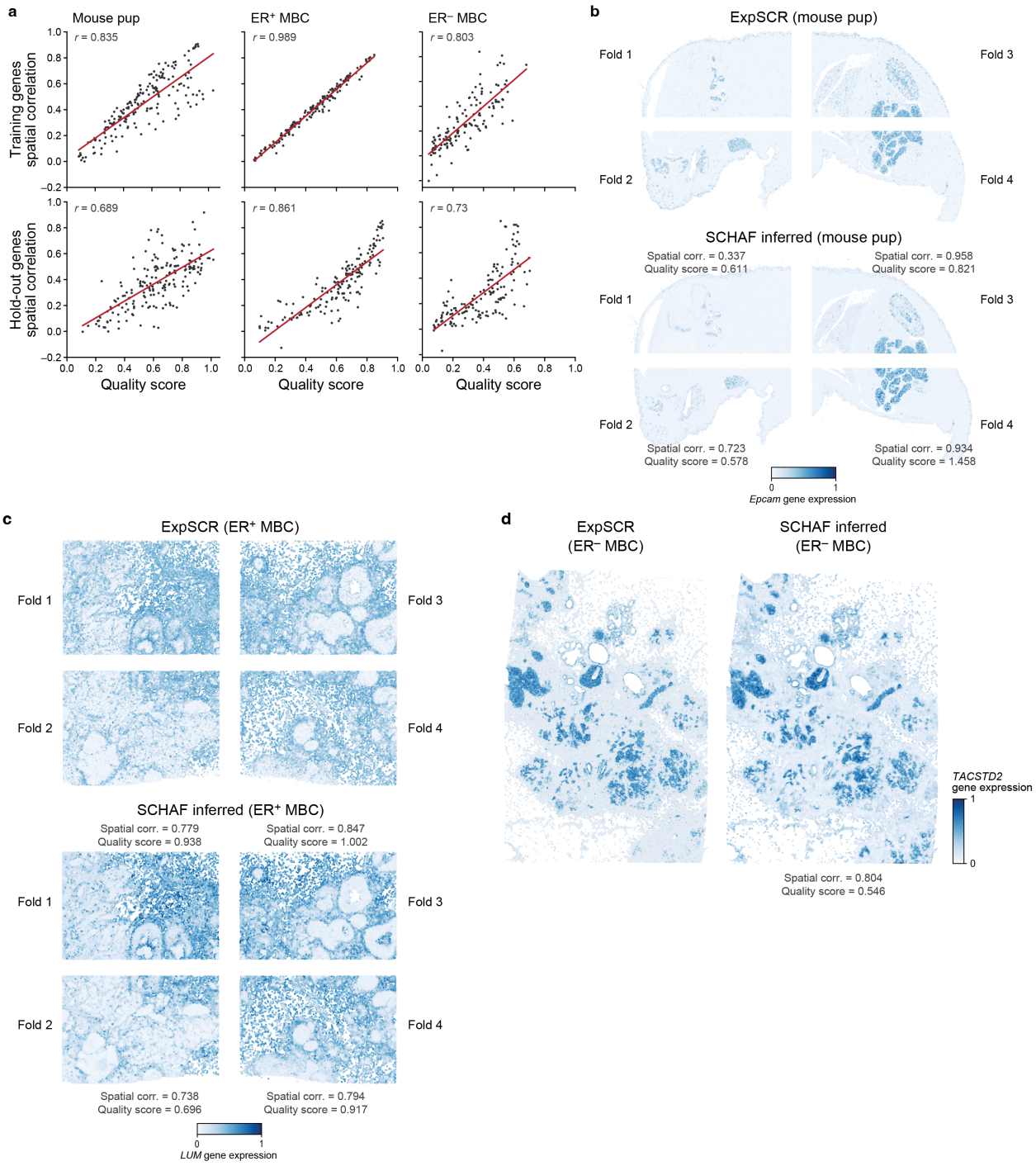

a
Mouse pup
ER+ MBC
ER− MBC
r = 0.835
r = 0.989
r = 0.803
r = 0.689
r = 0.861
r = 0.73
Training genes spatial correlation
Hold-out genes spatial correlation
Quality score
b
ExpSCR (mouse pup)
Fold 1
Fold 2
Fold 3
Fold 4
SCHAF inferred (mouse pup)
Spatial corr. = 0.337
Quality score = 0.611
Spatial corr. = 0.958
Quality score = 0.821
Spatial corr. = 0.723
Quality score = 0.578
Spatial corr. = 0.934
Quality score = 1.458
Epcam gene expression
c
ExpSCR (ER+ MBC)
SCHAF inferred (ER+ MBC)
Spatial corr. = 0.779
Quality score = 0.938
Spatial corr. = 0.847
Quality score = 1.002
Spatial corr. = 0.738
Quality score = 0.696
Spatial corr. = 0.794
Quality score = 0.917
LUM gene expression
d
ExpSCR (ER− MBC)
SCHAF inferred (ER− MBC)
TACSTD2 gene expression
Spatial corr. = 0.804
Quality score = 0.546

Figure 4

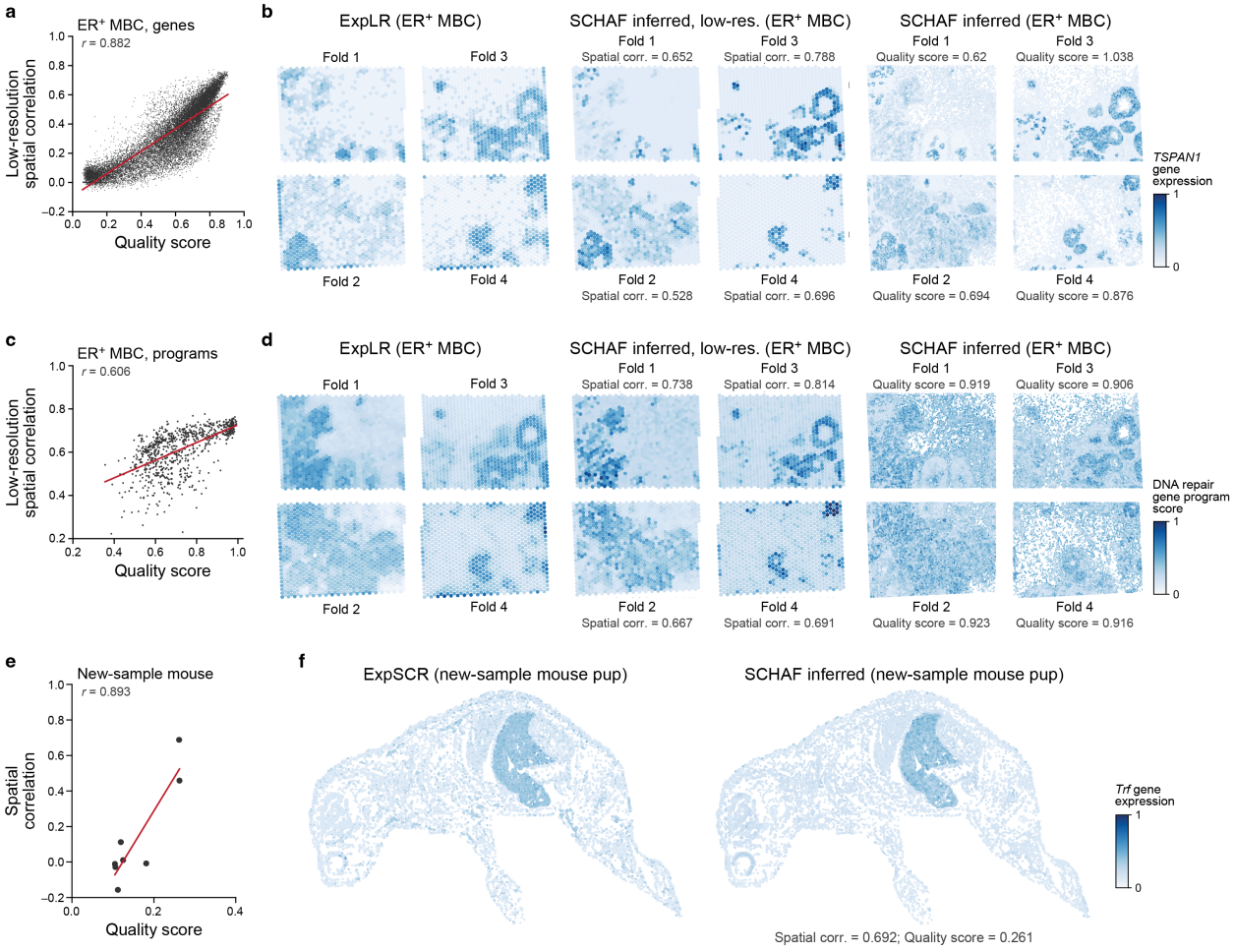

a
ER+ MBC, genes
r = 0.882
Low-resolution spatial correlation
Quality score
b
ExpLR (ER+ MBC)
SCHAF inferred, low-res. (ER+ MBC)
SCHAF inferred (ER+ MBC)
Fold 1
Fold 3
Fold 2
Fold 4
Spatial corr. = 0.652
Spatial corr. = 0.788
Spatial corr. = 0.528
Spatial corr. = 0.696
Quality score = 0.62
Quality score = 1.038
Quality score = 0.694
Quality score = 0.876
TSPAN1 gene expression
c
ER+ MBC, programs
r = 0.606
Low-resolution spatial correlation
Quality score
d
ExpLR (ER+ MBC)
SCHAF inferred, low-res. (ER+ MBC)
SCHAF inferred (ER+ MBC)
Spatial corr. = 0.738
Spatial corr. = 0.814
Spatial corr. = 0.667
Spatial corr. = 0.691
Quality score = 0.919
Quality score = 0.906
Quality score = 0.923
Quality score = 0.916
DNA repair gene program score
e
New-sample mouse
r = 0.893
Spatial correlation
Quality score
f
ExpSCR (new-sample mouse pup)
SCHAF inferred (new-sample mouse pup)
Trf gene expression
Spatial corr. = 0.692; Quality score = 0.261

Figure 5

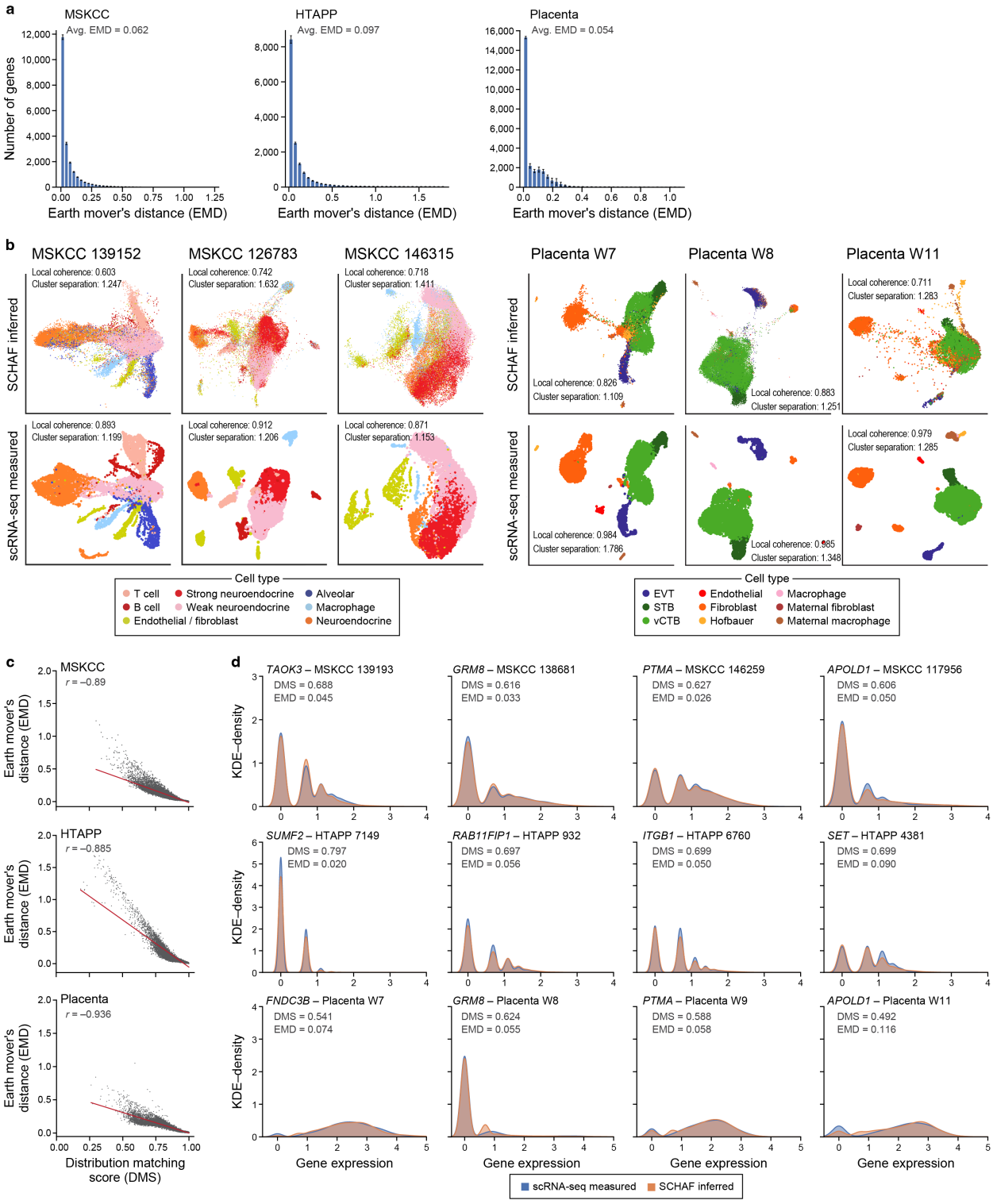
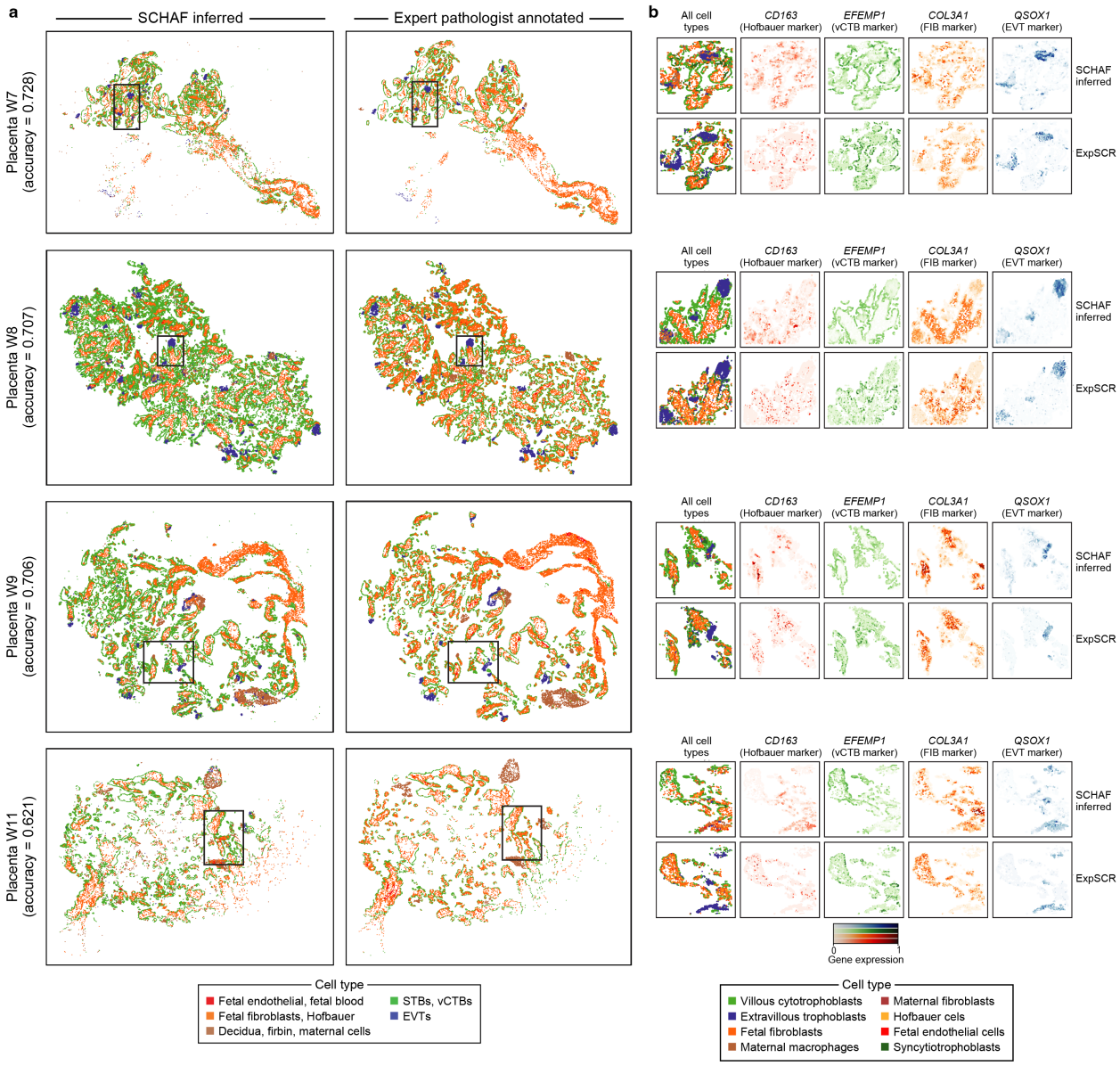
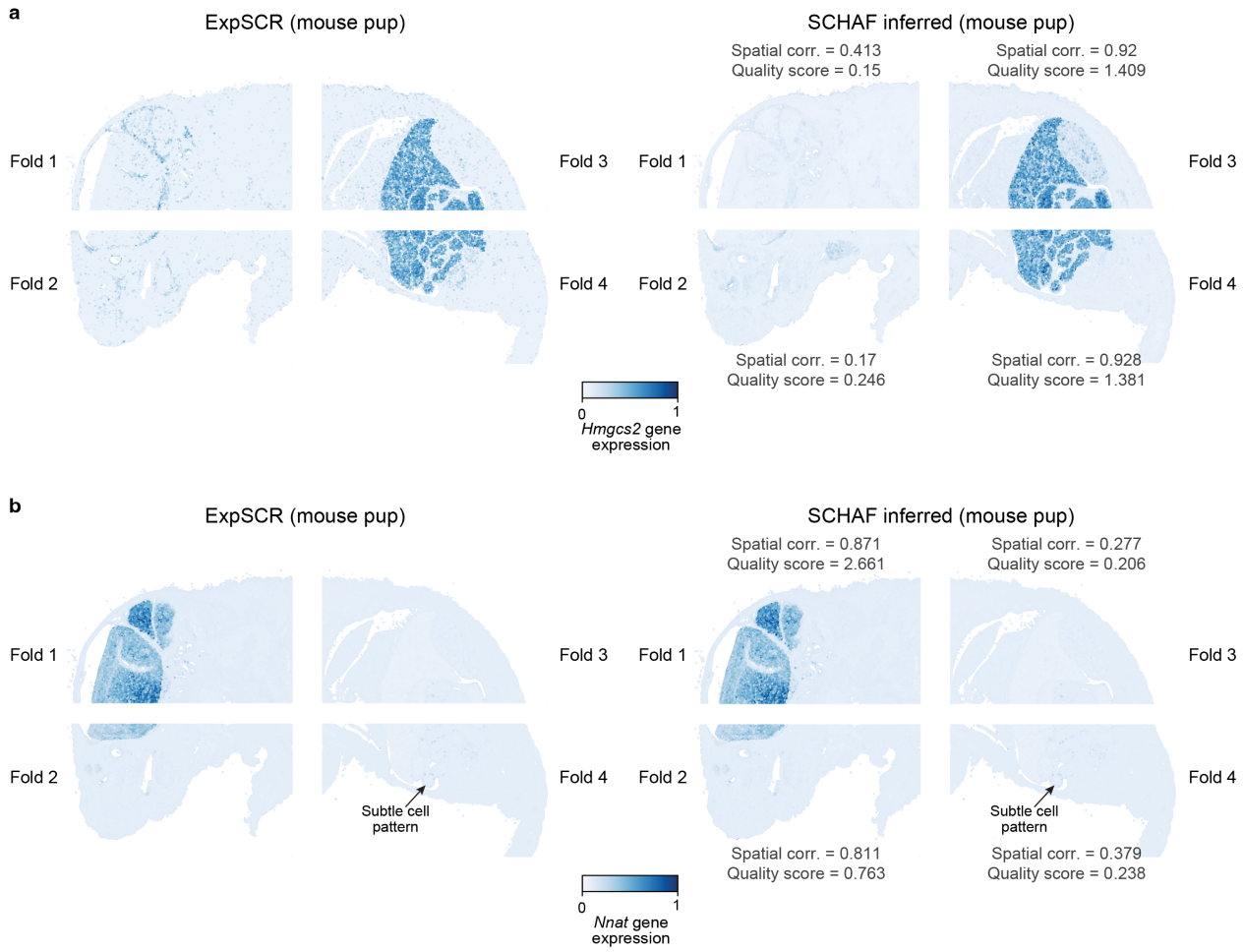

a
MSKCC
Avg. EMD = 0.062
HTAPP
Avg. EMD = 0.097
Placenta
Avg. EMD = 0.054
Number of genes
Earth mover's distance (EMD)
b
MSKCC 139152
MSKCC 126783
MSKCC 146315
Placenta W7
Placenta W8
Placenta W11
SCHAF inferred
scRNA-seq measured
Local coherence: 0.603
Cluster separation: 1.247
Local coherence: 0.742
Cluster separation: 1.632
Local coherence: 0.718
Cluster separation: 1.411
Local coherence: 0.826
Cluster separation: 1.109
Local coherence: 0.883
Cluster separation: 1.251
Local coherence: 0.711
Cluster separation: 1.283
Local coherence: 0.893
Cluster separation: 1.199
Local coherence: 0.912
Cluster separation: 1.206
Local coherence: 0.871
Cluster separation: 1.153
Local coherence: 0.984
Cluster separation: 1.786
Local coherence: 0.985
Cluster separation: 1.348
Local coherence: 0.979
Cluster separation: 1.285
Cell type
T cell
B cell
Endothelial / fibroblast
Strong neuroendocrine
Weak neuroendocrine
Alveolar
Macrophage
Neuroendocrine
Cell type
EVT
STB
vCTB
Endothelial
Fibroblast
Hofbauer
Macrophage
Maternal fibroblast
Maternal macrophage
c
MSKCC
r = −0.89
HTAPP
r = −0.885
Placenta
r = −0.936
Earth mover's distance (EMD)
Distribution matching score (DMS)
d
TAOK3 – MSKCC 139193
DMS = 0.688
EMD = 0.045
GRM8 – MSKCC 138681
DMS = 0.616
EMD = 0.033
PTMA – MSKCC 146259
DMS = 0.627
EMD = 0.026
APOLD1 – MSKCC 117956
DMS = 0.606
EMD = 0.050
SUMF2 – HTAPP 7149
DMS = 0.797
EMD = 0.020
RAB11FIP1 – HTAPP 932
DMS = 0.697
EMD = 0.056
ITGB1 – HTAPP 6760
DMS = 0.699
EMD = 0.050
SET – HTAPP 4381
DMS = 0.699
EMD = 0.090
FNDC3B – Placenta W7
DMS = 0.541
EMD = 0.074
GRM8 – Placenta W8
DMS = 0.624
EMD = 0.055
PTMA – Placenta W9
DMS = 0.588
EMD = 0.058
APOLD1 – Placenta W11
DMS = 0.492
EMD = 0.116
KDE–density
Gene expression
scRNA-seq measured
SCHAF inferred

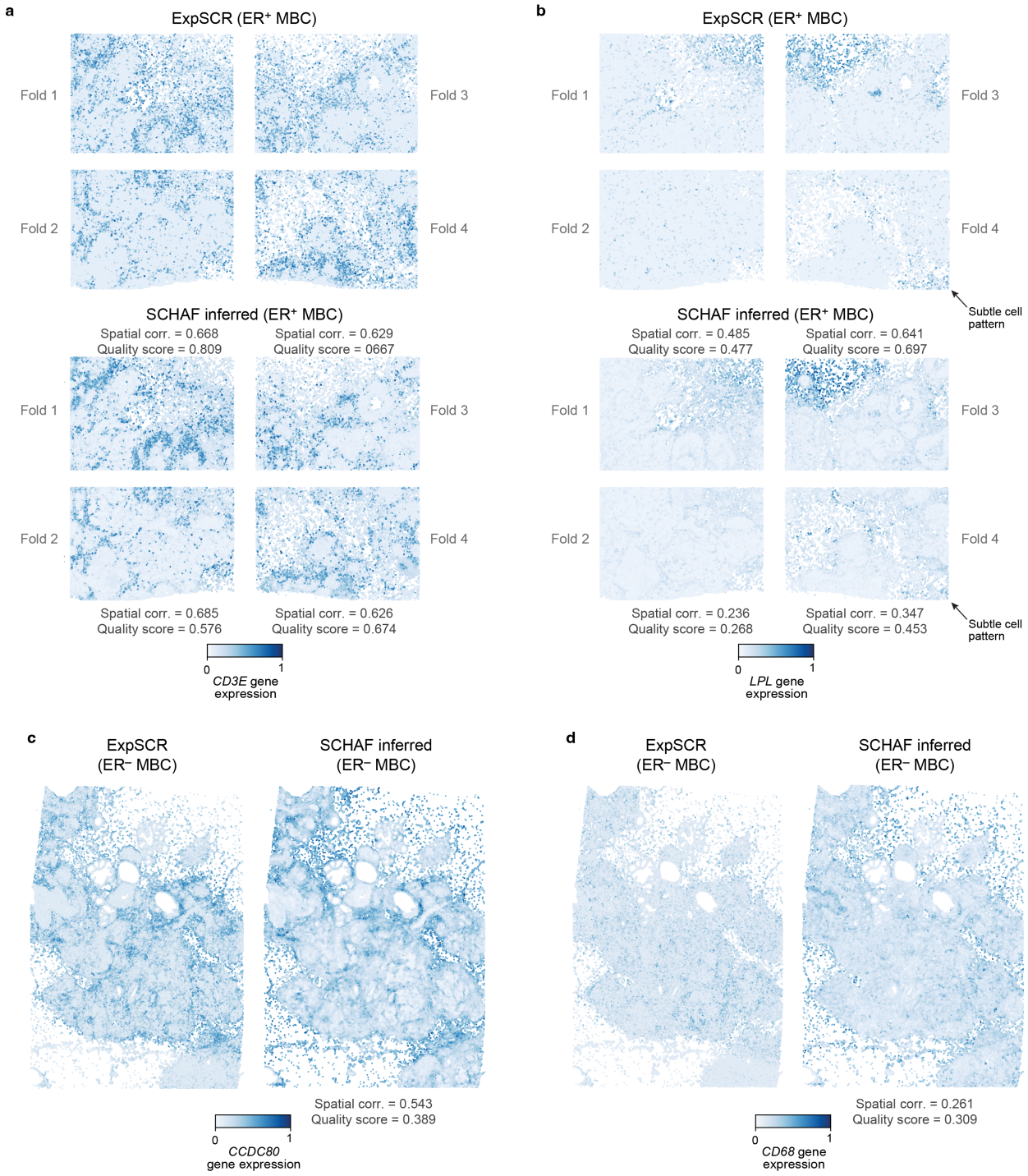
Figure 6


a
SCHAF inferred
Expert pathologist annotated
Placenta W7 (accuracy = 0.728)
Placenta W8 (accuracy = 0.707)
Placenta W9 (accuracy = 0.706)
Placenta W11 (accuracy = 0.621)
Cell type
Fetal endothelial, fetal blood
Fetal fibroblasts, Hofbauer
Decidua, firbin, maternal cells
STBs, vCTBs
EVTs
b
All cell types
CD163 (Hofbauer marker)
EFEMP1 (vCTB marker)
COL3A1 (FIB marker)
QSOX1 (EVT marker)
SCHAF inferred
ExpSCR
0
1
Gene expression
Cell type
Villous cytotrophoblasts
Extravillous trophoblasts
Fetal fibroblasts
Maternal macrophages
Maternal fibroblasts
Hofbauer cels
Fetal endothelial cells
Syncytiotrophoblasts

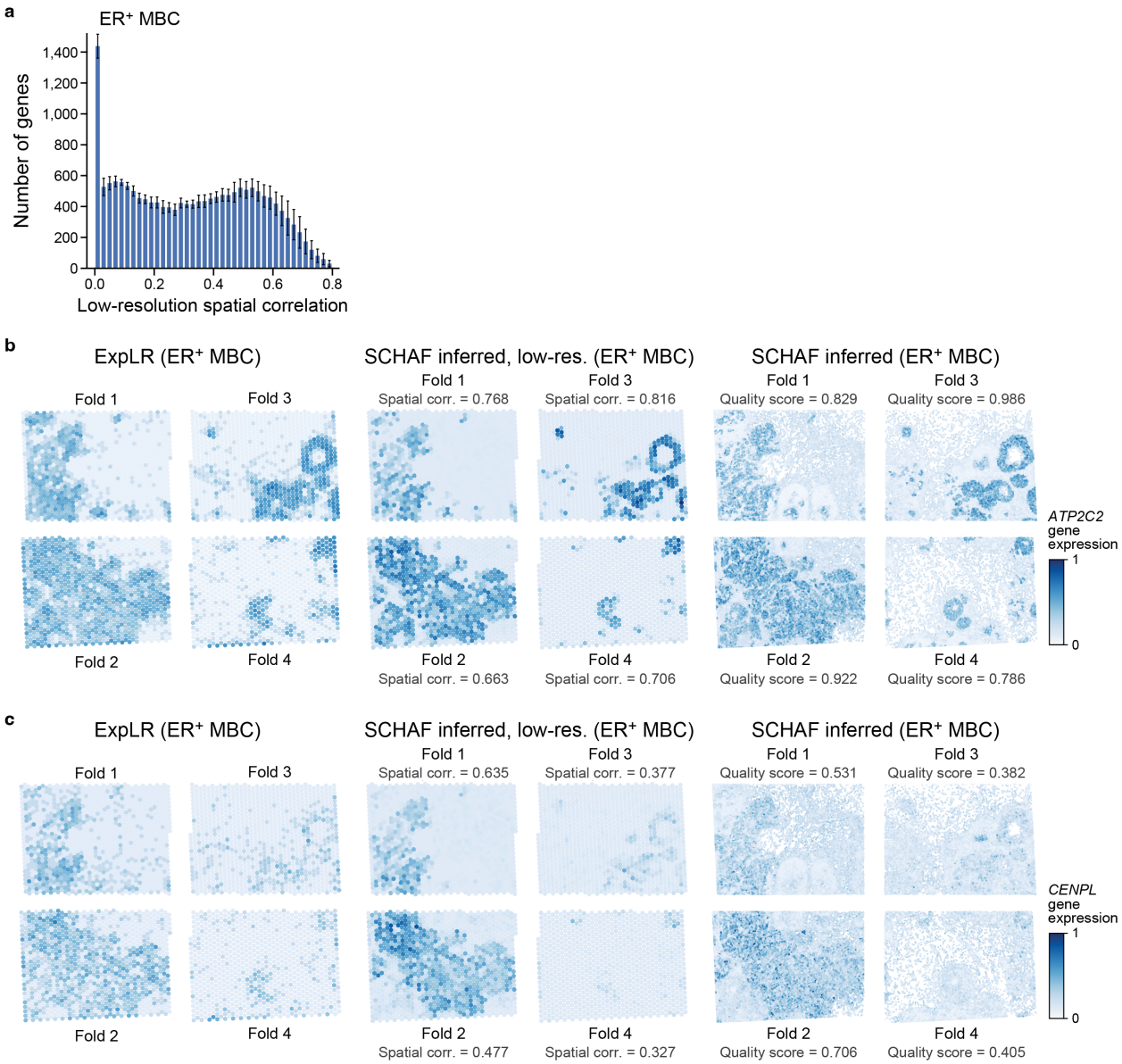
Ext. Data 1


a
ExpSCR (mouse pup)
SCHAF inferred (mouse pup)
Spatial corr. = 0.413
Quality score = 0.15
Spatial corr. = 0.92
Quality score = 1.409
Fold 1
Fold 3
Fold 1
Fold 3
Fold 2
Fold 4
Fold 2
Fold 4
Spatial corr. = 0.17
Quality score = 0.246
Spatial corr. = 0.928
Quality score = 1.381
0
1
Hmgcs2 gene expression
b
ExpSCR (mouse pup)
SCHAF inferred (mouse pup)
Spatial corr. = 0.871
Quality score = 2.661
Spatial corr. = 0.277
Quality score = 0.206
Fold 1
Fold 3
Fold 1
Fold 3
Fold 2
Fold 4
Fold 2
Fold 4
Subtle cell pattern
Subtle cell pattern
Spatial corr. = 0.811
Quality score = 0.763
Spatial corr. = 0.379
Quality score = 0.238
0
1
Nnat gene expression

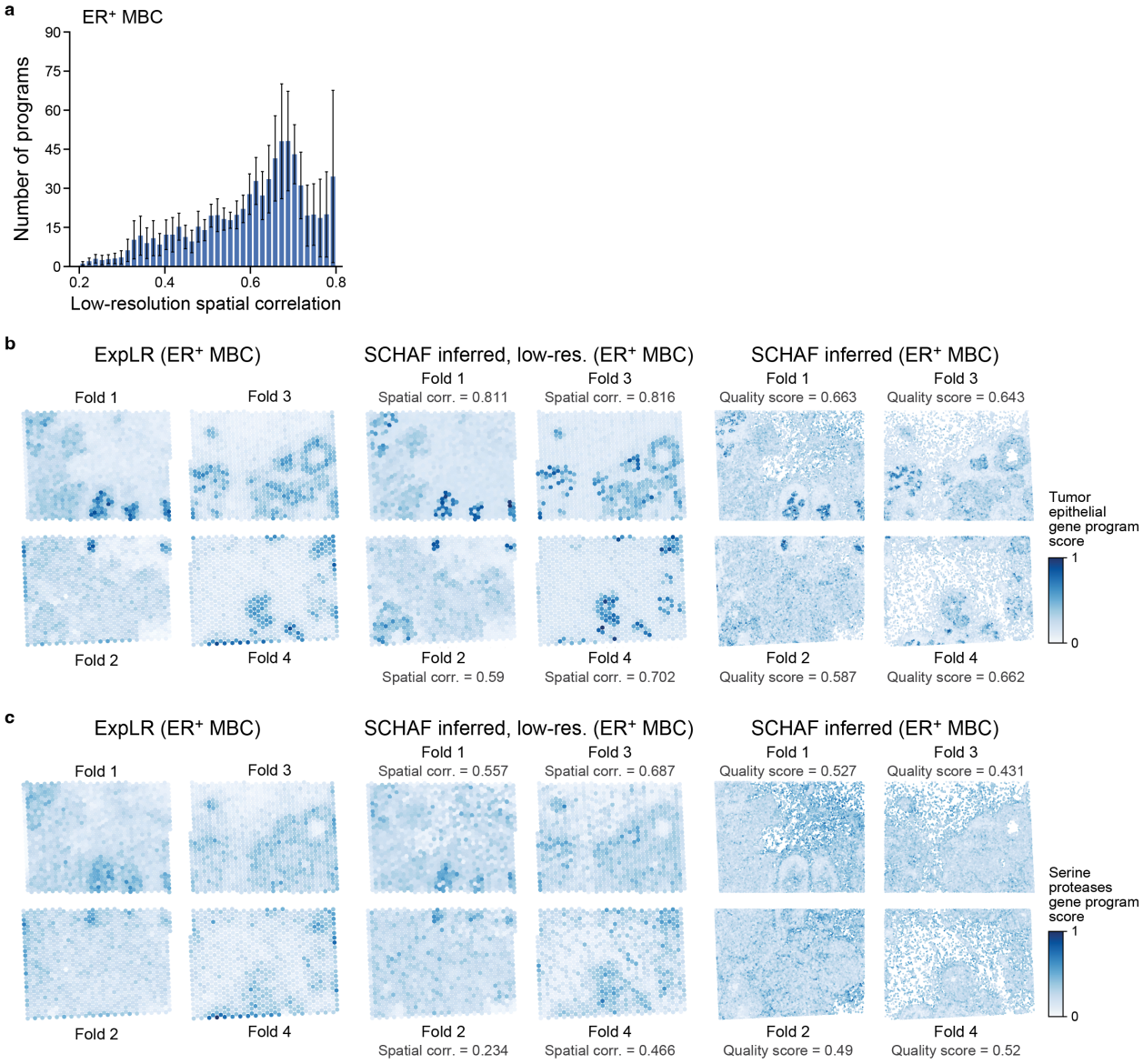
Ext. Data 2


a
ExpSCR (ER+ MBC)
Fold 1
Fold 3
Fold 2
Fold 4
SCHAF inferred (ER+ MBC)
Spatial corr. = 0.668
Quality score = 0.809
Spatial corr. = 0.629
Quality score = 0667
Fold 1
Fold 3
Fold 2
Fold 4
Spatial corr. = 0.685
Quality score = 0.576
Spatial corr. = 0.626
Quality score = 0.674
0
1
CD3E gene expression
b
ExpSCR (ER+ MBC)
Fold 1
Fold 3
Fold 2
Fold 4
Subtle cell pattern
SCHAF inferred (ER+ MBC)
Spatial corr. = 0.485
Quality score = 0.477
Spatial corr. = 0.641
Quality score = 0.697
Fold 1
Fold 3
Fold 2
Fold 4
Spatial corr. = 0.236
Quality score = 0.268
Spatial corr. = 0.347
Quality score = 0.453
Subtle cell pattern
0
1
LPL gene expression
c
ExpSCR (ER− MBC)
SCHAF inferred (ER− MBC)
Spatial corr. = 0.543
Quality score = 0.389
0
1
CCDC80 gene expression
d
ExpSCR (ER− MBC)
SCHAF inferred (ER− MBC)
Spatial corr. = 0.261
Quality score = 0.309
0
1
CD68 gene expression

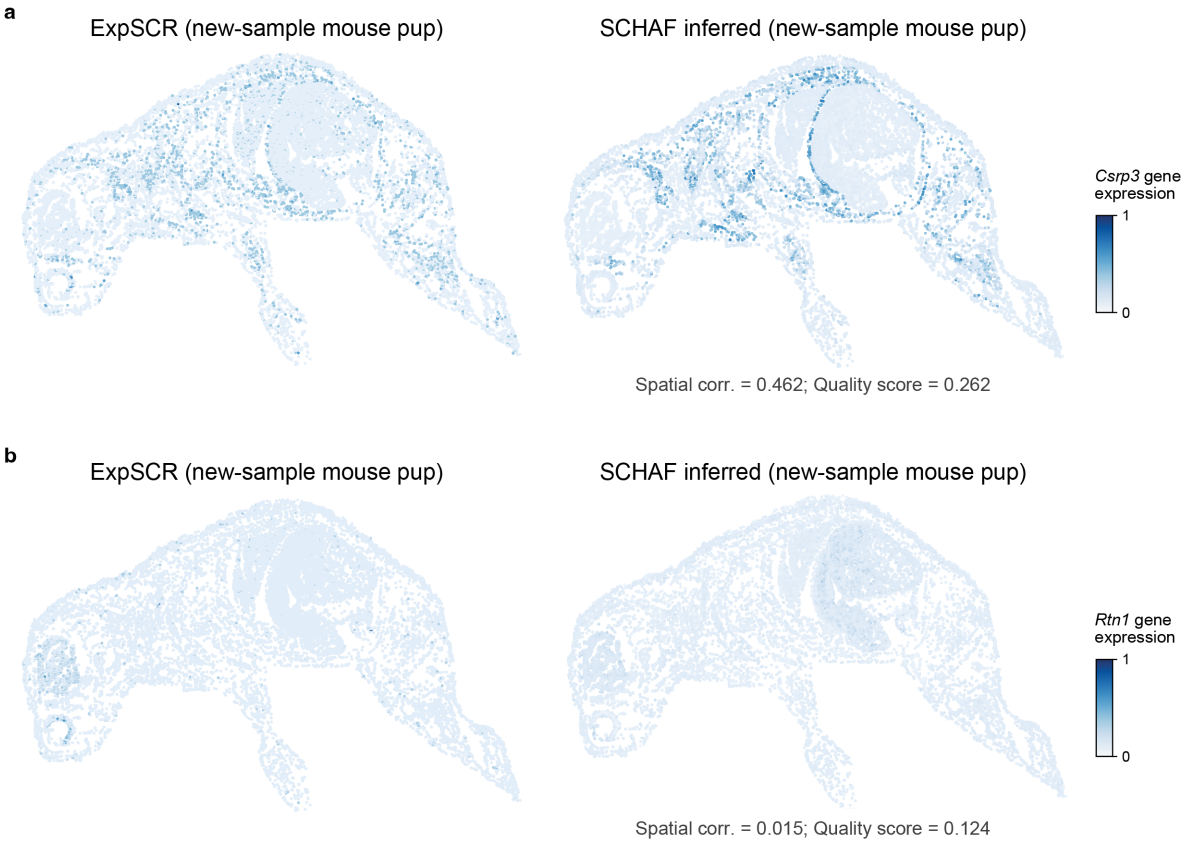
Ext. Data 3


a
ER+ MBC
Number of genes
1,400
1,200
1,000
800
600
400
200
0
0.0
0.2
0.4
0.6
0.8
Low-resolution spatial correlation
b
ExpLR (ER+ MBC)
SCHAF inferred, low-res. (ER+ MBC)
SCHAF inferred (ER+ MBC)
Fold 1
Fold 3
Fold 2
Fold 4
Spatial corr. = 0.768
Spatial corr. = 0.816
Spatial corr. = 0.663
Spatial corr. = 0.706
Quality score = 0.829
Quality score = 0.986
Quality score = 0.922
Quality score = 0.786
ATP2C2 gene expression
1
0
c
ExpLR (ER+ MBC)
SCHAF inferred, low-res. (ER+ MBC)
SCHAF inferred (ER+ MBC)
Fold 1
Fold 3
Fold 2
Fold 4
Spatial corr. = 0.635
Spatial corr. = 0.377
Spatial corr. = 0.477
Spatial corr. = 0.327
Quality score = 0.531
Quality score = 0.382
Quality score = 0.706
Quality score = 0.405
CENPL gene expression
1
0

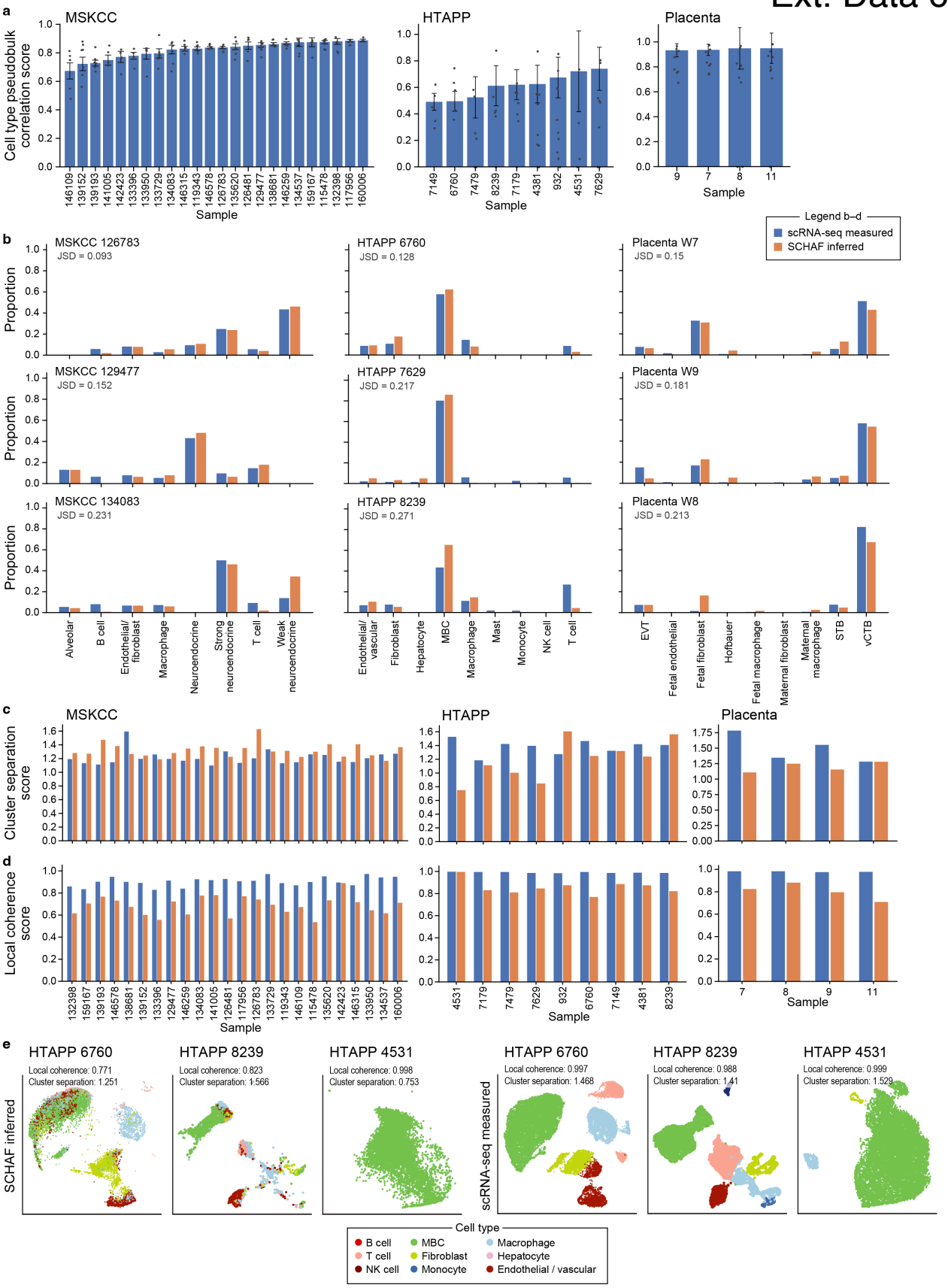
Ext. Data 4


a
ER+ MBC
Number of programs
Low-resolution spatial correlation
b
ExpLR (ER+ MBC)
SCHAF inferred, low-res. (ER+ MBC)
SCHAF inferred (ER+ MBC)
Fold 1
Fold 3
Fold 2
Fold 4
Spatial corr. = 0.811
Spatial corr. = 0.816
Spatial corr. = 0.59
Spatial corr. = 0.702
Quality score = 0.663
Quality score = 0.643
Quality score = 0.587
Quality score = 0.662
Tumor epithelial gene program score
c
Spatial corr. = 0.557
Spatial corr. = 0.687
Spatial corr. = 0.234
Spatial corr. = 0.466
Quality score = 0.527
Quality score = 0.431
Quality score = 0.49
Quality score = 0.52
Serine proteases gene program score

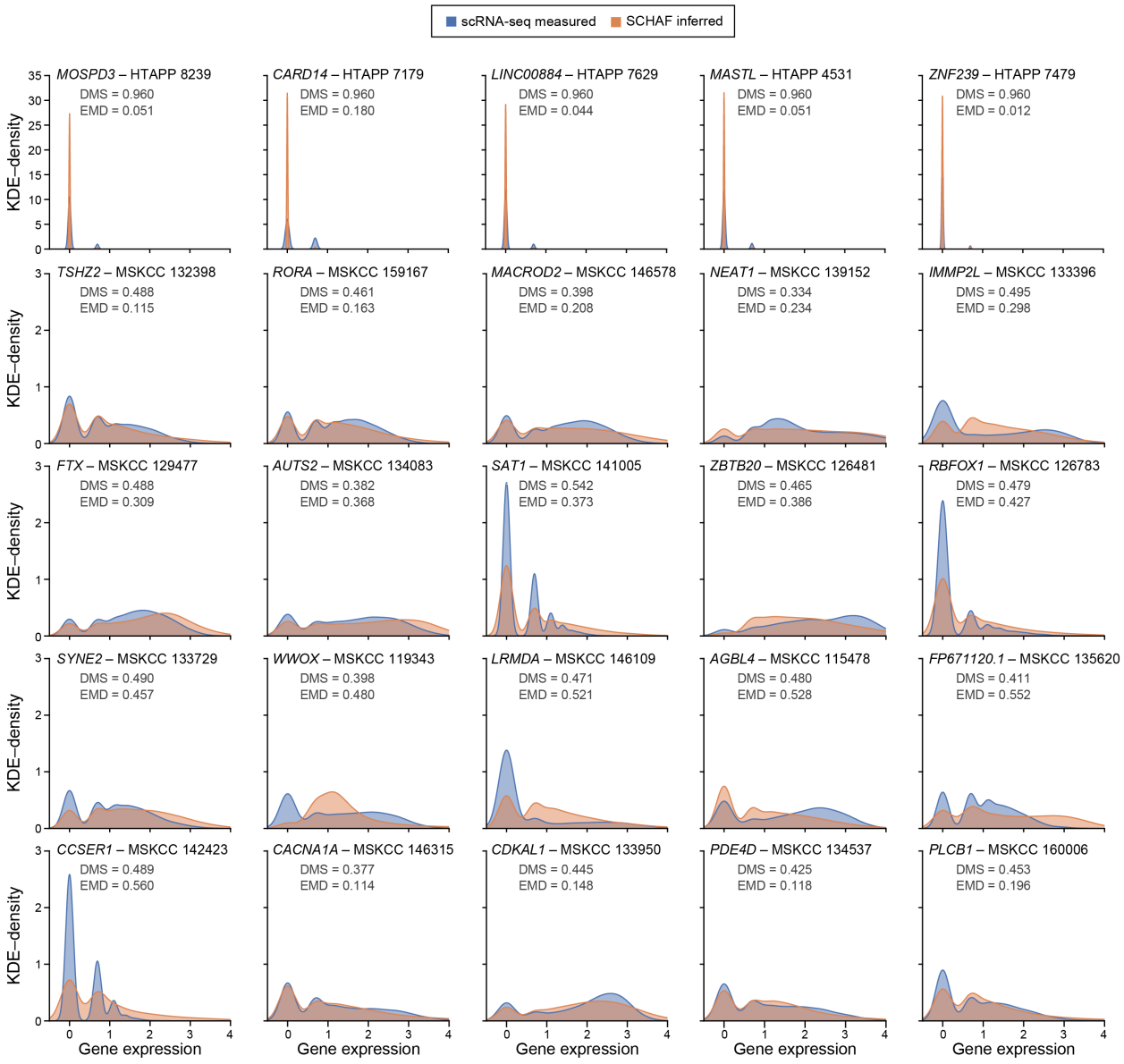
Ext. Data 5


a
ExpSCR (new-sample mouse pup)
SCHAF inferred (new-sample mouse pup)
Csrp3 gene expression
1
0
Spatial corr. = 0.462; Quality score = 0.262
b
ExpSCR (new-sample mouse pup)
SCHAF inferred (new-sample mouse pup)
Rtn1 gene expression
1
0
Spatial corr. = 0.015; Quality score = 0.124

Ext. Data 6


a
MSKCC
HTAPP
Placenta
Cell type pseudobulk correlation score
Sample
b
Legend b–d
scRNA-seq measured
SCHAF inferred
MSKCC 126783
JSD = 0.093
HTAPP 6760
JSD = 0.128
Placenta W7
JSD = 0.15
MSKCC 129477
JSD = 0.152
HTAPP 7629
JSD = 0.217
Placenta W9
JSD = 0.181
MSKCC 134083
JSD = 0.231
HTAPP 8239
JSD = 0.271
Placenta W8
JSD = 0.213
Proportion
Alveolar
B cell
Endothelial/ fibroblast
Macrophage
Neuroendocrine
Strong neuroendocrine
T cell
Weak neuroendocrine
Endothelial/ vascular
Fibroblast
Hepatocyte
MBC
Macrophage
Mast
Monocyte
NK cell
T cell
EVT
Fetal endothelial
Fetal fibroblast
Hofbauer
Fetal macrophage
Maternal fibroblast
Maternal macrophage
STB
vCTB
c
MSKCC
HTAPP
Placenta
Cluster separation score
d
Local coherence score
Sample
e
HTAPP 6760
Local coherence: 0.771
Cluster separation: 1.251
HTAPP 8239
Local coherence: 0.823
Cluster separation: 1.566
HTAPP 4531
Local coherence: 0.998
Cluster separation: 0.753
HTAPP 6760
Local coherence: 0.997
Cluster separation: 1.468
HTAPP 8239
Local coherence: 0.988
Cluster separation: 1.41
HTAPP 4531
Local coherence: 0.999
Cluster separation: 1.529
SCHAF inferred
scRNA-seq measured
Cell type
B cell
T cell
NK cell
MBC
Fibroblast
Monocyte
Macrophage
Hepatocyte
Endothelial / vascular

Ext. Data 7


scRNA-seq measured
SCHAF inferred
MOSPD3 – HTAPP 8239
DMS = 0.960
EMD = 0.051
CARD14 – HTAPP 7179
DMS = 0.960
EMD = 0.180
LINC00884 – HTAPP 7629
DMS = 0.960
EMD = 0.044
MASTL – HTAPP 4531
DMS = 0.960
EMD = 0.051
ZNF239 – HTAPP 7479
DMS = 0.960
EMD = 0.012
TSHZ2 – MSKCC 132398
DMS = 0.488
EMD = 0.115
RORA – MSKCC 159167
DMS = 0.461
EMD = 0.163
MACROD2 – MSKCC 146578
DMS = 0.398
EMD = 0.208
NEAT1 – MSKCC 139152
DMS = 0.334
EMD = 0.234
IMMP2L – MSKCC 133396
DMS = 0.495
EMD = 0.298
FTX – MSKCC 129477
DMS = 0.488
EMD = 0.309
AUTS2 – MSKCC 134083
DMS = 0.382
EMD = 0.368
SAT1 – MSKCC 141005
DMS = 0.542
EMD = 0.373
ZBTB20 – MSKCC 126481
DMS = 0.465
EMD = 0.386
RBFOX1 – MSKCC 126783
DMS = 0.479
EMD = 0.427
SYNE2 – MSKCC 133729
DMS = 0.490
EMD = 0.457
WWOX – MSKCC 119343
DMS = 0.398
EMD = 0.480
LRMDA – MSKCC 146109
DMS = 0.471
EMD = 0.521
AGBL4 – MSKCC 115478
DMS = 0.480
EMD = 0.528
FP671120.1 – MSKCC 135620
DMS = 0.411
EMD = 0.552
CCSER1 – MSKCC 142423
DMS = 0.489
EMD = 0.560
CACNA1A – MSKCC 146315
DMS = 0.377
EMD = 0.114
CDKAL1 – MSKCC 133950
DMS = 0.445
EMD = 0.148
PDE4D – MSKCC 134537
DMS = 0.425
EMD = 0.118
PLCB1 – MSKCC 160006
DMS = 0.453
EMD = 0.196
KDE–density
Gene expression