Modelling cell-cell interactions from spatial molecular data with spatial variance component analysis

Damien Arnol[1], Denis Schapiro[4,5], Bernd Bodenmiller[4], Julio Saez-Rodriguez[1,2], Oliver Stegle[1,3]

1. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, CB10 1SD, Hinxton, Cambridge, UK
2. Joint Research Center for Computational Biomedicine, RWTH Aachen University, Faculty of Medicine, Pauwelsstrasse 19, D-52074 Aachen, Germany
3. European Molecular Biology Laboratory, Genome Biology Unit, Heidelberg, Germany
4. Institute of Molecular Life Sciences, University of Zurich, Zurich, Switzerland.
5. Life Science Zurich Graduate School, ETH Zurich and University of Zurich, Zurich, Switzerland.

Method: SVCA (spatial variance component analysis)

Abstract

Technological advances allow for assaying multiplexed spatially resolved RNA and protein expression profiling of individual cells, thereby capturing physiological tissue contexts of single cell variation. While methods for assaying spatial expression profiles are increasingly accessible, there is a lack of computational approaches that allow for studying the relevance of the spatial organization of tissues on cell-cell heterogeneity. Here, we present *spatial variance component analysis (SVCA),* a computational framework for the analysis of spatial molecular data. SVCA estimates *signatures of spatial variance components*, thereby quantifying the effect of cell-cell interactions, as well as environmental and intrinsic cell features on the expression levels of individual molecules. In application to a breast cancer Imaging Mass Cytometry dataset, our model yields robust spatial variance signatures, identifying cell-cell interactions as a major driver of expression heterogeneity. We also apply SVCA to high-dimensional imaging-derived RNA data, where we identify molecular pathways that are linked to cell-cell interactions.

Introduction

Experimental advances have enabled assaying RNA and protein abundances of single cells in spatial contexts, thereby allowing to study single cell variation in tissues. Already, these technologies have delivered new insights into tissue systems and the sources of transcriptional variation (Battich, Stoeger, & Pelkmans, 2013; Bodenmiller, 2016), with potential use as biomarkers for human health (Bodenmiller, 2016). Spatial expression variation can reflect interactions between adjacent cells, or can be caused by cells that migrate to specific locations in a tissue to perform their functions (e.g. immune cells).

Different technologies allow for assaying spatially resolved expression profiles. Imaging Mass Cytometry (IMC) (Giesen et al., 2014) and Multiplexed Ion Beam Imaging (MIBI) (Angelo et al., 2014) rely on protein labeling with antibodies coupled with metal isotopes of specific masses followed by high-resolution tissue ablation and ionisation. IMC enables profiling of over 40 targeted proteins with subcellular resolution. Other methods such as MxIF and CycIF use immunofluorescence for protein quantification of tens of markers at a time (Gerdes et al., 2013; Lin, Fallahi-Sichani, & Sorger, 2015). Increasingly, there also exist optical imaging-based assays to measure single cell RNA levels. Mer-FISH and seq-FISH use a combinatorial approach of fluorescence-labeled small RNA probes to identify and localise single RNA molecules (Chen, Boettiger, Moffitt, Wang, & Zhuang, 2015; Gerdes et al., 2013; Lin et al., 2015; Shah, Lubeck, Zhou, & Cai, 2017)), which allows for measuring larger numbers of readouts (currently between 140 and 250). In addition, there exist other spatial expression profiling techniques such as Spatial Transcriptomics (Ståhl et al., 2016), which currently however do not offer single cell resolution and are therefore not adequate to study cell-to-cell variation.

The availability of spatially resolved expression data represents an unprecedented opportunity to disentangle largely unexplored sources of single cell variations: i) intrinsic sources of variation due to difference in cell types or states (e.g. cell-cycle state), ii) environmental effects from the cell microenvironment and iii) interactions between adjacent cells. Yet, although experimentally spatial omics profiles can be generated with increasingly high throughput, the required computational strategies for interpreting the resulting data are only beginning to emerge. On the one hand, there exist methods for spatial clustering based on single cell expression profiles (Achim et al., 2015), and to assess the overall relevance of the spatial topology on gene expression (Svensson, Teichmann, & Stegle, n.d.). However, these methods do not allow for directly assessing cell-cell interactions. On the other hand, methods exist that can capture cell-cell interactions based on manual cell-type assignments and predefined cellular neighbourhoods (Goltsev et al., 2018; Schapiro et al., 2017; Schulz et al., 2018)(Levine et al., 2015). While these methods have already identified qualitative insights into the interaction of cell types, they do not allow for quantifying the impact of spatial effects on individual genes or proteins.

Here, we present *Spatial variance component analysis (SVCA),* a computational framework to model spatial sources of variation of individual genes or proteins. SVCA allows for decomposing the sources of variation into intrinsic effects, environmental effects and cell-cell interactions. The model is parameterized based on continuous expression profiles, and in particular avoids the need to define discrete cell types. We illustrate SVCA using data from multiple technologies and biological domains, including IMC proteomics profiles data from human cancer tissue and spatial single-cell profiles from the mouse hippocampus based on seqFISH. Across these applications, we find pronounced effects due to cell-cell interactions, and we identify biologically relevant genes and pathways that participate in these processes.

Results and Discussion

SVCA considers random effect components that capture i) *intrinsic* sources of variation due to differences in cell types or states, ii) *environmental* sources of variation due to local extracellular factors iii) source of variations due to *cell-cell interactions* (**Fig. 1a**). The effect of these factors on molecular expression variability is modelled using a random effect model (**Fig. 1b**). Importantly, SVCA does not require assigning cells to discrete cell types, but instead is based on a continuous measure of cell-cell similarities estimated from the covariance of their expression profiles (**Fig. S1**). Our random effect model also circumvents the need to define discrete neighbours but instead weights interactions between pairs of cells as a function of their distance (**Fig. S1**). The fitted SVCA model provides a decomposition of these sources of variation of individual genes and proteins, and the model can be used to assess the significance of these respective components. The resulting spatial variance signatures provides a compact summary of the drivers of expression variation in a dataset (**Fig. 1c**).

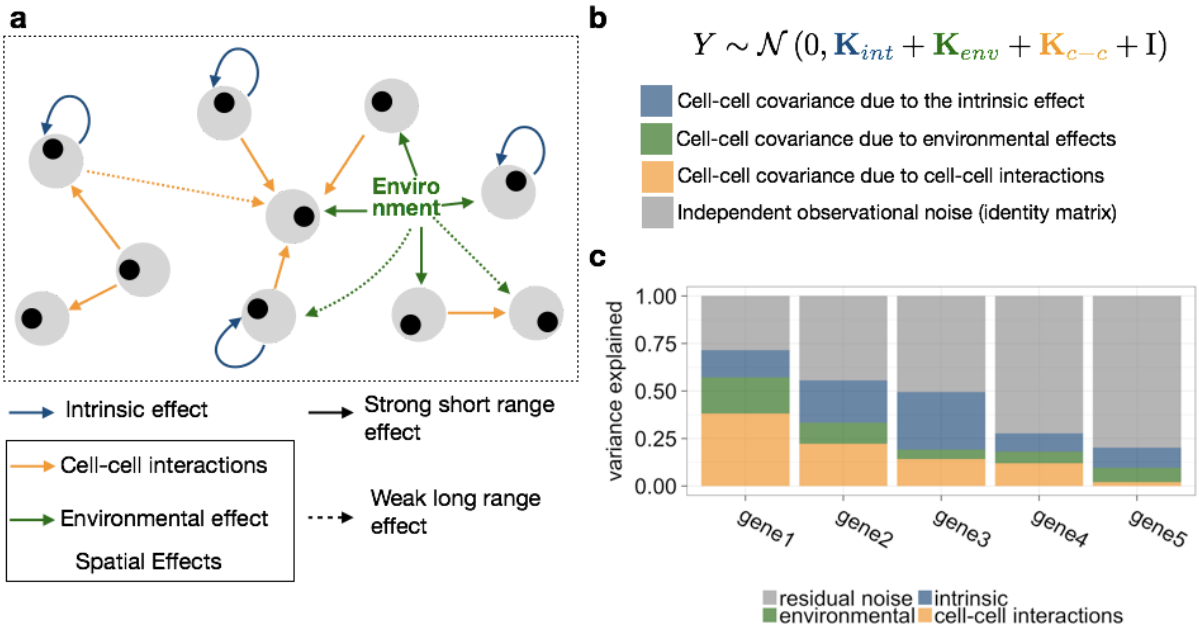

Figure 1 | Spatial variance component analysis (SVCA): approach and overview.
(**a**) SVCA allows for decomposing the variability of individual genes and proteins into i) cell intrinsic effects (due to differences in intrinsic cell type or state, blue), ii) general environmental effects that capture expression differences due to local extracellular factors (green) and iii) a cell-cell interaction component that captures differences in expression level attributable to different cellular composition of a cell's neighbourhood (yellow). (**b**) SVCA is based on a random effect framework, considering additive contributions of these different components (**Methods**). See **Fig. S1** and **Methods** for details of how these random effect components are defined.(**c**) SVCA output: gene-level break down of the proportion of variance attributable to different components.

Initially, we used simulated data to assess the statistical calibration and the accuracy of SVCA variance estimates. Briefly, we generated expression profiles by sampling from the model, using empirical parameters derived from 11 real datasets (**Methods**), including the position of cells and the cell state covariance. First, we simulated expression profiles assuming no interaction effects to assess the calibration of the corresponding test, finding that the model yields conservative estimates (**Fig. 2a**). We also assessed the detection power of the interaction test when simulating cell-cell interactions that explain increasing proportions of the gene expression variance (**Fig. 2c**), and when varying the number of cells in the dataset (**Fig. 2d**). Finally, we compared the estimated variance components for cell-cell interactions with the simulated variance components, again observing that the model estimates are conservative (**Fig. 2b**). Taken together, these results demonstrate that SVCA can be used to estimate and test for spatial drivers of single cell variability, in particular cell-cell interactions.

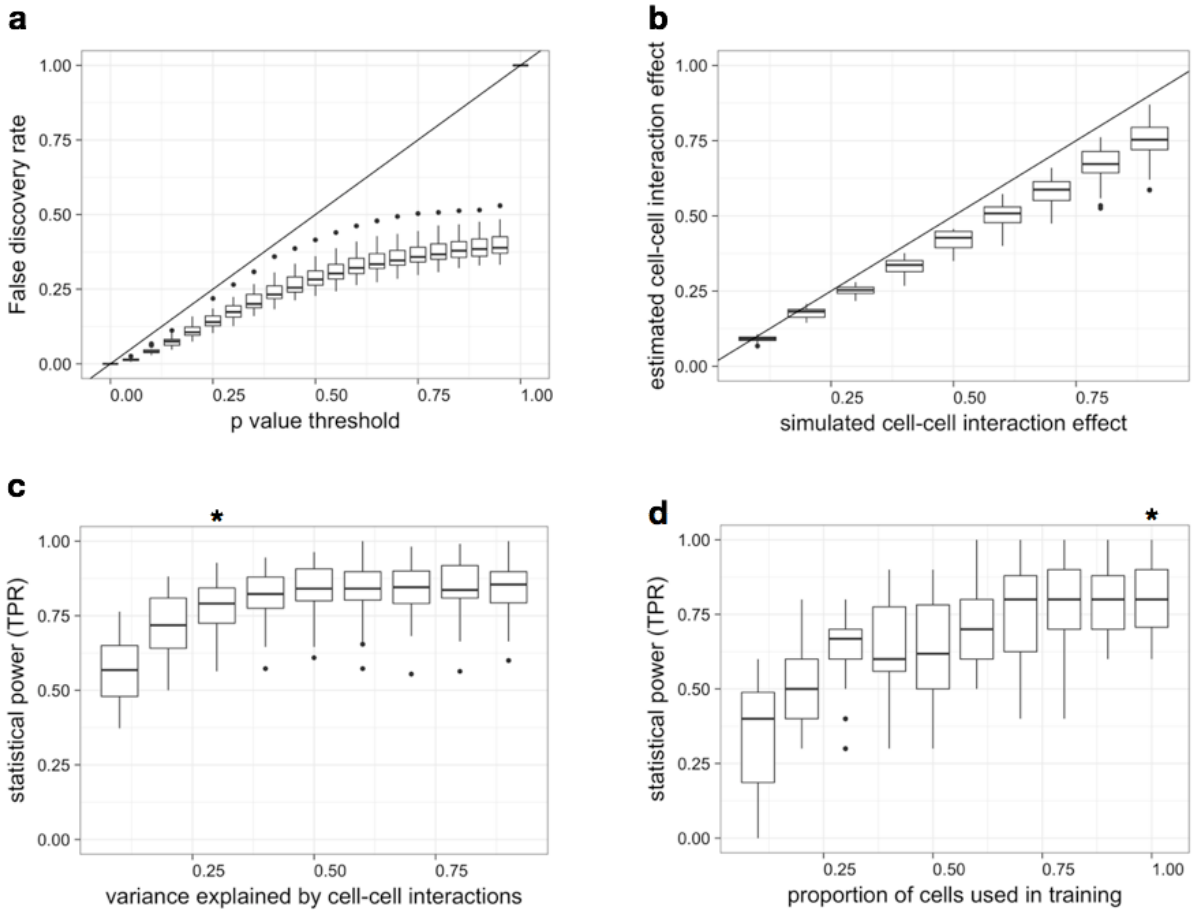

Figure 2 | Validation of cell-cell tests and variance components using simulated data. **(a)** Empirical false discovery rate for the cell-cell interaction test using data simulated from the null. Shown is the empirical false discovery rate (FDR) as a function of the applied P value threshold. **(b)** Proportion of variance due to cell-cell interactions estimated by SVCA when varying the true proportion of variance explained (**Methods**). **(c,d)** Statistical power

for the cell-cell interaction test (at family-wise error rate <1%), when varying the proportion of variance explained by cell-cell interactions (**c**) and when considering a subset of all cells for model fitting (**d**). All cells were used for model fitting in **c** (indicated using the asterix symbol). For each panel, boxplots show the distribution of results across 110 simulated settings (11 images times 10 independent runs) and for 26 proteins. Rates in panel **a**, **c**, **d** (True Positive Rate - TPR and False Discovery Rate - FDR) are computed across the 110 simulations for each protein and estimated effects in panel **b** are averaged across simulations.

Application of SVCA to spatial proteomics data of breast cancer tissues

Next, we applied SVCA to a recent Imaging Mass Cytometry (IMC) dataset from human breast cancers, consisting of 52 breast biopsies from 27 breast cancer patients with variable disease grade and from different cancer subtypes, sampled from different tumour locations (Schapiro et al., 2017). SVCA revealed substantial differences of cell-cell interaction component across proteins, explaining up to half of the expression variability (**Fig 3a**). Immune cell markers were enriched among the set of proteins with the strongest cell-cell interaction effects: CD44, CD20, CD3 and CD68 had significant cell-cell interaction components in 35, 34, 34 and 38 out of the 52 images respectively (FDR<1%, Benjamini-Hochberg adjusted, **Fig. 3a**). We used bootstrapping to confirm the robustness of the variance estimates (**Fig. S5**), and we observed that a variance component model that accounts for cell-cell interactions was able to more accurately impute gene expression profiles than models that ignore such effects (**Fig. 3b, Fig. S2)**.

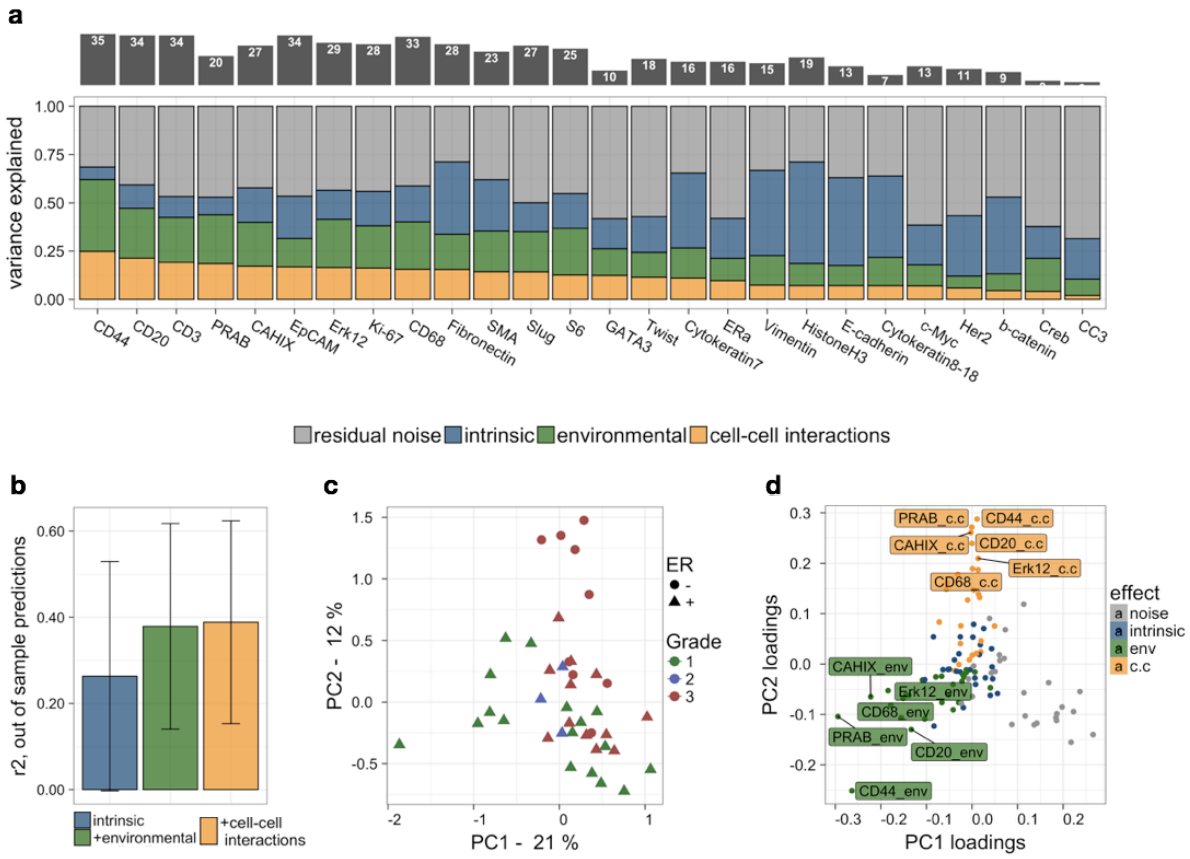

Figure 3 | Application of SVCA to 52 breast cancer samples profiled using IMC. (**a**) Bottom panel: SVCA signatures for 26 proteins. Shown are averages of the proportion of

variance explained by intrinsic effects, environmental effects and cell-cell interactions, across 52 images. Proteins are ordered by the magnitude of the cell-cell interaction component. Top panel: The number of images with significant cell-cell interaction effects for the corresponding proteins (FDR<1%, **Methods**, **Fig S7**). (**b**) Out of sample prediction accuracy using SVCA and models of lower complexity, using 5 fold cross validation. Shown are average coefficients (r^2) between predicted and observed gene expression values, averaged across proteins and images. Error bars correspond to plus and minus one standard deviation indicating the variation across images and proteins. (**c**) First two principal components for 42 clinically annotated images, calculated based on the full variance component signature, with individual images coloured by tumour grade. (**d**) Loadings of the variance components for the two first PCs, showing the relevance of individual proteins and variance components. Variations of the second principal component is driven by the relative values of the environmental component and the local components for some key proteins.

We also observed substantial variations in the spatial variance signatures between individual images (**Fig. S4**), motivating investigating the relationship between these differences and clinical covariates, including tumour grade, ER status, PR status, HER2 status. Using Principal Component Analysis, we identified tumour grade as the most important explanatory variable for differences in spatial variance components (**Fig 3c**), followed by ER status and PR status (**Fig. S8**). Inspection of the PCA loadings (**Fig. 3d**) identified the cell-cell interaction component and the environmental component for a subset of proteins (including PR and CD44) as the most informative SVCA feature affected by tumour grade. This was also evident when comparing environmental and cell-cell interaction effects across tumours of different grades (**Fig. S9**), showing a relative increase in the cell-cell interaction component in higher grade tumours. This effect could in part be due to the increased cell density in later stage tumours.

RNA-based datasets based on imaging technologies

SVCA can be used for the analysis of data from a broad range of spatially resolved technology, including imaging-based assays. To explore this, we considered a mouse hippocampus datasets, profiling 240 RNA expression levels in 21 distinct brain regions of a single animal, using seqFISH (Shah et al., 2017). Leveraging the high-dimensionality of the data, we sought to identify individual genes and pathways that are most linked to cell-cell interactions.

Similarly to the IMC dataset, SVCA signatures were robust and could accurately predict gene expression levels out of sample (**Fig S17**, **S18**, **S19**). After averaging SVCA signatures across images, we used the reactome database (Croft et al., 2014) to identify pathways that were enriched among genes with large cell-cell interaction components (**Methods**). 5 pathways showed a significant enrichment, with *transmission across chemical synapses, neurotransmitter release cycle* and *neuronal system* among the ones with highest significance (FDR<5%, considering 60 reactome pathways with three or more genes contained in this dataset, **Methods**, **Fig. 4c**). For comparison, we also considered gene set enrichment based on the average expression level of genes across images, which identified distinct GO categories (*transmembrane transport* and *SLC-mediated transmembrane transport*), with only little overlap to the pathways identified from the cell-cell interactions.

This demonstrates the complementarity of spatial and abundance-based signatures for studying biological activity in tissues, as only the former accounts for the spatial structure of the cells' native environment.

Similarly to results obtained on the IMC datasets, spatial variance signatures were variable across images, which is consistent with the high biological variability between the different regions of the hippocampus imaged here reported in (Shah et al., 2017). Principal components of the global spatial variance signature for the dorsal region clustered together, irrespective of their CA1/CA3 location. Similarly, images from the Dentate Gyrus (DG) also clustered together, and there was some proximity between signatures from the ventral region, although with more variation between them (**Fig. 4d,e**). This is consistent with Shah et al's observation that the ventral and dorsal regions of the CA1 and the CA3 mirrored each other with respect to their cellular compositions, and that ventral regions are much more heterogeneous in their cellular composition. Spatial Variance signatures for intermediate regions, however, did not show much resemblance (**Fig. 4d,f**).

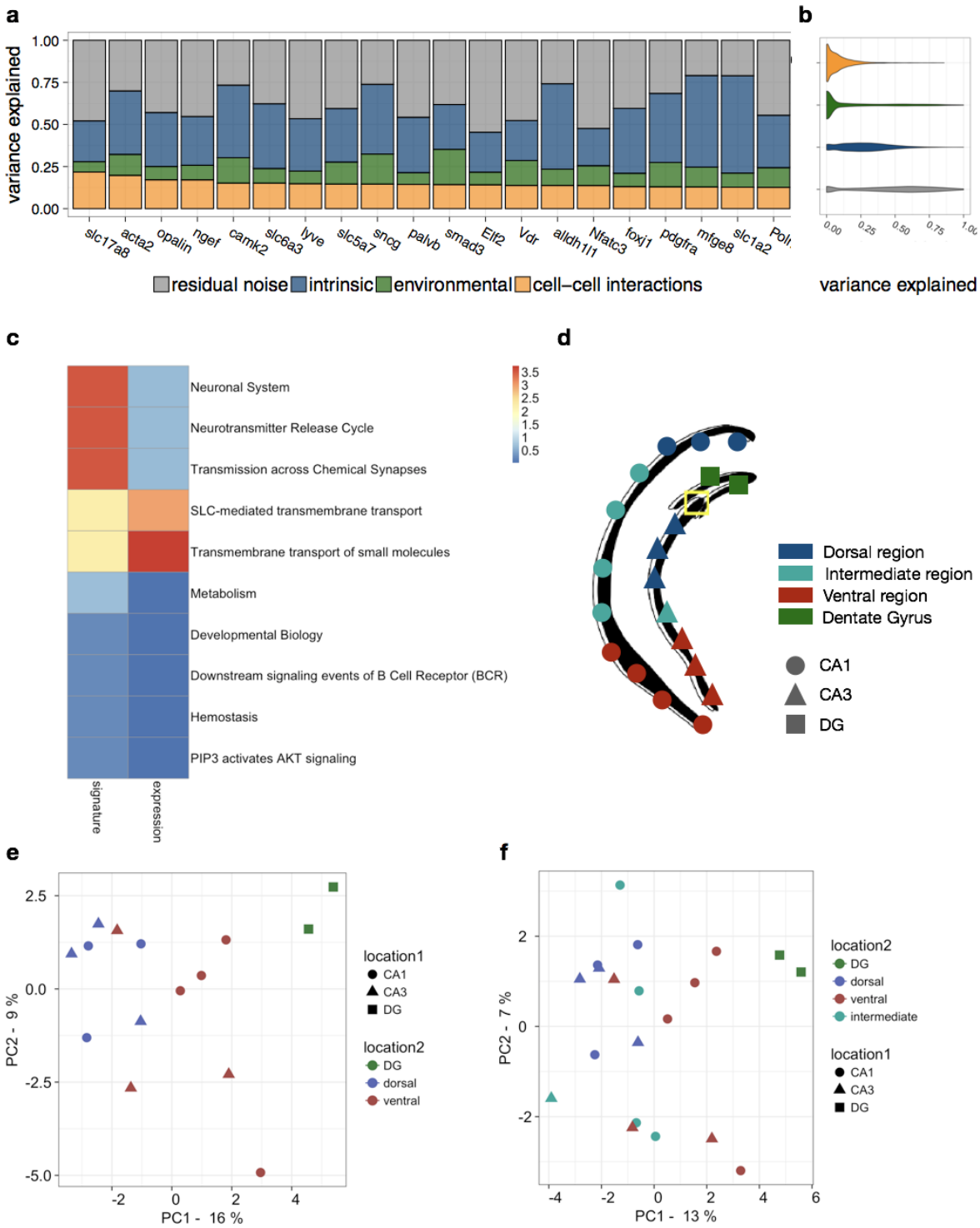

Figure 4: seqFISH results:

(a) Spatial Variance signature averaged across images: top 20 genes for cell-cell interactions. **(b)** Variance estimates distribution across images and proteins **(c)** Top 10 enriched pathways out of 60 reactome pathways using a rank-based enrichment analysis (see **Methods**) in the cell-cell interactions variance component averaged across images (left) compared to the corresponding gene expression enrichments, averaged across images (right). Colours denote statistical significance (negative log Benjamini-Hochberg adjusted P values). **(d)** Spatial organisation of the mouse hippocampus Each dot corresponds to an

image. Colours and shapes denote regions using the classification as in (Shah et al., 2017). **(e)** First two principal components of the spatial variance signatures for individual images from the DG, the dorsal region and the ventral region. The colours and shapes represent the location of the biopsy in the hippocampus. **(f)** First two principal components of the spatial variance signatures for all images.

We also applied SVCA to a breast cancer cell culture dataset profiled using the mer-FISH technique (140 RNA expression levels) (Moffitt et al., 2016). The large spatial domain and the number of cells imaged from a biological homogenous cell culture system means that multiple field of views can be defined as technical replicates. Consistent with this, the variance components of SVCA were highly consistent across images (**Fig. S10**, **Fig. S11**, **Fig. S12**, **Fig. S14**), with average coefficients of variation for cell-cell interactions of 20-40%, compared to typically 75-150% (**Fig. S13**) for the IMC and seq-fish data (**Fig. S4**, **Fig. S20**). These results further support the general applicability of the model to data from different technology and the robustness of SVCA variance signatures.

Conclusion

We presented Spatial Variance Component Analysis (SVCA), a framework for the analysis of spatially resolved molecular expression data. Our model computes a spatial variance signature for individual mRNA or protein levels, decomposing their sources of variation into spatial and non-spatial components. Most prominently, SVCA provides a quantitative assessment of the effect of cell-cell interactions on the expression profile of individual molecules. The model avoids the definition of cell types and neighbourhoods, instead using continuous measure of cell state and euclidean distances between cells.

We have applied SVCA to multiple datasets that were generated using alternative technologies, probing either RNA transcripts or proteins, thereby demonstrating the broad applicability of the model. Across these applications, we observed that cell-cell interactions substantially contributed to gene expression variation, which is consistent with previous reports (Battich, Stoeger, & Pelkmans, 2015) and supports the concept that studying single cell expression in the native context is important to understanding the sources of these variations. We also showed that the variability of spatial variance signatures across samples of the same biological system could be linked to different clinical contexts or to the internal structure of a given organ, which underlines the biological relevance of these signatures. Finally, using higher dimensionality data obtained from optical technologies, we found that genes with largest cell-cell interaction components were enriched in specific pathways, suggesting that spatial variance signatures could help us understanding biological activity in tissue. These results suggest that analysing single cell data in their native context, with SVCA or related methods, will help us understand how the spatial organisation of tissues impact single cell biology.

Although we have tested the calibration and robustness of SVCA, the model is not free of limitations. At present, the model does not account for technology specific noise and instead assumes Gaussian distributed residuals, thus requiring suitable processing of the raw data

such that these assumptions are sufficiently met (See **Methods**). Further development could consider generalized LMMs, for example to couple the random effect component with a negative-binomial likelihood. A second limitation of SVCA is that the model is univariate, which means that individual genes or proteins are modelled independently from each other. A multivariate model could account for relationships between genes involved in the same pathways, either in an unsupervised manner or using prior knowledge (Buettner, Pratanwanich, McCarthy, Marioni, & Stegle, 2017). Such approaches could give a more comprehensive understanding of how biological processes are affected by tissue structure.

There is a growing appreciation of the role of spatial distribution of proteins, transcripts and other molecules in determining tissues functioning and its deregulation in disease, with potential value as predictors of clinical outcomes. This is largely driven by vigorous development of novel technologies that enable us to capture such data (Aichler & Walch, 2015; Bodenmiller, 2016; Goltsev et al., 2018; Lin et al., 2017; Schulz et al., 2018). We believe that the SVCA framework and extensions thereof will be of broad use to analyze this burgeoning spatially-resolved molecular data to advance our understanding of the pathophysiology of multiple diseases.

Supplementary materials

https://drive.google.com/open?id=1iCZXGcORlodr8gLZQaPlv0LlSW6RNUSG

Author contributions

Author contributions: DA, OS developed the statistical method. DA implemented the model and analysed all the data. DS, BB contributed to the interpretation of the data and the design of the method. DA, DS, OS, JSR wrote the manuscript with input from all authors. JSR and OS conceived the project and supervised the work.

Acknowledgments

DA, JSR, OS acknowledge EMBL core funding. DS was supported by the Forschungskredit of the University of Zurich, grant FK-74419-01-01, and the BioEntrepreneur-Fellowship of the University of Zurich, reference no. BIOEF-17-001. B.B.'s research is funded by an SNSF R'Equip grant, an SNSF Assistant Professorship grant, the SystemsX Transfer Project "Friends and Foes", the SystemsX MetastasiX and PhosphoNetX grant, NIH grant (UC4 DK108132), and the European Research Council (ERC) under the European Union's Seventh Framework Program (FP/2007-2013)/ERC grant agreement no. 336921.

We thank R. Argelaguet, V. Svensson, R. Vento, H. Jackson, H. Jackson, A. Baud, N. Cai, F.P. Casale and D. Horta for discussions on data processing, model design and implementation and results visualisation.

Competing financial interest statement

The authors declare no competing financial interests.

References

Achim, K., Pettit, J.-B., Saraiva, L. R., Gavriouchkina, D., Larsson, T., Arendt, D., & Marioni, J. C. (2015). High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nature Biotechnology*, *33*(5), 503–509. https://doi.org/10.1038/nbt.3209

Aichler, M., & Walch, A. (2015). MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, *95*(4), 422. https://doi.org/10.1038/labinvest.2014.156

Angelo, M., Bendall, S. C., Finck, R., Hale, M. B., Hitzman, C., Borowsky, A. D., … Nolan, G. P. (2014). Multiplexed ion beam imaging of human breast tumors. *Nature Medicine*, *20*(4), 436–442. https://doi.org/10.1038/nm.3488

Battich, N., Stoeger, T., & Pelkmans, L. (2013). Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nature Methods*, *10*(11), 1127–1133. https://doi.org/10.1038/nmeth.2657

Battich, N., Stoeger, T., & Pelkmans, L. (2015). Control of Transcript Variability in Single Mammalian Cells. *Cell*, *163*(7), 1596–1610. https://doi.org/10.1016/j.cell.2015.11.018

Bodenmiller, B. (2016). Multiplexed Epitope-Based Tissue Imaging for Discovery and Healthcare Applications. *Cell Systems*, *2*(4), 225–238. https://doi.org/10.1016/j.cels.2016.03.008

Buettner, F., Pratanwanich, N., McCarthy, D. J., Marioni, J. C., & Stegle, O. (2017). f-scLVM: scalable and versatile factor analysis for single-cell RNA-seq. *Genome Biology*, *18*(1), 212. https://doi.org/10.1186/s13059-017-1334-8

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., & Zhuang, X. (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*, *348*(6233), aaa6090. https://doi.org/10.1126/science.aaa6090

Croft, D., Mundo, A. F., Haw, R., Milacic, M., Weiser, J., Wu, G., … D'Eustachio, P. (2014). The Reactome pathway knowledgebase. *Nucleic Acids Research*, *42*(Database issue), D472–7. https://doi.org/10.1093/nar/gkt1102

Gerdes, M. J., Sevinsky, C. J., Sood, A., Adak, S., Bello, M. O., Bordwell, A., … Ginty, F. (2013). Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(29), 11982–11987. https://doi.org/10.1073/pnas.1300136110

Giesen, C., Wang, H. A. O., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., … Bodenmiller, B. (2014). Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nature Methods*, *11*(4), 417–422. https://doi.org/10.1038/nmeth.2869

Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vasquez, G., … Nolan, G. (2018, February 5). *Deep profiling of mouse splenic architecture with CODEX multiplexed imaging*. *bioRxiv*. https://doi.org/10.1101/203166

Levine, J. H., Simonds, E. F., Bendall, S. C., Davis, K. L., Amir, E.-A. D., Tadmor, M. D., … Nolan, G. P. (2015). Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell*, *162*(1), 184–197. https://doi.org/10.1016/j.cell.2015.05.047

Lin, J.-R., Fallahi-Sichani, M., & Sorger, P. K. (2015). Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nature Communications*, *6*, 8390. https://doi.org/10.1038/ncomms9390

Lin, J.-R., Izar, B., Mei, S., Wang, S., Shah, P., & Sorger, P. (2017, June 19). *A simple open-source method for highly multiplexed imaging of single cells in tissues and tumours*. *bioRxiv*. https://doi.org/10.1101/151738

Moffitt, J. R., Hao, J., Wang, G., Chen, K. H., Babcock, H. P., & Zhuang, X. (2016). High-throughput single-cell gene-expression profiling with multiplexed error-robust

fluorescence in situ hybridization. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(39), 11046–11051. https://doi.org/10.1073/pnas.1612826113

Schapiro, D., Jackson, H. W., Raghuraman, S., Fischer, J. R., Zanotelli, V. R. T., Schulz, D., … Bodenmiller, B. (2017). histoCAT: analysis of cell phenotypes and interactions in multiplex image cytometry data. *Nature Methods*, *14*(9), 873–876. https://doi.org/10.1038/nmeth.4391

Schulz, D., Zanotelli, V. R. T., Fischer, J. R., Schapiro, D., Engler, S., Lun, X.-K., … Bodenmiller, B. (2018). Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. *Cell Systems*, *6*(1), 25–36.e5. https://doi.org/10.1016/j.cels.2017.12.001

Shah, S., Lubeck, E., Zhou, W., & Cai, L. (2017). seqFISH Accurately Detects Transcripts in Single Cells and Reveals Robust Spatial Organization in the Hippocampus. *Neuron*, *94*(4), 752–758.e1. https://doi.org/10.1016/j.neuron.2017.05.008

Ståhl, P. L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., … Frisén, J. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, *353*(6294), 78–82. https://doi.org/10.1126/science.aaf2403

Svensson, V., Teichmann, S. A., & Stegle, O. (n.d.). SpatialDE - Identification of spatially variable genes. https://doi.org/10.1101/143321