



Universiteit
Leiden

The Netherlands

Integration and disentanglement of single-cell and spatial transcriptomics in health and disease

Novella Rausell, C.

Citation

Novella Rausell, C. (2025, May 28). *Integration and disentanglement of single-cell and spatial transcriptomics in health and disease*. Retrieved from <https://hdl.handle.net/1887/4247894>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/4247894>

Note: To cite this publication please use the final published version (if applicable).

INTEGRATION AND DISENTANGLEMENT OF SINGLE-CELL AND SPATIAL TRANSCRIPTOMICS IN HEALTH AND DISEASE

**WITH APPLICATIONS IN AUTOSOMAL DOMINANT
POLYCYSTIC KIDNEY DISEASE**

Claudio Novella Rausell

INTEGRATION AND DISENTANGLEMENT OF SINGLE-CELL AND SPATIAL TRANSCRIPTOMICS IN HEALTH AND DISEASE

**WITH APPLICATIONS IN AUTOSOMAL DOMINANT
POLYCYSTIC KIDNEY DISEASE**

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof. dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op woensdag 28 mei 2025
klokke 10:00 uur

door

Claudio Novella Rausell

geboren te Valencia, Spanje

in 1996

promotoren:	prof.dr. D.J.M. Peters
copromotoren:	dr. A. Mahfouz
promotiecommissie:	prof. dr. B.P.F. Lelieveldt
	prof. dr. A.J. Rabelink
	prof. dr. B. T. Heijmans
	dr. J.P. Roelands
	dr. ir. P.D. Moerland
	<i>Universitair Medisch Centrum Amsterdam</i>
	dr. L. Cantini
	<i>Institut Pasteur, France</i>

Promoter:

Prof.dr. Dorien J.M Peters

Leiden University Medical Center, The Netherlands

Copromoter:

Dr. Ahmed Mahfouz

Leiden University Medical Center, The Netherlands

Chairman:

Rector Magnificus,

Prof. dr.ir. H. Bijl

Leiden University Medical Center, The Netherlands

Committee:

Prof. dr. B.P.F Lelieveldt

Leiden University Medical Center, The Netherlands

Prof. dr. A.J. Rabelink

Leiden University Medical Center, The Netherlands

Prof. dr. B.T. Heijmans

Leiden University Medical Center, The Netherlands

Dr. J.P. Roelands

Leiden University Medical Center, The Netherlands

Dr. ir. P.D. Moerland

Amsterdam University Medical Center, The Netherlands

Dr. L. Cantini

Institut Pasteur, France

This work has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 860977.

Printing: ridderprint.nl

Cover and layout: Claudio Novella Rausell

ISBN: 978-94-6522-243-1

© 2025, C. Novella Rausell, Leiden, the Netherlands.

Problems worthy of attack prove their worth by fighting back.

Piet Hein (Danish polymath; 20th century)

CONTENTS

1	Introduction	1
1.1	Autosomal Dominant Polycystic Kidney Disease	3
1.2	Challenges to be addressed by single-cell and spatial techniques in ADPKD.	4
1.3	Representation learning.	5
1.3.1	Deep representation learning	7
1.3.2	Integration across studies	9
1.3.3	Integration across technologies	9
1.4	Representation disentanglement	10
1.5	Contributions of this thesis	11
2	A comprehensive mouse kidney atlas enables rare cell population characterization and robust marker discovery	13
2.1	Introduction	14
2.2	Results	16
2.2.1	Integrated atlas accounts for technical differences among seven independent studies.	16
2.2.2	Integration highlights annotation inconsistencies across studies . . .	17
2.2.3	Manual curation of annotations significantly improves hierarchy learning	19
2.2.4	MKA accurately classifies unseen cells	22

2.2.5	Mouse Kidney Atlas facilitates the identification of robust cell population markers	25
2.3	Discussion	28
2.4	Limitations of the Study	31
2.5	Acknowledgements.	31
2.6	Author contributions	31
2.7	Methods	32
2.7.1	Key resources table.	32
2.7.2	Resource availability	32
2.7.3	Method details	33
2.8	Supplementary materials	38
3	Spatial transcriptomics reveals injured cells, signature genes, and communication patterns in the cyst microenvironment of Polycystic Kidney Disease	49
3.1	Introduction	50
3.2	Methods	50
3.2.1	Animals	50
3.2.2	10 x Genomics Visium spatial gene expression experiment	51
3.2.3	10 x Genomics Visium spatial gene expression raw data processing	51
3.2.4	Gene set enrichment analysis	51
3.2.5	Spatial pattern analysis	52
3.2.6	Visium spot deconvolution	52
3.2.7	Cyst Border Identification and Gene Expression Models	52

3.2.8	Cell Communication	52
3.2.9	Immunohistochemistry	53
3.3	Results	53
3.3.1	Spatial transcriptomics of wild-type and PKD kidneys and the identification of unsupervised major cell type clusters	53
3.3.2	Deconvolution defines cell-type composition of the spots in wild-type and PKD kidneys	55
3.3.3	Identification of spatial gene expression patterns in wild-type and PKD kidneys	55
3.3.4	Characterization of gene expression and cell type distribution in the cyst microenvironment	57
3.3.5	Unraveling cell-cell communication events in the cyst microenvironment	57
3.3.6	Identification and validation of key ligand-receptor pairs	60
3.4	Discussion	63
3.5	Disclosures.	65
3.6	Funding	65
3.7	Acknowledgements.	65
3.8	Author contributions	65
3.9	Data sharing statement	65
3.10	Supplementary Methods	67
4	Early PKD alterations at the single-cell level	71
4.1	Introduction	72

4.2	Results	73
4.2.1	Experiment design and power analysis	73
4.2.2	Antibody-based multiplexing	74
4.2.3	Probe-based multiplexing.	76
4.3	Discussion	80
4.4	Methods	80
4.4.1	Mouse model and experimental groups	80
4.4.2	Single-nuclei isolation and library preparation	81
4.4.3	Single-nuclei RNA-seq data processing	82
4.4.4	Over representation analysis	82
4.4.5	Transcription factor and pathway activities	83
5	Integrative learning of disentangled representations from single-cell RNA-sequencing datasets	85
5.1	Introduction	86
5.2	Methods	87
5.2.1	spVIPES inference	87
5.2.2	spVIPES generative process	88
5.2.3	Learning objective	89
5.2.4	Network architecture	89
5.2.5	Model training	89
5.2.6	Evaluation metrics	90
5.2.7	Simulated dataset experiment.	90

5.2.8	Multi-species integration experiment	91
5.2.9	Acute kidney injury experiment.	91
5.2.10	Stimulated PBMCs experiment	91
5.3	Results	92
5.3.1	spVIPES model overview.	92
5.3.2	Performance on a simulated dataset.	93
5.3.3	spVIPES significantly improves integration of distant species	94
5.3.4	spVIPES accurately identifies genes associated with maladaptive regeneration after Acute Kidney Injury	96
5.3.5	spVIPES helps identify proximal tubule populations associated with an injury-recovery transition after short IRI	97
5.3.6	spVIPES identifies cells and genes associated with stimulation	98
5.4	Discussion	101
5.5	Data availability statement	103
5.6	Acknowledgements.	103
5.7	Supplementary materials	104
6	Discussion	107
6.1	Cross-study reference building	108
6.1.1	Standardizing quality control	108
6.1.2	Leveraging biological information to enhance cross-study integra- tion and cell type identification	109
6.1.3	Improving reference atlases.	110

6.2	Cell-to-cell communication	111
6.2.1	Detecting changes in communication	111
6.2.2	mRNA-derived CCC is insufficient	112
6.3	Disentanglement	113
6.3.1	Using disentangled representations	113
6.3.2	Unsupervised disentanglement of datasets with unpaired features . .	113
6.3.3	What affects disentanglement?	114
6.3.4	Evaluating disentanglement.	115
6.3.5	Is disentanglement possible?	115
6.4	Future Perspectives and Biological Implications	116
	References	118
	Abstract	137
	Samenvatting	139
	Resumen	141
	Resumen	143
	Cell type glossary	149
	Acknowledgments	151