

Master's thesis in Bioinformatics

CHARACTERIZING HUMAN BROWN
ADIPOCYTES IN WHITE ADIPOSE
TISSUE WITH SINGLE-CELL
TRANSCRIPTOMICS

Author:

Natalia Vargas Díaz

Student number: 202402137

Supervisor:

Yonglun Luo

PhD, Professor

Department of biomedicine

Aarhus university

June 2026

Abbreviations	i
Declaration of GAI tools	iii
Acknowledgements	iv
ABSTRACT	v
1. INTRODUCTION AND BACKGROUND	1
1.1. Obesity as a global health concern	1
1.2. Human adipose tissues	1
1.3. Brown adipocytes: a potential target to treat obesity	3
1.4. Pheochromocytoma and paraganglioma patients as a model to study brown adipocytes	5
1.5. Single-cell RNA sequencing	7
1.5.1. Sparsity and dropouts.....	7
1.5.2. Batch effect and integration	8
Canonical Correlation Analysis.....	8
Reciprocal Principal Component Analysis.....	8
Harmony.....	9
Single cell Variational Inference	9
1.5.3. Data representation	9
Principal Component Analysis (PCA).....	9
Uniform Manifold Approximation and Projection (UMAP).....	10
1.5.4. Cell annotation.....	10
Manual annotation.....	10
Automated annotation	10
1.6. Aim of the project	11
2. METHODS	12
2.1. Public data: sample description.....	12
2.2. Collection of human samples and sample description	12
2.3. Library preparation and sequencing.....	12
2.4. Workflow for scRNA-seq analysis.....	13
2.5. Demultiplexing.....	14
2.5.1. Cell-snp lite.....	14
2.5.2. Vireo	14
2.6. Data pre-processing.....	14
2.7. Data processing	15
2.7.1. Normalization and batch effect detection	15

2.7.2.	Integration.....	16
2.7.3.	Clustering and cell annotation	17
2.7.3.1.	Annotation of the major populations of the case data	17
	Probability-based annotation approach	17
	Manual annotation.....	17
	Integrating both approaches	17
2.7.3.2.	Annotation of the subpopulations of the case data	18
	Manual annotation.....	18
2.7.3.3.	Annotation of the major populations of the control data.....	18
	Label transfer	18
2.7.3.4.	Annotation of the subpopulations of the control data.....	18
2.8.	Adipocyte subset	18
2.9.	Compositional analysis	18
2.10.	Differential Gene Expression	19
2.11.	Gene-set expression analysis.....	19
2.12.	Trajectory analysis.....	20
2.13.	Regulatory analysis	21
2.14.	Cell-cell interactions	22
3.	RESULTS.....	24
3.1.	Pre-processing	24
3.1.1.	Case data.....	24
3.1.2.	Control data	25
3.2.	Data processing	25
3.2.1.	Major cell populations of case data	25
3.2.2.	Cell subpopulations of case data.....	26
3.2.3.	Major cell populations of control data	29
3.2.4.	Cell subpopulations of control data	29
3.3.	Compositional analysis	30
3.3.1.	Across tissue depots.....	30
3.3.2.	Across conditions.....	33
3.4.	Differential expressed genes	34
3.5.	Gene-set analysis.....	35
3.6.	Trajectory analysis	37
3.7.	Regulatory analysis	38

3.8. Cell-cell interactions	39
3.8.1. Omental depots (case versus control)	40
4. DISCUSSION	42
Identification of brown-like adipocytes in adipose tissue	42
Functional analysis confirms brown-like adipocytes population	44
Trajectory analysis suggests a transition toward brown-like adipocytes	45
Signalling pathways associated with adipocytes	46
Limitations and future perspectives	47
5. CONCLUSIONS	49
Bibliography	50
Supplementary 1	60
Supplementary 2	61
Supplementary 3	62
Supplementary 4	63
Supplementary 5	64
Supplementary 6	65
Supplementary 7	66
Supplementary 8	67

Abbreviations

- **APC**: Adipocyte Progenitor Cells.
- **ATP**: Adenosine Triphosphate.
- **ASW**: average silhouette width.
- **AUC**: area under the recovery curve.
- **BMI**: Body Mass Index.
- **BAT**: brown adipose tissue.
- **BMP**: bone morphogenetic proteins.
- **cAMP**: cyclic Adenosine Monophosphate.
- **CCA**: canonical correlation analysis.
- **C/EBP**: CCAAT-enhancer binding protein.
- **cLISI**: cell-type Local Inverse Simpson's Index.
- **EC**: endothelial cells.
- **FA**: fatty acids.
- **FAP**: fibroblasts and adipocyte progenitors.
- **FDG-PET**: [18F]-fluodeoxyglucose positron emission tomography.
- **FDR**: False Discovery Rate.
- **FFA**: free fatty acids.
- **GTP**: Guanosine Triphosphate.
- **iLISI**: integration Local Inverse Simpson's Index.
- **LEC**: lymphatic endothelial cells.
- **MSL**: mesothelial cells.
- **NE**: norepinephrine.
- **NST**: non-shivering thermogenesis.
- **PC**: principal component.
- **PCA**: Principal Component Analysis.
- **PET**: Positron Emission Tomography.
- **PPARA**: Peroxisome Proliferator-Activated Receptor Alpha.
- **PPARG**: Peroxisome Proliferator-activated Receptor gamma.
- **PPARGC1A**: PPARG coactivator 1 alpha.
- **PRDM16**: PR Domain Containing 16.
- **PPGL**: pheochromocytoma and paraganglioma.
- **rPCA**: reciprocal Principal Component Analysis.
- **SCENIC**: Single-Cell rEgulatory Network Inference and Clustering.
- **scIB**: single-cell Integration Benchmarking.
- **scRNA-seq**: single cell RNA sequencing.
- **scVI**: single cell Variational Inference.
- **snRNA-seq**: single nuclei RNA sequencing.
- **SNP**: Single Nucleotide Polimorfism.
- **SVF**: stromal vascular fraction.
- **TF**: transcription factor

- **UCP1**: Uncoupling Protein 1.
- **UMAP**: uniform manifold approximation and projection.
- **UMI**: Unique Molecular Identifier.
- **Vireo**: Variational Inference for Reconstructing Ensemble Origins.
- **VEC**: vascular endothelial cells.
- **VEGF**: Vascular Endothelial Growth Factor.
- **WAT**: white adipose tissue.
- **WHO**: World Health Organization.

Declaration of GAI tools

I used generative artificial intelligence (GAI) to complete this project, specifically OpenAI GPT-5.3-mini, Perplexity AI (2026) and Anthropic Claude Opus 4.8.

1. Perplexity AI and OpenAI GPT-5.3-mini were used for feedback on own text and alternative ways of formulating text: rephrasing sentences and restructuring paragraphs.
2. Anthropic Claude Opus 4.8 was used for programming tasks.
3. Perplexity AI was additionally used to support me in my reading process by summarizing papers on key statements.

There has been no other usage of GAI for this project besides than for the purposes stated above.

Acknowledgements

There are many people I would like to thank for accompanying me over these past two years, both academically and personally. However, if I tried to say everything I want to say, I would probably need an extension similar to this thesis (including the bibliography), so I will try to keep it brief.

First, I would like to thank Yonglun Luo (Alun) for his trust and for the opportunity to carry out this project. To Lin Lin, because without her this work would not have been possible, for all her knowledge, and especially for being an exceptional mentor. This project reignited a spark in me that I thought had gone out: the desire to learn, to ask questions, and to understand how the world works.

I would especially like to thank Sowmiya Kalaiselvan and Aisha Nadukkandy for their work on library preparation and nuclei isolation. I would also like to thank Esben Søndergaard for his support throughout the project and for his contribution to the cohort sample collection together with Lars Rolighed, Andreas Ebbenhøj, and Amanda Bæk.

I would also like to thank Mohammed Hassan, who was my first mentor and gave me the tools and foundations I needed to begin this project. And to the members of the DREAM lab, for creating such a warm and welcoming environment; I have truly enjoyed the Friday breakfasts and the shared anecdotes.

Moving out on my own at 22, to a country I had never visited before, with an unfamiliar language and very few certainties, was not an easy experience. Having a strong support network has been essential, and I feel very fortunate to be surrounded by people who have accompanied, supported, and held me through every decision (some more daring than others) that I have made. Even from the distance, they have continued to be present throughout these two years, which would not have been possible without my parents, my sisters Elena and Sofia, and my friends María, Nerea, Paula, Noe, and Estrella.

I cannot finish without thanking the person who has made my time in this country kinder and much more enjoyable. To my friend Emilio Sopprani, for all the afternoons spent discussing science, for reading every single piece of work I have produced so far and giving his honest opinion, which has helped me grow as a scientist, and for being a constant source of support throughout this time, beyond academic matters. I hope we continue to share life and our passion for science together.

ABSTRACT

Obesity is a chronic multifactorial disease associated with numerous comorbidities, including cardiovascular disease, type 2 diabetes, and several forms of cancer. Despite the availability of lifestyle interventions, pharmacological treatments, and bariatric surgery, long-term weight management remains challenging. Brown adipose tissue (BAT) has emerged as a promising therapeutic target due to its ability to increase energy expenditure through non-shivering thermogenesis mediated by uncoupling protein 1 (UCP1). Understanding the molecular mechanisms underlying brown adipocyte differentiation and activation may contribute to the development of new obesity treatments.

This study aimed to characterize brown adipocytes in human adipose tissue and identify molecular features associated with the differentiation of brown adipocytes. Single-nucleus RNA sequencing (snRNA-seq) was performed on adipose tissue samples collected from five patients diagnosed with catecholamine-secreting paraganglioma, a condition known to promote BAT recruitment through chronic adrenergic stimulation. Samples were obtained from three adipose tissue depots: subcutaneous, visceral (particularly, omental), and perirenal. Publicly available snRNA-seq datasets from non-paraganglioma donors were incorporated as controls. Following quality control, demultiplexing, doublet removal, and batch-effect correction, cell populations were identified through clustering and were annotated following different approaches. Downstream analyses included compositional analysis, differential gene expression, gene-set enrichment, trajectory inference, gene regulatory network reconstruction, and cell-cell communication analysis.

A comprehensive cellular atlas of human adipose tissue was generated, revealing the presence of multiple immune, vascular and adipocyte populations across depots. Within the adipocyte compartment, distinct adipocyte subpopulations were identified, including a population displaying transcriptional features consistent with brown adipocytes. These cells exhibited expression patterns associated with thermogenic activity and mitochondrial metabolism such as *UCP1*. Functional analyses suggested enrichment of marker genes related to glycolysis, oxidative phosphorylation, and fatty acid metabolism. Trajectory inference indicated the existence of transitional cellular states starting in control-specific adipocytes and ending in brown-like adipocytes. Regulatory network analysis identified candidate transcriptional regulators potentially involved in brown adipocyte identity, while cell-cell communication analyses highlighted signalling pathways that may contribute to adipocyte differentiation.

Overall, this study provides a detailed single-nucleus transcriptomic characterization of brown adipocytes in human adipose tissue. The findings contribute to the understanding of brown adipocyte biology and identify candidate molecular pathways involved in the differentiation of brown adipocytes. These results may support future efforts to promote thermogenic adipocyte recruitment as a therapeutic strategy for obesity, after experimental validation.