CSE 527 Project Proposal
Re-analysis of the Human Endometrial Cell Atlas

Olaf B. W. H. Corning[1] and Garrett McPheron[1]

[1]Department of Bioengineering, University of Washington

October 2024

Background

The acceleration of machine learning methods in the modern age has rapidly opened many doors for biomedical research, including accelerating high throughput analyses, the creation of cell atlases, and precision medicine. These methods have been applied across biomedical applications such as creating the Allen Institute Atlases, developing individualized cancer immunotherapies, and the source of the focus dataset, the Human Endometrial Cell Atlas (HECA) [1].

The endometrium is the innermost epithelial lining of the uterus and provides a thick granular tissue layer for embryonic implantation while preventing adhesions to the myometrium. It is composed of two layers: the functional columnar epithelium, which is built up and subsequently shed during menstruation, and the stromal basal layer, which contains the progenitors that replace the shed functional layer. The process of building up the functional columnar epithelium — decidualization — enables implantation of the embryo into the endometrial surface and coordinates the invasion of extra-embryonic trophoblast lineages. As part of decidualization, endometrial fibroblasts cells can differentiate into decidual secretory cells that have the ability to regulate trophoblast invasion, to resist oxidative stress, and to protect the placental semi-allograft against maternal immune responses [2]. The function of the endometrium is heavily regulated by hormones, such as progesterone and estradiol, and pathways such as cAMP [3, 4, 5].

Several diseases impact the endometrium, including adenomyosis, endometrial hyperplasia, endometrial cancer, asherman's syndrome, and endometriosis. Endometriosis, the growth of endometrial tissue outside the uterus, affects up to 10% of women between the ages of 15 and 44 [6], causing pelvic pain, decreased fertility, and diminished quality of life [7]. Improved understanding of endometrial development, regeneration, cell niches, and function can improve health outcomes for individuals with endometrial diseases and can inform practices in regenerative medicine. Additionally, this work can contribute towards reducing the well documented disparity in research on women's health [8].

Method

HECA, a high-resolution single-cell reference atlas (313,527 cells) combining published and new endometrial single-cell transcriptomics datasets of 63 women with and without endometriosis. In our re-analysis of the HECA data, we will be using two R packages to visualizes connections

of the scRNA-seq data and to try to pull out cell state and gene regulatory network data. The former will be achieved using cellTree, a package that performs inference and visualization of scRNA-seq in a hierarchal tree structure using a latent dirichlet allocation model. The latter will be achieved using SCENIC, a package for simultaneous gene regulatory network reconstruction via a regression-per-target approach, motif discovery, and quantifying regulon enrichment from scRNA-seq data. [9, 10].

Goals

With both of us coming from a more biology-heavy background with more limited experiences of machine learning algorithms, our project goals are getting more experience applying newer and less widely used methods to interrogate large biological datasets. In terms of project questions, we are looking to gain a better understanding of the gene regulatory networks involved in endometriosis and healthy endometrium tissue. This dataset lacks ChIP-seq data which makes traditional gene regulatory network construct challenging. Furthermore, we hope to test whether generic scRNA-seq algorithm performance can be comparable or even outperform custom-made algorithms used for cell atlases.

Challenges

There are several challenges that are anticipated: Batch effects, cell cycling effects, perturbations due to hormonal signaling and or uterine cycle, and challenges with connecting our results to the HECA as a pseudo-ground truth. By integrating a priori knowledge of gene pathway associations, e.g., genes associated with cell cycling, some of these effects can be regressed out. Additionally, we believe the methods selected are robust enough to manage some of these challenges as they can be of biological relevance, e.g. hormonal signalling. Should there be an insurmountable incompatibility, there are other scRNA-seq analysis packages and algorithms that we might revert to if we run into too many issues, such as scVI.

Timeline

Milestone – October 31st (Thurs): Integrate Peer Review Feedback
Milestone – November 1st (Fri): Processing and integrating the data into local environments
Milestone – November 6th (Wed): Gain preliminary results from SCENIC
Due Date – November 7th (Thurs): Project Checkpoint Writeup Due
Milestone – November 15th (Fri): Gain preliminary results from CellTree
Milestone – November 17th (Sun): Integrate Final Peer Review Feedback
Due Date – November 26th (Tues): Final Report Drafts Due.
Due Date – December 5th (Thurs): Final Project Poster Session.
Due Date – December 10th (Tues): Final Project Reports.

Resources

HECA: Nature article
cellTree: Bioconductor page
SCENIC: Nature article

References

[1] Magda Marečková, Luz Garcia-Alonso, Marie Moullet, Valentina Lorenzi, Robert Petryszak, Carmen Sancho-Serra, Agnes Oszlanczi, Cecilia Icoresi Mazzeo, Frederick CK Wong, Iva Kelava, et al. An integrated single-cell reference atlas of the human endometrium. *Nature Genetics*, pages 1–13, 2024.

[2] Birgit Gellersen, Ivo Brosens, and Jan Brosens. Decidualization of the human endometrium: Mechanisms, functions, and clinical perspectives. *Seminars in Reproductive Medicine*, 25(6):445–453, Nov 2007.

[3] Dominique de Ziegler, Renato Fanchin, Béatrice de Moustier, and Carlo Bulletti. The hormonal control of endometrial receptivity: estrogen (e2) and progesterone. *Journal of reproductive immunology*, 39(1-2):149–166, 1998.

[4] Anoop K Brar, Graeme R Frank, Cherie A Kessler, Marcelle I Cedars, and S Handwerger. Progesterone-dependent decidualization of the human endometrium is mediated by camp. *Endocrine*, 6:301–307, 1997.

[5] Birgit Gellersen and Jan Brosens. Cyclic amp and progesterone receptor cross-talk in human endometrium: a decidualizing affair. *The Journal of endocrinology*, 178(3):357–372, 2003.

[6] Johns Hopkins Medicine - Medical Conditions: Endometriosis.

[7] Neil P Johnson, Lone Hummelshoj, G David Adamson, Jörg Keckstein, Hugh S Taylor, Mauricio S Abrao, Deborah Bush, Ludwig Kiesel, Rulla Tamimi, Kathy L Sharpe-Timms, et al. World endometriosis society consensus on the classification of endometriosis. *Human reproduction*, 32(2):315–324, 2017.

[8] Alicia Armstrong and Yvonne T Maddox. Health disparities and women's reproductive health. *Ethnicity & Disease*, 17:4–7, 2007.

[9] David A duVerle, Sohiya Yotsukura, Seitaro Nomura, Hiroyuki Aburatani, and Koji Tsuda. Celltree: an r/bioconductor package to infer the hierarchical structure of cell populations from single-cell rna-seq data. *BMC bioinformatics*, 17:1–17, 2016.

[10] Sara Aibar, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, Jean-Christophe Marine, Pierre Geurts, Jan Aerts, et al. Scenic: single-cell regulatory network inference and clustering. *Nature methods*, 14(11):1083–1086, 2017.