Finding the connections and channels of communication between types of endocrine cells in the human gut

Liron Gershuny
November 2020
Advisor: Dr. Danny Ben-Zvi

Research question:

Our research interest focuses on mapping and finding the channels of communication between different types of endocrine cells from the human gut according to the hormones, cytokines, and proteins secreted by the cell.

Background:

The enteroendocrine system produces peptide hormones secreted in the stomach, colon, and small intestine, comprising <1% of the epithelial cells[1], that regulate metabolism through the coordination of digestion, absorption, nutrient disposal, and appetite. Gut hormones have a wide range of targets, both within and outside the gastrointestinal tract, several hormones have more than one physiological role and most physiological roles are regulated by more than one hormone. Circulating gut hormones often trigger responses in target tissues that express the receptor for the hormone, lasting from several minutes to a few hours, reflecting both the temporal profile of the hormone in the bloodstream and the complement of receptors expressed by the target cells. Enteroendocrine hormones include gastrin, GIP, Glp1, Glp2, Ghrelin, Pyy, Serotonin, Somatostatin, and others. The importance of these hormones is reflected in many drugs that agonize or antagonize the actions of these hormones.

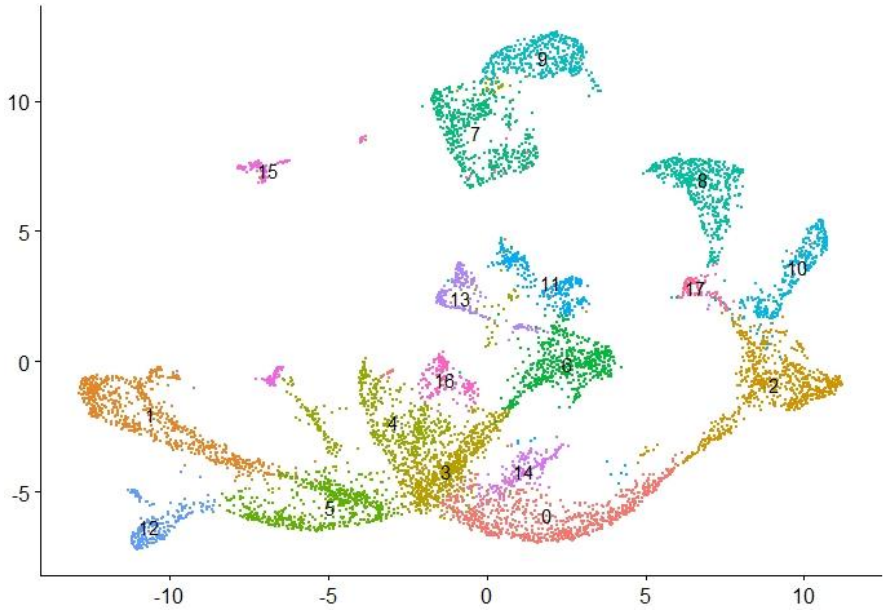
Gut hormones are produced by enteroendocrine cells (EECs), which are scattered throughout the epithelial layer of the gut wall. EECs sense the contents of the intestinal lumen via sensory proteins, including G protein-coupled receptors (GPCRs), transporters, and ion channels. Although some inducers of hormone secretion have been described in mice, there has been no experimental model to systematically assess such secretagogues for human endocrine cells[1].

EECs do not form glands but are spread throughout the digestive tract which makes studying these cells challenging. In particular, there is a debate on the exact number of endocrine cell types in the gut and it appears that unlike endocrine glands, there is a spectrum of cell types. With the advent of single-cell analysis, research of these cells became more accessible.

We propose to use data from single-cell sequencing of EECs in humans to study the communication between EECs. Through this analysis, we will treat the EECs as a distributed endocrine organ, and determine the interactions between its various cell types.

Work plan:

1. Single Cell mRNA clustering- This stage will be done by downloading the data from GEO database, arrange it and filter it for our needs, preform PCA processing and then clustering it to different groups. (October 2020)



Umap of data downloaded from Beumer et. al analyzed by Seurat. Currently at phase 2 to optimizing the clusters according to protocol and identify the biological significance of each cluster.

2. Clustering Characterization- We will decipher each cluster by significant gene expressions using previous researches. (November-December 2020)
3. Description of the communication network between the cells- Generating visual interpretation of the data as network[4] in which cell types are the nodes and drawing edges between the types that share biological correlation. (January-February 2021)
4. Buffer- As a result of the exam period and unexpected difficulties. (March 2021)
5. Integration with data generated from the lab- Combining former data from the lab, produce insights and a unified view of the data. (April-May 2021)
6. Network Analysis – module separation & Biological explanation- At this stage, we will figure the strength of interactions between the endocrine gut cells and find a biological explanation for the results. As a result, we will try and find if the endocrine gut cells are independent or not. (June-July 2021)

Total of 10 months.

Data:

The data for this research was taken from Joep Beumer[1] et.al. study which performed experiments on human intestinal mucosal biopsies that were obtained from patients undergoing colonoscopy at Addenbrooke's Hospital, Cambridge, UK.
Organoids were dissociated to single cells using a 10-min incubation with TrypLE and repeated mechanical disruption by pipetting. Single-cell sequencing was performed according to the Sort-seq method[2]. Reads were mapped to the human GRCh37 genome assembly.

All bulk and single-cell RNA sequencing data of Joep Beumer's study have been deposited in the Gene Expression Omnibus (GEO) under accession code GSE146799.

Data is formed by a CSV data frame in which every column represents a cell type and each row represents a gene.

Computational tools:

Most of the computational work will be written primarily in R language and will be focused on the analysis of single-cell sequencing under the package of 'Seurat'. In particular, we will perform:

a. Pre-processing using PCA- a statistical method that allows us to reduce the dimension of the data we work with
b. Clustering- classify groups of cells that secrets a similar group of hormones
c. Data normalization- normalization of the data generated from the lab and by Beumer et.al. to allow comparison. We will consider batch effects in differences in sequencing depth used by normalization algorithms in Seurat
d. Gene enrichment of specific clusters to determine the identity of each cluster
e. Data mining of papers in 'Pub-Med' to associate each cluster with cell identity

Later on, we would infer cell-cell communication using the CellPhoneDB[3] protocol published in April 2020 using Python language whose input is a single-cell sequencing and a chart of ligands and receptors to map all possible interactions. I will integrate CellPhoneDB within the Seurat package to perform this analysis, and model the interactions as a directed graph.

References:

1. High-Resolution mRNA and Secretome Atlas of Human Enteroendocrine Cells
   Joep Beumer, Jens Puschhof, Julia Bauza´ -Martinez, Adriana Martı´nez-Silgado, Rasa Elmentaite, Kylie R. James, Alexander Ross, Delilah Hendriks, Benedetta Artegiani, Georg A. Busslinger, Bas Ponsioen, Amanda Andersson-Rolf, Aurelia Saftien, Charelle Boot, Kai Kretzschmar, Maarten H. Geurts, Yotam E. Bar-Ephraim, Cayetano Pleguezuelos-Manzano, Yorick Post, Harry Begthel, Franka van der Linden, Carmen Lopez-Iglesias, Willine J. van de Wetering, Reinier van der Linden, Peter J. Peters, Albert J.R. Heck, Joachim Goedhart, Hugo Snippert, Matthias Zilbauer, Sarah A. Teichmann, Wei Wu, and Hans Clevers. Cell 181, 1291–1306, June 11, 2020.

2. A Single-Cell Transcriptome Atlas of the Human Pancreas
Mauro J. Muraro, Gitanjali Dharmadhikari, Dominic Gru¨ n, Nathalie Groen, Tim Dielen, Erik Jansen, Leon van Gurp, Marten A. Engelse, Francoise Carlotti, Eelco J.P. de Koning, and Alexander van Oudenaarden. Cell Systems 3, 385–394, October 26, 2016.

3. CellPhoneDB: inferring cell–cell communication from combined expression of multi-subunit ligand–receptor complexesMauro
Mirjana Efremova, Miquel Vento-Tormo, Sarah A. Teichmann and Roser Vento-Tormo. NATURE PROTOCOLS |VOL 15 | APRIL 2020 | 1484–1506.

4. How to visually interpret biological data using networks
Daniele Merico, David Gfeller & Gary D Bader. Nature biotechnology volume 27 number 10 october 2009.