Multicellular molecular characterisation of inflammatory bowel disease

Group Leaders:
Carl Anderson, Wellcome Sanger Institute (co-supervisor)
Julio Saez-Rodriguez, EMBL-EBI (host supervisor)

MOTIVATION

Inflammatory bowel disease (IBD) encompasses multiple chronic gastrointestinal disorders mediated by the immune system via a complex set of interacting risk factors including genetic predisposition and environmental exposures(Rudbaek *et al*, 2024; Graham & Xavier, 2020; Sazonovs *et al*, 2022). Single-cell RNA sequencing (scRNAseq) has delivered a high-resolution taxonomy of cell types and states along the gastrointestinal tract in both health and disease. This cell taxonomy represents a reference from which pathogenic cellular remodeling can be traced by performing comparative analyses between distinct patient groups of clinical relevance. In IBD, single cell studies have found associations of cell-type specific gene programs and tissue compositional changes to disease progression and treatment response (Zheng *et al*, 2021; Krzak *et al*, 2023), however the study of multicellular coordinated processes from single-cell transcriptomics data sets is still limited. While major efforts inferring potential cell-cell communication events via ligand-receptor interactions exist (Dimitrov *et al*, 2023), multicellular definitions of tissue function where the molecular characteristics of a cell are understood as responses conditioned to a higher order coordinated process with all other cells are generally lacking in the single-cell transcriptomics field (Ramirez Flores *et al*, 2024).

A multicellular analysis of IBD physiopathology could increase our understanding of the division of labor of cells in the intestinal mucosa and generate representations of cellular dependencies that could be the basis of mechanistic hypotheses paving the way of novel therapies. In addition, multicellular descriptions of tissues facilitate cross-condition comparisons of multiple patient cohorts by summarizing the molecular profiles of collections of cells into interpretable joint programs. This approach changes the focus of meta-analyses from a cell-centric perspective to a tissue-centric one, allowing for a direct comparison of inflammatory processes across the clinical spectrum of IBD. We propose to combine (i) ongoing single-cell integration efforts of the gastrointestinal tract, (ii) large scale profiling of IBD patient cohorts with comprehensive clinical data, (iii) novel multicellular integration and (iv) microbiome inference computational methods to meta-analyze disease processes of IBD from a multicellular tissue-centric perspective.

SPECIFIC OBJECTIVES

Our proposed work aims to answer the following questions:

1) What are the molecular footprints of the multicellular processes associated with IBD?
2) To what extent IBD disease processes are the product of specific cell-types or tissue changes in composition and multicellular coordination?

3) How conserved are the inferred molecular changes in IBD patients across distinct and independent human single-cell studies of ulcerative colitis and Crohn's disease of variable clinical status and tissue sampling locations?
4) Can microbiome compositions be deconvoluted from scRNAseq data and associated with multicellular responses of the host?

We propose to shed light on these questions related to the meta-analysis of IBD from a multicellular perspective by establishing a computational framework that expands our previous work on multicellular integration (Ramirez Flores *et al*, 2023) and repurposing a large-scale profiling of IBD patients. The overarching aim of our project is to generate multicellular molecular descriptions of IBD disease processes from the combination of multiple patient cohorts available in cellxgene, the Broad Single-Cell data portal, and other public repositories focusing on cross-condition single-cell transcriptomics data sets. We will initially infer multicellular programs and tissue taxonomies of disease from a large-scale patient cohort of IBD patients to generate a patient map that describes the molecular diversity of tissue samples across cell-types. Then, we will project smaller datasets into this patient map to assess its generalizability and power to predict treatment response. Finally, we will estimate microbiome compositions from single cell atlases using novel methods for the identification of microbial reads from scRNAseq data and associate them with multicellular disease programs in the gut (Robinson *et al*, 2024).

Specifically, we will address the following aims.

Aim 1: Infer a multicellular patient map of IBD informed by clinical covariates and tissue sampling location.

Aim 2: Infer putative bacterial compositions of the microbiome of tissue samples from independent patient cohorts.

Aim 3: Combine the descriptive power of multicellular programs and bacterial compositions to refine unsupervised patient groupings of IBD.

This multidisciplinary project combines the expertise on the genetics and molecular pathology of IBD of the Anderson lab, that has profiled large patient cohorts across scales and -omics, with the computational systems biology experience of the Saez-Rodriguez lab in combining the information of independent sources and data types under the framework of mechanistic modeling. This first interaction between the Anderson and Saez-Rodriguez groups complements the objectives of ongoing efforts of individual labs in the study of multicellular processes in IBD and their associations with the microbiome.

REFERENCES

Bibliography

Dimitrov D, Schäfer PSL, Farr E, Rodriguez Mier P, Lobentanzer S, Dugourd A, Tanevski J, Ramirez Flores RO & Saez-Rodriguez J (2023) LIANA+: an all-in-one cell-cell communication framework. *BioRxiv*

Graham DB & Xavier RJ (2020) Pathway paradigms revealed from the genetics of inflammatory bowel disease. *Nature* 578: 527–539

Krzak M, Alegbe T, Taylor DL, Ghouraba MH, Strickland M, Satti R, Thompson T, Arestang K, Przybilla MJ, Ramirez-Navarro L, *et al* (2023) Single-cell RNA sequencing reveals dysregulated cellular programmes in the inflamed epithelium of Crohn's disease patients. *medRxiv*

Ramirez Flores RO, Lanzer JD, Dimitrov D, Velten B & Saez-Rodriguez J (2023) Multicellular factor analysis of single-cell data for a tissue-centric understanding of disease. *eLife* 12

Ramirez Flores RO, Schäfer PSL, Küchenhoff L & Saez-Rodriguez J (2024) Complementing Cell Taxonomies with a Multicellular Analysis of Tissues. *Physiology (Bethesda)* 39: 0

Robinson W, Stone JK, Schischlik F, Gasmi B, Kelly MC, Seibert C, Dadkhah K, Gertz EM, Lee JS, Zhu K, *et al* (2024) Identification of intracellular bacteria from multiple single-cell RNA-seq platforms using CSI-Microbes. *Sci Adv* 10: eadj7402

Rudbaek JJ, Agrawal M, Torres J, Mehandru S, Colombel J-F & Jess T (2024) Deciphering the different phases of preclinical inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol* 21: 86–100

Sazonovs A, Stevens CR, Venkataraman GR, Yuan K, Avila B, Abreu MT, Ahmad T, Allez M, Ananthakrishnan AN, Atzmon G, *et al* (2022) Large-scale sequencing identifies multiple genes and rare variants associated with Crohn's disease susceptibility. *Nat Genet* 54: 1275–1283

Zheng HB, Doran BA, Kimler K, Yu A, Tkachev V, Niederlova V, Cribbin K, Fleming R, Bratrude B, Betz K, *et al* (2021) A Treatment-Naive Cellular Atlas of Pediatric Crohn's Disease Predicts Disease Severity and Therapeutic Response. *medRxiv*