Single-Cell Transcriptomic Analyses

Group Convenor: Dr Sascha Ott
External Partner: Dr Sam Clokie, West Midlands Regional Genetics Laboratory

In recent years it has become possible to obtain quantitative measurements of gene expression levels (transcript numbers) for thousands of individual cells in a single assay. This enables researchers to experimentally process tissue samples and computationally decompose the tissue into individual cell types. It also enables the detection of gene expression changes that are only occurring in one or few cell types and would, therefore, not be detectable if traditional, aggregate measurements across all cells in a tissue sample were taken instead.

This technological advance has created considerable potential for clinical applications, but also brought with it statistical challenges. Examples of relevant questions that can, in principle, be answered include “How many cancerous cell types are present in this tumour?”, “Are tumour-infiltrating killer cells in an activated or inactivated state?”, or “Is the uterine lining of recurrent miscarriage patients different in some way?”.

On the statistical side, the main challenges are 1) the sparseness of the data as each cell is only covered by typically a few thousands reads for about 30,000 genes implying that the measurement for most genes is zero in any given cell, 2) distinguishing unwanted variability as a result of lab procedures from the clinically relevant signal, 3) managing the “computational variability” coming about by a host of different parameter-dependent methods yielding significantly different results.

In the project, we are going to analyse the performance of a range of dimensionality reduction techniques, methods for mapping data sets to each other, methods for mapping data of individual cells to a given library of single-cell data, and investigate how these different aspects of data analysis depend on each other, and how these dependencies are best resolved in order to avoid going in circles in this type of data analysis.