Single-cell Analysis: Biology at High Resolution

Every cell in the human body contains exactly the same sequence of genes. Nevertheless, in each cell different genes are active and these genes determine the cell type and the behavior of the cell. Only recently the analysis of single cells has become technically possible. With this one can, for example, predict the further development of stem cells or identify cancer cells in an early stage.

For some years now, the idea of personalized medicine has been present in science and media. It means a patient specific treatment. Thus, one has realized that people react very differently to active pharmaceutical ingredients. One major factor is the human genome - a double helix of genes (DNA), which is unique for every person and present in every single cell. At a higher level of resolution even cells from the same organism look differently. The reasons can be errors in cell division or gene mutation due to radiation. More often than not, cell-to-cell differences occur naturally: In each cell type (e.g. skin cells and neurons) different genes are active. Only these genes are transcribed to RNA and further translated to functional proteins. The active genes determine the expression profile of a cell.

Fewer Cells, More Insights
In the past, gene expression was typically determined for an aggregation of multiple cells. A high amount of genetic material was necessary to get reliable measurement results with low technical noise. Only recently, technologies have become extremely sensitive and the analysis of single cells is now possible. Instead of a mean expression profile one can determine differences between cells. Moreover, cells can be assigned to specific cell types, even if the shapes of the cells or marker genes are undistinguishable. Especially rare cell types become available for the analysis - at least if it is possible to separate the cells from the tissue. Thus, fewer cells provide more insights. The journal Nature nominated single-cell sequencing as method of the year 2013.

Methods to Analyze Single Cells
One possibility to quantify gene activity is to determine the amount of messenger RNA (mRNA) in the cell.
An important technology is qPCR (quantitative polymerase chain reaction). Similarly to a normal cell division, mRNA is amplified and fluorescence markers are introduced in multiple cycles. Measurements of the fluorescence intensity then provide information about the amount of mRNA present in the cells.

To visualize (and quantify) genes or gene expression in cells one can use the FISH technology (fluorescence in situ hybridization). A fluorescence marker binds to the genes on the chromosomes or to the mRNA, respectively. Fluorescence microscopy then yields an image of the genes and their expression (Fig. 1).

CyTOF (cytometry by time-of-flight) is a variant of mass spectrometry and is produced by DVS Sciences. Over 30 proteins can be quantified simultaneously in single cells. Therefore, antibodies are labeled with heavy metal ions that bind target-specifically to the proteins. The cells are then introduced into the spectrometer by droplet nebulization. For each separated component one determines the time-of-flight. With this one can identify the respective protein and determine its amount in the cell.

**Results from Single-Cell Analysis**

Cells divide into daughter cells, they differentiate into specific cell types or go to apoptosis. These cell-decisions are an important application area of single cell analysis. It is of interest to predict a cell decision as early as possible. Marker genes, for example, give good evidence. These genes show a very high or very low expression when pointing to a (later) cell decision.

In his project Dr. Florian Büttner (Helmholtz Zentrum München) [1] investigates the first developmental stages of mouse embryos. After three cell divisions, the embryo consists of eight cells and the cells begin to differentiate into specific cell types (Fig. 1A). In the experiments qPCR analysis of 48 genes was performed after each cell division. The data consists of the expression profiles of different cell types,
namely the different cell stages. A common approach to identify differences between groups is to embed the 48-dimensional space of genes into a smaller 2-dimensional subspace. Each cell is then represented by a point in the plane and one is interested in cells that are relatively close to each other. Using standard projections, only cells at the 64-cell-stage have been distinguishable so far. A non-linear extension that also incorporates the grouping of the cells shows that already cells at the 16-cell-stage split into two subgroups (Fig. 1B-C). Marker genes for these new subgroups can be determined by comparison of the expression measurements.

**Stochastic Profiling**

Single-cell sequencing becomes more and more applicable. However, the techniques remain expensive and technical noise is a problem when the amount of genetic material is small. Dr. Christiane Fuchs (co-author) [2] and her collaboration partners circumvent this problem using advanced statistics.

The aim of their project is to identify different cell types (type I and type II) in a heterogeneous mixture of cells. The cell types can for example represent healthy cells and cancer cells. Instead of single-cell measurements, they perform stochastic profiling: From several tissues one selects a random sample of a few cells. For these samples one measures the collective gene expression (Fig. 3). Thus, one has more genetic material and more reliable measurements. To model the collective gene expression a stochastic mixing model is introduced. One of the parameters provides the probability for a randomly chosen cell to be of type I or type II, respectively. Three groups of genes were analyzed. The model predicts that in these groups 25%, 10% and 2.3% of the cells show a distinctive expression profile. The last group is of special interest. Since only 2.3% of the cells are distinctive, expression measurements at a population level would have failed. Nevertheless, this group contains genes that are indeed related to breast cancer. Since the number of tissues was limited, stochastic profiling here provides even better results compared to single-cell analysis.

**References**

Authors

Katrin Illner, Dr. Christiane Fuchs, Prof. Fabian J. Theis, Institute of Computational Biology, Helmholtz Zentrum München

Contact

Helmholtz Zentrum München
Ingolstädter Landstr. 1
85764 Oberschleißheim
Germany
Phone: +49 89 3187 1587
Telefax: +49 89 31873324