Analysis of multi-modal cell image time series for drug effect quantification

M2 internship 6 months

Keywords: machine learning, cell biology,
Institution(s): Cairn Biosciences and CBIO MINES ParisTech
Location: Station F, 5 Parvis Alan Turing, 75013 Paris / 60 boulevard Saint-Michel. 75006 Paris
Name and email of the internship mentors
Thomas Walter: thomas.walter@mines-paristech.fr
Amin Fehri: afehri@cairnbio.com
Susana Ribeiro: sribeiro@cairnbio.com

Summary

Cairn Biosciences is developing next-generation therapeutics that address significant challenges in the treatment of serious diseases. We combine cutting-edge advances in microfluidics, cell engineering and cell biology in order to study in detail the effect of anti-cancer drugs on cellular populations. Such studies allow us to better understand the mechanism of action of these drugs and identify those that are most promising to develop further.

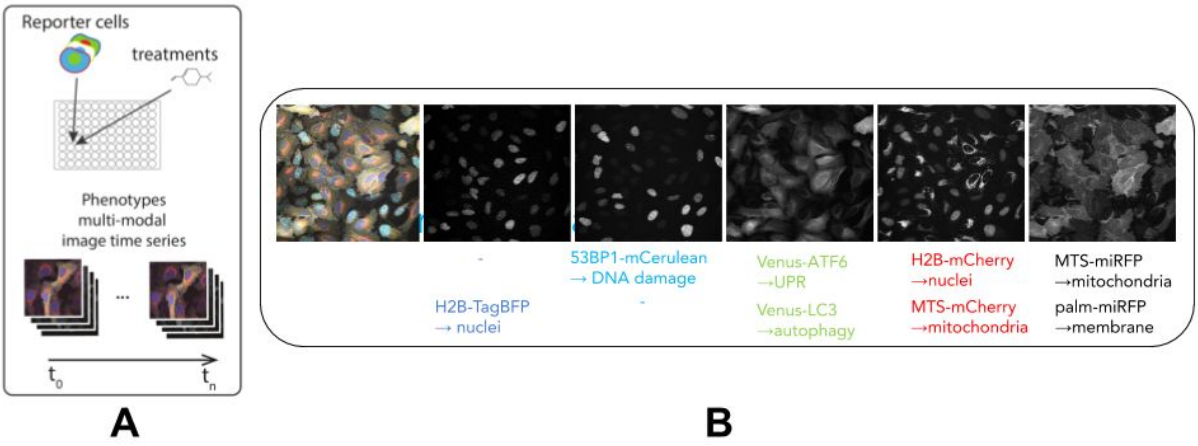

Figure 1 A: The experiments consist in studying *in vitro* the effects of potential treatments on reporter cells. Recent advances in microscopy and genetics enable to obtain rich data, in the form of multi-modal image time series. **B:** The multi-modality lies in the biological domain, as each channel highlights different parts of the cell. This unlocks new ways of understanding the effects of a drug, hence the need for suited algorithms and methods.

Our unique technology allows us to perform drug screens of unprecedented level of detail, where we test thousands of drugs at different concentrations with respect to their effect on cellular populations. For each of the cells, we use a set of markers that allows us to track different biological processes inside living cells over time. In particular, we are interested in studying cell morphology, protein kinetics and spatial distribution of proteins inside cells in response to treatment. In our case, the assays we generate are unique for three reasons: (i) they are live, meaning that we can monitor the evolution of cells over time; (ii) they contain pooled cells that have been engineered to highlight different parts of the cells or relevant biological pathways (6 different markers from 2 different cell lines in the current case); (iii) they contain individual cell data as opposed to whole population assays where data at single cell level cannot be extracted (see Figure 1).

These data also pose important challenges in terms of automatic analysis at both the cellular and the population level. At the cellular level, the main task is to segment, represent, classify and track individual cells, typically by using Deep Neural Networks. Each cell is then represented by a time series of either classification results or multidimensional feature vectors. The entire set of time-series corresponds to the description of the population phenotype. These primary analysis results can be either used in order to monitor changes in some biological property upon drug treatment, to predict the mechanism of action of a drug, to cluster drug effects or to develop concentration dependent profiles of the drug response. All of these questions involve development of novel methods, mainly in Image Analysis, Machine Learning and time series analysis.

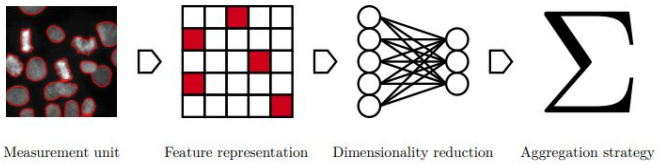

Figure 2: The development of a phenotypic profile spans four ordered stages. Each stage may be accomplished by a variety of algorithms, the combination of which define a unique pipeline [2].

Specific goals of the internship:

(1) Representation learning at single cell and population level
(2) Phenotype clustering at cell and population level
(3) Domain adaptation in order to ensure the transferability
(4) Compare developed method to other methods using public datasets
(5) Propose visualization tools for time-resolved data to present results to biology experts

Profile

M2 student, machine learning, computer vision, math/info

Required skills

Proficiency in at least one relevant programming language (e.g. Python)
Machine Learning, Statistical analyses, Image processing, Computer vision
Fluent in English

Desired skills

Experience with Deep Neural Networks
Knowledge or interest in learning about cell biology, microscopy and drug discovery

Bibliography

[1] Scheeder, C., Heigwer, F., & Boutros, M. (2018). Machine learning and image-based profiling in drug discovery. *Current opinion in systems biology*, *10*, 43-52.

[2] Boyd, J. C., Pinheiro, A., Del Nery, E., Reyal, F., & Walter, T. (2019). Domain-invariant features for mechanism of action prediction in a multi-cell-line drug screen. BioRxiv, 656025.

[3] Singh, S., BRAY, M. A., Jones, T. R., & Carpenter, A. E. (2014). Pipeline for illumination correction of images for high-throughput microscopy. Journal of microscopy, 256(3), 231-236.

[4] Lafarge, M. W., Caicedo, J. C., Carpenter, A. E., Pluim, J. P., Singh, S., & Veta, M. (2019, May). Capturing Single-Cell Phenotypic Variation via Unsupervised Representation Learning. In International Conference on Medical Imaging with Deep Learning (pp. 315-325).

To apply please send resume and motivation letter detailing your interest and experience to the emails shown above.