

# Semi-supervised learning for phenotypic profiling of high-content screens\*

Róger Bermúdez-Chacón<sup>†</sup>  
Supervisor: Peter Horvath<sup>‡</sup>

July 17, 2013

## Introduction

Recent advancements in high-throughput microscopy and data analysis made possible to perform large scale biological experiments and automatically evaluate them. For the detection of sub-cellular changes caused by different perturbations in the cell (RNAi or drugs) often supervised machine learning (SML) is used. Reliable training of an SML method requires significant effort from a field expert. The aim of this project is to investigate whether and to which extent intelligent control experiment design combined with semi-supervised learning can reach the accuracy of a human annotator and/or in certain cases substitute it.

## Materials, Tools and Methods

[TODO]

## Results

[TODO]

## Discussion

[TODO]

## Conclusions

[TODO]

---

\*This project was held as a Lab Rotation in Computer Science, as required by the Master program in Computational Biology and Bioinformatics - ETH Zürich

<sup>†</sup>Computational Biology and Bioinformatics Master program, ETH Zürich

<sup>‡</sup>Light Microscopy and Screening Centre, ETH Zürich

# Bibliography

- [1] D. Charif and J. Lobry, “SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis.,” in *Structural approaches to sequence evolution: Molecules, networks, populations* (U. Bastolla, M. Porto, H. Roman, and M. Vendruscolo, eds.), Biological and Medical Physics, Biomedical Engineering, pp. 207--232, New York: Springer Verlag, 2007. ISBN : 978-3-540-35305-8.