

HCS-Analyzer: Open source software for High-Content Screening data correction and analysis



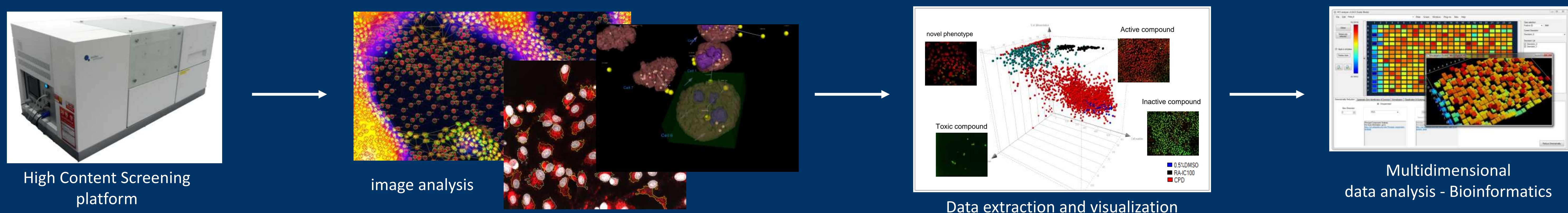
T. Dorval & A. Ogier
Cellular Differentiation - IPK

High-Throughput Screening is a powerful technology principally used by pharmaceutical industries allowing the identification of molecules of interest within large libraries. Originally target based, cellular assays provide a way to test compounds (or other biological material such as small interfering RNA) in a more physiologically realistic *in vitro* environment.

High Content Screening (HCS) platforms are giving the opportunity for universities or research institutes to access those technologies for research purposes. However, the amount of information extracted from each experiment is highly multiplexed and then difficult to handle. In such context, there is an important need for an easy to use, but still powerful software able to manage multidimensional screening data by performing adapted quality controls and classification.

HCS analyzer [1] includes: a user-friendly interface specifically dedicated to HCS readouts, an automated approach to identify systematic errors potentially occurring in a screening and a set of tools to classify, cluster and identify phenotypes of interest among large and multivariate data.

<http://hcs-analyzer.ip-korea.org>



Data import & plate design

SQL database
CSV files
Generate artificial screening data (cell-by-cell, biased, multivariate)
User friendly plate design (controls, classes, descriptor management)
Screening plates management and visualization

Quality control and correction

Data normalization
Screening and plate quality control
Systematic error identification and correction [2]

Descriptor relationships analysis

High level descriptors correlation analysis (Pearson, Spearman, MINE [3], ranking, etc.)
Statistical analysis (PCA, LDA, etc.) and visualization
Dimensionality reduction (supervised, unsupervised)

Clustering and classification

Automated multivariate phenotypic clustering (EM, K-Means, hierarchical)
Well based classification (J48, Neural Networks, KNN, Support Vector Machines)
Cell-by-cell classification

Genomic screening

Gene and pathway information
Pathway redundancy over a class

DRC and drug-to-drug interaction

Dose response designer, analysis and visualization
Drug-to-drug surface interaction visualization and quantification (isoboles)

Multi-format export functionalities

Visual report
Weka format

Open source and plug-in management

.NET C# source code
GitHub repository

References & Acknowledgement

- [1] A. Ogier and T. Dorval, HCS-Analyzer: Open source software for High-Content Screening data correction and analysis, *Bioinformatics*, May 2012.
[2] Carralot et al., A Novel Specific Edge Effect Correction Method for RNA Interference Screenings, *Bioinformatics*, Nov. 2011.
[3] Reshef et al. *Detecting Novel Associations in Large Data Sets*, Science, Dec. 2011.
- This work was supported by the National Research foundation of Korea (NRF) grant funded by the Korea government(MEST)(No.2011-00244), Gyeonggi-do and KISTI.