

# Merging high-content screening and in silico approaches for compound profiling and mode-of-action analysis

A. Bender<sup>1,2</sup>

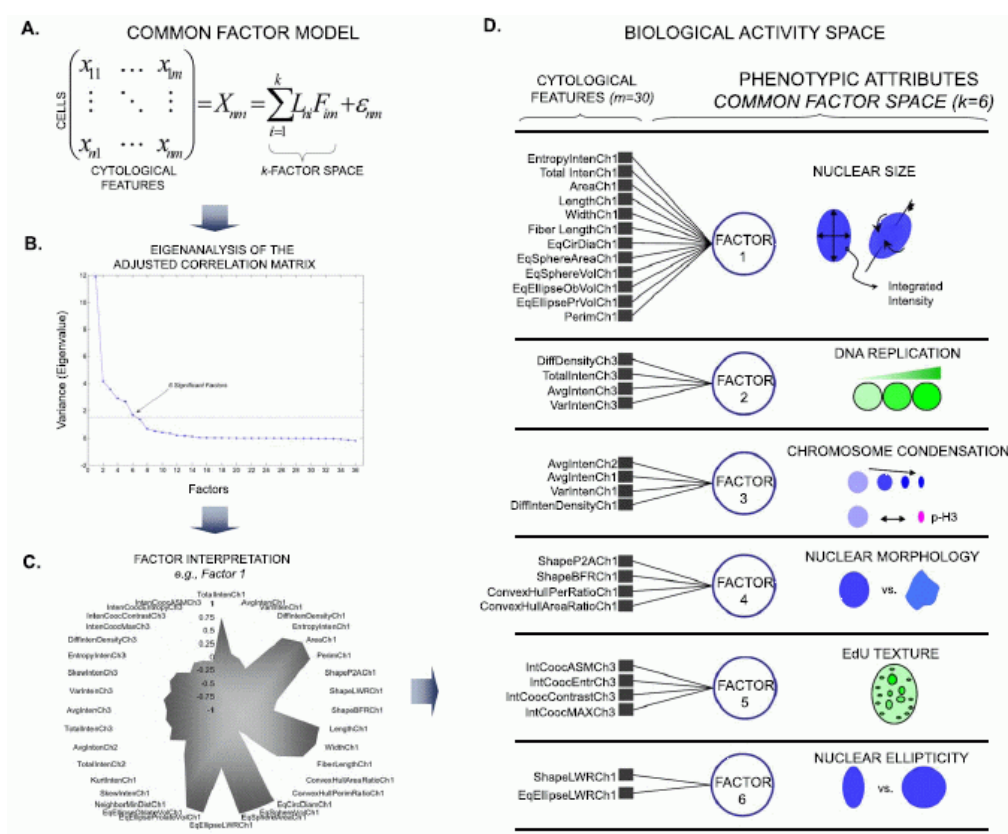
D. W. Young<sup>2</sup>, J. L. Jenkins<sup>2</sup>, Y. Feng<sup>2</sup>

<sup>1</sup>Leiden / Amsterdam Center for Drug Research, Division for Medicinal Chemistry, Leiden University, Einsteinweg 55, 2333 CC Leiden, The Netherlands

<sup>2</sup>Novartis Institutes for Biomedical Research, Inc., 250 Massachusetts Avenue, Cambridge, MA 02139, United States of America

High-content screening observes the reaction of a cell to an administered compound by multidimensional microscopy and it provides a potentially more information-rich complement to single-readout conventional assays. On the other hand, microscopy-based screening can also be more 'opaque' in the way that no mechanistic explanation for the observed effect is provided per se.

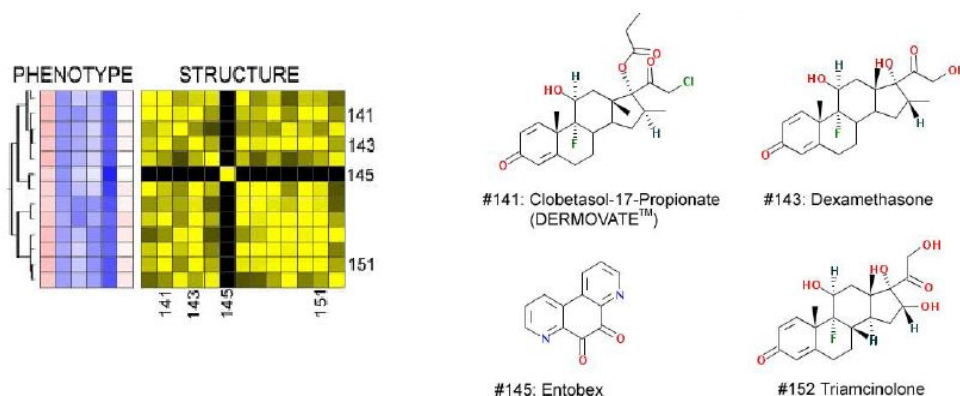
The work we performed to improve our ability to handle and understand high-content screening data consists of two parts<sup>1</sup>. Firstly, in order to reduce the vast amount of information obtained from microscopy based data, we performed factor analysis to reduce the amount of data to analyze, while at the same time retaining most of the information. We were able to define a six-dimensional factor space that defines cell state variables such as nuclear size and DNA replication, as depicted in Figure 1.



**Figure 1.** Factor analysis employed to project high-dimensional HCS readout space into low-dimensional space, using 6 variables to describe a cell state.

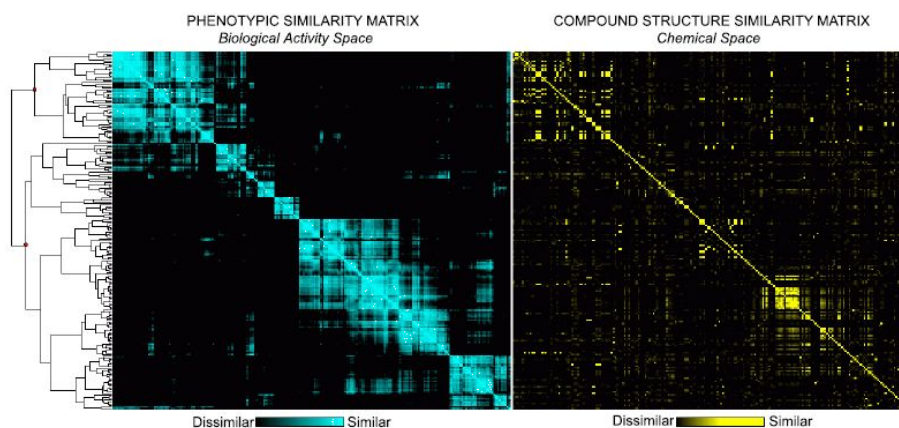
Next, by merging high-content screening with in silico target prediction, we merge both phenotypic and mechanistic approaches: by high-content screening we are able to observe the systems response, while at the same time providing hypotheses for the observed effects via the predicted targets of compounds.

We screened more than 6k compounds in high-content screenings and discuss cases where the phenotypic response and the predicted targets agree with each other, but also the even more interesting 'atypical' cases where similar phenotypes are observed by very different predicted targets (which might for example be located in the same pathway; Figure 2).



**Figure 2.** Four compounds, three of which steroids, which generate very similar phenotypes despite two distinct scaffolds present. Shown to the left are phenotypic readouts in six dimensions as well as a ligand structure similarity matrix.

Looking at the full compound set screened, one now has the opportunity to compare phenotypic similarity ('systems response') to structural similarity on a larger scale. This is shown in the similarity matrices in Figure 3 – while overall similar structures give similar readouts, also clear deviations from the rule are present. This analysis gives us the opportunity to compare compounds by not only using single or a defined set of targets, but the complete systems response instead.



**Figure 3.** Comparison of phenotypic and ligand structure similarities of all ligand pairs. While both matrices show similarities, also clear differences are present, giving the possibility for example to define 'phenotypic' compound similarities.

1. Young, D.W.; Bender, A.; Hoyt, J. McWhinnie, E.; Chirn, G.W.; Tao, C.Y.; Tallarico, J.A.; Labow, M.; Jenkins, J.L.; Mitchison, T.J.; Feng, Y. Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nature Chem. Biol.* **2007**, 4, 59 – 68.