Electronic Supplementary Material (ESI) for Natural Product Reports.

Table S1. List of commercially available automated microscopes

Company	Device name	Confocal	Link
Acquifer AG (Karlsruhe)	IM03	No	http://www.acquifer.de/high_content_screening/
GE Healthcare Europe GmbH (Freiburg)	IN CELL 2200/6000	No/ Yes	http://www.gelifesciences.com/webapp/wcs/stores/servlet/catalog/en/GELifeSciences-de/products/AlternativeProductStructure_12997/29043323
Leica Microsystems GmbH (Wetzlar)	HCS-A module for Leica microscopes	Yes	http://www.leica-microsystems.com/de/produkte/lichtmikroskope/biowissenschaften/fluoreszenzmikroskope/details/product/leica-las-x-widefield-systems/
Molecular Devices, LLC (CA, USA)	Image Express Micro/Ultra	No/Yes	http://www.moleculardevices.com/systems/high-content-imaging/imagexpress-ultra-confocal-high-content-analysis-system
Nikon GmbH (Düsseldorf)	Ti HCS	Yes	http://nikon.com/products/instruments/lineup/bioscience/biological-microscopes/inverted/ti_eus/index.htm
Olympus Deutschland GmbH (Hamburg)	ScanR	No	http://www.olympus-europa.com/microscopy/en/microscopy/components/component_details/component_detail_21320.jsp
PerkinElmer Inc. (MA, USA)	Operetta/Opera	No/yes	http://www.perkinelmer.com/pages/020/cellularimaging/products/operetta.xhtml
Thermo Fisher Scientific Inc. (MA, USA)	CellInsight/ Arrayscan XTI Infinity	No/Yes	https://www.lifetechnologies.com/order/catalog/product/ASN00004F
Visitron Systems GmbH (Puchheim)	CellVoyager CV100 (preliminary)	Yes	http://www.visitron.de/Products/Confocal/Cell_Voyager_CV1000/cell_voyager_cv1000.html
Wako Automation (CA, USA)	Yokogawa CV7000s/CQ1	Yes	https://www.lifetechnologies.com/order/catalog/product/ASN00004F

Table S2. Bioinformatical approaches for compound profiling

Approach	**Description**	**References** See main text
Principal component analysis (PCA)	PCA aims at reducing data complexity (dimension reduction by linear combination of descriptors/features) of multivariate data sets and improving interpretation and visualization. For this only principal components are considered that cover most of the variability within the dataset.	41, 49, 51
Kohonen neural networking	A Kohonen network is a self-organizing map (SOM) that produces a low dimensional representation for comparison of multiparametric phenotypic responses. First, multivariate measurements of a training set (compounds with known MOA) are analyzed for constructing a map. In the second step, measurements from unknown compounds are mapped to the predefined classes.	49
Distance/ similarity of mean profiles	For each feature the average is calculated. Feature responses derived from untreated or with known compounds treated cells are scaled to cover a common range. Compound profiles are then generated consisting of all features represented by their mean values. Using a distance measure the similarity of the profiles can be calculated and further analyzed by e.g. rank based approaches or cluster analysis.	41, 47, 51
Kolmogorov-Smirnov (KS)-statistic	The KS-algorithm is a nonparametric hypothesis test for comparing population responses. It is based on the maximum difference that is calculated between two cumulative distribution functions (CDF) of each feature. Profiles of scored KS-statistics can be used for profiling mode of actions.	9, 47
Normal vector to support-vector machine (SVM) hyperplanes	The SVM algorithm is a supervised data mining method that can be used, for pattern recognition. For constructing a SVM, objects for training are needed. In case of image analysis the SVM has to be trained to discriminate compound-treated from control cells. The SVM is then able to generate a hyperplane that divides both conditions in the feature space. The normal vector of the hyperplane is included from every feature for generating a profile.	44, 47
Gaussian mixture modeling (GMM)	In GMM it is assumed that compound treatment induces a limited number of cellular phenotypes. Features are represented here as a mixture of Gaussian distributions. Subsamples of the data set are used to generate GMM profiles for untreated cells/ known compounds. Unknown samples can be assigned to a MOA class by the probability of their features to belong to one of the GMM profiles.	45, 47
Factor analysis	The application of factor analysis for phenotypic profiling is based on the assumption that a certain biological response (induced by compound treatment) is represented by several features of which subsets can exhibit high correlation Correlated features are grouped and represent a common underlying factor. The factors allow for the interpretation of the biological effect of a compound.	47, 48