COMAS
Compound Management and Screening Center

MAX PLANCK INSTITUTE
OF MOLECULAR PHYSIOLOGY

Cell Painting

an unbiased multiparameter phenotypic screening assay at COMAS

Axel Pahl

AIDD Summer School, May 2022

What is „unbiased“
and
Why is it important?

How Is Activity Identified in New Chemical Matter?

- By dedicated, target- or pathway-based assays
  - mostly in-vitro or in cells
  - often reasonably easy to perform
  - BUT: requires a target hypothesis for the tested compound
  - most appropriate when working in compound series

But What To Do When There Is No Prior Information About The Activity?

- E.g. when generating new compound series / classes
- The target- / pathway-based approach would not work
  - most of the activity would be missed

- Target prediction would be an option
  - but needs experimental confirmation

An Assay Is Needed That Identifies Biological Activity Without Requiring a Prior Target Hypothesis

- This assay should not only find bioactivity in cpds. but also be able to…
- Identify a broad range of activities
  - without requiring a hypothesis (→ “Unbiased”)
- Be still reasonably easy to perform
  - medium to high throughput

- Does the Cell Painting fullfill these criteria?

Introduction to the Cell Painting Assay

Cell Painting Assay - Principle

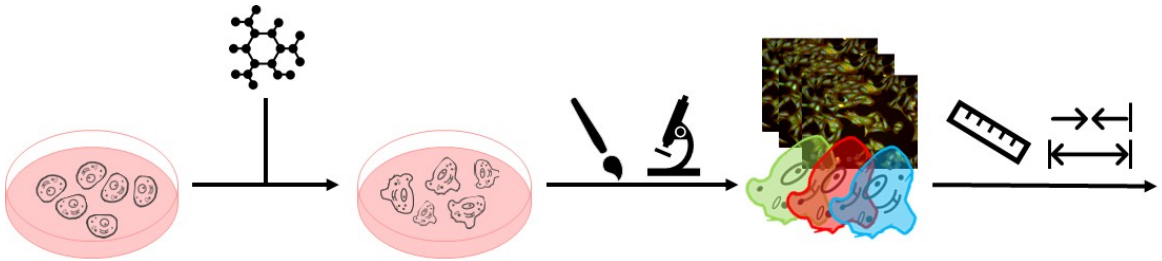
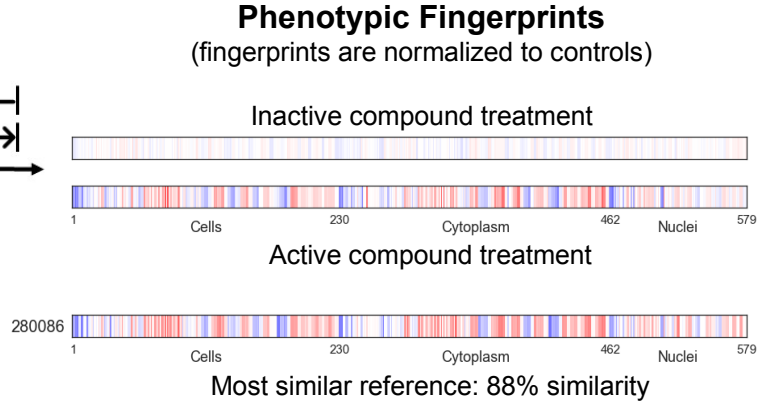


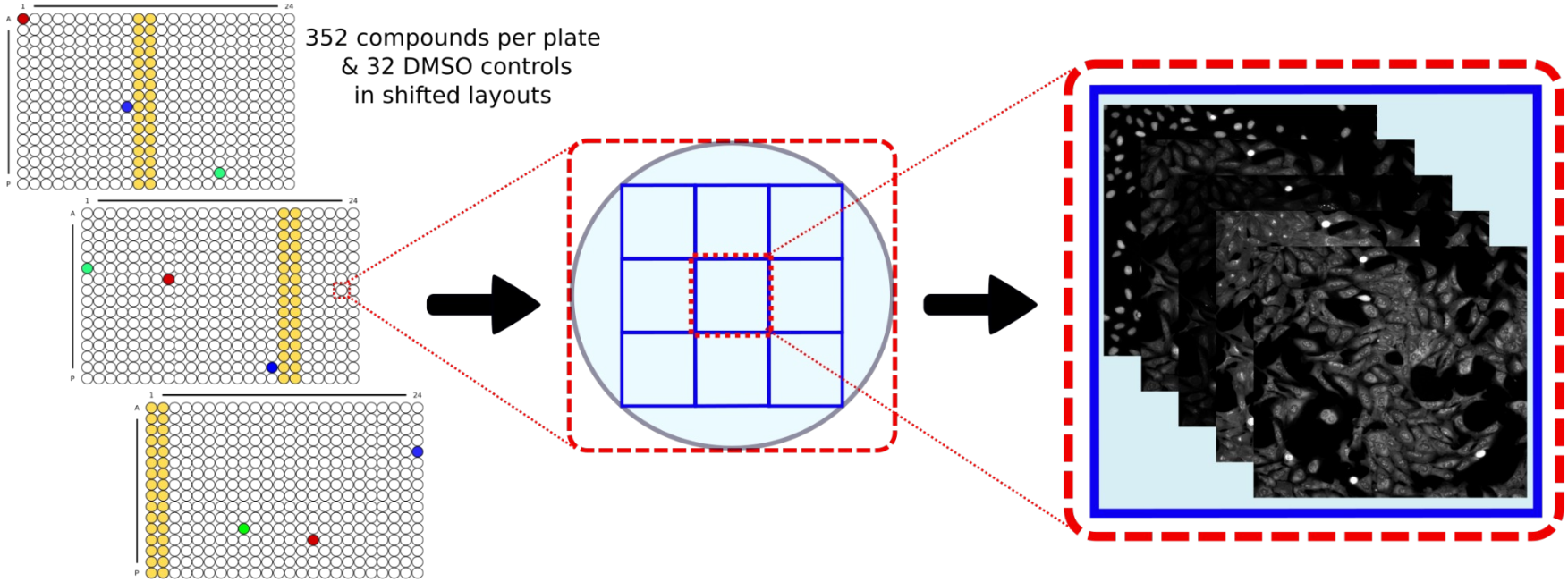
Figure by Sarah Zinken



- Developed by the Carpenter group from Broad institute
  - Bray, M.-A, Carpenter, A. et al. Cell Painting, a High-Content Image-Based Assay for Morphological Profiling Using Multiplexed Fluorescent Dyes. Nature Protocols 2016, 11 (9), 1757–1774. https://doi.org/10.1038/nprot.2016.105.
- Unbiased monitoring of changes in numerous cellular features and biological processes
  - staining of cellular compartments in five different fluorescent channels
  - image analysis
  - hundreds of parameters
  - numeric fingerprint of the cellular phenotype
- Deviations from control fingerprints define activity
- Comparison of fingerprints to reference compounds may reveal possible mode of action
- Service for scientists at the institute to identify activity in new chemical matter

Data Acquisition – Assay & Imaging

Default: U2OS; 20h compound incubation



3
replicates

x 384
wells

x 9
sites
per well
(binning: 2)

x 5
channels
per site

= 51480 images
= 12 h imaging time
= 120 GByte data

Data Processing

- Images are processed on a SLURM-managed in-house computing cluster (96 cores)
- Processing with CellProfiler 3.0.0
  → **>1700** features for ***each cell*** / ~1800 cells per microtiter well
- Processing of each replicate plate is distributed over 96 parallel jobs
  → each job takes a slice of 36 sites (180 images (36 * 5 channels))
  → ~5 h per 384 well replicate plate
    → **15 h compute time for 3 replicates**
- Results from the individual jobs are concatenated into one result file and aggregated per microscope site as Medians
  → spreadsheet of 3456 rows x 1700 columns per replicate plate

Data Analysis and Reporting

- All further downstream processing also on the computing cluster

  → Data aggregation per well (Median) over all replicates
  → Determination of phenotypic profiles (fingerprints)

  → Calculation of phenotypic fingerprint similarity to references
  → Generation of static HTML-based reports

- In addition: interactive web tools for flexible querying and visualization of the data by the users
- In total: ~5000 lines of in-house written Python code
  - with some performance-critical code written in Rust (e.g. calculation of profile-similarity)

Determination of Relevant Features

- Not all collected features are suitable for analysis
- The set of relevant features has to be determined
- This is done only once
  - one plate of reference compounds and one repeat*

Features remaining:	
>1700	Determined by CellProfiler
⇩	Remove features that have a high or a very low variability among the controls
~1300	
⇩	Keep features that have a minimum correlation of 0.8 between repeats for all cpds..
579	Final set of relevant features. Used for all further analyses

*) Selecting CP features based on biological reproducibility adapted from *Woehrmann et al, Mol. BioSyst.*, 2013, **9**, 2604

Activity Profile – Z-Scores

- For each feature the Median and Median Absolute Deviation (MAD) of the controls are calculated
- Z-score of measured feature value from test compound =
  how many times the MAD of the controls the measured value deviates from the Median of the controls:

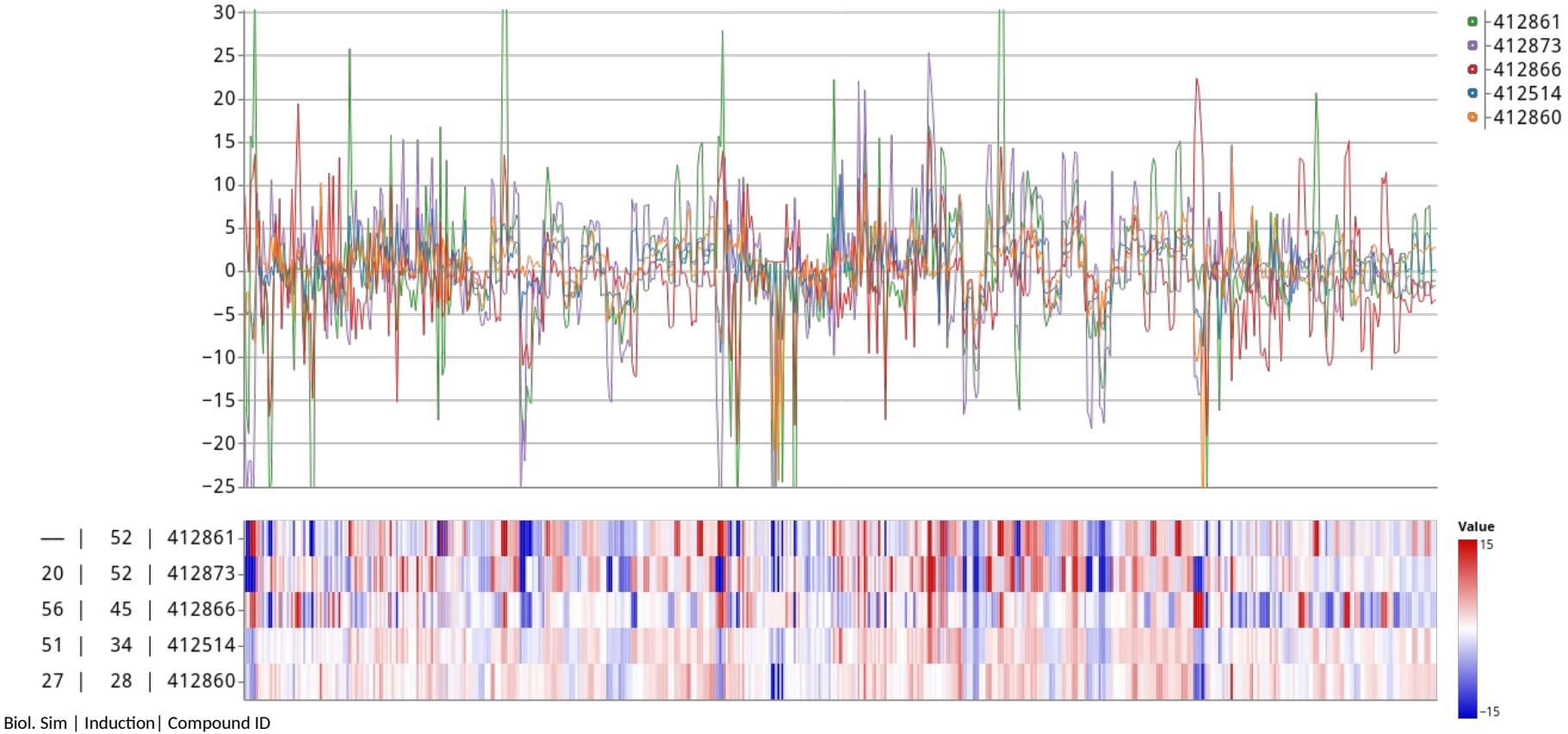
$$z\text{-}score = \frac{value_{meas.} - Median_{Controls}}{MAD_{Controls}}$$

The phenotypic profile is then the list of z-scores for a given test compound

- Induction: number of features with an abs(z-score) > 3 divided by the total number of features
  → expressed in %

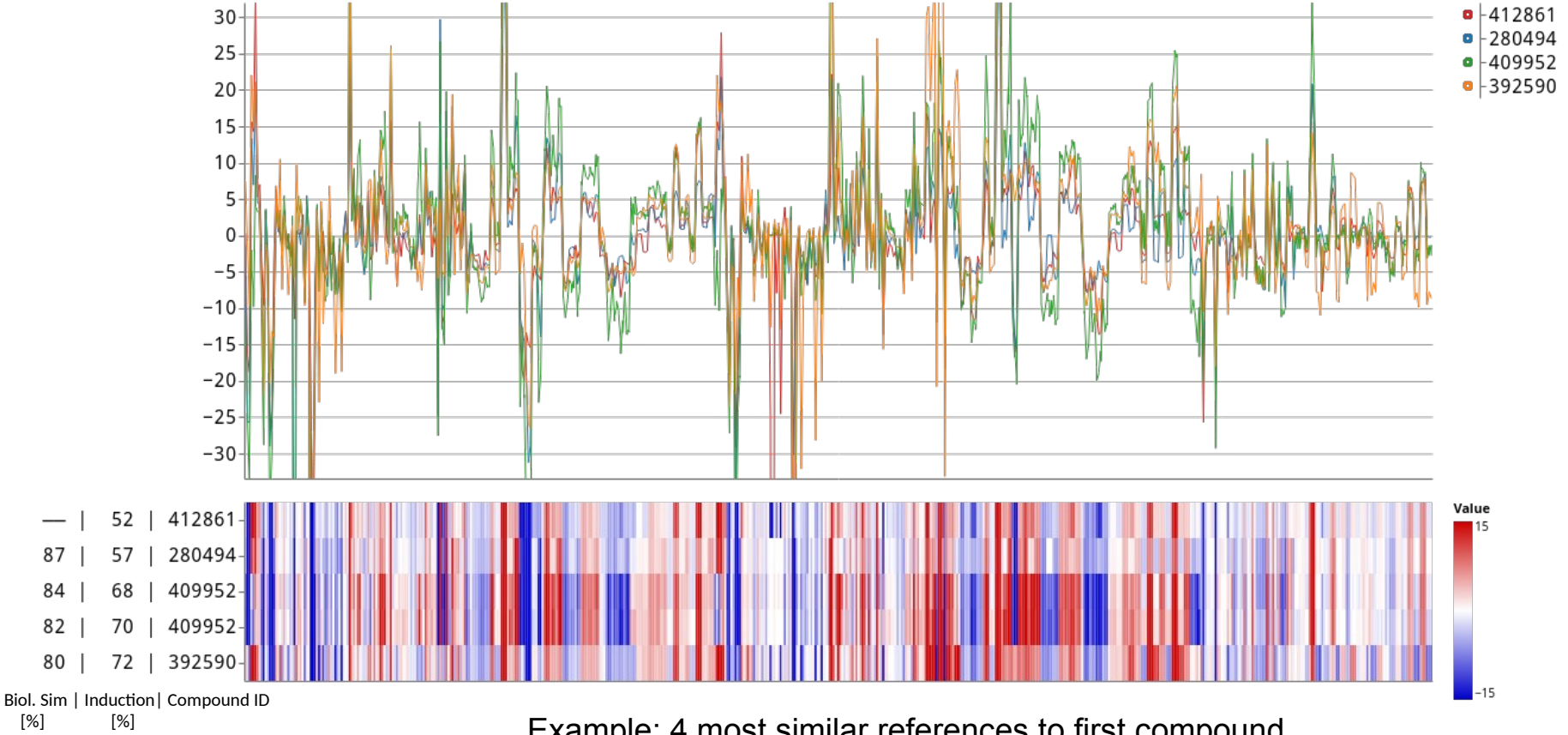
$$Induction\ [\%] = \frac{number\ of\ features\ with\ abs(z\text{-}score) > 3}{total\ number\ of\ features}$$

Z-score Profiles are Represented as Heat Maps or Line Plots



Example: 5 different test compounds

Z-score Profiles are Represented as Heat Maps or Line Plots



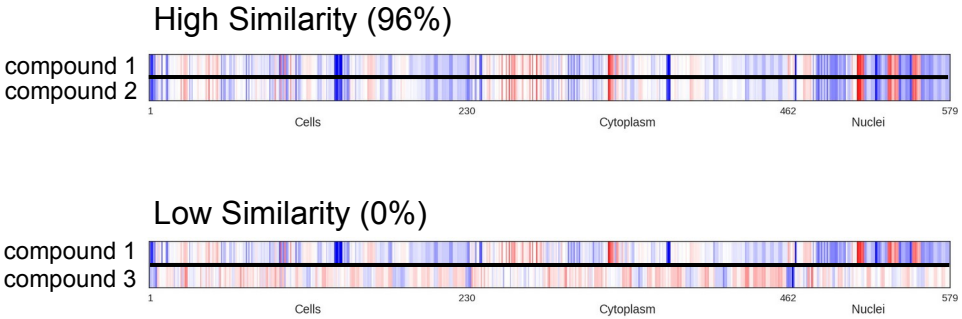
Example: 4 most similar references to first compound

Profile Similarity

- Similarity by Correlation Distance*

$$CorrDist = 1 - \frac{(u - \bar{u}) \cdot (v - \bar{v})}{Norm(u - \bar{u}) Norm(v - \bar{v})}$$

  - 0: low distance, 1: large distance
  - Similarity = 1 - CorrDist



Profiles can have **similar shapes**

→ High profile similarity, even when Z-scores differ in their absolute values.

→ BUT: better not compare compounds with very different inductions.

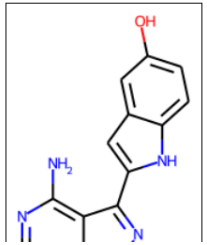
→ Robust against dose-dependent effects

*) implementation in scipy: https://docs.scipy.org/doc/scipy/reference/generated/scipy.spatial.distance.correlation.html

Data Reporting

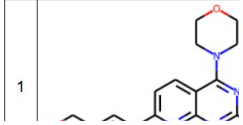
Detailed Report

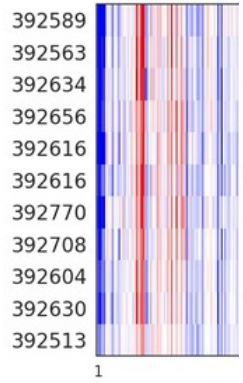
Compound Id 392589



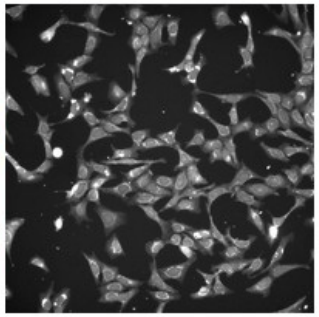
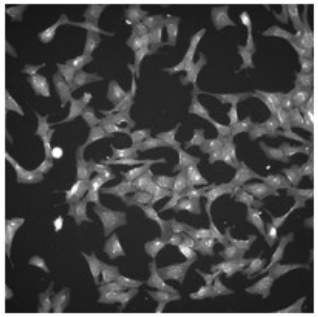
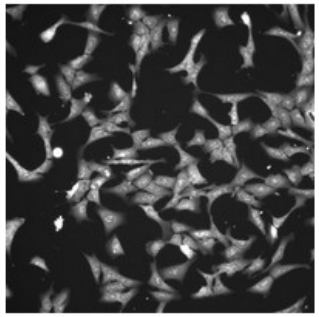
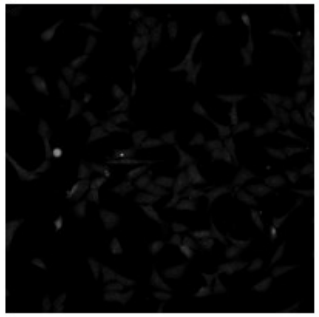
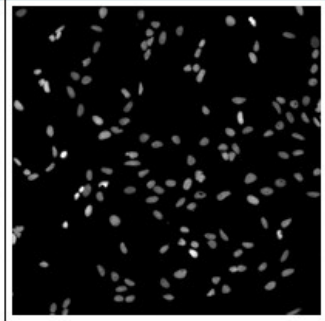
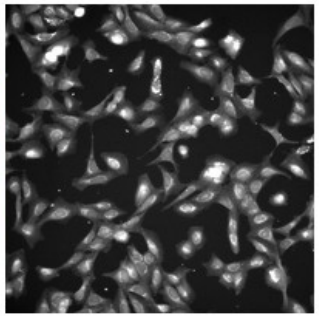
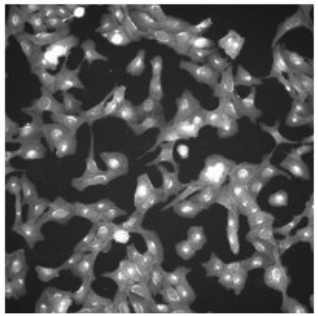
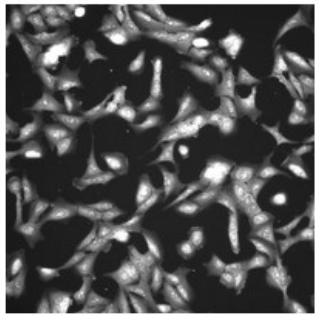
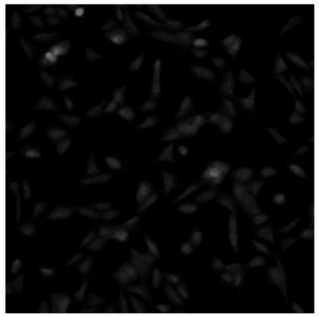
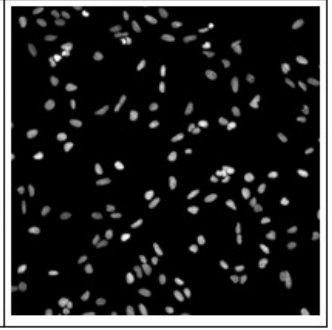
Well Id:	392589:01:10_02.00
Producer:	SIEVERS
Chiral:	False
Purity Flag:	Ok
Conc:	2.0 µM
Induction:	45.9 %
Similarity to Lysosomotropic Profile:	0.0
Cell Count:	100 %Ctrl

Similar References

Idx	Mol
1	



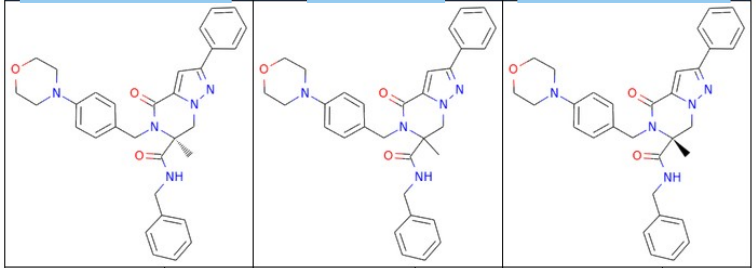
Sample images from site 5.

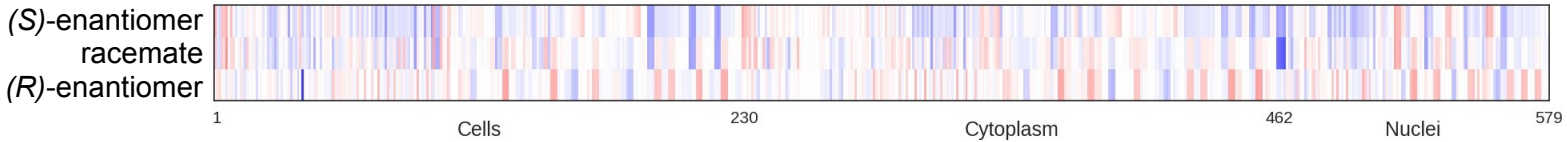
	Mitochondria	Golgi / Cell Membrane / Cytoskeleton	Cytoplasmic RNA / Nucleoli	Endoplasmatic Reticulum	Nuclei
Compound					
Control					

Selected Results

Cell Painting Profiles of Mixtures (Racemate):
can be (Partially) Additive

CP Profiles Can Be (Partially) Additive

	(S)-enantiomer	racemate	*(R)*-enantiomer
			
	Induction 14.5	Induction 11.4	Induction 16.4
Glucose uptake inhibition (pIC50 / IC50 [µM])	6.4 / 0.4	7.1 / 0.08	< 3 / > 1000



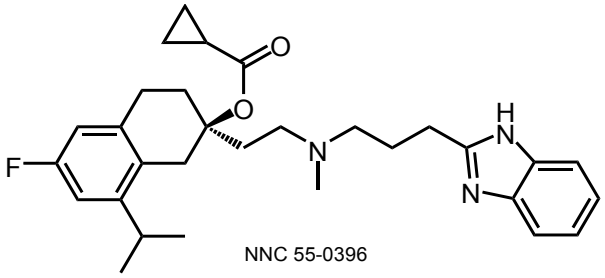
- The pure enantiomers and the racemate show comparable CP inductions
- Only the *(S)*-enantiomer and the racemate are active GLUT inhibitors
- The CP profiles of the two enantiomers are very different (0% similarity)
- The racemate shows (in part) additive features from both enantiomers

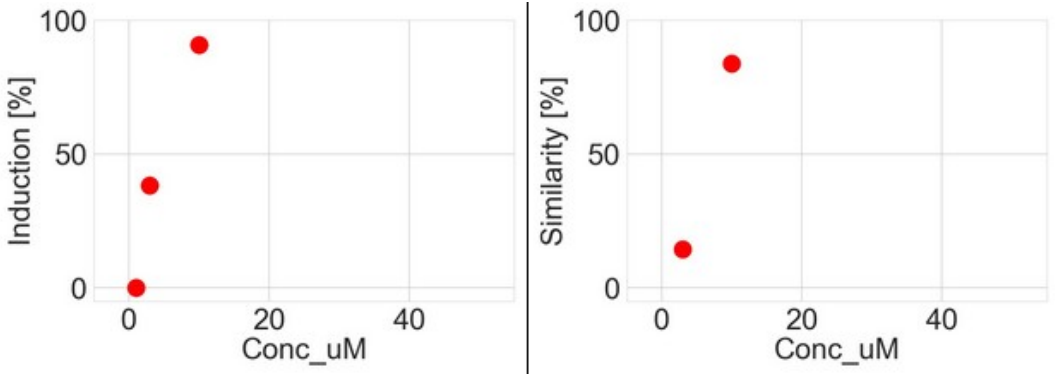
Selected Results

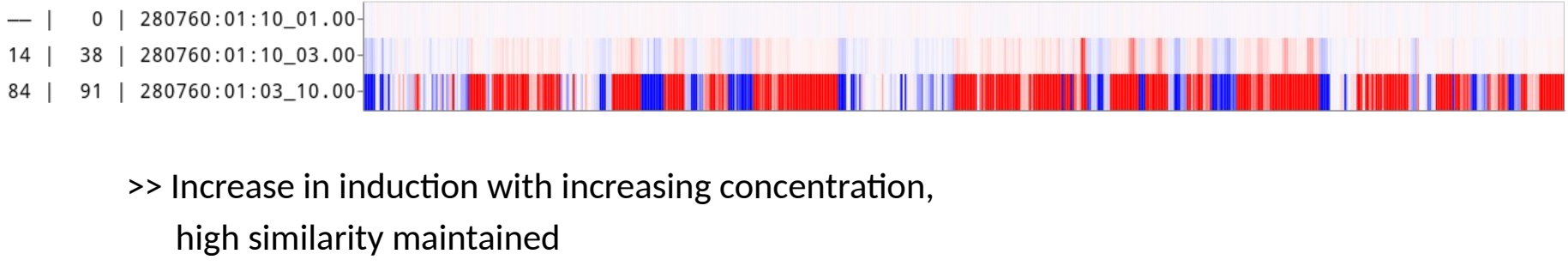
Profiles are Dose-Dependent

Concentration dependant phenotypes

LOPAC: selective T-type calcium channel inhibitor.



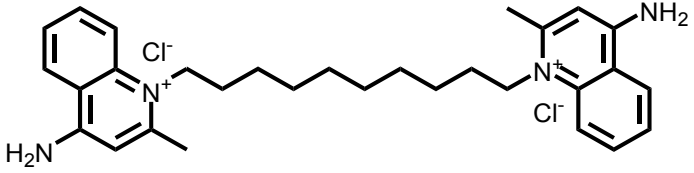


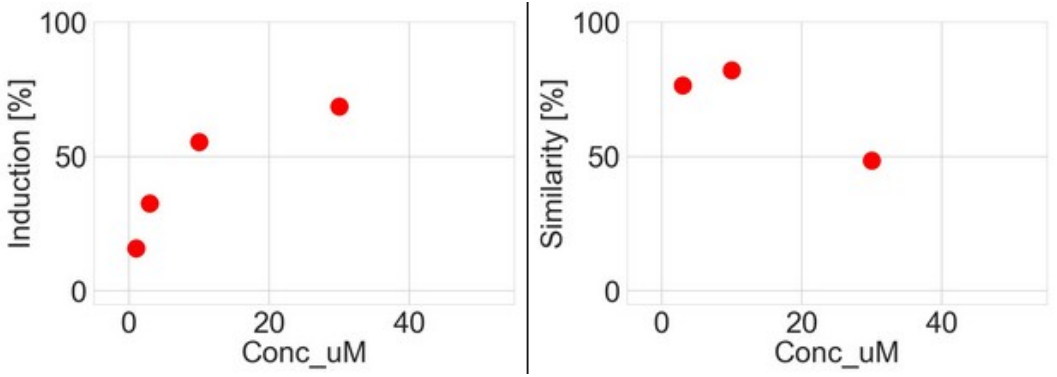


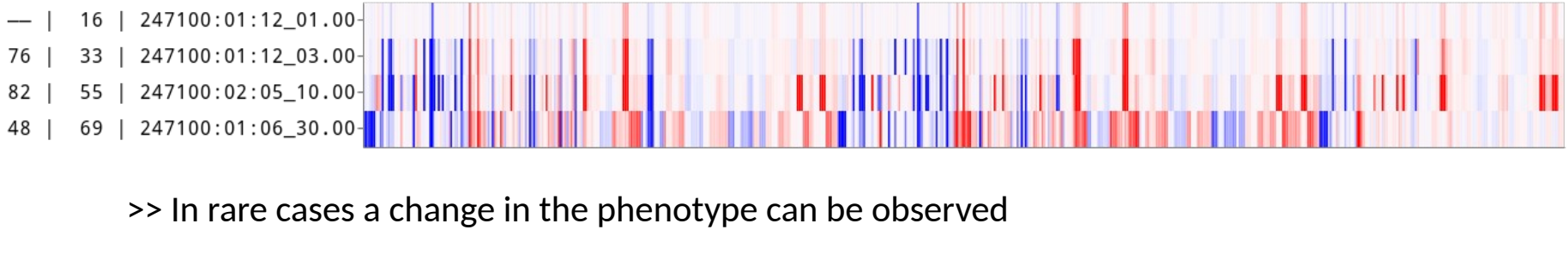
>> Increase in induction with increasing concentration,
high similarity maintained

Concentration dependant phenotypes

LOPAC: Dequalinium dichloride; K^+-channel blocker





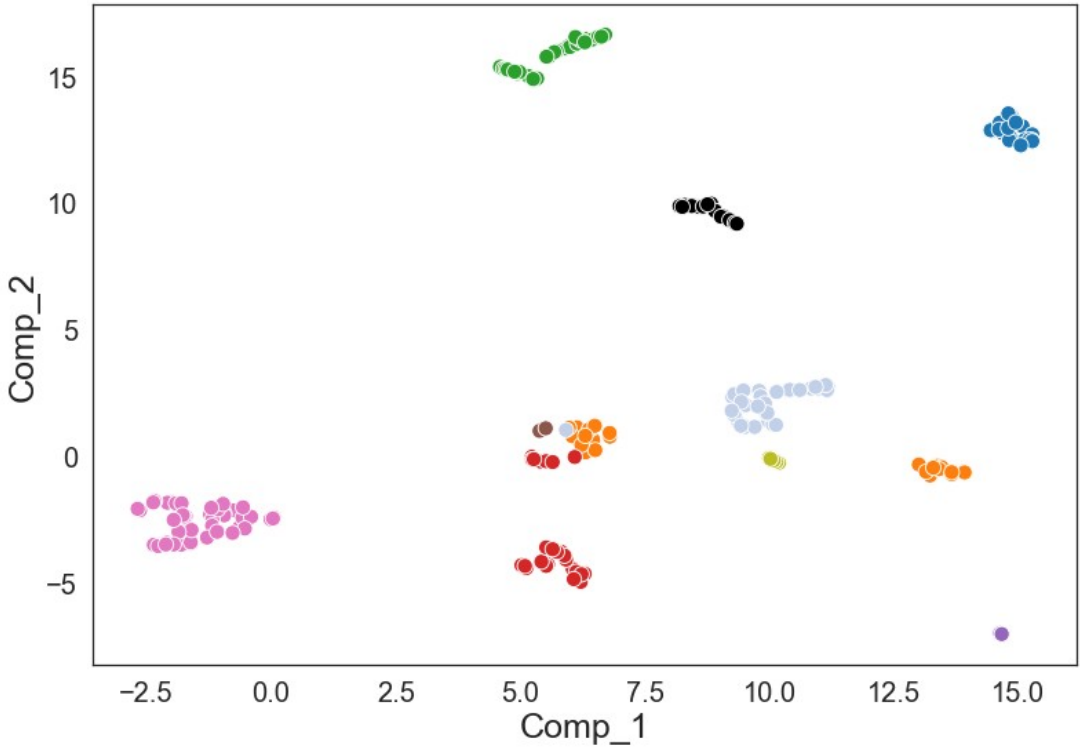


>> In rare cases a change in the phenotype can be observed

Selected Results

Clustering by Dimension Reduction
(UMAP)

Uniform Manifold Approximation and Projection (UMAP)



1. LCH [1]
2. AKT_PI3K_MTOR
3. Aurora
4. BET [2]
5. DNA_Synth
6. HDAC
7. HSP90
8. Protein_Synth
9. Tubulin
10. Uncoupler

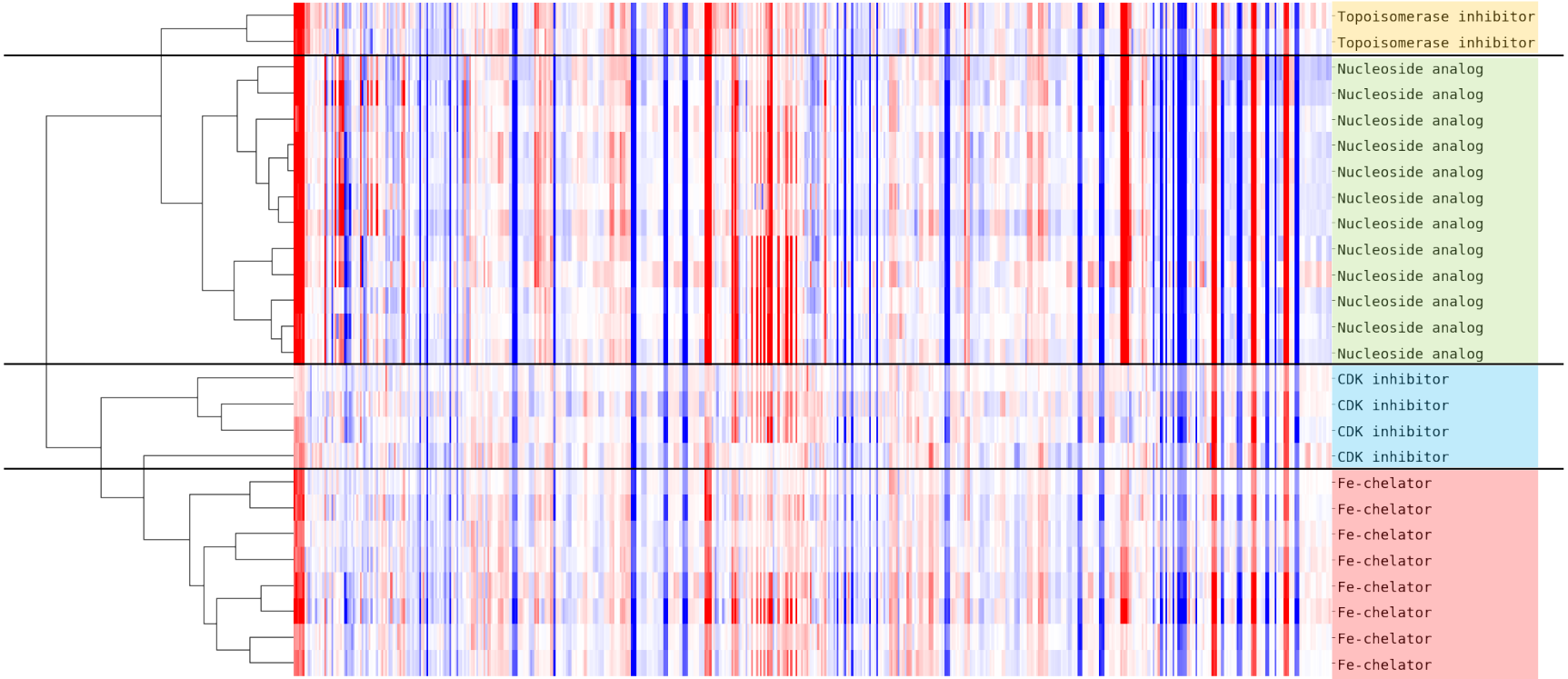
1) Lysosomotropic Cholesterol homeostasis
2) Bromodomain and Extra-Terminal motif

- Reduction of the 579 features to 2-3 dimensions using UMAP (or PCA, t-SNE) allows distinction of biological clusters

Selected Results

Hierarchical Clustering

Hierarchical Clustering



- Hierarchical clustering of the feature profiles allows distinguishing between different mechanism of actions within the same biological pathway (here: DNA synthesis cluster)

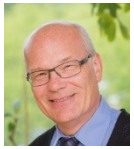
Schneidewind, T.; Brause, A.; Pahl, A.; Burhop, A.; Mejuch, T.; Sievers, S.; Waldmann, H.; Ziegler, S. DOI: 10.1002/cbic.202000381.

Limitations

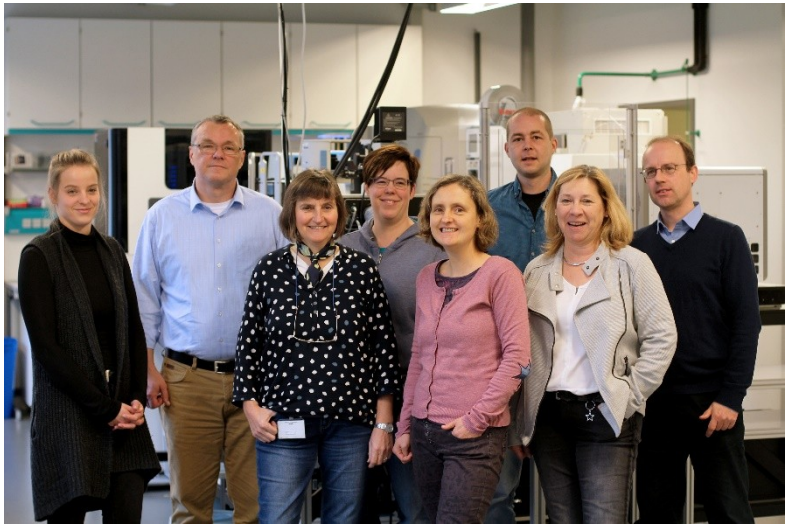
- Works only for compounds that induce a phenotypic change in the cells
  - Only ~1/3 of reference cpds. and of internal research cpds. show significant effect (>=5% Induction)
- Higher induction values do not identify higher compound activity
  - just means the phenotype was changed in more features
  - but: experience shows induction often is concentration dependent
- Relies on known / published annotations of the references
  - limited annotation of polypharmacology!
- Target / MoA-identification rely on representation in a reference
  - target bias of reference libraries!
- Changes in the fingerprints cannot easily be translated back into changes of cell morphology

Acknowledgements

Prof. Dr. Dr. Herbert Waldmann



Dr. Sonja Sievers (Head of COMAS)



COMAS Team (left to right):
Carina Birke, Axel Pahl, Heike Rimpel, Carla Brinkmann, Sonja Sievers (Head), Matthias Bischoff, Christiane Pfaff, Claude Ostermann; Philipp Lampe (not on picture)

Selected Publications

- Akbarzadeh, M.; Deipenwisch, I.; Schoelermann, B.; Pahl, A.; Sievers, S.; Ziegler, S.; Waldmann, H.
  Morphological Profiling by Means of the Cell Painting Assay Enables Identification of Tubulin-Targeting Compounds.
  Cell Chemical Biology 2021 (accepted, in press).
  https://doi.org/10.1016/j.chembiol.2021.12.009.

- Schneidewind, T.; Brause, A.; Schölermann, B.; Sievers, S.; Pahl, A.; Sankar, M. G.; Winzker, M.; Janning, P.; Kumar, K.; Ziegler, S.; Waldmann, H.
  Combined Morphological and Proteome Profiling Reveals Target-Independent Impairment of Cholesterol Homeostasis.
  Cell Chemical Biology 2021. https://doi.org/10.1016/j.chembiol.2021.06.003.

- Grigalunas, M.; Burhop, A.; Zinken, S.; Pahl, A.; Gally, J.-M.; Wild, N.; Mantel, Y.; Sievers, S.; Foley, D. J.; Scheel, R.; Strohmann, C.; Antonchick, A. P.; Waldmann, H.
  Natural Product Fragment Combination to Performance-Diverse Pseudo-Natural Products.
  Nat Commun 2021, 12 (1), 1883. https://doi.org/10.1038/s41467-021-22174-4.

- Liu, J.; Cremosnik, G. S.; Otte, F.; Pahl, A.; Sievers, S.; Strohmann, C.; Waldmann, H.
  Design, Synthesis, and Biological Evaluation of Chemically and Biologically Diverse Pyrroquinoline Pseudo Natural Products.
  Angewandte Chemie International Edition 2021, 60 (9), 4648–4656.
  https://doi.org/10.1002/anie.202013731.

- Christoforow, A.; Wilke, J.; Binici, A.; Pahl, A.; Ostermann, C.; Sievers, S.; Waldmann, H.
  Design, Synthesis, and Phenotypic Profiling of Pyrano-Furo-Pyridone Pseudo Natural Products.
  Angewandte Chemie International Edition 2019, 58 (41), 14715–14723.
  https://doi.org/10.1002/anie.201907853.