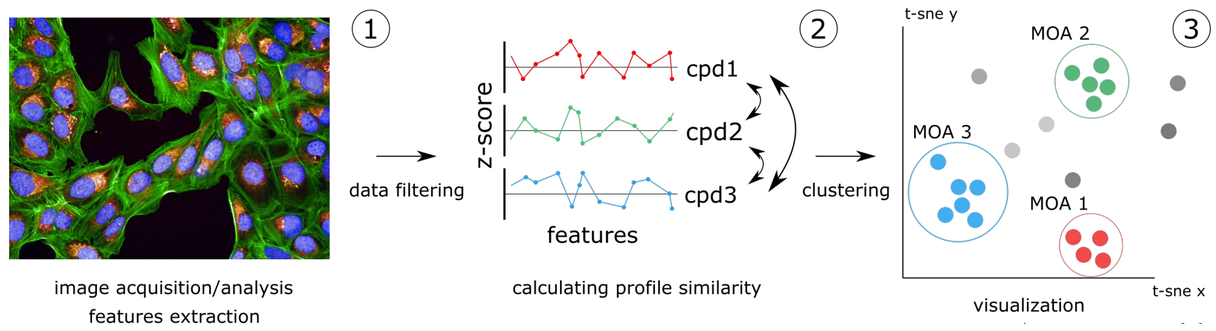
Cell Painting for compounds clustering and Mechanism Of Actions characterization

Julien Bellis[1], Quentin Guichard[1], Bruno Chilian[2], Karsten Kottig[2], Pierrick Rochard[1], Guillaume Mondesert[1]

[1]Evotec (France) SE, 195 route d'Espagne, 31036 Toulouse Cedex, France; [2]Evotec (Germany) SE, Manfred Eigen Campus Essener Bogen 7, 22419 Hamburg, Germany

Graphical abstract



Cell Painting methodology overview: Automated image analysis generates compound profiles. After comparison of all profiles, t-SNE visualization is used to evidence clustering of compounds with similar effect.

Introduction

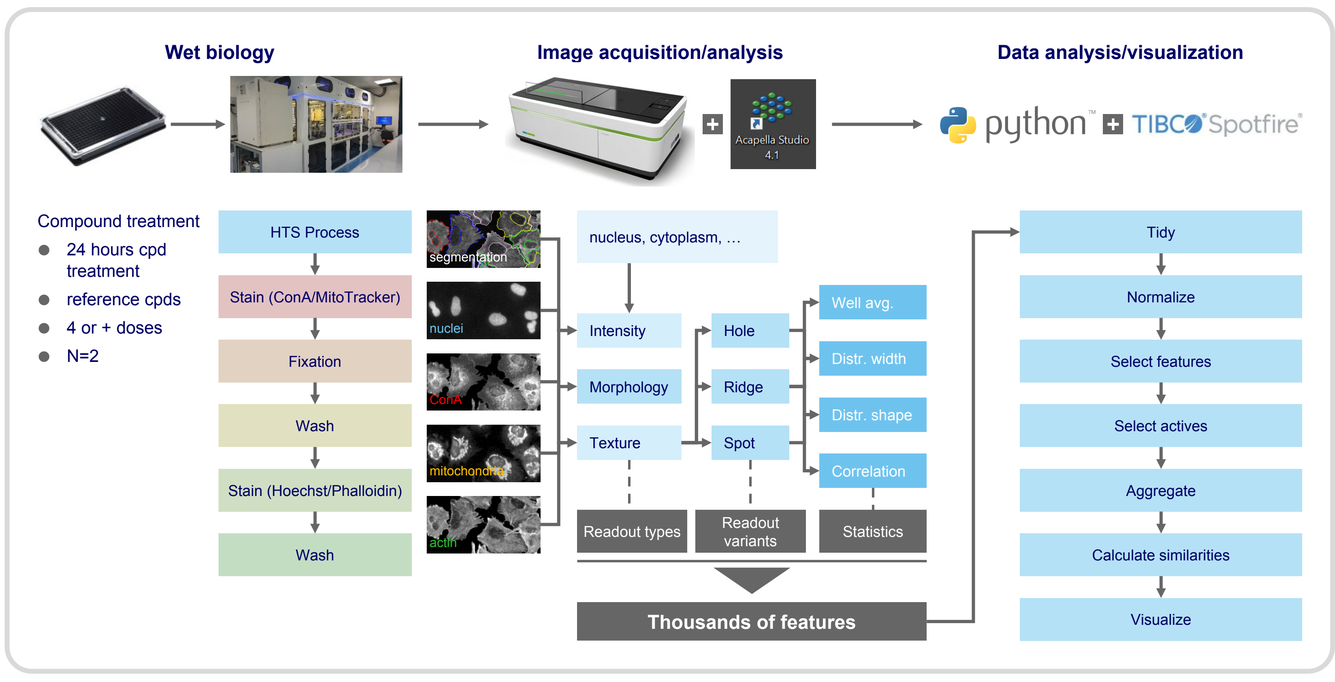
Cell Painting is a multi-parameter image-based description of the cell response to any perturbator condition: treatment with a compounds, a siRNA, CRISPR engineering, …

Treated cells are labelled with general fluorescent markers, and thousands of features are extracted by image analysis to build a profile, specific of a given perturbation. Assuming that similar perturbations lead to similar profiles, distance measurement between profiles is used to cluster compounds with similar impact on the cells.

We have implemented Cell Painting workflow at Evotec to allow for the study of several thousands of compounds, using automated process for cell treatment and labelling, image analysis, data processing and quality control.

We applied this workflow to the characterization of an internal compound collection, built with ~5,000 bioactive and annotated compounds. This work confirmed the power of Cell Painting approach to group compounds with similar MOAs and led to the selection of ~250 compounds to be used as a reference set in new studies. It also paved the way for further optimization, in particular to improve sensitivity of the technology.

Methods



We used U-2 OS cells seeded in 384 well plates and stained with 4 markers: Hoechst, Phalloidin, Mitotracker, ConcanavalinA. Imaging was performed on the Phenix® automated microscope (Perkin Elmer) using a 20x water objective and 6 fields per well.

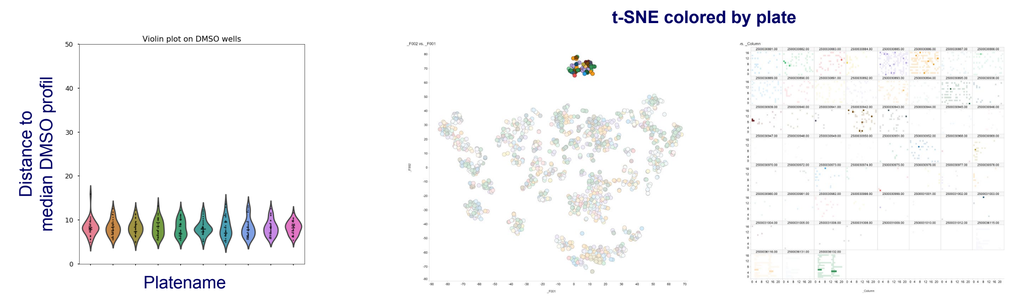
Typical throughput: 15k compounds in duplicate treated and imaged per week.

Applications of Cell Painting

- Predict the MOA of unannotated compounds by comparison with a set of reference compounds
- Group a large collection of unannotated compounds into clusters that harbor the same MOA → help maximize compounds profile diversity during hit triaging
- Identify compounds with potentially new MOA
- Identify compounds that revert to a control condition from a treated condition
- Identify cell line specific effect by comparing the compounds' profiles across different cell lines
- Follow evolution of similarity to reference compounds upon compound concentration increase
- Identify concentrations of compounds that have off-target effects

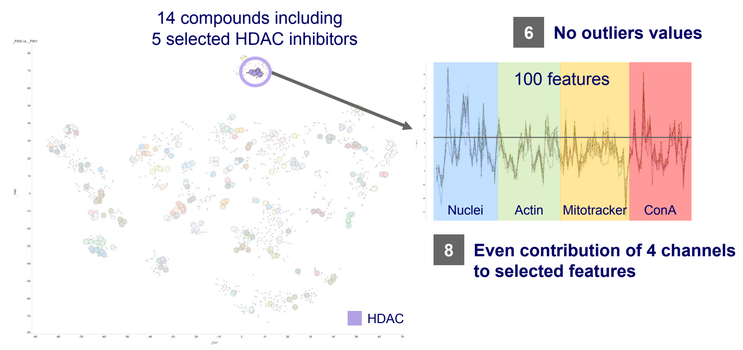
Quality Control

1. Sufficient cell count
2. Low variability for control wells
3. Control conditions mostly inactive
4. No time wise shift
5. No pattern due to cell seeding
6. No outliers values in profiles
7. No correlation between plate compound localization and t-SNE output
8. Even contribution of 4 channels to selected features
9. No spread of replicates in the t-SNE (non aggregated values)
10. Grouping of similar MOAs (aggregated values)



2 **Low variability for control wells**

7 **No correlation between plate compound localization and t-SNE output**



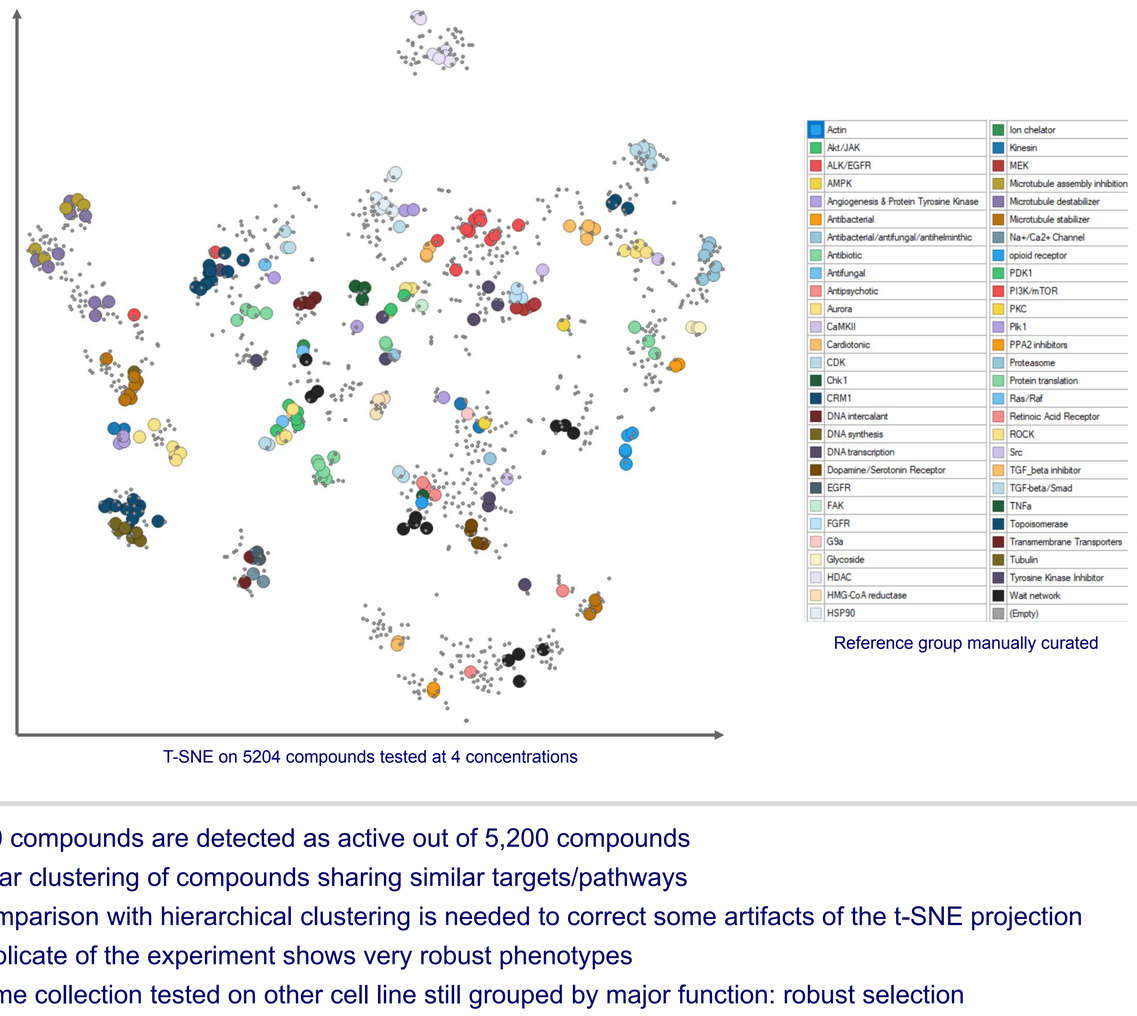
10 **Grouping of similar MOAs**

6 **No outliers values**

8 **Even contribution of 4 channels to selected features**

Cell Painting experiment quality control assessment: several parameters are controlled before any results interpretation

Results



Reference group manually curated

T-SNE on 5204 compounds tested at 4 concentrations

- 850 compounds are detected as active out of 5,200 compounds
- Clear clustering of compounds sharing similar targets/pathways
- Comparison with hierarchical clustering is needed to correct some artifacts of the t-SNE projection
- Duplicate of the experiment shows very robust phenotypes
- Same collection tested on other cell line still grouped by major function: robust selection
- 225 compounds, representing 57 major groups were selected as reference set

Cell Painting on the Bioannotated collection: selection of reference compounds from all phenotypic clusters

Conclusions and next steps

We have established a robust workflow to characterize with Cell Painting several thousands of compounds. Automated processes have been developed for the wet biology part as well as for data processing. After extensive quality control, data can be manipulated and visualized under different views in Spotfire.

We first confirmed the power of the technology to evidence groups of compounds with similar mechanisms of action. Among 5,200 bioactive compounds, we selected a set of 225 molecules representing 57 major phenotypic groups. Interestingly, when tested against different cell lines and at different time points, this set has shown reproducible clustering of compounds targeting similar pathways.

We are now using Cell Painting to support hit triage at the end of a High Throughput Screening, in order to select series with optimized phenotypic characteristics, for example to avoid major off-target effect or keep some degree of biological diversity.

We can also efficiently screen several thousands of compounds to select a few hundreds for characterization in a lower throughput assay.

Surprisingly, we found in our study with the bioactive collection that only 850 compounds out of 5,200 were considered as active, i.e. showing a general profile statistically different from the one of control vehicle. This suggests that the sensitivity of the assay could be improved.

We are therefore testing different approaches looking for improved sensitivity:

- Use new combinations of markers, including multiplexing
- Replace well level by cell level analysis
- Use artificial intelligence for image analysis