Automatic end-to-end analysis of high-throughput in vitro cell culture screening by HTSplotter

Carolina Nunes[1,2], Jasper Anckaert[1,2], Fanny De Vloed[1,2], Jolien De Wyn[1,2], Kaat Durinck[1,2], Jo Vandesompele[1,2], Frank Speleman[1,2,], Vanessa Vermeirssen[1,2,3*]

1 Department of Biomolecular Medicine, Ghent University, Ghent, Belgium
2 Cancer Research Institute Ghent (CRIG), Ghent, Belgium
3 Department of Biomedical Molecular Biology, Ghent University, Ghent, Belgium

* vanessa.vermeirssen@ugent.be

Abstract

Biomedical researchers are moving towards high-throughput screening, as this allows for automatization, better reproducibility and more and faster results. High-throughput screening experiments encompass drug, drug combination, genetic perturbagen or a combination of genetic and chemical perturbagen screens. These experiments are conducted in real-time assays over time or in an endpoint assay. The data analysis consists of data cleaning and structuring, as well as further data processing and visualisation, which, due to the amount of data, can easily become laborious, time consuming, and error-prone. Therefore, several tools have been developed to aid researchers in this data analysis, but they focus on specific experimental set-ups and are unable to process data of several time points and genetic-chemical perturbagen screens together. To meet these needs, we developed HTSplotter, available as web tool and Python module, that performs automatic data analysis and visualisation of either endpoint or real-time assays from different high-throughput screening experiments: drug, drug combination, genetic perturbagen and genetic-chemical perturbagen screens. HTSplotter implements an algorithm based on conditional statements in order to identify experiment type and controls. After appropriate data normalization, HTSplotter executes downstream analyses such as dose-response relationship and drug synergism by the Bliss independence method. All results are exported as a text file and plots are saved in a PDF file. The main advantage of HTSplotter over other available tools is the automatic analysis of genetic-chemical perturbagen screens and real-time assays where results are plotted over time. In conclusion, HTSplotter allows for the automatic end-to-end data processing, analysis and visualisation of various high-throughput *in vitro* cell culture screens, offering major improvements in terms of versatility, convenience and time over existing tools.

Author summary

Academic researchers are moving towards high-throughput screenings, where the experiments execution follows an automatic approach, such as robotic seeding, liquid dispensing and/or automatics readouts. This grants more flexible and reproducible experimental set ups. The type of high-throughput experiment can vary from drug, drug combination, genetic perturbagen to genetic-chemical perturbagen screens. These can be assessed through endpoint assays, measuring the effect at one time point, or real-time assays, assessing the phenotypic effect over time. High-throughput screening results in large amounts of data, requiring laborious, time consuming and error-prone data handling and analysis. These challenges hamper the biological interpretation of data and the fast progress in drug development programs. Hence, we developed HTSplotter, a web tool and Python module, to allow researchers to conduct a fast and flexible end-to-end data processing, analysis and visualisation of different high-throughput experiments, assessed either by endpoint or

real-time assays, relieving the high-throughput analysis bottleneck. In this way, HTSplotter, directly contributes to a faster identification of drugs and drug combinations, with potential practical applications, or novel targets for therapies.

Introduction

Diseases such as cancer, bacterial and viral infections, metabolic disorders, and neurodegenerative diseases need more effective drugs and new therapeutic approaches[1–4]. Therefore, biomedical researchers are focusing on identifying new druggable targets and combination treatments. In parallel, this presses researchers to pursuit reproducible, effective and efficient experiments leading to more and faster results. In order to address these needs, researchers are implementing experiment automatization, such as robotic cell seeding and liquid dispensing, and/or automatic readouts [2–5]. Whilst academic research is moving towards this new experimental paradigm of high-throughput screening (HTS), industrial settings already deal with 10^6 screenings per day [6].

HTS aids researchers to perform experiments in one or more cell lines with a large number of drugs, alone or in combination, and genetic perturbagen screens, such as RNAi and CRISPR screens [6–8]. Additionally, a genetic perturbagen (gene knockdown or overexpression) can be combined with a drug as to evaluate the response in a specific disease context, leading to the identification of certain genes that offer resistance to a set of drugs, or as to identify candidate genes that might work synergistically with drugs [9, 10]. The assessment of drug, drug combination and genetic perturbagen screens can be conducted in endpoint or real-time assays. The latter monitors cell proliferation at regular time intervals over days or weeks (e.g. using Incucyte or CytoSPAR machines), while the endpoint assay allows for the assessment of cell viability at single time points [6, 11]. Although both assessments can be applied to medium/high-throughput screens, the endpoint assay is the standard option for large HTS scale [6]. Regardless of the assay, the cornerstone is to measure the impact of the perturbagen on the cellular response in a specific disease context. Before proceeding to the biological interpretation of the results, the raw data must be processed and analysed. However, the large number of conditions and/or time points generated, by either endpoint or real-time assays, may result in a laborious, time consuming and error-prone data processing and analysis, which may delay and hinder drug development programs [12, 13].

At the very minimum, downstream processing of HTS consists of comparing control and treatment conditions. However, this might not be sufficient to achieve biological interpretation of the data. When a drug is tested in a dosage range, the dose-response relationship is commonly evaluated at a specific time point. Amongst several models used model to fit the dose-response curve, the most common is the four-parameter logistic. It allows the determination of the maximal effect at the highest dose tested (Emax), the dosage at which half of the total efficacy is observed (EC50) and the dosage that provokes 50% of inhibition (IC50). The EC50 and IC50 values can differ, since the EC50 is the inflection point of the dose-response curve and the IC50 is the concentration at which half of the response is achieved. The area under the curve (AUC) is an additional metric used to evaluate the drug's potency, where potency and AUC are inversely proportional [14–16].

When combining different drugs, one typically aims to quantify the degree of synergism/antagonism. Highest single agent (HSA), Loewe additivity, Zero interaction potency (ZIP) and Bliss independence (BI) are commonly used models that assume a non-interaction between drugs and compare the observed and expected combination responses [12, 17]. These are, however, based on different assumptions. The Loewe model compares the dose-response relationship of individual drugs. HSA compares the effect of drugs in combination to the highest inhibition effect of a single drug, while the BI model compares the combination effect to the expected effect of independent drugs. Lastly, the ZIP model assumes that two non-interacting drugs are expected to incur minimal changes in their dose-response curves upon combination [12, 17–19]. Based on the requirements for each approach, we characterized the Loewe and ZIP as dose-effect models and the HSA and BI as effect-based models [18, 19]. Dose-effect based models work if monotherapy dose-response curves are well characterized, which, depending on the response of a cell line to a drug dose range treatment, might not be achievable [19, 20]. As for the effect-based models, knowledge of the inhibition effect of its individual components is necessary to predict the combined effect. Hence, effect-based models can be applied in case of combinations of genetic-chemical perturbagens [9]. Additionally, the HSA method is less conservative, given that it focuses on the response of a single compound,

disregarding effects of competing compounds, thus increasing the likelihood of detecting false synergism/antagonism.

In order to allow researchers to get the most out of HTS, specific software has been developed to enable high-throughput data analysis, such as BREEZE for drug screens, CellHTS2 for siRNA, and SynergyFinder for drug combinations [12, 21, 22]. Depending on the type of experiment, one must organize the data as to fulfil the tool required input structure. Since these tools are designed for endpoint assays, data resulting from real-time assays implies repetitive work from the user.Moreover, these tools are not suited to analyse combinations of chemical and genetic perturbagens, requiring researcher to conduct own data analysis manually. To the best of our knowledge, there is currently no software with the ability of implementing automatic end-to-end analyses irrespective of HTS experiment and irrespective of assay type, endpoint or real-time. To address these limitations, we developed HTSplotter as a freely available web tool and Python module. HTSplotter is tailored to analyse drug, drug combination, genetic perturbagen and combinations of genetic-chemical perturbagen screens, either in real-time or as endpoint. HTSplotter identifies the type of experimental setup through a conditional statement algorithm. It performs a normalization and, in case of a drug screen, drug combination or genetic-chemical perturbagen experiment, identifies the dose-response relationship for each drug alone. Additionally, synergism/antagonism of drug or genetic-chemical combination screens is determined through the BI method. Finally, results are plotted and exported as PDF files, allowing a fast biological interpretation of the data.

Design and implementation

HTSplotter: an end-to-end tool for HTS

HTSplotter is written in Python 3.9 using SciPy, h5py, NumPy, Matplotlib, math, os and sys libraries. It automatically identifies four experiment types: drug, drug combination, genetic perturbagen and genetic-chemical perturbagen screens. A genetic perturbagen screen can be a simple knockdown or overexpression of a gene, a CRISPR/Cas9 screen or a siRNA library, while the genetic-chemical perturbagen screen consist of a genetic perturbagen in combination with a drug. These can be conducted either in endpoint or real-time assays. The evaluation of each experiment type consists of a normalization relative to the control. Additionally, in case of drug, drug combination and genetic-chemical perturbagen screens, the dose-response relationship is calculated for each drug. For drug combination and genetic-chemical perturbagen screens, the BI method is applied to identify synergism/antagonism on multiple drug combinations.Combinations can be arranged in a matrix $\Pi_{i=0}^{N} n_i$, where N is the number of drugs, n_i is the number of concentrations of drug i and $n_i \geqslant 1$.For all experiment types, HTSplotter has no limits regarding the number of cell lines tested on each experiment. Fig 1 summarizes the analyses and visualisations provided by HTSplotter.

HTSplotter: from input file to analysis

HTSplotter utilizes text files directly exported from real-time devices, such as Incucyte and xCELLigence, as inputs, where comment lines above the data to be processed are allowed. "Date Time" and "Elapsed" strings must be present at the first and second data columns respectively, so that the software detects the data headers row (Fig 2), holding the description of the experimental conditions. These columns have date and time points corresponding to each measurement, respectively, e.g. apoptosis, confluency or impedance. In case of more than one time point, called in this paper as time point intervals, the measurements correspond to time increments from the start of the experiment e.g. 2h, 12h, etc. The information regarding each experimental condition should contain drug name, gene knockdown or overexpression, concentration, units, cell line name and seeding density in one column separated by commas, thus "." should be used as decimal separator. More detailed information regarding data and header structure are available at the HTSplotter website (`https://htsplotter.cmgg.be/`).

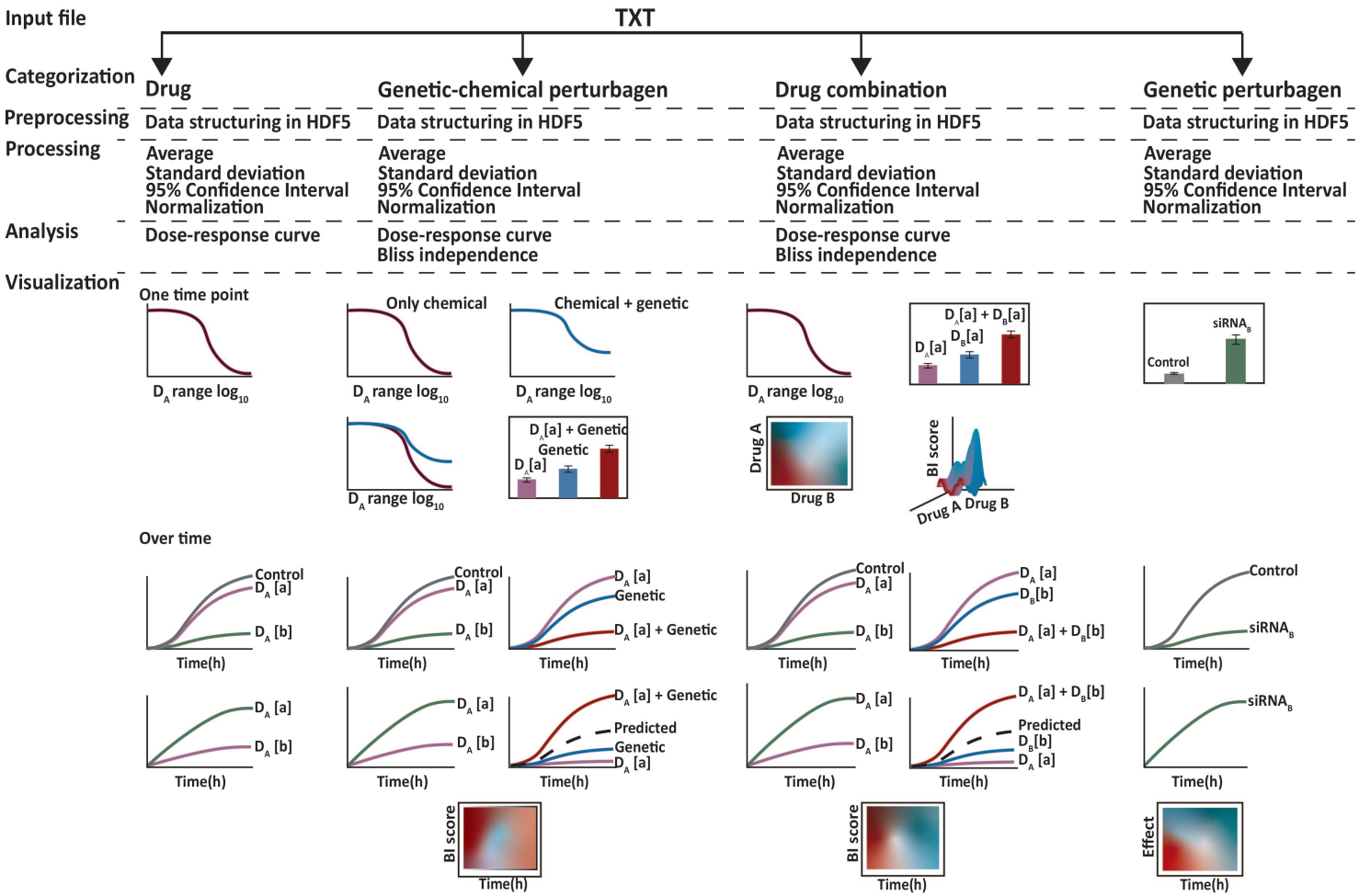


Fig 1. Overview of HTSplotter steps for each type of HTS experiment. The input file, directly imported from HTS machines like Incucyte S3 as a TXT file, is automatically processed and analysed by HTSplotter. As output, TXT and PDF files are generated. The PDF file contains the plots from each type of analysis. The y-axis title is defined by the user, where readout and the expected effect should be provided, e.g. confluency and inhibition, respectively. **Genetic perturbagen screen:** a heatmap and bar plots are provided with all the perturbagens tested, for specific time points. Additionally, a XY-plot is created for each perturbagen. **Drug screen:** a dose-response curve of each drug alone is plotted with associated statistical parameters. XY-plots are also returned with all concentrations of each tested drug over time. **Drug combinations and genetic-chemical perturbagen screens:** a dose-response curve of each drug alone is provided with their statistical parameters. In case of genetic-chemical perturbagen, as to allow for a direct comparison of drug response without and with the perturbagen, a dose-response curve for both situations is plotted. Also, for specific time points, a bar plot containing information about each condition is given. Moreover, XY-plots visualize all concentrations of each tested drug over time as well as each combination with respective drugs alone. A heatmap with BI scores over time is created, for either drug combination or genetic-chemical combination screens. Additionally, in case of a drug combination, resulting in a combination matrix ($m \times n$, where $m \geqslant 2$), a heatmap for each main time point (2D and 3D) is given, whilst in case of $m = 1$ only a 2D heatmap is plotted.

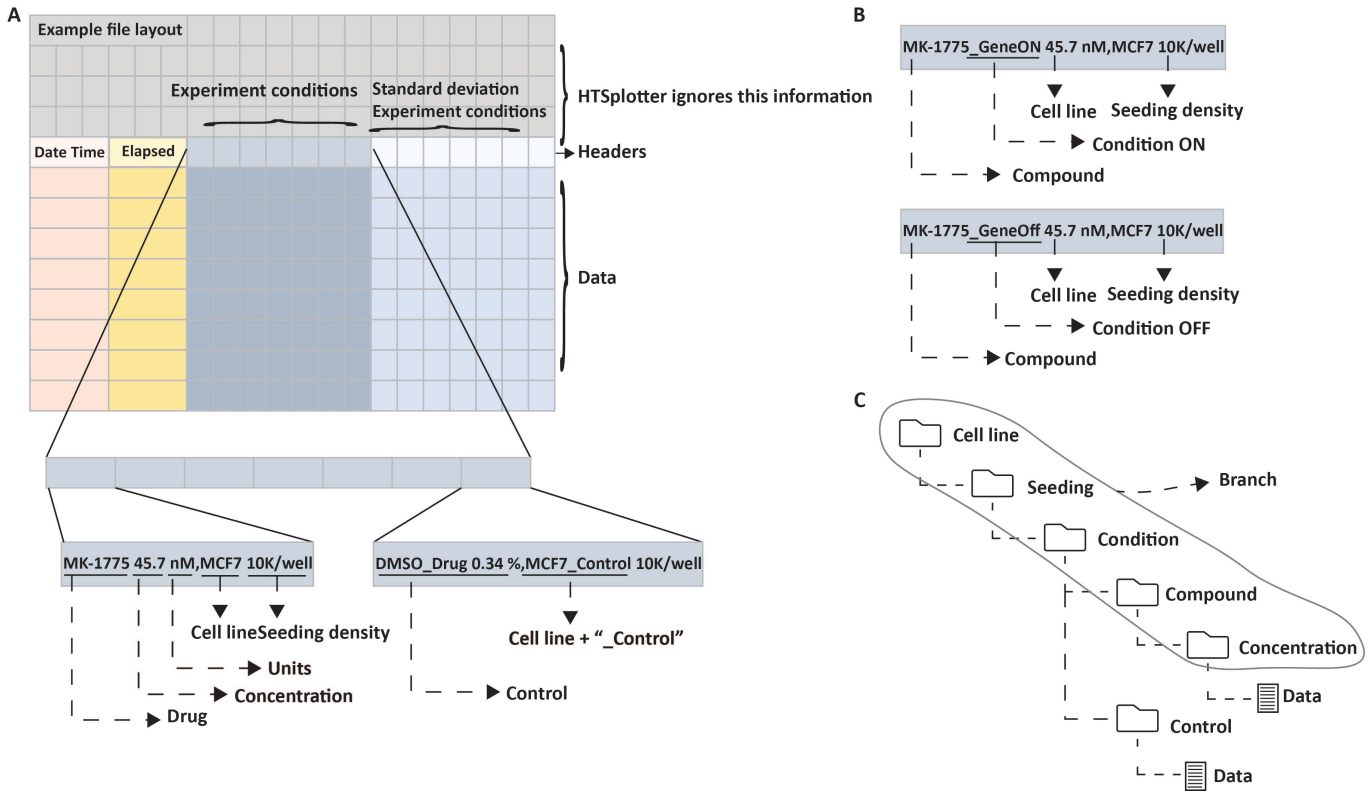


Fig 2. An example input file structure, experimental condition and data structure. A: Example of an Input file structure with a detailed view on the experiment condition description. Header and data consist of four sub-groups: "Date Time", "Elapsed", "Experiment". On the left is the description of the MK-1775 drug at a dosage of 45.7 nM on the MCF-7 cell line with 10k cell seeding per well and on the right the control condition, DMSO at 0.34% on the MCF-7 cell line with 10K seeding per well. B: Example of an experimental condition description for a genetic-chemical perturbagen. C: Data storage structure in the HDF5 file. A branch is created when cell line, seeding, condition and drug are grouped in the same experimental set-up

HTSplotter implements an algorithm based on conditional statements that extracts information about the number of cell lines, seeding density, conditions, drugs and concentrations tested. In this way, information about similar experiments in a cell line with the same seeding density and conditions for multiple tested drugs, are grouped creating a branch. The conditions can also be gene knockdown or overexpression. The experiment is then categorized as drug, drug combination, genetic perturbagen or genetic-chemical perturbagen screen according to the number of branches, number of identified controls and number of concentrations tested per drug. Categorization is a crucial step for HTSplotter, as it leads to a proper evaluation of the perturbagen of each condition and to the implementation of downstream analyses required for each experiment type, therefore requiring user confirmation about predicted categorization and identified control(s). In case of incorrect categorization, an error file with information regarding data headers and possible missing information is created. Otherwise, the experiment information is stored in a hierarchical data format file (HDF5), the structure of which is determined by the type of experiment.

Subsequently, HTSplotter normalizes the data to the proper control(s). The resulting phenotypic measurement (e.g. confluency or apoptosis) can be inhibition or enhancement, this option being given to the user before the analysis. If the standard deviation is present in the input file, the tool rescales it. In case of repetitive conditions, being these technical or biological replicates in the same file or in independent files, HTSplotter computes averages, standard deviations and 95% confidence intervals. For genetic perturbagen screens with more than one control, HTSplotter calculates averages, standard deviations and 95% confidence intervals of all controls as to have a unique control, and only then the normalization takes place. The HDF5 file includes all information from the independent experiments.

Upon a minimum of two concentrations in a drug, drug combination or genetic-chemical perturbagen screen, a dose-response curve of each tested drug is calculated. Additionally, in the case of a drug combination or genetic-chemical perturbagen screen, the BI method is applied to determine synergism/antagonism. First, HTSplotter determines the predicted combination effect by the BI method [23], Eq (1), where $P_{A_\partial B_\delta}$ is the predicted inhibition effect produced by the combination of drug A and B, at doses ∂ and δ, being A_∂, B_δ the inhibition effect of drugs A, B at doses ∂, δ, respectively.Next, the BI score is determined by the difference between $O_{A_\partial B_\delta}$, observed combination inhibition effect, and $P_{A_\partial B_\delta}$, Eq (2).If BI score is higher than zero, the combination is synergistic, if below zero then it is antagonistic. Following these analyses, data is exported as text and results are plotted in a PDF file.

$$\mathrm{P}_{\mathrm{A}_\partial \mathrm{B}_\delta} = (A_\partial + B_\delta) - (A_\partial B_\delta), \tag{1}$$

$$\mathrm{BI}_{\mathrm{score}} = (O_{A_\partial B_\delta}) - P_{A_\partial B_\delta}, \tag{2}$$

HTSplotter: visualisation

HTSplotter allows the analyses of endpoint and real-time assays, regardless of the experiment type. If the input file contains more than one time point, additional plots with time features are provided, as well as plots at preselected time points (e.g. 24h, 48h, etc.). A feature of HTSplotter is the XY-plot, where the read-out (e.g. confluency or apoptosis) and either inhibition or enhancement effects (Y) are plotted over time (X) leading to better biological interpretations of the chemical or genetic perturbation effect. Upon drug or genetic-chemical combinations, the predicted combination effect is plotted together with the observed combined and single effect. In this way, XY-plots and bar-plots allow a detailed observation of the BI synergism/antagonism scores. Another unique feature of HTSplotter is the visualisation of BI scores over time in a heatmap, allowingto identify the onset of synergism/antagonism. In case of drug combinations, when the number of concentrations is higher or equal to two, 2D and 3D heatmaps are shown for main time points. Whilst in the case of a single concentration, only a 2D heatmap is shown. For all HTS drugging experiments, dose-response curves are displayed for main time points, regardless of the experiment type.

Results

Applying HTSplotter on a continuous assessment of in vitro cell based drug combinations: a case study

Drug combinations require multiple conditions, however, when monitored over time, the number of data points rapidly increases. Consequently, increasing the number of required computations. HTSplotter organizes the data according to experimental set-up to easily implement further data processing, analysis and visualisation. Data processing consists of a normalization of all conditions to the proper control. Analyses include the determination of the dose-response relationship (Fig 3) and synergism/antagonism (Figs 4, 5).

To illustrate the advantage of HTSplotter on drug combination evaluation assessed in real-time, we applied our tool to an in-house assessment of the human breast cancer cell line MCF-7 in response to single and combination exposure to three inhibitors targeting WEE1, CHK1 and ATR (MK-1775, prexasertib and BAY1895344, respectively). The combination of these inhibitors has been shown to be synergistic in different cancer entities, such as breast cancer[24–26]. The drug combination experiment consisted of 136 conditions and 37 time points, resulting in 5032 data points. The data used in HTSplotter was directly exported from Incucyte S3.

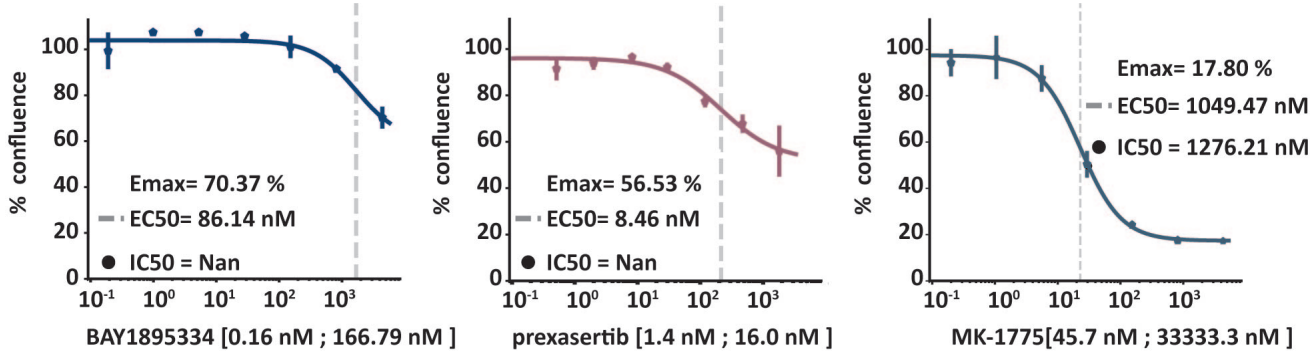


Fig 3. Dose-response curve provided by HTSplotter for BAY1895334, prexasertib and MK-1775 at 72h. A: Dose-response curve of BAY1895334, in which the Emax is 70.37% of confluency, the EC50 is 86.14 nM and the IC50 could not be determined. B: Dose-response curve of prexasertib, in which the Emax is 56.33% of confluency, the EC50 is 8.46 nM and the IC50 could not be determined. C: Dose-response curve of MK-1775, in which the Emax is 17.80% of confluency, the EC50 is 1049.47 nM and the IC50 is 1776.71 nM.

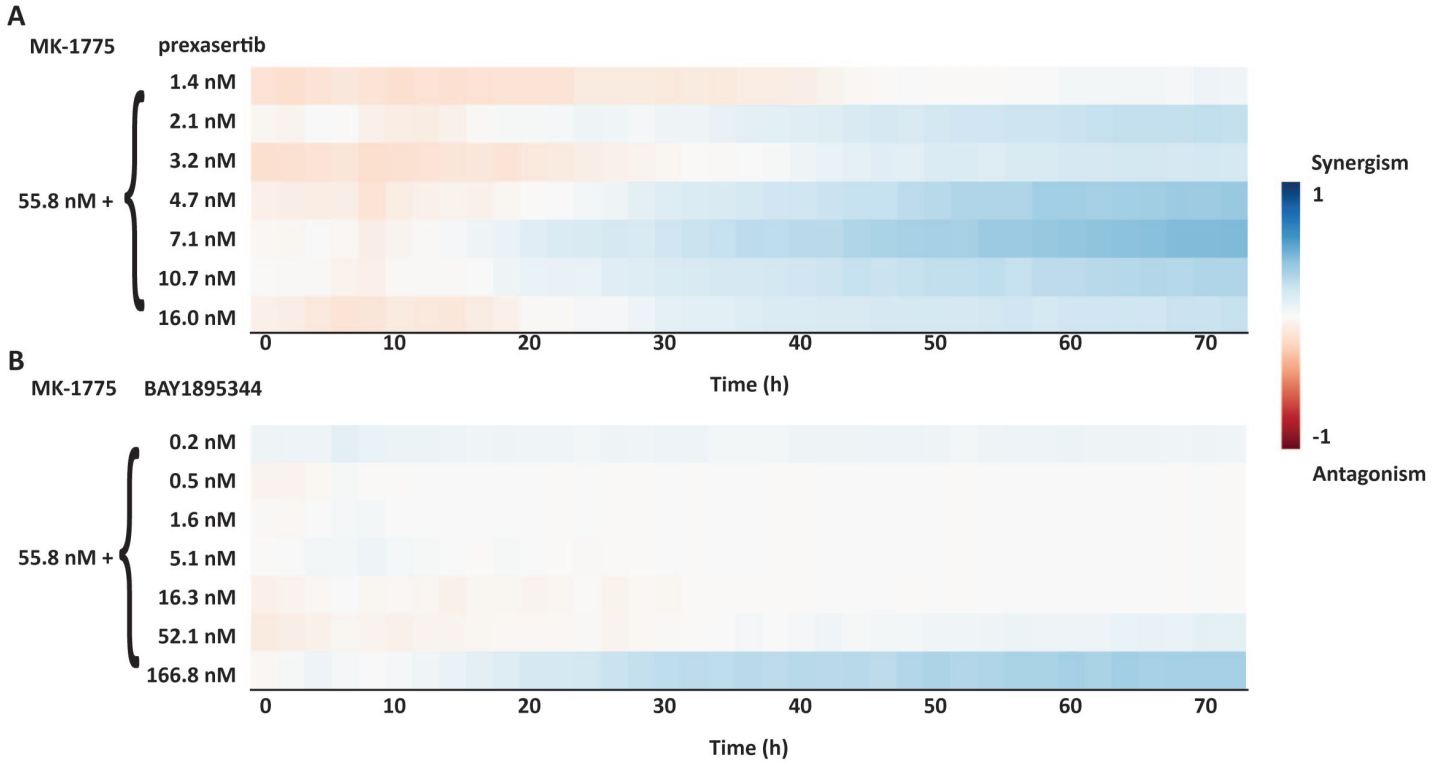


Fig 4. BI score heatmap of effect combination overtime, the blue gradient indicates synergism and the red gradient indicates the antagonism. A: Effect combination heatmap overtime of MK-1775 (55.8 nM) with a concentration range of prexasertib (1 x 7 matrix). B: Effect combination heatmap overtime of MK-1775 (55.8 nM) combined with a concentration range of BAY1895344 (1 x 7 matrix).

Each drug alone was assessed by the dose-response curve and their corresponding statistical metrics (Fig 3). At 72h, the tested dose-range of MK-1775 allowed to determine the IC50 (1276.21 nM), as opposed to other drugs. Accordingly,

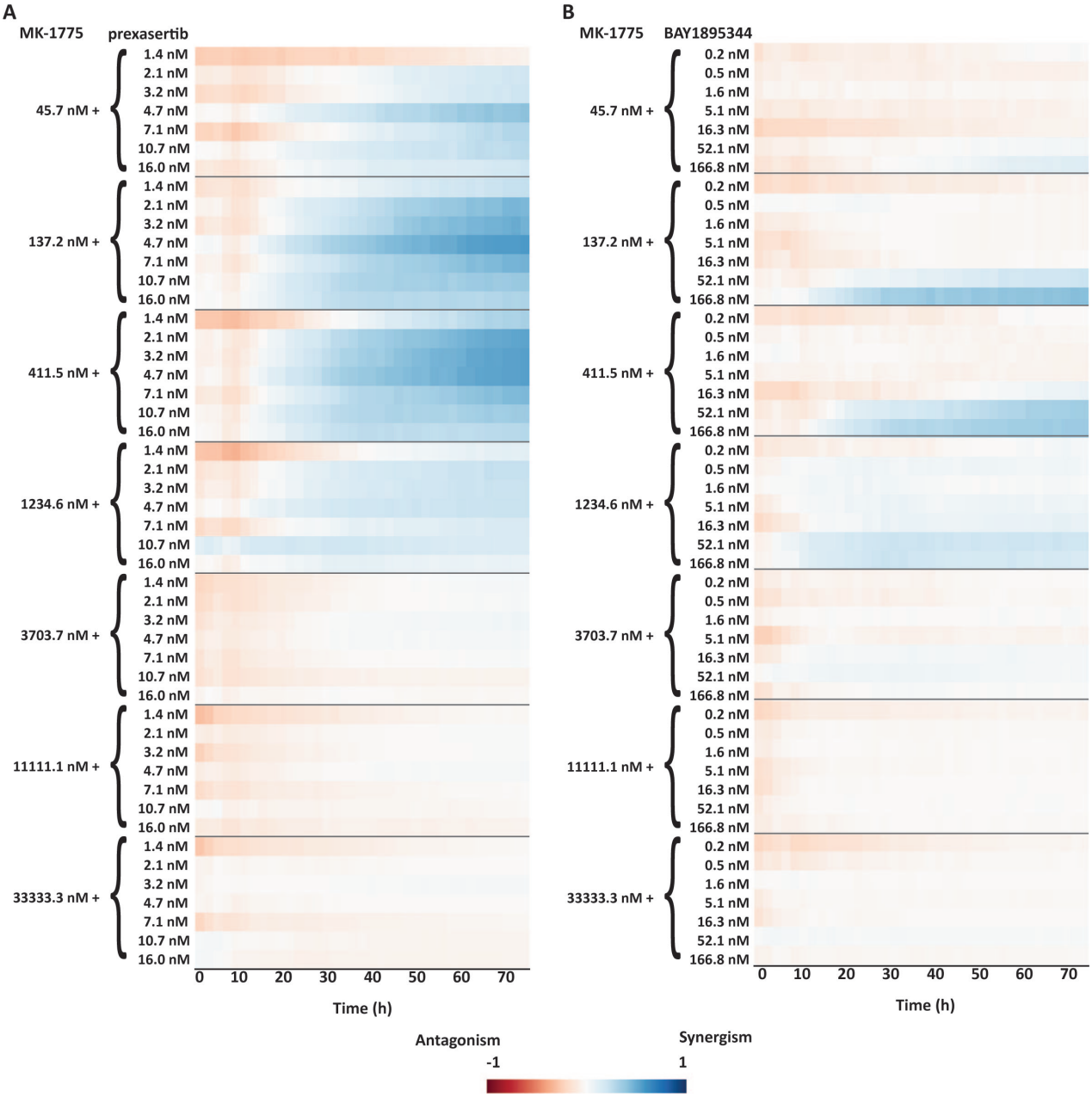


Fig 5. BI score heatmap of dose-effect combination overtime, in which the blue gradient indicates synergism and the red gradient indicates the antagonism. A: Dose-effect combination heatmap over time of a combination range of MK-1775 with a range of prexasertib (7 x 7 matrix). B: Dose-effect combination heatmap over time of a combination range of MK-1775 with range of BAY1895344 (7 x 7 matrix).

the Emax of both BAY1895334 and prexasertib, was above 50% of confluency, confirming that the IC50 was not achieved by the tested dose-range. Prexasertib and MK-1775 showed the lowest and highest EC50, respectively. Thus, prexasertib revealed to be the most potent drug, followed by BAY1895344 and MK-1775.

The drug combination setup encompassed a single effect of MK-1775 (55.7 nM) with a dose range of BAY1895344 or prexasertib (Fig 4), later referred as effect combinations, and a dose-response combination matrix of MK-1775 with

BAY1895344 or prexasertib, referred as dose-effect combinations (Fig 5). The effect combination is a similar experimental set-up as a genetic-chemical combination, which was analysed by HTSplotter in an analogous way.

HTSplotter revealed that prexasertib, for both combination setups, showed stronger synergism than BAY1895344 over time. Both heatmaps and XY plots illustrate nicely that the synergistic effects emerged at 20h of treatment for all combinations. Moreover, for some synergistic combinations, the BI score increased over time, indicating an increasing synergistic effect in this cell line, unlike the other synergistic combinations, where the increase of BI score was less pronounced. Assessing in detail the highest BI score of each dose-effect combination (Fig 6), BAY1895344 alone had a higher and increasing inhibitory effect when compared to MK-1775 and prexasertib, which both had almost no impact on the growth inhibition over time. Consequently, the predicted combination effect of MK-1775 with BAY1895344 increased over time, while the MK-1775 with prexasertib had a small and constant inhibition effect.

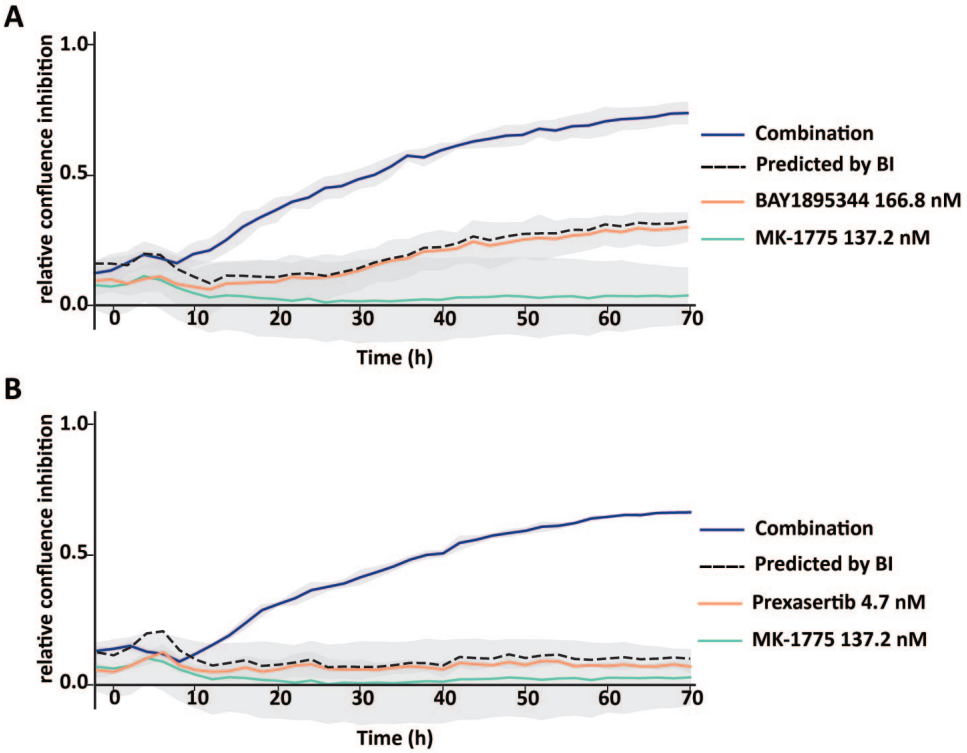


Fig 6. XY- plot of the best dose-effect combination of MK-1775 with BAY1895344 (A) or MK-1775 with Prexasertib (B). The dash line indicates the predicted combination effect, computed according to the BI method Eq (1). A: Relative confluence inhibition over time of BAY1895344 at 166.8 nM combined with MK-1775 at 137.2 nM. The BI score at 24h, 48h and 72h is 0.28, 0.39 and 0.42, respectively. B: Relative confluence inhibition plot overtime of prexasertib at 4.7 nM combined with MK-1775 at 137.2 nM. The BI score at 24h, 48h and 72h is 0.25, 0.47 and 0.56, respectively.

Comparison to related software

Over the years, several tools were developed for the analysis of different types of HTS, such as drug combination and genetic-chemical perturbagen screens. Some of those, such as RNAither, IncucyteDRC and DRC, require programming knowledge [14, 27, 28]. Also, more user friendly tools were developed (Table 1) [13, 15, 21, 22, 29, 30].

The available web tools and software are designed for specific experiment types, mostly endpoint assays. Often, one

Table 1. Open source software for HTS data analysis and their characteristics.

	BREEZE	GRCalculator	Combenefit	SynergyFinder 2.0	CellHTS 2	HTSplotter
Web tool	✓	✓	×*	✓	✓	✓
Endpoint assay	✓	✓	✓	✓	✓	✓
Real-time assay	×	✓	×	×	×	✓
Drug screen	✓	✓	×	×	×	✓
Drug combination#	×	×	✓	✓	×	✓
Genetic perturbagen	×	×	×	×	✓	✓
Genetic-chemical perturbagen	×	×	×	×	×	✓
Replicates	×	✓	✓	✓	×	✓
Dose-response matrix	✓	✓	✓	×	×	✓
Plate quality	✓	×	✓	×	✓	×
$\geqslant 1$ cell line/file	✓	✓	×	×	×	✓

* Available as a standalone application for Windows.

The Combenefit only allows pairwise drug combination, while SynergyFinder and HTSplotter allow multiple drug combination.

has to organize, structure and clean the data specifically to the tool. Depending on the selected tool, the user might have to normalize the data. Once the tool has been chosen, this procedure has to be repeated for each selected time point. Therefore, if a real-time assay was performed, at least the last step must be repeated as many times as selected time points. Additionally, in case of drug combination, some tools only provide a heatmap with synergism/antagonism scores, missing important information regarding the perturbation effect.

To compare HTSplotter on drug combination with the available tools, we resorted to SynergyFinder, a tool widely used for synergism/antagonism evaluation. SynergyFinder was applied to the dose-effect combination at 72h using the BI method. The tailoring of the data to the specific input format of SynergyFinder was performed by HTSplotter. As for the case of effect combination (1×7) (S1 Fig), where one dosage of MK-1775 was combined with seven dosages of prexasertib or BAY1895344, SynergyFinder was not able to evaluate synergism/antagonism, since two concentrations of each drug is the minimum requirement for a pairwise combination.The dose-response curve profile matched the one provided by HTSplotter, however, the statistical metrics of the dose-response relationship of the single agent are not calculated by SynergyFinder.

SynergyFinder identified the same synergistic dose range as the HTSplotter using the BI method (Fig 7), with minor differences in scores due to different approaches: SynergyFinder computes the averaged over the full dose-response matrix, while HTSplotter computes synergism between tested concentrations. When multiple time points are assessed, HTSplotter provides a BI score heatmap over time, indicating if and when synergism starts and ends,a feature not provided by SynergyFinder, where if one is interested in another time point, one has to prepare a new input file. Moreover, HTSplotter allows a detailed comparison between combination effect and each drug effect, over time (XY-plots, 6) and for specific time points (bar plot, Fig S2 Fig). In this way, the researcher observes the effect of each drug alone and predicted and observed effects upon combination, which may aid the optimization of further experiments as well as the identification of the best concentrations for combination experiments.

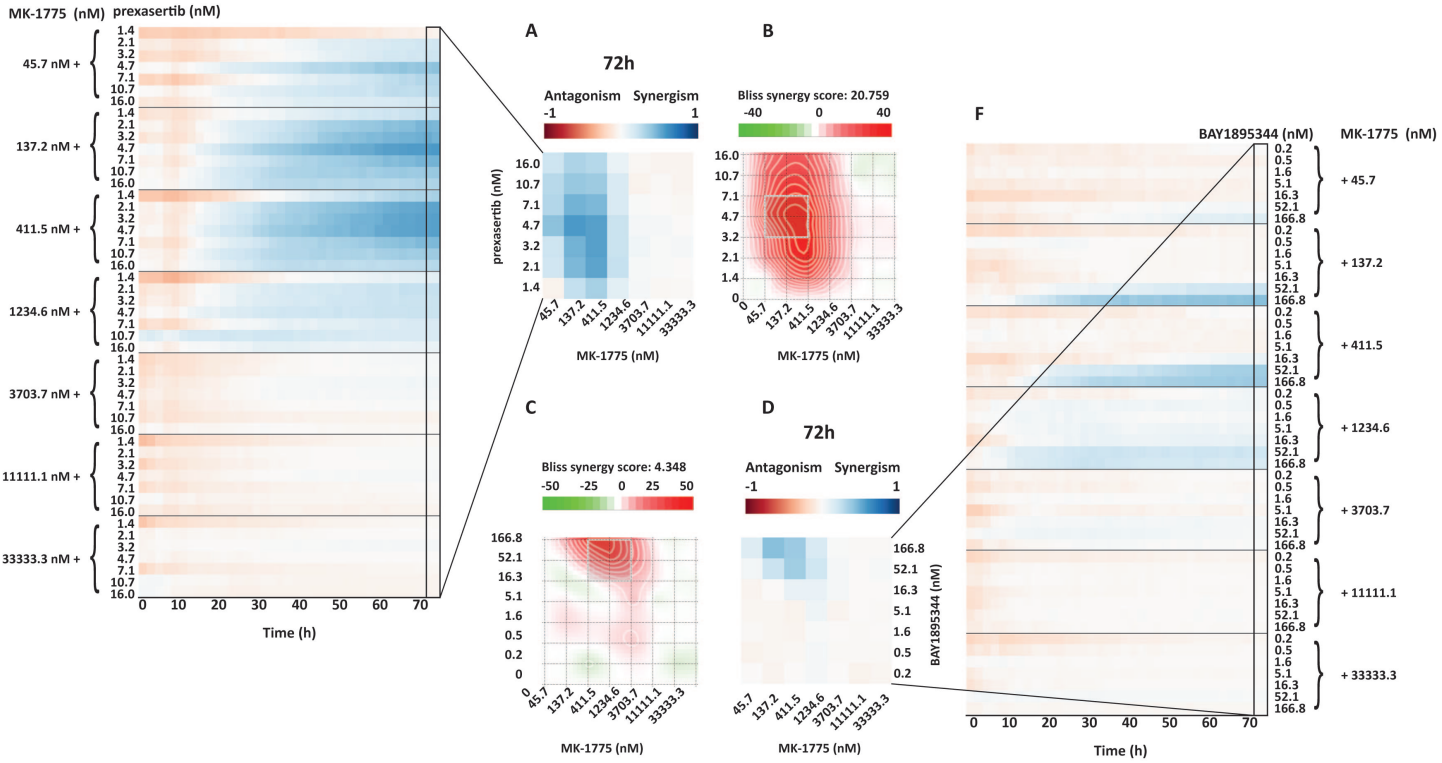


Fig 7. SynergyFinder 2.0 and HTSplotter heatmap at 72h. A: HTSplotter heatmap over time and at final time point (72h) of dose-effect combination of prexasertib with MK-1775, maximum BI score of 0.56 at 72h. B: SynergyFinder heatmap of dose-effect combination of prexasertib with MK-1775, with BI score of 20.759. C: SynergyFinder heatmap of dose-effect combination of BAY1895344 with MK-1775. D: HTSplotter heatmap over time and at final time point (72h) of dose-effect combination of BAY1895344 with MK-1775, maximum BI score of 0.42 at 72h. HTSplotter has a fixed legend scale from -1 to 1.

Conclusion

Several tools have been developed to aid researchers in the data analysis of HTS. However, these lack analytical flexibility on the processing and downstream analyses for several time points and for genetic-chemical perturbagen screens. Therefore, we developed HTSplotter to analyse drug, drug combination, genetic perturbagen and genetic-chemical perturbagen screens, both in endpoint and in real-time assays. In this paper, we demonstrated that HTSplotter is able to perform a tailored end-to-end analysis, including processing, dose-response and synergism/antagonism analyses, as well as high quality visualisations, requiring minimum user intervention. Moreover, it provides a unique, detailed visualisation of the drug combination effect over time, going beyond the state-of-the-art, for a better biological interpretation. Furthermore, in contrast to other tools, HTSplotter analyses a varying number of concentrations in pairwise and multi-drug combinations. Due to this flexibility, HTSplotter is also fit for genetic-chemical combination screens, such as siRNA/drug or CRISPR/drug experiments. Hence, HTSplotter is a major contribution to biomedical research with efficient and effective data analysis for high-throughput screening.

Availability and Future Directions

HTSplotter is freely available as a web-application, `https://htsplotter.cmgg.be/`, and as a Python module, `https://github.com/CBIGR/HTSplotter`, together with an outlined documentation including step-by-step user instructions and examples. Data examples for drugs, drug combination, genetic perturbagen and genetic-chemical perturbagen screens are provided at the website.

Given the importance of determining synergism/antagonism using different models, future HTSplotter updates will focus on the integration of the Loewe and ZIP models.

Material and methods

Dose-response curve and drug combination analysis

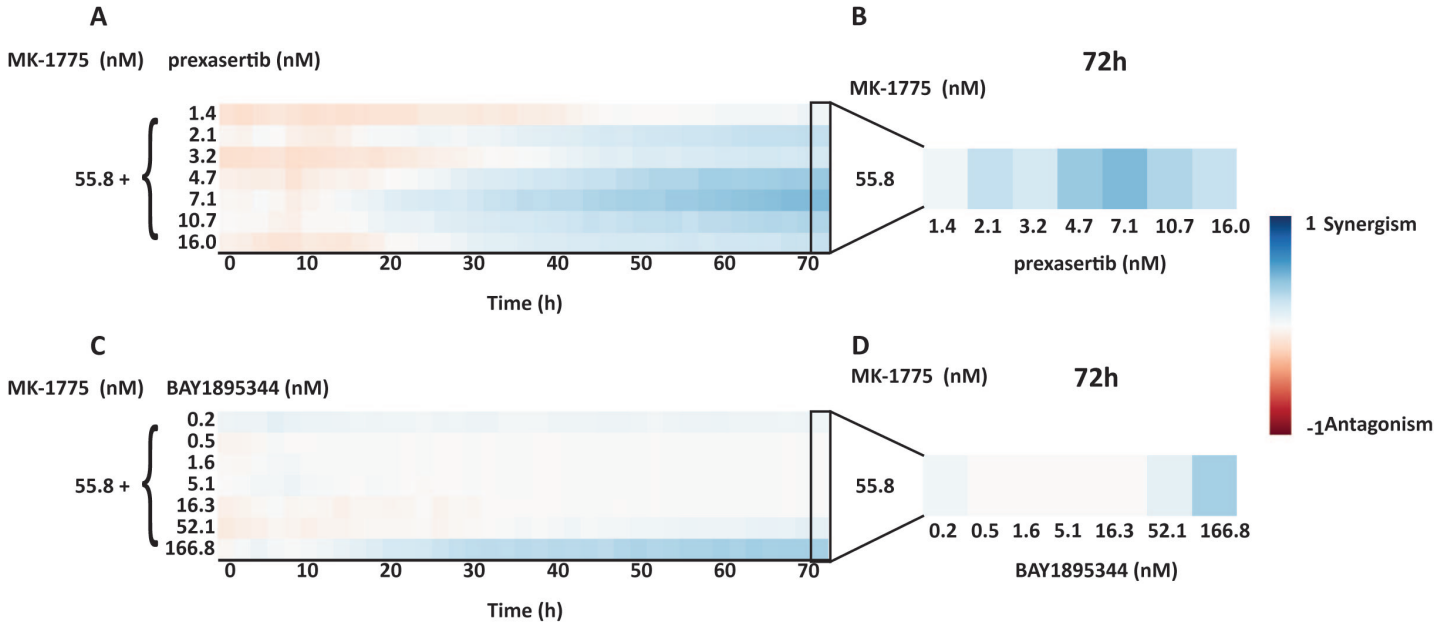
HTSplotter implemented a four-parameter logistic function, testing the curve fitness usingthe Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting (SciPy 1.6.0). The curve fitting statistical parameters, such as Chi-square, residuals and R-squared were calculated. For the calculation of the AUC, the trapezoidal rule was applied (NumPy 1.19.5). To compute the different ICs, HTSplotter determined the four-parameter logistic function coefficients.

HTSplotter applies the BI method to calculate synergism/antagonism scores for pairwise or multiple combinations. Before computing the BI score for each time point, HTSplotter verified for each condition if the cell viability effect was higher than the control. If so, the effect was corrected to the maximum of the control, to avoid false BI scores. The predicted combined effect and the BI score were calculated following the Eq (1) and (2) [23]. For the comparison study, SynergyFinder 2.0 was applied following their guidelines [30].

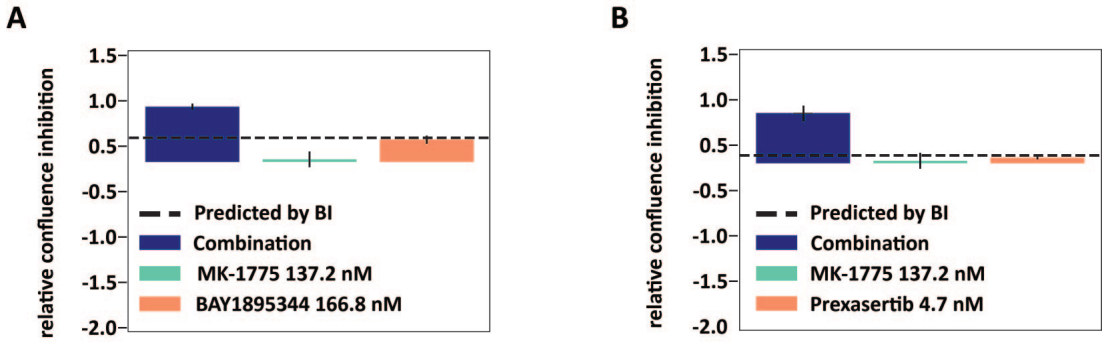
Case study

For the case study, human breast cancer cell line MCF-7 was grown in RPMI medium supplemented with 10% FCS and 2mM L-Glutamine and 100 IU/mL penicillin/streptomycin. To evaluate synergism and the dose-response relationship, MCF-7 cells were seeded in a 384-well plate (Corning COS3764), at a density of 3x103 cells per well. Cells were allowed to adhere overnight, after which these were exposed to the respective treatment.. The drugs used were: AZD2281 (Selleckchem, S1060), MK-1775 (Selleckchem, S1525), prexasertib (MedChem Express, HY-18174A) and BAY 1895344 (Selleckchem, S8666). The treatment was performed by the D300 TECAN. Cell proliferation was monitored for 72h, in which pictures were taken through IncuCyte S3 Live Cell Imaging System each 2h.Each image was analysed through the IncuCyte S3 Software. Cell proliferation was monitored by analysing the occupied area (% confluence) over time.

Supporting information



S1 Fig. BI score heatmap of effect combination overtime and for 72h. The blue indicates synergism and the red indicates antagonism. A: Effect combination heatmap overtime of MK-1775 (55.8 nM) with a concentration range of prexasertib (1 x 7 matrix). B: 2D heatmap for 72h of effect combination, MK-1775 with a concentration range of prexasertib. C: Effect combination heatmap overtime of MK-1775 (55.8 nM) combined with a concentration range of BAY1895344 (1 X 7 matrix). D: 2D heatmap for 72h of effect combination, MK-1775 with a concentration range of BAY1895344.



S2 Fig. Bar plot, at 72h, of the best dose-effect combination of MK-1775 with BAY1895344 (A) or MK-1775 with prexasertib (B). The dash line indicates the predicted combination effect, computed according to the BI method, Eq (1). A: Relative confluence inhibition of BAY1895344 at 166.8 nM combined with MK-1775 at 137.2 nM, with a BI score of 0.42. B: Relative confluence inhibition of prexasertib at 4.7 nM combined with MK-1775 at 137.2 nM, with BI a score of 0.56.

Acknowledgments

The authors acknowledge the following individuals for providing data from their projects: Eric de Bony de Lavergne, Jolien Van Laere, Siebe Loontiens and Suzanne Vanhauwaert.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author contributions

Conceptualization - Carolina Nunes, Frank Speleman, Jo Vandesompele, Jolien De Wyn, and Vanessa Vermeirssen

Formal Analysis – Carolina Nunes

Investigation - Carolina Nunes and Fanny De Vloed

Methodology - Carolina Nunes

Software - Carolina Nunes and Jasper Anckaert

Supervision – Frank Speleman and Vanessa Vermeirssen

Visualisation - Carolina Nunes and Vanessa Vermeirssen

Writing- original draft - Carolina Nunes and Vanessa Vermeirssen

Writing- review & editing - Carolina Nunes, Frank Speleman, Jo Vandesompele, Kaat Durinck and Vanessa Vermeirssen

References

[1] Luca Falzone, Salvatore Salomone, and Massimo Libra. "Evolution of cancer pharmacological treatments at the turn of the third millennium". In: *Frontiers in Pharmacology* 9.NOV (2018). ISSN: 16639812. DOI: 10.3389/fphar.2018.01300.

[2] Bryan T. Mott et al. "High-throughput matrix screening identifies synergistic and antagonistic antimalarial drug combinations". In: *Scientific Reports* 5.August (2015), pp. 1–14. ISSN: 20452322. DOI: 10.1038/srep13891.

[3] Minji Jeon et al. "In silico drug combination discovery for personalized cancer therapy". In: *BMC Systems Biology* 12.Suppl 2 (2018). ISSN: 17520509. DOI: 10.1186/s12918-018-0546-1.

[4] Hussin A. Rothan and Teow Chong Teoh. "Cell-Based High-Throughput Screening Protocol for Discovering Antiviral Inhibitors Against SARS-COV-2 Main Protease (3CLpro)". In: *Molecular Biotechnology* 63.3 (2021), pp. 240–248. ISSN: 15590305. DOI: 10.1007/s12033-021-00299-7.

[5] Christian T. Meyer et al. "Charting the Fragmented Landscape of Drug Synergy". In: *Trends in Pharmacological Sciences* 41.4 (2020), pp. 266–280. ISSN: 18733735. DOI: 10.1016/j.tips.2020.01.011.

[6] James Inglese et al. "High-throughput screening assays for the identification of chemical probes". In: *Nature Chemical Biology* 3.8 (2007), pp. 466–479. ISSN: 15524469. DOI: 10.1038/nchembio.2007.17.

[7] Ophir Shalem, Neville E. Sanjana, and Feng Zhang. "High-throughput functional genomics using CRISPR-Cas9". In: *Nature Reviews Genetics* 16.5 (2015), pp. 299–311. ISSN: 14710064. DOI: 10.1038/nrg3899.

[8] Xiaochen Sun, Santiago Vilar, and Nicholas P. Tatonetti. "High-throughput methods for combinatorial drug discovery". In: *Science Translational Medicine* 5.205 (2013). ISSN: 19466234. DOI: 10.1126/scitranslmed.3006667.

[9] Steven P. Williams et al. "Data Descriptor: High-throughput RNAi screen for essential genes and drug synergistic combinations in colorectal cancer". In: *Scientific Data* 4 (2017), pp. 1–13. ISSN: 20524463. DOI: 10.1038/sdata.2017.139.

[10] Joseph Lehár et al. "Combination chemical genetics". In: *Nature Chemical Biology* 4.11 (2008), pp. 674–681. ISSN: 15524469. DOI: 10.1038/nchembio.120.

[11] Andrew Single et al. "A comparison of real-time and endpoint cell viability assays for improved synthetic lethal drug validation". In: *Journal of Biomolecular Screening* 20.10 (2015), pp. 1286–1293. ISSN: 1552454X. DOI: 10.1177/1087057115605765.

[12] Aleksandr Ianevski et al. "SynergyFinder: a web application for analyzing drug combination dose–response matrix data". In: *Bioinformatics* 33.15 (Aug. 2017). Ed. by Oliver Stegle, pp. 2413–2415. ISSN: 1367-4803. DOI: 10.1093/bioinformatics/btx162.

[13] Giovanni Y. Di Veroli et al. "Combenefit: An interactive platform for the analysis and visualization of drug combinations". In: *Bioinformatics* 32.18 (2016), pp. 2866–2868. ISSN: 14602059. DOI: 10.1093/bioinformatics/btw230.

[14] Christian Ritz et al. "Dose-response analysis using R". In: *PLoS ONE* 10.12 (2015), pp. 1–13. ISSN: 19326203. DOI: 10.1371/journal.pone.0146021.

[15] Nicholas A. Clark et al. "GRcalculator: an online tool for calculating and mining dose-response data". In: *BMC cancer* 17.1 (2017), p. 698. ISSN: 14712407. DOI: 10.1186/s12885-017-3689-3.

[16] Bhagwan Yadav et al. "Quantitative scoring of differential drug sensitivity for individually optimized anticancer therapies". In: *Scientific Reports* 4 (2014), pp. 1–10. ISSN: 20452322. DOI: 10.1038/srep05193.

[17] Bhagwan Yadav et al. "Searching for drug synergy in complex dose–response landscapes using an interaction potency model". In: *Computational and Structural Biotechnology Journal* 15 (2017), p. 387. ISSN: 20010370. DOI: 10.1016/j.csbj.2017.07.003.

[18] Julie Foucquier and Mickael Guedj. "Analysis of drug combinations: current methodological landscape". In: *Pharmacology Research Perspectives* 3.3 (2015), e00149. ISSN: 20521707. DOI: 10.1002/prp2.149.

[19] Mark Sinzger et al. "Comparison of null models for combination drug therapy reveals Hand model as biochemically most plausible". In: *Scientific Reports* 9.1 (2019), pp. 1–15. ISSN: 20452322. DOI: 10.1038/s41598-019-38907-x.

[20] Wei Zhao et al. "A new bliss independence model to analyze drug combination data". In: *Journal of Biomolecular Screening* 19.5 (2014), pp. 817–821. ISSN: 1552454X. DOI: 10.1177/1087057114521867.

[21] Swapnil Potdar et al. "Breeze: An integrated quality control and data analysis application for high-throughput drug screening". In: *Bioinformatics* 36.11 (2020), pp. 3602–3604. ISSN: 14602059. DOI: 10.1093/bioinformatics/btaa138.

[22] Oliver Pelz, Moritz Gilsdorf, and Michael Boutros. "Web cellHTS2: A web-application for the analysis of high-throughput screening data". In: *BMC Bioinformatics* 11 (2010). ISSN: 14712105. DOI: 10.1186/1471-2105-11-185.

[23] Kyle R. Roell, David M. Reif, and Alison A. Motsinger-Reif. "An introduction to terminology and methodology of chemical synergy-perspectives from across disciplines". In: *Frontiers in Pharmacology* 8.APR (2017), pp. 1–11. ISSN: 16639812. DOI: 10.3389/fphar.2017.00158.

[24] Cody W. Lewis et al. "Prolonged mitotic arrest induced by Wee1 inhibition sensitizes breast cancer cells to paclitaxel". In: *Oncotarget* 8.43 (2017), pp. 73705–73722. ISSN: 19492553. DOI: 10.18632/oncotarget.17848.

[25] Amirali B. Bukhari et al. "Inhibiting Wee1 and ATR kinases produces tumor-selective synthetic lethality and suppresses metastasis". In: *Journal of Clinical Investigation* 129.3 (2019), pp. 1329–1344. ISSN: 15588238. DOI: 10.1172/JCI122622.

[26] Laura Carrassa et al. "Combined inhibition of Chk1 and Wee1: In vitro synergistic effect translates to tumor growth inhibition in vivo". In: *Cell Cycle* 11.13 (2012), pp. 2507–2517. ISSN: 15514005. DOI: 10.4161/cc.20899.

[27] Nora Rieber et al. "RNAither, an automated pipeline for the statistical analysis of high-throughput RNAi screens". In: *Bioinformatics* 25.5 (2009), pp. 678–679. ISSN: 13674803. DOI: 10.1093/bioinformatics/btp014.

[28] Philip J. Chapman et al. "IncucyteDRC: An R package for the dose response analysis of live cell imaging data [version 1; referees: 2 approved]". In: *F1000Research* 5.May 2016 (2016), pp. 1–9. ISSN: 1759796X. DOI: 10.12688/F1000RESEARCH.8694.1.

[29] Markus List et al. "Comprehensive analysis of high-throughput screens with HiTSeekR". In: *Nucleic Acids Research* 44.14 (2016), pp. 6639–6648. ISSN: 13624962. DOI: 10.1093/nar/gkw554.

[30] Aleksandr Ianevski, Anil K. Giri, and Tero Aittokallio. "SynergyFinder 2.0: Visual analytics of multi-drug combination synergies". In: *Nucleic Acids Research* 48.1 (2021), W488–W493. ISSN: 13624962. DOI: 10.1093/NAR/GKAA216.