# Fine tuning a logical model of cancer cells to predict drug synergies: combining manual curation and automated parameterization

Åsmund Flobak[1,2]*†, John Zobolas[3]†, Miguel Vazquez[4],
Tonje S. Steigedal[1], Liv Thommesen[5], Asle Grislingås[5],
Barbara Niederdorfer[1], Evelina Folkesson[1] and Martin Kuiper[3]

[1]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology,
Trondheim, Norway, [2]The Cancer Clinic, St. Olav's Hospital, Trondheim, Norway, [3]Department of Biology,
Norwegian University of Science and Technology, Trondheim, Norway, [4]Barcelona Supercomputing
Center, Barcelona, Spain, [5]Department of Biomedical Laboratory Science, Norwegian University of
Science and Technology, Trondheim, Norway

Treatment with combinations of drugs carries great promise for personalized
therapy for a variety of diseases. We have previously shown that synergistic
combinations of cancer signaling inhibitors can be identified based on a logical
framework, by manual model definition. We now demonstrate how automated
adjustments of model topology and logic equations both can greatly reduce the
workload traditionally associated with logical model optimization. Our
methodology allows the exploration of larger model ensembles that all obey a
set of observations, while being less restrained for parts of the model where
parameterization is not guided by biological data. We benchmark the synergy
prediction performance of our logical models in a dataset of 153 targeted drug
combinations. We show that well-performing manual models faithfully represent
measured biomarker data and that their performance can be outmatched by
automated parameterization using a genetic algorithm. Whereas the predictive
performance of a curated model is strongly affected by simulated curation errors,
data-guided deletion of a small subset of regulatory model edges can significantly
improve prediction quality. With correct topology we find evidence of some
tolerance to simulated errors in the biomarker calibration data, yet
performance decreases with reduced data quality. Moreover, we show that
predictive logical models are valuable for proposing mechanisms underpinning
observed synergies. With our framework we predict the synergy of joint inhibition
of PI3K and TAK1, and further substantiate this prediction with observations in
cancer cell cultures and in xenograft experiments.

KEYWORDS

logical modeling, cancer, synergy, genetic algorithm & optimization technique, xenograft
analysis, cell line, cancer signal processing

# Introduction

Combining several specific and targeted drugs in one therapy to fight disease increases chances of treatment success (Hanahan and Weinberg, 2011). Drug combinations that together act in synergy are especially attractive because they allow for pushing treatment effects beyond those obtainable by each drug alone (Al-Lazikani et al., 2012), with drug dosages that can be well below levels where individual drugs begin to cause adverse effects. In addition, synergistic drug combinations may have reduced side-effects by improved selectivity in a specific biological context, for instance by tailoring therapy to individual patients (Eduati et al., 2017; Eduati et al., 2020), and by allowing targeting of only certain cell types in an organism (Lehár et al., 2009). Lastly, searching for new combination therapies has the additional benefit that already approved drugs can act beneficially in novel combinations, and thus even allow bypassing initial drug development phases.

While the development of rational drug combination treatment has become a major priority due to hopes of increased treatment potency, a grand challenge remains in dealing with their identification in the vast space of potential drug combinations. Currently, more than two hundred drugs have been approved by the FDA to treat cancer (https://www.cancer.gov/about-cancer/treatment/drugs). The testing of drugs in all combinations with other drugs in a panel needs assays in numbers that increase exponentially with increasing drug numbers. Even a modestly sized drug panel of 150 drugs corresponds to over 10,000 pairwise drug combinations. Testing high numbers of drug combinations in high throughput screens on cell lines or other patient-specific model systems is costly and at some point prohibitively expensive and cumbersome. Therefore, help is needed from *in silico* drug effect simulations to produce high quality predictions that can guide drug combination screens or therapy choices for testing in cell lines or patients. *In silico* simulations may help identify those combinations that are unlikely to produce synergies, which can be of significant help to reduce the large experimental search space that otherwise would need to be covered in exhaustive screens. As many drug synergies can be seen as emergent properties arising from molecular causality networks, analytical frameworks from computational systems biology seem to be well suited to the task.

Several mathematical frameworks have already been tested to mechanistically model drug combination effects, including continuous, discrete, and hybrid modeling approaches. Published approaches generally depend on molecular causalities downloaded from prior knowledge databases, extracted from large-scale data, or obtained by a combination of the two. Based on a dataset capturing proteomic responses of 14 targeted drugs, Miller et al. used ordinary differential equations to study mechanisms of synergy between inhibitors of CDK4 and IGF1R, revealing that the mechanisms rely on the activity of AKT (Miller et al., 2013). Nelander et al. explored ODE models derived from observations on phospho-proteins and cell cycle markers following 21 pairwise applications of targeted drugs, with the aim to use best-performing pairs for design of new combination therapies (Nelander et al., 2008). In a semi-qualitative modeling approach, Klinger et al. used a perturbation dataset for MAPK, PI3K and NF-κB signaling to inform a model showing that combined inactivation of MEK and

EGFR could inactivate endpoints of RAS, ERK and AKT signaling (Klinger et al., 2013). Jin et al. explored enhanced Petri nets to describe molecular processes for the synergy of an EGFR inhibitor (gefitinib) with chemotherapy (docetaxel) and identified KRT8 as a candidate gene to explain the synergy (Jin et al., 2011). Saez-Rodriguez et al. have explored logic models configured by perturbation data to identify novel drug synergy mechanisms and combinations (Saez-Rodriguez et al., 2009; Tognetti et al., 2021). However, all of these approaches rely on extensive and costly combinatorial drug perturbation data, be it transcriptomic, proteomic, viability etc., for describing mechanisms of synergy, and therefore they require large investments in data production and do not provide a scalable solution for testing of the large drug combination space.

In order to reduce dependence on *a priori* perturbation experiments, attempts have been made to predict drug synergies from data obtained in a marginal experimental search space, rather than the full combinatorial space. Fröhlich et al. used ODE models informed by transcriptomic and viability data to predict drug combination responses, finding that highly accurate predictions could be produced for those drugs for which they had viability response data (Fröhlich et al., 2018). In the DREAM7 - NCI-DREAM, Drug Sensitivity and Drug Synergy Challenges (Bansal et al., 2014; Goswami et al., 2015; Sun et al., 2015) (NCI-DREAM), pairwise drug responses were predicted from response data obtained for each drug alone. The best performing teams in the NCI-DREAM challenge obtained a probabilistic concordance (PC) index of 0.61, on a scale ranging from 0.9 (perfect prediction) to 0.1 (perfect opposite prediction). Although this is better than random (PC index of 0.5), it clearly illustrates that obtaining accurate synergy predictions is far from trivial, due to a variety of reasons that will be discussed in this paper. In the more recent AstraZeneca-Sanger Drug Combination Prediction DREAM Challenge (Menden et al., 2019) (AZ-DREAM), one of the aims was to develop and demonstrate drug combination response predictability independent of extensive perturbation data. The development of such powerful prediction approaches has the advantage of being relevant not only to preclinical drug screens, but also to bed-side applications. Drug perturbation data clearly will not be trivial to obtain for individual patients, unless patient-derived experimental assays that mimic patient responses can be developed (e.g., xenografts, explants etc.). Sobering results from the AZ-DREAM challenge showed that most teams had balanced accuracies of 0.5–0.6, with the best performing team obtaining a balanced accuracy of only 0.69.

With the availability of training data, machine learning algorithms have also been explored to predict drug synergies (Gayvert et al., 2017; Preuer et al., 2017; Tang et al., 2019). However, major limitations of such approaches include the lack of mechanistic insight (Yu et al., 2018), and dependence on high quantities of training data. Despite some increase in their availability, such large scale datasets are still largely missing, in part due to great experimental complexity and high economic cost. Efficient *in silico* therapy based on patient-specific models should ultimately be integrated into the clinical decision process. In this study, we investigate the performance of logical modeling in predicting drug responses of cancer cell lines. In order to reduce the dependency on large training datasets, we explore the use of cell lines measurements of biomolecules obtained at a single time point:

at steady state proliferation. To reduce data dependency and to improve mechanistic insights, these measurements are combined with prior knowledge to construct logical model ensembles to simulate drug combination effects. Logical model building is known to require meticulous involvement of curators and bioinformaticians, with substantial commitment to manual tinkering of models before the behavior of one model matches that of its experimental counterpart. We have previously published logical models of cancer cell lines, named CASCADE 1.0, CASCADE 2.0 and CASCADE 3.0, which demonstrate the potential of logical modeling for the prediction of drug synergies (Flobak et al., 2015; Niederdorfer et al., 2020; Tsirvouli et al., 2020). However, the curation effort required to assemble a cancer signaling network and the dedicated interactive efforts needed to optimally modify logical rule definitions becomes a clear obstacle when constructing larger models.

If patient-specific logical models are to be used routinely, it is preferable that such logical models can be automatically constructed for any cell line or patient-derived cell culture, and for any repertoire of targeted drugs (Dorier et al., 2016; Gjerga et al., 2020). We therefore set out to automate processes required to calibrate a logical model from a set of molecular causative statements, i.e., a prior knowledge network (Dorier et al., 2016; Gjerga et al., 2020). A software pipeline developed to that end would have to 1) assemble a network topology from structured data obtained from prior knowledge databases, 2) interpret baseline cancer cell line biomarker data into a signaling entity activity score, 3) calibrate generic logical models, created from prior knowledge data, by modifying logic equations to match the observed activity scores, and 4) predict phenotypic consequences of combinatorial interventions to the simulated model behavior. Our software solution for realizing points 3 and 4 is available at https://github.com/druglogics. We use a genetic algorithm to automatically parameterize a set of logic equations representing cancer growth-promoting signaling in the AGS gastric adenocarcinoma cell line. We demonstrate our approach by reproducing results from a previous manual effort and next test its utility with a larger model that was benchmarked against a dataset from a drug effect screen of 153 drug combinations. Experiments that simulate different levels of curation quality and biomarker data quality indicate the need for a reliable PKN, while still allowing for model improvement by network link pruning and parameter optimization.

# Methods

## Boolean modeling and attractor computation

Boolean models rely on the formalism initially proposed by Stuart Kauffman (Kauffman, 1969) and René Thomas (Thomas, 1973), and are based on the Boolean operators AND, OR and NOT for computing the activity states of variables (nodes). Their approach first defines a regulatory graph consisting of nodes representing signaling entities (model components), and signed and directed edges representing regulatory interactions that connect signaling entities. The activities of all model components

are then associated with the Boolean values "True" and "False," represented by 1 and 0, corresponding to activity and inactivity, respectively. This dichotomy of activity levels can be interpreted as activity being above or below a "threshold": a component is "active" when its activity level is sufficiently high to influence a target component's activity level. In model simulations, specific model components can be defined as "output nodes," whose activity values serve as a proxy for a phenotype of interest. This allows us to compute an overall "growth" value in our simulations, by integrating all activity values of model output nodes, and scaling this sum to [0,1]. For example, if the anti-survival model output nodes CASP8, CASP9 and FOXO3 are inactive and the pro-survival model output nodes MYC, CCND1 and TCF7 are active, the global output "growth" is 1. Inhibitory effects of a drug are simulated by fixing the drug target to an activity level of 0, i.e., simulating a block in signaling activity of the drug target node.

The ability of a logical model to represent biological reality and, in consequence, reliably predict the result of perturbations by drugs that specifically target model nodes, requires an optimization of both the logical rules (parametrization) and the regulatory topology. The effect of both optimizations on drug simulations needs an efficient process to identify model attractors. Model attractors represent asymptotic behavior of the system, from which it cannot escape, either as fixed points, or as complex attractors involving cycles of iterative activated/inactivated nodes. We identified fixed points using the bioLQM software, which, among others, provides a fast algorithm for finding such stable states and other complex attractors (Naldi, 2018). Due to the large number of simulations required for the "Randomizing regulatory edges of the curated model reduces predictiveness" Results section, we used a modified version of the algorithm BNReduction (Veliz-Cuba et al., 2014), which allows the identification of single stable state phenotypes that are most prevalent with our self-contained CASCADE topologies.

## Model calibration by parameterization optimization

A genetic algorithm is applied to automate model parameterization, as follows:

The input for model calibration consists of:

- Interactions: signed and directed binary interactions (SIF format (Shannon et al., 2003)).
- Steady state: Boolean vector containing states of nodes in interactions, where nodes that should be active are assigned the value 1, and nodes that should be inactive are assigned the value 0. For nodes whose state cannot be determined a dash (−) can optionally be used to explicitly declare that the node state is undetermined. This steady state vector will be used for evaluating the correctness of a model's stable state.

Acknowledging that there are multiple optimal parameterization solutions possible, the output of an automated model calibration is an ensemble of models each with a stable state optimized to match the input steady state.

To run the parameter configuration, interactions are first assembled to logic equations, based on a default equation (Mendoza and Xenarios, 2006) relating a node with its regulators, for instance:

$$\text{Target} *= (A \text{ or } B \text{ or } C) \text{ and not} (D \text{ or } E \text{ or } F)$$

where activating regulators A, B and C of a target are combined with logic "or" operators, and inhibitory regulators D, E and F are combined with "and not" operators to determine the state of the target node in the next time step. The operator "and not," which directs the integration of activating and inactivating regulators, is referred to as the link-operator (Zobolas et al., 2022). The topology is "self-contained," meaning that any regulator is itself regulated by one or more components from within the network topology, effectively meaning there are no "user-controlled" inputs to the network through e.g., hormone receptors.

Next, a genetic algorithm is used to iteratively refine the parameterization to produce a logical model with a stable state matching the specified input steady state. First, an initial generation of models is formulated, where a large number of mutations to the parameterization is introduced: randomly selected equations are mutated from "and not" to "or not", or *vice versa*. Hence, the two possible equations for the example above would be:

$$\text{Target} *= (A \text{ or } B \text{ or } C) \text{ and not} (D \text{ or } E \text{ or } F)$$

Or

$$\text{Target} *= (A \text{ or } B \text{ or } C) \text{ or not} (D \text{ or } E \text{ or } F)$$

For each model a fitness score is computed: each matching Boolean value between the vector of a stable state and the steady state improves the fitness score, which is scaled to the range [0,1]. A fitness score of 1 is achieved when all states match the given steady states of specific nodes within the model. Models without a stable state, i.e., models that only have complex attractors, are assigned a fitness of zero. The fitness score for models with $n$ stable states, assuming that the steady state vector ($ss$) and the stable states ($fix$) have $m$ nodes in common, is as follows:

$$fitness = \sum_{i=1}^{n} \frac{\left( \sum_{j=1}^{m} I\left[ s_{fix}^i(j) = s_{ss}(j) \right] \right) / m}{n}$$

where $s(j) \in \{0, 1\}$ denotes the state of the $j$ node, $n$ is the number of fixed points (stable states), $i$ is the fixed point index and $I: X \rightarrow \{0, 1\}$ the indicator function. Thus, models with multiple stable states are penalized.

The stable state vector used to compute fitness in the reported experiments was based on manual annotation of signaling activities in AGS cancer cells (Flobak et al., 2015).

For each generation, a user-defined number of models were selected for populating the next-generation of models. For our simulations three models were selected, specifically the ones that achieved the highest fitness scores in each generation, to populate a next-generation of 20 models. First, crossover was performed, where each selected model would exchange logic equations with other selected models (including itself, thus also enabling asexual reproduction). Then a number of mutations were introduced as described above. For our simulations 3 mutations were introduced. Before a stable state was obtained, the number of mutations introduced per generation was increased by a user-defined factor. The large number of mutations in the initial phase ensured that a large variation in parameterization could be explored. For our simulations we chose this multiplication factor to be 1000 (i.e., introducing a total of 3000 mutations), effectively ensuring that the initial generations were randomly sampled from all possible

model configurations (a total of $2^n$ models, where $n$ is the number of equations with a link-operator). The genetic algorithm's evolutionary process kept the best 3 models with the higher fitness scores. Evolution was halted upon reaching a user-specified threshold fitness of 0.99, which was compared with the minimum fitness among the top 3 models. In case this fitness could not be reached, evolution was halted when a user-defined maximum number of generations had been spanned. For our simulations we allowed for a maximum of 20 generations. The model evolution process was repeated 50 times (user-defined value) resulting in a total of $3 \times 50 = 150$ models at the end of the genetic algorithm optimization.

## Model calibration by topology optimization

In order to introduce variations to the topology, the genetic algorithm modified a whitelist and a blacklist of regulators of the prior knowledge network, while always preserving at least one regulator for each target, so as to not break the self-contained property of the network. Based on the same formula as shown above,

$$\text{Target} *= (A \text{ or } B \text{ or } C) \text{ and not} (D \text{ or } E \text{ or } F)$$

this means that the genetic algorithm takes out some subset of the regulators A, B, C, D, E or F (blacklisting). After a regulator has been eliminated, the genetic algorithm is also allowed to bring back regulators originally defined in the PKN (whitelisting). For our simulations, we introduced 50 such topology mutations during the initialization phase and when models with stable states were found, we reduced this number to 10, so as to not severely reduce the number of PKN edges.

## Model simulation and synergy prediction

After repeating evolution a specified number of times, model ensembles were analyzed in a third step of the software pipeline, as follows:

The input to model simulation and synergy prediction consists of:

- An ensemble of logical models
- A drug panel: List of drugs and their target node(s) in the model
- Perturbations: the single and drug pair combinations to be analyzed.
- Model output nodes with weighted score to evaluate global output (i.e., 'growth')

The output from model simulation and synergy prediction:

- Drug synergy predictions from ensembles of models

For each model, all specified perturbations were simulated. For each perturbation, the drug panel was consulted to fix the state of the associated node(s) to the value 0. A node state could in principle also be fixed to the value 1 for a drug that *activates* a signaling entity, but this feature was not used here as all drugs *inhibit* nodes in the model, thereby representing inhibition of their target in a cell. After simulating a perturbation, the global output parameter "growth" was computed by integrating the weighted score derived from the states of model output nodes. The output nodes RSK_f, CCND1 and MYC contributed to cell proliferation with a positive weight of

1 while CASP8, CASP9 and FOXO3 contributed to cell death with a negative weight of −1. For example, if two output nodes A (weight 1) and B (weight −1) were found to have the states A = 1, B = 1 for a perturbation, the global output would result in $A_{state} \times A_{weight} + B_{state} \times B_{weight} = 1 \times 1 + 1 \times (-1) = 0$. This value was then scaled from 0 to 1 based on the theoretical minimum and theoretical maximum "growth," for this example, the range $[-1,1]$, the global output would be 0.5 (which could be interpreted to represent a cytostatic cell state). The scaled global output ("growth") was then used to compute synergies (see below).

All steps of the software pipeline were implemented in the OpenJDK Java v1.8 language and run on Linux 4.15.0–122-generic/Ubuntu 18.04.4 LTS. The different pipeline modules can be accessed at https://github.com/druglogics. The main package used for the simulations was druglogics-synergy v1.2.1 (https://github.com/druglogics/druglogics-synergy/). It takes ~1 h to calculate the drug synergy predictions using the default configuration options and the CASCADE 2.0 topology. Changing the attractor calculation to the modified BNReduction script can result up to 90% less computation time. Software documentation (including installation and various configuration documentation) is available online at https://druglogics.github.io/druglogics-doc/. A detailed tutorial was made as an introduction guide to the software pipeline (https://druglogics.github.io/synergy-tutorial/). For an extensive documentation of the methods used in this work, see https://github.com/druglogics/ags-paper.

## In silico definition of synergy

Synergy is defined as an additional effect beyond what is expected from a reference model of drug combination responses. Both for *in silico* simulations and *in vitro* experiments an observed combination effect can be formally defined as the effect E observed for two drugs *a* and *b,* where $E(a,b)$ is the observed effect in a combination experiment, $A(a,b)$ is the drug combination effect expected from each individual drug's properties as based on a reference model for combination responses, and $S(a,b)$ denotes any difference between the observed and the expected drug combination effect, such that $E(a,b) = A(a,b) + S(a,b)$ (Li et al., 2018). In the case of excess effects observed for a combination, $S(a,b)$ is positive and synergy is called, and conversely for attenuated effects, $S(a,b)$ is negative and antagonism is called. Finally, for drug combinations where $E(a,b)$ equals $A(a,b)$, the drug combination effect can fully be anticipated by each drug response independently, and neither synergy nor antagonism is called.

In model simulations the expected drug combination response is defined as the product of the two global output "growth" values for each single drug, similarly to the Bliss independence (Bliss, 1939) synergy metric used in lab experiments: when a combinatorial perturbation in simulations is found to predict a lower growth than expected, i.e., $growth(a,b) < growth(a) * growth(b)$, the combinatorial perturbation response is declared synergistic.

## Gold standard synergies

Our previously published dataset of targeted drug combinations (Flobak et al., 2019) was used to benchmark the algorithms. The

drugs included comprised the inhibitors 5Z-7-oxozeaenol (5Z), AKTi-1,2 (AK), BIRB0796 (BI), CT99021 (CT), PD0325901 (PD), PI103 (PI), PKF118-310 (PK), JNK Inhibitor XVI (JN), BI-D1870 (D1), BI605906 (BIX02514) (60), Ruxolitinib (INCB18424) (SB), SB-505124 (RU), D4476 (D4), KU-55933 (KU), 10058-F4 (F4), Stattic (ST), GSK2334470 (G2), GSK-429286 (G4), P 505-15 (P5). For the drug synergy calling in the 153 combinations drug screen, three curators were asked to evaluate growth curves and decide on which showed interesting combination effects that could have warranted further investigations. A consensus list was then used to identify a threshold for drug synergy assessment using the software Cimbinator (Flobak et al., 2017) and configured to compute synergies per the Bliss metric. The analysis identified six drug synergies (AK-BI, PI-D1, BI-D1, PI-G2, PD-PI, 5Z-PI). Note that two drug synergies in the drug screen performed in 2015 were not captured by this analysis, probably relating to the different readouts used in the drug screen performed in 2019 (xCELLigence and CellTiter Glo, respectively).
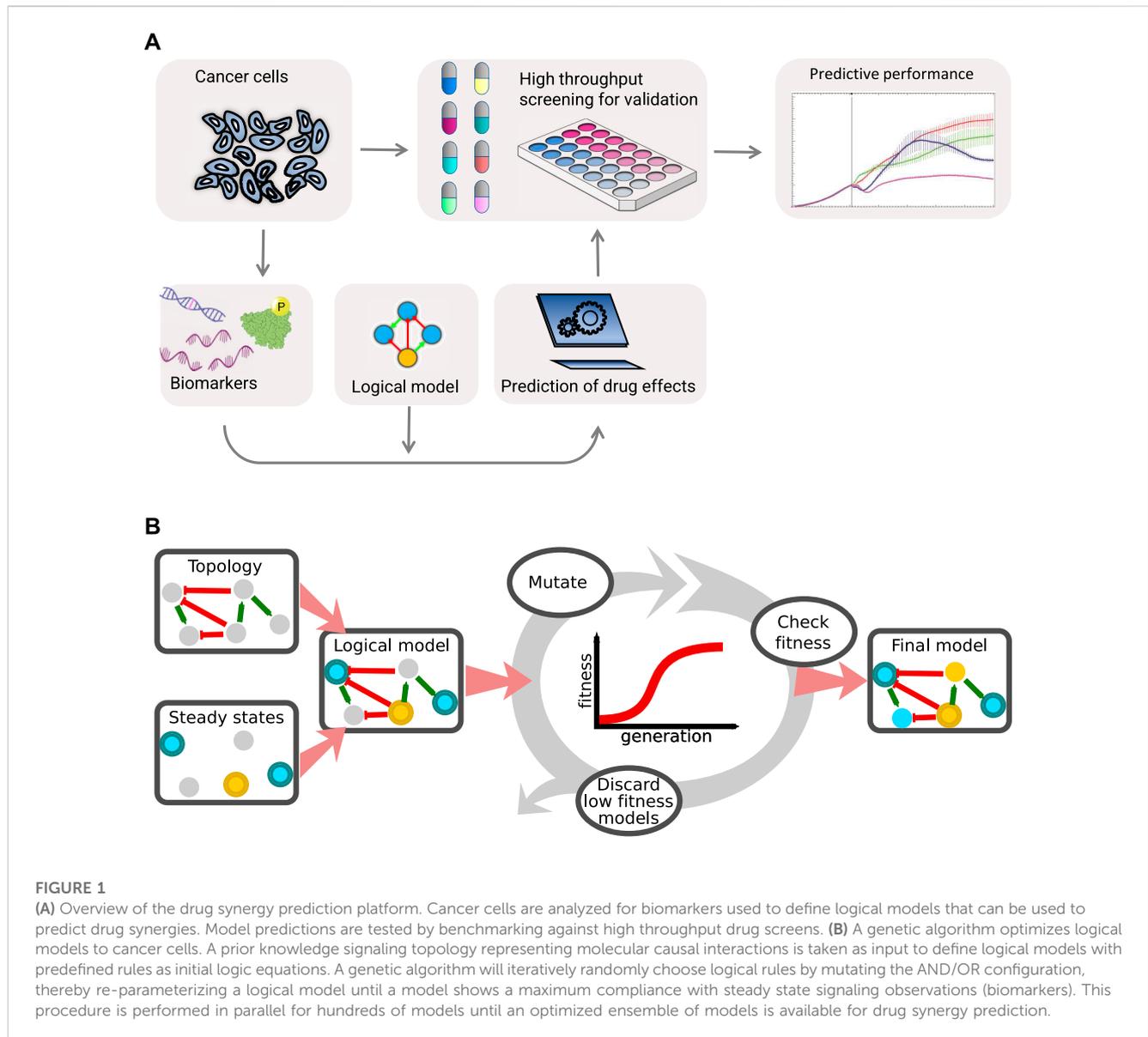
## Normalization

Normalization of synergy predictions for drugs a and b was performed by computing the exponential fold change for the ratio of output from models *calibrated* to steady state biomarker data (x) to output from models calibrated to a *random* yet proliferative phenotype (y):

$$\text{synergy} = \exp \big( \big( \text{growth}_x(a,b) - \text{growth}_x(a)^*\text{growth}_x(b) \big) - \big( \text{growth}_y(a,b) - \text{growth}_y(a)^*\text{growth}_y(b) \big) \big)$$

Our random proliferative phenotype corresponds to a cell with all anti-survival signals inactivated, and at least one pro-survival signal active. All references in the main text to "Calibrated" model performance refer to normalized synergy predictions as explained above.

## Mouse xenograft experiments

40 female Balb/c mice 4–5 weeks old (Taconic) were inoculated with two million AGS cells subcutaneously in the right dorsal flank. Cells were mixed with Matrigel to improve probability of successfully establishing a xenograft model: in a small pilot (n = 3) we observed that none of three mice injected with cells in medium (DMEM) developed tumors, while two of three mice injected with cells in medium and Matrigel developed tumors. 100 μL of cell suspension in HAM'S F12 medium (Invitrogen, Carlsbad, CA) with 10% fetal calf serum (FCS; Euroclone, Devon, UK), and 10 U/mL penicillin-streptomycin (Invitrogen) was mixed with 100 μL of ECM Gel from Engelbreth-Holm-Swarm murine sarcoma (Sigma-Aldrich). After 4 weeks, minuscule but palpable tumors had formed in 30 mice, which were randomized to four groups and subjected to treatment: 1) 5Z-7-oxozeaenol (3 mg/kg/d), 2) PI103 (5 mg/kg/d), 3) 5Z-7-oxozeaenol (3 mg/kg/d) + PI103 (5 mg/kg/d), 4) vehicle. Randomization ensured that average tumor volume was similar in the four groups. Weights of mice ranged from 14.9 to 20.0 g at onset of treatment, with average weight 17.66 g and standard deviation of 1.06 g. All mice received the same dose of

**FIGURE 1**
**(A)** Overview of the drug synergy prediction platform. Cancer cells are analyzed for biomarkers used to define logical models that can be used to predict drug synergies. Model predictions are tested by benchmarking against high throughput drug screens. **(B)** A genetic algorithm optimizes logical models to cancer cells. A prior knowledge signaling topology representing molecular causal interactions is taken as input to define logical models with predefined rules as initial logic equations. A genetic algorithm will iteratively randomly choose logical rules by mutating the AND/OR configuration, thereby re-parameterizing a logical model until a model shows a maximum compliance with steady state signaling observations (biomarkers). This procedure is performed in parallel for hundreds of models until an optimized ensemble of models is available for drug synergy prediction.

drugs, and the dose was adjusted for a body weight one standard deviation below average, i.e. 16.6 g. Drugs were diluted in medium with 40% DMSO for a total injection volume of 250 μL and injected intraperitoneally three times per week for a total of seven injections. Maximum (a) and minimum (b) tumor diameters were measured twice weekly with a caliper, and the volume V of the tumor was estimated from the formula $V = 0.5 \, a \times b^2$.
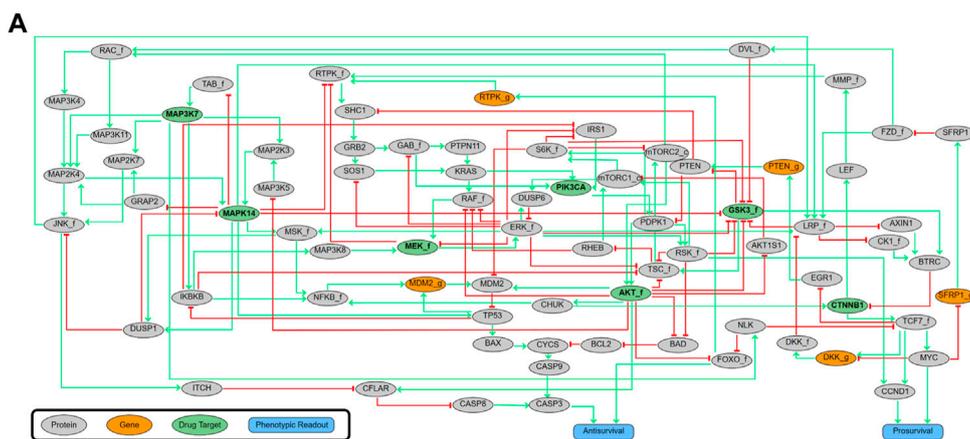
## Results

For reliable simulation of drug responses of cancer cell lines, computational models must adequately represent the regulatory network (topology) underlying cell fate decisions, meaning that high quality molecular causal relationship data must exist and be converted to regulatory graphs. In addition, the activity states of molecular regulatory components must be measured, demanding high quality biomarker data. From the regulatory graph the response

of components to upstream source nodes and influence on downstream target nodes needs to be specified in the form of logical rules and calibrated so as to accurately represent the biological decision mechanisms of these cells in a Boolean framework. Finally, good benchmarking data must exist to evaluate the performance of the model, e.g., for our purposes in the form of drug synergy data, see Figure 1A.

## Design of an automated model parameterization module

Previously we have shown the feasibility of logical model predictions of drug synergies (Flobak et al., 2015) using the cancer cell line AGS, chosen due to known deregulations of several core cell survival signaling pathways. A Prior Knowledge Network (PKN) was curated to represent these signaling pathways, and converted to a set of mathematical rules formulated in Boolean

**FIGURE 2**
**(A)** The CASCADE 1.0 prior knowledge network. 75 signaling nodes with signed and directed regulatory influences annotated (activating interactions in green, inhibiting interactions in red). All signaling components receive input from other signaling components from within the network, and ultimately influence the two phenotypic output nodes Antisurvival and Prosurvival. **(B)** Evolution of fitness of calibrated models. Overall fitness is plotted as a function of generation, with average fitness and standard deviation indicated. The data for this figure was produced by running the genetic algorithm for 1,000 times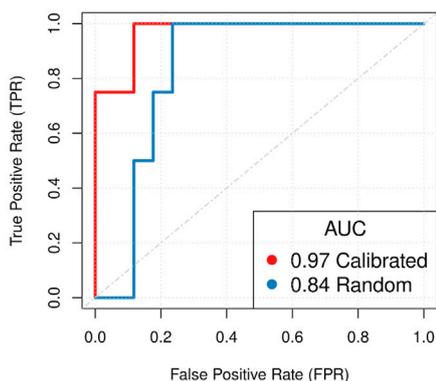 (evolutions), with 20 generations per evolution and 20 models per generation. We observe that the average fitness and standard deviation follow a sigmoidal increase and stabilize after 10−15 generations. The persistence of the standard deviation across generations including those late in the evolution shows that new models still explore variations to the model parametrization while selection keeps the fitness score of the trained models at a constant plateau. **(C)** Predictive performance of ensembles of logical models. ROC curves are in the left panel, and PR curves are in the right panel. Random model predictions were generated by collecting predictions from ensembles of models trained to a random yet proliferative phenotype.

(*Continued*)

**FIGURE 2 (Continued)**

Calibrated predictions were generated by model ensembles, trained to steady state data, and normalized to the random model predictions (see Methods for more details). The genetic algorithm modified the balance of influence between positive and negative regulators of a target node, while topology features (edges, nodes) were not modified. Both ROC and PR curves show very good performance across all model sets for the calibrated models, similar to results from Flobak et al. (2015). Comparison of AUC performance between "Calibrated" and "Random" model predictions using 10,000 bootstrap resamples, revealed no statistically significant difference for ROC ($p$-value = 0.173) but slightly significant difference for the PR performance ($p$-value = 0.014).

logic. This model, available from https://github.com/druglogics/cascade as CASCADE 1.0 (Figure 2A), consists of 75 nodes representing cancer signaling entities and 149 edges representing regulatory interactions, and it could predict five synergies of which four were experimentally confirmed (Flobak et al., 2015). Since our model is based on prior knowledge, amenable to interpretation by molecular biologists, the model also can be used for inspection as to which signaling pathways are important for particular drug response observations, e.g., we have suggested that FOXO signaling was crucial to the drug synergy effect of joint PI3K and MEK inhibition (Flobak et al., 2015). Whereas for many of the logical rules the definition of the logical operators (AND, OR, NOT) was more or less evident from literature and database knowledge, analysis of Boolean model attractors indicated that some logical rules needed further manual optimization in a stepwise manner so that ultimately the model stable state behavior matched the observed pattern of signaling entities at steady state in proliferating AGS cells. We now report on how we automated and generalized these steps required to parameterize an *in silico* model of a cancer cell line, by employing a genetic algorithm for deriving logical rules from prior knowledge and steady state signaling observations, see Figure 1B. From a curated network topology a set of standardized logic equations are obtained by defining one logic equation for each model target node, with model source nodes as operands (Mendoza and Xenarios, 2006). For example, if protein T is activated by proteins A and B, while protein C inhibits protein T, the equation could read as T = (A *OR* B) *AND NOT* C. Subsequently, the parameterization (choices of logic operators) is optimized by a genetic algorithm, specifically modifying the AND NOT/OR NOT parameter. This translates to adjusting the balance of influence from activating and inhibiting regulators; if the operator is OR NOT then either activation or absence of inhibition is required for activation of the target (more permissive), while if the operator is AND NOT then both activation and absence of inhibition is required (less permissive). The genetic algorithm iteratively modifies the parameterization of a small subset of the equations, and selects best performing models for defining a new generation of candidate models. Best performing models are chosen based on their ability to reproduce known baseline cell signaling states, as far as the available cell line data allows it. Evaluating fitness from a match with baseline observations also means that our models are defined independently of perturbation data. Our software solution for automatic parameterization is available at http://github.com/druglogics/.

## Automated logical rule definitions perform on par with manually curated rules

To compare results from our manually constructed logical rules with automated rule definitions, our software translates the graph, as encoded in a SIF file format, to a set of 75 logic equations in a standardized format (Mendoza and Xenarios, 2006). Logic equations are then optimized using a genetic algorithm in a process where the fitness of each model is calculated by comparing the matches of its stable state nodes with observed activities of signaling entities for proliferating cancer cells. For the AGS cells, this process comprised 20 generations, in which each generation received mutations to a small subset of logic equations iteratively, with 20 models per generation tested for fitness. In order to adequately cover the space of local optima, the evolutionary process was repeated 50 times and the three best performing models from each evolution were retained, which resulted in an ensemble of 150 models. A theoretical maximal fitness of one would be reached if all nodes have a state matching the observed activity state of the corresponding protein. As can be seen in Figure 2B, the population average fitness of each generation increases exponentially before plateauing at a fitness close to the theoretical maximal fitness, per Holland's Schema Theorem (Holland, 1992), indicating that the theoretical models can be parameterized so as to be compliant with experimentally observed signaling states. While a genetic algorithm cannot guarantee a global optimum, our results clearly indicate that we quickly achieve convergence to a local optimum.

Whereas these model ensembles provide the testing ground for the *in silico* drug effect simulations, it is to be expected that certain motifs of the network topology itself will create "blind spots" causing some synergies to escape prediction. For example, if two directly sequential signaling nodes are targeted by two different drugs, while no other influences from other signaling entities are allowed by the topology, then these two drugs cannot be predicted to act synergistically in our logical modeling framework. To remedy such situations, extensions to the prior knowledge network are necessary, or conversion to non-discrete modeling. On the other hand, if two drug targets are active and are the only (activating) source nodes of a joint downstream target node, with their joint effect on the target governed by an *OR* logic operator, this may constitute a synergy that is highly likely to stand out in an analysis, since the OR operator would cause either drug alone to not affect a joint downstream node. Between these two extremes, the topology will have varying ability to produce a particular synergy prediction for a given combination perturbation. In order to correct for topology-intrinsic propensities for predicting some synergies we next employed a normalization strategy where synergy predictions for an automatically parameterized model ensemble are normalized to a *randomly* parameterized model, meaning a model ensemble that covers many different selections of OR and AND operators, irrespective of any particular stable state. This means that in our further analyses we used the fold change of the predicted synergy score of a test model against a randomly, yet proliferative, parameterized model (see Methods).
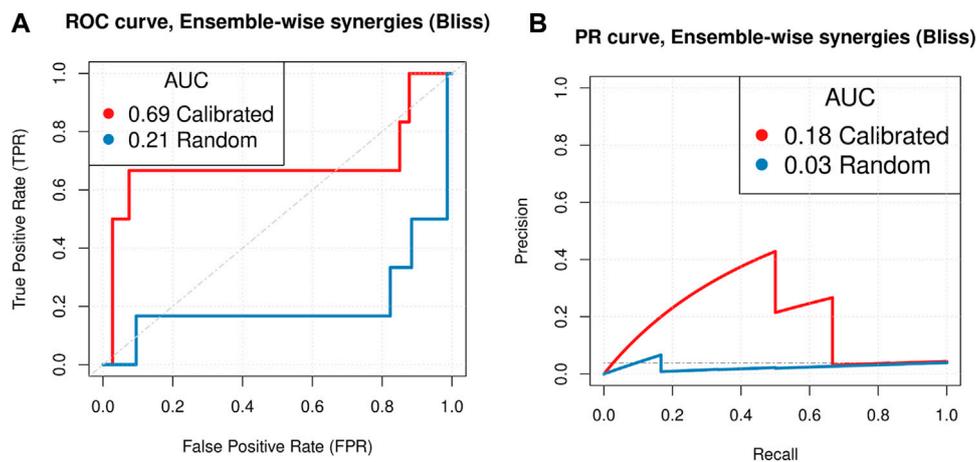
**FIGURE 3**
Model performance. Predictive performance of calibrated and random models based on the CASCADE 2.0 topology was tested against data from a corresponding drug screen (Flobak et al., 2019). Calibrated predictions were generated by model ensembles, trained to steady state data, and normalized to the random model predictions. Random model predictions were generated by collecting predictions from ensembles of models trained to a random yet proliferative phenotype. Models have logical rule mutations only and Bliss Independence was used to assess the model performance. The dashed diagonal **(A)** and horizontal **(B)** lines represent the performance of a random drug synergy classifier. We observe that correctly calibrated models perform substantially better than random models as indicated by a p-value of 0.007 for testing the null hypothesis of no difference in ROC AUCs, and a p-value of 0.18 for the PR AUC (based on 10000 bootstrap resamples).
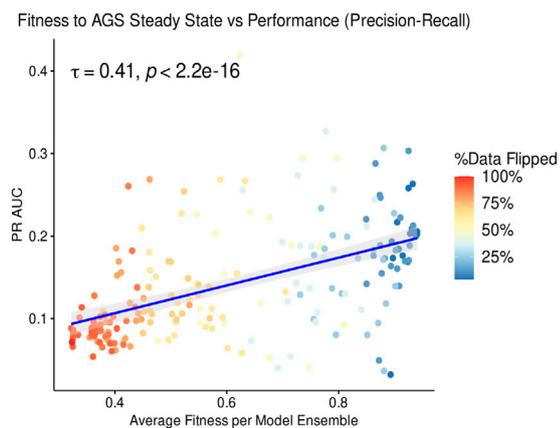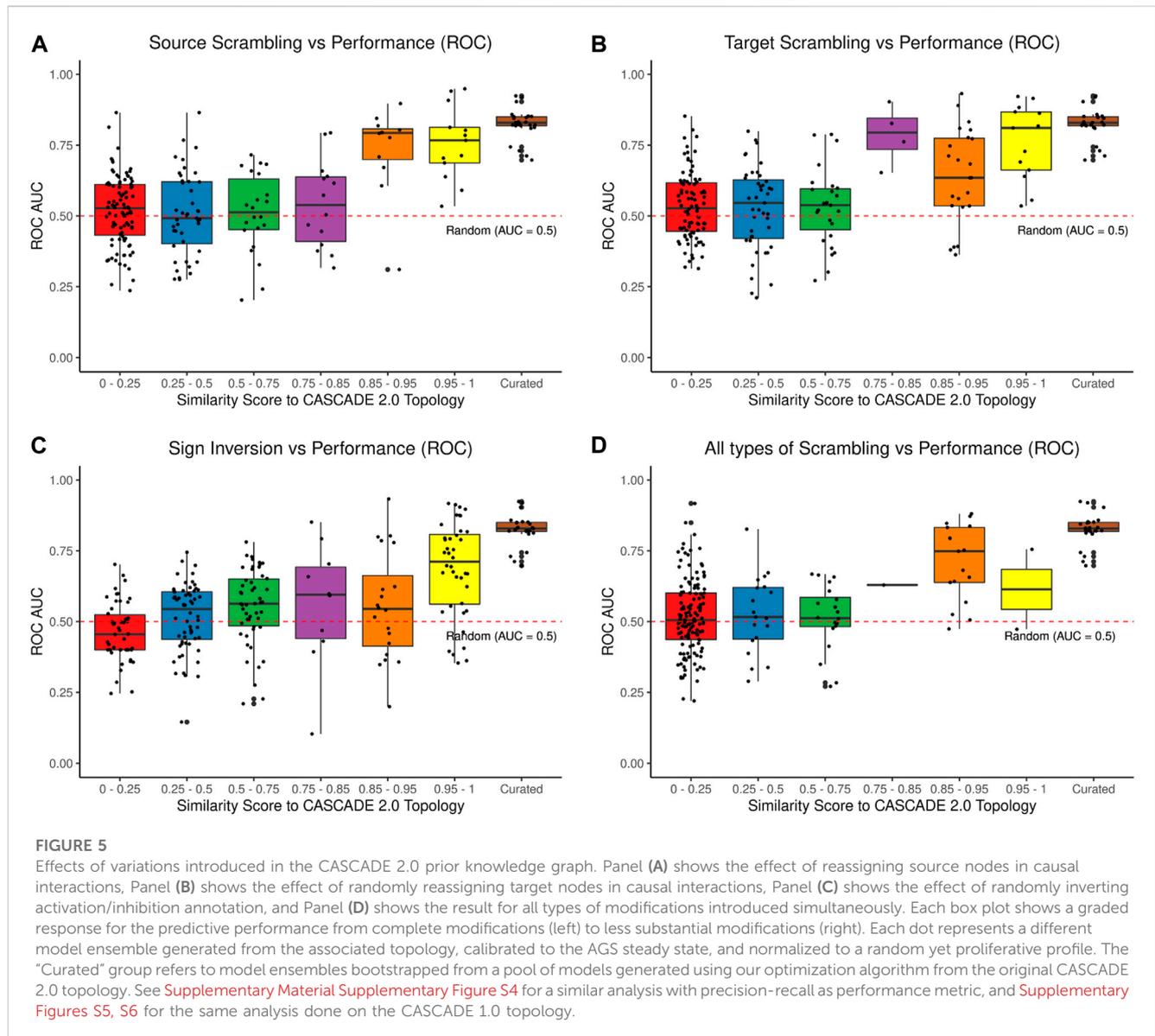


**FIGURE 4**
PR AUC performance dependence on fitness. Each model ensemble, displayed as one dot in the scatterplot, was trained to a partially incorrect steady state signaling profile derived from the biological phenotype of the AGS cell line (Flobak et al., 2015). A total of 205 training profiles were created, each one used to generate one model ensemble consisting of 60 models. The x-axis reports the average fitness of each model ensemble as evaluated to the curated steady state. The colorbar reports the percentage of steady state nodes whose binary value was flipped to train each model ensemble. Because of the non-normality of the data, the Kendall rank-based correlation (Kendall, 1948) test is used to derive the proposed association.

We first tested our software pipeline by considering predictions from model ensembles for simulating the 21 drug combinations that were analyzed previously (Flobak et al., 2015). Synergies were defined as a predicted "growth" output for two drugs together being lower than the product of each individual drug's "growth,"

analogous to the Bliss synergy metric (Bliss, 1939) used in cell culture lab experiments (see Methods). Among 21 drug pairs, 15 were predicted to act synergistically by this definition, exhibiting a range of synergy strengths, and quantified performance of these models was surprisingly high: by selecting different thresholds for synergy predictions a receiver-operating characteristic (ROC) curve (sensitivity vs. 1-specificity) shows a ROC area-under-curve (AUC) of 0.97 and a precision-recall (PR) AUC of 0.91, see Figure 2C. The analysis shows that the top six predictions comprise all four experimentally validated synergies. Notably, the automatically parameterized models produced no false negatives. This effectively means that we could in principle have reduced our full drug screen to only test 29% percent of the combinations (6 out of 21), had we guided our experiments by model simulations, which is comparable to the performance in our manually parameterized model (Flobak et al., 2015).

## Automated model optimization as a solution for larger model topologies

The benefits of automatic parameterization become more apparent in calibration of models with larger topologies. We demonstrate this with the CASCADE 2.0 model, which is a manually curated cancer signaling topology comprising 144 nodes and 367 edges. CASCADE 2.0 includes pathways with TGF-beta, JAK-STAT, and Rho GTPases, as well as extensions of pathways already present in CASCADE 1.0, to enable simulation of a larger set of drug combinations (Niederdorfer et al., 2020). Starting with this large curated model, we analyzed in more detail the effects of automated model training while randomly mutating logical rule configurations and network connectivity, and assessed the results against the biological regulatory mechanisms that were affected. We varied three aspects in the training protocol, each

**FIGURE 5**
Effects of variations introduced in the CASCADE 2.0 prior knowledge graph. Panel **(A)** shows the effect of reassigning source nodes in causal interactions, Panel **(B)** shows the effect of randomly reassigning target nodes in causal interactions, Panel **(C)** shows the effect of randomly inverting activation/inhibition annotation, and Panel **(D)** shows the result for all types of modifications introduced simultaneously. Each box plot shows a graded response for the predictive performance from complete modifications (left) to less substantial modifications (right). Each dot represents a different model ensemble generated from the associated topology, calibrated to the AGS steady state, and normalized to a random yet proliferative profile. The "Curated" group refers to model ensembles bootstrapped from a pool of models generated using our optimization algorithm from the original CASCADE 2.0 topology. See Supplementary Material Supplementary Figure S4 for a similar analysis with precision-recall as performance metric, and Supplementary Figures S5, S6 for the same analysis done on the CASCADE 1.0 topology.

time assessing the effect on the performance of the models for correct synergy prediction:
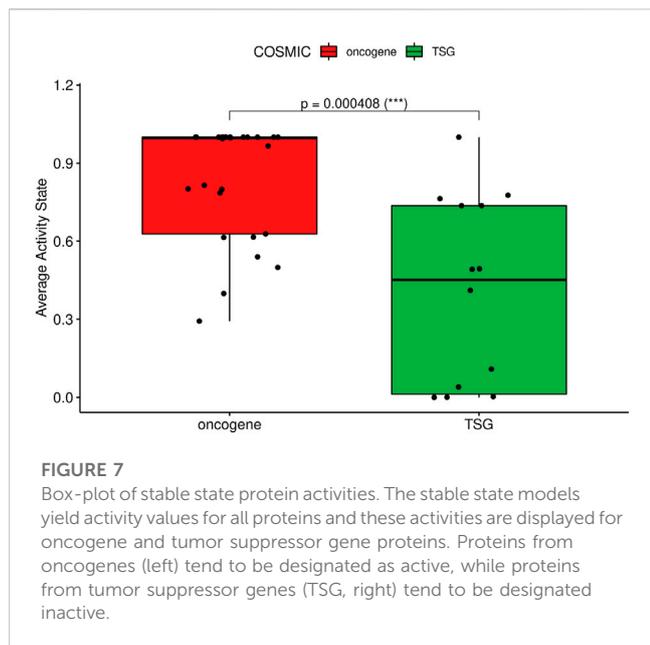
- Optimizing logical rules against partially incorrect calibration data
- Optimizing the regulatory network by stepwise, random removal/inclusion of edges
- Checking the effect of random rewiring of the regulatory network

For each of these model alteration strategies, we not only looked for overall fitness but also in more detail at the represented biological mechanisms that were affected, to judge whether improved or reduced simulation performance could be reconciled with involvement of proteins of regulatory interactions in the context of cancer. The hypothesis was that, while taking the overall value of curated prior knowledge as a given, the relevance of individual regulatory interactions and the precise mathematical representation

of their regulatory effects in specific cancer environments might be difficult to infer from papers and therefore could be algorithmically improved. The effects of the network connectivity and rule mutations were judged in model ensembles and compared with observed synergies. For each model able to reach a stable state, mutations introduced to logic equations could also be reviewed to assess mutual dependencies between edges (or subsets of edges) and corresponding rules. This allowed us to identify parameters and edges that appeared to be essentially fixed and thereby of fundamental importance for model performance.

We compared simulation results from automatically trained ensembles of 450 models to drug synergies in a new drug screen of the AGS cancer cell line comprising 153 combinations of 18 targeted drugs (Flobak et al., 2019). Model training was performed as described in Methods - Model calibration by parameterisation optimization. We found that, in contrast to the CASCADE 1.0 predictions, normalization of topology-intrinsic prediction propensities was critical to the predictive performance

**FIGURE 6**
Combined stable states and parameterization heatmaps. A total of 4500 Boolean models were used for this analysis. Only the CASCADE 2.0 target nodes that have a link-operator in their respective Boolean equation are shown. The 52 link-operator nodes have been assigned to 3 clusters with K-means using the stable states matrix data. The link-operator data heatmap has the same row order as the stable states heatmap. Steady state data (Calibration, where possible activity states are inhibited (red) and active (green)), COSMIC classification of tumor suppressor genes (TSG) and oncogenes, pathway association (Pathway), annotation of nodes which were used as drug targets in the model simulation (Drug Targets, denoted by purple bars), in-degree connectivity (Target connectivity), out-degree connectivity (Source connectivity) and percent agreement between parameterization and stable state annotations (Agreement) are indicated below the heatmaps.

(see Supplementary Figures S1, S2 and Methods - Normalization). We find that models obtained by automated optimization, as described above, could predict drug synergies with a ROC AUC of 0.69 and a PR AUC of 0.18, clearly outperforming random model predictions (see Figure 3). This means that our drug screen could have been reduced from blindly testing all 153 combinations to only the screening of 35 combinations, which would increase the synergy prevalence of tested combinations from 4% (6 of 153) to 11% (4 of 35). We would dismiss 117 drug combinations but at the expense of missing two observed synergies.

## Topology and calibration data needs to be correct

From here on we focus on the CASCADE 2.0 topology, for additional analyses see Supplementary Material, which has similar

experiments with the CASCADE 1.0 topology, underpinning conclusions analogous to those drawn here.

## Impact of data quality on model performance

Since our models are derived from prior knowledge and calibrated based on sample-specific measurements (calibration data), modifications to both the prior knowledge and data must be expected to affect the predictive performance of the models. We first checked how the quality of the calibration data affected model predictions. The performance of models trained to partially incorrect calibration data was expressed as PR AUC and, when plotted against the fitness of these models against the fully correct calibration data, we observe that higher PR AUC correlates with higher fitness of models, indicating that calibration of models indeed improved synergy predictions for our dataset (see Figure 4). However, even models trained to highly incorrect data, with roughly 50% of calibration data flipped (meaning a true fitness

**FIGURE 7**
Box-plot of stable state protein activities. The stable state models yield activity values for all proteins and these activities are displayed for oncogene and tumor suppressor gene proteins. Proteins from oncogenes (left) tend to be designated as active, while proteins from tumor suppressor genes (TSG, right) tend to be designated inactive.

around 0.5), perform better than a random classifier (PR AUC 0.04), indicating that model topology alone already carries information that can be leveraged to predict drug synergies. Note that due to the stochasticity of model calibration, the model ensemble average fitness never reached the extreme fitnesses below ~0.3.

## Randomizing regulatory edges of the curated model reduces predictiveness

As the quality of the calibration data does impact model performance, but not obviate it even if these data are highly incorrect, we next explored the value of the quality of the curated regulatory graph topology. We generated a series of models with various degrees of "scrambled" topologies with modified causal interactions in the regulatory graph (of the type: source - effect - target), by randomly exchanging a particular "source" in a causal interaction with the source of another interaction and investigated the
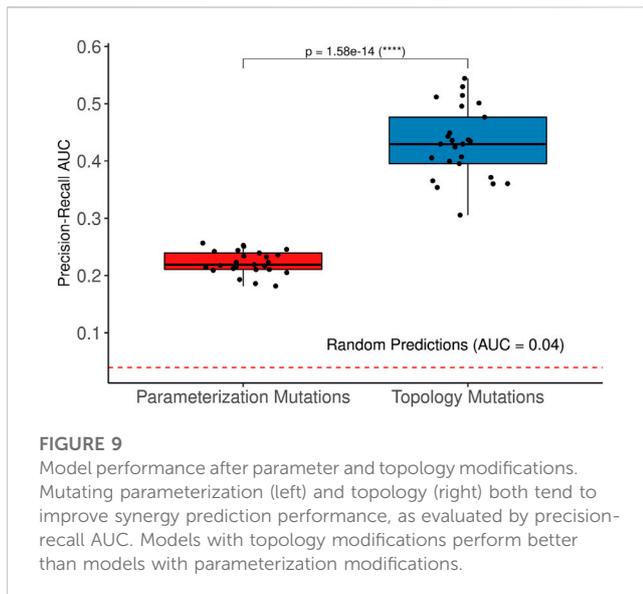
performance of models with these incorrect regulatory interactions. Similarly, we investigated the impact of randomly reassigning target nodes, as well as the impact of inverting signed effects, i.e., from inhibition to activation, and *vice versa*. Note that we will later explore the effect of missing prior knowledge (simple deletions), while here we present results for incorrect prior knowledge. The results (Figure 5) show that even low levels of randomization in the curated knowledge significantly reduce the predictive power of the models, quickly approaching random performance. Overall, we conclude that both calibration data and prior knowledge quality are important to correctly predict drug synergies, and that errors can be detrimental.

## Model ensemble heterogeneity and mechanistic insight

To appreciate the heterogeneity amongst models in model ensembles obtained through the parameter optimization, we studied both attractor and parameterization heterogeneity against model fitness, in subsets of these ensembles selected for specific features (model sub-ensembles). In the heatmap grouped by K-means clustering (Figure 6), calibrated models to a relatively large extent obey the calibration data, with states of steady state nodes mostly identical to the data to which they were trained (the subset of nodes (24 of 144) that were specified in the calibration data). Model stable state vectors (rows in Figure 6) have notable areas of homogeneity, as judged by large stretches of nodes (indicated on the X-axis) that are either all activated (green) or inhibited (red) in all models, but in other areas (e.g., the upper-middle panel of the heatmap) the heterogeneity and discrepancies with calibration data is quite substantial. This heterogeneity was much more widespread in the parameterization space. For some nodes there is high correlation between their parameterization (link operator AND-NOT vs. OR-NOT) and stable state (Inactive or Active, respectively), but for many the correlation is surprisingly low (see Agreement panel in Figure 6). These observations indicate that a) a limited set of training nodes was sufficient to provide homogeneity in parts of the attractor space, and b) large heterogeneity in the parameterization space still can be compliant with homogeneity in the attractor space (see in particular the large green (active) area of the stable-state heatmap). In other words: there are many logical rule configurations that yield models properly
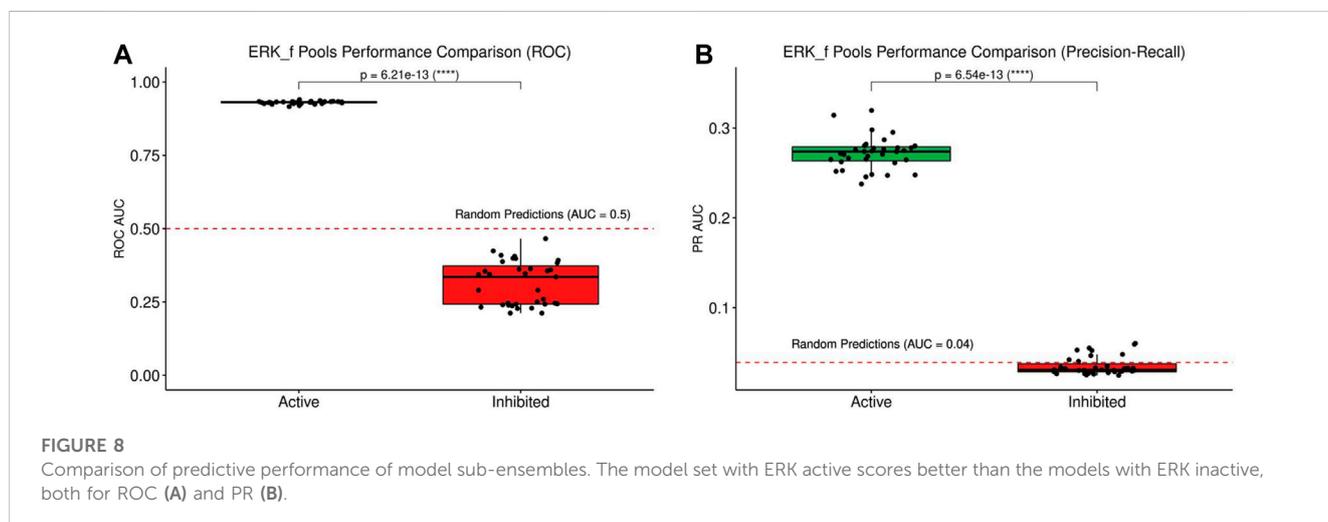


**FIGURE 8**
Comparison of predictive performance of model sub-ensembles. The model set with ERK active scores better than the models with ERK inactive, both for ROC **(A)** and PR **(B)**.

**FIGURE 9**
Model performance after parameter and topology modifications. Mutating parameterization (left) and topology (right) both tend to improve synergy prediction performance, as evaluated by precision-recall AUC. Models with topology modifications perform better than models with parameterization modifications.

representing biological observations compliant with calibration data. This underpins the decision to use model ensembles rather than single models, since these ensembles cover a larger set of parameterizations (behavior) that are all compliant with the input data.

The analysis of node states and parameterization allows us to investigate mechanisms underlying observed behavior and to look for biological explanations for some of the observations. As shown in Figure 6, indicated by panel "COSMIC," the models allow the prediction of activity of several proteins implicated in cancer. Figure 7 shows the analysis of their activity state, and it appears that proteins from genes annotated as oncogenes in COSMIC tend to be active, while proteins from genes annotated as tumor suppressor genes (TSG) tend to be inactive, in overall steady states. This is biologically plausible and attests to the capability of our mechanistic model to generate hypotheses about the underlying molecular biology.

The training of the models to biomarker data never results in the absolute maximum fitness (1), and the stable state analysis (Figure 6) shows that three data points are most often violated: JNK signaling (JNK_f), ERK signaling (ERK_f) and p38 signaling (MAPK14), all members of the MAPK pathway (see Figure 6, stable state panel, black rectangles). These network nodes are all clustered in the highly heterogeneous section in the steady state heatmap. In the manual curation of the CASCADE 1.0 topology (Flobak et al., 2015) it was noted that reports on the activity of ERK in AGS cells varied frequently, with only slightly more than half of the publications reporting ERK to be active. We found that the predictive performance of model versions with ERK being active was significantly higher than the sub-ensemble where ERK is inactive (see Figure 8), suggesting that from a functional point of view ERK should be considered active in AGS cells.

## Rule and edge optimization identifies key regulatory mechanisms
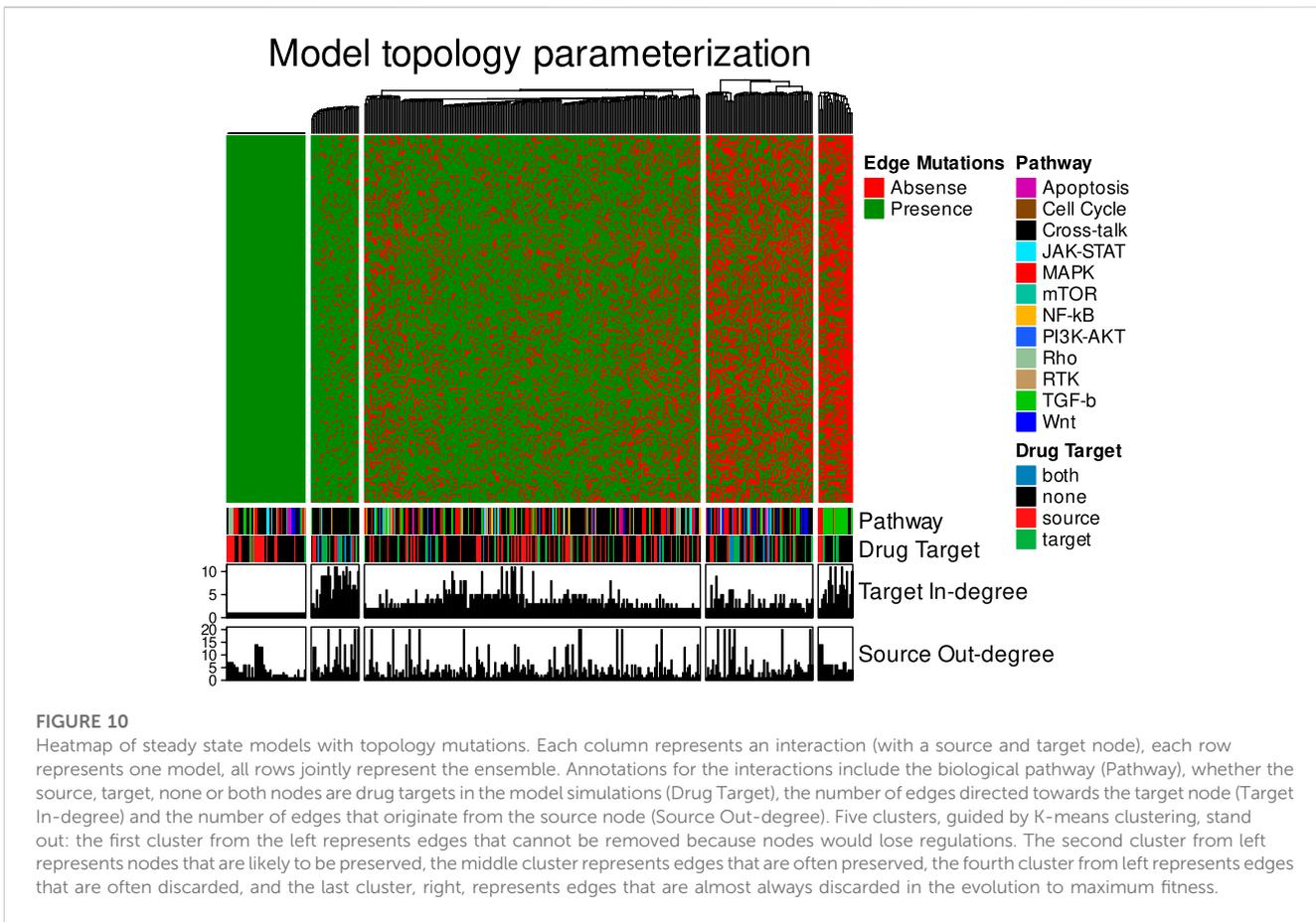
Traditionally, logical model definitions start out with a prior knowledge graph, after which a most optimal parameterization is

sought based on experimental evidence and model behavior. We asked which was most influential to accurately predict synergies: alterations to the topology or to the parameterization, in the evolution to maximum fitness. We configured the genetic algorithm to modify edges in the topology (see Methods - Model calibration by topology optimization), by either removing or subsequently restoring edges from the initial prior knowledge network (as long as no signaling component lost all its source inputs), or to modify the parameterization of logical rules (see Methods—Model calibration by parameterisation optimization). We found that modifications of the edges and the parameterization both resulted in substantially improved prediction performance, significantly better than a random classifier. In particular, the performance (as evaluated by precision-recall) was very high for topology-mutated models, even outperforming models trained by parameterization optimization. While ROC AUC was consistently high, around 0.8 (Supplementary Figure S7), the PR curve showed clear dependency by displaying very high positive predictive values at very conservative sensitivity thresholds, meaning that a predicted drug synergy is highly likely to represent an actual synergy, see Figure 9. Note that there is a major difference between missing data and incorrect data: as was previously demonstrated (Figure 5), model predictive performance suffered severely from including incorrect prior knowledge, while model predictive performance can improve by omitting putatively correct prior knowledge.

Looking at the topology modifications, we hypothesized that deletion of certain edges could be favored by the genetic algorithm, to obtain maximum fitness. Every node in CASCADE 2.0 is annotated to a specific pathway (Niederdorfer et al., 2020), allowing us to assign all edges to a specific pathway, if both source and target node belong to the same pathway, or to crosstalk for edges that link nodes from different pathways, see Figure 10. We observed that certain edge groups are always preserved (the left-most cluster), while other edges are very likely to be removed (cluster on the right). Interestingly, a majority of these removed edges belong to the TGF-beta pathway, in particular representing inhibitory effects of the protein SKI and other inhibitors of SMADs. In the model, SKI itself is inhibited by active AKT signaling, and thus removal of inhibitory edges from SKI allows restricting the activity of some of the SMAD proteins in the model, in particular to SMAD1, SMAD3 and SMAD4, which tend to be inactivated in the topology mutated models. It is evident from Figure 10 that crosstalk is largely preserved during model optimization, potentially relating not only to sparse knowledge of crosstalk in the prior knowledge network, but also to the biological importance of signaling that is not confined to what was more or less arbitrarily viewed as pathways.
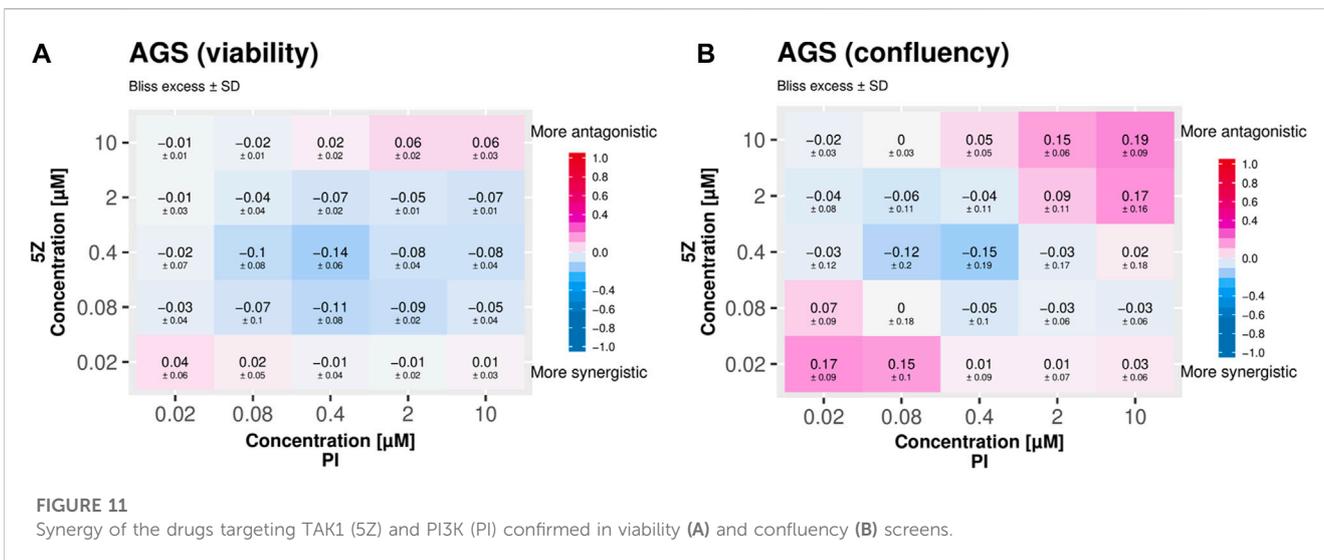
## Validation of synergy predictions *in vivo*

In order to test the translational relevance of our drug synergy prediction platform we performed *in vivo* validation for one of the proposed novel drug synergies. The synergies of TAK1 inhibition combined with PI3K inhibition, already identified in our previous logical modeling work, had not been reported earlier and thus
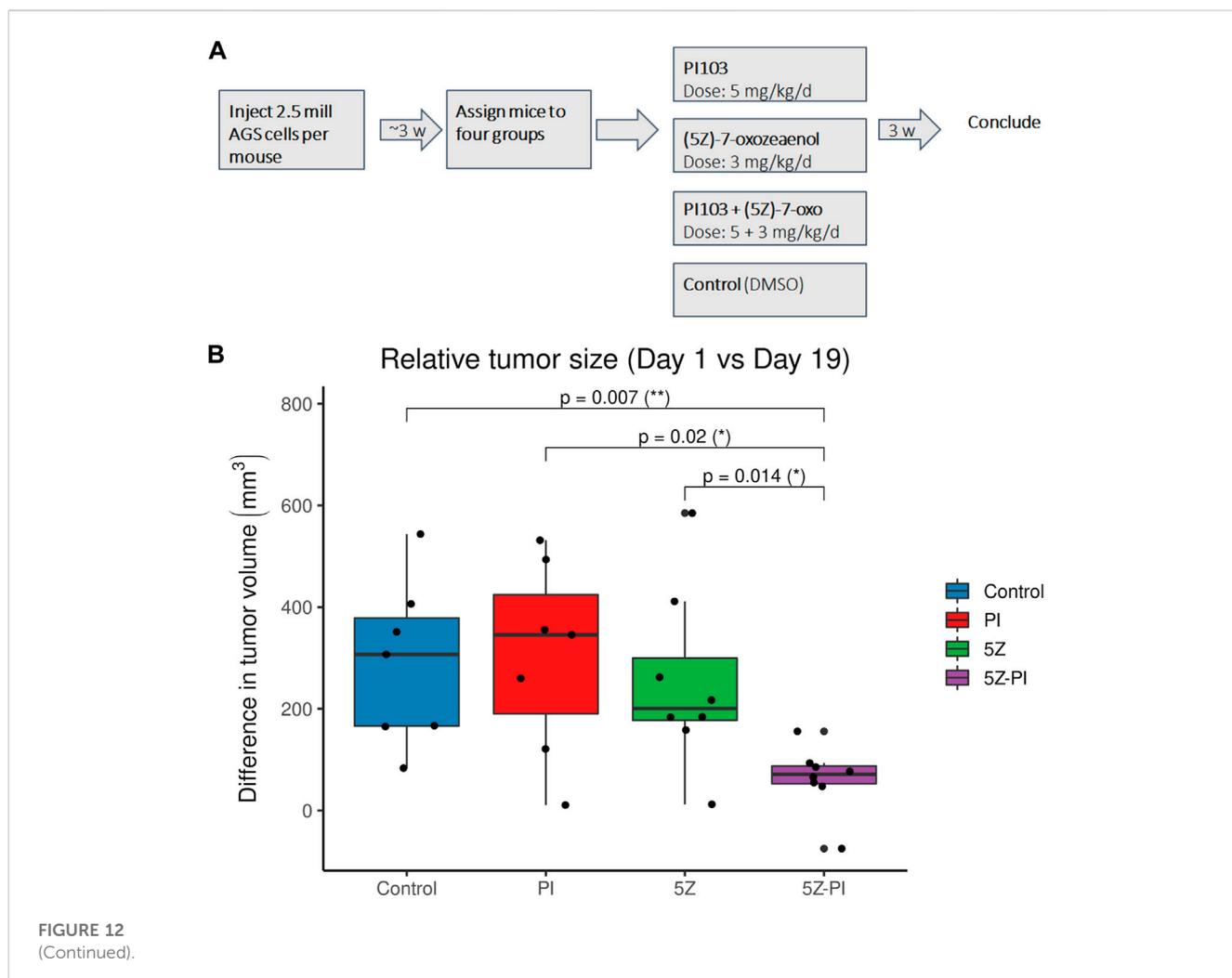
**FIGURE 10**
Heatmap of steady state models with topology mutations. Each column represents an interaction (with a source and target node), each row represents one model, all rows jointly represent the ensemble. Annotations for the interactions include the biological pathway (Pathway), whether the source, target, none or both nodes are drug targets in the model simulations (Drug Target), the number of edges directed towards the target node (Target In-degree) and the number of edges that originate from the source node (Source Out-degree). Five clusters, guided by K-means clustering, stand out: the first cluster from the left represents edges that cannot be removed because nodes would lose regulations. The second cluster from left represents nodes that are likely to be preserved, the middle cluster represents edges that are often preserved, the fourth cluster from left represents edges that are often discarded, and the last cluster, right, represents edges that are almost always discarded in the evolution to maximum fitness.

represent novel synergies of potential interest in future cancer therapy. The robustness of this synergy is evident as we detect the same synergy of combined TAK1 and PI3K inhibition in our framework that integrates automated model parameterization.

In order to test reproducibility across different high throughput drug screen readouts we subjected AGS cancer cells to combined TAK1 and PI3K inhibition and monitored the response both by ATP content measurement (viability) and by microscopy

(confluence), see Figure 11. For both readouts there is a region of synergistic response to drugs applied at medium doses as indicated by the drug concentration gradients. We subcutaneously injected AGS gastric adenocarcinoma xenograft tumors in Balb/c mice to test the synergy of combined inhibition of TAK1 (5Z-7-oxozeaenol) and PI3K (PI103) *in vivo*. The xenograft tumors ($n = 30$) were randomized to four groups: control, PI103, (5Z)-7-oxozeaenol and a combination group which received both PI103 and (5Z)-7- oxozeaenol, see



**FIGURE 11**
Synergy of the drugs targeting TAK1 (5Z) and PI3K (PI) confirmed in viability **(A)** and confluency **(B)** screens.
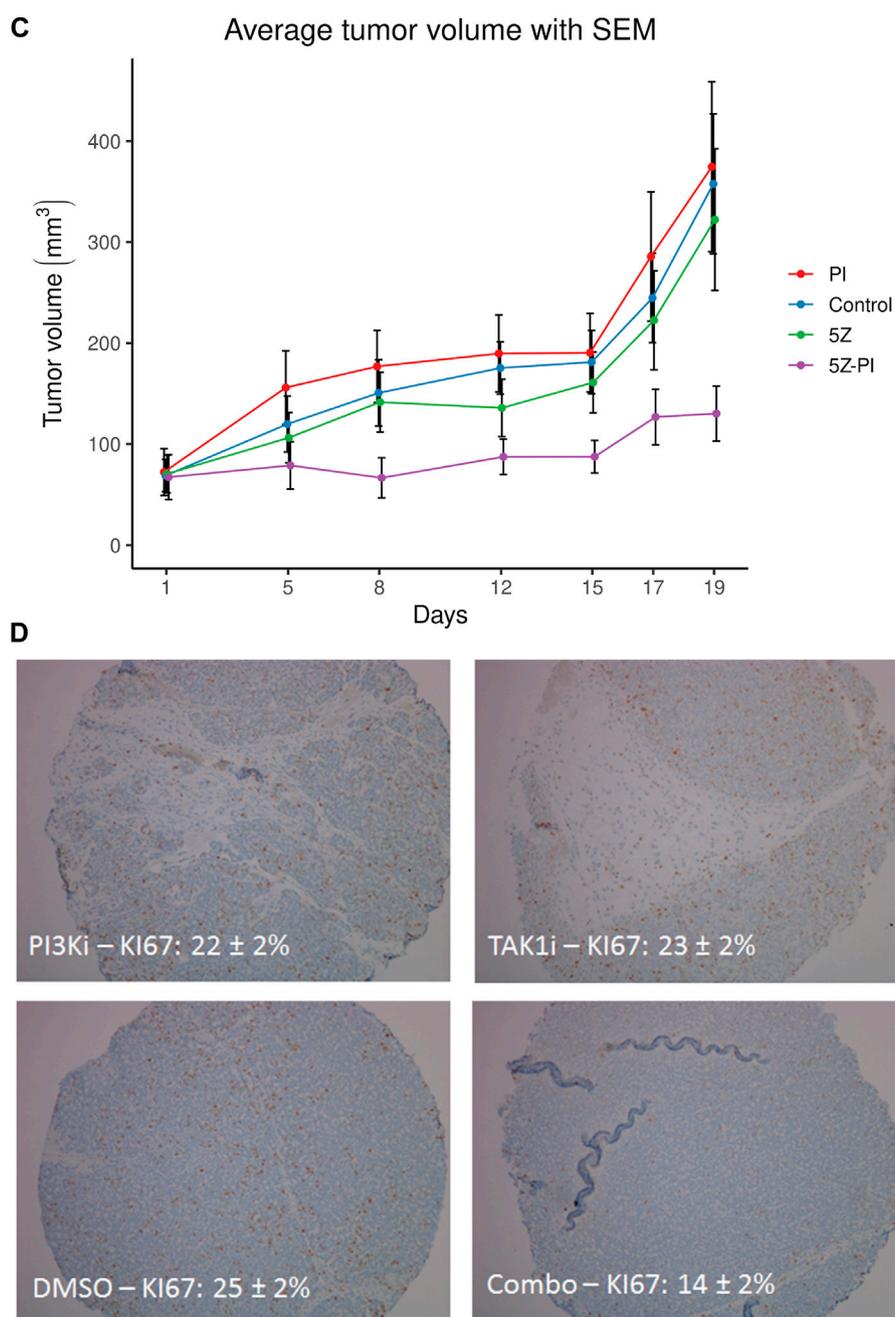
**FIGURE 12**
(Continued).

Figure 12A. At the end of the experiment the combination group displayed significant changes (Wilcoxon rank-sum test) in relative tumor size compared to either single-treatment group (Figures 12B, C). We observed a similar reduced proliferative capacity for tumors in mice treated jointly with TAK1 and PI3K inhibitors, as indicated by tumor proliferation marker Ki67 (Figure 12D). The clear difference between the significant tumor growth inhibitory effect of the combination and the non-significant activity of individual agents strongly indicates a synergistic anti-tumor effect of the two agents. A concern for drug synergies is that possible side-effects might also be expressed in synergistic ways. We therefore chose doses to be at the lower end of effective concentrations, compared to previously published *in vivo* use of the inhibitors (Donev et al., 2011; Bhattacharya et al., 2012; Singh et al., 2012; Fan et al., 2013). Despite low dosage the inhibitors together reduced tumor growth, without signs of pain or weight loss.

## Discussion

Our results show that a curated prior knowledge network with an initial set of logical rules can be automatically parameterized by a genetic algorithm, using a fitness score reflecting how well the global stable state of a model matches the experimentally determined local states of signaling components of a cell line in its native growing state. We showed that the predictive performance from an automatically parameterized ensemble of models was on par with our original, curated CASCADE 1.0 model. We next used our parameterization software to calibrate the larger CASCADE 2.0 network topology. With larger topologies, benefits of automation become more apparent, and can be used to enable simulation for larger numbers of drugs and numbers of cell lines.

Finding drug synergies among the vast set of possible combinations of drugs calls for new approaches. To rationally reduce the prohibitively large experimental search space, we found that our approach can be highly useful to identify sets of drugs that are unlikely to display synergy and that could be omitted from testing. Even tackling the combinatorial complexity for standardized cancer models is already challenging, as exemplified by the AZ-DREAM Challenge (Menden et al., 2019), where pairwise combinations of 118 drugs (6903 possible drug-drug combinations) are tested against a panel of 85 cell lines. If all combinations are screened in $6 \times 6$ matrices this corresponds to over 200,000 384-well plates for four technical replicates, clearly indicating that a trial-and-

**FIGURE 12**
(Continued). **(A)** Xenograft experimental design. Cancer cells were injected subcutaneously in mice and allowed to grow until a visible tumor could be identified, after which the mice were randomized into four groups. Mice were then treated with single drugs and the combination for 19 days after which the experiment was stopped and tumor sizes evaluated. **(B)** Testing of drug synergy in a mouse xenograft model. Tumor volumes determined at the end of the study were compared with tumor volumes at treatment onset. Tumors in mice receiving both inhibitors 5Z-7-oxozeaenol 3 mg/kg/d (5Z) and PI103 5 mg/kg/d (PI) show a smaller increase in size compared to either of the groups receiving only single inhibitors, and the control group. The combination effect was statistically significantly different from either single drug therapy as evaluated by Mann-Whitney U tests with corresponding $p$-values shown above the boxplots. Similar results were obtained using t-tests with a chosen significance level of $p = 0.05$. **(C)** Average tumor volume for the four groups of mice with standard error of the mean (SEM) indicated by the error bars. The group receiving both inhibitors (5Z + PI) displays a more inhibited tumor growth than either of the groups receiving each single inhibitor and the control group. **(D)** Ki67 proliferation index (count of Ki67 positive cells) reduced upon joint combination of TAK1 and PI3K inhibition.

error approach is not economic for drug synergy discovery. Whereas most approaches for drug synergy predictions rely on perturbation data for training a classifier (Yang et al., 2015; Preuer et al., 2017; Fröhlich et al., 2018; Ianevski et al., 2019; Tognetti et al., 2021), our approach works well with calibration data obtained from an unperturbed system, which greatly reduces the cost of data acquisition and opens possibilities for clinical applications in which such data is less easily available.

Assessing the performance of drug synergy predictions is met with several challenges. First, targeted drugs, as those employed here ('small molecule' inhibitors), can affect a number of other targets in addition to their intended target, known as 'off-target' effects (Klaeger et al., 2017). Since our simulation is based on canonical drug targets annotated for each drug, any information missing about off-target effects must be expected to impact simulations. In addition, drug synergy is an elusive concept itself, with different mathematical reference models producing different synergy scores (Vlot et al., 2019). Finally, high throughput drug screens, as employed here, typically reports drug responses based on measured residual ATP content after drug exposure, which is known not to capture all growth-reducing drug responses (Gautam et al., 2016; Bae et al., 2020; Folkesson et al., 2020). Despite these limitations, which must be expected to reduce the performance of any drug synergy prediction approach, our logical simulation-based *in silico* pre-selection approach performs immensely better than a blinded screen that would assay the same numbers of candidates: at a sensitivity of 50%, roughly 35%-40% of a pre-selected set of predicted synergies will be observed in follow-up drug synergy experiments in a drug screen, where only 4% of drug combinations acted synergistically overall.

Our choice of a logical framework for computational simulations comes both with some benefits and limitations. Logic equations are very quick to evaluate, with high simulation speed enabling extensive simulations even on regular desktop computers. However, logic equations as employed here only allow two activity states for model components: active and inactive, and the inference of these activity states from measurements of e.g., RNA expression or DNA mutations is not trivial, and typically involves manual interpretation (Niederdorfer et al., 2020). Moreover, only two interaction strengths between components are allowed: full interaction or no interaction. These limitations still to a large extent meet the demands and possibilities offered by experiments with present day laboratory techniques. Our implementation of logical model simulations only computes stable states, thereby discarding potential complex attractors of models that would require asynchronous simulation for full characterization. Palli et al. chose to characterize model behavior by synchronous updating, which can allow identification of some complex attractors, but at the risk of identifying also artificial attractors (Palli et al., 2019). Our choice of focusing on stable states was motivated by considering the computational efficiency, as computation of complex attractors in addition to stable states would severely tax our simulations. One possible avenue for future research will be to account for also complex attractors, either by approximations as offered by e.g., trap space analysis, or by a full characterization of model behavior. Other solutions could include sampling stochastically a part of the model behavior space, as done by Béal et al. (2019) for personalization of logical models by the implementation of Monte-Carlo kinetic algorithms, also allowing to infer network transition probabilities. Park et al. (2020) computed the behavior of models during one simulation consisting of multiple time steps. Here, partial node inhibition was approximated by setting the drug target node inactive in a fraction of simulation steps corresponding to the fraction of target inhibition.

We observed that our approach is somewhat sensitive to errors in the calibration data, and even more sensitive to errors in the prior knowledge, indicating that curation quality is paramount to our modeling approach. This demands for adequate causal statement curation protocols and standards that feed into high quality general and cancer signaling databases, a demand that is materializing (Tü et al., 2016; Cristobal Monraz Gomez et al., 2019; Licata et al., 2020; Touré et al., 2020; Touré et al., 2021).

Interestingly, we found that deleting a small fraction of the regulatory links from the prior knowledge network can be very powerful in optimizing models for drug synergy predictions, whereas randomly rearranging regulatory links is very detrimental to model performance. Logical model construction can be performed by curating data resources or the literature, or by relying on high quality curated databases like Signor (Licata et al., 2020), SignaLink (Csabai et al., 2022) or IntAct (Orchard et al., 2014). If quality of these resources, or the *ad hoc* curation of literature, would not be of high standard, this would significantly limit the performance of the resulting model. On the other hand, calling a prior knowledge network complete is essentially a judgment call, as all models are limited. This demonstrates that, while lacking in completeness, high quality curation can produce models that can predict drug synergies. Furthermore, whereas the inclusion of a regulatory link ideally needs evidence from observations about the functional relevance of such a link in the cell that is modeled, our observations about ERK activity highlight the variability of experimental data concerning such observations. It is therefore not unreasonable to accept that an optimization algorithm can choose to dismiss regulatory links for the benefit of improved model performance. The increasing availability of high quality curated molecular causal interaction data opens a perspective toward fully automated model building, where algorithmic topology optimization can fine tune a model to perform adequately, for any target node requirements and cell type for which baseline biomarker data is available. A welcome feature of automatically parameterized logical models is their innate ability to suggest mechanisms underpinning a particular observation. Such model-driven hypotheses can lay foundations for targeted follow-up experiments that provide observations for directed model revisions, resulting in a model with higher validity.

It has been suggested that network topology alone already explains drug synergy (Cokol et al., 2011; Jaeger et al., 2017), and that parameterization to a lesser extent defines or refines synergy predictions (Yin et al., 2014). Experimentally it has been observed that drug synergies tend to vary between cell lines, with the most frequently observed synergistic drug pair only effective in about half of the cell lines analyzed (Axelrod et al., 2013; Amzallag et al., 2019). In our analysis we observed that drug synergy predictions depend both on the specific parameterization of a given topology and on the topology itself. When we put our manually defined topology to the test, drug synergy predictions are more accurate for models optimized to represent cell-specific baseline biomarkers in their local states, compared to unconstrained local states. One may speculate that the interactions relevant to describe drug combination effects represent a subset of all potential (general) interactions, and that this subset varies from cell line to cell line. From a completeness perspective, given the limited knowledge of

molecular biology today, any model representation will be a major simplification of reality, yet some of these models work.

## Summary

Boolean models representing cell fate signaling can be used to predict drug combination responses. However, identifying exact logical operators for computer models that are based on prior knowledge can be challenging. We here present a computational pipeline that can automatically configure ensembles of Boolean models compliant with prior knowledge and observed baseline activity data.

We find that algorithmic-guided deletion of regulatory edges leads to increased predictive performance, while incorrect prior knowledge can drastically decrease it. Moreover, with correct prior knowledge, our models show tolerance to incorrect activity data. Overall, our simulations demonstrate that prior topological knowledge is the most important factor for the construction of accurate predictive models capable of providing mechanistic biological insights.

We benchmark our software pipeline against a dataset reporting on responses to 153 drug combinations, showing that we can effectively enrich the prevalence of synergies from 4% in an un-guided drug screen to 11% in a set of model-guided predicted synergies. Considering the many limiting factors that hamper the development of effective rational combinatorial cancer therapies, including the lack of prior knowledge, incorrect activity assessments, drug target promiscuity, and the absence of consensus for drug synergy quantification, our pipeline can contribute to the improvement of personalized cancer treatment.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

## Ethics statement

Ethical approval was not required for the studies on humans in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used. The animal study was approved by the Unit of Comparative Medicine, Faculty of Medicine, Norwegian University of Science and Technology. The study was conducted in accordance with the local legislation and institutional requirements.

## Author contributions

ÅF: Devised the project and developed the main conceptual ideas. Designed and developed prototype for simulation pipeline.

Developed and implemented drug response combination simulation software. Analysed drug response prediction data and model performance. Designed and executed cell line experiments. Manually curated drug combination response data. Designed and executed xenograft experiment. Supervised drug combination matrix experiments. Supervised tissue microarray experiments. JZ: Developed and implemented production-ready drug response combination simulation software. Executed various simulations, analysed drug response prediction data and model performance, produced resulting figures, produced Supplementary Material, produced github repositories, made all analyses reproducible. MV: Analysed drug combination data. Developed and implemented drug response combination simulation software. TS: Supervised xenograft experiments. LT: Supervised xenograft experiments. AG: Prepared and analysed tissue microarrays of xenograft tumours. BN: Manually curated drug combination response data. Designed and executed drug combination matrix experiments. EF: Manually curated drug combination response data. Designed and executed drug combination matrix experiments. MK: Analysed drug response prediction data and model performance. Supervised the project. All authors contributed to the article and approved the submitted version.

## Funding

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fsysb.2023.1252961/full#supplementary-material

# References

Al-Lazikani, B., Banerji, U., and Workman, P. (2012). Combinatorial drug therapy for cancer in the post-genomic era. *Nat. Biotechnol.* 30, 679–692. doi:10.1038/nbt.2284

Amzallag, A., Ramaswamy, S., and Benes, C. H. (2019). Statistical assessment and visualization of synergies for large-scale sparse drug combination datasets. *BMC Bioinforma.* 20, 83. doi:10.1186/s12859-019-2642-7

Axelrod, M., Gordon, V. L., Conaway, M., Tarcsafalvi, A., Neitzke, D. J., Gioeli, D., et al. (2013). Combinatorial drug screening identifies compensatory pathway interactions and adaptive resistance mechanisms. *Oncotarget* 4, 622–635. doi:10.18632/oncotarget.938

Bae, S. Y., Guan, N., Yan, R., Warner, K., Taylor, S. D., and Meyer, A. S. (2020). Measurement and models accounting for cell death capture hidden variation in compound response. *Cell Death Dis.* 11, 255. doi:10.1038/s41419-020-2462-8

Bansal, M., Yang, J., Karan, C., Menden, M. P., Costello, J. C., Tang, H., et al. (2014). A community computational challenge to predict the activity of pairs of compounds. *Nat. Biotechnol.* 32, 1213–1222. doi:10.1038/nbt.3052

Béal, J., Montagud, A., Traynard, P., Barillot, E., and Calzone, L. (2019). Personalization of logical models with multi-omics data allows clinical stratification of patients. *Front. Physiol.* 9, 1965. doi:10.3389/fphys.2018.01965

Bhattacharya, B., Akram, M., Balasubramanian, I., Tam, K. K. Y., Koh, K. X., Yee, M. Q., et al. (2012). Pharmacologic synergy between dual phosphoinositide-3-kinase and mammalian target of rapamycin inhibition and 5-fluorouracil in PIK3CA mutant gastric cancer cells. *Cancer Biol. Ther.* 13, 34–42. doi:10.4161/cbt.13.1.18437

Bliss, C. I. (1939). The toxicity of poisons applied jointly. *Ann. Appl. Biol.* 26, 585–615. doi:10.1111/j.1744-7348.1939.tb06990.x

Cokol, M., Chua, H. N., Tasan, M., Mutlu, B., Weinstein, Z. B., Suzuki, Y., et al. (2011). Systematic exploration of synergistic drug pairs. *Mol. Syst. Biol.* 7, 544. doi:10.1038/msb.2011.71

Cristobal Monraz Gomez, L., Kondratova, M., Ravel, J. M., Barillot, E., Zinovyev, A., and Kuperstein, I. (2019). Application of atlas of cancer signalling network in preclinical studies. *Brief. Bioinform* 20, 701–716. doi:10.1093/bib/bby031

Csabai, L., Fazekas, D., Kadlecsik, T., Szalay-Bekő, M., Bohár, B., Madgwick, M., et al. (2022). SignaLink3: a multi-layered resource to uncover tissue-specific signaling networks. *Nucleic Acids Res.* 50, D701–D709. doi:10.1093/nar/gkab909

Donev, I. S., Wang, W., Yamada, T., Li, Q., Takeuchi, S., Matsumoto, K., et al. (2011). Transient PI3K inhibition induces apoptosis and overcomes HGF-mediated resistance to EGFR-TKIs in EGFR mutant lung cancer. *Clin. Cancer Res.* 17, 2260–2269. doi:10.1158/1078-0432.CCR-10-1993

Dorier, J., Crespo, I, Niknejad, A., Liechti, R., Ebeling, M., and Xenarios, I. (2016). Boolean regulatory network reconstruction using literature based knowledge with a genetic algorithm optimization method. *BMC Bioinforma.* 17, 410. doi:10.1186/s12859-016-1287-z

Eduati, F., Doldàn-Martelli, V., Klinger, B., Cokelaer, T., Sieber, A., Kogera, F., et al. (2017). Drug resistance mechanisms in colorectal cancer dissected with cell type-specific dynamic logic models. *Cancer Res.* 77, 3364–3375. doi:10.1158/0008-5472.CAN-17-0078

Eduati, F., Jaaks, P., Wappler, J., Cramer, T., Merten, C. A., Garnett, M. J., et al. (2020). Patient-specific logic models of signaling pathways from screenings on cancer biopsies to prioritize personalized combination therapies. *Mol. Syst. Biol.* 16, 86644–e8713. doi:10.15252/msb.20188664

Fan, Y., Cheng, J., Vasudevan, S. a., Patel, R. H., Liang, L., Xu, X., et al. (2013). TAK1 inhibitor 5Z-7-oxozeaenol sensitizes neuroblastoma to chemotherapy. *Apoptosis* 18, 1224–1234. doi:10.1007/s10495-013-0864-0

Flobak, Å., Baudot, A., Remy, E., Thommesen, L., Thieffry, D., Kuiper, M., et al. (2015). Discovery of drug synergies in gastric cancer cells predicted by logical modeling. *PLoS Comput. Biol.* 11, e1004426. doi:10.1371/journal.pcbi.1004426

Flobak, Å., Niederdorfer, B., Nakstad, V. T., Thommesen, L., Klinkenberg, G., and Lægreid, A. (2019). A high-throughput drug combination screen of targeted small molecule inhibitors in cancer cell lines. *Sci. Data* 6, 237. doi:10.1038/s41597-019-0255-7

Flobak, Å., Vazquez, M., Lægreid, A., and Valencia, A. (2017). CImbinator: a web-based tool for drug synergy analysis in small-and large-scale datasets. *Bioinformatics* 33, 2410–2412. doi:10.1093/bioinformatics/btx161

Folkesson, E., Niederdorfer, B., Nakstad, V. T., Thommesen, L., Klinkenberg, G., Lægreid, A., et al. (2020). High-throughput screening reveals higher synergistic effect of MEK inhibitor combinations in colon cancer spheroids. *Sci. Rep.* 10, 11574. doi:10.1038/s41598-020-68441-0

Fröhlich, F., Kessler, T., Weindl, D., Shadrin, A., Schmiester, L., Hache, H., et al. (2018). Efficient parameter estimation enables the prediction of drug response using a mechanistic pan-cancer pathway model. *Cell Syst.* 7, 567–579. doi:10.1016/j.cels.2018.10.013

Gautam, P., Karhinen, L., Szwajda, A., Jha, S. K., Yadav, B., Aittokallio, T., et al. (2016). Identification of selective cytotoxic and synthetic lethal drug responses in triple negative breast cancer cells. *Mol. Cancer* 15, 34–16. doi:10.1186/s12943-016-0517-3

Gayvert, K. M., Aly, O., Platt, J., Bosenberg, M. W., Stern, D. F., and Elemento, O. (2017). A computational approach for identifying synergistic drug combinations. *PLoS Comput. Biol.* 13, e1005308–e1005311. doi:10.1371/journal.pcbi.1005308

Gjerga, E., Trairatphisan, P., Gabor, A., Koch, H., Chevalier, C., Ceccarelli, F., et al. (2020). Converting networks to predictive logic models from perturbation signalling data with CellNOpt. *Bioinformatics* 36, 4523–4524. doi:10.1093/bioinformatics/btaa561

Goswami, C. P., Cheng, L., Alexander, P. S., Singal, a, and Li, L. (2015). A new drug combinatory effect prediction algorithm on the cancer cell based on gene expression and dose-response curve. *CPT pharmacometrics Syst. Pharmacol.* 4, e9. doi:10.1002/psp4.9

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. doi:10.1016/j.cell.2011.02.013

Holland, J. H. (1992). "Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. MIT press. doi:10.7551/mitpress/1090.001.0001

Ianevski, A., Giri, A. K., Gautam, P., Kononov, A., Potdar, S., Saarela, J., et al. (2019). Prediction of drug combination effects with a minimal set of experiments. *Nat. Mach. Intell.* 1 (12), 568–577. doi:10.1038/s42256-019-0122-4

Jaeger, S., Igea, A., Arroyo, R., Alcalde, V., Canovas, B., Orozco, M., et al. (2017). Quantification of pathway cross-talk reveals novel synergistic drug combinations for breast cancer. *Cancer Res.* 77, 459–469. doi:10.1158/0008-5472.CAN-16-0097

Jin, G., Zhao, H., Zhou, X., and Wong, S. T. C. (2011). An enhanced Petri-net model to predict synergistic effects of pairwise drug combinations from gene microarray data. *Bioinformatics* 27, i310–i316. doi:10.1093/bioinformatics/btr202

Kauffman, S. a. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.* 22, 437–467. doi:10.1016/0022-5193(69)90015-0

Kendall, M. G. (1948). *Rank correlation methods.* London: Griffin.

Klaeger, S., Heinzlmeir, S., Wilhelm, M., Polzer, H., Vick, B., Koenig, P.-A., et al. (2017). The target landscape of clinical kinase drugs. *Science* 358, eaan4368. doi:10.1126/science.aan4368

Klinger, B., Sieber, A., Fritsche-Guenther, R., Witzel, F., Berry, L., Schumacher, D., et al. (2013). Network quantification of EGFR signaling unveils potential for targeted combination therapy. *Mol. Syst. Biol.* 9, 673. doi:10.1038/msb.2013.29

Lehár, J., Krueger, A. S., Avery, W., Heilbut, A. M., Johansen, L. M., Price, E. R., et al. (2009). Synergistic drug combinations tend to improve therapeutically relevant selectivity. *Nat. Biotechnol.* 27, 659–666. doi:10.1038/nbt.1549

Li, H., Li, T., Quang, D., and Guan, Y. (2018). Network propagation predicts drug synergy in cancers. *Cancer Res.* 78, 5446–5457. doi:10.1158/0008-5472.CAN-18-0740

Licata, L., Lo Surdo, P., Iannuccelli, M., Palma, A., Micarelli, E., Perfetto, L., et al. (2020). SIGNOR 2.0, the SIGnaling network open resource 2.0: 2019 update. *Nucleic Acids Res.* 48, D504–D510–D510. doi:10.1093/nar/gkz949

Menden, M. P., Wang, D., Mason, M. J., Szalai, B., Bulusu, K. C., Guan, Y., et al. (2019). Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. *Nat. Commun.* 10, 2674. doi:10.1038/s41467-019-09799-2

Mendoza, L., and Xenarios, I. (2006). A method for the generation of standardized qualitative dynamical systems of regulatory networks. *Theor. Biol. Med. Model* 3, 13–18. doi:10.1186/1742-4682-3-13

Miller, M. L., Molinelli, E. J., Nair, J. S., Sheikh, T., Samy, R., Jing, X., et al. (2013). Drug synergy screen and network modeling in dedifferentiated liposarcoma identifies CDK4 and IGF1R as synergistic drug targets. *Sci. Signal* 6, ra85. doi:10.1126/scisignal.2004014

Naldi, A. (2018). BioLQM: a Java toolkit for the manipulation and conversion of logical qualitative models of biological networks. *Front. Physiol.* 9, 1605–1610. doi:10.3389/fphys.2018.01605

Nelander, S., Wang, W., Nilsson, B., She, Q.-B., Pratilas, C., Rosen, N., et al. (2008). Models from experiments: combinatorial drug perturbations of cancer cells. *Mol. Syst. Biol.* 4, 216. doi:10.1038/msb.2008.53

Niederdorfer, B., Touré, V., Vazquez, M., Thommesen, L., Kuiper, M., Lægreid, A., et al. (2020). Strategies to enhance logic modeling-based cell line-specific drug synergy prediction. *Front. Physiol.* 11, 862. doi:10.3389/fphys.2020.00862

Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., et al. (2014). The MIntAct project--IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* 42, D358–D363. doi:10.1093/nar/gkt1115

Palli, R., Palshikar, M. G., and Thakar, J. (2019). Executable pathway analysis using ensemble discrete-state modeling for large-scale data. *PLoS Comput. Biol.* 15, e1007317–e1007321. doi:10.1371/journal.pcbi.1007317

Park, S. M., Hwang, C. Y., Choi, J., Joung, C. Y., and Cho, K. H. (2020). Feedback analysis identifies a combination target for overcoming adaptive resistance to targeted cancer therapy. *Oncogene* 39, 3803–3820. doi:10.1038/s41388-020-1255-y

Preuer, K., Lewis, R. P. I., Hochreiter, S., Bender, A., Bulusu, K. C., and Klambauer, G. (2017). DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics* 34, 1538–1546. doi:10.1093/bioinformatics/btx806

Saez-Rodriguez, J., Alexopoulos, L. G., Epperlein, J., Samaga, R., Lauffenburger, D. a., Klamt, S., et al. (2009). Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. *Mol. Syst. Biol.* 5, 331. doi:10.1038/msb.2009.87

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi:10.1101/gr.1239303

Singh, A., Sweeney, M. F., Yu, M., Burger, A., Greninger, P., Benes, C., et al. (2012). TAK1 inhibition promotes apoptosis in KRAS-dependent colon cancers. *Cell* 148, 639–650. doi:10.1016/j.cell.2011.12.033

Sun, Y., Sheng, Z., Ma, C., Tang, K., Zhu, R., Wu, Z., et al. (2015). Combining genomic and network characteristics for extended capability in predicting synergistic drugs for cancer. *Nat. Commun.* 6, 8481. doi:10.1038/ncomms9481

Tang, J., Gautam, P., Gupta, A., He, L., Timonen, S., Akimov, Y., et al. (2019). Network pharmacology modeling identifies synergistic Aurora B and ZAK interaction in triple-negative breast cancer. *npj Syst. Biol. Appl.* 5, 20. doi:10.1038/s41540-019-0098-z

Thomas, R. (1973). Boolean formalization of genetic control circuits. *J. Theor. Biol.* 42, 563–585. doi:10.1016/0022-5193(73)90247-6

Tognetti, M., Gabor, A., Yang, M., Cappelletti, V., Windhager, J., Rueda, O. M., et al. (2021). Deciphering the signaling network of breast cancer improves drug sensitivity prediction. *Cell Syst.* 12, 401–418.e12. doi:10.1016/j.cels.2021.04.002

Touré, V., Vercruysse, S., Acencio, M. L., Lovering, R. C., Orchard, S., Bradley, G., et al. (2020). The minimum information about a molecular interaction causal statement (MI2CAST). *Bioinformatics* 36, 5712–5718. doi:10.1093/bioinformatics/btaa622

Touré, V., Zobolas, J., Kuiper, M., and Vercruysse, S. (2021). CausalBuilder: bringing the MI2CAST causal interaction annotation standard to the curator. *Database (Oxford)* 2021, 1–6. doi:10.1093/database/baaa107

Tsirvouli, E., Touré, V., Niederdorfer, B., Vázquez, M., Flobak, Å, and Kuiper, M. (2020). A middle-out modeling strategy to extend a colon cancer logical model improves drug synergy predictions in epithelial-derived cancer cell lines. *Front. Mol. Biosci.* 7, 502573. doi:10.3389/fmolb.2020.502573

Türei, D., Korcsmáros, T., and Saez-Rodriguez, J. (2016). OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat. Methods* 13, 966–967. doi:10.1038/nmeth.4077

Veliz-Cuba, A., Aguilar, B., Hinkelmann, F., and Laubenbacher, R. (2014). Steady state analysis of Boolean molecular network models via model reduction and computational algebra. *BMC Bioinforma.* 15, 221. doi:10.1186/1471-2105-15-221

Vlot, A. H. C., Aniceto, N., Menden, M. P., Ulrich-Merzenich, G., and Bender, A. (2019). Applying drug synergy metrics to oncology combination screening data: agreements, disagreements and pitfalls. *Drug Discov. Today* 00. doi:10.1016/j.drudis.2019.09.002

Yang, J., Tang, H., Li, Y., Zhong, R., Wang, T., Wong, S. T. C., et al. (2015). DIGRE: drug-induced genomic residual effect model for successful prediction of multidrug effects. *CPT Pharmacometrics Syst. Pharmacol.* 4, 91–97. doi:10.1002/psp4.1

Yin, N., Ma, W., Pei, J., Ouyang, Q., Tang, C., and Lai, L. (2014). Synergistic and antagonistic drug combinations depend on network topology. *PLoS One* 9, e93960. doi:10.1371/journal.pone.0093960

Yu, M. K., Ma, J., Fisher, J., Kreisberg, J. F., Raphael, B. J., and Ideker, T. (2018). Visible machine learning for biomedicine. *Cell* 173, 1562–1565. doi:10.1016/j.cell.2018.05.056

Zobolas, J., Monteiro, P. T., Kuiper, M., and Flobak, Å. (2022). Boolean function metrics can assist modelers to check and choose logical rules. *Journal of Theoretical Biology* 538, 111025. doi:10.1016/J.JTBI.2022.111025