Patient-specific Boolean models of signalling networks guide personalised treatments

Arnau Montagud$^{1,2,3,4,*}$, Jonas Béal$^{1,2,3}$, Luis Tobalina$^{5,5}$, Pauline Traynard$^{1,2,3}$, Vigneshwari Subramanian$^{5,5}$, Bence Szalai$^{5,5}$, Róbert Alföldi$^{7}$, László Puskás$^{7}$, Alfonso Valencia$^{4,6}$, Emmanuel Barillot$^{1,2,3}$, Julio Saez-Rodriguez$^{5,5,9,9}$, Laurence Calzone$^{1,2,3,*}$

1 Institut Curie, PSL Research University, Paris, France
2 INSERM, U900, Paris, France
3 MINES ParisTech, PSL Research University, CBIO-Centre for Computational Biology, Paris, France
4 Barcelona Supercomputing Center (BSC), Barcelona, Spain
5 Faculty of Medicine, Joint Research Centre for Computational Biomedicine (JRC–COMBINE), RWTH Aachen University, 52074 Aachen, Germany
6 Semmelweis University, Faculty of Medicine, Department of Physiology, Budapest, Hungary
7 Astridbio Technologies Ltd., 6728 Szeged, Hungary
8 Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010, Barcelona, Spain
9 Faculty of Medicine and Heidelberg University Hospital, Institute of Computational Biomedicine, Heidelberg University, Heidelberg, Germany
10 $ Current address: Bioinformatics and Data Science, Research and Early Development, Oncology R&D, AstraZeneca, Cambridge, UK
11 § Current address: Data Science & Artificial Intelligence, Imaging & Data Analytics, Clinical Pharmacology & Safety Sciences, R&D, AstraZeneca, Gothenburg, Sweden

* corresponding authors: arnau.montagud@bsc.es, laurence.calzone@curie.fr
# These authors contributed equally to this work and should be considered co-senior authors.

Abstract

Prostate cancer is the second most occurring cancer in men worldwide. To better understand the mechanisms of tumorigenesis and possible treatment responses, we developed a mathematical model of prostate cancer which considers the major signalling pathways known to be deregulated. We personalised this Boolean model to molecular data to reflect the heterogeneity and specific response to perturbations of cancer patients. 488 prostate samples were used to build patient-specific models and compared to available clinical data. Additionally, eight prostate cell-line-specific models were built to validate our approach with dose-response data of several drugs. The effects of single and combined drugs were tested in these models under different growth conditions. We identified 15 actionable points of interventions in one cell-line-specific model whose inactivation hinders tumorigenesis. To validate these results, we tested nine small molecule inhibitors of five of those putative targets and found a dose-dependent effect on four of them, notably those
targeting HSP90 and PI3K. These results highlight the predictive power of our personalised Boolean models and illustrate how they can be used for precision oncology.

Introduction

Like most cancers, prostate cancer arises from mutations in single somatic cells that induce deregulations in processes such as proliferation, invasion of adjacent tissues and metastasis. Not all prostate patients respond to the treatments in the same way, depending on the stage and type of their tumour (Chen and Zhou, 2016) and differences in their genetic and epigenetic profiles (Toth et al., 2019; Yang et al., 2018). The high heterogeneity of these profiles can be explained by a large number of interacting proteins and the complex cross-talks between the cell signalling pathways that can be altered in cancer cells. Because of this complexity, understanding the process of tumorigenesis and tumour growth would benefit from a systemic and dynamical description of the disease. At the molecular level, this can be tackled by a simplified mechanistic cell-wide model of protein interactions of the underlying pathways, dependent on external environmental signals.

Although continuous mathematical modelling has been widely used to study cellular biochemistry dynamics (e.g., ordinary differential equations) (Goldbeter, 2002; Kholodenko et al., 1995; Le Novère, 2015; Sible and Tyson, 2007; Tyson et al., 2019), this formalism does not scale up well to large signalling networks, due to the difficulty of estimating kinetic parameter values (Babtie and Stumpf, 2017). In contrast, the logical (or logic) modelling formalism represents a simpler means of abstraction where the causal relationships between proteins (or genes) are encoded with logic statements, and dynamical behaviours are represented by transitions between discrete states of the system (Kauffman, 1969; Thomas, 1973). In particular, Boolean models, the simplest implementation of logical models, describe each protein as a binary variable (ON/OFF). This framework is flexible, requires in principle no quantitative information, can be hence applied to large networks combining multiple pathways, and can also provide a qualitative understanding of molecular systems lacking detailed mechanistic information.

In the last years, logical and, in particular, Boolean modelling has successfully been used to describe the dynamics of human cellular signal transduction and gene regulations (Calzone et al., 2010; Cho et al., 2016; Flobak et al., 2015; Grieco et al., 2013; Helikar et al., 2008; Traynard et al., 2016) and their deregulation in cancer (Fumiã and Martins, 2013; Hu et al., 2015). Numerous applications of logical modelling have shown that this framework is able to delineate the main dynamical properties of complex biological regulatory networks (Abou-Jaoudé et al., 2011; Faure et al., 2006).

However, the Boolean approach is purely qualitative and does not consider the real time of cellular events (half time of proteins, triggering of apoptosis, etc.). To cope with this issue, we developed the MaBoSS software to compute continuous Markov Chain simulations on the model state transition graph (STG), in which a model state is defined as a vector of nodes that are either active or inactive. In practice, MaBoSS associates transition rates for activation and inhibition of each node of the network, enabling it to account for different time scales of the processes described by the model. Given some initial conditions, MaBoSS applies a Monte-Carlo kinetic algorithm (or Gillespie algorithm) to the STG to produce time trajectories (Stoll et al., 2017, 2012) such that time evolution of the model state probabilities.
can be estimated. Stochastic simulations can easily explore the model dynamics with different initial conditions by varying the probability of having a node active at the beginning of the simulations and by modifying the model such that it accounts for genetic and environmental perturbations (e.g., presence or absence of growth factors, or death receptors). For each case, the effect on the probabilities of selected read-outs can be measured (Cohen et al., 2015; Montagud et al., 2017).

When summarising the biological knowledge into a network and translating it into logical terms, the obtained model is generic and cannot explain the differences and heterogeneity between patients’ responses to treatments. Models can be trained with dedicated perturbation experiments (Dorier et al., 2016; Saez-Rodriguez et al., 2009), but such data can only be obtained with non-standard procedures such as microfluidics from patients’ material (Eduati et al., 2020). To address this limitation, we developed a methodology to use different omics data that are more commonly available to personalise generic models to individual cancer patients or cell lines and verified that the obtained models correlated with clinical results such as patient survival information (Béal et al., 2019). In the present work, we apply this approach to prostate cancer to suggest targeted therapy to patients based on their omics profile (Figure 1). We first built 488 patient- and eight cell line-prostate-specific models using data from The Cancer Genome Atlas (TCGA) and the Genomics of Drug Sensitivity in Cancer (GDSC) projects, respectively. Simulating these models with the MaBoSS framework, we identified points of intervention that diminish the probability of reaching pro-tumorigenic phenotypes. Lastly, we developed a new methodology to simulate drug effects on these data-tailored Boolean models and present a list of viable drugs and regimes that could be used on these patient- and cell-line-specific models for optimal results. Experimental validations were performed on the LNCaP prostate cell line with two predicted targets, confirming the predictions of the model.

Results

Prostate Boolean model construction

A network of signalling pathways and genes relevant for prostate cancer progression was assembled to recapitulate the potential deregulations that lead to high-grade tumours. Dynamical properties were added onto this network to perform simulations, uncover therapeutic targets and explore drug combinations. The model was built upon a generic cancer Boolean model by Fumiã and Martins (2013), which integrates major signalling pathways and their substantial cross-talks. The pathways include the regulation of cell death and proliferation in many tumours.

This initial generic network was extended to include prostate-cancer-specific genes (e.g., SPOP, AR, etc.), pathways identified using ROMA (Martignetti et al., 2016), OmniPath (Türei et al., 2021) and up-to-date literature. ROMA is applied on omics data, either transcriptomics or proteomics. In each pathway, the genes that contribute the most to the overdispersion are selected. ROMA was applied to the TCGA transcriptomics data using gene sets from cancer pathway databases (Appendix 1, Section 1.1.3, Appendix figure 1). These results were used as guidelines to extend the network to fully cover the alterations found in prostate cancer patients. OmniPath was used to complete our network finding connections between the
proteins of interest known to play a role in the prostate and the ones identified with ROMA,
and the list of genes already present in the model (Appendix 1, Sections 1.1.3 and 1.1.4,
Appendix figures 2 and 3). The final network includes pathways such as androgen receptor,
MAPK, Wnt, NFkB, PI3K/AKT, MAPK, mTOR, SHH, the cell cycle, the epithelial-
mesenchymal transition (EMT), apoptosis and DNA damage pathways.

This network was then converted into a Boolean model where all variables can take two
values: 0 (inactivate or absent) or 1 (activate or present). Our model aims at predicting
prostate phenotypic behaviours for healthy and cancer cells in different conditions. Nine
inputs that represent some of these physiological conditions of interest were considered:
Epithelial Growth Factor (EGF), Fibroblast Growth Factor (FGF), Transforming Growth
Factor beta (TGFbeta), Nutrients, Hypoxia, Acidosis, Androgen, Tumour Necrosis Factor
alpha (TNF alpha) and Carcinogen. These input nodes have no regulation. Their value is
fixed according to the simulated experiment to represent the status of the
microenvironmental characteristics (e.g., the presence or absence of growth factors, oxygen,
etc.). A more complex multiscale approach would be required to consider the dynamical
interaction with other cell types.

We defined six variables as output nodes that allow the integration of multiple phenotypic
signals and simplify the analysis of the model. Two of these phenotypes represent the
possible growth status of the cell: Proliferation and Apoptosis. Apoptosis is activated by
Caspase 8 or Caspase 9, while Proliferation is activated by cyclins D and B (read-outs of the
G1 and M phases, respectively). The Proliferation output is described in published models
as specific stationary protein activation patterns, namely the following sequence of activation
of cyclins: Cyclin D, then Cyclin E, then Cyclin A, and finally Cyclin B (Traynard et al., 2016).
Here, we considered a proper sequence when Cyclin D activates first, allowing the release
of the transcriptional factor E2F1 from the inhibitory complex it was forming with RB
(retinoblastoma protein), and then triggering a series of events leading to the activation of
Cyclin B, responsible for the cell’s entry into mitosis (Appendix 1, Section 2.2, Appendix
figure 5). We also define several phenotypic outputs that are readouts of cancer hallmarks:
Invasion, Migration, (bone) Metastasis and DNA repair. The final model accounts for 133
nodes and 449 edges (Figure 2, Supplementary File 1, and in GINsim format at the address:
Prostate Boolean model simulation

The model can be considered as a model of healthy prostate cells when no mutants (or fused genes) are present. We refer to this model as the wild type model. These healthy cells mostly exhibit quiescence (neither proliferation nor apoptosis) in the absence of any input (Figure 3A). When Nutrients and growth factors (EGF or FGF) are present, Proliferation is activated (Figure 3B). Androgen is necessary for AR activation and helps in the activation of Proliferation, even though it is not necessary when Nutrients or growth factors are present. Cell death factors (such as Caspase 8 or 9) trigger Apoptosis in the absence of SPOP, while Hypoxia and Carcinogen facilitate apoptosis but are not necessary if cell death factors are present (Figure 3C).

In our model, the progression towards metastasis is described as a stepwise process. Invasion is first activated by known pro-invasive proteins: either β-catenin (Francis et al., 2013) or a combination of CDH2 (De Wever et al., 2004), SMAD (Daroqui et al., 2012) or EZH2 (Ren et al., 2012). Migration is then activated by Invasion and EMT and with either AKT or AR (Castoria et al., 2011). Lastly, (bone) Metastasis is activated by Migration and one of three nodes: RUNX2 (Altieri et al., 2009), ERG (Adamo and Ladomery, 2016) or ERG fused with TMPRSS2 (St John et al., 2012), FLI1, ETV1 or ETV4 (The Cancer Genome Atlas Research Network, 2015).

This prostate Boolean model was simulated stochastically using MaBoSS (Stoll et al., 2017, 2012) and validated by recapitulating known phenotypes of prostate cells under physiological conditions (Figure 3 and Appendix 1, Sections 2.2 and 2.3, Appendix figures 5-7). In particular, we tested that combinations of inputs lead to non-aberrant phenotypes such as growth factors leading to apoptosis in wild type conditions; we also verified that the cell cycle events occur in proper order: as CyclinD gets activated, RB1 is phosphorylated and turned OFF, allowing E2F1 to mediate the synthesis of CyclinB (see Supplementary File 2 for the jupyter notebook and the simulation of diverse cellular conditions).

Personalisation of the prostate Boolean model

Personalised TCGA prostate cancer patient Boolean models

We tailored the generic prostate Boolean model to a set of 488 TCGA prostate cancer patients (Appendix 1, Section 4, Appendix figure 9) using our personalisation method (PROFILE, (Béal et al., 2019)), constructing 488 individual Boolean models, one for each patient. Personalised models were built using three types of data: discrete data such as mutations and copy number alterations (CNA) and continuous data such as RNAseq data. For discrete data, the nodes corresponding to the mutations or the CNA were forced to 0 or 1 according to the effect of alterations, based on a priori knowledge (i.e., if the mutation was reported to be activating or inhibiting the gene’s activity). For continuous data, the personalisation method modifies the value for the transition rates of model variables and their initial conditions to influence the probability of some transitions. This corresponds, in a biologically-meaningful way, to translating genetic mutations as lasting modifications making the gene independent of regulation, and to translating RNA expression levels as modulation
of a signal but not changing the regulation rules (see Materials and Methods and in Appendix 1, Section 4.1, Appendix figure 10-14).

We assess the general behaviour of the individual patient-specific models by comparing the model outputs (i.e., probabilities to reach certain phenotypes) with clinical data. Here, the clinical data consist of a Gleason grade score associated with each patient, which in turn corresponds to the gravity of the tumour based on its appearance and the stage of invasion (Chen and Zhou, 2016; Gleason, 1992, 1977). We gathered output probabilities for all patient-specific models and confronted them to their Gleason scores. The phenotype DNA_repair, which can be interpreted as a sensor of DNA damage and genome integrity which could lead to DNA repair, seems to separate low and high Gleason scores (Figure 4A and Appendix 1, Section 4.1, Appendix figures 15-18), confirming that DNA damage pathways are activated in patients (Marshall et al., 2019) but may not lead to the triggering of apoptosis in this model (Appendix 1, Section 4.1, Appendix figure 11). Also, the centroids of Gleason grades tend to move following Proliferation, Migration and Invasion variables. We then looked at the profiles of the phenotype scores across patients and their Gleason grade and found that the density of high Proliferation score (close to 1, Figure 4B) tends to increase as the Gleason score increases (from low to intermediate to high) and these distributions are significantly different (Kruskal-Wallis rank sum test, p-value=0.00207; Appendix 1, Section 4.1). The Apoptosis phenotype, however, does not have a clear trend across grades’ probabilities (Figure 4C), even though the distributions are significantly different (Kruskal-Wallis rank sum test, p-value=2.83E-6; Appendix 1, Section 4.1).

Personalised drug predictions of TCGA Boolean models

Using the 488 TCGA-patient-specific models, we looked in each patient for genes that, when inhibited, hamper Proliferation or promote Apoptosis in the model. We focused on these inhibitions as most drugs interfere with the protein activity related to these genes, even though our methodology allows us to study increased protein activity related to over-expression of genes as well (Béal et al., 2019; Montagud et al., 2017). Interestingly, we found several genes that were found as suitable points of intervention in most of the patients (MYC_MAX complex and SPOP were identified in more than 80% of the cases) (Appendix 1, Section 4.2, Appendix figure 19 and 20), but others were specific to only some of the patients (MXI1 was identified in only 4 patients, 1% of the total, GLI in only 7% and WNT in 8% of patients). All the TCGA-specific personalised models can be found in Supplementary File 3, and the TCGA mutants and their phenotype scores can be found in Supplementary File 4.

Furthermore, we explored the possibility of finding combinations of treatments that could reduce the Proliferation phenotype or increase the Apoptosis one. To lower the computational power need, we narrowed down the list of potential candidates to a set of selected genes that are targets of already-developed drugs relevant in cancer progression (Table 1) and analysed the simulations of the models with all the single and combined perturbations.

We used the models to grade the effect that the combined treatments have in each one of the 488 TCGA-patient-specific models’ phenotypes. This list of combinations of treatments can be used to compare the effects of drugs on each TCGA patient and allows us to propose
some of them for individual patients and to suggest drugs suitable to groups of patients (Supplementary File 4). Indeed, the inactivation of some of the targeted genes had a greater effect in some patients than in others, suggesting the possibility for the design of personalised drug treatments. For instance, for the TCGA-EJ-5527 patient, the use of MYC_MAX complex inhibitor reduced Proliferation to 66%. For this patient, combining MYC_MAX with other inhibitors, such as AR or AKT, did not further reduce the Proliferation score (67% in these cases). Other patients have MYC_MAX as an interesting drug target, but the inhibition of this complex did not have such a dramatic effect on their Proliferation scores as in the case of TCGA-EJ-5527. Likewise, for the TCGA-H9-A6BX patient, the use of SPOP inhibitor increased Apoptosis by 87%, while the use of a combination of cFLAR and SPOP inhibitors further increased Apoptosis by 89%. For the rest of this section, we focus on the analysis of clinical groups rather than individuals.

Studying the decrease of Proliferation, we found that AKT is the top hit in Gleason Grades 1, 2, 3, and 4, seconded by EGFR and SPOP in Grade 1, by SPOP and PIP3 in Grade 2, by PIP3 and AR in Grade 3, and by CyclinD and MYC_MAX in Grade 4. MYC_MAX is the top hit in Grade 5, seconded by AR (Appendix 1, Section 4.2, Appendix figure 19). In regards to the increase of Apoptosis, SPOP is the top hit in all grades, seconded by SSH in Grades 1, 2 and 3 and by AKT in Grade 4 (Appendix 1, Section 4.2, Appendix figure 20). It is interesting to note here that many of these genes are targeted by drugs (Table 1). Notably, AR is the target of the drug Enzalutamide, which is indicated for men with an advanced stage of the disease (Scott, 2018), or that MYC is the target of BET bromodomain inhibitors and are generally effective in castration-resistant prostate cancer cases (Coleman et al., 2019).

The work on patient data provided some possible insights and suggested patient- and grade-specific potential targets. To validate our approach experimentally, we personalised the prostate model to different prostate cell lines, where we performed drug assays to confirm the predictions of the model.

Personalised drug predictions of LNCaP Boolean model

We applied the methodology for personalisation of the prostate model to eight prostate cell lines available in GDSC (Iorio et al., 2016): 22RV1, BPH-1, DU-145, NCI-H660, PC-3, PWR-1E and VCaP (results in Appendix File, Section 5 and are publicly available in Supplementary File 5). We decided to focus the validation on one cell line, LNCaP.

LNCaP, first isolated from a human metastatic prostate adenocarcinoma found in a lymph node (Horoszewicz et al., 1983), is one of the most widely used cell lines for prostate cancer studies. Androgen-sensitive LNCaP cells are representative of patients sensitive to treatments as opposed to resistant cell lines such as DU-145. Additionally, LNCaP cells have been used to obtain numerous subsequent derivatives with different characteristics (Cunningham and You, 2015).

The LNCaP personalisation was performed based on mutations as discrete data and RNA-Seq as continuous data. The resulting LNCaP-specific Boolean model was then used to identify all possible combinations of mutations (interpreted as effects of therapies) and to study the synergy of these perturbations. For that purpose, we automatically performed single and double mutant analyses on the LNCaP-specific model (knock-out and overexpression) (Montagud et al., 2017) and focused on the model phenotype probabilities
as read-outs of the simulations. The analysis of the complete set of simulations for the 32258 mutants can be found in the Appendix 1, Section 6.1 and in Supplementary File 6, where the LNCaP-cell-line-specific mutants and their phenotype scores are reported for all mutants. Among all combinations, we identified the top 20 knock-out mutations that depleted \textit{Proliferation} or increased \textit{Apoptosis} the most. As some of them overlapped, we ended up with 29 nodes: \textit{AKT}, \textit{AR}, \textit{ATR}, \textit{AXIN1}, \textit{Bak}, \textit{BIRC5}, \textit{CDH2}, \textit{cFLAR}, \textit{CyclinB}, \textit{CyclinD}, \textit{E2F1}, \textit{eEF2K}, \textit{eEF2}, \textit{EGFR}, \textit{ERK}, \textit{HSPs}, \textit{MED12}, \textit{mTORC1}, \textit{mTORC2}, \textit{MYC}, \textit{MYC\_MAX}, \textit{PHDs}, \textit{PI3K}, \textit{PIP3}, \textit{SPOP}, \textit{TAK1}, \textit{TWIST1}, and \textit{VHL}. We used the scores of these nodes to further trim down the list to have 10 final nodes (\textit{AKT}, \textit{AR}, \textit{cFLAR}, \textit{EGFR}, \textit{ERK}, \textit{HSPs}, \textit{MYC\_MAX}, \textit{SPOP} and \textit{PI3K}) and added 7 other nodes whose genes are considered relevant in cancer biology, such as \textit{AR\_ERG} fusion, \textit{Caspase8}, \textit{HIF1}, \textit{GLUT1}, \textit{MEK1\_2}, \textit{p14ARF}, \textit{ROS} and \textit{TERT} (Table 1). We did not consider the overexpression mutants as they have a very difficult translation to drug uses and clinical practices.

To further analyse the mutant effects, we simulated the LNCaP model with increasing node inhibition values to mimic the effect of drugs’ dosages using a methodology we specifically developed for these purposes (PROFILE\_v2 and available at https://github.com/ArnauMontagud/PROFILE\_v2). Six simulations were done for each inhibited node, with 100\% of node activity (no inhibition), 80\%, 60\%, 40\%, 20\% and 0\% (full knock-out) (see Methods). A nutrient-rich media with EGF was used for these simulations, and we show results on three additional sets of initial conditions in the Appendix 1, Section 7.3.1, Appendix figure 35). Notably, the \textit{Apoptosis} score was slightly promoted when knocking out \textit{SPOP} under all growth conditions (Appendix 1, Section 7.3.1, Appendix figure 35). Likewise, \textit{Proliferation} depletion was accomplished when \textit{HSPs} or \textit{MYC\_MAX} were inhibited under all conditions and, less notably, when \textit{ERK}, \textit{EGFR}, \textit{SPOP} or \textit{PI3K} were inhibited (Appendix 1, Section 7.3.1, Appendix figure 35).

Additionally, these gradual inhibition analyses can be combined to study the interaction of two simultaneously inhibiting nodes (Appendix 1, Section 7.3.2, Appendix figure 36 and 37). For instance, the combined gradual inhibition of \textit{ERK} and \textit{MYC\_MAX} nodes affects the \textit{Proliferation} score in a balanced manner (Figure 5A) even though \textit{MYC\_MAX} seems to affect this phenotype more, notably at low activity levels. By extracting subnetworks of interaction around \textit{ERK} and \textit{MYC\_MAX} and comparing them, we found that the pathways they belong to have complementary downstream targets participating in cell proliferation through targets in MAPK and cell cycle pathways. This complementarity could explain the synergistic effects observed (Figure 5A and 5C).

Lastly, drug synergies can be studied using Bliss Independence using the results from single and combined simulations with gradual inhibitions. This score compares the combined effect of two drugs with the effect of each one of them, with a synergy when the value of this score is lower than 1. We found that the combined inhibition of \textit{ERK} and \textit{MYC\_MAX} nodes on the \textit{Proliferation} score was synergistic (Figure 5C). Another synergistic pair is the combined gradual inhibition of \textit{HSPs} and \textit{PI3K} nodes that also affects the \textit{Proliferation} score in a joint manner (Figure 5B), with some Bliss Independence synergy found (Figure 5D). A complete
study on the Bliss Independence synergy of all the drugs considered in the present work on
Proliferation and Apoptosis phenotypes can be found in Appendix 1, Section 7.3.2, Appendix
figures 38 and 39.

Experimental validation of predicted targets

Drugs associated with the proposed targets

To identify drugs that could act as potential inhibitors of the genes identified with the Boolean
model, we explored the drug-target associations in DrugBank (Wishart et al., 2018) and
ChEMBL (Gaulton et al., 2017). We found drugs that targeted almost all genes
corresponding to the nodes of interest in Table 1, except for cFLAR, p14ARF and SPOP.
However, we could not identify experimental cases where drugs targeting both members of
the proposed combinations were available (Appendix 1, Section 7.1 and in Supplementary
File 6). One possible explanation is that the combinations predicted by the model suggest, in
some cases, to overexpress the potential target and most of the drugs available act as
inhibitors of their targets.

Using the cell-line specific models, we tested if the LNCaP cell line was more sensitive than
the rest of the prostate cell lines to the LNCaP-specific drugs identified in Table 1. We
compared GDSC’s Z-score of these drugs in LNCaP with their Z-scores in all GDSC cell
lines (Figure 6 and Appendix 1, Section 7.2, Appendix figure 33). We observed that LNCaP
is more sensitive to drugs targeting AKT or TERT than the rest of the studied prostate cell
lines. Furthermore, we saw that the drugs that targeted the genes included in the model
allowed the identification of cell line specificities (Appendix 1, Section 7.1). For instance,
target enrichment analysis showed that LNCaP cell lines are especially sensitive to drugs
targeting PI3K/AKT/mTOR, hormone-related (AR targeting) and Chromatin (bromodomain
inhibitors, regulating Myc) pathways (adjusted p-values from target enrichment: 0.001, 0.001
and 0.032, respectively, Appendix 1, Section 7.1, Appendix table 2), which corresponds to
the model predictions (Table 1). Also, the LNCaP cell line is more sensitive to drugs
targeting model-identified nodes than to drugs targeting other proteins (Appendix 1, Section
7.1, Appendix figure 32, Mann–Whitney p-value 0.00041), and this effect is specific for
LNCaP cell line (Mann–Whitney p-values ranging from 0.0033 to 0.38 for other prostate
cancer cell lines).

Overall, the drugs proposed through this analysis suggest the possibility to repurpose drugs
that are used in treating other forms of cancer for prostate cancer and open the avenue for
further experimental validations based on these suggestions.

Experimental validation of drugs in LNCaP

To validate the model predictions of the candidate drugs, we selected four drugs that target
HSPs and PI3K and tested them in LNCaP cell line experiments by using endpoint cell
viability measurement assays and real-time cell survival assays using the xCELLigence
system (see Methods). The drug selection was a compromise between the drugs identified
by our analyses (Table 1) and their effect in diminishing LNCaP’s proliferation (see the
previous section). In both assays, drugs that target HSP90AA1 and PI3K/AKT pathway
genres retrieved from the model analyses were found to be effective against cell proliferation.
The Hsp90 chaperone is expressed abundantly and plays a crucial role in the correct folding of a wide variety of proteins such as protein kinases and steroid hormone receptors (Schopf et al., 2017). Hsp90 can act as a protector of less stable proteins produced by DNA mutations in cancer cells (Barrott and Haystead, 2013; Hessenkemper and Banaihmad, 2013). Currently, Hsp90 inhibitors are in clinical trials for multiple indications in cancer (Chen et al., 2019; Iwai et al., 2012; Le et al., 2017). The PI3K/AKT signalling pathway controls many different cellular processes such as cell growth, motility, proliferation, and apoptosis and is frequently altered in different cancer cells (Carceles-Cordon et al., 2020; Shorning et al., 2020). Many PI3K/AKT inhibitors are in different stages of clinical development, and some of them are approved for clinical use (Table 1).

Notably, Hsp90 (NMS-E973,17-DMAG) and PI3K/AKT pathway (PI-103, Pictilisib) inhibitors showed a dose-dependent activity in the endpoint cell viability assay determined by the fluorescent resazurin after a 48-hour incubation (Figure 7). This dose-dependent activity is more notable in Hsp90 drugs (NMS-E973,17-DMAG) than in PI3K/AKT pathway (Pictilisib) ones and very modest for PI-103.

We studied the real-time response of LNCaP cell viability upon drug addition and saw that the LNCaP cell line is sensitive to Hsp90 and PI3K/AKT pathway inhibitors (Figure 8 and 9, respectively). Both Hsp90 inhibitors tested, 17-DMAG and NMS-E973, reduced the cell viability 12 hours after drug supplementation (Figure 8A for 17-DMAG and Figure 8B for NMS-E973), with 17-DMAG having a stronger effect and in a more clear concentration-dependent manner than NMS-E973 (Appendix 1, Section 8, Appendix figure 40, panels B-D for 17-DMAG and panels F-H for NMS-E973).

Likewise, both PI3K/AKT pathway inhibitors tested, Pictilisib and PI-103, reduced the cell viability immediately after drug supplementation (Figure 9A for Pictilisib and Figure 9B for PI-103), in a concentration-dependent manner (Appendix 1, Section 8, Appendix figure 41, panels B-D for Pictilisib and panels F-H for PI-103). In addition, Hsp90 inhibitors had a more prolonged effect on the cells’ proliferation than PI3K/AKT pathway inhibitors.

Discussion

Clinical assessment of cancers is moving towards more precise, personalised treatments, as the times of one-size-fits-all treatments are no longer appropriate, and patient-tailored models could boost the success rate of these treatments in clinical practice. In this study, we set out to develop a methodology to investigate drug treatments using personalised Boolean models. Our approach consists of building a model that represents the patient-specific disease status and retrieving a list of proposed interventions that affect this disease status, notably by reducing its pro-cancerous behaviours. In this work, we have showcased this methodology by applying it to TCGA prostate cancer patients and to GDSC prostate cancer cell lines, finding patient- and cell-line-specific targets and validating selected cell-line-specific predicted targets (Figure 1).

First, a prostate cancer Boolean model that encompasses relevant signalling pathways in cancer was constructed based on already published models, experimental data analyses and pathway databases (Figure 2). The influence network and the assignment of logical rules for each node of this network were obtained from known interactions described in the
literature (Figure 3). This model describes the regulation of invasion, migration, cell cycle, apoptosis, androgen and growth factors signalling in prostate cancer (Appendix File, Section 1).

Second, from this generic Boolean model, we constructed personalised models using the different datasets, i.e. 488 patients from TCGA and eight cell lines from GDSC. We obtained Gleason-score-specific behaviours for TCGA’s patients when studying their Proliferation and Apoptosis scores, observing that high Proliferation scores are higher in high Gleason grades (Figure 4). Thus, the use of these personalised models can help rationalise the relationship of Gleason grading with some of these phenotypes.

Likewise, GDSC data was used with the prostate model to obtain prostate-specific cell-line models (Figure 6). These models show differential behaviours, notably in terms of Invasion and Proliferation phenotypes (Appendix 1, Section 5, Appendix figure 21). One of these cell-line-specific models, LNCaP, was chosen, and the effects of all its genetic perturbations were thoroughly studied. We studied 32258 mutants, including single and double mutants, knock-out and over-expressed, and their phenotypes (Appendix 1, Section 6.1, Appendix figures 28 and 29). 32 knock-out perturbations that depleted Proliferation and/or increased Apoptosis were identified, and 16 of them were selected for further analyses (Table 1). The LNCaP-specific model was simulated using different initial conditions that capture different growth media’s specificities, such as RPMI media with and without androgen or epidermal growth factor (Appendix 1, Section 6, Appendix figure 27).

Third, these personalised models were used to simulate the inhibition of druggable genes and proteins, uncovering new treatment’s combination and their synergies. We developed a methodology to simulate drug inhibitions in Boolean models, termed PROFILE_v2, as an extension of previous works (Béal et al., 2019). The LNCaP-specific model was used to obtain simulations with nodes and pairs of nodes corresponding to the genes of interest inhibited with varying strengths. This study allowed us to compile a list of potential targets (Table 1) and to identify potential synergies among genes in the model (Figure 5). Some of the drugs that targeted these genes, such as AKT and TERT, were identified in GDSC as having more sensitivity in LNCaP than in the rest of the prostate cancer cell lines (Figure 6). In addition, drugs that targeted genes included in the model allowed the identification of cell line specificities (Appendix 1, Section 5).

Fourth, we validated the effect of Hsp90 and PI3K/AKT pathway inhibitors on the LNCaP cell line experimentally, finding a concentration-dependent inhibition of the cell line viability as predicted, confirming the role of the drugs targeting these proteins in reducing LNCaP’s proliferation (Figure 7 and 8). Notably, these targets have been studied in other works on prostate cancer (Chen et al., 2019; Le et al., 2017).

The study presented here enables the study of drug combinations and their synergies. One reason for searching for combinations of drugs is that these have been described for allowing the use of lower doses of each of the two drugs reducing their toxicity (Bayat Mokhtari et al., 2017), evading compensatory mechanisms and combating drug resistances (Al-Lazikani et al., 2012; Krzyszczyk et al., 2018).

Even if this approach is attractive and promising, it has some limitations. The scope of present work is to test this methodology on a prostate model and infer patient-specific
prostate cancer treatments. The method need to be adapted if it were to be expanded to
study other cancers by using other models and target lists. The analyses performed with the
mathematical model do not aim to predict drug dosages per se but to help in the
identification of potential candidates. The patient-specific changes in Proliferation and
Apoptosis scores upon mutation are maximal theoretical yields that are used to rank the
different potential treatments and should not be used as a direct target for experimental
results or clinical trials. Our methodology suggests treatments for individual patients, but the
obtained results vary greatly from patient to patient, which is not an uncommon issue of
personalised medicine (Ciccarese et al., 2017; Molinari et al., 2018). This variability is an
economic challenge for labs and companies to pursue true patient-specific treatments and
also poses challenges in clinical trial designs aimed at validating the model based on the
selection of treatments (Cunanan et al., 2017). Nowadays, and because of these constraints,
it might be more commercially interesting to target group-specific treatments, which can be
more easily related to clinical stages of the disease.

Mathematical modelling of patient profiles helps to classify them in groups with differential
characteristics, providing, in essence, a grade-specific treatment. We, therefore, based our
analysis on clinical grouping defined by the Gleason grades, but some works have
emphasised the difficulty to properly assess them (Chen and Zhou, 2016) and, as a result,
may not be the perfect predictor for the patient subgrouping in this analysis, even though it is
the only available one for these datasets. The lack of subgrouping that stratifies patients
adequately may undermine the analysis of our results and could explain the Proliferation and
Apoptosis scores of high-grade and low-grade Gleason patients.

Moreover, the behaviours observed in the simulations of the cell-lines-specific models do not
always correspond to what is reported in the literature. The differences between simulation
results and biological characteristics could be addressed in further studies by including other
pathways, for example, better describing the DNA repair mechanisms, or by tailoring the
model with different sets of data, as the data used to personalise these models do not allow
for clustering these cell lines according to their different characteristics (Appendix 1, Section
5, Appendix figure 24 and 25). In this sense, another limitation is that we use static data (or a
snapshot of dynamic data) to build dynamic models and to study its stochastic results. Thus,
these personalised models would likely improve their performance if they were fitted to
dynamic data (Saez-Rodriguez and Blüthgen, 2020) or quantitative versions of the models
were built, such as ODE-based, that may capture more fine differences among cell lines. As
perspectives, we are working on integrating these models in multiscale models to study the
effect of the tumour microenvironment (Ponce-de-Leon et al., 2021, 2022), on including
information to simulate multiple reagents targeting a single node of the model, on scaling
these multiscale models to exascale high-performance computing clusters (Montagud et al.,
2021; Saxena et al., 2021), and on streamlining these studies using workflows in computing
clusters to fasten the processing of new, bigger cohorts, as in the PerMedCoE project
(https://permedcoe.eu/).

The present work contributes to efforts aimed at using modelling (Eduati et al., 2020; Rivas-
Barragan et al., 2020; Gómez Tejeda Zañudo et al., 2017) and other computational methods
(Madani Tonekaboni et al., 2018; Menden et al., 2019) for the discovery of novel drug
targets and combinatorial strategies. Our study expands the prostate drug catalogue and
improves predictions of the impact of these in clinical strategies for prostate cancer by
proposing and grading the effectiveness of a set of drugs that could be used off-label or
repurposed. The insights gained from this study present the potential of using personalised models to obtain precise, personalised drug treatments for cancer patients.
Materials and Methods

Data acquisition

Publicly available data of 489 human prostate cancer patients from TCGA described in (Hoadley et al., 2018) were used in the present work. We gathered mutations, CNA, RNA and clinical data from cBioPortal (https://www.cbioportal.org/study/summary?id=prad_tca_g_pan_can_atlas_2018) for all of these samples resulting in 488 with complete omics datasets.

Publicly available data of cell lines used in the present work were obtained from the Genomics of Drug Sensitivity in Cancer database (GDSC) (Iorio et al., 2016). Mutations, CNA and RNA data, as well as cell lines descriptors, were downloaded from (https://www.cancerrxgene.org/downloads). In this work, we have used 3- and 5-stage Gleason grades. Their correspondence is the following: GG Low is GG 1, GG Intermediate is GG 2 and 3, and GG High is GG 4 and 5.

All these data were used to personalise Boolean models using our PROFILE method (Béal et al., 2019).

Prior knowledge network construction

Several sources were used in building this prostate Boolean model and, in particular, the model published by Fumià and Martins (2013). This model includes several signalling pathways such as the ones involving receptor tyrosine kinase (RTKs), phosphatidylinositol 3-kinase (PI3K)/AKT, WNT/b-Catenin, transforming growth factor-b (TGF-b)/Smads, cyclins, retinoblastoma protein (Rb), hypoxia-inducible transcription factor (HIF-1), p53 and ataxia-telangiectasia mutated (ATM)/ataxia-telangiectasia and Rad3-related (ATR) protein kinases. The model includes these pathways as well as the substantial cross-talks among them. For a complete description of the process of construction, see Appendix 1, Section 1.

The model also includes several pathways that have a relevant role in our datasets identified by ROMA (Martignetti et al., 2016), a software that uses the first principal component of a PCA analysis to summarise the coexpression of a group of genes in the gene set, identifying significantly overdispersed pathways with a relevant role in a given set of samples. This software was applied to the TCGA transcriptomics data using the gene sets described in the Atlas of Cancer Signaling Networks, ACSN (Kuperstein et al., 2015) (www.acsn.curie.fr) and in Hallmarks (Liberzon et al., 2015) (Appendix 1, Section 1.1.3, Appendix figure 1) and highlighted the signalling pathways that show high variance across all samples, suggesting candidate pathways and genes. Additionally, OmniPath (Türei et al., 2021) was used to extend the model and complete it, connecting the nodes from Fumià and Martins and the ones from ROMA analysis. OmniPath is a comprehensive collection of literature-curated human signalling pathways, which includes several databases such as Signor (Perfetto et al., 2016) or Reactome (Fabregat et al., 2016) and that can be queried using pypath, a Python module for molecular networks and pathways analyses.
Fusion genes are frequently found in human prostate cancer and have been identified as a specific subtype marker (The Cancer Genome Atlas Research Network, 2015). The most frequent is TMPRSS2:ERG, as it involves the transcription factor ERG, which leads to cell-cycle progression. ERG fuses with the AR-regulated TMPRSS2 gene promoter to form an oncogenic fusion gene that is especially common in hormone-refractory prostate cancer, conferring androgen responsiveness to ERG. A literature search reveals that ERG directly regulates EZH2, oncogene c-Myc and many other targets in prostate cancer (Kunderfranco et al., 2010).

We modelled the gene fusion with activation of ERG by the decoupling of ERG in a special node \(AR_{\text{ERG}}\) that is only activated by the \(AR\) when the \(\text{fused}_\text{event}\) input node is active. In the healthy case, \(\text{fused}_\text{event}\) (that represents TMPRSS2:ERG fusion event) is fixed to 0 or inactive. The occurrence of the gene fusion is represented with the model perturbation where \(\text{fused}_\text{event}\) is fixed to 1. This \(AR_{\text{ERG}}\) node is further controlled by tumour suppressor NKX3-1 that accelerates DNA repair response, and avoids the gene fusion TMPRSS2:ERG. Thus, loss of NKX3-1 favours recruitment to the ERG gene breakpoint of proteins that promote error-prone non-homologous end-joining (Bowen et al., 2015).

The network was further documented using up-to-date literature and was constructed using GINsim (Chaouiya et al., 2012), which allowed us to study its stable states and network properties.

**Boolean model construction**

We converted the network to a Boolean model by defining a regulatory graph, where each node is associated with discrete levels of activity (0 or 1). Each edge represents a regulatory interaction between the source and target nodes and is labelled with a threshold and a sign (positive or negative). The model is completed by logical rules (or functions), which assign a target value to each node for each regulator level combination (Abou-Jaoudé et al., 2016; Chaouiya et al., 2012). The regulatory graph was constructed using GINsim software (Chaouiya et al., 2012) and then exported in a format readable by MaBoSS software (see below) in order to perform stochastic simulations on the Boolean model.

The final model has a total of 133 nodes and 449 edges (Supplementary File 1) and includes pathways such as androgen receptor and growth factor signalling, several signalling pathways (Wnt, NFkB, PI3K/AKT, MAPK, mTOR, SHH), cell cycle, epithelial-mesenchymal transition (EMT), Apoptosis, DNA damage, etc. This model has 9 inputs (\(EGF, FGF, TGF\beta, Nutrients, Hypoxia, Acidosis, Androgen, TNF\alpha\) and Carcinogen presence) and 6 outputs (\(\text{Proliferation, Apoptosis, Invasion, Migration, (bone) Metastasis and DNA repair}\)). Note that a node in the network can represent complexes or families of proteins (e.g., AMPK represents the genes PRKAA1, PRKAA2, PRKAB1, PRKAB2, PRKAG1, PRKAG2, PRKAG3). The correspondence can be found in “Montagud2021_interactions_sources.xlsx” and “Montagud2021_nodes_in_pathways.xlsx” in Supplementary File 1.

This model was deposited in the GINsim Database with identifier 252 (http://ginsim.org/model/signalling-prostate-cancer) and in BioModels (Malik-Sheriff et al., 2019) with identifier MODEL2106070001 (https://www.ebi.ac.uk/biomodels/MODEL2106070001). Supplementary File 1 is provided as
a zipped folder with the model in several formats: MaBoSS, GINsim, SBML, as well as images of the networks and their annotations. An extensive description of the model construction can be found in the Appendix 1, Section 1.

Stochastic Boolean model simulation

MaBoSS (Stoll et al., 2017, 2012) is a C++ software for stochastically simulating continuous/discrete-time Markov processes defined on the state transition graph (STG) describing the dynamics of a Boolean model (for more details, see (Abou-Jaoudé et al., 2016; Chaouiya et al., 2012)). MaBoSS associates transition rates to each node's activation and inhibition, enabling it to account for different time scales of the processes described by the model. Probabilities to reach a phenotype (to have value ON) are thus computed by simulating random walks on the probabilistic STG. Since a state in the STG can combine the activation of several phenotypic variables, not all phenotype probabilities are mutually exclusive (like the ones in Appendix 1, Section 6.1, Appendix figure 28). Using MaBoSS, we can study an increase or decrease of a phenotype probability when the model variables are altered (nodes status, initial conditions and transition rates), which may correspond to the effect of particular genetic or environmental perturbation. In the present work, the outputs of MaBoSS focused on the readouts of the model, but this can be done for any node of a model.

MaBoSS applies Monte-Carlo kinetic algorithm (i.e. Gillespie algorithm) to the STG to produce time trajectories (Stoll et al., 2017, 2012), so time evolution of probabilities are estimated once a set of initial conditions are defined and a maximum time is set to ensure that the simulations reach asymptotic solutions. Results are analysed in two ways: (1) the trajectories for particular model states (states of nodes) can be interpreted as the evolution of a cell population as a function of time, and (2) asymptotic solutions can be represented as pie charts to illustrate the proportions of cells in particular model states. Stochastic simulations with MaBoSS have already been successfully applied to study several Boolean models (Calzone et al., 2010; Cohen et al., 2015; Remy et al., 2015). A description of the methods we have used for the simulation of the model can be found in the Appendix 1, Section 2.

Data tailoring the Boolean model

Logical models were tailored to a dataset using PROFILE to obtain personalised models that capture the particularities of a set of patients (Béal et al., 2019) and cell lines (Béal et al., 2021). Proteomics, transcriptomics, mutations and CNA data can be used to modify different variables of the MaBoSS framework, such as node activity status, transition rates and initial conditions. The resulting ensemble of models is a set of personalised variants of the original model that can show great phenotypic differences. Different recipes (use of a given data type to modify a given MaBoSS variable) can be tested to find the combination that better correlates to a given clinical or otherwise descriptive data.

In the present case, TCGA-patient-specific models were built using mutations, CNA and/or RNA expression data. After studying the effect of these recipes in the clustering of patients
according to their Gleason grouping (Appendix 1, Section 4.1, Appendix figure 10-14), we chose to use mutations and CNA as discrete data and RNA expression as continuous data. Likewise, we tried different personalisation recipes to personalise the GDSC prostate cell lines models, but as they had no associated clinical grouping features, we were left with the comparison of the different values for the model’s outputs among the recipes (Appendix 1, Section 5, Appendix figure 23). We used mutation data as discrete data and RNA expression as continuous data as it included the most quantity of data and reproduced the desired results (Appendix 1, Section 5, Appendix figure 23). We decided not to include CNA as discrete data as it forced LNCaP proliferation to be zero by forcing the E2F1 node to be 0 and the SMAD node to be 1 throughout the simulation (for more details, refer to Appendix 1, Section 5).

More on PROFILE’s methodology can be found in its own work (Béal et al., 2019) and at its dedicated GitHub repository: https://github.com/sysbio-curie/PROFILE. A description of the methods we have used for the personalisation of the models can be found in the Appendix 1, Section 3. The analysis of the TCGA personalisations and their patient-specific drug treatments can be found in Appendix 1, Section 4. The analysis of the prostate cell lines personalisations can be found in Appendix 1, Section 5, with a special focus on the LNCaP cell line model analysis in Section 6.

High-throughput mutant analysis of Boolean models

MaBoSS allows the study of knock-out or loss-of-function (node forced to 0) and gain-of-function (node forced to 1) mutants as genetic perturbations and of initial conditions as environmental perturbations. Phenotypes’ stabilities against perturbations can be studied and allow to determine driver mutations that promote phenotypic transitions (Montagud et al., 2017).

Genetic interactions were thoroughly studied using our pipeline of computational methods for Boolean modelling of biological networks (available at https://github.com/sysbio-curie/Logical_modelling_pipeline). LNCaP-specific Boolean model was used to perform single and double knock-out (node forced to 0) and gain-of-function (node forced to 1) mutants for each one of the 133 nodes, resulting in a total of 32258 models. These were simulated under the same initial conditions, their phenotypic results were collected, and a PCA was applied on the wild-type-centred matrix (Appendix 1, Section 6.1, Appendix figure 28 and 29). In addition, we found that the LNCaP model is very robust against perturbations of its logical rules by systematically changing an AND for an OR gate or vice versa in all of its logical rules (Appendix 1, Section 6.2, Appendix figure 30 and 31).

The 488 TCGA-patient-specific models were studied in a similar way, but only perturbing 16 nodes shortlisted for their therapeutic target potential (AKT, AR, Caspase8, cFLAR, EGFR, ERK, GLUT1, HIF-1, HSPs, MEK1_2, MYC_MAX, p14ARF, PI3K, ROS, SPOP and TERT). Then, the nodes that mostly contributed to a decrease of Proliferation (Appendix 1, Section 4.2, Appendix figure 19) or an increase in Apoptosis (Appendix 1, Section 4.2, Appendix figure 20) were gathered from the 488 models perturbed. Additionally, the results of the LNCaP model’s double mutants were used to quantify the level of genetic interactions (epistasis or otherwise (Drees et al., 2005)) between two model
genetic perturbations (resulting from either the gain-of-function mutation of a gene or from its knock-out or loss-of-function mutation) with respect to wild type phenotypes’ probabilities (Calzone et al., 2015). The method was applied to the LNCaP model studying Proliferation and Apoptosis scores (Appendix 1, Section 7.3.2, Appendix figure 34 and 35).

This genetic interaction study uses the following equation for each gene pair, which is equation 2 in Calzone et al., (2015):

\[
\epsilon_\phi(A, B) = f_\phi^{AB} - \psi(f_\phi^A, f_\phi^B)
\]  

(1)

Where \( f_\phi^A \) and \( f_\phi^B \) are phenotype \( \phi \) fitness values of single gene defects, \( f_\phi^{AB} \) is the phenotype \( \phi \) fitness of the double mutant, and \( \psi(x, y) \) is one of the four functions:

\[
\psi^{ADD}(x, y) = x + y \quad \text{ (additive)}
\]
\[
\psi^{LOG}(x, y) = \log_2((2^x - 1)(2^y - 1) + 1) \quad \text{ (log)}
\]
\[
\psi^{MLT}(x, y) = x \times y \quad \text{ (multiplicative)}
\]
\[
\psi^{MIN}(x, y) = \min(x, y) \quad \text{ (min)}
\]

(2)

To choose the best definition of \( \psi(x, y) \), the Pearson correlation coefficient is computed between the fitness values observed in all double mutants and estimated by the null model (more information on (Drees et al., 2005)). Regarding \( f_\phi \) fitness value, to a given phenotype \( \phi \), \( f_\phi <1 \) represents deleterious, \( f_\phi >1 \) beneficial and \( f_\phi \approx1 \) neutral mutation.

Drug simulations in Boolean models

Logical models can be used to simulate the effect of therapeutic interventions and predict the expected efficacy of candidate drugs on different genetic and environmental backgrounds by using our PROFILE_v2 methodology. MaBoSS can perform simulations changing the proportion of activated and inhibited status of a given node. This can be determined in the configuration file of each model (see, for instance, the “istate” section of the CFG files in the Supplementary File 1, 3 and 5). For instance, out of 5000 trajectories of the Gillespie algorithm, MaBoSS can simulate 70% of them with an activated AKT and 30% with an inhibited AKT node. The phenotypes’ probabilities for the 5000 trajectories are averaged, and these are considered to be representative of a model with a drug that inhibits 30% of the activity of AKT. The same applies for a combined drug inhibition: a simulation of 50% AKT activity and 50% PI3K will have 50% of them with an activated AKT and 50% with an activated PI3K. Combining them, this will lead to 25% of the trajectories with both AKT and PI3K active, 25% with both nodes inactive, 25% with AKT active and 25% with PI3K active.

In the present work, the LNCaP model has been simulated with different levels of node activity, with 100% of node activity (no inhibition), 80%, 60%, 40%, 20% and 0% (proper knock-out), under four different initial conditions, a nutrient-rich media that simulates RPMI Gibco® media with DHT (androgen), with EGF, with both and with none. In terms of the model, the initial conditions are Nutrients is ON and Acidosis, Hypoxia, TGF beta,
Carcinogen and TNF alpha are set to OFF. EGF and Androgen values vary upon simulations. We simulated the inhibition of 17 nodes of interest. These were the 16 nodes from Table 1 with the addition of the fused AR-ERG (Appendix 1, Section 7.3.1, Appendix figures 34 and 35) and their 136 pairwise combinations (Appendix 1, Section 7.3.2, Appendix figures 36 and 37). As we used 6 different levels of activity for each node, the resulting Appendix figures 36 and 37 comprise a total of 4998 simulations for each phenotype (136 x 6 x 6 + 17 x 6).

Drug synergies have been studied using Bliss Independence. The Combination Index was calculated with the following equation (Foucquier and Guedj, 2015):

\[
CI = \frac{(E_a + E_b - E_a \cdot E_b)}{E_{ab}}
\]

Where \( E_a \) and \( E_b \) is the efficiency of the single drug inhibitions and \( E_{ab} \) is the inhibition resulting from the double drug simulations. A Combination Index (CI) below 1 represents synergy among drugs (Appendix 1, Section 7.3.2, Appendix figures 36 and 37).

This methodology can be found in its own repository: https://github.com/ArnauMontagud/PROFILE_v2

Identification of drugs associated with proposed targets

To identify drugs that could act as potential inhibitors of the genes identified with our models (Table 1), we explored the drug-target associations in DrugBank (Wishart et al., 2018). For those genes with multiple drug-target links, only those drugs that are selective and known to have relevance in various forms of cancer are considered here.

In addition to DrugBank searches, we also conducted exhaustive searches in ChEMBL (Gaulton et al., 2017) (http://doi.org/10.6019/CHEMBL.database.23) to suggest potential candidates for genes whose information is not well documented in Drug Bank. From the large number of bioactivities extracted from ChEMBL, we filtered human data and considered only those compounds whose bioactivities fall within a specific threshold (IC50/Kd/ Ki<100 nM).

We performed a target set enrichment analysis using the fgsea method (Korotkevich et al., 2016) from the piano R package (Väremo et al., 2013). We targeted pathway information from the GDSC1 and GDSC2 studies (Iorio et al., 2016) as target sets and performed the enrichment analysis on the normalised drug sensitivity profile of the LNCaP cell line. We normalised drug sensitivity across cell lines in the following way: cells were ranked from most sensitive to least sensitive (using ln(IC50) as drug sensitivity metrics), and the rank was divided by the number of cell lines tested with the given drug. Thus, the most sensitive cell line has 0, while the most resistant cell line has 1 normalised sensitivity. This rank-based metric made it possible to analyse all drug sensitivities for a given cell line without drug-specific confounding factors, like mean IC50 of a given drug, etc. (Appendix 1, Section 7.1 and 7.2).

Cell culture method

For the in vitro drug perturbation validations, we used the androgen-sensitive prostate adenocarcinoma cell line LNCaP purchased from American Type Culture Collection (ATCC,
Manassas, WV, USA). ATCC found no Mycoplasma contamination and the cell line was identified using STR profiling. Cells were maintained in RPMI-1640 culture media (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) containing 4.5 g/L glucose, 10% foetal bovine serum (FBS, Gibco), 1X GlutaMAX (Gibco), 1% PenStrep antibiotics (Penicillin G sodium salt, and Streptomycin sulfate salt, Sigma-Aldrich, St. Louis, MI, USA). Cells were maintained in a humidified incubator at 37 °C with 5% CO₂ (Sanyo, Osaka, Japan).

Drugs used in the cell culture experiments

We tested two drugs targeted at Hsp90 and two targeted at PI3K complex. 17-DMAG is an Hsp90 inhibitor with an IC50 of 62 nM in a cell-free assay (Pacey et al., 2011). NMS-E973 is an Hsp90 inhibitor with DC50 of <10 nM for Hsp90 binding (Fogliatto et al., 2013). Pictilisib is an inhibitor of PI3Kα/δ with IC50 of 3.3 nM in cell-free assays (Zhan et al., 2017). PI-103 is a multi-targeted PI3K inhibitor for p110α/β/δ/γ with IC50 of 2 to 3 nM in cell-free assays and less potent inhibitor to mTOR/DNA-PK with IC50 of 30 nM (Raynaud et al., 2009). All drugs were obtained from commercial vendors and added to the growth media to have concentrations of 2, 8, 32, 128 and 512 nM for NMS-E973 and 1, 5, 25, 125 and 625 nM for the rest of the drugs in the endpoint cell viability and of 3.3, 10, 30 µM for all the drugs in the RT-CES cytotoxicity assay.

Endpoint cell viability measurements

In vitro toxicity of the selected inhibitors was determined using the viability of LNCaP cells, determined by the fluorescent resazurin (Sigma-Aldrich, Germany) assay as described previously (Szebeni et al., 2017). Briefly, the LNCaP cells (10000) were seeded into 96-well plates (Corning Life Sciences, Tewksbury, MA, USA) in 100 µl RPMI media and incubated overnight. Test compounds were dissolved in dimethyl sulfoxide (DMSO, Sigma-Aldrich, Germany), and cells were treated with an increasing concentration of test compounds: 2, 8, 32, 128 and 512 nM for NMS-E973 and 1, 5, 25, 125 and 625 nM for the rest of the drugs. The highest applied DMSO content of the treated cells was 0.4%. Cell viability was determined after 48 hours of incubation. Resazurin reagent (Sigma–Aldrich, Budapest, Hungary) was added at a final concentration of 25 µg/mL. After 2 hours at 37°C 5%, CO₂ (Sanyo) fluorescence (530 nm excitation/580 nm emission) was recorded on a multimode microplate reader (Cytofluor4000, PerSeptive Biosystems, Framingham, MA, USA). Viability was calculated with relation to blank wells containing media without cells and to wells with untreated cells. Each treatment was repeated in 2 wells per plate during the experiments, except for the PI-103 treatment with 1 nM in which only one well was used.

In these assays, a deviation of 10-15% for in vitro cellular assays is an acceptable variation as it is a fluorescent assay that detects the cellular metabolic activity of living cells. Thus, in our analyses, we consider changes above 1.00 to be the same value as the controls.

Real-time cell electronic sensing (RT-CES) cytotoxicity assay

A real-time cytotoxicity assay was performed as previously described (Ozsvári et al., 2010). Briefly, RT-CES 96-well E-plate (BioTech Hungary, Budapest, Hungary) was coated with gelatin solution (0.2% in PBS, phosphate buffer saline) for 20 min at 37 °C; then gelatin was
washed twice with PBS solution. Growth media (50 μL) was then gently dispensed into each well of the 96-well E-plate for background readings by the RT-CES system prior to the addition of 50 μL of the cell suspension containing 2x10^4 LNCaP cells. Plates were kept at room temperature in a tissue culture hood for 30 min prior to insertion into the RT-CES device in the incubator to allow cells to settle. Cell growth was monitored overnight by measurements of electrical impedance every 15 min. The next day cells were co-treated with different drugs with concentrations of 3.3, 10 and 30 μM. Treated and control wells were dynamically monitored over 72 h by measurements of electrical impedance every 5 min. Each treatment was repeated in 2 wells per plate during the experiments, except for the 3.3 μM ones in which only one well was used. Continuous recording of impedance in cells was used as a measurement of the cell growth rate and reflected by the Cell Index value (Solly et al., 2004).

Note that around hour 15, our RT-CES reader had a technical problem caused by a short blackout in our laboratory and the reader detected a minor voltage fluctuation while the uninterruptible power supply (UPS) was switched on. This caused differences that are consistent across all samples and replicates: all wild type and drug reads decrease at that time point, except Pictilisib that slightly increases. For the sake of transparency and as the overall dynamic was not affected, we decided to not remove these readings.
Acknowledgements

The authors acknowledge the help provided by Jelena Čuklina at ETH Zurich, Vincent Noël at Institut Curie, Annika Meert at Barcelona Supercomputing Center and Aurélien Naldi at INRIA Saclay. XXX thank reviewers

The authors acknowledge the technical expertise and assistance provided by the Spanish Supercomputing Network (Red Española de Supercomputación), as well as the computer resources used: the LaPalma Supercomputer, located at the Instituto de Astrofísica de Canarias and MareNostrum4, located at the Barcelona Supercomputing Center.

This work has been partially supported by the European Commission under the PrECISE project (H2020-PHC-668858), the INFORE project (H2020-ICT-825070) and the PerMedCoE (H2020-ICT-951773).


doi:10.1016/j.ctrv.2017.02.001


doi:10.1038/s41388-019-0815-5


doi:10.3892/or.2012.1813


doi:10.15252/msb.20188664

doi:10.1093/nar/gkv1351

Hessenkemper W, Baniahmad A. 2013. Targeting Heat Shock Proteins in
Hayward SW, Dahiya R, Cunha GR, Bartek J, Deshpande N, N
Grieco L, Calzone L, Bernard
Gómez
doi:10.1016/0021-9991(76)90041-3
Gleason DF. 1977. The Veteran's Administration Cooperative Urologic Research Grou
Gleason DF. 1977. The Veteran’s Administration Cooperative Urologic Research Group:
histologic grading and clinical staging of prostatic carcinomaUrologic Pathology: The
doi:10.1038/nature01259
Gómez Tejeda Zañudo J, Scaltriti M, Albert R. 2017. A network modeling approach to
elucidate drug resistance mechanisms and predict combinatorial drug treatments in
Integrative modelling of the influence of MAPK network on cancer cell fate decision.
Establishment and characterization of an immortalized but non-transformed human
doi:10.1007/BF02631333
Helikar T, Konvalina J, Heidel J, Rogers JA. 2008. Emergent decision-making in biological
doi:10.1073/pnas.070508105
Weinstein JN, Zhang J (Julia), Chudamani S, Liu
I de, Gao J, Gross BE, Heins ZJ, Kundra R, La K, Ladanyi M, Luna A, Nissan MG,
1029 Linehan WM, Metwalli AR, Pacak K, Pinto PA, Schiffman M, Schmidt LS, Vocke CD,
1030 Wentzensen N, Worrell R, Yang H, Moncrieff M, Goparaju C, Melamed J, Pass H,
1031 Botnar Iuc N, Caraman I, Cernat M, Chemencedij I, Clipca A, Doruc S, Gorincioi G,
1032 Mura S, Pirtac M, Stancul I, Tcaciu D, Albert M, Alexopoulou I, Arnaout A, Bartlett J,
1033 Engel J, Gilbert S, Parfitt J, Sekhon H, Thomas G, Rassl DM, Rintoul RC, Bifulco C,
1034 Tamakawa R, Urba W, Hayward N, Timmers H, Antenucci A, Facchiolo F, Grazi G,
1035 Marino M, Merola R, Krijger R de, Gimenez-Roqueplo A-P, Piché A, Chevalier S,
1037 Pool M, Seder CW, Junker K, Omerberg L, Dinkin M, Manikhas G, Alvaro D, Bragazzi
1038 Abdel-Rahman MH, Aziz D, Bell S, Cebulla CM, Davis A, Duell R, Elder JB, Eldad Rham
1039 Horoszewicz JS, Leong SS, Kawinski E, Karr JP, Rosenthal H, Chu TM, Mirand EA, Murphy
doi:10.1038/srep17280
1041 Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves
doi:10.1038/srep17280
1044 Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves
1045 E, Barthorpe S, Lightfoot H, Cokelaer T, Greninger P, van Dyk E, Chang H, de Silva
1047 Botnar Iuc N, Caraman I, Cernat M, Chemencedij I, Clipca A, Doruc S, Gorincioi G,
1048 Mura S, Pirtac M, Stancul I, Tcaciu D, Albert M, Alexopoulou I, Arnaout A, Bartlett J,
1050 Tamakawa R, Urba W, Hayward N, Timmers H, Antenucci A, Facchiolo F, Grazi G,
1051 Marino M, Merola R, Krijger R de, Gimenez-Roqueplo A-P, Piché A, Chevalier S,
1053 Pool M, Seder CW, Junker K, Omerberg L, Dinkin M, Manikhas G, Alvaro D, Bragazzi
1054 Abdel-Rahman MH, Aziz D, Bell S, Cebulla CM, Davis A, Duell R, Elder JB, Eldad Rham
1055 Horoszewicz JS, Leong SS, Kawinski E, Karr JP, Rosenthal H, Chu TM, Mirand EA, Murphy
doi:10.1038/srep17280
1057 Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves
1058 E, Barthorpe S, Lightfoot H, Cokelaer T, Greninger P, van Dyk E, Chang H, de Silva
1060 Botnar Iuc N, Caraman I, Cernat M, Chemencedij I, Clipca A, Doruc S, Gorincioi G,
1061 Mura S, Pirtac M, Stancul I, Tcaciu D, Albert M, Alexopoulou I, Arnaout A, Bartlett J,
1063 Tamakawa R, Urba W, Hayward N, Timmers H, Antenucci A, Facchiolo F, Grazi G,
1064 Marino M, Merola R, Krijger R de, Gimenez-Roqueplo A-P, Piché A, Chevalier S,
1066 Pool M, Seder CW, Junker K, Omerberg L, Dinkin M, Manikhas G, Alvaro D, Bragazzi
1067 Abdel-Rahman MH, Aziz D, Bell S, Cebulla CM, Davis A, Duell R, Elder JB, Eldad Rham
1068 Horoszewicz JS, Leong SS, Kawinski E, Karr JP, Rosenthal H, Chu TM, Mirand EA, Murphy
doi:10.1038/srep17280
1070 Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves
1071 E, Barthorpe S, Lightfoot H, Cokelaer T, Greninger P, van Dyk E, Chang H, de Silva
1073 Botnar Iuc N, Caraman I, Cernat M, Chemencedij I, Clipca A, Doruc S, Gorincioi G,
1074 Mura S, Pirtac M, Stancul I, Tcaciu D, Albert M, Alexopoulou I, Arnaout A, Bartlett J,
1075 Engel J, Gilbert S, Parfitt J, Sekhon H, Thomas G, Rassl DM, Rintoul RC, Bifulco C,
1076 Tamakawa R, Urba W, Hayward N, Timmers H, Antenucci A, Facchiolo F, Grazi G,
1077 Marino M, Merola R, Krijger R de, Gimenez-Roqueplo A-P, Piché A, Chevalier S,
1079 Pool M, Seder CW, Junker K, Omerberg L, Dinkin M, Manikhas G, Alvaro D, Bragazzi
1080 Abdel-Rahman MH, Aziz D, Bell S, Cebulla CM, Davis A, Duell R, Elder JB, Eldad Rham
1081 Horoszewicz JS, Leong SS, Kawinski E, Karr JP, Rosenthal H, Chu TM, Mirand EA, Murphy
doi:10.1038/srep17280
1083 Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves
1086 Botnar Iuc N, Caraman I, Cernat M, Chemencedij I, Clipca A, Doruc S, Gorincioi G,


### Table 1: List of selected nodes, their corresponding genes and drugs that were included in the drug analysis of the models tailored for TCGA patients and LNCaP cell line.

<table>
<thead>
<tr>
<th>Node</th>
<th>Gene</th>
<th>Compound / Inhibitor name</th>
<th>Clinical stage</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKT</td>
<td>AKT1, AKT2, AKT3</td>
<td>PI-103</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Enzastaurin</td>
<td>Phase 3</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Archexin, Pictilisib</td>
<td>Phase 2</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>AR</td>
<td>AR</td>
<td>Abiraterone, Enzalutamide, Formestane, Testosterone propionate</td>
<td>Approved</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5alpha-androstan-3beta-ol</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>Caspase8</td>
<td>CASP8</td>
<td>Bardoxolone</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>cFLAR</td>
<td>CFLAR</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>EGFR</td>
<td>EGFR</td>
<td>Afatinib, Osimertinib, Neratinib, Erlotinib, Gefitinib</td>
<td>Approved</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varlitinib</td>
<td>Phase 3</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Olmutinib, Pelitinib</td>
<td>Phase 2</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>ERK</td>
<td>MAPK1</td>
<td>Isoprenaline</td>
<td>Approved</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Perifosine</td>
<td>Phase 3</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Turpentine, SB220025, Olomoucine</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>Pathway</td>
<td>Target</td>
<td>Drug</td>
<td>Phase</td>
<td>Status</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>MAPK3, MAPK1</td>
<td>MAPK3, MAPK1</td>
<td>Arsenic trioxide</td>
<td>Approved</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ulixertinib, Seliciclib</td>
<td>Phase 2</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Purvalanol</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>MAPK3</td>
<td>MAPK3</td>
<td>Sulindac, Cholecystokinin</td>
<td>Approved</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5-iodotubercidin</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>GLUT1</td>
<td>GLUT1</td>
<td>Resveratrol</td>
<td>Phase 4</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>HIF-1</td>
<td>HIF1A</td>
<td>CAY-10585</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>HSPs</td>
<td>HSP90AA1, HSP90AB1, HSP90B1, HSPA1A, HSPA1B, HSPB1</td>
<td>Cladribine</td>
<td>Approved</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>17-DMAG</td>
<td>Phase 2</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NMS-E973</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>MEK1_2</td>
<td>MEK1_2</td>
<td>Trametinib, Selumetinib</td>
<td>Approved</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Perifosine</td>
<td>Phase 3</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PD184352 (CI-1040)</td>
<td>Phase 2</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>MYC_MAX</td>
<td>complex of MYC and MAX</td>
<td>10058-F4 (for MAX)</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>p14ARF</td>
<td>CDKN2A</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PI3K</td>
<td>PIK3CA, PIK3CB, PIK3CG, PIK3CD,</td>
<td>PI-103</td>
<td>Preclinical</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pictilisib</td>
<td>Phase 2</td>
<td>Drug Bank</td>
</tr>
<tr>
<td>Gene</td>
<td>Protein(s)</td>
<td>Inhibitor(s)</td>
<td>Status</td>
<td>Database</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>PIK3R1, PIK3R2, PIK3R3, PIK3R4, PIK3R5, PIK3R6, PIK3C2A, PIK3C2B, PIK3C2G, PIK3C3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOX1, NOX3, NOX4</td>
<td>Fostamatinib</td>
<td>Approved</td>
<td>Drug Bank</td>
<td></td>
</tr>
<tr>
<td>NOX2</td>
<td>Dextromethorphan</td>
<td>Approved</td>
<td>Drug Bank</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tetrahydroisoquino lines</td>
<td>Preclinical</td>
<td>ChEMBL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(CHEMBL3733336,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHEMBL3347550,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CHEMBL3347551)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPOP</td>
<td>SPOP</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TERT</td>
<td>TERT</td>
<td>Grn163I</td>
<td>Phase 2</td>
<td>Drug Bank</td>
</tr>
<tr>
<td></td>
<td>BIBR 1532</td>
<td>Preclinical</td>
<td>ChEMBL</td>
<td></td>
</tr>
</tbody>
</table>
Figure titles and their legends:

Figure 1: Workflow to build patient-specific Boolean models and to uncover personalised drug treatments from present work. We gathered data from Fumiã and Martins (2013) Boolean model, Omnipath (Türei et al., 2021) and pathways identified with ROMA (Martignetti et al., 2016) on the TCGA data to build a prostate-specific prior knowledge network. This network was manually converted into a prostate Boolean model that could be stochastically simulated using MaBoSS (Stoll et al., 2017) and tailored to different TCGA and GDSC datasets using our PROFILE tool to have personalised Boolean models. Then, we studied all the possible single and double mutants on these tailored models using our logical pipeline of tools (Montagud et al., 2017). Using these personalised models and our PROFILE_v2 tool presented in this work, we obtained tailored drug simulations and drug treatments for 488 TCGA patients and eight prostate cell lines. Lastly, we performed drug-dose experiments on a shortlist of candidate drugs that were particularly interesting in the LNCaP prostate cell line. Created with BioRender.com.

Figure 2: Prostate Boolean model used in present work. Nodes (ellipses) represent biological entities, and arcs are positive (green) or negative (red) influences of one entity on another one. Orange rectangles correspond to inputs (from left to right: Epithelial Growth Factor (EGF), Fibroblast Growth Factor (FGF), Transforming Growth Factor beta (TGFbeta), Nutrients, Hypoxia, Acidosis, Androgen, fused_event, Tumour Necrosis Factor alpha (TNFalpha), SPOP, Carcinogen) and dark blue rectangles to outputs that represent biological phenotypes (from left to right: Proliferation, Migration, Invasion, Metastasis, Apoptosis, DNA_repair), the read-outs of the model. This network is available to be inspected as a Cytoscape file in the Supplementary File 1.

Figure 3: Prostate Boolean model MaBoSS simulations. (A) The model was simulated with all initial inputs set to 0 and all other variables random. All phenotypes are 0 at the end of the simulations, which should be understood as a quiescent state, where neither proliferation nor apoptosis is active. (B) The model was simulated with growth factors (EGF and FGF), Nutrients and Androgen ON. (C) The model was simulated with Carcinogen, Androgen, TNFalpha, Acidosis, and Hypoxia ON.

Figure 4: Associations between simulations and Gleason grades (GG). A) Centroids of the Principal Component Analysis of the samples according to their Gleason grades (GG). The personalisation recipe used was mutations and copy number alterations (CNA) as discrete data and RNAseq as continuous data. Density plots of Proliferation (B) and Apoptosis (C) scores according to GG; each vignette corresponds to a specific sub-cohort with a given GG. Kruskal-Wallis rank sum test across GG is significant for Proliferation (p-value=0.00207) and Apoptosis (p-value=2.83E-6).

Figure 5: Phenotype score variations and synergy upon combined ERK and MYC_MAX (A and C) and HSPs and PI3K (B and D) inhibition under EGF condition. Proliferation score variation (A) and Bliss Independence synergy score (C) with increased node activation of nodes ERK and MYC_MAX. Proliferation score variation (B) and Bliss Independence synergy score (D) with increased node activation of nodes HSPs and PI3K. Bliss Independence synergy score < 1 is characteristic of drug synergy, grey colour means one of the drugs is absent, and thus no synergy score is available.
Figure 6: Model-targeting drugs’ sensitivities across prostate cell lines. GDSC z-score was obtained for all the drugs targeting genes included in the model for all the prostate cell lines in GDSC. Negative values mean that the cell line is more sensitive to the drug. Drugs included in Table 1 were highlighted. "Other targets" are drugs targeting model-related genes that are not part of Table 1.

Figure 7: Cell viability assay determined by the fluorescent resazurin after a 48-hours incubation showed a dose-dependent response to different inhibitors. A) Cell viability assay of LNCaP cell line response to 17-DMAG HSP90 inhibitor. B) Cell viability assay of LNCaP cell line response to PI-103 PI3K/AKT pathway inhibitor. C) Cell viability assay of LNCaP cell line response to NMS-E973 HSP90 inhibitor. D) Cell viability assay of LNCaP cell line response to Pictilisib PI3K/AKT pathway inhibitor. Concentrations of drugs were selected to capture their drug-dose response curves. The concentrations for the NMS-E973 are different from the rest as this drug is more potent than the rest (see Material and methods).

Figure 8: Hsp90 inhibitors resulted in dose-dependent changes in the LNCaP cell line growth. A) Real-time cell electronic sensing (RT-CES) cytotoxicity assay of Hsp90 inhibitor, 17-DMAG, that uses the Cell Index as a measurement of the cell growth rate (see the Material and Methods section). The yellow dotted line represents the 17-DMAG addition. B) RT-CES cytotoxicity assay of Hsp90 inhibitor, NMS-E973. The yellow dotted line represents the NMS-E973 addition.

Figure 9: PI3K/AKT pathway inhibition with different PI3K/AKT inhibitors shows the dose-dependent response in LNCaP cell line growth. A) Real-time cell electronic sensing (RT-CES) cytotoxicity assay of PI3K/AKT pathway inhibitor, PI-103, that uses the Cell Index as a measurement of the cell growth rate (see the Material and Methods section). The yellow dotted line represents the PI-103 addition. B) RT-CES cytotoxicity assay of PI3K/AKT pathway inhibitor, Pictilisib. The yellow dotted line represents the Pictilisib addition.

Appendix, Supplementary Files, Source code and Source data files

Appendix 1, a document with supplemental analyses, extended results and introduction to the methodologies used in present work.

Supplementary File 1, a zipped folder with the generic prostate model in several formats: MaBoSS, GINsim, SBML, as well as images of the networks and their annotations.

Supplementary File 2, a jupyter notebook to inspect Boolean models using MaBoSS. This notebook can be used as source code with the model files from Supplementary File 1 to generate Figure 3.

Supplementary File 3, a zipped folder with the TCGA-specific personalised models and their Apoptosis and Proliferation phenotype scores.
Supplementary File 4, a TSV file with all the phenotype scores, including *Apoptosis* and *Proliferation*, of the TCGA-patient-specific mutations. In the mutation list "_oe" stands for an overexpressed gene and "_ko" for a knocked out gene.

Supplementary File 5, a zipped folder with the cell-lines-specific personalised models.

Supplementary File 6, a TSV file with all the phenotype scores, including *Apoptosis* and *Proliferation*, of all 32258 LNCaP-cell-line-specific mutations and the wild type LNCaP model. In the mutation list "_oe" stands for an overexpressed gene and "_ko" for a knocked out gene.

Supplementary File 7, a spreadsheet with the Key Resources Table of this work.

Source code 1, file needed to obtain Figure 4. Processed datasets needed are Source data 1 and 2 and are located in the corresponding folder of the repository: https://github.com/ArnauMontagud/PROFILE_v2/tree/main/Analysis%20of%20TCGA%20patients' simulations

Source code 2, file needed to perform the drug dosage experiments and obtain Figure 5 from the main text and Figures 27 and 34-39 from Appendix 1. Processed datasets needed is Source data 3 and is located in the corresponding folder of the repository: https://github.com/ArnauMontagud/PROFILE_v2/tree/main/Gradient%20inhibition%20of%20nodes

Source code 3, file needed to obtain Figure 6. Processed datasets needed are Source data 4 and 5 and are located in the corresponding folder of the repository: https://github.com/ArnauMontagud/PROFILE_v2/tree/main/Analysis%20of%20drug%20sensitivities%20across%20cell%20lines

Source code 4, file needed to obtain Figures 7, 8 and 9. Processed datasets needed are Source data 6, 7 and 8 and are located in the corresponding folder of the repository: https://github.com/ArnauMontagud/PROFILE_v2/tree/main/Analysis%20of%20experimental%20validation
Figure 1: Workflow to build patient-specific Boolean models and to uncover personalised drug treatments from present work.
Figure 2: Prostate Boolean model used in present work.
Figure 3: Prostate Boolean model MaBoSS simulations.
Figure 4: Associations between simulations and Gleason grades (GG).
Figure 5: Phenotype score variations and synergy upon combined ERK and MYC_MAX (A and C) and HSPs and PI3K (B and D) inhibition under EGF growth condition.
Figure 6: Model-targeting drugs’ sensitivities across prostate cell lines.
Figure 7: Cell viability assay determined by the fluorescent resazurin after a 48-hours incubation showed a dose-dependent response to different inhibitors.
Figure 8: Hsp90 inhibitors resulted in dose-dependent changes in the LNCaP cell line growth.
Figure 9: PI3K/AKT pathway inhibition with different PI3K/AKT inhibitors shows the dose-dependent response in LNCaP cell line growth.