Dissertation
submitted to the
Combined Faculty of Natural Sciences and Mathematics
of the Ruperto Carola University Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by
Vladislav Kim, M. Sc.
born in Tashkent, Uzbekistan
Oral examination: October 25, 2021

Analysis of high-throughput screening data for precision oncology and antibiotic combination modelling

Referees:
Prof. Dr. Michael Boutros
Prof. Dr. Oliver Stegle

Acknowledgements

I am grateful for the opportunity to have completed my PhD at one of the best research institutes in the world. Over the years I have met so many inspiring people at EMBL Heidelberg. I enjoyed my scientific collaborations and social interactions with my amazing colleagues. This experience would not have been possible, had I not been recruited by my supervisor — Wolfgang Huber. I am thankful for Wolfgang's mentorship and kindness through all these years.

I would like to thank all members of Huber group, past and present. Having friendly, fun and knowledgeable colleagues made this journey much easier than it could have been. I am especially thankful for Simone's kind support and generosity.

Next, I would like to thank Elisabetta, whose passion for antibiotic resistance research was inspiring. I learned a lot from our collaboration and we made a pretty good research team. I would also like to thank Sophie with whom I collaborated on a compound screen in leukemia. I enjoyed being part of an interdisciplinary team of doctors, biologists and bioinformaticians.

Working in multiple scientific collaborations can be both challenging and motivating. I appreciate all the feedback that Nassos and Sascha provided in our numerous meetings. I would also like to thank my TAC members Michael, Oliver and Anna for our insightful discussions.

I am grateful for having caring and supportive friends in my life. A big thank you goes to my Heidelberg friends: Fabio, Diana, Steve, Miriam, Mark, Anna, Rachel, Julian, Florian, Robin and Gwen for brightening my life with their positivity and happy vibes. I'd like to give special thanks to my best friends from Bremen: Linda and Henrike, who supported me from my first year on in Germany. I am thanking my favorite couple: Mathias and Zoe, with whom I probably spent the most time quarantining in 2020. I am thanking Jan for his support and Instagram memes. And of course special thanks to my EMBL fellows and EMBL alumni: Jakob, Ben, Karen, Martine, Nils, Nikos, Can, Erica, Carol, Gabi, Florian, Stefan, Chelsea, Jen, Alejandro, Almut, Emma, Holly for all the good times we shared.

Last but not least, I would like to thank my mother and grandmother for their infinite love and patience. Thank you for believing in me.

Abstract

The recent success of targeted anticancer therapeutics has propelled cancer genomics to the forefront of clinical oncology. Patients with actionable mutations can now be treated with compounds that target cancer-specific pathways while minimizing damage to healthy tissues. Precision oncology aims to match treatment to a tumor's mutational composition. Beyond genomic sequencing technologies, drug sensitivity assays play a crucial role in functional profiling of cancers. Tumor sample accessibility makes blood cancers particularly amenable to *ex-vivo* drug screening. Given the number of functional assays established in leukemias and lymphomas, critics often question the diagnostic value of *ex-vivo* drug sensitivities.

One limitation of compound screening techniques in liquid cancers is the omission of microenvironment signals secreted by the bone marrow *in vivo*. To assess the impact of the microenvironment on drug response, I analyzed the high-throughput imaging data obtained from a compound screen conducted in primary leukemias and lymphomas ($n = 108$). In this study, patient-derived cancer cells were exposed to compound perturbations both alone and in coculture with a bone marrow stromal cell line. In total, 50 compounds were probed at 3 different concentrations across 2 culture conditions. I developed an automated analysis workflow, which was applied to $> 700,000$ confocal microscopy images. To extract multivariate phenotypes, I implemented a Python package (`bioimg`) for single-cell morphological profiling and performed the statistical analysis of compound effects in mono- and coculture.

One of the key findings of the leukemia-stroma coculture study was that about 50% of the probed compounds were less effective in coculture compared to monoculture. Stratifying by compound class, I found that the efficacy of chemotherapeutics, BET and proteasome inhibitors was diminished by stroma-mediated protection. JAK inhibitors were the only compounds in the screen that reduced stromal protection. However, pro-survival effects of stroma were not uniform and stroma-induced morphology changes observed in cancer cells varied among samples. To understand this variability, I explored drug-gene associations in the presence of microenvironment signals and found that IGHV-unmutated and trisomy-12-positive samples gained stronger stromal protection when treated with BCR inhibitors.

In addition to precision medicine, antibiotic resistance research makes common use of high-throughput screening techniques. Misuse of antibiotics is driving the evolution and expansion of antibiotic-resistant pathogens, posing a significant risk to global public health. The current drug

development efforts to combat this urgent threat are inadequate outside of academia. To address this gap, researchers have developed drug combination profiling systems to identify synergistic and antagonistic drug pairs. Species and strain specificity of synergies and antagonisms necessitates high-throughput investigations in multiple bacterial strains. To tackle these challenges, I analyzed the largest antibiotic combination screen in Gram-positive bacteria that exists to date. In this study, about 2000 drug pairs were probed in *B. subtilis*, *S. pneumoniae*, and two *S. aureus* strains. I developed a computational analysis pipeline that processed high-throughput bacterial growth data. My analysis revealed the landscape of drug interactions in Gram-positive species. I observed high within-class synergistic rates, especially for cell wall targeting compounds and inhibitors of protein and DNA synthesis. Compared with drug interactions observed in Gram-negatives, both abundance and interspecies conservation rates of synergies and antagonisms were lower in the Gram-positive organisms.

Currently, high-throughput screening remains the only feasible method of uncovering drug-drug interactions on a large scale, as only a minority of synergy and antagonism mechanisms have been fully elucidated. To facilitate the rational design of combinatorial therapies, I analyzed drug-drug and drug-gene interaction data in *E. coli* and *S. typhimurium* with the goal of identifying genetic determinants of drug-drug interactions. Consistent with the previous findings in *E. coli*, my analysis revealed that ATP and lipopolysaccharide biosynthesis were among the most important biological processes for drug-drug interactions. In *E. coli* I found that ATP synthesis was particularly important for synergistic interactions among cell wall targeting compounds. Furthermore, I could show that it is possible to predict novel drug-drug interactions using chemogenomic data in *E. coli* and *S. typhimurium*. Finally, using a trained machine learning model I was able to identify genes predictive of drug-drug interactions within specific drug classes and calculate feature importance for individual predictions.

Zusammenfassung

Dank jüngster Erfolge von gezielten Krebstherapien ist die Krebsgenomik ein wesentlicher Bestandteil der klinischen Onkologie geworden. Patienten mit bestimmten Mutationen können mit Arzneimitteln behandelt werden, die im Gegensatz zu Chemotherapeutika nur krebsspezifische biologische Prozesse angreifen. Die Präzisionsonkologie strebt an, Krebstherapie möglichst auf das Mutationsprofil eines Tumors abzustimmen. Neben Gensequenzierungstechnologien spielen Wirkstoffscreeningverfahren eine bedeutende Rolle in der funktionellen Krebsdiagnostik. Ex-vivo-Wirkstoffscreens werden wegen der leichten Verfügbarkeit der Tumorproben insbesondere bei Blutkrebs eingesetzt. Angesichts der Vielzahl der Ex-vivo-Screening-Assays, die bei Leukämien und Lymphomen Anwendung finden, wird oft die Aussagekraft solcher diagnostischer Verfahren in Frage gestellt.

Eine mögliche Schwachstelle von vielen Wirkstoffscreeningmethoden in hämatologischen Krebsarten ist die Abwesenheit von Signalen der Mikroumgebung des Tumors im Knochenmark. Um den Einfluss der Mikroumgebung auf die In-vitro-Wirksamkeit zu untersuchen, habe ich Hochdurchsatzmikroskopiedaten von einem Wirkstoffscreen analysiert, welcher in Leukämie- und Lymphomproben ($n = 108$) durchgeführt wurde. In dieser Studie wurden von Patienten stammende Krebszellen verschiedenen Wirkstoffen ausgesetzt, sowohl in Monokultur als auch in Kokultur mit einer Knochenmarkstromazelllinie. Insgesamt 50 Wirkstoffe in 3 verschiedenen Konzentrationen wurden in Mono- und Kokultur getestet. Für die Auswertung von über 700.000 Konfokalmikroskopiebildern habe ich eine Analyse-Pipeline entwickelt. Um das Auslesen multivariater Phänotypen zu ermöglichen, habe ich ein Python-Paket (`bioimg`) für Einzelzellmorphologieanalyse implementiert. Des Weiteren habe ich die statistische Analyse von In-vitro-Wirkstoffwirkungen in Mono- und Kokultur durchgeführt.

Eine der wichtigsten Erkenntnisse der Kokultur-Studie war, dass circa 50% der getesteten Wirkstoffe ihre Wirksamkeit in Kokultur eingebüßt haben. Mit Blick auf Wirkstoffgruppen habe ich festgestellt, dass die Zytotoxizität von Chemotherapeutika sowie BET- und Proteasominhibitoren von stromavermittelter Schutzwirkung verringert wurde. JAK-Inhibitoren waren die einzige Wirkstoffgruppe, welche die Schutzwirkung des Stromas in Kokultur reduziert hat. Allerdings war die überlebensfördernde Wirkung von Stromazellen nicht einheitlich und unterschied sich von einer Probe zur anderen. Ferner variierten stromainduzierte Morphologieveränderungen, die in Krebszellen beobachtet wurden, je nach Probe. Um diese Variabilität zu erforschen, habe ich zusätzlich Wirkstoff-Gen-Zusammenhänge in Kokultur untersucht, mit dem Ergebnis, dass sowohl IGHV-unmutierte als

auch Trisomie-12-positive CLL-Proben stärkeren Schutz vor BCR-Inhibitoren in Stromakokultur erhalten haben.

Über die Präzisionsmedizin hinaus werden Hochdurchsatzverfahren auch in der Antibiotikaresistenzforschung eingesetzt. Antibiotikamissbrauch in humaner Medizin und Tierhaltung hat zur Evolution und Ausbreitung von antibiotikaresistenten Keimen geführt, was eine Gefahr für das globale Gesundheitssystem darstellt. Die Wirkstoffentwicklung jenseits der akademischen Forschung setzt sich kaum mit dieser Problematik auseinander. Um diesen Bedarf zu decken, haben Forscher Screening-Verfahren für synergistische Wirkstoffkombinationen entwickelt, die im Kampf gegen resistente Bakterien eingesetzt werden könnten. Erschwerend dazu kommt, dass die Mehrheit der synergistischen und antagonistischen Wirkstoffinteraktionen eine Artspezifität aufweist. Dementsprechend müssen Screenings in mehreren Bakterienarten durchgeführt werden. Um diesen Herausforderungen gerecht zu werden, habe ich die Daten des bisher größten kombinatoriellen Screens in Gram-positiven Bakterien ausgewertet. In dieser Studie wurden 2000 Wirkstoffpaare in *B. subtilis*, *S. pneumoniae* und 2 Stämmen von *S. aureus* getestet. Ich habe eine Bioinformatik-Pipeline entwickelt mit deren Hilfe die Hochdurchsatzdaten ausgewertet wurden. Die Analyse lieferte eine umfangreiche Zusammenstellung von Wirkstoffinteraktionen in Gram-positiven Bakterienarten. Unter Wirkstoffen, die zu denselben Antibiotikaklassen gehören, habe ich hohe synergistische Raten festgestellt, insbesondere unter Zellwandsynthese hemmenden Antibiotika und Protein- und DNA-Synthese-Inhibitoren. Im Vergleich zu Gram-negativen Bakterien war das Vorkommen und interspezifische Erhaltung der Wirkstoffinteraktionen geringer in den Gram-positiven Bakterienstämmen.

Hochdurchsatz-Screenings sind momentan das einzig mögliche Verfahren zum Entdecken von Wirkstoffinteraktionen im großen Stil, weil nur ein Bruchteil der Synergien und Antagonismen in Bezug auf ihren Entstehungsmechanismus aufgeklärt ist. Um das rationale Design von Wirkstoffkombinationen zu ermöglichen habe ich Wirkstoff-Wirkstoff- und Wirkstoff-Gen-Interaktionen in *E. coli* und *S. typhimurium* analysiert, damit ich genetische Auslöser von Wirkstoff-Wirkstoff-Interaktionen identifizieren konnte. Die Analyse hat gezeigt, dass den Proteinen der ATP- und Lipopolysaccharidsynthese eine wichtige Rolle bei der Entstehung der Synergien zukommt. In *E. coli* war insbesondere die ATP-Synthese wichtig für Synergien zwischen Zellwandsynthese hemmenden Antibiotika. Des Weiteren konnte ich zeigen, dass man neue Wirkstoff-Wirkstoff-interaktionen in *E. coli* und *S. typhimurium* anhand von chemogenomischen Daten vorhersagen kann. Abschließend habe ich anhand des verwendeten Machine-Learning-Modells Gene identifiziert, die für die Vorhersage von Wirkstoff-Wirkstoff-Interaktionen von größter Bedeutung waren.

Table of contents

Chapter 1

Introduction

1.1 Crisis in drug development

Since early 1990's drug discovery has been stagnating with higher attrition rates and longer development phases reported in recent years (Pammolli et al., 2011). The total cost of developing a new drug skyrocketed in the past decades and crossed the billion-euro mark in early 2010s (Wouters et al., 2020). Only a small fraction of therapeutic candidates clears all 3 phases of clinical trials, with the metaphorical "valley of death" separating basic and translational research (Seyhan, 2019).

The unprecedented success of high-throughput technologies advanced our understanding of complex diseases and produced numerous putative drug targets. In light of this tremendous progress, it is seemingly paradoxical that drug development is stalling. There are a number of reasons for the current crisis in drug development. The simplest explanation proposed by some researchers is that "easy" targets were exhausted in the past (Evenson, 1993). As pharmaceutical companies are taking on more challenging targets, the probability of success decreases (Pammolli et al., 2011). The shift in focus towards more difficult therapeutic areas can be further explained by market forces, which incentivize companies to adopt "high-risk, high-return" strategies.

Economic incentives also played an important role in many companies abandoning antibiotic development efforts (Piddock, 2012). Antibiotics have a low return on investment (Nathan and Goldberg, 2005; Plackett, 2020) compared to drugs treating cancer or chronic diseases, which typically have longer treatment periods. With lagging antibacterial drug development, the crisis of antibiotic resistance is looming (Ventola, 2015).

Drug repurposing (Pushpakom et al., 2018) and combinatorial therapies (Mokhtari et al., 2017; Sun et al., 2016a) are attractive alternatives to *de-novo* drug development. Academic research can be instrumental in identifying synergistic drug combinations, screening natural compounds (Atanasov et al., 2021; Harvey, 2008) or repositioning already approved drugs. Furthermore, lack of reproducibility (Baker, 2016; Scannell and Bosley, 2016) is another obstacle to translation of basic research findings. More accurate screening techniques are crucial for boosting drug discovery success rates and for bridging the gap between bench and bedside.

1.2 Drug screening and precision oncology

Cancer is a disease with considerable interpatient and intratumor heterogeneity (Bedard et al., 2013). Compound perturbation screens in cancer cell lines (Barretina et al., 2012; Garnett et al., 2012; Iorio et al., 2016) confirmed the importance of known cancer gene mutations and uncovered novel genetic markers associated with *in-vitro* drug response. Leveraging genomic and transcriptomic information to determine the optimal course of treatment is one of the central objectives of precision oncology (Senft et al., 2017). Tumor sequencing is widely used in oncological practice (Fittall and Van Loo, 2019; van der Velden et al., 2017) and in presence of actionable genetic alterations (Bungartz et al., 2018; Kumar-Sinha and Chinnaiyan, 2018) cancer therapy can be tailored to the molecular profile of a patient.

In addition to next generation sequencing (NGS) technologies, oncologists can use *in-vitro* compound screening for functional profiling of patients (Friedman et al., 2015; Letai, 2017). Drug sensitivity assays coupled with NGS technologies played a key role in genotype-phenotype mapping as many novel drug-gene associations were uncovered in multiomic studies combining drug response assays with genomic or transcriptomic profiling (Barretina et al., 2012; Dietrich et al., 2018; Garnett et al., 2012; Meyer et al., 2015; Schütte et al., 2017; Tyner et al., 2018).

Perturbation assays in patient-derived samples provide direct evidence of compound efficacy without relying on the prior knowledge of genetic biomarkers of drug sensitivity (Friedman et al., 2015). In hematological cancers, short-term cell cultures of cancer cells derived from peripheral blood samples were successfully used for establishing drug sensitivity assays (Pemovska et al., 2013; Tyner et al., 2013). Moreover, Dietrich et al. (2018) showed that using drug sensitivity data alone one can stratify leukemia patients into subgroups with distinct molecular pathway dependencies. The results of *ex-vivo* drug screening are now being incorporated into clinical practice and used for therapeutic guidance in blood cancers (Liebers et al., 2019; Snijder et al., 2017; Swords et al., 2018).

One criticism that functional diagnostics in liquid and especially solid tumors face is whether *ex-vivo* culture conditions used for drug screening accurately mirror the *in-vivo* tumor context (Friedman et al., 2015). To add microenvironment signals, coculture models can be used to grow tumor cells with stroma or immune cells (Kämpfer et al., 2017; Kirkpatrick et al., 2011; Kurtova et al., 2009; Zhang et al., 1999). To overcome the limitations of 2D culture, patient-derived organoids (Sachs and Clevers, 2014) and organotypic cultures (Ridky et al., 2010) were established in a number of disease models. Many of these next-generation drug screening assays rely on high-throughput microscopy to generate rich multivariate readouts (Broutier et al., 2017; Czerniecki et al., 2018; Lukonin et al., 2020; Sachs et al., 2018).

1.3 Imaging-based compound profiling

Microscopy-based phenotyping was successfully used in numerous perturbation studies in bacteria, yeast, human cell lines and primary samples (Bray et al., 2016; Brodin and Christophe, 2011; Fuchs et al., 2010; Kraus et al., 2017; Snijder et al., 2017). Even with the simplest imaging modality — brightfield microscopy — one can obtain multivariate readouts that include dozens of morphological features (Borten et al., 2018). In confocal microscopy, a multiplexed panel of fluorescent dyes is used to stain various cellular compartments (Bray et al., 2016) and morphological properties are read out for each color channel. This high-throughput imaging technique is commonly referred to as high-content screening.

Applications of image-based phenotyping are manifold. In compound screens, microscopy provides a means to determine mechanism of action based on phenotypic similarity, as similar compounds are expected to have similar image-based profiles (Perlman et al., 2004). Using microscopy, on- and off-target effects of probed compounds can be unveiled (Loo et al., 2007; MacDonald et al., 2006). High-content screening was also used in a number of genetic perturbation screens (Fuchs et al., 2010; Laufer et al., 2013; Liberali et al., 2014). Drug-gene and gene-gene relationships can be inferred from image-based phenotypes produced by genetic perturbations.

A plethora of computational algorithms were developed for image analysis of microscopy-based perturbation screens. Most analysis pipelines require that individual cells be labelled in all images before single-cell morphological profiling is performed. Both conventional computer vision and deep learning techniques are employed for image segmentation (Caicedo et al., 2017; Mattiazzi Usaj et al., 2016), with a recent trend seeing a stronger shift towards machine-learning-based bioimage analysis (Chandrasekaran et al., 2021).

1.4 Computer vision tasks in high-content screening

In high-content screening thousands of images are collected and quantified to obtain multivariate characterizations of probed conditions (Boutros et al., 2015). Typically, individual cells are identified in each image so that morphological properties can be computed for every cell in the dataset (Caicedo et al., 2017; Mattiazzi Usaj et al., 2016). In biology, one commonly refers to this step as image segmentation.

Definition (Image segmentation) Given an input image I, an image segmentation model $\mathcal{F}$ partitions the image by assigning each pixel I_{ij} to one of the pixel sets $S_k,\ k = 1,\ldots,n$:

$$\mathcal{F}(I) = \{\mathbb{1}(I_{ij} \in S_k) \mid \forall I_{ij},\ k = 1,\ldots,n\}$$

It must be noted that the term “segmentation” is very broad and there are many different approaches that achieve image segmentation. To be precise, a number of computer vision tasks used for image segmentation will be defined and discussed.

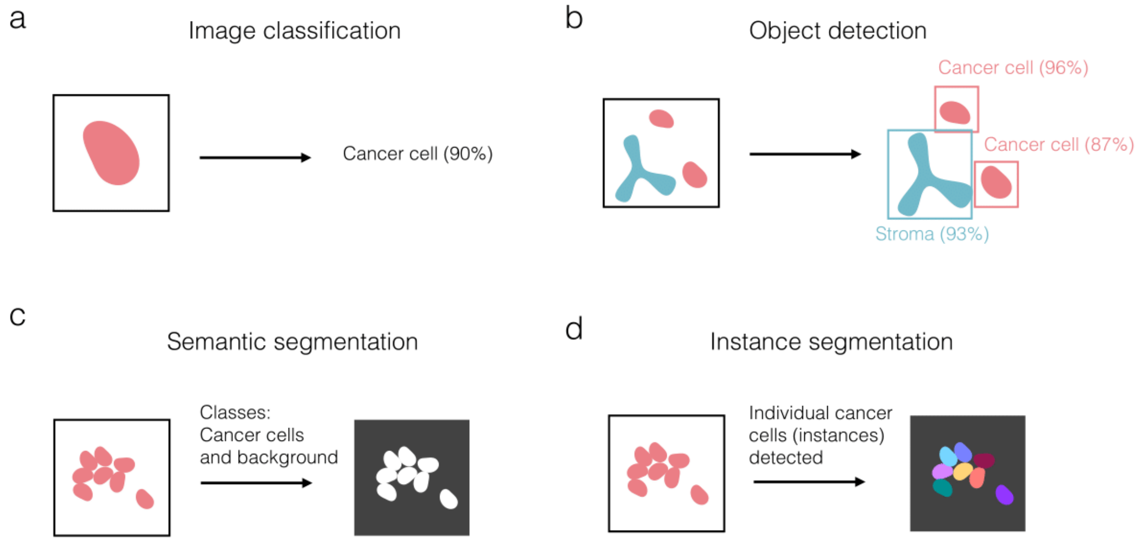


Fig. 1.1 Analysis of large-scale microscopy data can be automated using supervised machine learning methods. a) Image classification outputs a single image-level label. b) Object detection simultaneously localizes and classifies multiple objects in the same image. c) Semantic segmentation classifies image pixels (e.g. foreground and background pixels) d) Instance segmentation goes a step further and assigns pixels to individual class instances

In the past, unsupervised computer vision approaches, such as watershed (Beucher, 1992), were successfully used for image segmentation in cell-line-based high-throughput screens (Fuchs et al., 2010; Laufer et al., 2013). With ever increasing complexity of imaged systems, supervised learning algorithms are becoming more popular for image segmentation (Caicedo et al., 2017; Moen et al., 2019). Most state-of-the-art image analysis tools allow the user to train a machine learning model by providing labelled examples (Arganda-Carreras et al., 2017; McQuin et al., 2018; Sommer et al., 2011). One of the most straightforward trainable models is whole-image classification (Fig. 1.1a).

Definition (Image classification) Given an input image I, with C defined image classes, an image classifier $\mathcal{F}$ outputs a class label $\hat{y} \in 1,\ldots,C$:

$$\mathcal{F}(I) \rightarrow \hat{y}$$

While image classification is frequently used in the natural image domain (animals, urban scenes, etc), it is rarely used directly 'as is' in the biological setting, as microscopy images typically contain many objects (cells, organelles, etc) in one frame that need to be separated ('segmented') before image classification can be run. However, some segmentation-free approaches (Godinez et al., 2017) provide 'weak' image-level labels (e.g. compound class) and perform whole-image classification to obtain probabilistic vectors.

For images with complex backgrounds object detection models (Fig. 1.1b) can be useful for identifying and locating objects of interest. Object detection is an example of multi-task learning that performs simultaneous classification and localization.

Definition (Object detection) Given an input image I, with C defined object classes, an object detection model $\mathcal{F}$ outputs a class label $\hat{y}_o \in 1,\ldots,C$ and bounding box coordinates $\hat{\mathbf{b}}_o = (x_{min},\, x_{max},\, y_{min},\, y_{max})_o$ for each detected object $o \in 1,\ldots,n$ in the image:

$$\mathcal{F}(I) \rightarrow \{(\hat{y_o},\, \hat{\mathbf{b}}_o) \mid \forall o \in 1,\ldots,n\}$$

Training examples for object detection may have multiple overlapping instances in a single frame. Class labels and bounding box coordinates are provided as input for training.

A common task in bioimage analysis is pixel classification, known as semantic segmentation (Fig. 1.1)

Definition (Semantic segmentation) Given an input image I, with C defined pixel classes, a semantic segmentation model $\mathcal{F}$ assigns a class label $\hat{y}_{ij} \in 1,\ldots,C$ to each pixel I_{ij}:

$$\mathcal{F}(I) \rightarrow \{\hat{y}_{ij} \mid \forall I_{ij}\}$$

The simplest form of semantic segmentation is separation of foreground from background (Fig. 1.1c). In more complex biological applications, such as electron micrograph analysis, convolutional neural networks (CNNs) are used for semantic segmentation, with the U-Net architecture being the most prominent example (Ronneberger et al., 2015).

Semantic segmentation partitions an input image into regions belonging to each class. However, within the same class, individual objects have to be separated from one another to achieve instance segmentation (Fig. 1.1d).

Definition (Instance segmentation) Given an input image I with C defined pixel classes, an instance segmentation model $\mathcal{F}$ assigns a class label $\hat{y}_{ij} \in 1,\ldots,C$ and object index $\hat{o}_{ij} \in 1,\ldots,n$ to each pixel I_{ij}

$$\mathcal{F}(I) \to \{(\hat{y}_{ij},\, \hat{o}_{ij}) \mid \forall I_{ij}\}$$

Instance segmentation can be designed as a multistep pipeline that first performs semantic segmentation or object detection and subsequently refines the output. State-of-the art instance segmentation models in the natural image domain extend object detection algorithms directly (He et al., 2020) or adapt other convolutional neural networks (CNNs) to output object masks (Pinheiro et al., 2015, 2016).

1.5 Chemical genetics

Chemical genetics combines compound and genetic perturbations to uncover drug-gene interactions. In chemogenomic screens, drugs are probed in gene deletion, knockdown or overexpression libraries (Roemer et al., 2012). Chemogenomic profiles provide a comprehensive characterization of compound effects, as drugs can be interrogated in genome-wide mutant libraries.

The primary motivation for chemogenomic experiments is to find drug-gene interactions, infer drug mode of action and identify off-target effects of probed compounds. Multiple chemogenomic datasets are available in yeast (Giaever et al., 2002; Jones et al., 2008), bacteria (Nichols et al., 2011; Peters et al., 2016; Santiago et al., 2018) and mammalian organisms (Boutros and Ahringer, 2008; Moffat et al., 2006).

Drug-gene relationships inferred from chemogenomic screens can inform the design of synergistic drug combinations (Breinig et al., 2015), as many drug-drug interactions result from epistatic interactions between targeted genes. In addition to single perturbations, combinations of compounds can be probed in genetic mutants (Lehár et al., 2008). Combination chemogenomic data can be used directly to dissect mechanisms of drug-drug interactions. Probing drug combinations in gene deletions can help elucidate mechanisms of antagonisms and synergies, since drug interactions become neutral in the absence of genes underlying the interaction (Cacace et al., 2017).

Chapter 2

In-vitro modelling of leukemia microenvironment

Summary

In this chapter I will present my contributions to the large-scale imaging-based screen (Subsection 2.3.1) conducted in coculture of primary leukemia cells with bone marrow stroma. I designed an image analysis pipeline (Subsection 2.2.3 and 2.3.2) for segmentation of terabyte-scale leukemia coculture images. Furthermore, I developed a Python package (see Subsection 2.2.12) for single-cell profiling of segmented leukemia and lymphoma cells in coculture. I performed the analysis of drug sensitivity profiles in coculture (Subsection 2.3.6–2.3.8), which produced the most comprehensive reference set of compound profiles in presence of bone marrow stroma (Herbst and Kim, 2021).

In-vitro investigation of immunotherapeutics is the natural ground for microscopy-based coculture screening, as the interaction of cancer and immune cells can be directly observed. To study the effect of bi-specific T cell engagers (BiTE) *ex-vivo*, I analyzed microscopy images of autologous B and T cell coculture (Subsection 2.3.10). Microscopy provided valuable insights that could explain the heterogeneity of patient response to BiTE antibodies (Roider et al., 2021).

The imaging-based coculture screen in primary leukemia samples was conducted by S. Herbst and detailed experimental methods are described in (Herbst, 2020). I carried out all computational analysis for the leukemia-stroma coculture screen presented in this chapter. T. Roider generated FACS and microscopy data for the bispecific T cell engager (BiTE) antibody screen. I performed image analysis, while T. Roider processed FACS data.

2.1 Introduction

Leukemias, lymphomas and myelomas collectively referred to as hematopoietic cancers (HCs) comprise up to 10% of all cancer types (Hsieh, 2014, Chapter 20). These clonal disease entities arise when hematopoietic cells in the blood or lymph nodes undergo malignant transformation (Hsieh, 2014, Chapter 20). Some leukemias may also originate from hematopoietic precursor cells in the bone marrow.

There are many subtypes of leukemia such as chronic lymphocytic leukemia (CLL), acute myeloid leukemia (AML), hairy cell leukemia (HCL), T-cell prolymphocytic leukemia (T-PLL), which among other factors differ in what cell types, i.e. whether those of lymphoid or myeloid lineage, are affected (Hsieh, 2014, Chapter 20). Some leukemia cells, such as those of AML and HCL patients, have distinctive cytological features (Hasserjian, 2013; Summers and Jaffe, 2011).

Even within the same disease entity there is a high degree of heterogeneity that results in different outcomes and many prognostic markers were developed for each subtype of leukemia and lymphoma (Döhner and Gaidzik, 2011; Landau et al., 2013; Moreno and Montserrat, 2008; Paulsson et al., 2010; Stengel et al., 2016). Perhaps, the most well studied class of leukemia is chronic lymphocytic leukemia (CLL), which happens to be the most common leukemia in adults and primarily affects the elderly (Zenz et al., 2010b). Within CLL, a number of genetic markers were identified that allow researchers to stratify patients into subgroups with different expected outcomes and drug sensitivity profiles (Dietrich et al., 2018). In fact, a number of *ex-vivo* studies made the personalized approach feasible: one can test approved compounds and drug combinations *in vitro* using patient material to determine the most promising therapeutic course (Kirtonia et al., 2020; Montserrat et al., 2016; Rozovski et al., 2014; Thimiri et al., 2018). One significant limitation of these approaches is that *ex-vivo* screens are usually conducted in absence of microenvironment signals (Dietrich et al., 2018; Eriksson et al., 2015; Malani et al., 2014; Norberg et al., 2012; Pemovska et al., 2013; Snijder et al., 2017; Tyner et al., 2013).

It has recently become evident that the bone marrow microenvironment plays an important role in disease progression (Duarte et al., 2018). Bone marrow stroma protects leukemia cells from spontaneous and drug-induced apoptosis *in vitro* (Bendall et al., 1994; Konopleva et al., 2002; Kurtova et al., 2009; Lagneaux et al., 1998; Panayiotidis et al., 1996). In light of these findings it is crucial to determine to what extent the findings of monoculture compound screens reflect “real” patient-specific drug sensitivity profiles. Studying the effects of stroma on leukemia cells can advance our understanding of how compound efficacy is altered by the bone marrow microenvironment.

2.1.1 Role of bone marrow microenvironment in leukemia

The interaction of CLL cells with bone marrow stromal cells has been extensively covered in the literature (Burger et al., 1999; Ding et al., 2009; Dubois et al., 2020; Kay et al., 2007; Kurtova et al., 2009). CLL cells and mesenchymal stromal cells (MSCs) interact primarily via direct physical contact and soluble factor exchange (Dubois et al., 2020).

The adhesion of CLL cells to neighboring stromal cells is mediated by a number of ligands and receptors (Fig. 2.1a). Multiple pathways such as B cell receptor (BCR) signalling, Toll-like receptor (TLR) signalling and Wnt signalling are stimulated upon direct contact, which leads to survival and proliferation (Fig. 2.1a). Physical contact with stromal cells appears to be important, as placing a barrier between MSCs and CLL cells diminishes the protective effect of stroma (Dubois et al., 2020).

CLL cells and MSCs engage in a bi-directional exchange of soluble factors (Fig. 2.1b). The best-characterized chemokine axis is CXCR4/ CXCL12, which plays an important role in migration of leukemia cells to the bone marrow and acts as a pro-survival signal (Burger et al., 1999). Moreover, the CXCR4/CXCL12 axis and B cell receptor (BCR) signalling are interlinked (Montresor et al., 2018), with BTK and PI3K activated by CXCL12 in CLL (Fig. 2.1b).

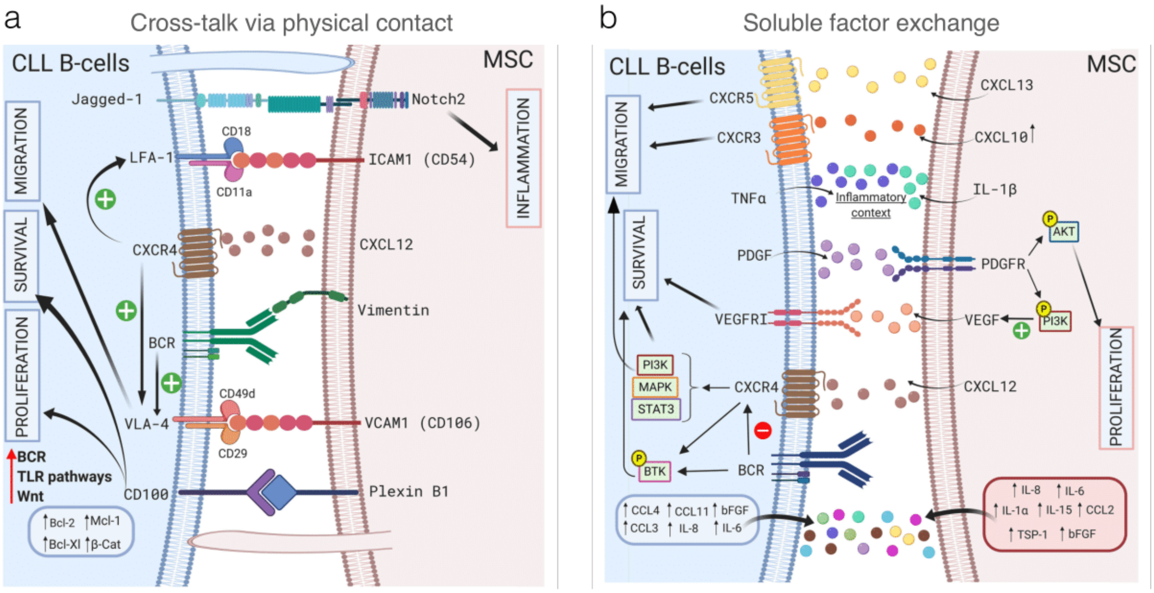


Fig. 2.1 CLL microenvironment schematic. Adapted from Dubois et al. (2020). Copyright 2020 Dubois, Crompot, Meuleman, Bron, Lagneaux and Stamatopoulos. License: Creative Commons Attribution License (CC BY 4.0).

Various strategies were proposed to counter the protective effect of the bone marrow microenvironment. Inhibition of CXCR4 with plerixafor can be used to cut off the

CXCR4/CXCL12 soluble axis (Stamatopoulos et al., 2012) and indeed plerixafor combinations were used in clinical trials to treat acute myeloid leukemia (AML) (Sison et al., 2014; Uy et al., 2012).

Since B cell receptor (BCR) signalling is activated by the bone marrow microenvironment, BTK inhibition was suggested to reduce the protective effect of stroma (Chen et al., 2016; Tissino et al., 2018). However, the CXCR4 / CXCL12 chemokine axis does not solely depend on BTK activation and may still promote the survival of CLL cells via PI3K, MAPK or STAT3 (Fig. 2.1).

Targeting anti-apoptotic proteins is another promising approach, since stromal cells lead to upregulation of pro-survival molecules such as Bcl-2 and Mcl-1 (Fig. 2.1). BH3-only proteins are apoptosis regulators that are sequestered by anti-apoptotic Bcl2-family proteins (Delbridge and Strasser, 2015). BH3 mimetics directly target Bcl2-family proteins (Billard, 2013) and these compounds (e.g. venetoclax) could potentially reduce stroma-mediated resistance (Dubois et al., 2020).

Up until now sample sizes of coculture compound screens were very modest, typically with $n < 10$ (Ding et al., 2018; Gupta et al., 2013; Kamdje et al., 2012; Karjalainen et al., 2017; Konopleva et al., 2002; Mudry et al., 2000; Quintarelli et al., 2014; Vianello et al., 2010; Zeng et al., 2009; Zhang et al., 2019). Therefore it is important to develop scalable high-throughput coculture screening systems to assess the effectiveness of the proposed strategies against stroma protection.

2.1.2 Bispecific T cell engager (BiTE) antibodies

Using immune system against neoplastic cells has emerged as an attractive therapeutic strategy (Mellman et al., 2011). Monoclonal antibodies, such as rituximab, are used in a number of haemotological malignancies (Weiner et al., 2009). Rituximab, CD20 antibody, is one of the first-line treatments against B-cell non-Hodgkin's lymphoma (B-NHL). Rituximab binds to CD20, expressed on the surface of healthy and NHL B cells, and recruits cytotoxic components of the immune system. The exact mechanism of rituximab's cytotoxic effect on B cells is unclear (Keating, 2010).

Despite the overall success of rituximab in treating B-cell NHL, relapsed and refractory cases remain a major point of concern (Goebeler et al., 2016). A novel class of bivalent antibodies, known as bispecific T cell engagers (BiTEs), proved their efficacy in relapsed and refractory acute lymphoblastic leukemia (Curran and Stock, 2019) and showed promising results in phase I trial for NHL (Goebeler et al., 2016). Bispecific T cell engagers (BiTEs) have an antibody that binds to CD3, a T cell coreceptor. The CD3 antibody is conjugated with the target-specific antibody forming a single-chain antibody that activates CD8+ and

CD4+ T cells upon simultaneous binding of an antigen-carrying target cell and a cytotoxic T cell (Baeuerle and Reinhardt, 2009).

Unlike CAR T cell therapy, which requires extraction and genetic editing of autologous T cells (Mellman et al., 2011), BiTEs are easier to engineer, which makes BiTE treatment more practical in aggressive malignancies, for which time is a critical variable. However, just like other cell-mediated immunotherapeutics, BiTE antibodies can cause severe adverse effects and treatment with blinautomomab (CD3-CD19 BiTE) results in collateral damage to healthy B cells (Baeuerle and Reinhardt, 2009). *In-vitro* investigations of BiTE antibodies in autologous B and T cell coculture may shed light on factors underlying variable responses to BiTE antibodies observed in different patient subgroups.

2.2 Methods

2.2.1 Image acquisition and preprocessing

Leukemia-stroma coculture screen

Images were acquired using Opera Phoenix High Content Screening System (Perkin Elmer). Primary samples were screened in 384-well plates and detailed experimental methods are described in (Herbst, 2020). Each well was imaged at 3 positions with 10 optical sections per field of view, resulting in $384 \cdot 3 \cdot 10 = 11520$ images per sample and color channel.

For all images I applied maximum intensity projection (MIP) along the z-axis:

$$I_{xy} = \max_{z} I_{xyz}$$

The dimensions of each maximum-intensity projected (MIP) image were 2160×2160. Additionally, I applied gamma correction to all MIP images ($\gamma = 0.3$).

Before providing non-CLL entity images to an object detection model, I combined all color channels and cut RGB-overlayed images into 9 tiles of size ($720 \times 720 \times 3$).

BiTE antibody screen

Images were acquired using Opera Phoenix High Content Screening System (Perkin Elmer). 96-well plates were used for BiTE antibody screening and detailed experimental methods are described in (Roider et al., 2021). Each well was imaged at 4 positions with 8 optical sections per field of view, resulting in $96 \cdot 4 \cdot 8 = 3072$ images per plate and color channel.

I used maximum intensity projection (MIP) to flatten the image stack and applied gamma correction with $\gamma = 0.5$ to the Hoechst channel only. Each MIP image had dimensions 2160×2160.

2.2.2 Segmentation of CLL-stroma coculture images

CLL coculture images were segmented using Perkin Elmer software Harmony. CLL and stroma cells were separated based on nucelus area (Herbst, 2020).

Apoptotic CLL nuclei tend to shrink and become brighter as they accumulate the Hoechst dye. As described in (Herbst and Kim, 2021), viable CLL nuclei were identified based on the minimum area threshold $A > 23.8\,\mu\text{m}^2$, while apoptotic CLL nuclei had area in the range $10\,\mu\text{m}^2 < A \leq 23.8\,\mu\text{m}^2$.

I used the bounding box positions of viable CLL nuclei that were output by Harmony for single-cell characterization as described in Subsection 2.2.4.

2.2.3 Object detection for non-CLL entities

For microscopy images of non-CLL entities, I trained an object detection model that located only leukemia cells and ignored the stroma in coculture wells. I used Faster R-CNN object detection model (Ren et al., 2017) implemented in TensorFlow (Huang et al., 2017). The classes that I specified were:

$$y \in \{\text{viable leukemia cell, apoptotic leukemia cell}\}$$

I chose 5 AML plates[1] from which I randomly sampled 5 DMSO and 5 drug-treated coculture well images. These $5 \cdot 10 \cdot 9 = 450$ images[2] were labelled using `labelimg`[3] tool, i.e. viable and apoptotic leukemia cells were makred by bounding boxes. I split the labelled image data into train (70%), validation (10%) and test (20%) sets.

There are a number of pretrained models available with different convolutional architectures (e.g. Inception v2, ResNet, etc). I downloaded 4 pre-trained models from the TensorFlow detection model zoo[4] and used their weights for hyperparameter initialization:

- `faster_rcnn_inception_v2_Oxford-IIIT`

---

[1] From 5 distinct primary samples

[2] Factor 9 due to each image being cut into 9 tiles prior to training

[3] https://github.com/tzutalin/labelImg

[4] https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf1_detection_zoo.md

- `faster_rcnn_inception_v2_coco`
- `faster_rcnn_resnet50_coco`
- `faster_rcnn_resnet101_coco`

I trained the downloaded object detection models on the labelled coculture images and compared their performance on the validation set in terms of classification and localization loss. I chose Faster R-CNN with the Inception v2 module pretrained on the Oxford-IIIT image set.

2.2.4 Single-cell morphological profiling

Once the microscopy images were segmented, I computed morphological properties and pixel-based measures such as area, extent, equivalent diameter and mean intensity. Additionally, I computed some common shape descriptors including eccentricity, Euler characteristic and moment invariants.

I used `skimage.measure.regionprops` function from the `scikit-image` Python package (van der Walt et al., 2014) to compute size and intensity features as well as moment invariants. In particular, I calculated central and normalized moments up to order $(3,3)$ and the complete set of Hu moment invariants. Additionally, I evaluated the first 12 Zernike moments ($Z_{0,0}$ through $Z_{5,5}$) using `mahotas` Python package (Coelho, 2013). Texture features were obtained using `skimage.feature.greycoprops` function (van der Walt et al., 2014) and additional Haralick features were estimated using `mahotas` package (Coelho, 2013).

I implemented a function `compute_props` for computing all morphological features simultaneously (Subsection 2.2.12). Mathematical details of image moment and texture feature calculation are provided below and can be skipped on first reading.

Size, intensity and shape features

Image moments capture the shape of pixel intensity distributions. Raw image moments of order (p,q) are weighted averages of pixel intensities of an image I:

$$\mu_{p,q} = \sum_{x,y} x^p \, y^q \, I(x,y)$$

Raw moments are not used in practice as they are not invariant under geometric transformations. Instead, I used translation-invariant central moments:

$$\overline{\mu}_{p,q} = \sum_{x,y} (x-\bar{x})^p (y-\bar{y})^q \, I(x,y)$$

as well as normalized central moments, which are translation- and scale-invariant:

$$\eta_{p,q} = \frac{\overline{\mu}_{p,q}}{\overline{\mu}_{00}^{(p+q)/2+1}},\ p+q \geq 2$$

Based on normalized central moments, Hu (1962) constructed 7 image descriptors $\{H_i, i = 1,\ldots,7\}$, known as Hu moments, that are translation-, rotation- and scale-invariant (Huang and Leng, 2010):

$$\begin{aligned}
&H_1 = \eta_{20} + \eta_{02}\\
&H_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2\\
&H_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2\\
&H_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2\\
&H_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]\\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]\\
&H_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})\\
&H_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2]\\
&\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned}$$

For morphological profiling I included the Hu descriptors and Zernike moments, which project pixel intensities $I(x,y)$ onto a set of orthogonal polynomials on the unit disk (Khotanzad and Hong, 1990). Zernike polynomials of (p,q)-order are parameterized in terms of radial distance r and azimuthal angle θ:

$$\Phi_{p,q}(r,\theta) = R_{p,q}(r)\ \exp(ip\theta)\ \text{ with } q \geq 0,\ |p| \leq q$$

with radial polynomials $R_{p,q}(r)$ defined as follows:

$$R_{p,q}(r) = \sum_{k=0}^{\frac{q-p}{2}} \frac{(-1)^k\,(n-k)!}{k!\,\left(\frac{p+q}{2}-k\right)!\,\left(\frac{q-p}{2}-k\right)!}\, r^{q-2k}$$

Zernike moments are computed by placing the origin at the center of an image I and mapping pixel coordinates onto the unit disk range (Khotanzad and Hong, 1990):

$$Z_{p,q} = \sum_{x,y} I(x,y)\ \Phi_{p,q}(r,\theta),\ x^2 + y^2 \leq 1$$

Texture features

Texture features describe local intensity variation (‘smoothness’ or ‘coarseness’) at scales smaller than the size of a typical object in an image (Petrou and García Sevilla, 2006, Chapter 1). One of the most common statistical descriptions of texture is based on features derived from grey-level co-occurence matrices (Bharati et al., 2004).

A grey-level co-occurence matrix (GLCM) can be thought of as a higher-order image histogram with entries $P_{ij}^{d,\theta}$ counting the co-occurence of greyscale levels i and j at distance d and angle θ (Petrou and García Sevilla, 2006, Chapter 3.8):

$$P_{ij}^{d,\theta} = \sum_x \sum_y \delta(i - I(x,y))\delta(j - I(x + d\cos\theta, y + d\sin\theta))$$

Normalized GLCM ($P_{ij} := \frac{P_{ij}}{\sum_{i,j} P_{ij}}$) can be interpreted as the joint probability mass function of greyscale levels $\{0,\dots,G-1\}$ and I computed the following six GLCM-based texture features:

- Entropy:
$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P_{ij}\log P_{ij}$$
- Angular second moment (ASM) / energy[5]:
$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P_{ij}^2$$
- Contrast:
$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P_{ij}(i-j)^2$$
- Dissimilarity:
$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P_{ij}|i-j|$$
- Homogeneity:
$$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} \frac{P_{ij}}{1+(i-j)^2}$$

---

[5] Note that in some texts energy is defined as $\sqrt{ASM}$

- Correlation:

  $$\sum_{i=0}^{G-1}\sum_{j=0}^{G-1} P_{ij}\left[\frac{(i-\mu_x)(j-\mu_y)}{\sigma_x\,\sigma_y}\right]$$

  where μ_x, μ_y, σ_x and σ_y are the first and second moments of marginal distributions $p_x(i) = \sum_j P_{ij}$ and $p_y(j) = \sum_i P_{ij}$.

In the original publication Haralick et al. (1973) proposed 14 GLCM-based texture features for pattern recognition tasks. Some of the additional features require computing probabilities of sums and differences:

$$p_{x+y}(k) = \sum_{\substack{i=0 \\ i+j=k}}^{G-1}\sum_{i=0}^{G-1} P_{ij}$$

$$p_{x-y}(k) = \sum_{\substack{i=0 \\ |i-j|=k}}^{G-1}\sum_{i=0}^{G-1} P_{ij}$$

Below I am providing the definitions of 3 Haralick features computed by `mahotas` package that were subsequently used in the downstream analysis (see Subsection 2.2.6).

- Sum average:

  $$\sum_{k=0}^{2G-2} k \cdot p_{x+y}(k)$$

- Difference variance:

  $$\mathrm{Var}\left(p_{x-y}\right) = \frac{\sum_{k=0}^{G-1}\left(p_{x-y}(k) - \overline{p_{x-y}}\right)^2}{G-1}$$

- Information measure of correlation 1 (InfoMeas1):

  $$\frac{S_{xy} - S_{x\cap y}}{\max(S_x, S_y)}$$

  with $S_{xy} = -\sum_i \sum_j P_{ij} \log P_{ij}$, entropy of the joint distribution; S_x and S_y, entropies of marginal distributions $p_x(i)$ and $p_y(j)$ and

  $$S_{x\cap y} = -\sum_i \sum_j P_{ij} \log(p_x(i) \cdot p_y(j))$$

2.2.5 Image feature preprocessing

I used the variance threshold of $\sigma^2 = 10^{-8}$ to filter out the covariates with little to no variation. Image feature distributions tend to be skewed and often span a wide range of values (Caicedo et al., 2017). Therefore I applied the generalized log-transformation, which additionally has variance-stabilizing properties (Huber et al., 2002):

$$g(x) = \log\left(\frac{x + \sqrt{x^2 + c^2}}{2}\right) \text{ with } c = \min(x)$$

The glog-transformed features were centered and scaled based on the plate mean and variance estimated from control wells.

2.2.6 Image feature selection

Filtering based on replicate correlation was used to select only reproducible features. For all features I computed Pearson correlations between plate replicate pairs. In the downstream analysis, I considered only those features that had replicate correlations of $r > 0.6$ in AML and $r > 0.5$ in CLL.

To construct a non-redundant feature set I used a similar approach described in Fischer et al. (2015). An initial set of 2-3 features was provided. At each iteration I regressed all other features against the set of selected features and added the feature with the highest replicate correlation between regression residuals. The iterative procedure continued as long as the number of features with positive correlations between residuals was exceeding the number of those with negative correlations.

2.2.7 Morphological changes in stroma coculture

To assess morphological changes in coculture, I selected only viable leukemia cells in no-drug control wells. After centering and scaling all image features, I estimated population medians of leukemia cells in monoculture and coculture. The change in morphology due to to stromal cells was quantified as the difference in medians: $m_C - m_M$.

I generated a lower-dimensional embedding of single-cell image feature data using t-SNE (Van Der Maaten and Hinton, 2008). Only reproducible image features $(r > 0.6)$ were used for this analysis (see Subsection 2.2.6). Since many of the image features were correlated, I used the first 20 principal components as input to dimension reduction with t-SNE.

2.2.8 Normalization and aggregation of morphological features

To account for baseline differences, I normalized the data separately for mono- and coculture. In particular, for each morphological feature I estimated the plate mean and variance $\{(\mu_M, \sigma_M^2), (\mu_C, \sigma_C^2)\}$ based on respective mono- and coculture controls. I centered and scaled all morphological properties:

$$\tilde{x} = \begin{cases} \frac{x-\mu_M}{\sigma_M} & \text{monoculture} \\ \frac{x-\mu_C}{\sigma_C} & \text{coculture} \end{cases}$$

I aggregated the normalized image data by taking the arithmetic mean of single-cell profiles in each well.

2.2.9 Drug response quantification

Viability is defined as the ratio of viable cell count to total cell count:

$$v = \frac{n_{viab}}{n_{total}}$$

I used the normalized viability $\tilde{v}$ as a measure of drug response in the leukemia-stroma coculture screen with viabilities normalized to control values $\{v_{0,M}, v_{0,C}\}$ in mono- and coculture:

$$\tilde{v} = \begin{cases} \frac{v}{v_{0,M}} & \text{monoculture} \\ \frac{v}{v_{0,C}} & \text{coculture} \end{cases}$$

2.2.10 Compound efficacy changes in stroma coculture

For each drug I selected the concentration with the highest variance in terms of normalized viability. I applied a t-test to detect statistically significant compound efficacy changes in coculture, with the null hypothesis H_0 assuming that mean normalized viabilities in mono- and coculture are equal ($\tilde{v}_M = \tilde{v}_C$).

To correct for multiple hypothesis testing, I applied the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995) with FDR < 0.01. For the effect size calculation I took the median of dose-response curves in mono- and coculture and computed the percentage change in area under the median dose-response curve in coculture (C) relative to monoculture (M):

$$\varepsilon = \frac{\mathrm{AUC}_C - \mathrm{AUC}_M}{\mathrm{AUC}_M} \cdot 100\%$$

I chose the following criteria for compound efficacy changes in CLL coculture:

$$\text{Efficacy} = \begin{cases} \text{Unchanged} & \text{if } q > 0.01 \\ \text{Decreased} & \text{if } q \leq 0.01 \text{ and } \varepsilon > 5\% \\ \text{Increased} & \text{if } q \leq 0.01 \text{ and } \varepsilon < -5\% \end{cases}$$

I applied the same thresholds for compound efficacy changes in AML coculture. However, since the sample size in AML was much smaller ($n = 17$), I relaxed the significance requirement for drugs with altered efficacy in CLL coculture:

$$\text{Efficacy} = \begin{cases} \text{Changed} & \text{if } q \leq 0.01 \text{ and } |\varepsilon| > 5\% \\ \text{Changed} & \text{if } |\varepsilon| > 5\% \text{ and Changed in CLL coculture} \\ \text{Unchanged} & \text{Otherwise} \end{cases}$$

2.2.11 BiTE screen analysis

Segmentation of B and T cell coculture

In order to segment the nuclei in B and T cell coculture, I first thresholded Hoechst channel images using Otsu's method (Otsu, 1979). I applied shape index (Koenderink and van Doorn, 1992) to enhance the circular appearance of nuclei and subsequently used Laplace-of-Gaussian spot detection algorithm (Kong et al., 2013) to segment individual nuclei. For single-cell profiling (described in Subsection 2.2.4) I output positions of bounding boxes of all detected nuclei with the padding of 5 pixels.

Cell cluster characterization in the BiTE screen

In the BiTE screen we were interested in characterizing clusters of B and T cells. For cluster detection, I used the "top-down" approach, i.e. I identified clusters of viable cells directly without using the results of cell segmentation. Similar to single-cell profiling, I characterized cell clusters in terms of their shape, intensity and texture.

To find clusters in B and T cell coculture, I thresholded the Calcein channel, which marks all viable cells, using Otsu's method (Otsu, 1979). I applied binary dilation with the disk radius of 2 pixels so that all viable cells within the distance range of $d \leq 4 < r/2$ were merged into clusters (where $r \approx 10\,\text{px} \approx 3.3\,\mu\text{m}$ is the estimated radius of a typical B or T cell). Afterward, I applied connected-component labelling (CCL) with 4-connectivity (Wu et al., 2005) to segment individual cluster candidates formed by binary dilation. I defined a cluster to have minimum area of 1000 pixels ($\approx 100\ \mu m^2$), corresponding to the combined area of two dilated viable cells. Binary masks of detected clusters were used for morphological profiling.

Since we observed cell clusters in control wells as well, I normalized cluster profiles with respect to controls to account for baseline cluster formation. I aggregated cluster profiles by taking the mean for each well.

2.2.12 *bioimg*: high-content screening analysis workflow

For morphological profiling I implemented a Python package *bioimg*[6], which accepts the segmented data as input and outputs morphological properties of all segmented objects (e.g. cells or clusters) in an image.

Given an image and bounding box positions of segmented objects, the labelled image class `ImgX` outputs a `pandas.DataFrame` with morphological properties (Code 2.1).

```python
from bioimg import ImgX, get_instance
# load an image and bounding boxes of segmented objects
img, bbox = get_instance(path=platedir, fname=im)
# initialize 'ImgX' class
imgx = ImgX(img=img, bbox=bbox,
            n_chan=['Lysosomal', 'Calcein', 'Hoechst'])
# compute morphological properties for every segmented object
imgx.compute_props()
# get pd.DataFrame with shape (n_cell x n_feature)
img_df = imgx.get_df().copy()
```

Code 2.1 Using an instance of labeled-image class for single-cell morphological profiling.

By default, the features are computed separately for each color channel, if `n_chan` > 1 or channel names are specified at initialization (Code 2.1).

In addition to morphological profiling, *bioimg* has a number of methods implemented for preprocessing and normalization of image data as well as feature selection. Code 2.2 illustrates how the package can be used to preprocess, normalize and aggregate the image data and compute the Pearson correlation between the compound profiles of replicate plates.

```python
from bioimg.singlecell import scale_data, preprocess_data
from bioimg.singlecell import aggregate_profiles
from bioimg.singlecell import get_repcor, select_features
# load viable cells of replicate plates
rep1_df, rep1_annot = load_viable_cells(platedir=os.path.join(path,
    plate1),
```

---

[6] https://github.com/vladchimescu/bioimg

```
                                        wells=all_wells,
                                        annot=annot_df)
rep2_df, rep2_annot = load_viable_cells(platedir=os.path.join(path,
   plate2),
                                        wells=all_wells,
                                        annot=annot_df)

# preprocess the image data (apply glog-transformation)
rep1_df = preprocess_data(df=rep1_df, glog=True)
rep2_df = preprocess_data(df=rep2_df, glog=True)

# center and scale the image data using control well mean / sd
scaler = StandardScaler().fit(rep1_df[rep1_annot['Drug']=='DMSO'])
rep1_scaled = scale_data(rep1_df, scaler=scaler)
scaler = StandardScaler().fit(rep2_df[rep2_annot['Drug']=='DMSO'])
rep2_scaled = scale_data(rep2_df, scaler=scaler)

# aggregate the profiles
prof_rep1 = aggregate_profiles(rep1_scaled, rep1_annot)
prof_rep2 = aggregate_profiles(rep2_scaled, rep2_annot)

# compute replicate correlation between compound profiles
repcor = get_repcor(prof_rep1, prof_rep2)

# select the featuers for which replicate correlation r > 0.6
sel_feats = repcor[repcor > 0.6].index.values
```

Code 2.2 Feature selection based on replicate correlations.

Feature selection can be refined beyond filtering out noisy covariates: I implemented the method based on residual correlation (see Subsection 2.2.6) and its usage is demonstrated in Code 2.3.

```
from bioimg.singlecell import select_residcor
# use the aggregated profiles of replicate plates
# and provide an initial set of features ('sel')
sel_feats = select_residcor(prof1=prof_rep1, prof2=prof_rep2,
                            sel = ['Calcein-eccentricity',
                                   'Hoechst-mean_intensity',
                                   'Lysosomal-mean_intensity'])
# once 'sel_feats', vector of feature names,
# is stored, we can use it to subset image data
imgdf = preprocess_data(df=imgdf, sel=sel_feats, glog=True)
```

Code 2.3 Generating a non-redundant set of features using `select_residcor`.

Finally, *bioimg* provides a number of utility functions for single-cell analysis of image data. In particular, dimension reduction plots can be easily visualized using `plot_dimred` or `facet_dimred` functions (Code 2.4). Jupyter notebooks with more detailed examples and use cases can be found in the *microscopy-notebooks* repository[7].

```python
from bioimg.singlecell import plot_dimred, facet_dimred
# provide the scaled image data as input to PCA
pcs = PCA(n_components=20).fit_transform(imgdf_scaled)
# use the first 20 PCs to compute the T-SNE embedding
X_tsne = TSNE(n_components=2,
              random_state=21,
              perplexity=50).fit_transform(pcs)
# add cell labels
X_df = pd.concat([pd.DataFrame(X_tsne, columns=['tsne1', 'tsne2']),
    cell_labels], axis=1)
# plot and specify color settings
plot_dimred(X_df, hue='Culture', title='DMSO control wells')
# select a subset of features
feat_subset = ['Calcein-eccentricity', 'Lysosomal-mean_intensity',
               'Hoechst-InfoMeas1', 'Calcein-convex_area']
# add the feature subset as columns
X_df = pd.concat([X_df, imgdf_scaled.loc[:,feat_subset]], axis=1)
# plot dimension-reduced data colored by feature values
facet_dimred(X_ctrl, feat_subset=feat_subset,
             nrows=2, ncols=2)
```

Code 2.4 Plotting dimension-reduced image data.

---

[7] https://github.com/vladchimescu/microscopy-notebooks

2.3 Results

2.3.1 Leukemia-stroma coculture screen

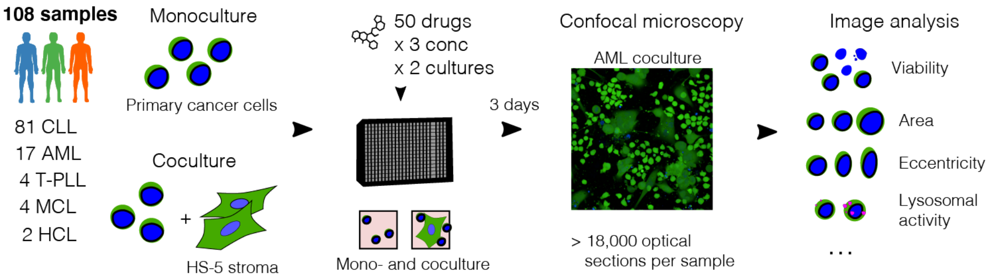


Fig. 2.2 Overview of the compound screen in leukemia-stroma coculture. Graphical abstract created jointly with S. Herbst.

50 compounds were screened at 3 concentrations in 108 blood cancer samples (Fig. 2.2). Mono- and coculture wells were imaged as described in Subsection 2.2.1. In total, 300 distinct conditions (50 drugs x 3 concentrations x 2 culture conditions) were probed in the screen. S. Herbst produced the experimental data, while I performed the analysis.

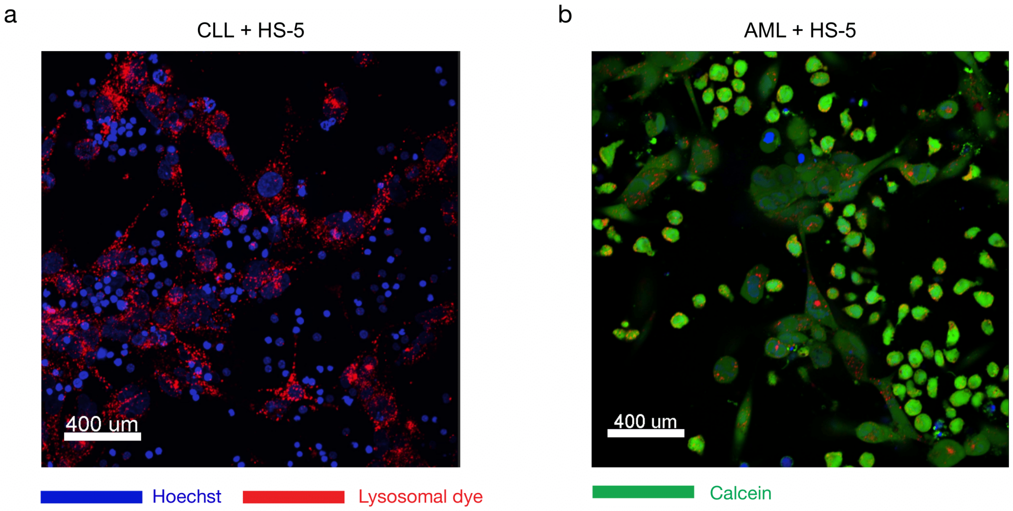


Fig. 2.3 Example images of a) chronic lymphocytic leukemia (CLL) cells with the HS-5 bone marrow stromal cell line. In CLL only nuclei and lysosomal compartments were stained. b) Acute myeloid leukemia (AML) cells in coculture with bone marrow stromal cells. In non-CLL entities (AML, T-PLL, MCL and HCL) a viability marker (Calcein) was used in addition to Hoechst (nuclei) and lysosomal dye. Microscopy images acquired by S. Herbst.

In CLL samples, only Hoechst (nuclei) and lysosomal dyes were used (Fig. 2.3a). In non-CLL entities, a viability marker, Calcein, was added (Fig. 2.3b). Segmentation of CLL microscopy images was performed as described in Subsection 2.2.2. Images of non-CLL entities were more challenging to analyze and I used a deep-learning based model for object detection.

2.3.2 Object detection for coculture images of non-CLL diseases

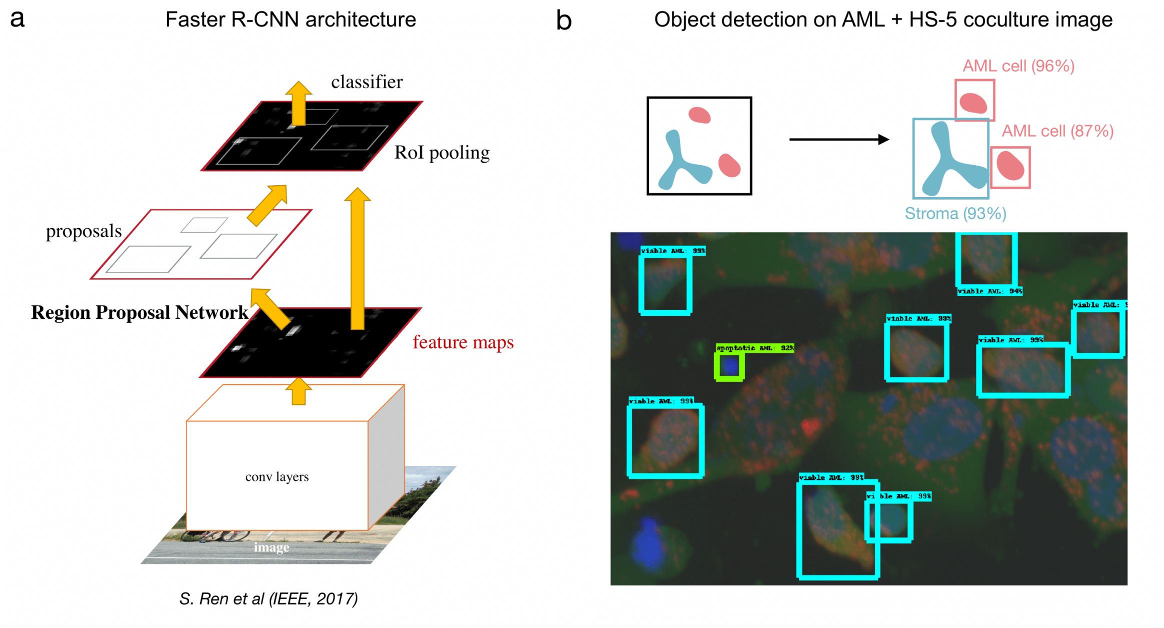


Fig. 2.4 a) Faster R-CNN architecture. Source: Ren et al. (2017) b) Object detection applied to an AML-stroma coculture image. Microscopy image acquired by S. Herbst.

For non-CLL entities (AML, HCL, MCL and T-PLL) I used a machine learning model to identify viable and apoptotic primary cancer cells (Fig. 2.4b). As described in Subsection 2.2.3, I trained 4 object detection models (Fig. 2.5) that differed in CNN architecture (Inception v2, ResNet-50, ResNet-101) and in terms of image sets used for model pretraining (COCO, Oxford-IIIT). I initialized the neural network weights with those obtained from pretraining and provided labelled coculture images as input to each object detection model (see Subsection 2.2.3).

First, I compared the validation losses of these 4 models (Fig. 2.5) and selected the model with the best performance on the validation set – Faster R-CNN with the Inception v2 module pretrained on the Oxford-IIIT image set.

Once the object detection model was chosen, I used the test set to evaluate the classification performance using precision-recall and ROC curves (Fig. 2.6). The average precision on

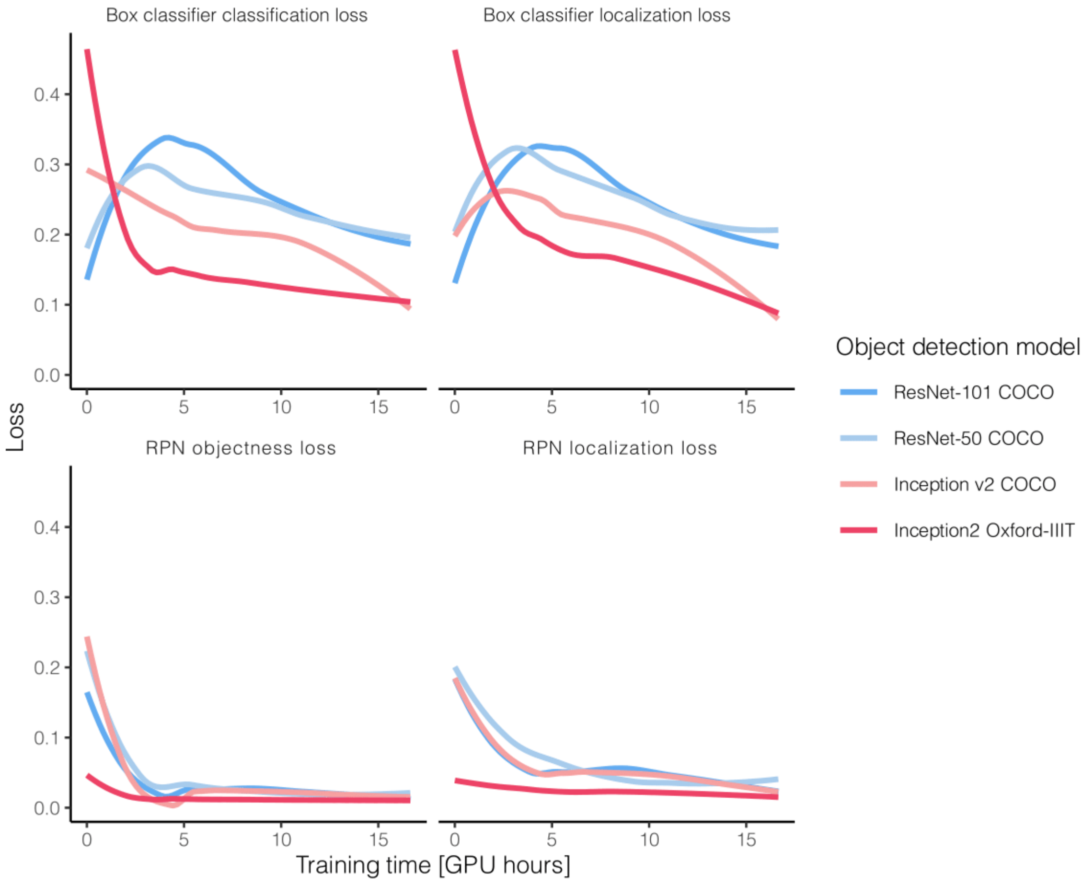


Fig. 2.5 Validation set losses of 4 object detection models that differed in CNN architecture (ResNet or Inception module) as well as in terms of which image set (COCO or Oxford-IIIT) was used to pretrain these models. Data produced jointly with S. Herbst.

the test set was 0.93 for apoptotic and 0.99 for viable leukemia cells, respectively. The test set AUCROC values were 0.98 for both apoptotic and viable leukemia cells.

I used intersection over union (IoU) to quantify the overlap between the predicted and ground-truth bounding boxes. The mean IoU on the test set was > 0.8 for both apoptotic and viable leukemia cells.

I applied the object detection model to all images of non-CLL entities, both mono- and coculture. After I computed viability as the proportion of viable leukemia cells in each well (see Subsection 2.2.9), I checked the replicate correlations of viabilities obtained from the object detection analysis. The results were highly reproducible with Pearson correlations $r > 0.8$ in AML and $r > 0.9$ in T-PLL (Fig. 2.7).

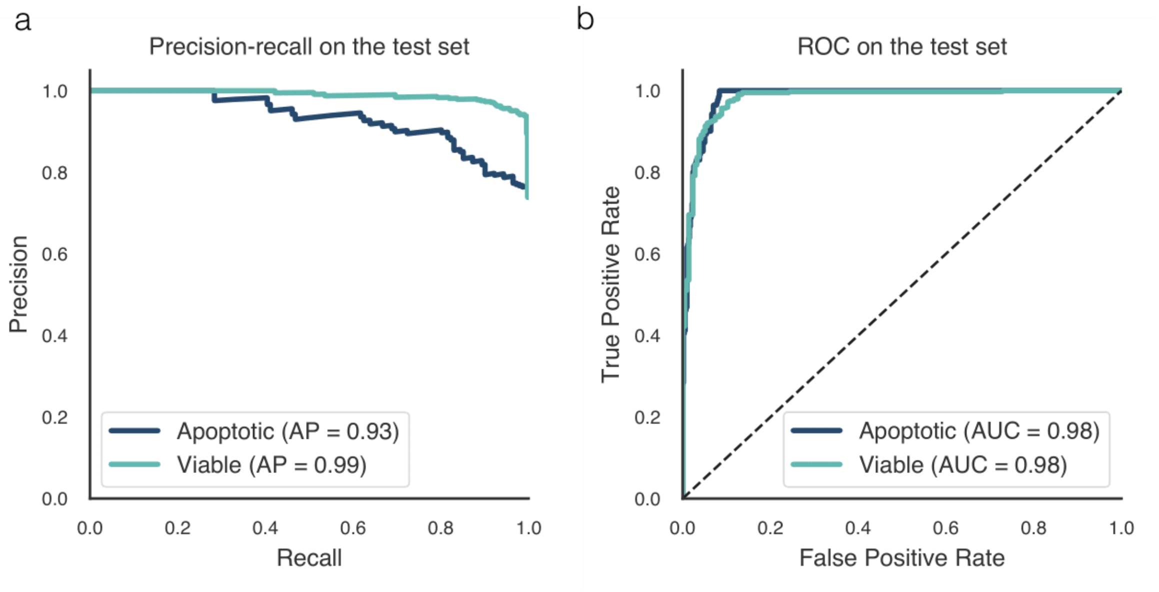


Fig. 2.6 Test set performance of the selected object detection model. Data produced jointly with S. Herbst.

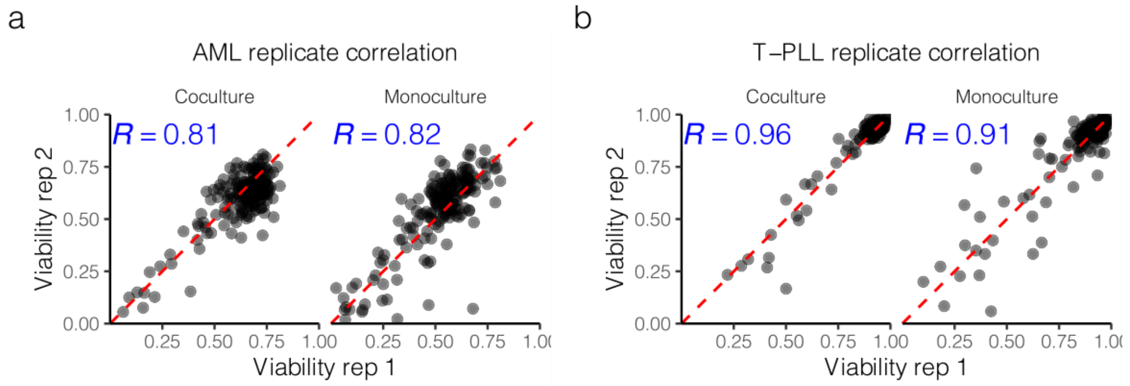


Fig. 2.7 Replicate correlation of the drug response readout in a) AML b) T-PLL. Data produced jointly with S. Herbst.

2.3.3 Image feature selection for morphological profiling

As described in Subsection 2.2.6, I used only reproducible features with replicate correlation thresholds of $r > 0.6$ in non-CLL diseases and $r > 0.5$ in CLL. After filtering out the features with low replicate correlations, lysosomal and nuclear (Hoechst) features were the most abundant morphological descriptors (Fig. 2.9).

For compound profiling I generated a non-redundant set of reproducible features (see Subsection 2.2.6). Table 2.1 shows the number of morphological properties selected with the method based on regression residuals. The detailed list of features used for morphological profiling is provided in Supplementary Tables S1 and S2.

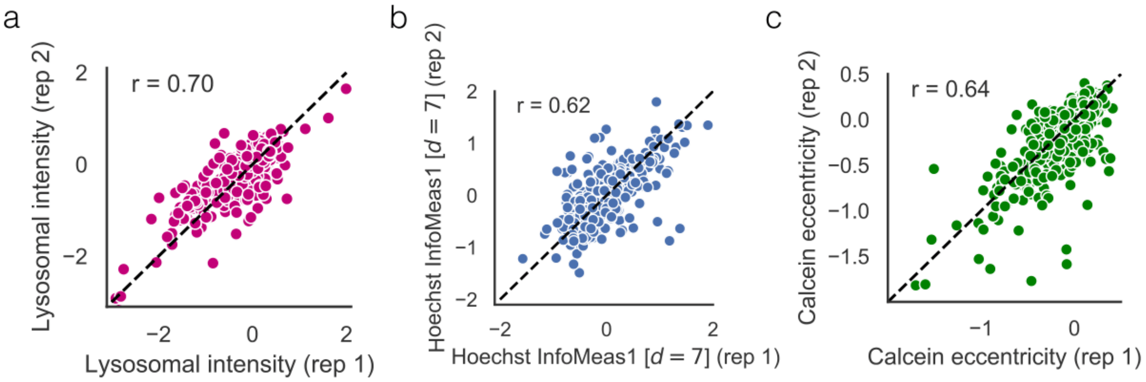


Fig. 2.8 Examples of reproducible features in AML: a) lysosomal intensity b) Hoehst InfoMeas1 c) Calcein eccentricity. Data produced jointly with S. Herbst.

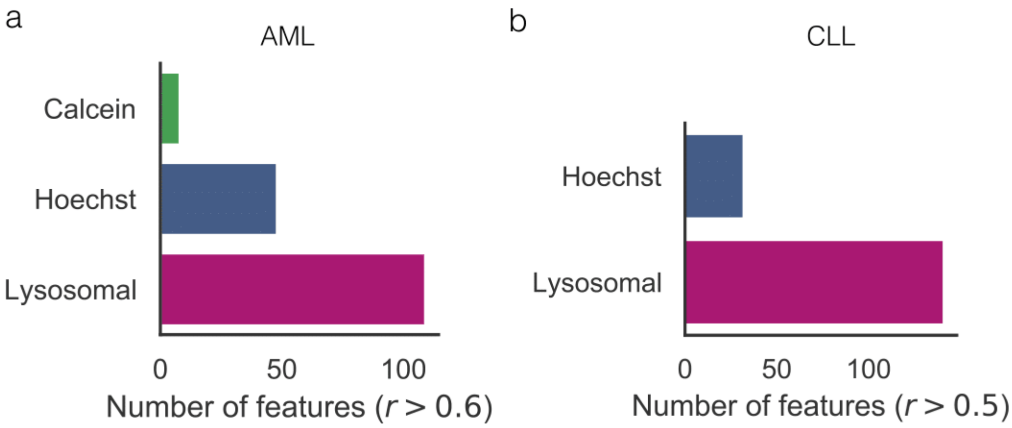


Fig. 2.9 Number of reproducible features in a) AML ($r > 0.6$) b) CLL ($r > 0.5$). Data produced jointly with S. Herbst.

Disease entity	Features (total)	Selected features
CLL	934	42
AML, T-PLL, MCL, HCL	1401	44

Table 2.1 Number of selected morphological features by disease entity in the coculture screen. Data produced jointly with S. Herbst.

2.3.4 Coculture and spontaneous apoptosis

One of the primary motivations for compound screening in stroma coculture is the occurrence of spontaneous apoptosis: in absence of pro-survival signals leukemia cells die even in untreated wells. Culturing leukemia cells with stromal cells reduces *in-vitro* spontaneous apoptosis (Lagneaux et al., 1998; Panayiotidis et al., 1996). To assess the protective effect of stroma, I compared no-drug control viability in mono- and coculture across all samples and found that spontaneous apoptosis rates showed considerable variability (Fig. 2.10a).

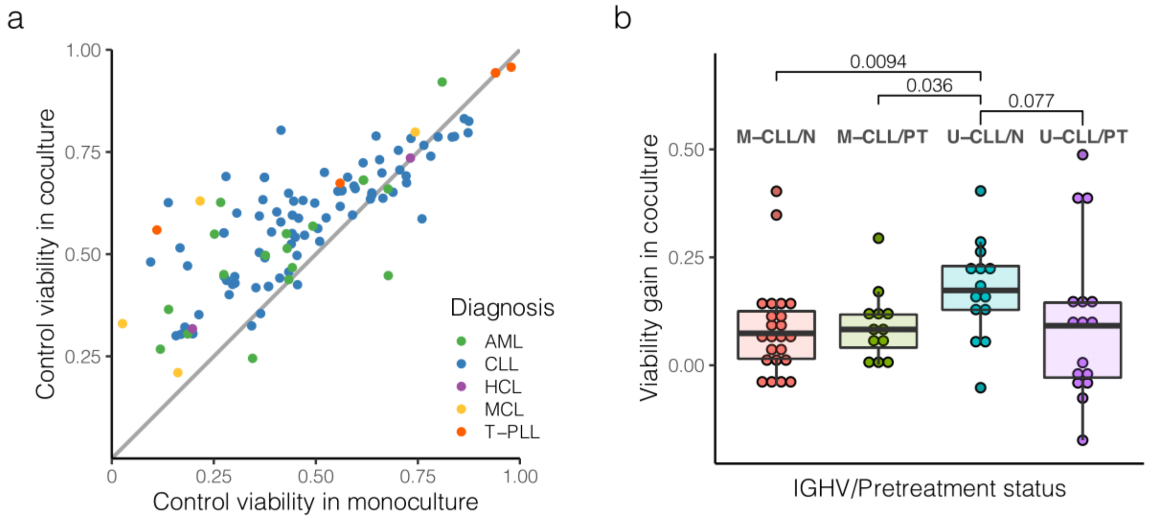


Fig. 2.10 Spontaneous apoptosis in primary leukemia and lymphoma samples. a) Control viability in mono- and coculture varied from sample to sample. Spontaneous apoptosis $(1-\text{viability})$ was reduced in stroma coculture, particularly in samples with low monoculture viability. b) Viability gain in coculture stratified by IGHV mutational status (U-CLL/M-CLL) and pretreatment (N = no pretreatment, PT = pretreated). U-CLL samples derived from patients with no prior treatment gained stronger protection from spontaneous apoptosis in coculture (Wilcoxon signed rank test p-values). Data produced jointly with S. Herbst.

Samples with low base viability in monoculture were protected to a greater extent from spontaneous apoptosis than samples that already fared well on their own without stroma (Fig. 2.10a). This suggests that leukemia cells with a high degree of spontaneous apoptosis rely more on microenvironment signals for survival. Furthermore, I found that U-CLL samples derived from patients with no prior therapy exhibited a higher viability gain in coculture compared with the M-CLL group (Fig. 2.10b).

To investigate the effect of spontaneous apoptosis on the drug sensitivity readout, I divided all samples into 3 groups with low (L), medium (M) and high (H) base viabilities (Fig. 2.11a). The raw (unnormalized) drug sensitivity readout in the low-viability group (L) was more restricted in monoculture (Fig. 2.11b) compared with the readout of other groups

(M and H). Interestingly, after normalization the drug response values of the L group became the most variable in monoculture (Fig. 2.11c) despite having the most limited range in terms of unnormalized values (Fig. 2.11b). By contrast, in coculture no such effect was observed (Fig. 2.11b,c). Thus coculture not only expands the range of the unnormalized drug response readout but also stabilizes the variance of normalized drug sensitivities.

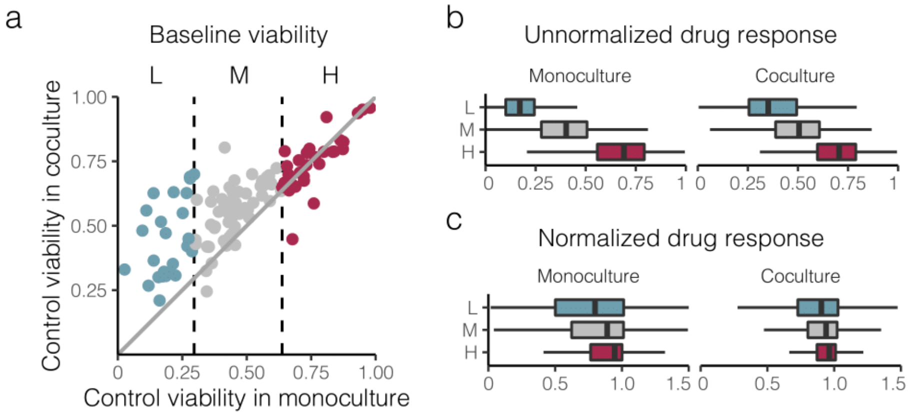


Fig. 2.11 Effect of spontaneous apoptosis on the drug response readout. a) Samples were assigned to low (L), medium (M) and high (H) groups based on control viability in monoculture. b) Unnormalized viability in drug-treated wells by culture condition. In monoculture the dynamic range of drug response in the L group suffers from spontaneous apoptosis. In coculture the interquartile range of the L group is comparable with those of M and H groups. c) Normalized viability in drug-treated wells by culture condition. Despite having a limited dynamic range in terms of unnormalized drug sensitivities, the L group in monoculture becomes the most variable upon normalization. Coculture alleviates this problem. Data produced jointly with S. Herbst.

2.3.5 Effect of microenvironment on unperturbed leukemia cells

The effect of the bone marrow microenvironment is not limited to pro-survival signals that leukemia cells receive from stromal cells. *In vivo*, leukemia cells follow cytokine and chemokine gradients to migrate to the bone marrow (Burger and Kipps, 2002). Leukemia cells in contact with stroma are reported to undergo cytoskeletal re-arrangements (Dubois et al., 2020). To determine whether stroma coculture induces cytological changes, I performed single-cell morphological profiling of no-drug controls (see Subsection 2.2.7).

Each screening plate had equal numbers ($n = 23$) of monoculture and coculture no-drug control wells. I used only viable patient-derived leukemia cells to compare the cell populations of mono- and coculture DMSO controls.

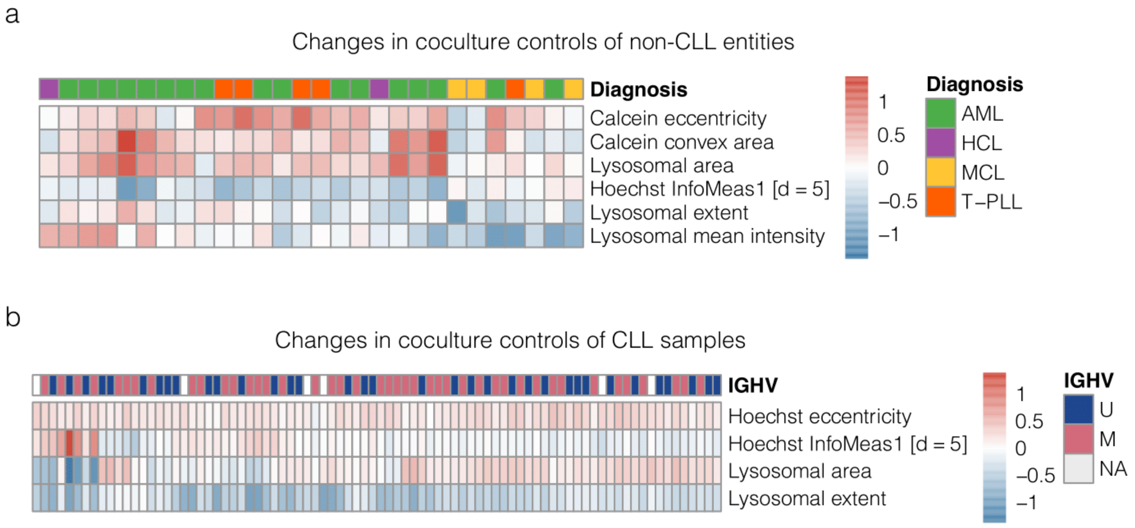


Fig. 2.12 Changes in leukemia and lymphoma cell morphology due to bone marrow stroma in no-drug controls. a) The heatmap shows the difference of medians of cell populations in coculture and monoculture for AML, HCL, MCL and T-PLL samples. All image features were scaled before the difference of medians was computed. b) Difference of medians of coculture and monoculture control populations in CLL samples. The column annotation bar indicates IGHV mutational status. Data produced jointly with S. Herbst.

The effect of stroma on cytological features was heterogeneous with certain primary samples exhibiting dramatic changes in cell morphology in presence of stroma (Fig. 2.12). Moreover, there was a pattern in how different leukemia subtypes responded to the bone marrow microenvironment. For example, AML samples in coculture showed an increase in both cell area and eccentricity, while T-PLL samples primarily showed an increase in cell eccentricity (Fig. 2.12a). There were some disease entities such as MCL and HCL that only presented modest changes in cell shape and size compared with their untperturbed state in monoculture (Fig. 2.12a).

Using multivariate image-based profiles, I performed dimension reduction to visualize the shifts in cell morphology observed in coculture. For each sample I pooled leukemia cells from monoculture and coculture controls and generated a joint two-dimensional embedding based on morphological features. In some samples, such as AML218, the cytological changes observed in coculture compared with monoculture controls were conscpicuous (Fig. 2.13a). Coculture and monoculture control populations of AML218 diverged in the two-dimensional embedding with Calcein (cell) convex area, eccentricity, lysosomal area and nucleus texture (Hoechst InfoMeas1) distinguishing these subpopulations (Fig. 2.13b).

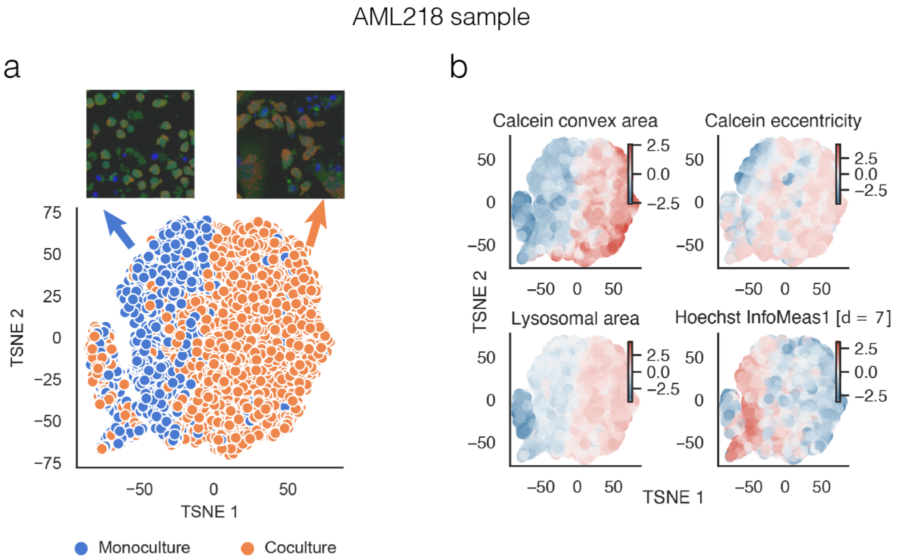


Fig. 2.13 Joint T-SNE embedding of acute myeloid leukemia cells (AML218 sample) in no-drug controls a) Leukemia cells are colored by culture condition. Example images show typical AML cells in mono- and coculture control wells. b) Leukemia cells are colored by centered and scaled morphological features that separate coculture and monoculture control subpopulations. Data produced jointly with S. Herbst.

Acute myeloid leukemia cells in coculture (see Fig. 2.12 and Fig. 2.13a) took on more elliptical shapes compared with the circular cell shape in monoculture controls. The microwell plate is an adhesive surface upon which moving cells appear to have elongated shapes (i.e. have higher cell eccentricity). Due to cell stretching the lysosomal area was also increased in most cocultures (Fig. 2.12). The soluble factors of the bone marrow microenvironment are known to induce a migratory phenotype in leukemia cells and indeed in AML and T-PLL coculture I observed cellular phenotypes consistent with increased motility.

CLL samples were stained only with the nuclear (Hoechst) and lysosomal dyes. Therefore, without cytosol staining, cell shape changes could not be observed directly in CLL coculture. However, CLL samples had a slightly higher nucleus eccentricity (Fig. 2.12b) in coculture. Moreover, lysosomal extent – ratio of lysosomal pixels to the bounding box area – decreased in coculture controls (Fig. 2.12b), while at the same time lysosomal area was somewhat larger or remained unaltered (Fig. 2.12b), providing indirect evidence that stroma coculture led to an increase in cell area and eccentricity in CLL. The morphological changes observed in CLL samples were independent of IGHV mutational status (Fig. 2.12b).

2.3.6 Compound efficacy in CLL and AML coculture

One of the primary goals of the coculture screen was to understand the effect of the bone marrow microenvironment on *ex-vivo* compound activity. For this purpose, 50 compounds and drug combinations were screened in mono- and coculture. I compared normalized drug sensitivities to determine efficacy changes in coculture (see Methods, Subsection 2.2.10).

About half of the screened compounds ($n = 26$) showed a significant decrease in efficacy in CLL-stroma coculture. Coculture induced an upward shift in dose-response curves of compounds with decreased efficacy, such as carfilzomib (Fig. 2.14a).

22 compounds showed no significant change in efficacy in CLL-stroma, which suggests that the bone marrow microenvironment does not interfere with the pathways targeted by these compounds. Dasatinib, an inhibitor targeting multiple kinases, was among the compounds not affected by the bone marrow stroma (Fig. 2.14b).

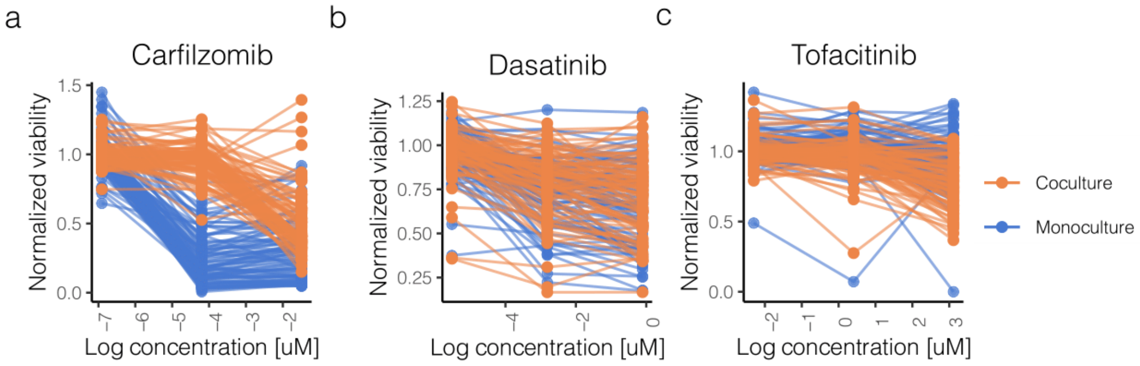


Fig. 2.14 Dose-response curves (CLL / CLL + stroma). a) Carfilzomib: decreased efficacy in coculture b) Dasatinib: unchanged efficacy in coculture c) Tofacitinib: increased efficacy in coculture. Data produced jointly with S. Herbst.

By contrast, compounds with increased efficacy in coculture were rare. Only JAK inhibitors, tofacitnib (Fig. 2.14c) and ruxolitinib, achieved stronger responses in coculture. The effect of JAK inhibitors was enhanced in coculture since JAK inhibition reduces stromal protection from drug-induced and spontaneous apoptosis (see Discussion).

I grouped all probed compounds by their mode of action (Fig. 2.15) and drug class (Fig. 2.16) to identify key patterns in drug-stroma interactions. Common to both AML and CLL was that BET inhibitors, chemotherapeutics and proteasome inhibitors all showed a significant reduction in efficacy in coculture (Fig. 2.15). Another trend conserved across AML and CLL was that JAK inhibitors, tofacitinib and ruxolitinib, induced stronger response in presence of stroma (Fig. 2.15 and Fig. 2.16).

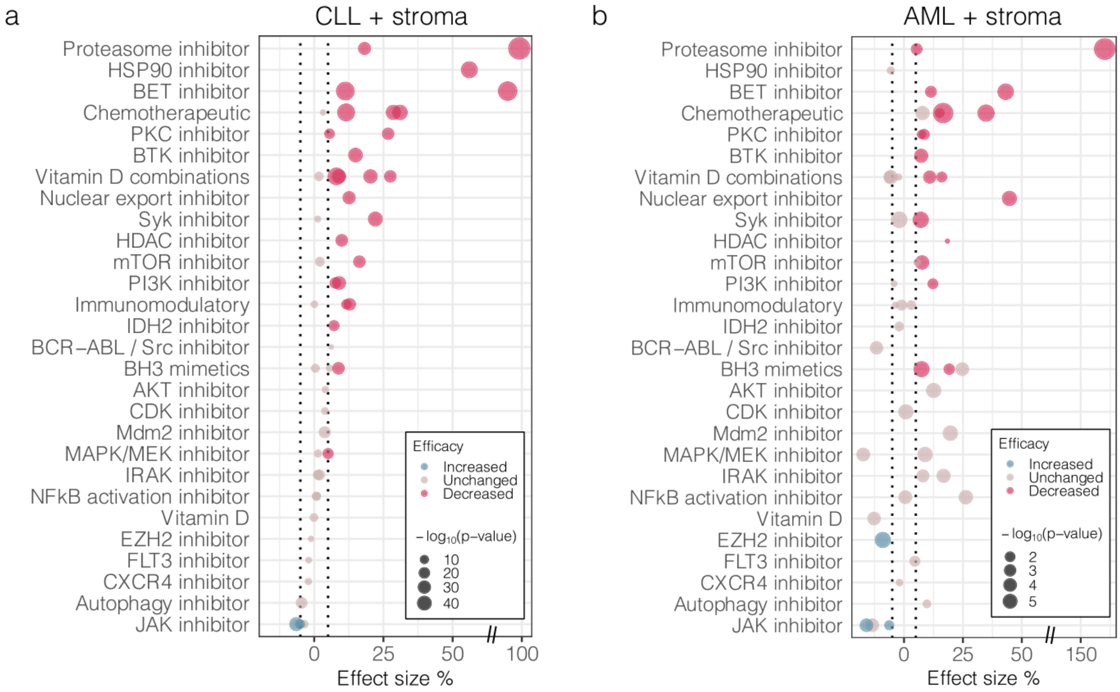


Fig. 2.15 Compound efficacy changes in a) CLL-stroma and b) AML-stroma coculture relative to monoculture. All compounds (points) were grouped by their mode of action (y-axis). Point size indicates the log-scaled p-value; dashed vertical lines mark the effect size cutoffs of $|5\%|$ (see Methods, Subsec. 2.2.10). Data produced jointly with S. Herbst.

Most of the drug-stroma interactions shared between AML and CLL have been uncovered in small-scale microenvironment studies. For example, stroma-mediated resistance to the proteasome inhibitor, carfilzomib was reported by Gupta et al. (2013).

The shielding effect of the bone marrow microenvironment in leukemias was first described for chemotherapeutic agents (Konopleva et al., 2002; Mudry et al., 2000). The upregulation of anti-apoptotic proteins is one of the mechanisms of stroma-mediated resistance (Konopleva et al., 2002). While early reports in B-ALL stressed the importance of direct physical contact (Mudry et al., 2000), more recent studies identified soluble factors crucial for protection from chemotherapy (Burger et al., 2005; Zeng et al., 2009). In accordance with the previous findings, all but one chemotherapeutic (fludarabine, cytarabine and doxorubicin) induced lower responses in coculture (Fig. 2.15 and 2.16).

BET inhibitors, JQ1 and I-BET-762, belong to a broader class of epigenetic drugs. Both JQ1 and I-BET-762 were less effective in AML and CLL coculture (Fig. 2.15). BET inhibition showed promising results in a number of malginancies (Doroshow et al., 2017; Shi and Vakoc, 2014), including hematological cancers (Fong et al., 2015). In uveal melanoma Chua et al. (2019) showed that resistance to BET inhibitors is mediated by the FGF2 factor

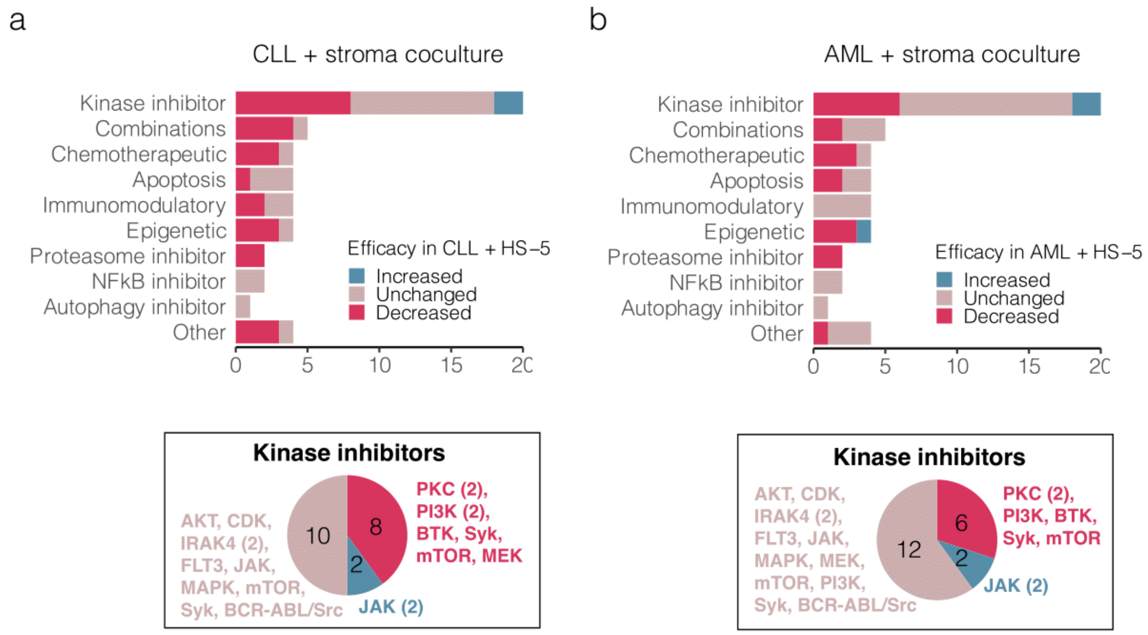


Fig. 2.16 Drug sensitivity changes in coculture relative to monoculture by compound class. The pie charts show the effect of coculture on kinase inhibitor response by target. a) Compound efficacy changes in CLL-stroma coculture by drug class. b) Compound efficacy changes in AML-stroma coculture by drug class. Data produced jointly with S. Herbst.

secreted by stromal cells. Stroma-associated decrease in BET inhibitor efficacy in leukemia has not been reported to date. It is plausible that the mechanism of BET resistance in CLL and AML is similar to that in UM, as FGF2 was also expressed in the bone marrow stromal cell line (Supplementary Fig. S4).

As mentioned in Subsection 2.1.1, B cell receptor signalling is activated by the bone marrow microenvironment – making BCR pathway inhibition an attractive therapeutic option. In the compound screen we used BTK, SYK and PI3K inhibitors that target various components of the BCR pathway (Fig. 2.17a). I found both in CLL and AML that the effect of BTK inhibition by ibrutinib was weakened in coculture (Fig. 2.17b). As Dubois et al. (2020) pointed out, BTK inhibition can be circumvented by direct activation via PI3K. By contrast, PI3K inhibitors showed only a modest decrease in efficacy when used in coculture (Fig. 2.17c), suggesting that targeting downstream components of the B cell receptor signalling pathway results in smaller efficacy reduction in coculture (Fig. 2.17a).

Some of the compounds with no significant change in efficacy target cellular processes that are unaffected by the leukemia-stroma cross-talk, while others such as the CXCR4 inhibitor, plerixafor (Supplementary Fig. S3b), block one of the major soluble axes in leukemia (Sison et al., 2014; Uy et al., 2012). Among the targeted therapeutics that had no

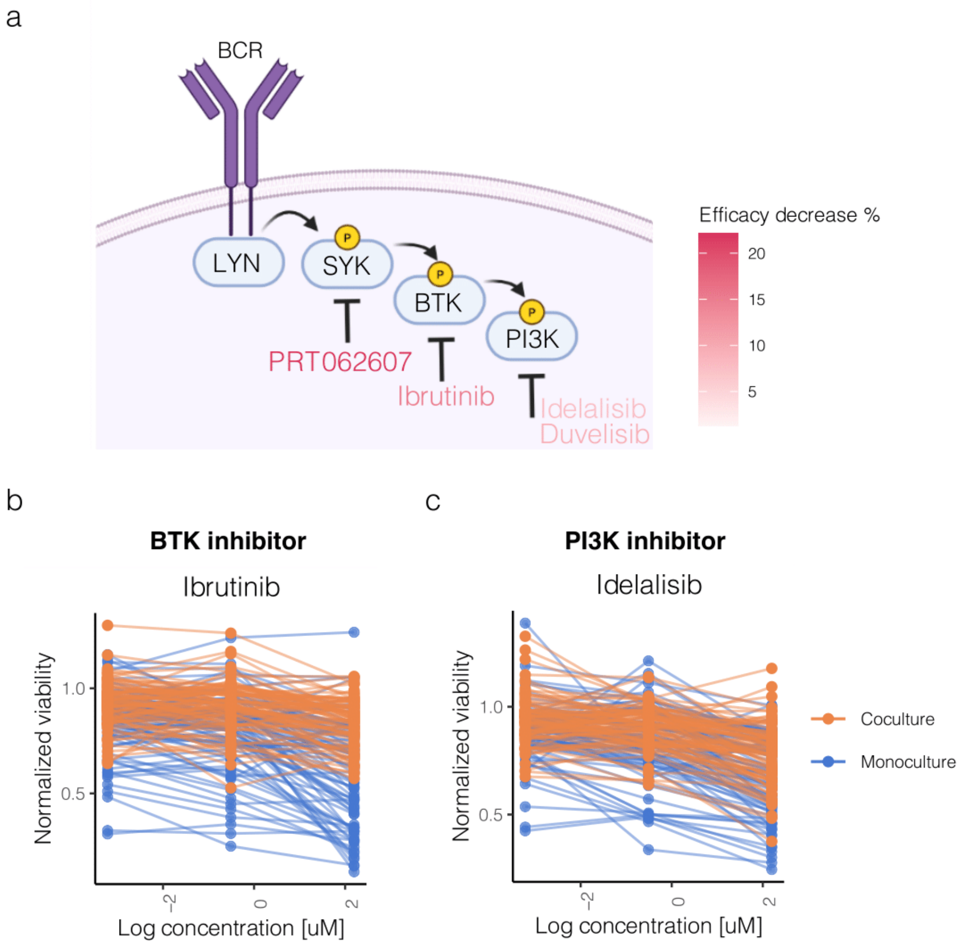


Fig. 2.17 BCR inhibitor response in CLL-stroma coculture. a) Schematic of the BCR signalling cascade. SYK, BTK and PI3K inhibitors are colored by their drug sensitivity decrease in CLL-stroma coculture. b)-c) Drug-response curves for ibrutinib and idelalisib in monoculture and coculture. Data produced jointly with S. Herbst.

altered efficacy in coculture were the inhibitors of Akt, CDK, IRAK4, FLT3, MAPK and mTOR kinases (Fig. 2.16), which suggests that these pathways are not rewired in leukemia cells upon their interaction with the bone marrow stroma.

JAK inhibitors, ruxolitinib and tofacitinib, were the only compounds in the screen with increased efficacy in CLL and AML coculture (Fig. 2.15 and 2.16). Another JAK inhibitor, pyridone 6, was equally effective in coculture and monoculture (Supplementary Fig. S3a). Protection from tyrosine kinase inhibitors in chronic myeloid leukemia (CML) was reported to be mediated by JAK1/STAT5 (Zhang et al., 2019). Moreover, STAT5 phosphorylation was implicated in increased resistance to quizartinib in AML (Dumas et al., 2019). In CLL samples the mechanistic follow-up showed that JAK inhibition abrogated STAT3 phosphorylation (Supplementary Fig. S1). Severin et al. (2019) also found that JAK2/STAT3 was responsible for the BMSC-mediated resistance to ibrutinib and provided the rationale for the JAK inhibitor + ibrutinib combination therapy.

The HS5 stromal cell line was used to emulate the bone marrow microenvironment in the compound screen. A number of drug-stroma interactions were additionally validated in coculture with primary mesenchymal stromal cells (pMSCs) extracted from 3 healthy donors (Supplementary Fig. S2) The validations in CLL + pMSC coculture were carried out for fludarabine (chemotherapeutic), JQ1 (BET inhibitor), tofacitinib (JAK inhibitor), ruxolitinib (JAK inhibitor) and idelalisib (PI3K inhibitor).

Vitamin D is an important prognostic marker in leukemia (Shanafelt et al., 2011) as well as in other diseases (Holick, 2004; Judd and Tangpricha, 2009; Plum and Deluca, 2010). Vitamin D alone had a modest cytotoxic effect and there was no significant change in sensitivity to vitamin D in stroma coculture (Supplementary Fig. S3c). We observed that adding vitamin D as an adjuvant in drug combinations with chemotherapeutics and BCR inhibitors partially diminished the stroma-mediated protection judging by the effect size (Fig. 2.18).

In summary, I observed a decrease in drug sensitivity in 52% of screened compounds in CLL coculture. The remaining compounds were mostly unaffected by the stromal microenvironment and only 2 compounds (JAK inhibitors) showed an increase in efficacy in coculture. The complete list of all probed compounds annotated by their efficacy in CLL-stroma and AML-stroma coculture is provided in Table 2.2.

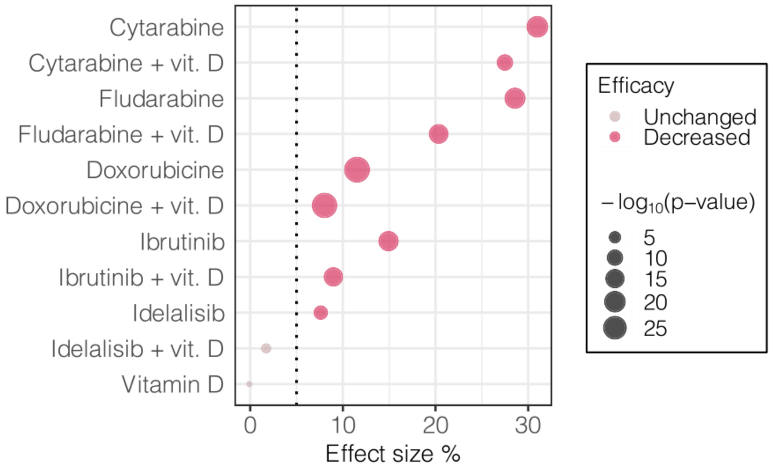


Fig. 2.18 Efficacy of vitamin D combinations in CLL-stroma coculture relative to monoculture. The effect size of stromal protection is reduced in vitamin D combinations compared with single agents. Data produced jointly with S. Herbst.

Table 2.2 Compounds by their drug efficacy in CLL and AML coculture relative to monoculture. Data produced jointly with S. Herbst.

		Compound efficacy	
Compound class	Drug	CLL + stroma	AML + stroma
AKT inhibitor	MK2206	Unchanged	Unchanged
Autophagy inhibitor	Bafilomycin A1	Unchanged	Unchanged
BCR-ABL / Src inhibitor	Dasatinib	Unchanged	Unchanged
BET inhibitor	I-BET-762	Decreased	Decreased
BET inhibitor	JQ1	Decreased	Decreased
BH3 mimetics	Obatoclax mesylate	Unchanged	Unchanged
BH3 mimetics	UMI-77	Unchanged	Decreased
BH3 mimetics	Venetoclax	Decreased	Decreased
BTK inhibitor	Ibrutinib	Decreased	Decreased
CDK inhibitor	Palbociclib	Unchanged	Unchanged
Chemotherapeutic	Cytarabine	Decreased	Decreased
Chemotherapeutic	Doxorubicine	Decreased	Decreased
Chemotherapeutic	Fludarabine	Decreased	Decreased
Chemotherapeutic	Thioguanine	Unchanged	Unchanged
CXCR4 inhibitor	Plerixafor	Unchanged	Unchanged
EZH2 inhibitor	CPI-169	Unchanged	Increased
FLT3 inhibitor	Quizartinib	Unchanged	Unchanged
HDAC inhibitor	Vorinostat	Decreased	Decreased
HSP90 inhibitor	Ganetespib	Decreased	Unchanged
IDH2 inhibitor	Enasidenib	Decreased	Unchanged
Immunomodulatory	Dexamethasone	Decreased	Unchanged
Immunomodulatory	Lenalidomide	Decreased	Unchanged
Immunomodulatory	Pomalidomide	Unchanged	Unchanged
IRAK inhibitor	IRAK1/4 Inhibitor I	Unchanged	Unchanged
IRAK inhibitor	IRAK4 inhibitor	Unchanged	Unchanged
JAK inhibitor	Pyridone 6	Unchanged	Unchanged
JAK inhibitor	Ruxolitinib	Increased	Increased
JAK inhibitor	Tofacitinib	Increased	Increased
MAPK/MEK inhibitor	LY2228820	Unchanged	Unchanged
MAPK/MEK inhibitor	Selumetinib	Decreased	Unchanged
Mdm2 inhibitor	Nutlin 3a	Unchanged	Unchanged
mTOR inhibitor	Everolimus	Decreased	Decreased
mTOR inhibitor	Rapamycin	Unchanged	Unchanged

NFkB activation inhibitor	BAY11-7085	Unchanged	Unchanged
NFkB activation inhibitor	EVP4593	Unchanged	Unchanged
Nuclear export inhibitor	Selinexor	Decreased	Decreased
PI3K inhibitor	Duvelisib	Decreased	Decreased
PI3K inhibitor	Idelalisib	Decreased	Unchanged
PKC inhibitor	Midostaurin	Decreased	Decreased
PKC inhibitor	Sotrastaurin	Decreased	Decreased
Proteasome inhibitor	Carfilzomib	Decreased	Decreased
Proteasome inhibitor	Ixazomib	Decreased	Decreased
Syk inhibitor	BAY61-3606	Unchanged	Decreased
Syk inhibitor	PRT062607	Decreased	Unchanged
Vitamin D	Vitamin D	Unchanged	Unchanged
Vitamin D combinations	Vitamin D + Cytarabine	Decreased	Decreased
Vitamin D combinations	Vitamin D + Doxorubicine	Decreased	Unchanged
Vitamin D combinations	Vitamin D + Fludarabine	Decreased	Decreased
Vitamin D combinations	Vitamin D + Ibrutinib	Decreased	Unchanged
Vitamin D combinations	Vitamin D + Idelalisib	Unchanged	Unchanged

2.3.7 Compound profiles in mono- and coculture

Morphological profiling was successfully used in compound screens to cluster similar drugs or identify on- and off-target effects (Bray et al., 2016; Loo et al., 2007; MacDonald et al., 2006; Perlman et al., 2004). I generated compound descriptors based on morphological profiles (see Subsections 2.2.4 and 2.3.3) and used these to characterize drug effects on cell/nucleus morphology and lysosomal activity.

Many compounds influenced lysosomal features (Fig. 2.19). As expected, Bafilomycin A1 induced strong changes in the lysosomal channel (Fig. 2.19). Bafilomycin A1 served as a positive control for autophagy inhibition as this drug leads to deacidification of lysosomal compartments (Mauvezin and Neufeld, 2015) resulting in decreased lysosomal intensity. There were a number of other compounds that affected lysosomal features such as proteasome inhibitor carfilzomib, BET inhibitor JQ1 and Bcl-2 inhibitor venetoclax, which suggests that these drugs could have off-target effects on autophagy.

Aside from the changes in lysosomal features and cell count, there was little variation in terms of other morphological properties (Fig. 2.19). None of the probed compounds showed a strong effect on cell shape or cytoplasmic texture (Calcein features). The strongest alterations in Calcein features were observed for JAK inhibitors.

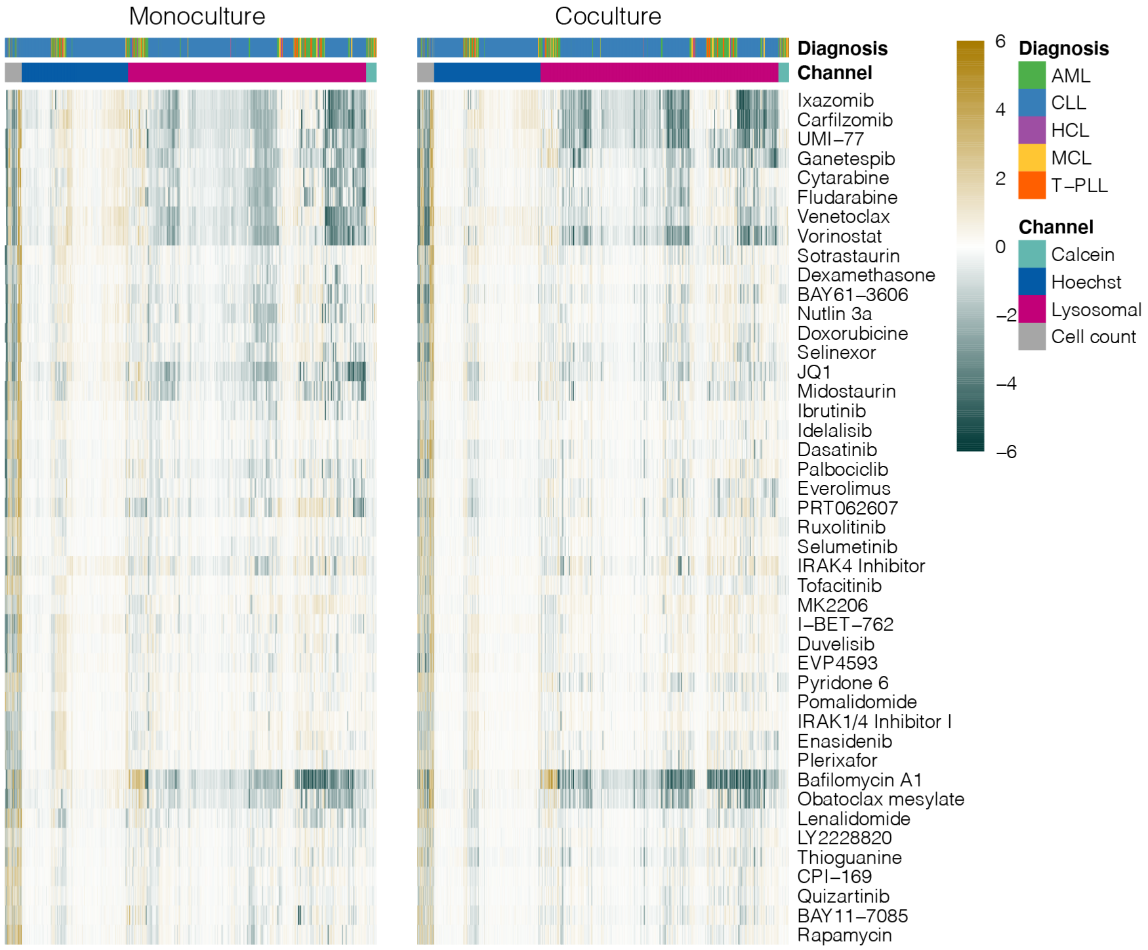


Fig. 2.19 Compound profiles aggregated across all samples. The heatmap shows morphological features normalized to controls separately for drugs in mono- and coculture. Each row consists of morphological features observed across all screened samples. The features in columns are ordered by channel/feature type (Calcein, Hoechst, lysosomal dye, cell count). Data produced jointly with S. Herbst.

Based on compound profiles in mono- and coculture, I performed hierarchical clustering analysis (Fig. 2.20). A number of important drug classes showed high similarity in terms of image-based profiles. In particular, BCR inhibitors, chemotherapeutics, JAK inhibitors, proteasome inhibitors and BH3 mimetics clustered together (Fig. 2.20).

Drug similarities observed in monoculture were generally preserved in coculture (Fig. 2.20). However, some subtle discrepancies were observed for dasatinib and sotrastaurin, which were similar to other BCR inhibitors only in monoculture (Fig. 2.20). Moreover, BH3 mimetics, venetoclax and UMI-77, had similar profiles in monoculture but not in coculture (Fig. 2.20).

While microscopy provided valuable insights into morphological changes in coculture (Subection 2.3.5), image features were not very informative for characterizing drug action beyond highlighting the changes in lysosomal activity (Fig. 2.19). Therefore it is justified to ask how important the microscopy readout was for compound clustering. For this purpose, I compared the hierarchical clustering based on morphological properties with that based on cell count alone (Supplementary fig. S5a). I found that certain clusters, such as BCR inhibitors, were driven by cell count changes and could be reproduced even without image features. However, morphological properties were crucial for the clustering of proteasome inhibitors and BH3 mimetics (Supplementary fig. S5b).

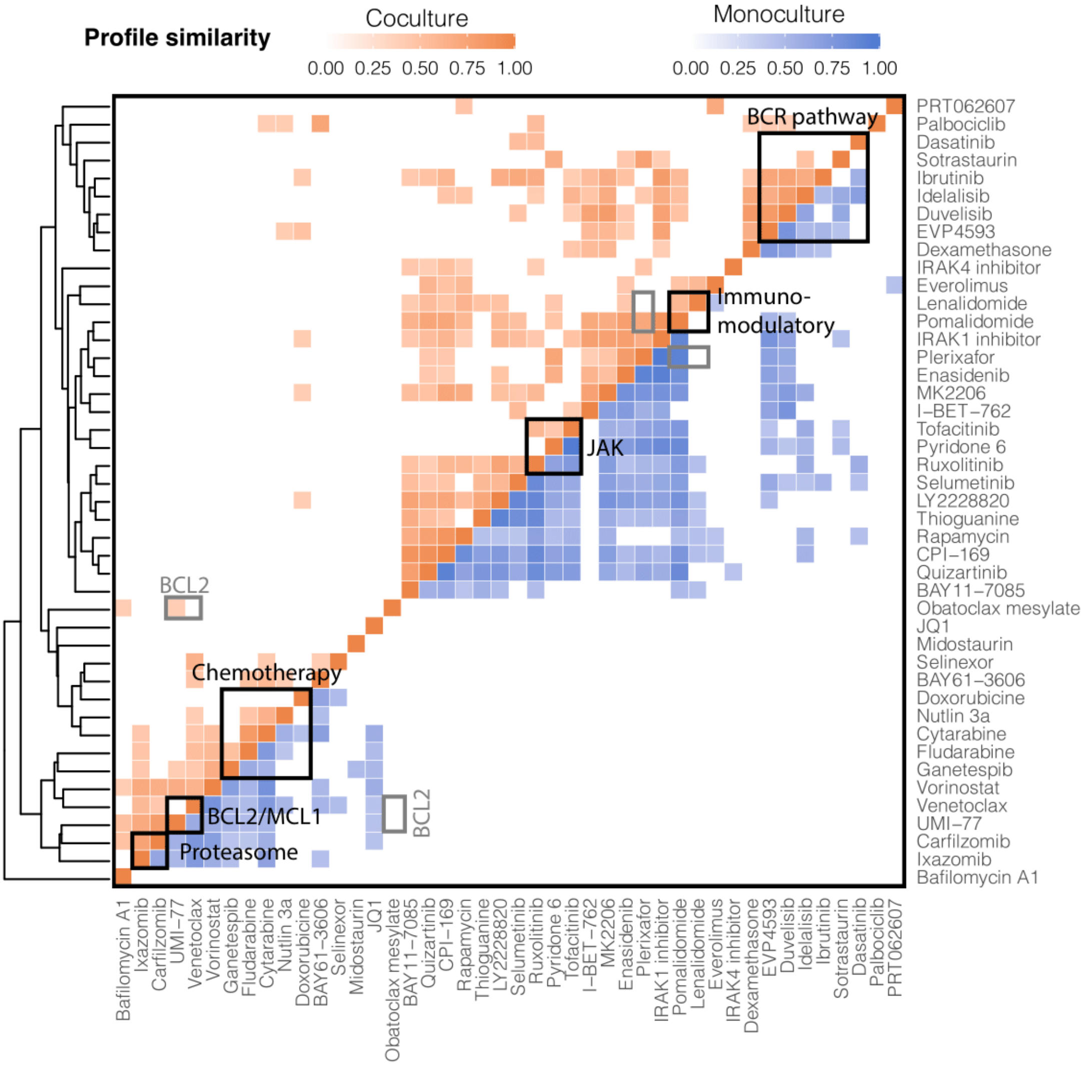


Fig. 2.20 Hierarchical clustering of compounds in the stroma-coculture screen. Joint mono- and coculture compound profiles were used to generate this clustering. Pairwise drug correlations were computed separately for mono- and coculture. Only high correlations ($r > 0.4$) are shown in the heatmap. Data produced jointly with S. Herbst.

2.3.8 Microenvironment and mutational background

Numerous studies in leukemias and lymphomas identified a number of somatic mutations associated with *ex-vivo* drug sensitivity (Barretina et al., 2012; Garnett et al., 2012; Iorio et al., 2016) and clinical outcome (Noguera et al., 2002; Shanafelt, 2009; Wiestner et al., 2003; Zenz et al., 2010a). For 80 CLL samples we had genetic information, including somatic mutations, structural variants and IGHV mutational status.

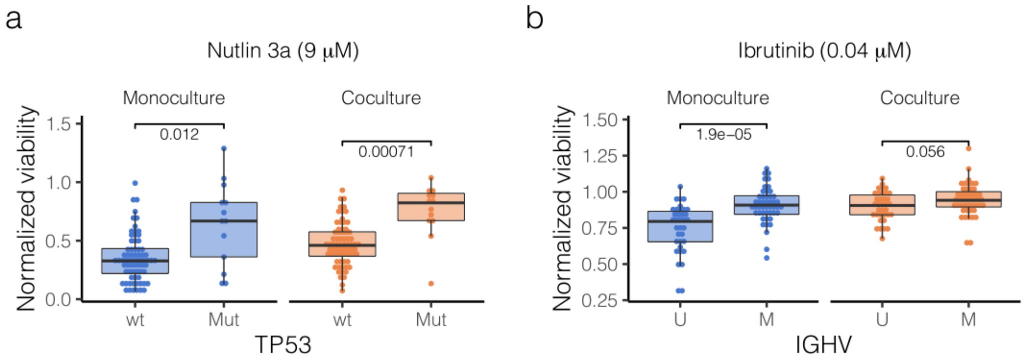


Fig. 2.21 Associations of drug response with genetic features. a) Nutlin 3a response associated with TP53 mutation b) Ibrutinib response associated with IGHV status. Data produced jointly with S. Herbst.

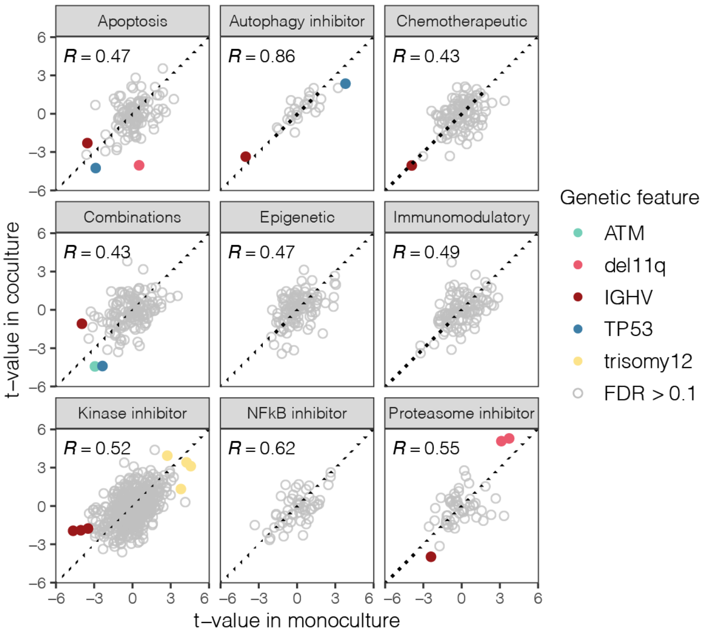


Fig. 2.22 Testing associations of *ex-vivo* drug response with genetic features. Comparison of t-statistic values in mono- and coculture by drug class. Data produced jointly with S. Herbst.

I tested associations of genetic features with drug response independently in mono- and coculture (Fig. 2.21). For each drug I applied a t-test with the null hypothesis H_0: $\mu_{Wt} = \mu_{Mut}$ (i.e. mean drug sensitivities of wildtype and mutated groups are equal). The t-values in mono- and coculture are shown in Fig. 2.22.

Generally, the sign of drug-gene association t-values was concordant between mono- and coculture (Fig. 2.22). In terms of the magnitude of association t-values, I observed some deviations among kinase inhibitors, apoptosis-targeting drugs and proteasome inhibitors (Fig. 2.22).

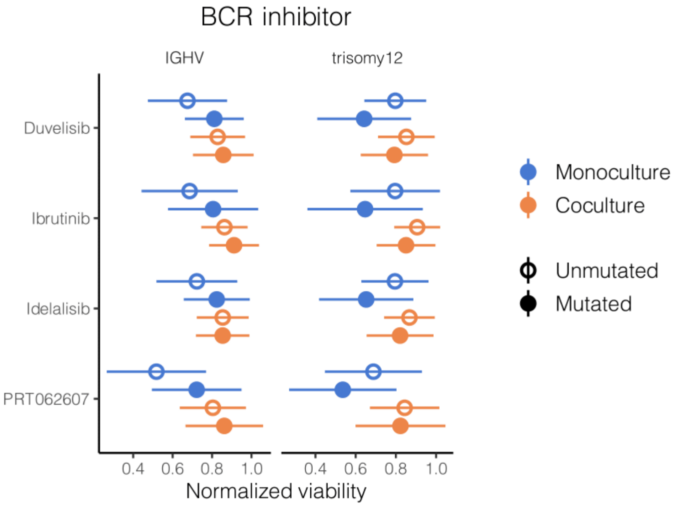


Fig. 2.23 BCR inhibitor response stratified by mutation (IGHV, trisomy12) and culture condition (mono-, coculture). Data produced jointly with S. Herbst.

BCR inhibitor associations with IGHV mutational status were weakened in coculture (Fig. 2.23). For example, ibrutinib-IGHV association (Fig. 2.21) was significant in monoculture (FDR < 0.1) but not in coculture. The IGHV-unmutated group, which is more sensitive to BCR inhibitors in monoculture, received more protection from BCR inhibitors in coculture than the mutated group (Fig. 2.23).

A similar pattern was observed for BCR inhibitor associations with trisomy12 (Fig. 2.23). The trisomy12-positive group, which was more susceptible to the action of BCR inhibitors in monoculture, received more stromal protection in coculture than trisomy12-negative samples (Fig. 2.23). Stronger protection observed in IGHV-unmutated and trisomy12-positive samples resulted in smaller effect sizes of BCR inhibitor associations in coculture (Supplementary fig. S6).

Conversely, associations of proteasome inhibitors with the deletion of the long arm of chromosome 11 (del11q) were amplified in coculture (Fig. 2.24). Coculture reduced the

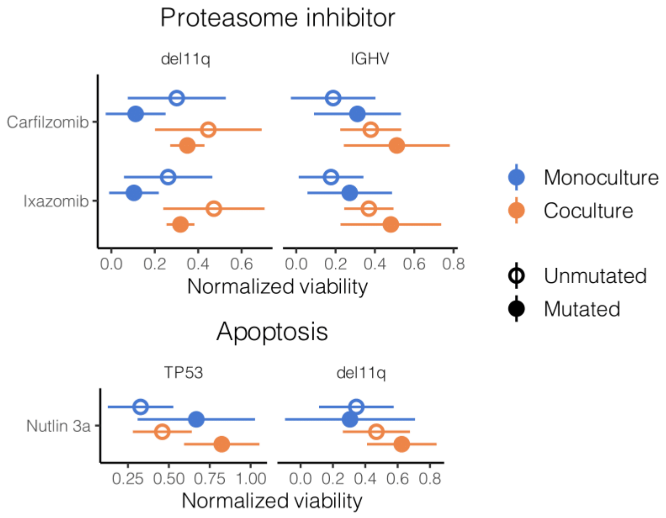


Fig. 2.24 Response to proteasome inhibitors (carfilzomib, ixazomb) and nutlin 3a stratified by mutation (del11q/IGHV, TP53/del11q) and culture condition (mono-/coculture). Data produced jointly with S. Herbst.

variance of the del11q-mutated group. Similarly, nutlin 3a - TP53 association was enhanced in coculture due to variance reduction in the TP53-mutated group (Fig. 2.24).

Overall, I found more drug-gene associations in monoculture (Fig. 2.25) showing that variance reduction was not sufficient to compensate for the decrease in effect sizes of drug-gene associations.

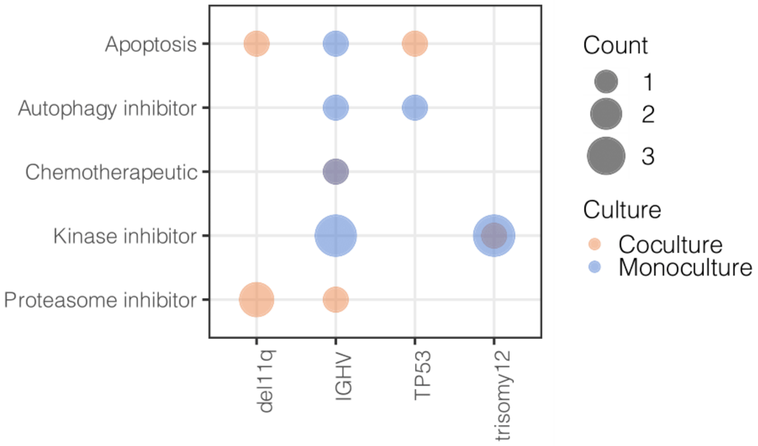


Fig. 2.25 Number of drug-gene associations in mono- and coculture (FDR < 0.1) by compound class and genetic feature. Data produced jointly with S. Herbst.

2.3.9 Using *ex-vivo* profiles to predict clinical outcomes

Dietrich et al. (2018) showed that *ex-vivo* drug sensitivities can be used to predict clinical outcomes in leukemia patients. In ongoing clinical trials, drug screening assays are being incorporated to inform therapeutic design (Liebers et al., 2019; Snijder et al., 2017). With BM microenvironment signals added, coculture drug sensitivities may better reflect *in-vivo* interactions and hence predict clinical outcomes with higher accuracy. I used the time to next treatment (TTNT) as a measure of patient outcome and performed survival analysis using mono- and coculture drug sensitivity data (Fig. 2.26).

I stratified patients by IGHV status and used maximally selected rank statistics (Lausen and Schumacher, 1992) to divide patients into sensitive and poor-responder groups based on *ex-vivo* drug response (Fig. 2.26b,c). I used the logrank test (Wellek, 1993) to assess whether time to next treatment (TTNT) differed significantly in responders and non-responders.

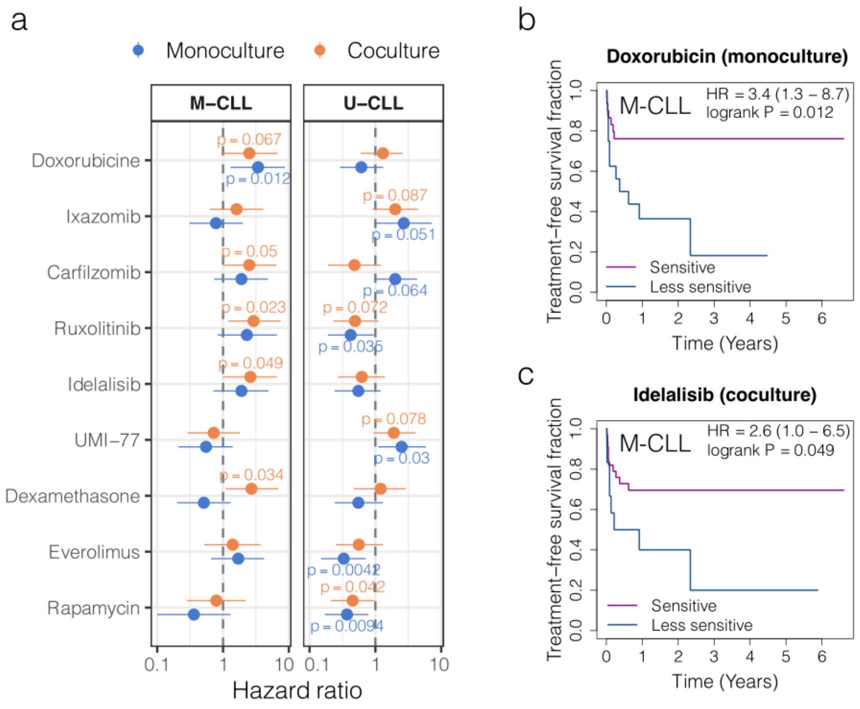


Fig. 2.26 Time-to-next-treatment (TTNT) analysis. a) Forest plot with hazard ratios stratified by IGHV status and culture condition (logrank test *p*-values). b) Kaplan-Meier curves with doxorubicin response in monoculture used for stratification of M-CLL patients. c) Kaplan-Meier curves with idelalisib response in coculture used for stratification of M-CLL patients. Data produced jointly with S. Herbst.

I found that *ex-vivo* response to clinically relevant compounds, such as doxorubicin (Fig. 2.26b) and idelalisib (Fig. 2.26c), was associated with time to next treatment (TTNT) either in mono-, coculture or both (Fig. 2.26a). There were some compounds (doxorubicin, UMI-77, everolimus, rapamycin) for which drug sensitivity in monoculture was more informative for patient stratification (Fig. 2.26a). There were, however, several counterexamples (idelalisib, ruxolitinib, dexamethasone, carfilzomib) for which coculture drug response was more indicative of outcome (Fig. 2.26a). These results suggest that coculture profiles are not more predictive of patient outcome than the drug sensitivity data in monoculture.

2.3.10 Spatial clustering induced by BiTE antibodies

In the BiTE screen, primary leukemia and lymphoma samples were screened in autologous B- and T-cell coculture. Samples were treated with bispecific antibodies (BsAb) alone and in combinations (Fig. 2.27). All cells were stained with Hoechst (nuclei) and Calcein (viability marker). T cells were stained with APC-CD3 dye and B cells were marked with PE-CD20 (Fig. 2.27).

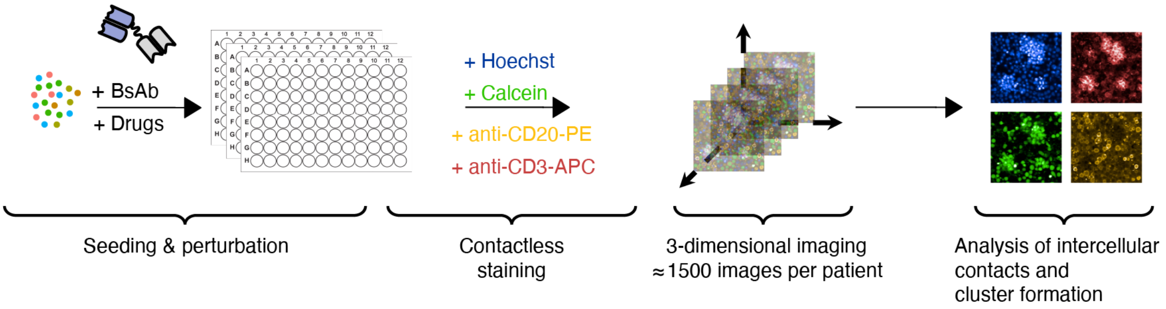


Fig. 2.27 Schematic of the BiTE screen. Graphical abstract made by T. Roider.

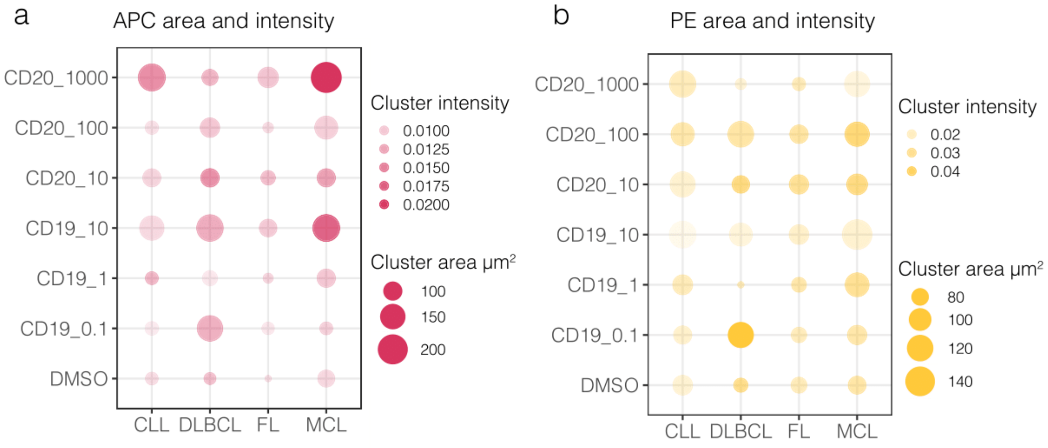


Fig. 2.28 Cell cluster area and intensity based on a) APC (T cell marker) and b) PE (B cell marker) signal by disease entity and BiTE treatment. Data produced jointly with T. Roider.

As described in Methods (Subsection 2.2.11), I identified clusters of viable cells in microscopy images. Subsequently, I compared the clusters in terms of their B and T cell composition across various probed conditions and disease entities.

Bispecific antibodies bind B and T cells simultaneously and induce cluster formation. As expected, cluster APC (T cell marker) area increased with higher concentration of bispecific antibodies in all screened disease entities (Fig. 2.28a). With increasing BsAb concentration the clusters also had higher APC intensity, which further indicates accumulation of T cells or alternatively T cell activation.

PE (B cell marker) area increased with higher antibody concentration (Fig. 2.28b), confirming that following BsAb treatment B and T cells formed clusters. Despite the observed increase in PE area, B cell marker intensity was very low at the highest BsAb concentration (Fig. 2.28b). This suggests that the PE signal at high BsAb concentrations stems from (pre-)apoptotic B cells with residual PE staining.

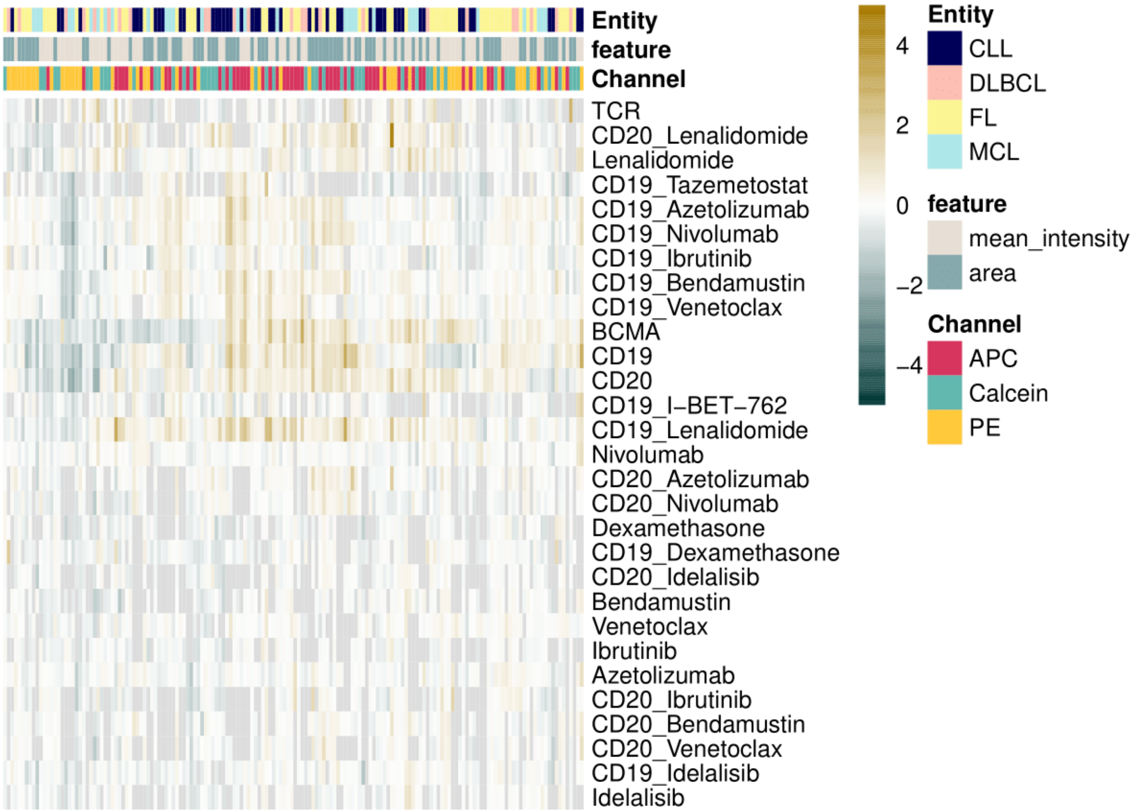


Fig. 2.29 Image-based profiling of bispecific antibodies (BsAb) and BsAb combinations. The heatmap shows mean cluster area and intensity values normalized to controls. Data produced jointly with T. Roider.

Additionally, I generated compound pofiles based on cluster morphologies (Fig. 2.29). CD19, CD20 and BCMA had similar effects on the screened samples. Of all BsAb combinations, the CD19 + Lenalidomide combination had the most promising profile with cluster APC area and intensity increased in many samples (Fig. 2.29). The remaining CD19 combinations induced weaker responses in terms of cluster APC area/intensity (Fig. 2.29), possibly due to the combination drugs' cytotoxic effect on both B and T cells.

2.4 Discussion

2.4.1 Imaging-based assays for coculture screening

When developing coculture screening assays, one has to choose between established phenotyping platforms to quantify the readout. The advantage of microscopy in the context of coculture is that high-content imaging obviates the need for proteolytic cell disassociation that can alter surface protein abundances (Donnenberg et al., 2018), potentially falsifying the results of cell sorting. However, analysis of coculture images is inherently more challenging as different cell types have to be identified, which could require additional fluorescent markers.

In the leukemia-stroma coculture screen (Subsection 2.3.1) a minimal staining palette was chosen that marked all viable and apoptotic cells, with no color channel distinguishing leukemia from stromal cells. In non-CLL entities, I used a deep learning model to detect only primary cancer cells (Subsection 2.2.3). Using supervised learning, I could achieve high precision, recall, localization accuracy and high reproducibility of the primary readout (Subsection 2.3.2), highlighting the importance of machine learning for large-scale bioimage analysis.

For object detection in leukemia-stroma coculture I leveraged transfer learning (Subsection 2.2.3) by initializing the network with the parameters learned from massive datasets of natural images. The abundance of labelled instances in the natural image domain makes this an attractive option (Kensert et al., 2019; Meijering, 2020; Morid et al., 2021).

One limitation of my analysis was that only leukemia and lymphoma cells were considered, while the morphology of bone marrow stromal cells was ignored. It is known that leukemia-stroma interactions are bidirectional (Dubois et al., 2020) and therefore future studies should focus on capturing morphology of all cell types in coculture.

In the bispecific antibody (BsAb) screen (Subsection 2.3.10) I identified clusters of B and T cells and showed that high BsAb concentrations induce cluster formation. In general, detection and quantification of cell interactions in coculture is challenging. Even though spatial point processes (Baddeley et al., 2006) can be used to model cell interactions, these methods need to be adapted for large-scale testing. It has been previously shown (Liberali et al., 2015; Snijder et al., 2009) that neighborhood properties and cell population context must be accounted for to disentangle direct and indirect effects of perturbations. Inferring and quantifying cell interactions in static coculture images is another future research direction.

2.4.2 Microenvironment and genetic context

In Subsection 2.3.5 I showed that morphological changes in leukemia cells in coculture varied depending on malignancy subtype. This indicates that microenvironment effects are disease-dependent. Not only cell morphology but also protection from spontaneous apoptosis varied across the screened samples (Subsection 2.3.4), which suggests that the BM niche may have variable pro-survival effects on an individual basis. In Subsection 2.3.8 I investigated the link between genomic profiles and bone marrow protective effects.

In chronic lymphocytic leukemia (CLL), IGHV mutational status is an important prognostic marker that helps clinicians stratify patients (Shanafelt, 2009). The unmutated IGHV group is characterized by a more aggressive course of disease and associated with poorer outcome (Rassenti et al., 2008). I observed (Subsection 2.3.4) that cells derived from U-CLL patients were more prone to spontaneous apoptosis than M-CLL cells. Viability gain in coculture was also higher in U-CLL samples (Fig. 2.10b), which corroborates the stronger dependence of U-CLL cells on the microenvironment (Coscia et al., 2011). Additionally, U-CLL cells in coculture were less sensitive to BCR inhibitors (Fig. 2.23) than M-CLL cells, demonstrating that certain genomic groups gain stronger protection from drug-induced cytotoxicity.

As it is the case with somatic mutations, chromosomal aberrations can also influence drug sensitivity. The response to certain kinase inhibitors is associated with trisomy12[8] (Dietrich et al., 2018). Trisomy12-positive CLL samples gained stronger protection from BCR inhibitors in coculture (Fig. 2.23) than those without trisomy12. Integrins (CD49d and CD38) important for cell adhesion and homing to the bone marrow are overexpressed in trisomy12-positive CLLs (Riches et al., 2014; Zucchetto et al., 2013). Increased resistance to BCR inhibitors in coculture observed in trisomy12-positive samples further confirms that trisomy12 CLLs rely stronger on the BM niche for survival.

---

[8] Presence of a third copy of chromosome 12

2.4.3 Coculture for compound screening: pros and cons

In Subsection 2.3.4 I discovered that samples with low viability in monoculture gain stronger protection from spontaneous apoptosis in coculture, which is advantageous since screening compounds in samples with low baseline viabilities may be challenging. I further showed that in such low-viability samples the dynamic range in monoculture may be severely limited (Fig. 2.11) as baseline viability places an upper limit on the drug response readout.

Spontaneous apoptosis may introduce other systematic biases. For instance, wells on the plate edge tend to have lower drug sensitivities due to solvent evaporation near the plate border, which further exacerbates spontaneous apoptosis in edge wells. Plate-positional effects occur frequently in high-throughput assays (Caicedo et al., 2017). In the leukemia-stroma coculture screen I also found evidence for edge effects in monoculture (Supplementary fig. S7). However, coculture wells were less susceptible to fluctuations in the medium (Supplementry fig. S7), which is another technical advantage of using coculture.

In Subsection 2.3.8 I showed that effect sizes of drug-gene associations were reduced (Supplementary fig. S6) due to preferential stromal protection of genomic groups that depend stronger on the bone marrow microenvironment. Despite drug response variance reduction, the number of drug-gene associations in monoculture still exceeded the number of those found in coculture (Fig. 2.25), which suggests that monoculture may be better suited for discovery of novel genetic markers associated with *ex-vivo* drug response.

Finally, I found (Subsection 2.3.9) that coculture drug sensitivity profiles were not more predictive of clinical outcome than monoculture profiles. Nevertheless, the undoubted advantage of the stroma coculture model is the ability to uncover microenvironment effects on drug action. Since spontaneous apoptosis is indicative of individual reliance on microenvironment siganls, *in-vitro* assays determining the degree of dependence on the bone marrow niche could be used to identify cases in which the microenvironment is a driver of drug resistance development.

Chapter 3

Antibiotic combination modelling

Summary

In this chapter I will describe the computational analysis pipeline that I developed for the antibiotic combination screen (Subsection 3.3.1) in Gram-positive organisms (*S. areus*, *B. subtilis* and *S. pneumoniae*). To help navigate the landscape of drug interactions in Gram-positive species (Cacace et al., 2021), I created an interactive screen explorer: https://combscreen.shiny.embl.de (Subsection 3.3.3).

To dissect mechanisms of drug interactions, I applied statistical methods to identify major determinants (Subsection 3.3.8) of drug interactions in *E. coli* and *S. typhimurium*. Furthermore, I used single-drug chemogenomic data in *E. coli* and *S. typhimurium* to predict novel drug interactions (Subsection 3.3.10) and to investigate the functional role of uncharacterized genes (Subsection 3.3.9). I assessed different algorithms and modelling strategies (Subsection 3.2.13 – 3.2.14) for drug interaction prediction using chemogenomic data.

In addition to using machine learning to classify drug combinations, I performed feature selection and ranking to identify genes important (Subsection 3.3.11) for drug interaction prediction in *E. coli* and *S. typhimurium*. I obtained feature importance for individual predictions (Subsection 3.2.16) and identified genes important for β-lactam synergies and for antagonisms involving DNA gyrase targeting compounds (Subsection 3.3.11).

Finally, using combination data in *E. coli*, *S. typhimurium* and *P. aeruginosa*, I constructed a directional network of drug interactions (Subsection 3.3.12) in Gram-negative species.

E. Cacace conducted the high-throughput combinatorial screen in Gram-positive bacteria and experimental methods are described in detail in (Cacace, 2021). Unless stated otherwise, I performed all computational analysis presented in this chapter, including drug interaction modelling and prediction using chemogenomic data.

3.1 Introduction

Antibiotic resistance has become one of the most urgent threats to the public health (Bush et al., 2011; Frieri et al., 2017; Neu, 1992). The rapid emergence of resistance mechanisms dwarfs the efforts in antimicrobial drug development (Aslam et al., 2018; Brown and Wright, 2016). Drug interactions, in particular synergistic drug combinations of approved compounds, can be leveraged in the fight against antibiotic resistance (Tyers and Wright, 2019). While single-agent therapies are susceptible to resistance development (Silver and Bostian, 1993; Toprak et al., 2012), multidrug approaches can be instrumental in overcoming antibiotic resistance or even slowing resistance evolution under certain conditions (Baym et al., 2016; Chait et al., 2007; Hernando-Amado et al., 2020; Imamovic and Sommer, 2013). Moreover, there is some evidence that drug interactions are preserved in a multitude of genetic backgrounds (Chevereau and Bollenbach, 2015).

Only a handful of well studied synergies have been mechanistically dissected (Sullivan et al., 2020). In recent years, high-throughput combinatorial screening played a major role in uncovering novel drug interactions (Brochado et al., 2018; Mott et al., 2015; Ramón-García et al., 2011; Sun et al., 2016b). Mode-of-action driven design of drug combinations remains feasible only for a small number of compound classes with elucidated synergy mechanisms.

To bypass complex mechanistic studies, researchers repurposed single-compound features to predict novel drug interactions using machine learning (Chandrasekaran et al., 2016; Mason et al., 2017; Wildenhain et al., 2015). Compound features used for drug interaction prediction could stem from chemogenomic profiling, which combines chemical and genetic perturbations (Subsection 1.5).

3.1.1 High-throughput drug combination screening in bacteria

High-throughput screening of antibiotic combinations shed light on the abundance of drug interactions and their cross-species conservation. In the largest screen in Gram-negative bacteria Brochado et al. (2018) profiled ≈ 3000 drug combinations in *E. coli*, *S. typhimurium* and *P. aeuruginosa* and confirmed that drug interactions in bacteria are rare, with $< 10\%$ of drug pairs interacting synergistically and $\approx 10\%$ antagonistically.

Most of the drug interactions reported in the literature are heavily biased towards synergies. Brochado et al. (2018) found that antagonisms were more common in Gram-negative bacterial species, highlighting the need for systematic screening of drug combinations. The degree of conservation among Gram-negative species was low with only 5% of drug interactions shared by all 3 species. The conservation rates across species varied between 13% and

25%, suggesting the possibility of designing species-specific drug combinations targeting pathogens and minimizing collateral damage to commensal organisms.

The experimental load of drug combination screening in several bacterial species is daunting. Screening pairwise drug combinations has quadratic complexity $O(n^2)$ in terms of drug library size n alone. Additionally, multiple concentration ratios and replicates have to be included to boost the statistical power and ensure reproducibility. Therefore drug interaction prediction is an attractive method to expand the currently known set of drug interactions based on single-compound features, such as chemogenomic profiles (Chandrasekaran et al., 2016; Wildenhain et al., 2015) or molecular fingerprints (Mason et al., 2017).

Due to species specificity, drug interaction prediction models cannot fully replace experimental screening as in each novel species initial training data is required for prediction. For organisms with large reference sets of drug interactions predictive modelling is a promising approach to guide the mechanistically informed design of drug combinations.

3.1.2 Mechanisms of drug interactions in bacteria

According to the parallel pathway inhibition model, the combined action of drugs produces an effect corresponding to the epistatic interaction between the targeted genes in parallel pathways (Cokol et al., 2011; Yeh et al., 2009). Numerous examples of anticancer synergistic combinations disrupting parallel signalling pathways are known (Dent et al., 2009; Dietlein et al., 2015; Rao et al., 2005; Woo et al., 2017) and some antibiotic synergies arise due to the parallel pathway inhibition (Sullivan et al., 2020).

Under the bioavailability model, synergies occur when the adjuvant drug increases the intracellular concentration of the potentiated drug (Cokol et al., 2011). Conversely, antagonisms arise due to the reduced uptake or increased efflux of the partner drug. There is some evidence (Brochado et al., 2018) that modulation of intracellular concentration underlies many antibiotic antagonisms.

Transmembrane proteins involved in xenobiotic transport play an important role in antibiotic resistance (Langton et al., 2005; Levy, 2002; Markham and Neyfakh, 2001). Multidrug efflux pumps are protein complexes with a broad substrate affinity range responsible for drug extrusion (Sun et al., 2014). Antagonistic interactions can occur due to efflux pump activation by one of the combination drugs. For instance, salicylate, a plant hormone, is known to induce the upregulation of the *acrAB-tolC* efflux system in *E. coli* (Martin and Rosner, 1995) and many antagonistic interactions with acetylsalicylic acid were observed in both Gram-negative and Gram-positive organisms (Brochado et al., 2018; Cacace, 2021).

Increasing the permeability of the bacterial cell wall lowers the drug barrier. β-lactams inhibit peptidoglycan synthesis and thereby permeabilize the Gram-positive bacterial cell wall and disrupt the murein layer in Gram-negatives. β-lactam - aminoglycoside synergy has been widely used by clinicians (Miller et al., 1987; Moellering et al., 1986).

Systematic investigations of drug interactions are carried out by probing drug combinations in gene deletion libraries, similar to chemogenomic profiling of single agents. Due to the genome-wide scale of these studies, the number of drug interactions interrogated in this manner is very limited. Chevereau and Bollenbach (2015) studied pairwise drug combinations of 6 antibiotics in *E. coli*. One of the central findings was that all probed drug interactions were preserved in the majority of gene deletion strains and only a small subset of gene deletions, governing these drug interactions, resulted in a loss of interaction. Furthermore, the authors identified polysaccharide and ATP biosynthesis as key biological processes underlying the probed antibiotic interactions.

3.1.3 Supervised learning for drug interaction prediction

Ensemble methods, such as random forests, have been successfully used for drug interaction prediction based on chemogenomic profiles (Chandrasekaran et al., 2016; Wildenhain et al., 2015). Random forests (Breiman, 2001) can be used for both regression and classification tasks. Predictions are aggregated across an ensemble of N estimators, with the simple mean used for regression and the majority vote for classification (Zhang and Ma, 2012, Chapter 5):

$$\hat{y} = \frac{1}{N}\sum_{i=1}^{N}\hat{f}_i(x) \qquad \text{(regression /classification)}$$

$$\hat{y} = \arg\max_{c}\sum_{i=1}^{N} I(\hat{f}_i(x) = c) \qquad \text{(classification)}$$

Each estimator in a random forest receives a bootstrap sample of the original data and a random subset of features (Zhang and Ma, 2012, Chapter 5). There are a number of hyperparameters that can be adjusted such as the number of estimators N, individual tree properties (depth, number of leaf nodes) or degree of randomization. It is common to use cross-validation to choose the best hyperparameter combination empirically (Probst et al., 2019).

Another popular class of ensemble methods used for both classification and regression is based on boosting (Freund and Schapire, 1997). Unlike random forests, boosted tree ensembles are not composed of independent learners and at each iteration more weight is given to the training instances misclassified by the previous estimator (Zhang and Ma, 2012, Chapter 2). In boosted ensembles, a base learner $h_t(x)$ at iteration t is added to the previous function approximation with weight ω_t (Natekin and Knoll, 2013):

$$\hat{f}_t(x) \leftarrow \hat{f}_{t-1}(x) + \omega_t h_t(x)$$

In gradient boosting, the base learners are parameterized as $h(x, \theta)$ with $\hat{\theta}$ to be determined by minimization of the loss function $\mathcal{L}$. The gradient descent algorithm is applied to obtain optimal estimates of $\hat{\omega}$ and $\hat{\theta}$ at each iteration (Natekin and Knoll, 2013):

$$\hat{\omega}_t,\ \hat{\theta}_t = \underset{\omega,\theta}{\arg\min}\, \mathcal{L}(y_i,\ \hat{f}_{t-1}(x) + \omega h_t(x, \theta))$$

Due to its scalability and flexibility, XGBoost (Chen and Guestrin, 2016) has become one of the most widely used variants of regularized gradient boosting. Recently, newer implementations such as CatBoost (Prokhorenkova et al., 2018) and LightGBM (Ke et al., 2017) advanced the computational efficiency of gradient tree boosting even further (Bentéjac et al., 2021).

Deep learning models achieved near human-level performance in a number of machine learning problems (Lecun et al., 2015; Taigman et al., 2014; Yasaka and Abe, 2018). Various domains from computer vision to natural language processing adopted neural networks as default supervised learning models (Lecun et al., 2015). A single neuron transforms input data using a linear transformation W followed by a non-linear activation function σ to generate activation a (Goodfellow et al., 2016, Chapter 6):

$$a = \sigma(WX)$$

When the sigmoid function $\sigma(x) = (1 + e^{-x})^{-1}$ is used , a single neuron performs logistic regression. In feed-forward networks a vector of activations a^l in layer l can be computed in terms of activations a^{l-1} in the preceding layer:

$$a^l = \sigma(W^{l-1} a^{l-1})$$

The power of neural networks comes from the composition of successive linear and non-linear transformations that can be used to construct universal function approximators (Goodfellow et al., 2016, Chapter 6). I trained a neural network for drug interaction prediction and compared its performance with other algorithms. Although never applied to chemogneomic data, neural networks were previously used to predict *in-vivo* drug interactions based on biomedical literature data (Sahu and Anand, 2018; Zhang et al., 2018).

In the DREAM community challenge (Menden et al., 2019) independent research teams used multiomic and single-drug data to predict drug interactions in various cancer cell lines. The comparison of prediction methods revealed that performance can be improved by integrating gene-gene interaction data and drug target-related pathway information. By contrast, the choice of machine learning algorithm played a subordinate role for prediction accuracy.

In another DREAM challenge (Bansal et al., 2014), researchers used transcriptional profiles of single-compound treatments to predict pairwise drug interactions in an unsupervised manner. In absence of labelled training data, prediction methods rely on drug profile similarity or dissimilarity. Submissions that used similarity to predict synergies performed slightly better in terms of sensitivity and probabilistic concordance index, whereas neither similarity nor dissimilarity hypotheses were useful for predicting antagonisms (Bansal et al., 2014).

Both DREAM community challenges (Bansal et al., 2014; Menden et al., 2019) highlighted the difficulty of drug interaction prediction based on single-drug properties. Moreover, these benchmark studies emphasized the importance of integrating structural data, genetic network information and pathway annotations.

3.1.4 Feature selection methods

Most machine learning algorithms operate as black-box predictors and little can be said about which features are driving predictions. One of the goals of feature selection is to make machine learning models more interpretable (Saeys et al., 2007). Some other objectives of feature selection methods may include (Saeys et al., 2007; Tang et al., 2014): i) eliminating redundancy and reducing the number of features ii) choosing only those features that are associated with the target variable iii) ranking features by their model-specific relevance iv) improving generalizability and preventing overfitting.

Statistical tests can be used to select features associated with the target variable. These feature selection methods are commonly referred to as filter methods (Saeys et al., 2007). For continuous X and Y, correlations between explanatory and target variables can be computed to select the top correlated features. Similarly, for discrete X and Y, one can use the χ^2-test to choose variables with high χ^2-statistic values (Saeys et al., 2007).

Many supervised learning algorithms are capable of internal feature ranking. Tree-based models such as random forests and boosted trees can output feature importance based on mean reduction in Gini impurity (Archer and Kimes, 2008). While computationally efficient, embedded feature selection methods are limited to specific models and may lead to overfitting (Saeys et al., 2007).

Wrapper methods are iterative procedures that grow a feature subset incrementally or eliminate one feature at a time until a termination criterion is reached (Saeys et al., 2007). For instance, recursive feature elimination (RFE) is a commonly used approach in high-content imaging screens (Caicedo et al., 2017). Wrapper feature selection methods often consider correlations among predictor variables to reduce redundancy. Some of the limitations of wrapper methods are their computational cost and model-specificity (Saeys et al., 2007).

While feature selection provides some intuition as to which features are important for the model as a whole, they do not explain individual predictions, e.g. why a particular drug combination is classified as synergistic or antagonistic. Shapley values (Shapley, 1953), a game-theoretic concept, are used in machine learning to determine local feature importances for individual predictions (Molnar, 2019, Chapter 5).

Lundberg and Lee (2017) showed that Shapley values uniquely satisfy the conditions of local accuracy, missingness and consistency. Furthermore, Lundberg and Lee (2017) recast the problem of Shapley value estimation as solving weighted linear regression, which made kernel SHAP much more computationally efficient than previous estimation methods. SHAP[1] is currently the most popular approach for model explainability as it is model-agnostic and can be used for ensemble (Lundberg et al., 2020) and deep learning (Lundberg and Lee, 2017) models.

3.2 Methods

3.2.1 Drug interaction modelling in the Gram-positive screen

I quantified drug interactions using the Bliss independence model (Greco et al., 1995). The expected combination fitness for drugs r and d according to the Bliss model is the product of single-drug effects f_r and f_d. Bliss score of a drug combination estimates the deviation of the measured combination response f_{rd} from the expected value f_{rd}^{Bliss}:

$$\varepsilon = f_{rd} - f_{rd}^{\text{Bliss}} = f_{rd} - f_r \cdot f_d$$

By definition, synergistic drug pairs have negative Bliss scores ($\varepsilon < 0$), while antagonistic combinations have positive interaction scores ($\varepsilon > 0$).

---

[1] https://github.com/slundberg/shap

3.2.2 Single-drug and combination effects in the Gram-positive screen

In the antibiotic combination screen in Gram-positive species, combination and single-drug effects on bacterial growth were assessed using optical density (OD) at $\lambda = 595$ nm. OD values were obtained over the course of > 7 hours and the OD value at the first stationary phase point of the growth curve was chosen as the screen readout. All OD values were normalized to no-drug control values. Experimental methods are described in detail in Cacace (2021).

I estimated single-drug effects by minimizing the sum of squared deviations of measured combination OD values f_{rd} from expected combination fitness $(f_r \cdot f_d)$:

$$\{f_r, f_d\} = \arg\min \sum_{r,d} (f_{rd} - f_r \cdot f_d)^2$$

I chose estimated single-drug effects over experimentally determined ones since using the latter was introducing systematic biases as explained in Cacace (2021).

3.2.3 Statistical analysis of drug interactions in Gram-positives

Each drug combination was screened at 3×3 concentration ratios with 2 well replicates and at least 2 biological replicates (Cacace, 2021). Thus for each combination we have a vector $\vec{\varepsilon}$ of Bliss scores estimated across concentration ratios and replicate measurements.

I determined the statistical significance of observed synergies and antagonisms using a resampling procedure described in (Brochado et al., 2018). Briefly, a Wilcoxon signed rank test was performed to compare the vector of Bliss scores $\vec{\varepsilon}$ of a drug combination with a sample $\vec{e}$ drawn randomly from the complete Bliss score distribution (null distribution). The procedure was iterated 10,000 times and the p-value of each combination was estimated as:

$$p = \frac{\sum_{i=1}^{N} (p_i^{wilc} > 0.1) + 1}{N+1}, \; N = 10^4$$

All p-values were corrected for multiple testing using Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

The effect size of drug interactions was quantified as the first $Q_1(\vec{\varepsilon})$ or third quartile $Q_3(\vec{\varepsilon})$ of a combination's Bliss score vector:

$$\hat{\varepsilon} = \begin{cases} Q_1(\vec{\varepsilon}) & \text{if } |Q_1(\vec{\varepsilon})| > |Q_3(\vec{\varepsilon})| \\ Q_3(\vec{\varepsilon}) & \text{if } |Q_3(\vec{\varepsilon})| \geq |Q_1(\vec{\varepsilon})| \end{cases}$$

Finally, I defined drug interactions as follows:

$$\hat{y} = \begin{cases} \text{synergy} & \text{if } q < 0.05 \text{ and } \hat{\varepsilon} < -0.1 \\ \text{antagonism} & \text{if } q < 0.05 \text{ and } \hat{\varepsilon} > 0.1 \\ \text{additive} & \text{else} \end{cases}$$

3.2.4 Combinatorial screen explorer

For data exploration in the Gram-positive screen, I developed an interactive web application in R using `shiny v1.5.0`. The source code of the Shiny app is available on EMBL Gitlab: https://git.embl.de/vkim/docker_combscreen.git.

3.2.5 Pairwise drug combination data in bacteria

Pairwise drug combination data in *E. coli*, *S. typhimurium*, and *P. aeruginosa* were obtained from the high-throughput screen (Brochado et al., 2018) in Gram-negative bacterial species. Using a similar experimental setting, Cacace (2021) screened drug combinations for antagonisms and synergies in Gram-positive species. An overview of drug combination data is provided in Table 3.1.

Species	Strain	*n* drugs	Source
E. coli	BW25113 iAi1	68	(Brochado et al., 2018), Gram-negative species
S. typhimurium	ST14028 LT2		
P. aeruginosa	PA01 PA14		
S. aureus	Newman DSM20231	65	(Cacace, 2021), Gram-positive species
B. subtilis	168		
S. pneumoniae	D39V		

Table 3.1 Tested drug pairs in Gram-negative and Gram-positive bacterial strains.

3.2.6 Chemogenomic data in *E. coli* and *S. typhimurium*

Chemogenomic data in *E. coli* (K12 strain) was available from the study by Nichols et al. (2011) that probed 90 compounds in $\approx$ 4000 gene deletion strains. Chemical genetics data in *S. typhimurium* (ST14028 strain) was generated by Pfalz (2017). As in *E. coli*, 140 compound and physico-chemical perturbations were probed in $\approx$ 4000 gene deletions in *Salmonella*.

For the subsequent analysis I selected only those drugs that were shared between chemogenomic (Nichols et al., 2011; Pfalz, 2017) and combinatorial screen (Brochado et al., 2018) datasets, resulting in 45 compounds remaining in *E. coli* and 64 in *S. typhimurium*. In both organisms I removed gene deletions with no significant S-scores (at FDR = 0.05). 1302 gene deletions in *E. coli* and 1146 in *S. typhimurium* had at least one significant S-score with any of the drugs probed in the cobminatorial screen by Brochado et al. (2018).

3.2.7 Multiomic factor analysis (MOFA) on chemogenomic and drug interaction data

I used multiomic factor analysis (MOFA) (Argelaguet et al., 2018) to integrate drug interaction and chemogenomic data in two Gram-negative species: *E. coli* (EC) and *S. typhimurium* (ST). 4 omic views — drug interactions in EC and ST as well as chemogenomic data in EC and ST — were provided as input to MOFA (`v1.4.0`) with the threshold of $> 1\%$ variance explained per factor (`DropFactorThreshold = 0.01`) and convergence tolerance of 0.01 (`tolerance = 0.01`).

After 10 random initializations, I chose the MOFA model with the highest ELBO. The selected model produced 10 latent factors characterizing 68 compounds. I performed a hierarchical clustering in the latent space using `pheatmap v1.0.12` package.

3.2.8 Compound similarity in Gram-negative species

Chemogenomic similarity between drugs a and b was defined as the Pearson correlation between chemogenomic profiles X_a and X_b. I evaluated chemogenomic similarity separately in *E. coli* and *S. typhimurium*.

Interaction profile similarity between drugs a and b was quantified as the Pearson correlation between Bliss score profiles $\vec{\varepsilon}_a$ and $\vec{\varepsilon}_b$ observed across all screened conditions and bacterial strains. I computed interaction profile similarity jointly for all 6 Gram-negative strains (see Subsection 3.2.5).

3.2.9 Combination chemogenomic profiles in *E. coli* and *S. typhimurium*

Single-drug chemogenomic profiles are composed of continuous S-scores measuring drug-gene interactions. First, I thresholded single-drug profiles so that only significant S-scores had non zero values:

$$X_{dg} = \mathrm{sign}(S_{dg}) \cdot \mathbb{1}(q_{dg} < 0.05)$$

where q_{dg} is the adjusted p-value of the interaction between drug d and gene g. Discrete drug-gene states X_{dg} were in the range $\{-1, 0, 1\}$.

I generated drug combination profiles using single-drug chemogenomic data. Given thresholded chemogenomic profiles X_{ag} and X_{bg} of drugs a and b in a gene deletion g, I encoded the drug combination ab as follows:

$$X_{abg} = \begin{cases} X_{ag} + X_{bg} & \text{if } \mathrm{sign}(X_{ag}) = \mathrm{sign}(X_{bg}) \\ (\pm) & \text{else} \end{cases}$$

Thus combination-gene states X_{abg} take on values in the range $\{-2, -1, 0, 1, 2, \pm\}$. The resulting matrix of combination chemogenomic profiles $\mathbf{X}$ had dimensions $N_{comb} \times N_{gene}$.

Combination chemogenomic matrix $\mathbf{X}$ consists of categorical variables. To ensure that discrete combination-gene states are not treated as continuous variables, I used the one-hot encoding scheme to transform $\mathbf{X}$ into a binary matrix of form:

$$\mathbf{X}^{\text{one-hot}} = \begin{bmatrix} gene = -2 & gene = -1 & gene = 0 & gene = 1 & gene = 2 & gene = \pm \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

3.2.10 Interaction-associated genes in *E. coli* and *S. typhimurium*

Drug combinations can be antagonistic, synergistic or neutral:

$$y \in \{\text{antagonism, synergy, neutral}\}$$

To identify genes associated with antagonisms and synergies, I applied a χ^2-test to each column of the combination chemogenomic matrix $\mathbf{X}^{\text{one-hot}}$ (see Subsection 3.2.9), with the null hypothesis that combination type y is independent of combination-gene state $\mathbf{X}^{\text{one-hot}}_{:,g}$

for every gene g:

$$H_0 : y \perp\!\!\!\perp \mathbf{X}^{\text{one-hot}}_{\cdot,g} \ \forall g$$

To account for multiple testing, I used the FDR cutoff of 0.1, with only significant associations reported. I conducted statistical tests independently in *E. coli* and *S. typhimurium*.

3.2.11 Gene ontology (GO) analysis

I performed gene ontology enrichment analysis on interaction-associated genes (see Subsection 3.2.10) in *E. coli* and *S. typhimurium*. To assess the statistical significance of GO enrichment, I used Fisher's exact test. The GO annotation map in *E. coli* was provided by G. Kritikos. Gene ontology annotation in *Salmonella* was generated by B. El Debs. For the sake of interpretability I considered only 'Biological Process' terms in this analysis.

Since only non-essential genes were probed by Nichols et al. (2011) and Pfalz (2017), I restricted the GO background to non-essential genes. Input genes with missing GO annotation were excluded from the analysis. In each species, I corrected GO enrichment p-values for multiple testing using Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995).

3.2.12 Gene-gene correlation networks

I extracted protein association networks for *E. coli* (K12 strain) and *S. typhimurium* (ST14028 strain) from `STRINGdb v11`. I used the combined evidence score threshold of ≥ 0.7 to remove low-confidence network links.

To embed uncharacterized genes in the genetic network, I constructed gene-gene correlation networks based on chemical genetics data. I augmented protein association networks by adding links from gene-gene correlation networks with correlation thresholds of $r > 0.5$ in *E. coli* and $r > 0.45$ in *S. typhimurium*.

To understand the functional role of uncharacaterized genes, I visualized ego networks centered on uncharacterized genes with their first-degree neighbors arranged in a circular layout. The guilt-by-association principle aids in functional mapping (Costanzo et al., 2010).

3.2.13 Supervised models for drug interaction prediction

To predict combination type $y \in \{\text{antagonism, synergy, neutral}\}$, I used the combination chemogenomic matrix $\mathbf{X}^{\text{one-hot}}$ (see Subsection 3.2.9) as predictor variables. As this was a multiclass problem ($N_{class} = 3$), I used the one-versus-rest strategy that effectively trains N_{class} binary classifiers to learn one class at a time versus all other classes.

To find the best model, I trained and benchmarked 3 supervised learning algorithms: random forest (`scikit-learn 0.21.3`), XGBoost (`xgboost 0.90`) and a deep neural network (DNN) (`Keras 2.3.1`). I evaluated the performance of supervised models based on average precision (AP) and area under the ROC curve (AUCROC).

As described in Subsection 3.2.14, I used 20-fold cross-validation to choose the best hyperparamter set. Training, cross-validation, model selection and prediction on the test set were carried out separately for *E. coli* and *S. typhimurium*.

3.2.14 Hyperparameter tuning and model selection for drug interaction prediction

I adapted the k-fold cross-validation strategy for drug combination data. In each iteration $m = 15$ drugs were randomly chosen so that all $\binom{m}{2}$ combinations of these were masked from the training set. The model was trained on $\binom{n-m}{2}$ chemogenomic profiles and validated on $\binom{m}{2}$ drug pairs. For each hyperparameter combination I used $k = 20$ cross-validation folds.

To find the best model hyperparameters for random forest, XGBoost and neural networks (see Subsection 3.2.13), I performed a grid search. The hyperparameter space that I explored is shown in code block 3.1.

```
# parameter grid for random forest classifier
# here 9 * 3 * 8 * 8 * 11 = 19008 combinations of hyperparameters
RF_grid = {"max_depth": list(range(2,10)) + [None],
           "n_estimators": [200, 500, 1000],
           "min_samples_split": range(2,10),
           "min_samples_leaf": range(2,10),
           "class_weight": [{0: 1, 1: i} for i in range(1,10)] +
                           ['balanced', 'balanced_subsample']
           }

# XGB has 20 * 20 * 8 * 3 * 10 = 96000 combinations of hyperparameters
XGB_grid = {"learning_rate": np.logspace(-2,np.log10(0.9), 20),
            "colsample_bytree": np.linspace(0.1,1,20),
            "max_depth": range(2,10),
            "n_estimators": [200, 500, 1000],
            "scale_pos_weight": range(1,10)
            }

# DNN has 10 * 10 * 5 * 5 * 5 * 10 * 10 = 2880 combinations of hyperparameters
DNN_grid = {"learning_rate_deep": [0.001,0.005,0.01,0.1,1],
```

```
                "layers": [1,3,5,7],
                "nodes": [16,32,64,128],
                "dropout": [0.1,0.3,0.5,0.7],
                "steps": [32,64,128],
                "epochs": [400, 600, 800]}
```

Code 3.1 Hyperparameter combinations were assessed using 20-fold cross-validation. Each cross-validation fold consisted of $\binom{m}{2}$ combinations withheld from the training set, with $m = 15$ randomly selected drugs.

3.2.15 Global feature importance in *E. coli* and *S. typhimurium*

For XGBoost classifier (`xgboost 0.90`, see Subsection 3.2.13), I obtained global feature importance to understand what genes were driving drug interaction predictions. From each cross-validation run I extracted 20 features with the highest variable importance. After aggregating top predictive features across all cross-validation folds, I ranked features by their occurrence in the list, with genes appearing in many cross-validation runs having higher global feature importance. This ranking was performed separately for antagonism and synergy predictions in each species.

3.2.16 SHAP values in *E. coli* and *S. typhimurium*

To obtain feature importance for individual drug interaction predictions, I used `SHAP v 0.39.0` Python package. From the selected XGBoost model (see Subsection 3.2.13) I extracted SHAP values for each training instance. SHAP values indicate how much each feature contributed to the prediction score (Lundberg and Lee, 2017) of a drug combination.

As I used one-versus-rest classification (see Subsection 3.2.13), I determined SHAP values separately for antagonism and synergy predictions in each species. I grouped all antibiotic combinations by drug class (see Supplementary Table S4), to which one or both drugs belong and aggregated SHAP values for every drug class.

3.2.17 Directional modelling of drug interactions in Gram-negative species

To infer directionality of drug interactions in Gram-negative species (*E. coli*, *S. typhimurium* and *P. aeruginosa*), I used a similar approach described in (Fischer et al., 2015). As multivariate phenotypes I used single-drug and combination fitness values at different concentration ratios and in different batches and Gram-negative strains.

First, I log-transformed single-drug and combination effects to generate an additive model with interaction score π_{ij}:

$$\text{Additive model}: \quad \pi_{ij} = \log f_{ij} - \log f_i - \log f_j$$

When $\pi_{ij} = 0$, the additive model follows directly from the log-transformed Bliss independence model :

$$\text{Bliss independence}: \quad f_{ij} = f_i \cdot f_j$$

Under the additive model, combinations in which the interaction vector π_{ij} is parallel or anti-parallel to single-drug phenotype $\log f_i$ implies modulation by the other drug j (Fischer et al., 2015).

To identify drug interactions, in which the interaction vector π_{ij} is either parallel or anti-parallel to the vector of the effect of a single drug, I fitted a linear model $\pi_{ij} \sim \log f_i + \log f_j$ and calculated proportions of variance explained (v_i, v_j) by drugs i and j. Drug interaction was assigned directionality $(i \rightarrow j)$ if

$$v_i < l \text{ and } v_j > u$$

where l is the lower quartile and u is the upper quartile over all proportions of variance explained.

3.2.18 Code availability

I developed an analysis pipeline for the combinatorial screen in Gram-positive species in R. The code for drug interaction modelling (Subsections 3.2.1–3.2.3) is available on EMBL Gitlab: https://git.embl.de/cacace/drug_comb_screen/-/tree/parallel.

I performed data integration (Subsection 3.2.6–3.2.12) of chemogenomic and drug interaction datasets in R. Likewise, conservation analysis in Gram-negative and Gram-positive species as well as directional network analysis (Subsection 3.2.17) in Gram-negatives were conducted in R. The code is available on EMBL Gitlab: https://git.embl.de/vkim/species_conservation.

For drug interaction prediction (Subsection 3.2.13–3.2.16), I used Python to train and cross-validate all machine learning models. The code is avaialble on Github: https://github.com/vladchimescu/chemgen.

3.3 Results

3.3.1 Combinatorial antibiotic screen in Gram-positive organisms

The antibiotic combination screen was conducted in 4 Gram-positive strains: *S. areus* Newman, *S. aureus* DSM 20231, *B. subtilis* 168 and *S. pneumoniae* D39V (Fig. 3.1). About 2000 drug combinations were profiled in each strain. The list of screened compounds is provided in Supplementary Table S3. E. Cacace (2021) produced the experimental data, while I developed the analysis pipeline as described in Methods.

The screen was carried out using 384-well plates and the primary readout was based on the time-resolved optical density (OD) at $\lambda = 595\,\text{nm}$ (Fig. 3.1). As described in (Cacace, 2021), bacterial fitness was quantified as the OD value at the first stationary phase point of the growth curve (point marked by a $*$ in Fig. 3.1). For each drug pair, 3×3 concentration ratios were used to determine Bliss scores (see Methods, Subsection 3.2.1) within a batch. There were at least 4 replicates for each combination (Cacace, 2021).

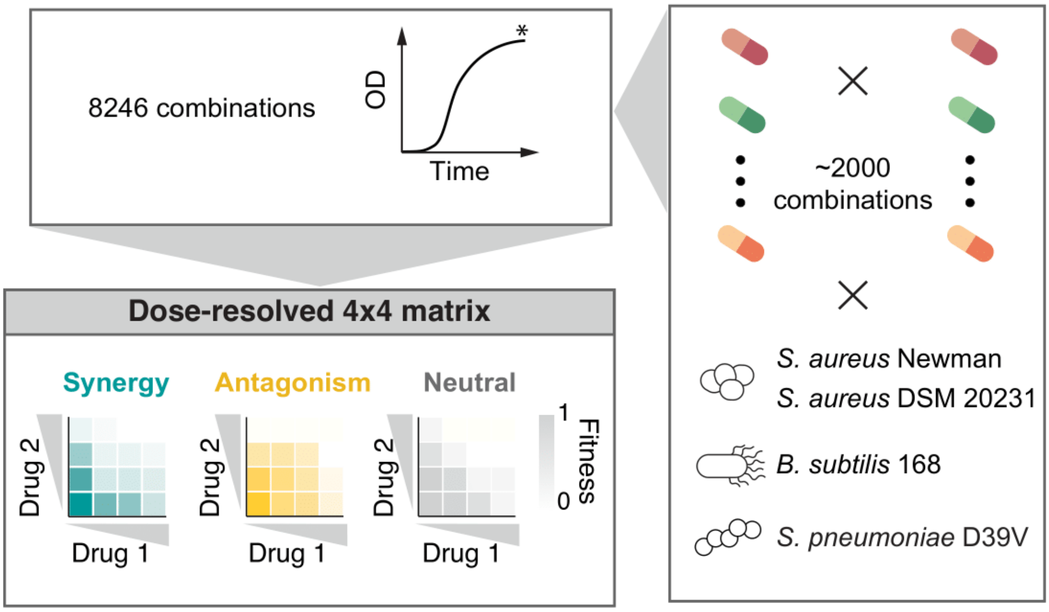


Fig. 3.1 Outline of the combinatorial antibacterial screen in Gram-positive species. Screen schematic created by E. Cacace.

OD values were measured both for single drugs and drug combinations. However, as explained in (Cacace, 2021) the single-drug fitness estimation method that I implemented (see Methods) yielded more reliable estimates than the experimentally measured values.

After I computed within-plate replicate correlations, I excluded all plates with Pearson correlation $r < 0.7$. The median value of OD replicate correlations was 0.92 for combination plates and 0.91 for single-drug plates (Fig. 3.2).

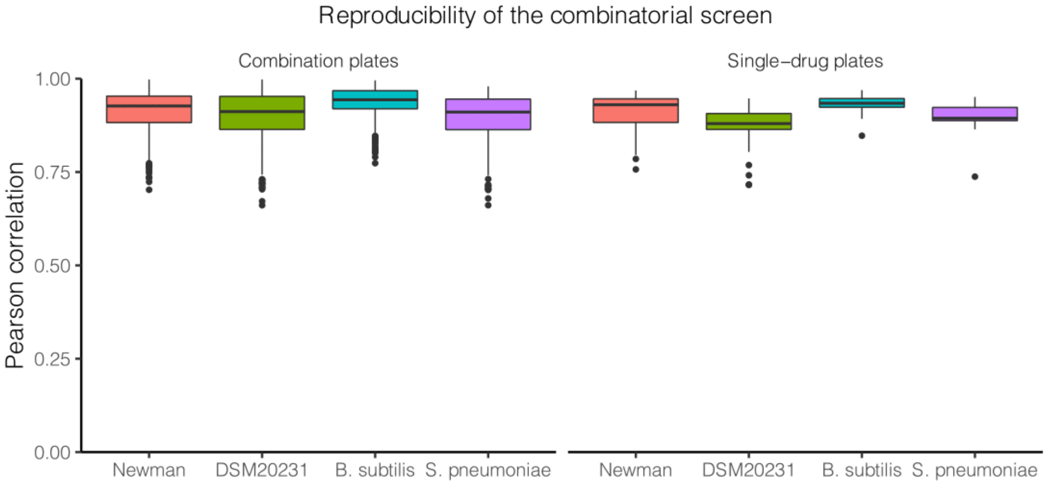


Fig. 3.2 Reproducibility of OD values in the combinatorial screen in 4 Gram-positive strains. Data produced jointly with E. Cacace.

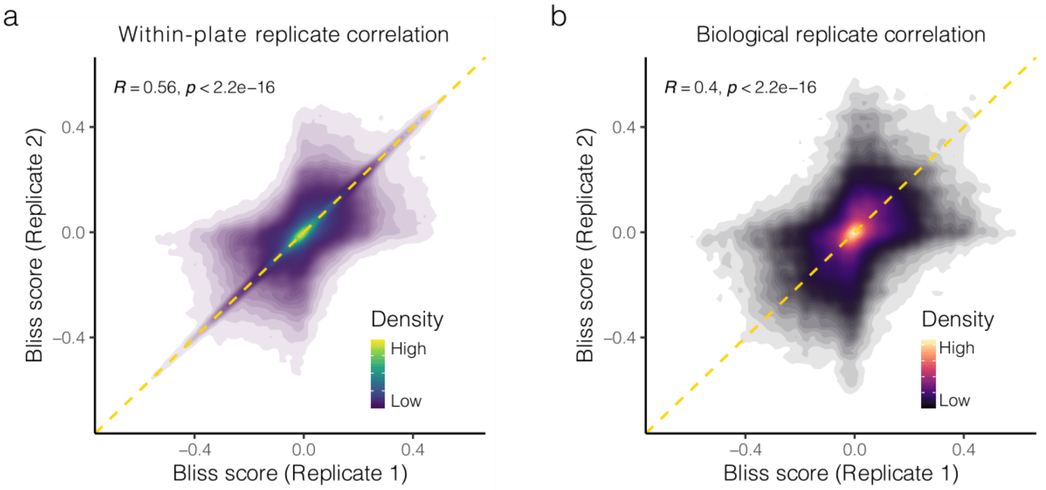


Fig. 3.3 Reproducibility of Bliss scores of drug interactions for a) within-plate replicates and b) biological replicates (well replicates in different batches). Data produced jointly with E. Cacace.

Upon normalization to controls and removal of noisy observations, I computed Bliss scores based on measured combination fitness and estimated single-drug effects (see Methods, Subsection 3.2.2). The Pearson correlation of Bliss scores between replicates within the same plate was $r = 0.56$ (Fig. 3.3a). For biological replicates the Bliss score correlation was much lower, $r = 0.4$ (Fig. 3.3b), due to drug concentration ratios varying across biological replicates. Some of the concentration changes were deliberate, while the remaining variation occurred due to batch effects.

In order to assess the quality of the high-throughput combinatorial screen, 123 drug combinations were validated using an 8×8 checkerboard assay (Fig. 3.4). The benchmarking set included 28 synergies, 26 antagonisms and 69 additive combinations.

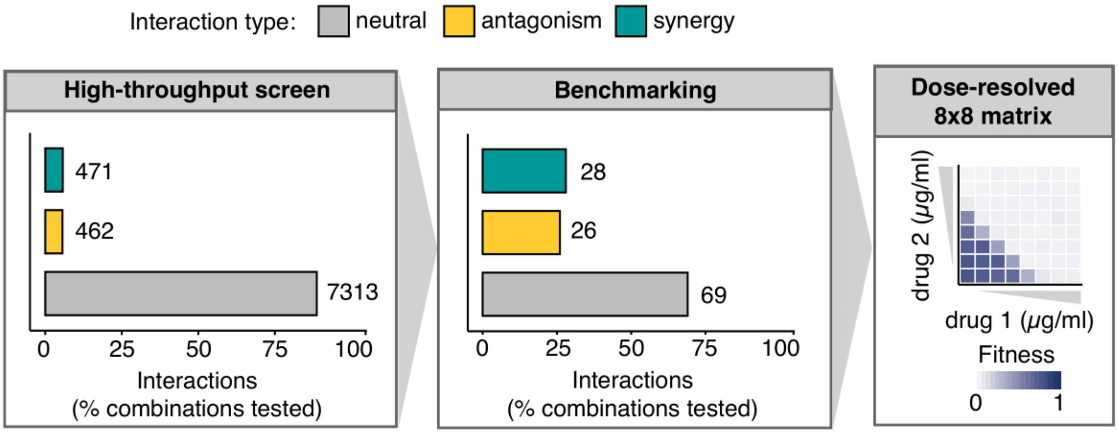


Fig. 3.4 Benchmarking set to assess sensitivity and specificity of the Gram-positive combinatorial screen. Schematic created by E. Cacace.

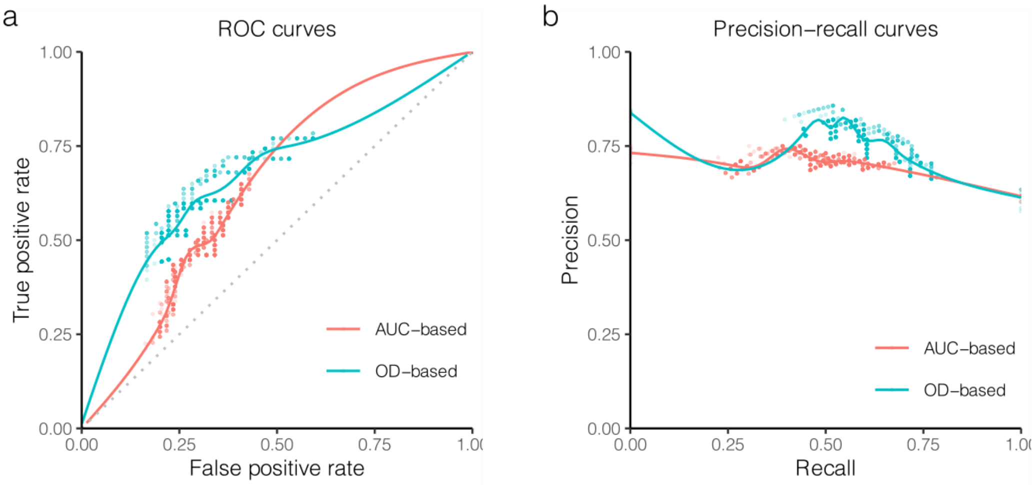


Fig. 3.5 Comparison of the OD- vs AUC-based readout in terms of a) sensitivity and false positive rate b) precision and recall on the benchmarking set. Points along the curves represent different FDR and effect size thresholds. Data produced jointly with E. Cacace.

The validation set was instrumental for making a number of important decisions. In particular, I compared drug interactions quantified based on single time-point OD values with those determined based on area under the growth curve (AUC) (Fig. 3.5). I found that the drug interaction quantification using single time-point OD values was superior in terms of receiver operator characteristic (ROC) and precision-recall curves (Fig. 3.5).

As described in Methods (Subsection 3.2.3), I estimated the effect size of interaction by using the 75th percentile (Q_3) of the Bliss score vector for antagonisms and the 25th percentile (Q_1) for synergies. Using the benchmarking set, I evaluated the influence of varying effect size thresholds on the screen's precision and recall (Fig. 3.6). I chose the effect size cutoff of 0.1 and FDR treshold of 0.05, which is indicated in Fig. 3.6.

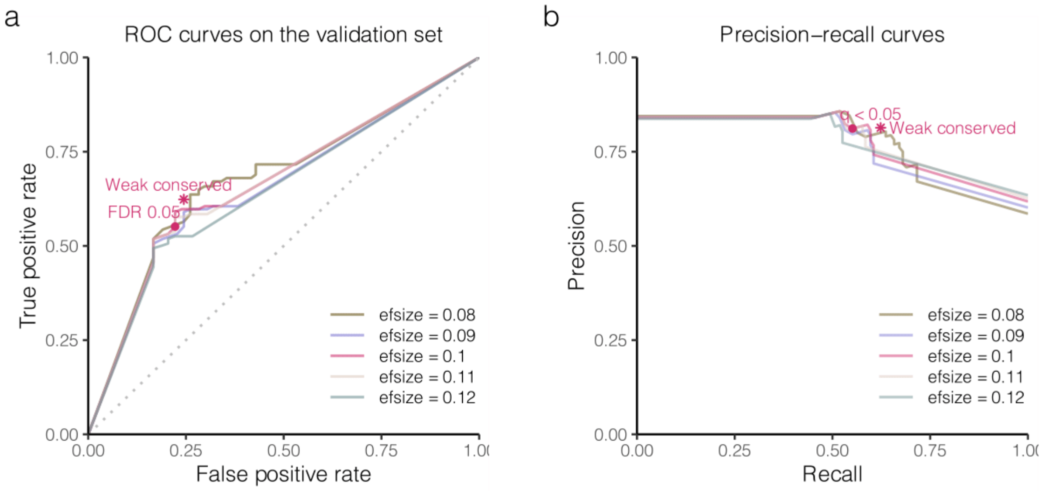


Fig. 3.6 Comparison of effect size thresholds based on a) ROC and b) precision-recall curves on the benchmarking set. The • point indicates the FDR cutoff of 0.05 for the curves with the effect size threshold of 0.1. Leveraging weak conservation (∗ point) boosts precision and recall as described in (Cacace, 2021). Data produced jointly with E. Cacace.

The true positive rate (TPR/recall), false positive rate (FPR) and precision calculations based on the benchmarking set are provided below:

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} = \frac{44}{44+34} = 0.56$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}} = \frac{10}{10+35} = 0.22$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FP}} = \frac{44}{44+10} = 0.81$$

3.3.2 Compendium of antibiotic interactions in Gram-positive species

After identifying drug interactions in the screen (see Methods), I observed comparable numbers of interactions detected across the species (Fig. 3.7). In all strains except *B. subtilis* the number of antagonisms was larger than that of synergies (Fig. 3.7).

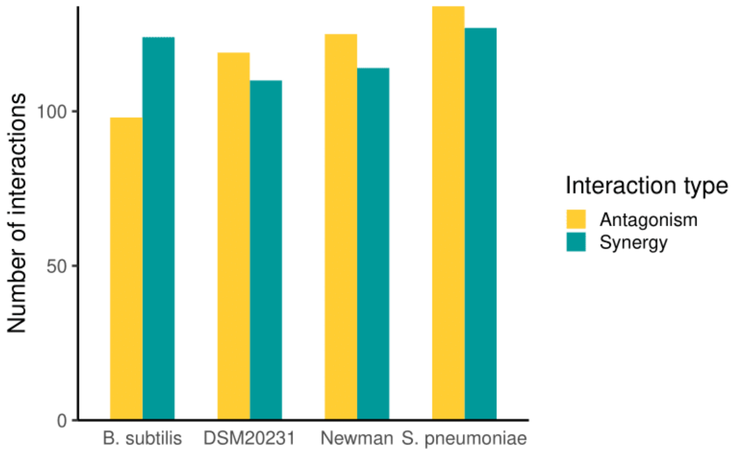


Fig. 3.7 Number of synergies and antagonisms observed in the combinatorial screen. Data produced jointly with E. Cacace.

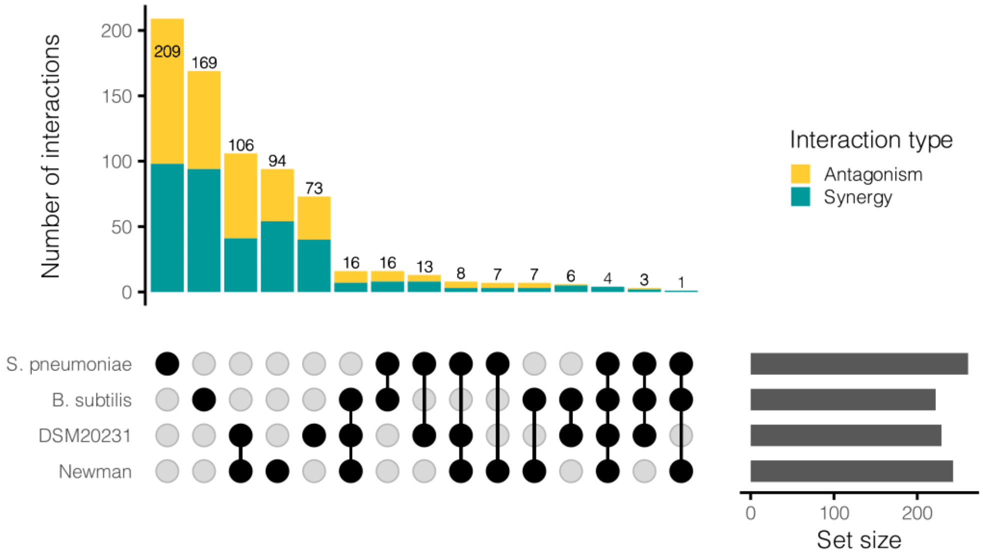


Fig. 3.8 Conservation of drug interactions among the Gram-positive strains. The upset plot shows the number of drug interactions (bar plot) for different strain intersections (dot plot). Data produced jointly with E. Cacace.

Conservation rates varied across the Gram-positive organisms (Fig. 3.8). As expected, *S. aureus* strains Newman and DSM20231 shared the highest number of drug interactions (Fig. 3.8), with slightly more conserved antagonisms. *B. subtilis* and the two *S. aureus* strains shared 37 interactions. 24 interactions were conserved between *B. subtilis* and *S. pneumoniae*. I found that only 4 drug interactions were conserved in all Gram-positive strains (Fig. 3.8), all of which were synergies. Overall, I observed low cross-species conservation rates among the Gram-positive organisms, ranging from 9% to 16%.

Comparing the conservation of drug interactions in both Gram-positive and Gram-negative organisms (Fig. 3.9), I found that Gram-negative species *E. coli* and *S. typhimurium* have considerably higher within- and cross-speices conservation rates than the Gram-positive organisms. However, within the Gram-negatives, conservation rates of the *Pseudomonas* strains were comparable with those observed for the Gram-positive organisms (Fig. 3.9). Across the Gram dichotomy, the conservation network was very sparse.

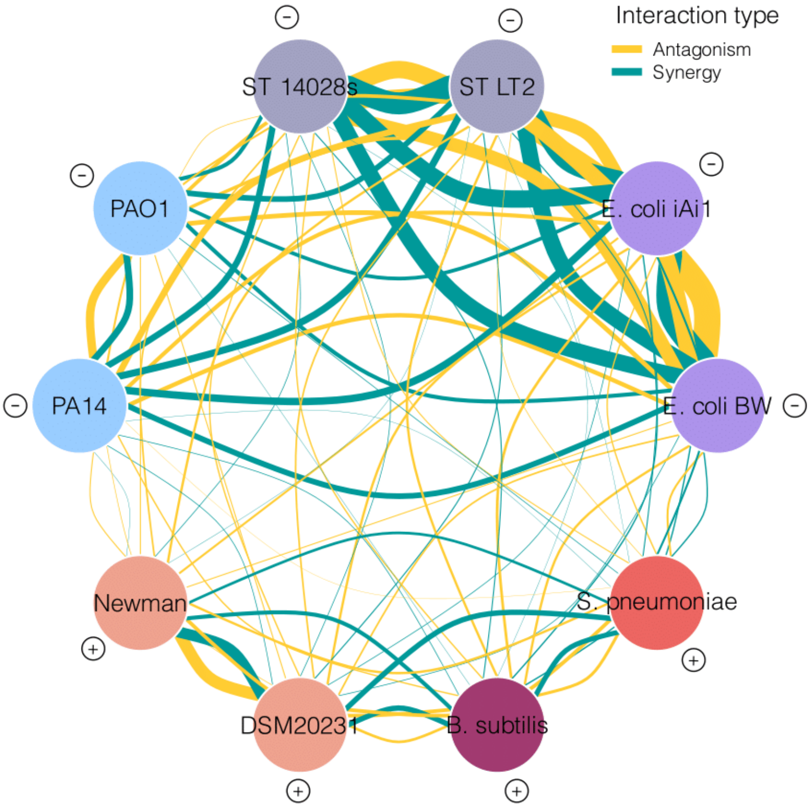


Fig. 3.9 Conservation of drug interactions among Gram-positive and Gram-negative species. Nodes represent bacterial strains, with ⊕ marking Gram-positive species. Edge width indicates the number of drug interactions conserved between the connected strains. Gram-positive data produced jointly with E. Cacace. Gram-negative data source: Brochado et al. (2018).

By integrating drug class information, I identified the key patterns of drug interactions in Gram-positive bacteria (Fig. 3.10). Drug classes with the most drug interactions in the screen were cell wall-targeting agents, protein and DNA synthesis inhibitors (Fig. 3.10). Within-class drug interactions occurred primarily within the top interacting classes and the majority of these interactions were synergistic (Fig. 3.10). Furthermore, I found a number of antagonizing drug class pairs that included {protein synthesis, cell wall}, {protein synthesis, human} and {protein synthesis, DNA} (Fig. 3.10). Drug classes that were likely to synergize included {cell wall, DNA}, {cell wall, membrane} and {cell wall, human} (Fig. 3.10).

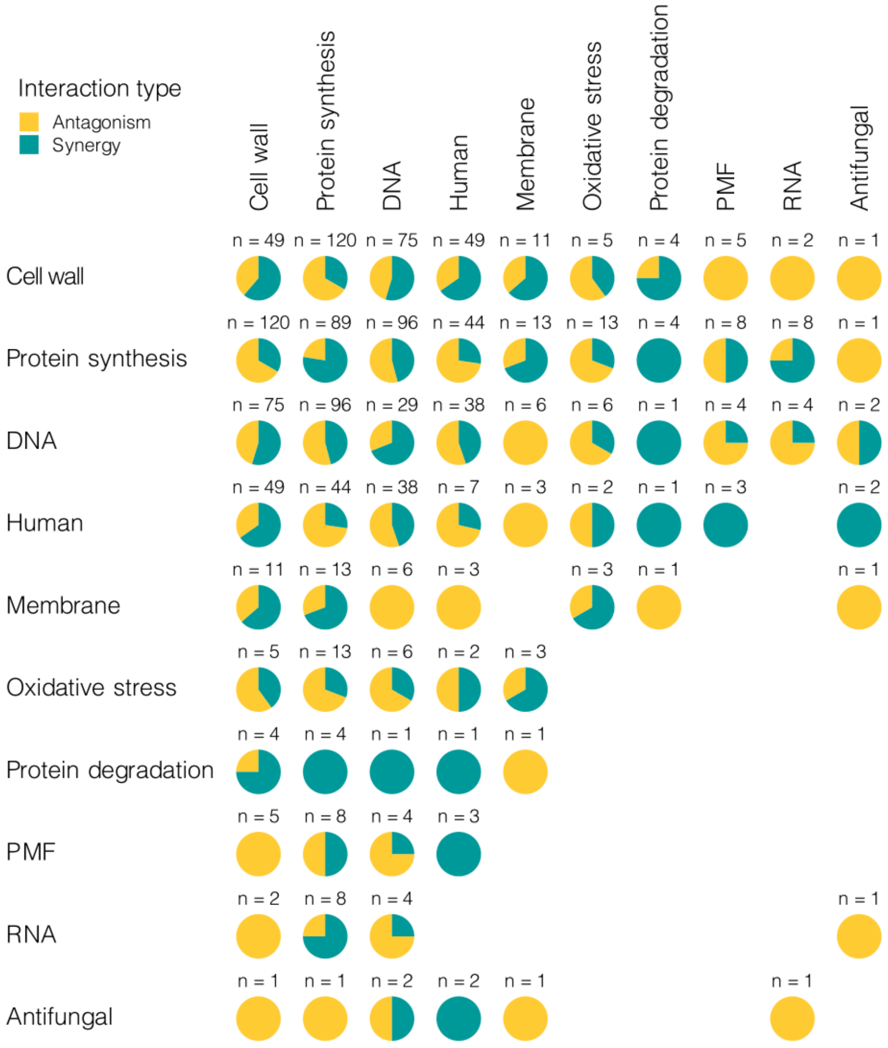
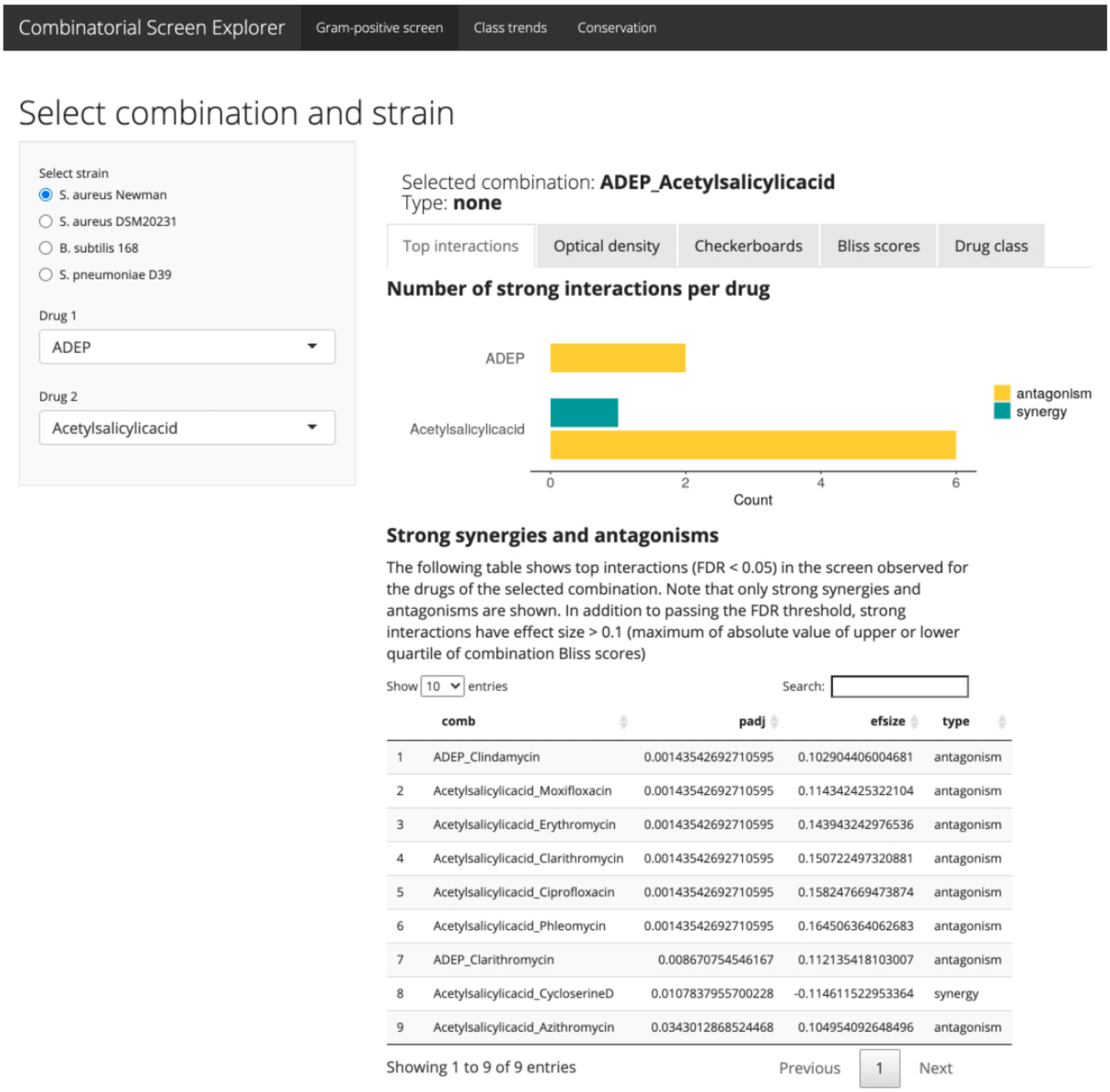


Fig. 3.10 Drug interactions ordered by compound class (class-class interaction matrix). Pie charts show relative abundances of antagonisms and synergies for a given pair of drug classes. Text labels ($n = \dots$) indicate numbers of drug interactions observed. Data produced jointly with E. Cacace.

3.3.3 Combinatorial screen explorer

To facilitate exploratory analysis, I developed an interactive web application that enables a user to query the combinatorial screen data (Fig. 3.11). The Shiny app is available online: https://combscreen.shiny.embl.de and can be used for inspection of the drug combination data in Gram-positive species.

The first dashboard (Fig. 3.11) shows combination type and number of drug interactions for a selected drug pair and bacterial strain as well as the table of all drug interactions involving the selected drugs.



Select combination and strain

Select strain
● S. aureus Newman
○ S. aureus DSM20231
○ B. subtilis 168
○ S. pneumoniae D39

Drug 1: ADEP

Drug 2: Acetylsalicylicacid

Selected combination: **ADEP_Acetylsalicylicacid**
Type: **none**

Number of strong interactions per drug

Strong synergies and antagonisms

The following table shows top interactions (FDR < 0.05) in the screen observed for the drugs of the selected combination. Note that only strong synergies and antagonisms are shown. In addition to passing the FDR threshold, strong interactions have effect size > 0.1 (maximum of absolute value of upper or lower quartile of combination Bliss scores)

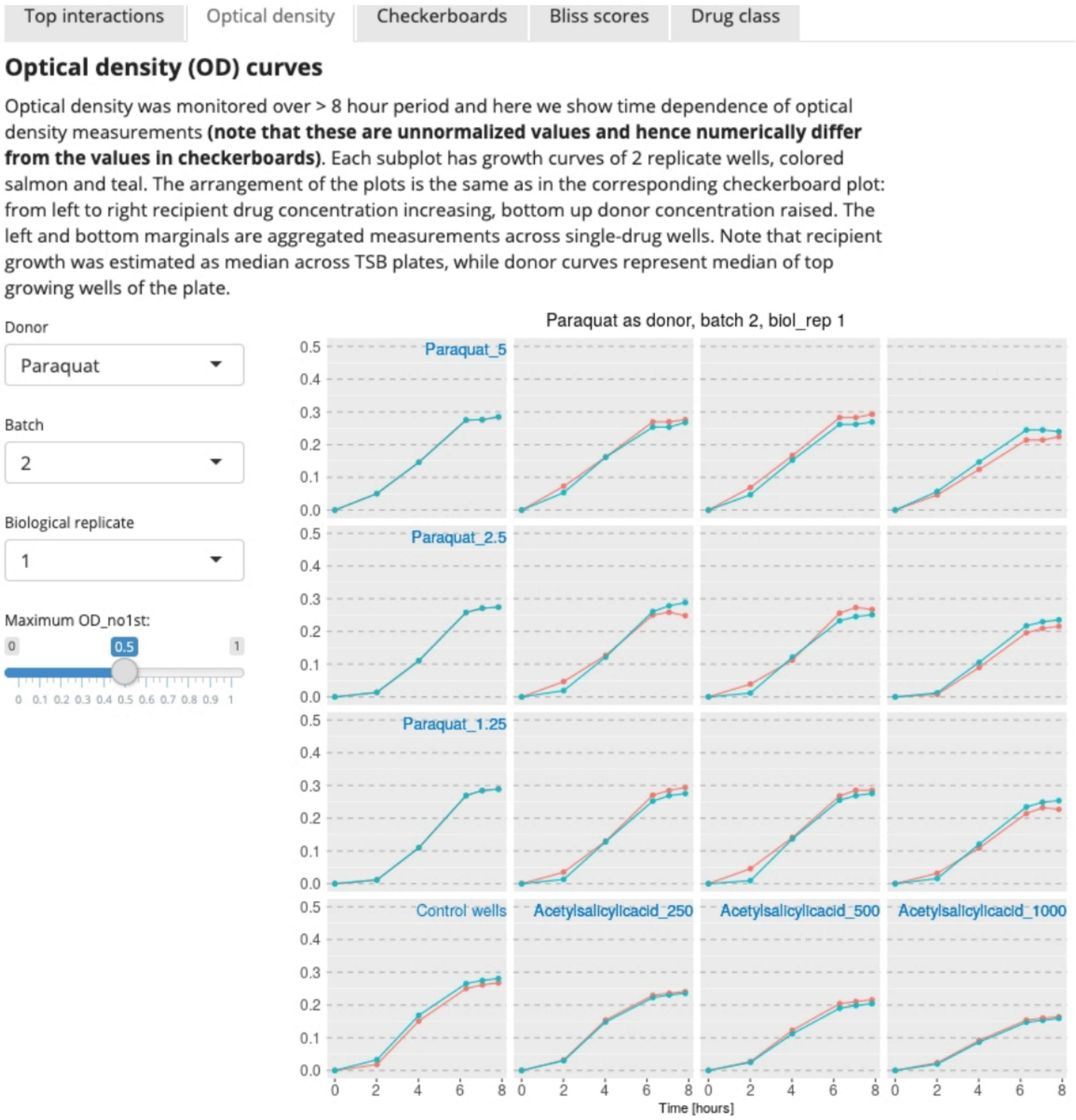
Show 10 entries Search:

	comb	padj	efsize	type
1	ADEP_Clindamycin	0.00143542692710595	0.102904406004681	antagonism
2	Acetylsalicylicacid_Moxifloxacin	0.00143542692710595	0.114342425322104	antagonism
3	Acetylsalicylicacid_Erythromycin	0.00143542692710595	0.143943242976536	antagonism
4	Acetylsalicylicacid_Clarithromycin	0.00143542692710595	0.150722497320881	antagonism
5	Acetylsalicylicacid_Ciprofloxacin	0.00143542692710595	0.158247669473874	antagonism
6	Acetylsalicylicacid_Phleomycin	0.00143542692710595	0.164506364062683	antagonism
7	ADEP_Clarithromycin	0.008670754546167	0.112135418103007	antagonism
8	Acetylsalicylicacid_CycloserineD	0.0107837955700228	-0.114611522953364	synergy
9	Acetylsalicylicacid_Azithromycin	0.0343012868524468	0.104954092648496	antagonism

Showing 1 to 9 of 9 entries

Fig. 3.11 The first panel of the interactive web application developed for exploration of the antibiotic combination data in Gram-positive species. Data produced jointly with E. Cacace.

A user can view bacterial growth curves (Fig. 3.12) in wells treated with a selected drug combination and compare these with the growth under single-drug treatment. I implemented this feature mainly for quality control as one can check for measurement artifacts in individual batches and replicates.



Optical density (OD) curves

Optical density was monitored over > 8 hour period and here we show time dependence of optical density measurements **(note that these are unnormalized values and hence numerically differ from the values in checkerboards)**. Each subplot has growth curves of 2 replicate wells, colored salmon and teal. The arrangement of the plots is the same as in the corresponding checkerboard plot: from left to right recipient drug concentration increasing, bottom up donor concentration raised. The left and bottom marginals are aggregated measurements across single-drug wells. Note that recipient growth was estimated as median across TSB plates, while donor curves represent median of top growing wells of the plate.

Donor: Paraquat

Batch: 2

Biological replicate: 1

Maximum OD_no1st: 0.5

Fig. 3.12 Interactive visualization of optical density (OD) curves in the combinatorial screen explorer. The marginal plots (bottom and leftmost rows) show bacterial growth under single-drug treatment. The inner plots of the matrix show OD curves in wells treated with drug combinations. Data produced jointly with E. Cacace.

For each combination 3 × 3 concentration ratios were probed in each batch. To compare biological replicates, I visualized combination data as heatmaps (Fig. 3.13) that show normalized optical density (OD) values for all concentration ratios, technical and biological replicates and batches.

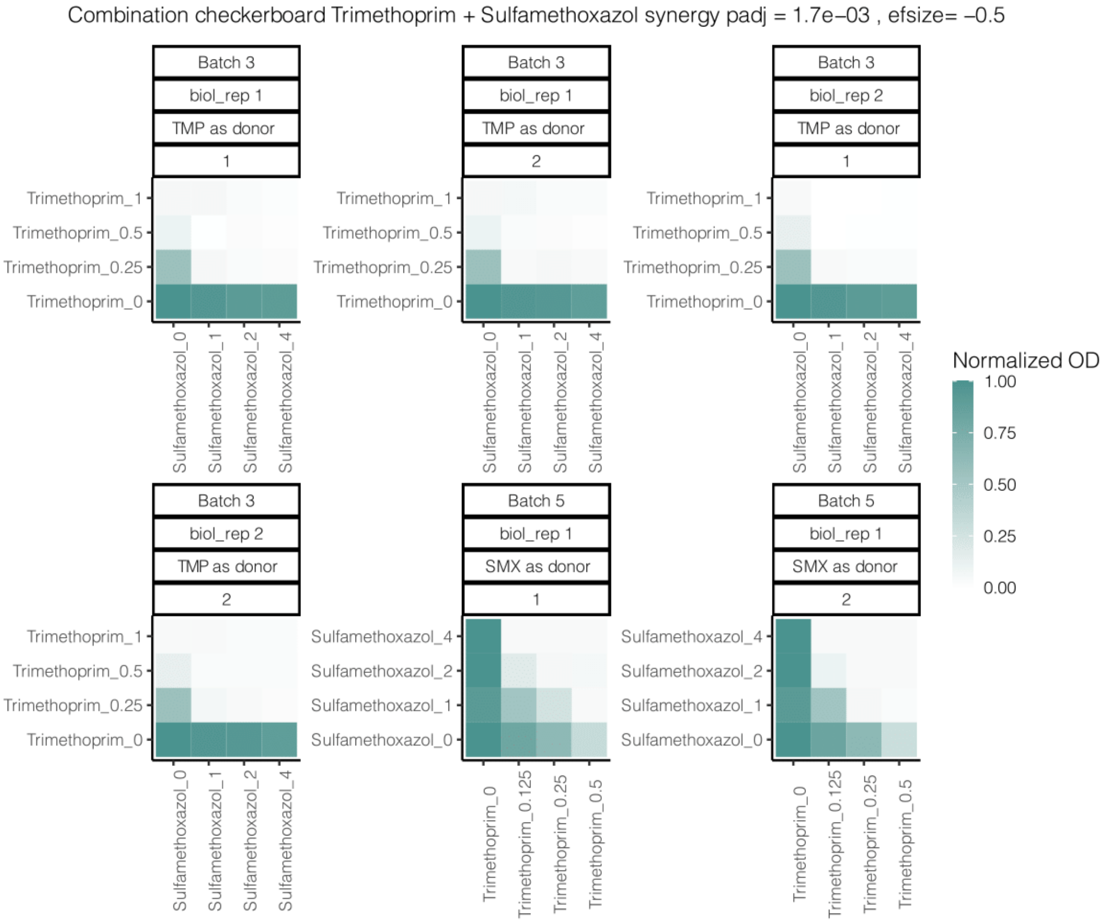


Fig. 3.13 Checkerboard plot of trimethoprim + sulfamethoxazol synergy in the combinatorial screen explorer. Heatmaps show normalized optical density (OD) values for all tested concentration ratios, biological replicates and batches. A "donor" drug is pipetted onto a drug dissolved in the medium to generate a drug combination. Data produced jointly with E. Cacace.

Additionally, one can use the web application to check class-class interactions in individual strains and examine synergies and antagonisms for a specific class pair. Furthermore, conservation in selected strain pairs and across all probed organisms can be explored interactively.

3.3.4 Hierarchical clustering of compounds

As described in (Yeh et al., 2006), drug interaction networks can be used for functional classification of compounds. Assuming that similar drugs have similar interactions, I clustered the probed compounds by their observed Bliss scores in all strains and combinations (Fig. 3.14).

The clustering recapitulated virtually all cell wall targeting compounds (Cluster 2 in Fig. 3.14), with the exception of nisin, teicoplanin and streptozotocin that did not cluster with the remaining cell wall drugs.

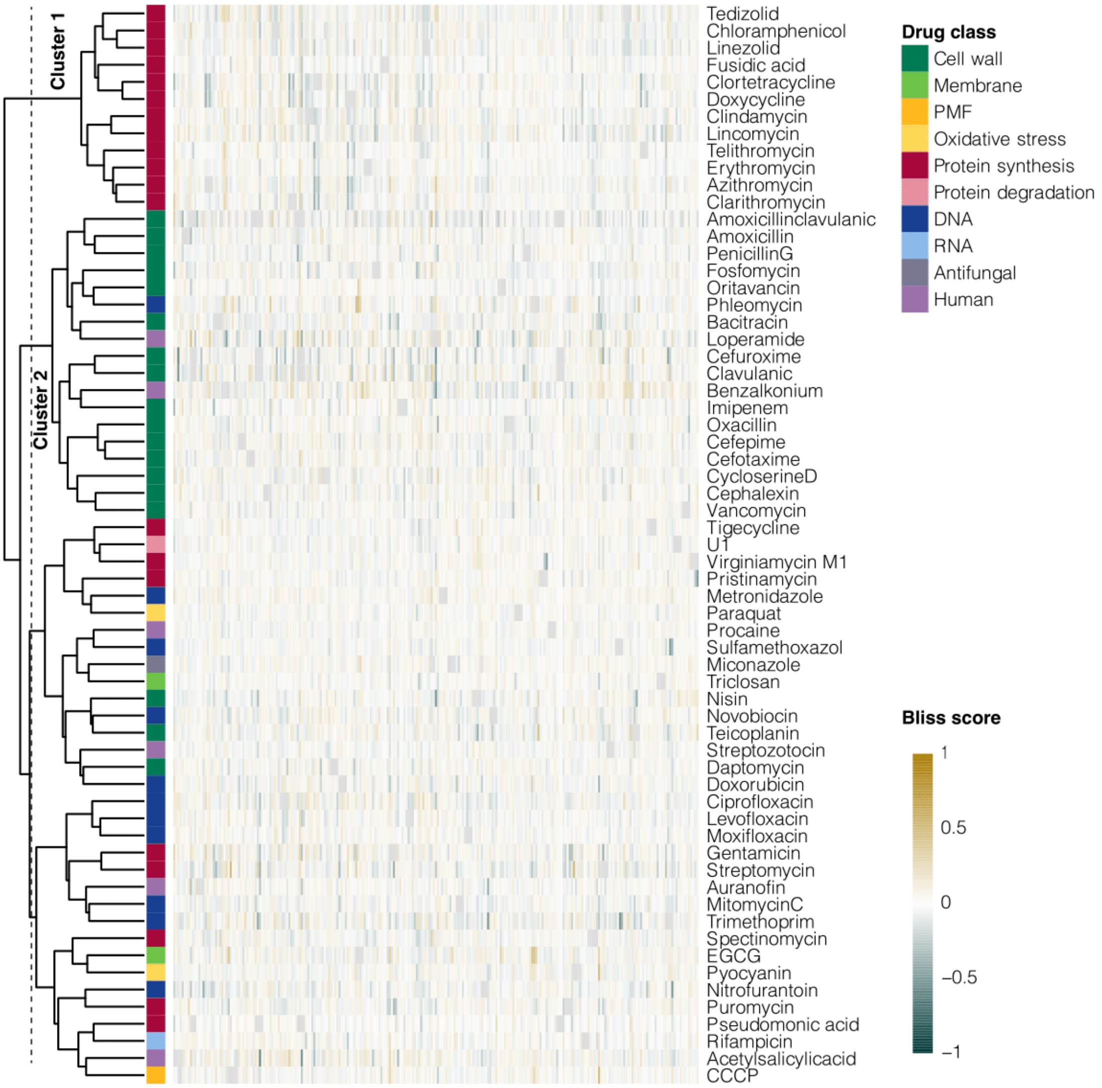


Fig. 3.14 Hierarchical clustering of compounds based on observed interaction scores in the Gram-positive combinatorial screen. Missing values are colored grey. Data produced jointly with E. Cacace.

Many protein synthesis inhibitors were grouped together (Cluster 1 in Fig. 3.14) reflecting that most compounds within this class had similar drug interaction profiles. Streptogramin antibiotics, virginamycin and pristinamycin, clustered together (Fig. 3.14) but not with other protein synthesis inhibitors. A similar pattern was observed for the aminoglycosides gentamicin and streptomycin (Fig. 3.14).

While several DNA synthesis inhibitors (doxorubicin, ciprofloxacin, levofloxacin and moxifloxacin) had similar interaction profiles (Fig. 3.14), they did not form a separate cluster. The class of human-targeted drugs ('Human') is heterogeneous and unites many different functional classes, which is also reflected by their dispersion in the heatmap (Fig. 3.14).

3.3.5 Data integration in Gram-negative species

First, I obtained the published chemogenomic data in *E. coli* (Nichols et al., 2011) and *S. typhimurium* (Pfalz, 2017). Drug combination data was available for the same species from the publication by Brochado et al. (2018). I integrated the chemogenomic and drug combination datasets (Fig. 3.15) to predict novel drug interactions and identify drug interaction-associated genes.

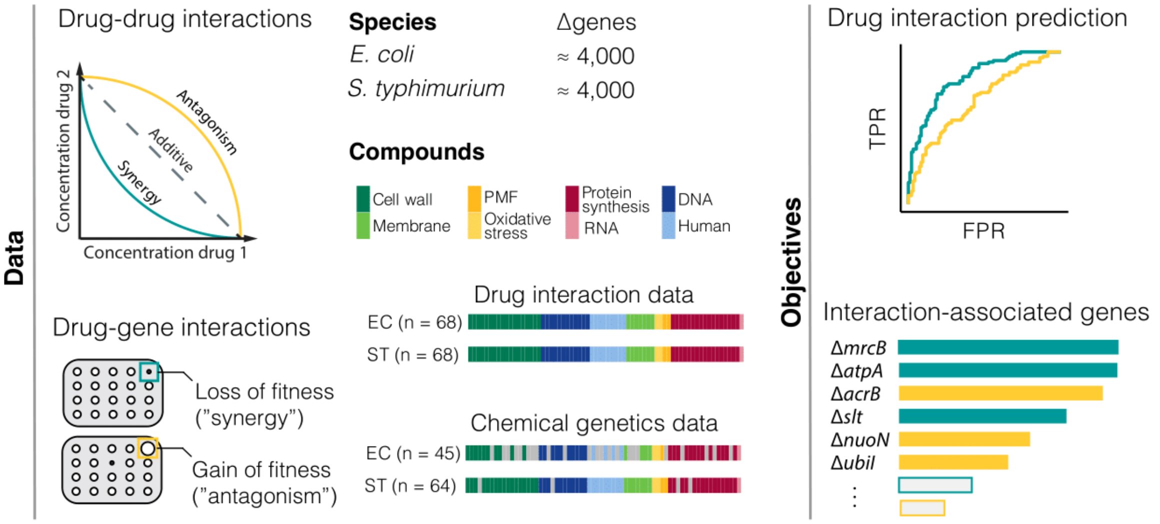


Fig. 3.15 Integration of chemogenomic and drug interaction data in Gram-negative species *E. coli* and *S. typhimurium* (EC = *E. coli*, ST = *S. typhimurium*).

Both *E. coli* and *S. typhimurium* chemogenomic datasets were generated by testing drugs in ≈ 4000 non-essential gene deletion strains (Fig. 3.15). The major compound classes represented in all datasets were cell wall targeting agents, DNA and protein synthesis inhibitors, human-targeted drugs and compounds acting on the membrane (Fig. 3.15).

All pairwise combinations of 68 drugs were screened in Gram-negative species by Brochado et al. (2018). When intersecting these compounds with those probed in gene deletions, I found (Fig. 3.15) that there were 45 compounds in *E. coli* and 64 compounds in *S. typhimurium* that were overlapping between chemogenomic and drug cobmination datasets. In the downstream analysis, I used only the common subset of drugs (*E. coli*: 45 and *S. typimurium*: 64).

As shown in the schematic (Fig. 3.15) loss of fitness in a gene deletion can be interpreted as a synergistic interaction between a drug and that gene. Conversely, gain of fitness in a deletion strain can be construed as an antagonistic drug-gene interaction. Given this correspondence, I checked the relative prevalance of synergistic/antagonistc interactions in chemogenomic and drug combinatorial domains. The number of synegistic drug-gene interactions was much higher than that of gain-of-fitness phenotypes observed in gene deletion strains (Fig. 3.16a), suggesting that there is potentially more room for synergies when non-essential genes and major compound classes are considered.

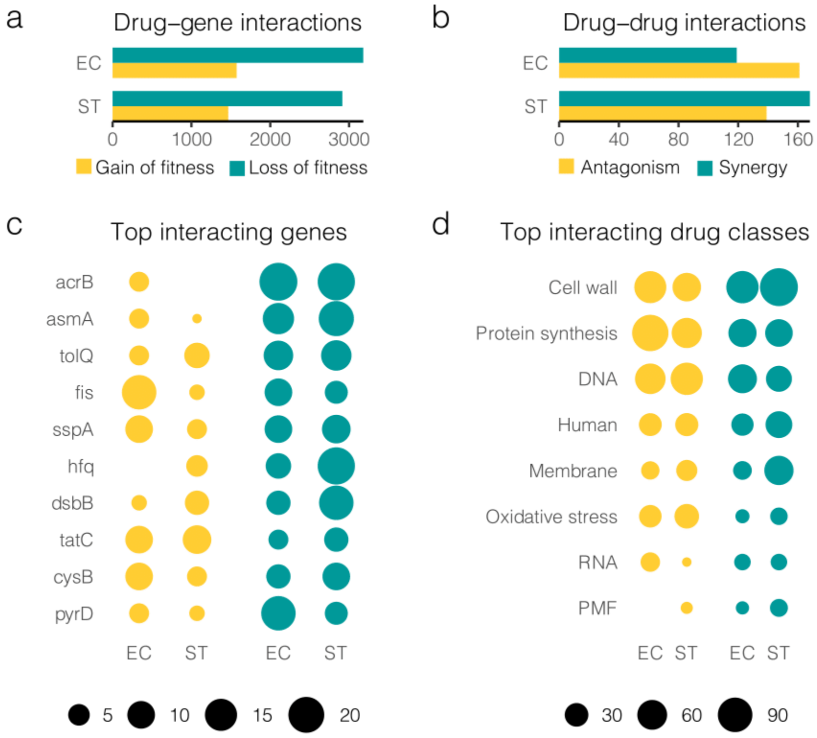


Fig. 3.16 Overview of chemogenomic and drug interaction data (EC = *E. coli*, ST = *S. typhimurium*). a) Drug-gene and b) drug-drug interactions in *E. coli* and *S. typhimurium*. c) Top interacting genes in the chemogenomic data. d) Top interacting drug classes in the drug combination data. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

Surveying drug combination data (Fig. 3.16b), I found that the numbers of synergies and antagonisms were comparable, with *E. coli* having slightly more antagonistic drug interactions. The top interacting drug classes (Fig. 3.16d) in the drug combination dataset were cell wall targeting compounds, protein and DNA synthesis inhibitors.

To investigate the synergy gap between drug-gene and drug-drug interactions, I examined the genes with the highest number of interactions (Fig. 3.16c). Among the synergy-generating genes were membrane proteins *acrB*, *tolQ*, *dsbB*, and *pyrD* – none of which is currently targetable by any of the approved drugs according to DrugBank. Similarly, all other top interacting genes are not targeted by any of the approved compounds, which partially explains the discrepancy in synergy rates between chemical genetics and drug combinations.

3.3.6 Multiomic factor analysis on chemogenomic and interaction data

As described in Methods (Subsection 3.2.7), I used multiomic factor analysis (MOFA) to analyze chemogenomic and drug interaction data jointly and perform compound profiling based on the observed drug-gene and drug-drug interactions. The omic layers (Fig. 3.17) describing each drug were drug interaction and chemogenomic profiles in *E. coli* and *S. typhimurium*.

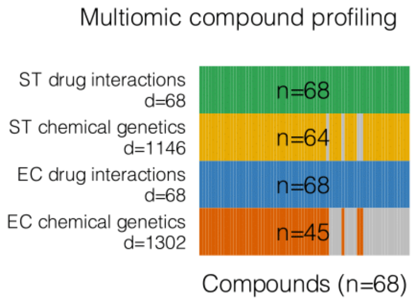


Fig. 3.17 Multiomic views used for unsupervised analysis of compounds (EC = *E. coli*, ST = *S. typhimurium*). Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

The unsupervised MOFA model produced 10 latent factors (Fig. 3.18) describing the variation of compounds across different omic layers. The first latent factor (LF1) had shared loadings on the drug interaction views in both *E. coli* and *Salmonella* as well as on the chemical genetics view in *E. coli* (Fig. 3.18). There were 4 latent factors (LF1, LF3, LF4 and LF5) that were shared between the EC and ST drug interaction views (Fig. 3.18). Additionally, 4 latent factors (LF3, LF4, LF7, LF8) were shared between the EC and ST chemical genetics views (Fig. 3.18).

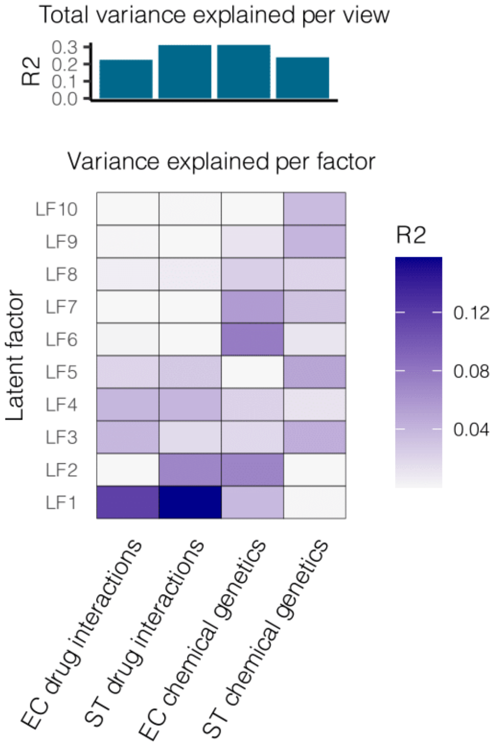


Fig. 3.18 Latent factors inferred by multiomic factor analysis (MOFA) and variance explained per factor (EC = *E. coli*, ST = *S. typhimurium*). Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

Next, I performed a hierarchical clustering of compounds in the inferred latent space (Fig. 3.19). The analysis revealed clusters corresponding to mainly cell wall targeting drugs, protein synthesis inhibitors and compounds acting on the membrane (Fig. 3.19). It is remarkable that one can classify compounds by using solely 10 latent factors. The advantage of this interpretable clustering is the ability to quickly identify which latent factors drive the differences and similarities of drugs.

For instance, membrane-targeting drugs (colistin, CHIR-90, chlorhexidine) clustered together, which was largely driven by LF1 (Fig. 3.19). By checking the loadings on the chemical genetics data in *E. coli* (Fig. 3.20a), I found that an outer membrane protein, *ompA*, was the feature with the highest absolute loading in EC chemical genetics.

In drugs with higher LF1 values (Fig. 3.20b) I observed primarily a gain of fitness in *ompA* gene deletions. Even though the membrane-targeting drugs chlorhexidine, CHIR-90 and colistin were not probed in *E. coli*, MOFA predicts that in these drugs one would observe a gain of fitness in an *ompA* deletion mutant. In *Salmonella*, CHIR-90 and colistin indeed gained fitness in the *ompA* gene deletion strain (Fig. 3.20c). In addition to their inferred antagonistic interaction with *ompA*, membrane drugs had similar drug interactions that were conserved between *E. coli* and *S. typhimurium* (Supplementary Fig. S8c,d).

Based on the latent space clustering (Fig. 3.19), I identified a number of other drug class specific commonalities. Cell wall targeting compounds were similar in terms of LF2 (Fig. 3.19), which had high absolute loadings on purine metabolism genes in *E. coli* (Supplementary fig. S9a). Additionally, the clustering of cell wall drugs was driven by synergies with other drugs within the cell wall class (Supplementary fig. S9c,d).

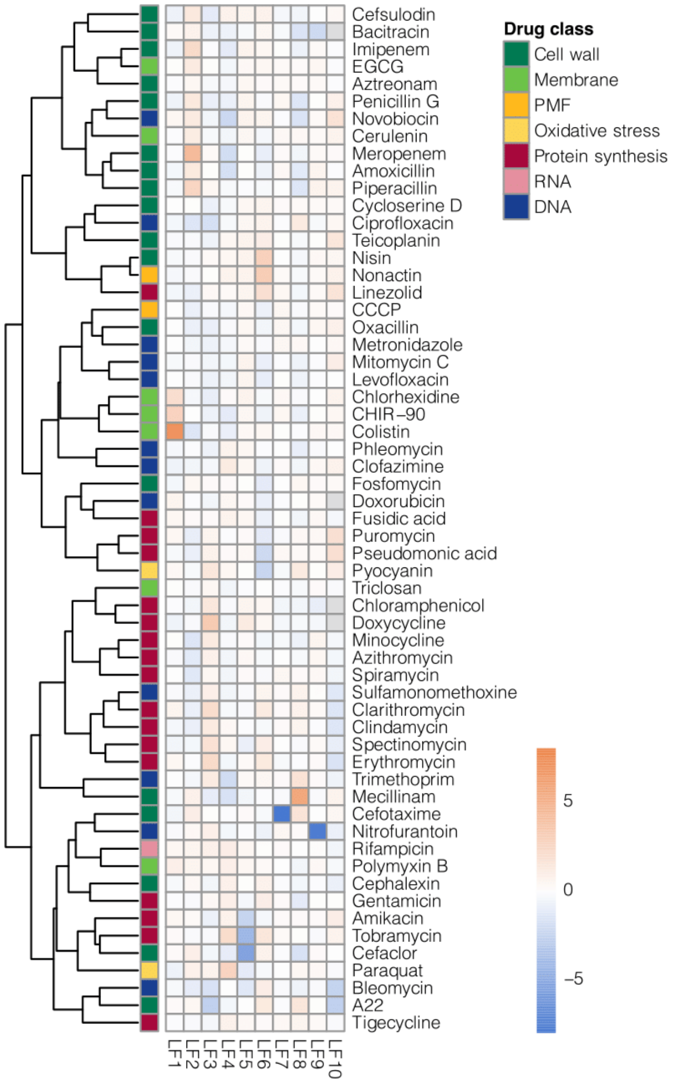


Fig. 3.19 Clustering of compounds in the inferred latent factor space. Row annotation provides drug class labels. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

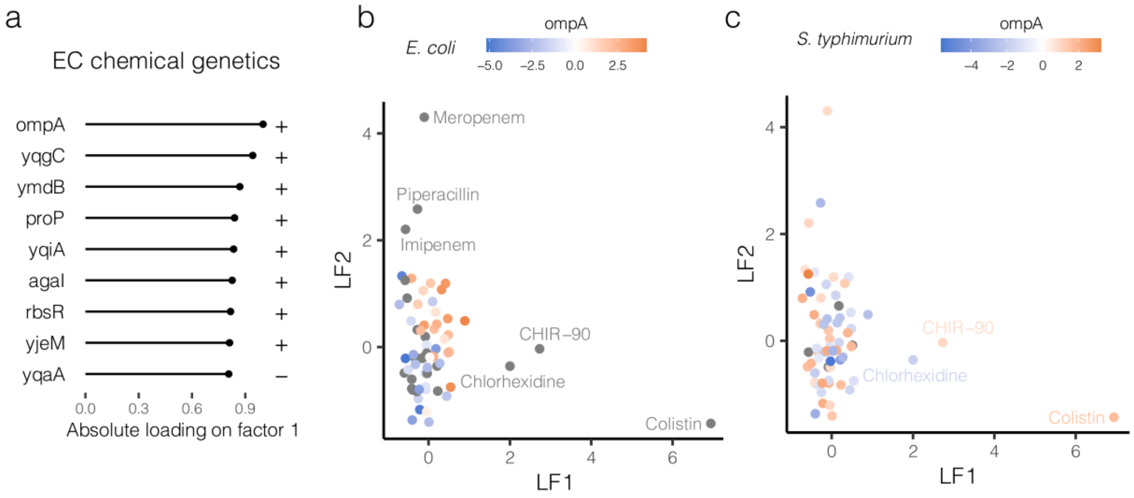


Fig. 3.20 Exploring latent factor 1 (EC = *E. coli*, ST = *S. typhimurium*). a) LF1 absolute loadings in EC chemical genetics. b) Drug embedding in the LF1-LF2 space. Points are colored by drug-gene interaction score in EC *ompA* deletion. c) Drug embedding in the LF1-LF2 space. Points are colored by drug-gene interaction score in ST *ompA* deletion. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

3.3.7 Compound similarity and drug interactions

The similarity hypothesis is a popular assumption, according to which similar compounds are more likely to interact synergistically (Bansal et al., 2014). To verify this hypothesis, I computed chemogenomic similarity for every drug pair and investigated the relationship between similarity and drug interaction score (Fig. 3.21). In general, there was no tendency for chemogenomically similar drugs to interact predominantly synergistically or antagonistically (Fig. 3.21). However, antagonisms had somewhat higher chemogenomic similarity compared with additive combinations (Fig. 3.21, Wilcoxon $p = 0.009$). Synergistic drug pairs also exhibited higher chemogenomic similarity than neutral combinations, although the shift was not as significant (Fig. 3.21, Wilcoxon $p = 0.02$).

As described in Subsection 3.3.4, drug interaction profiles can be used for compound classification. I computed drug similarity based on interaction profiles (see Methods, Subsection 3.2.8) and compared antagonisms and synergies with neutral combinations (Fig. 3.22). Synergistic drug pairs had significantly higher interaction profile similarity than additive combinations (Fig. 3.22, Wilcoxon $p = 4.6 \cdot 10^{-15}$), while the opposite was observed for antagonistic pairs (Fig. 3.22, Wilcoxon $p = 4.2 \cdot 10^{-4}$), which had lower interaction profile similarity than neutral combinations.

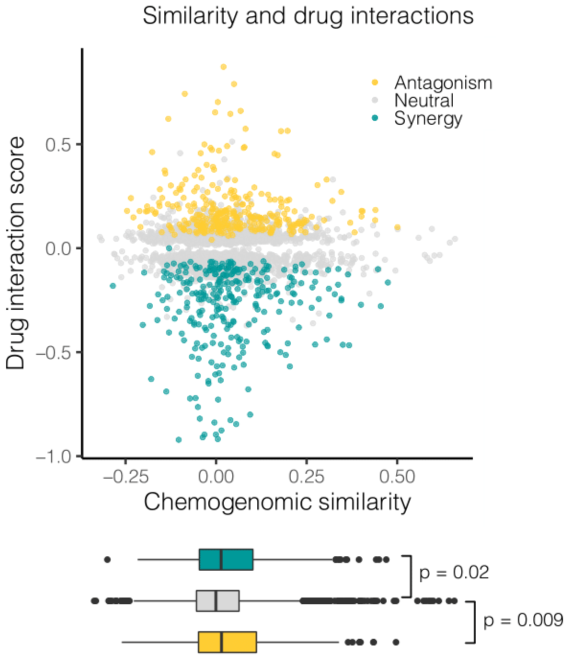


Fig. 3.21 Chemogenomic similarity (*x*-axis) and drug interaction scores (*y*-axis) in *E. coli* and *S. typhimurium*. Drug combinations are colored by interaction type. The boxplots compare chemogenomic similarity of antagonisms, synergies and neutral combinations. Wilcoxon test *p*-values are shown. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

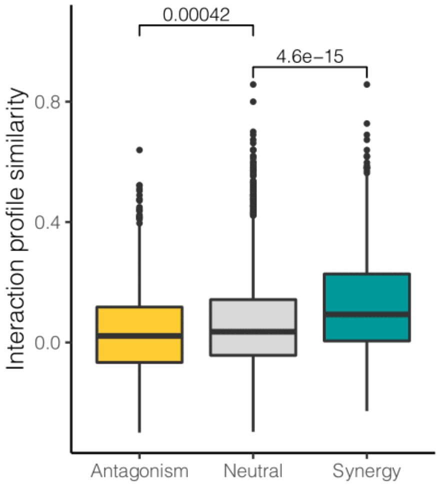


Fig. 3.22 Interaction profile similarity by drug combination type in Gram-negative species (*E. coli*, *S. typhimurium* and *P. aeruginosa*). Wilcoxon test *p*-values are shown. Drug interaction data source: Brochado et al. (2018).

For antagonisms and synergies I compared chemogenomic and interaction profile similarities (Fig. 3.23) and found a low correlation between the two similarity measures ($r = 0.2$ for synergies and $r = 0.3$ for antagonisms). This suggests that drug similarity estimated based on chemogenomic profiles differs from that determined using drug interaction profiles. Furthermore, the similarity hypothesis holds if interaction profile similarity is used (Fig. 3.22), but not when drugs are compared in chemogenomic space (Fig. 3.21).

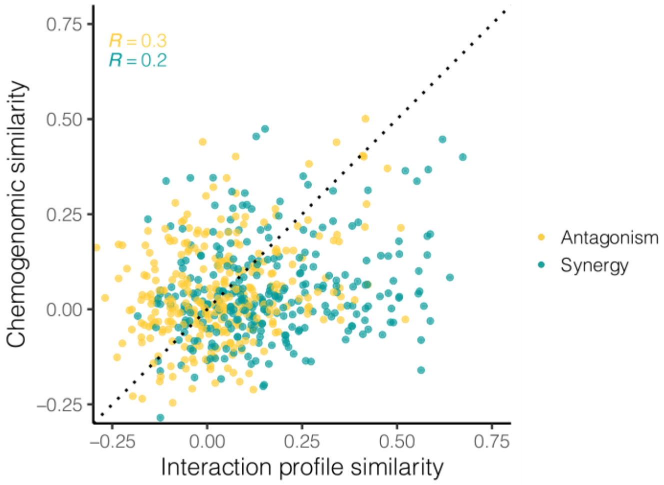


Fig. 3.23 Chemogenomic and interaction profile similarity of antagonisms and synergies in *E. coli* and *S. typhimurium*. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

3.3.8 Genes associated with drug-drug interactions

To identify genes associated with antagonisms and synergies, I applied a χ^2-test, with the null hypothesis being that combination type ($y_c \in \{\text{antagonism}, \text{synergy}, \text{neutral}\}$) is independent of gene-specific combination signature ($X_c \in \{-2, -1, 0, 1, 2, \pm\}$) (see Methods, Subsection 3.2.10). In the subsequent analysis, I used only those genes that were significantly associated with antagonisms or synergies.

After multiple testing correction (FDR < 0.1), I found that there were 438 genes associated with antagonisms and 224 associated with synergies in *E. coli*. In *Salmonella*, there were 321 antagonism-associated and 245 synergy-associated genes. I performed gene enrichment analysis in *E. coli* (Fig. 3.24) and *S. typhimurium* (Fig. 3.25) and found a number of biologial processes enriched among the interaction-associated genes.

In *E. coli*, ATP synthesis coupled proton transport and lipopolysaccharide biosynthesis were the top gene ontoloty (GO) terms (Fig. 3.24), which is consistent with the findings

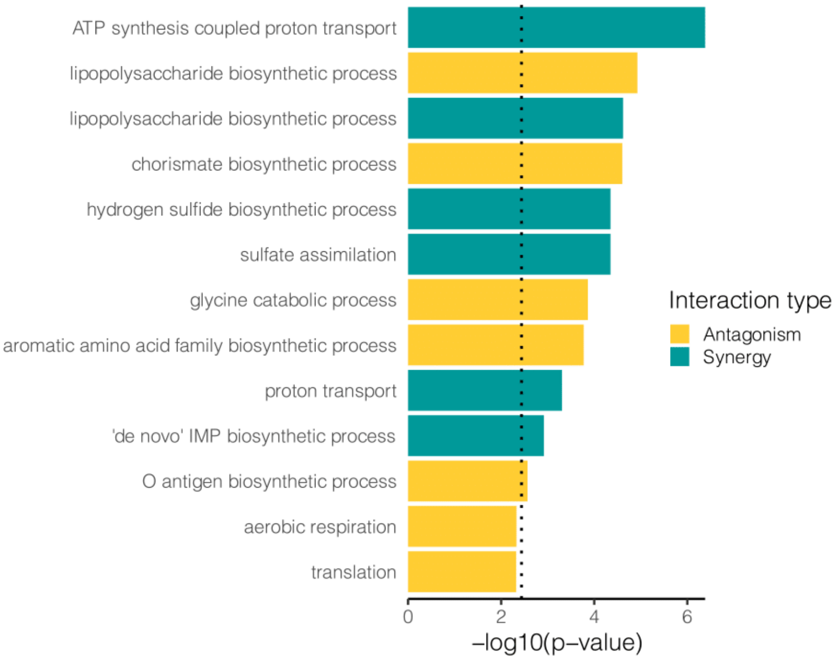


Fig. 3.24 Gene ontology (GO) enrichment in *E. coli*. The dotted line indicates FDR = 0.1. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

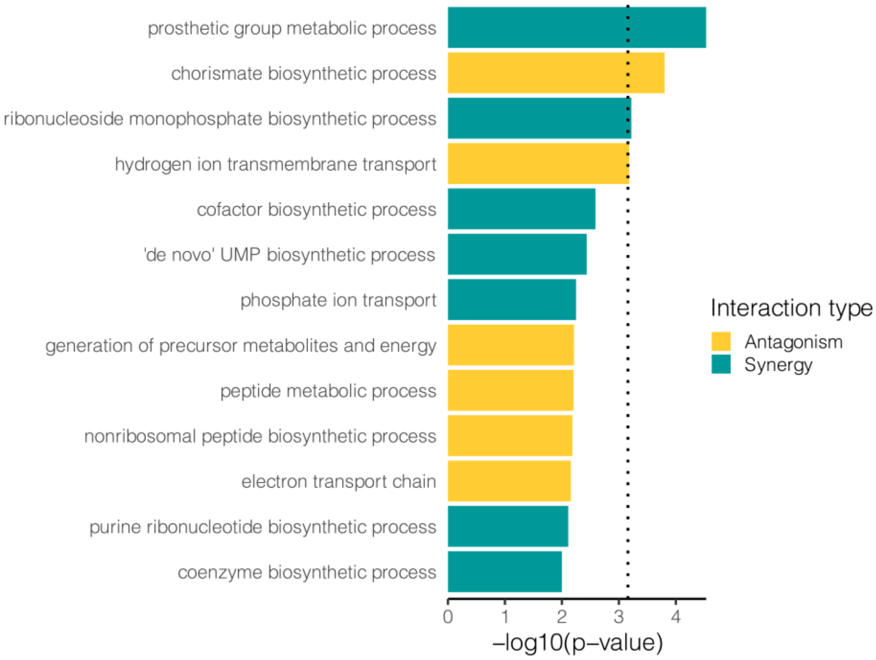


Fig. 3.25 Gene ontology (GO) enrichment in *S. typhimurium*. The dotted line indicates FDR = 0.1. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Pfalz (2017).

of the study by Chevereau and Bollenbach (2015), in which pairwise combinations of 6 antibiotics were screened in a genome-wide deletion library in *E. coli*. Some other processes enriched among the interaction-associated genes in *E. coli* were chorismate biosynthetic process, hydrogen sulfide biosynthesis, glycine catabolic process and proton transport.

Prosthetic group metabolic process (Fig. 3.25) was the top gene ontology (GO) term in *S. typhimurium*. The only shared GO term which was significant in *E. coli* and *S. typhimurium* was the chorismate biosynthetic process, which in both species was associated with antagonisms. The complete list of genes belonging to the top GO terms in *E. coli* and *Salmonella* is provided in Supplementary Table S5.

Since ATP synthesis was an important biological process for synergies, I chose to focus on the *atpA* gene, for which more synergies were observed than expected by chance (Fig. 3.26a) if both drugs of a combination were losing fitness in the *atpA* gene deletion in *E. coli* (i.e. if *atpA* = -2). I visualized the synergy network of the drugs losing fitness in Δ*atpA* (Fig. 3.26b) and found that all synergizing drugs were cell wall targeting compounds. While cell wall drugs have a substantial number of within-class synergies (Brochado et al., 2018), compounds losing fitness in Δ*atpA* formed a nearly complete graph (Fig. 3.26b). Another way of visualizing this trend is to plot drug interaction score against combination gene state (Fig. 3.27), which additionally shows that the combinations with *atpA* = -2 state had lower Bliss scores (were more synergistic).

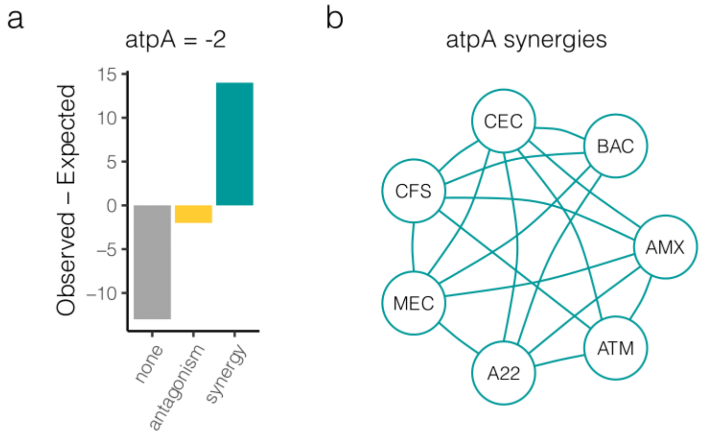


Fig. 3.26 Synergies associated with *atpA* in *E. coli* (Δ*atpA* = *atpA* gene deletion strain). a) More synergies were observed than expected by chance in combinations with drugs both losing fitness in Δ*atpA*. b) Synergies among the drugs losing fitness in Δ*atpA*. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

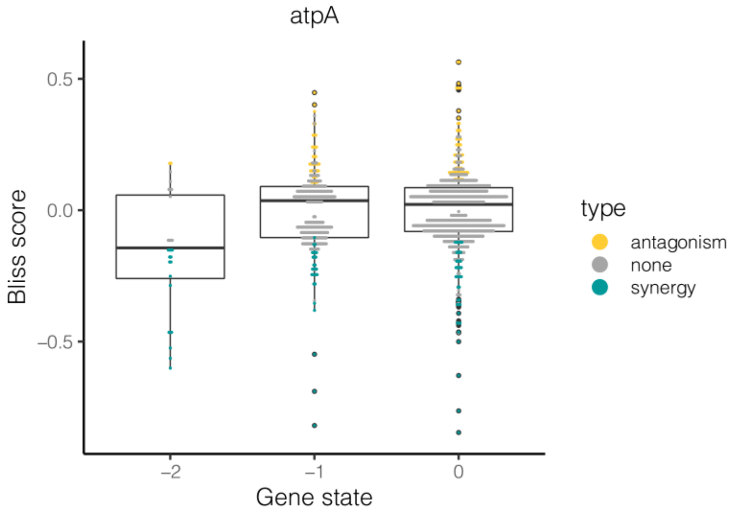


Fig. 3.27 Bliss scores stratified by *atpA* gene state in *E. coli* (Δ*atpA* = *atpA* gene deletion strain). Combination gene state indicates whether drugs of a combination gain (> 0) or lose (< 0) fitness in Δ*atpA*. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

Penicillin-binding protein *mrcB* was another gene that was important for cell wall synergies. In combinations with both drugs losing fitness in Δ*mrcB*, I observed more synergies than expected by chance (Fig. 3.28a). Synergies of cell wall targeting drugs with trimethoprim (TMP) were particularly overrepresented in such combinations (Fig. 3.28b).

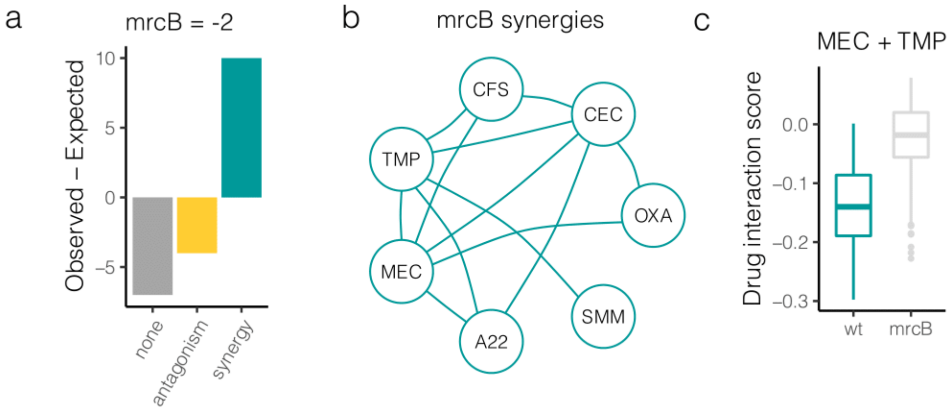


Fig. 3.28 Synergies associated with *mrcB* in *E. coli* (Δ*mrcB* = *mrcB* gene deletion strain). a) More synergies were observed than expected by chance in combinations with drugs both losing fitness in Δ*mrcB*. b) Synergies among the drugs losing fitness in Δ*mrcB*. c) Mecillinam (MEC) + trimethorpim (TMP) synergy becomes neutral in Δ*mrcB*. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011). Data in c) produced by E. Cacace.

We probed trimethoprim (TMP) + mecillinam (MEC) synergy in Δ*mrcB* and found (Fig. 3.28c) that this combination was neutral in the deletion strain, which means that *mrcB* gene is causal for TMP + MEC synergy. *mrcB* is expected to play a similar role in trimethoprim synergies with A22, cefsulodin (CFS) and cefaclor (CEC) (Fig. 3.28b).

Bliss scores of combinations with both drugs losing fitness in Δ*mrcB*, i.e. with gene state *mrcB* = -2, were lower than those of other combinations (Fig. 3.29). Furthermore, I found that combinations in which one drug gains and the other loses fitness in Δ*mrcB* were also primarily synergistic and had lower (Fig. 3.29) drug interaction scores.

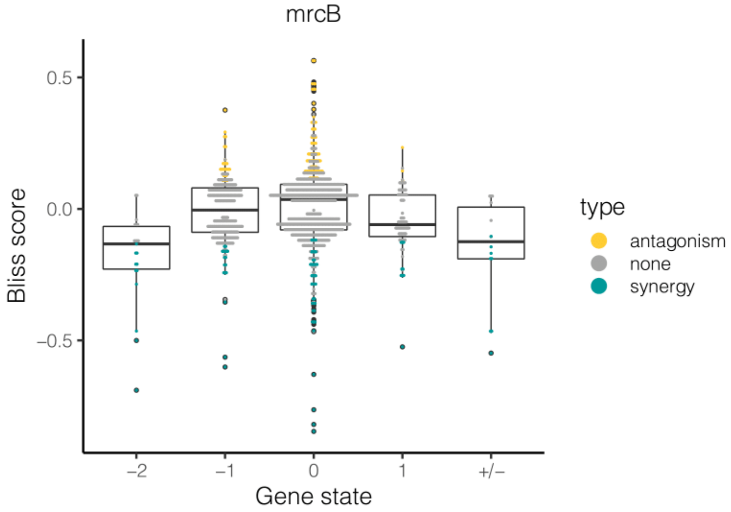


Fig. 3.29 Bliss scores stratified by *mrcB* gene state in *E. coli* (Δ*mrcB* = *mrcB* gene deletion strain). Combination gene state indicates whether drugs of a combination gain (> 0) or lose (< 0) fitness in Δ*mrcB*. Gene state +/- indicates that one drug of a combination gains and the other drug loses fitness in Δ*mrcB*. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

In both *E. coli* and *S. typhimurium*, chorismate and aerobic respiration genes were associated with antagonisms (Table S5). Therefore I explored the relationship between *ubiI*, an enzyme required for ubiquinone biosynthesis, and antagonisms. In combinations with one drug losing fitness in Δ*ubiI*, I observed more antagonisms than expected by chance (Fig. 3.30a). This association was primarily driven by antagonisms of fosfomycin (FOF) that was losing fitness in Δ*ubiI* (Fig. 3.30b). Fosfomycin interacted antagonistically with protein and DNA synthesis inhibitors (Fig. 3.30b). Similar to fosfomycin, another promiscuous antagonist was pyocyanin (Supplementary Fig. S10), which also loses fitness in Δ*ubiI* and interacts antagonistically with protein and DNA synthesis inhibitors. In general, combinations with drugs losing fitness in Δ*ubiI* had higher (more antagonistic) Bliss scores (Fig. 3.31).

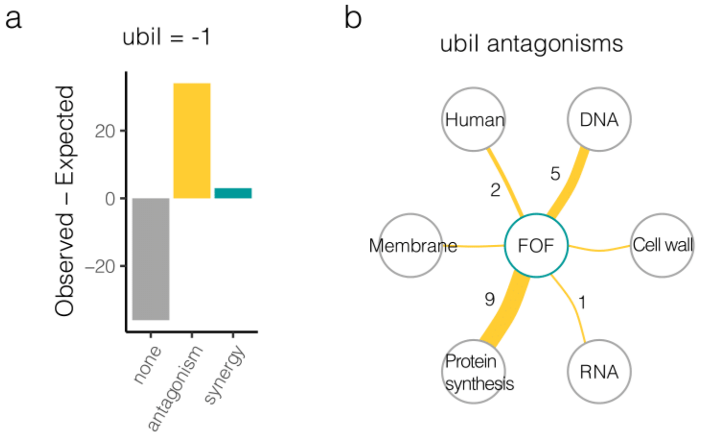


Fig. 3.30 Antagonisms associated with *ubiI* in *E. coli* (Δ*ubiI* = *ubiI* gene deletion strain). a) More antagonisms were observed than expected by chance in combinations in which one of the drugs loses fitness in Δ*ubiI*. b) Antagonisms of fosfomycin (FOF), a drug that loses fitness in Δ*ubiI*. Edge width is labelled and indicates the number of antagonisms occurring between fosfomycin and a drug class. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

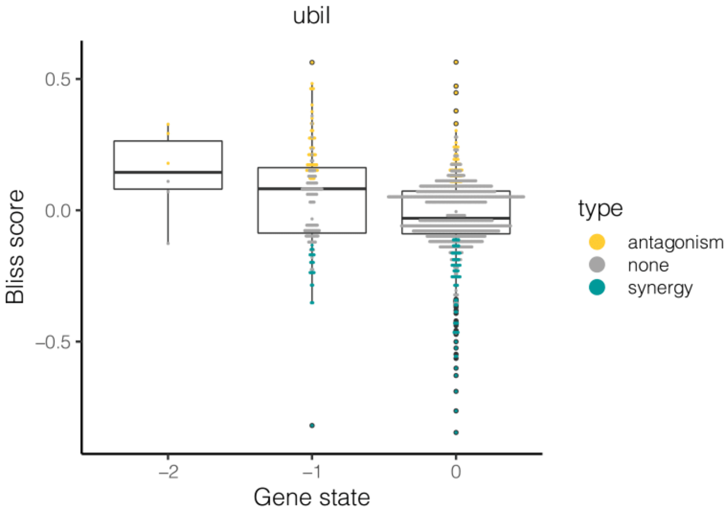


Fig. 3.31 Bliss scores stratified by *ubiI* gene state in *E. coli* (Δ*ubiI* = *ubiI* gene deletion strain). Combination gene state indicates whether drugs of a combination gain (> 0) or lose (< 0) fitness in Δ*ubiI*. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

In *E. coli*, nicotinate phosphoribosyltransferase *pncB* was the gene with the strongest association with antagonisms. As was the case for *ubiI*, combinations with one drug losing fitness in Δ*pncB* were enriched in antagonisms (Fig. 3.32a). The underlying antagonism network was identical to that of *ubiI*, with fosfomycin and pyocyanin interacting antagonistically with protein synthesis inhibitors, DNA targeting compounds and human-targeted drugs (Fig. 3.32b and Supplementary fig. S11). Similarly, combinations with drugs losing fitness in Δ*pncB* had higher (more antagonistic) Bliss scores (Fig. 3.33).

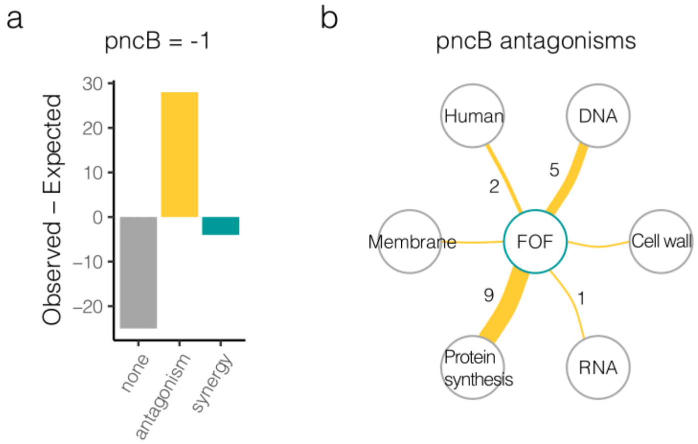


Fig. 3.32 Antagonisms associated with *pncB* in *E. coli* (Δ*pncB* = *pncB* gene deletion strain). a) More antagonisms were observed than expected by chance in combinations in which one of the drugs loses fitness in Δ*pncB*. b) Antagonisms of fosfomycin (FOF), a drug that loses fitness in Δ*pncB*. Edge width is labelled and indicates the number of antagonisms occurring between fosfomycin and a drug class. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

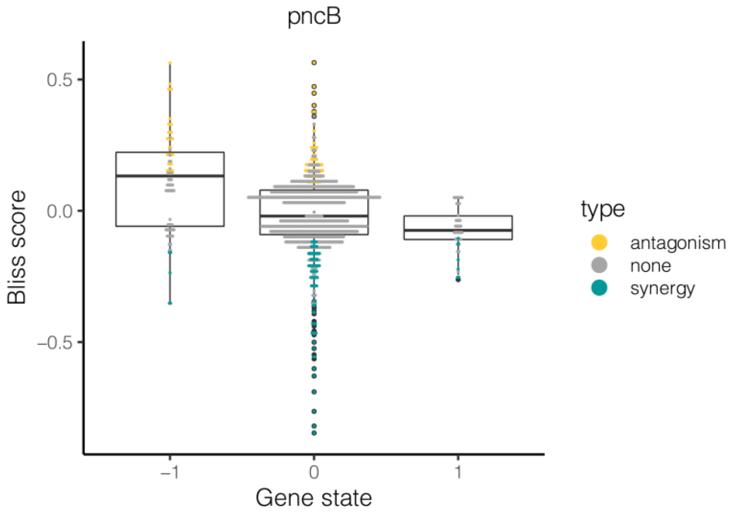


Fig. 3.33 Bliss scores stratified by *pncB* gene state in *E. coli* (Δ*pncB* = *pncB* gene deletion strain). Combination gene state indicates whether drugs of a combination gain (> 0) or lose (< 0) fitness in Δ*pncB*. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

3.3.9 Understanding the role of uncharacterized genes

Y-ome genes with little to no annotation make up about 33% of the genes in *E. coli* (Ghatak et al., 2019). I found that 34% of STM14 loci probed in the *S. typhimurium* chemogenomic screen did not have any annotation or had only putative functional roles assigned.

Overall uncharacterized genes were underrepresented among the interaction-associated hits with 21% in *E. coli* and 28% in *S. typhimurium* (Fig. 3.34a) – less than the proportion of uncharacterized genes in respective genomes. Nevertheless I observed an appreciable number of uncharacterized genes associated with drug interactions (Fig. 3.34a).

For the Y-ome genes in *E. coli*, I constructed a network (Fig. 3.34b) based on gene-gene correlations in the chemogenomic screen (see Materials, Subsection 3.2.12). Similarly, I embedded genes with no known function in the gene-gene correlation network of *S. typhimurium* (Fig. 3.34c). Many uncharacterized genes in *E. coli* were falling into the dense cluster of genes (Fig. 3.34b) important for aerobic respiration. In *Salmonella* the gene-gene correlation network was much sparser (Fig. 3.34c) and only few uncharacterized genes had annotated neighborhoods.

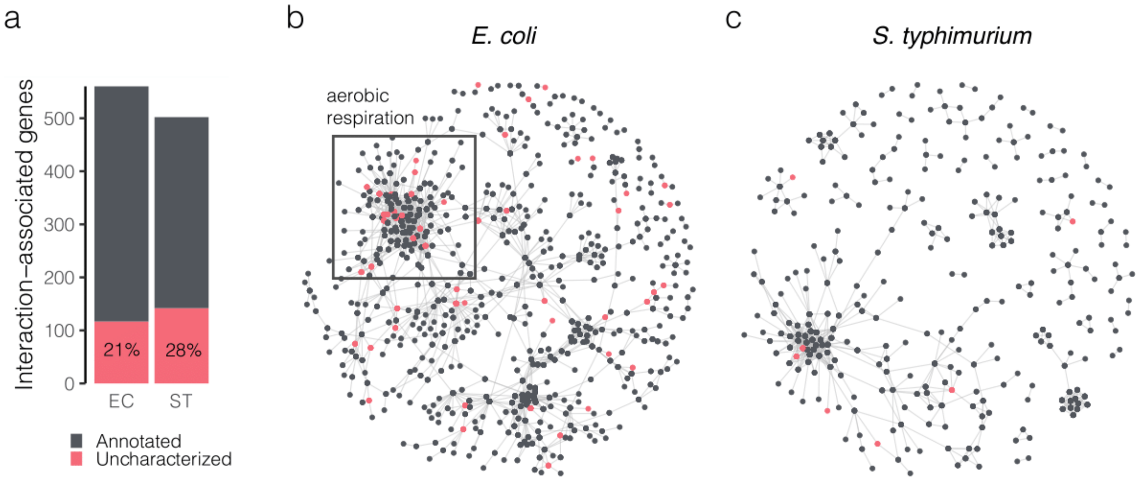


Fig. 3.34 Uncharacterized genes associated with drug-drug interactions (EC = *E. coli*, ST = *S. typhimurium*). a) Number of annotated and uncharacterized interaction-associated genes in EC and ST. Proportions are shown for uncharacterized genes. b) Gene-gene correlation network in *E. coli* and in c) *S. typhimurium*. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

The majority (99 out of 117) of uncharacterized genes in *E. coli* were associated with antagonisms. In the following analysis I prioritized putative transporter and transmembrane proteins, as this class of proteins can potentially promote drug extrusion.

A putative metabolite transporter *ydjK* was associated with antagonisms. The first-degree neighbors of *ydjK* (Fig. 3.35a) were *malS* and *sad*, genes with metabolic activity, and *yibA*, an uncharacterized gene involved in response to antibotics and radiation (Han et al., 2010; Sargentini et al., 2016). Combinations, in which both drugs gain fitness in Δ*ydjK*, were enriched in antagonisms (Fig. 3.35b). Additionally, drugs gaining fitness in Δ*ydjK*, such as trimethoprim (TMP), were promiscuous antagonists (Fig. 3.35b).

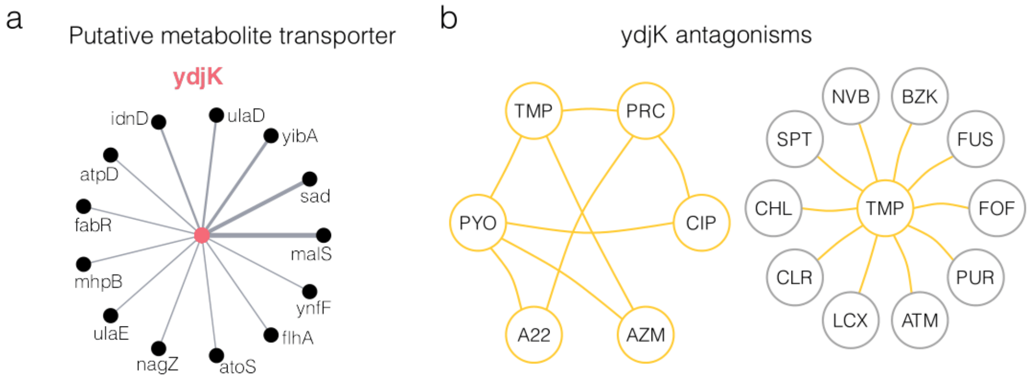


Fig. 3.35 Putative metabolite transporter *ydjK*. a) Gene-gene correlation network centered on *ydjK*. Only gene pairs with correlation $r > 0.4$ are linked. Edge thickness is proportional to Pearson correlation between the connected nodes. b) Antagonisms (edges) are shown for combinations of drugs gaining fitness in Δ*ydjK* (yellow nodes) in *E. coli*. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

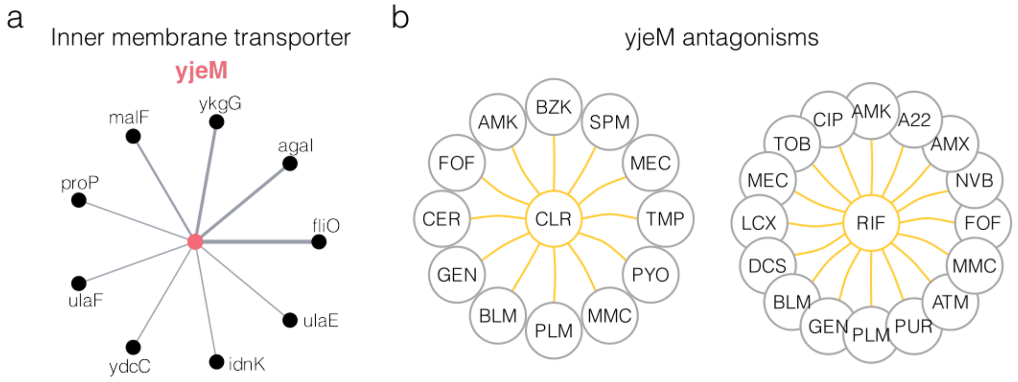


Fig. 3.36 Inner membrane transporter *yjeM*. a) Gene-gene correlation network centered on *yjeM*. Only gene pairs with correlation $r > 0.4$ are linked. Edge thickness is proportional to Pearson correlation between the connected nodes. b) Antagonisms (edges) are shown for combinations, in which one drug (CLR, RIF) gains fitness in Δ*yjeM* in *E. coli*. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

An inner membrane transporter *yjeM* was another uncharacterized gene that was associated with antagonisms. In *E. coli* chemical genetics, *yjeM* was correlated with genes with metabolic activity (*agaI*, *malF*, *ulaE*) and with proline/betaine transporter *proP* (Fig. 3.36a). Drugs gaining fitness in Δ*yjeM*, such as clarithromycin (CLR), had many antagonisms (Fig. 3.36b). In addition to clarithromycin, this trend was also observed for rifampicin (RIF), which also gained fitness in Δ*yjeM* (Fig. 3.36b).

As mentioned in Subsection 3.3.8, many genes important for aerobic respiration were associated with antagonisms. Transmembrane protein *ycaP* (Fig. 3.37) and inner membrane protein *yaiY* (Fig. 3.38) were among the uncharacterized genes located in the dense cellular respiration cluster in Fig. 3.34b. In the chemogenomic screen in *E. coli*, *ycaP* was strongly correlated with respiration genes *sdhE*, *ubiF* and *arcA* (Fig. 3.37a). The first-degree neighbors of *yaiY* (Fig. 3.38a) in the gene-gene correlation network were *ubiI* and aerobic respiration control genes *arcA* and *arcB*.

Similar to *ubiI*, drugs losing fitness in Δ*ycaP* and Δ*yaiY* were promiscuous antagonists. Pyocyanin (PYO) and benzalkonium (BZK), which both lose fitness in Δ*ycaP* (Fig. 3.37b), interacted antagonistically with protein synthesis inhibitors, DNA and cell wall targeting compounds. Likewise, fosfomycin (FOF) and benzalkonium (BZK), which lose fitness in Δ*yaiY* (Fig. 3.38b), had numerous antagonistic interactions with cell wall targeting compounds, protein and DNA synthesis inhibitors as well as human-targeted drugs.

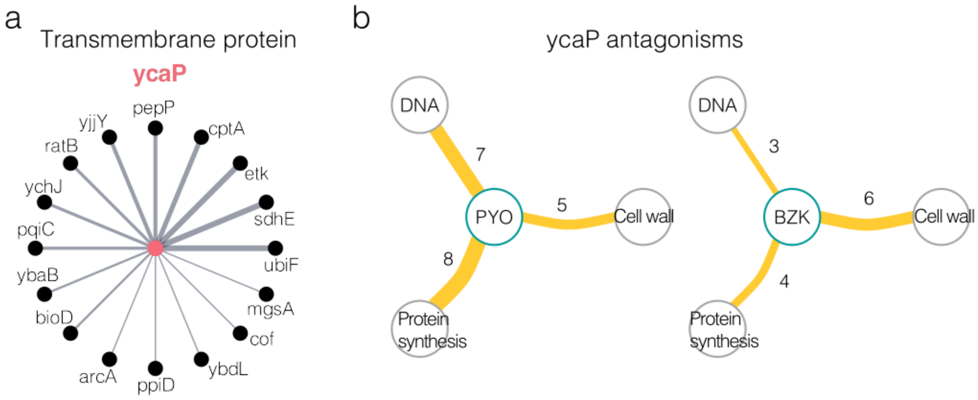


Fig. 3.37 Transmembrane protein *ycaP*. a) Gene-gene correlation network centered on *ycaP*. Only gene pairs with correlation $r > 0.4$ are linked. Edge thickness is proportional to Pearson correlation between the connected nodes. b) Both pyocyanin (PYO) and benzalkonium (BZK) lose fitness in Δ*ycaP* in *E. coli*. Numerous antagonisms were observed between these drugs and shown drug classes. Edge width and label indicate number of antagonisms. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

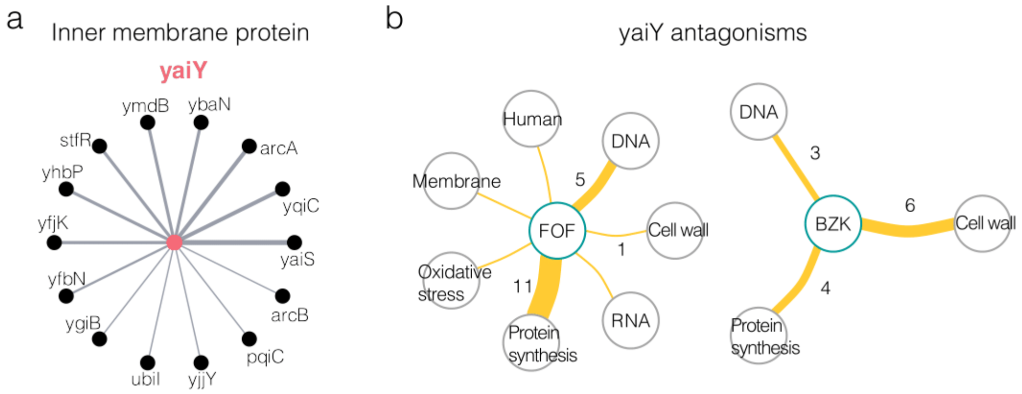


Fig. 3.38 Inner membrane protein *yaiY*. a) Gene-gene correlation network centered on *yaiY*. Only gene pairs with correlation $r > 0.4$ are linked. Edge thickness is proportional to Pearson correlation between the connected nodes. b) Both fosfomycin (FOF) and benzalkonium (BZK) lose fitness in Δ*yaiY* in *E. coli*. Numerous antagonisms were observed between these drugs and shown drug classes. Edge width and label indicate number of antagonisms. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

3.3.10 Prediction based on chemogenomic profiles

Having identified genes associated with antagonisms and synergies (Subsection 3.3.8), I next turned to the question whether one can predict drug interactions based on chemogenomic profiles, particularly on the subset of interaction-associated genes. I applied ensemble models (random forest, XGBoost) and neural networks (see Methods, Subsection 3.2.13) with the objective to find the best predictive model for drug interactions.

To select the best hyperparameter set for each model, I used 20-fold cross-validation (Fig. 3.39). In each cross-validation fold I chose 15 drugs at random, all pairwise combinations of which were withheld from training. The trained model was then validated on the held-out combinations (see Methods, Subsection 3.2.14).

Based on cross-validation with an extensive hyperparameter grid (see Methods, Subsection 3.2.14), I found that XGBoost had superior performance in terms of aggregated average precision (Fig. 3.40)) in *E. coli*. While I could achieve somewhat higher average precision using neural networks in *S. typhimurium* (Fig. 3.40), I observed much higher variability in terms of average precision when using neural nets. Due to the stronger dependence of deep learning models on a specific hyperparameter set, I chose XGBoost classifier, which was less sensitive to the choice of hyperparameters (Fig. 3.40).

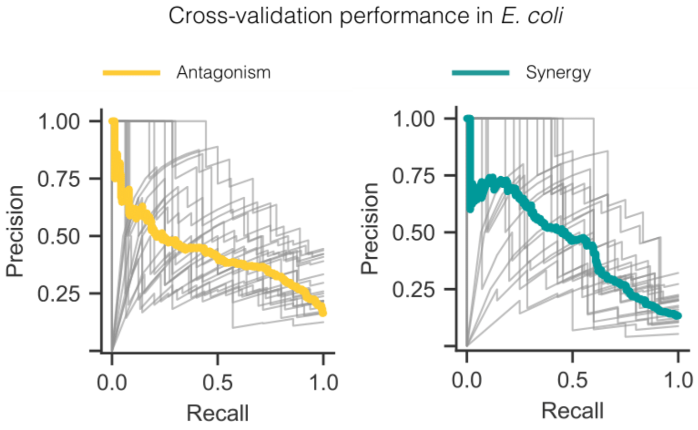


Fig. 3.39 Cross-validation performance of an XGBoost classifier for antagonism and synergy prediction in *E. coli*. 20 cross-validation folds were used to assess the performance of the supervised model. Average precision (AP) for antagonism prediction was 0.43 and 0.45 for synergy prediction. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011).

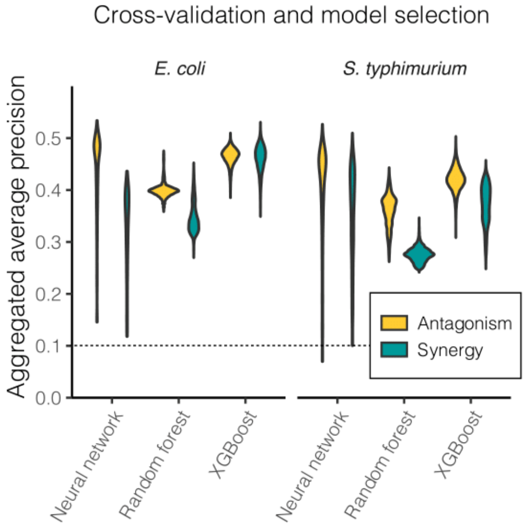


Fig. 3.40 Model selection in *E. coli* and *S. typhimurium*. Aggregated average precision (y-axis) of antagonism and synergy predictions is shown for 3 machine learning algorithms (x-axis) for various hyperparameter combinations (violin plots). 20-fold cross-validation was used to compare the performance of random forest, XGBoost and deep neural net classifiers. For each hyperparameter combination, average precision (AP) was aggregated across 20 cross-validation folds. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011) and Pfalz (2017). Neural network results produced by S. Bassler.

Using an independent test set, which consisted of pairwise combinations of 5 antibiotics (Fig. 3.41a) probed in *E. coli*, I assessed the performance of the chosen machine learning model (XGBoost). All 3 synergies in the test set were predicted correctly (Fig. 3.41a) and XGBoost classifier achieved AUCROC = 1 and AP = 1 for synergy prediction (Fig. 3.41b). None of the antagonisms in the test set was predicted by the model (Fig. 3.41a) when maximum scores of one-versus-rest predictions were considered. Nevertheless, XGBoost classifier achieved AUCROC = 0.75 and AP = 0.45 (Fig. 3.41b) for antagonism prediction on the test set.

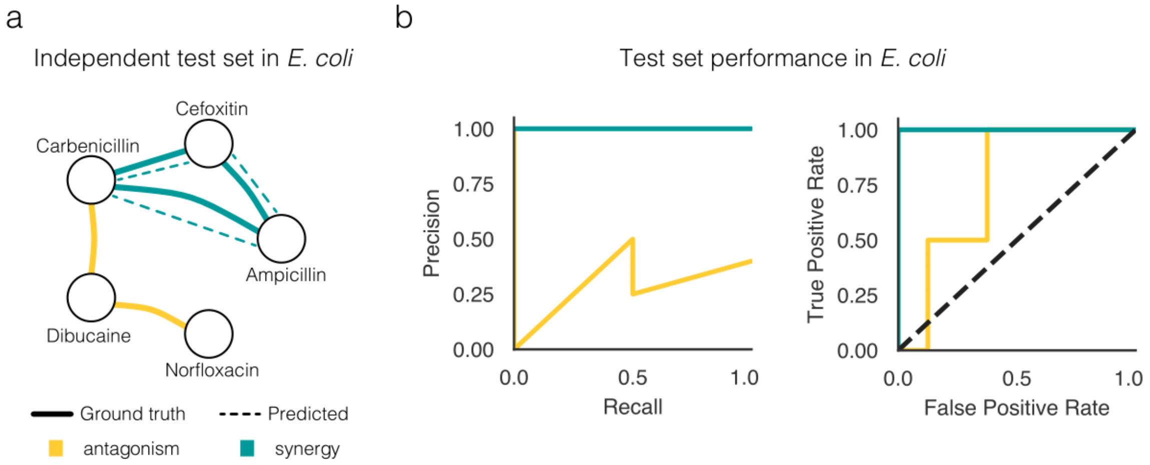


Fig. 3.41 Test set performance of the selected model in *E. coli*. a) XGBoost classifier with the best cross-validation performance was used to predict drug interactions (dashed lines in the network) in an independent test in *E. coli*. Drug interactions (solid edges) in the test set were determined experimentally. b) Precision-recall and ROC curves on the test set data in *E. coli*. Data in a) produced by E. Cacace. Chemogenomic data source: Nichols et al. (2011).

3.3.11 Feature importance and interpretable machine learning

In tree-based models, feature importance (see Subsection 3.1.4) may shed some light on how drug interaction prediction works. For the selected XGBoost model (see Subsection 3.3.10), I extracted top 20 predictive features from each cross-validation run and ranked all features (see Methods, Subsection 3.2.15) by their prevalence in the aggregated list in *E. coli* (Fig. 3.42) and *S. typhimurium* (Fig. 3.43). Since I used the one-hot encoding scheme, I could check what gene state was important for predictions. For antagonism predictions in *E. coli* (Fig. 3.42a) the top predictive features were genes in states {-1, 0, 1}, i.e. whether only one drug of a combination loses (-1) or gains (+1) fitness in a gene deletion or whether there is no drug-gene interaction (0). Conversely, for synergy prediction in *E. coli* (Fig. 3.42b) the most important features were genes in states {-2, 2, +/-}, i.e. whether both drugs of a combination lose or gain fitness in a certain gene deletion.

Consistent with the findings reported in Subsection 3.3.8, ATP genes (*atpA*, *atpF*, *atpC*) and proteins involved in cell wall organization and lipopolysaccharide biosynthesis (*slt*, *fepE*, *wcaF*) were important for synergy prediction in *E. coli* (Fig. 3.42b). Antagonism predictions in *E. coli* (Fig. 3.42a) were driven by the genes with the strongest associations with antagonisms (*ubiI*, *pncB*), which were described in Subsection 3.3.8.

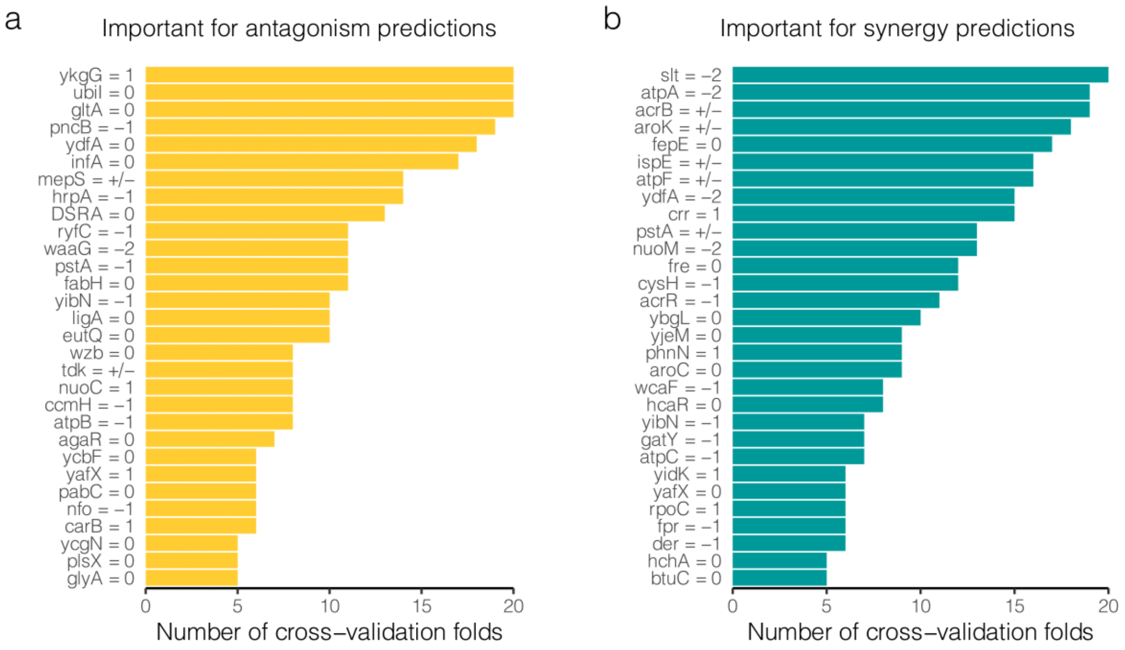


Fig. 3.42 Most important genes in *E. coli* for a) antagonism and b) synergy predictions. The *x*-axis shows the number of cross-validation folds, in which each feature was one of the top 20 features based on variable importance. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011).

Interestingly, genes important for antagonism and synergy prediction in *Salmonella* (Fig. 3.43) were distinct from those in *E. coli* (Fig. 3.42). For synergy prediction in *Salmonella* (Fig. 3.43b), an ATP binding gene *deaD*, an amino acid transporter *gltI*, an enterobactin synthetase *entF*, an ion transporter *fre* and a cellular respiration gene *cyoA* were among the top predictive features with gene state {-2}. For antagonism predictions in *S. typhimurium* (Fig. 3.43a), an effector protein *steB*, a transcriptional regulator *soxS*, a protease *prc*, an ATP-binding gene *pgtB* and a cellular respiration gene *nuoE* were important.

While Fig. 3.42–3.43 provide an overview of global feature importance, they do not reveal for which drug combinations these features were important. To obtain feature importance for

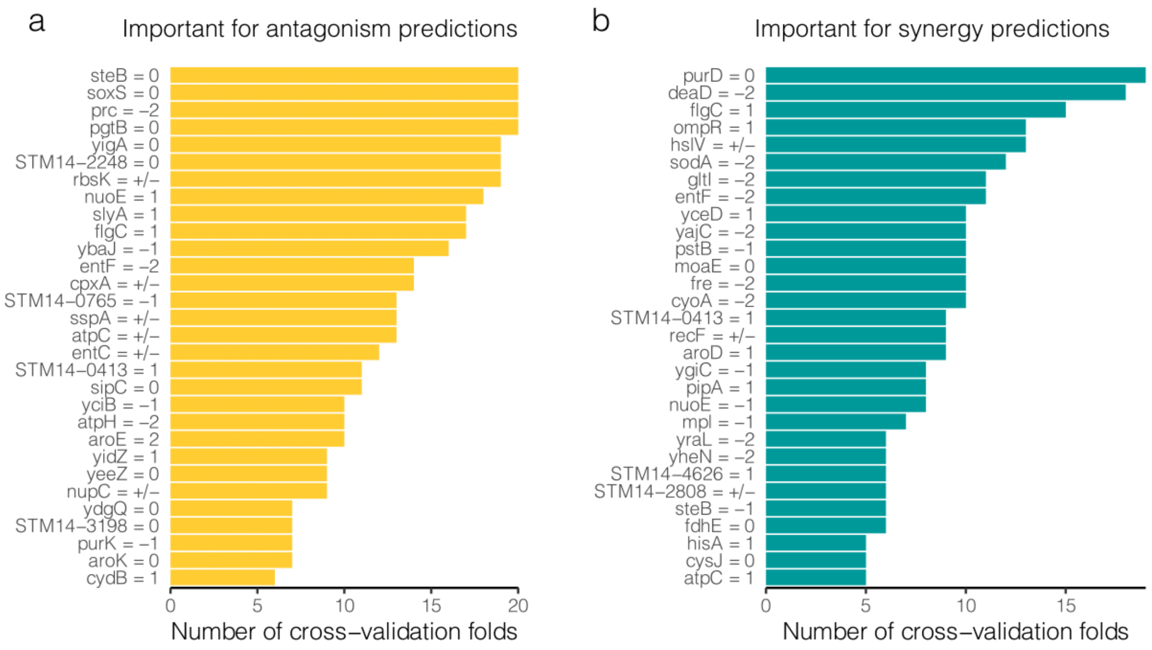


Fig. 3.43 Most important genes in *S. typhimurium* for a) antagonism and b) synergy predictions. The x-axis shows the number of cross-validation folds in which each feature was one of the top 20 features based on variable importance. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Pfalz (2017)

individual predictions, I used SHAP values estimated based on the trained XGBoost model (see Methods, Subsection 3.2.16).

In *E. coli*, prediction of synergies of β-lactam antibiotics (drug classes in Supplementary Table S4) was driven primarily by *atpA*, *slt* and *mrcB* (Fig. 3.44). In particular, β-lactam combinations with both drugs losing fitness in Δ*atpA*, Δ*slt* and Δ*mrcB* were more likely to be predicted synergistic (Fig. 3.44) by the trained XGBoost model.

Compounds targeting DNA gyrase (see Supplementary Table S4) had many antagonistic interactions in *E. coli*. DNA gyrase combinations with one drug losing fitness in Δ*pncB*, Δ*ligA* or having an interaction with *ubiI* were likelty to be predicted antagonistic (Fig. 3.45) by the XGBoost classifier.

Genes important for β-lactam synergy prediction in *S. typhimurium* (Fig. 3.46) differed from those in *E. coli* (Fig. 3.44). Double gene states {-2, 2, +/-} played a minor role in predicting β-lactam synergies in *Salmonella* (Fig. 3.46). *dacA*, a protein involved in peptidoglycan biosynthesis, was the only gene with state {-2} that was predictive of β-lactam synergies in *S. typhimurium* (Fig. 3.46). β-lactam combinations with one drug gaining fitness in Δ*prc* were more likely to be classified synergistic (Fig. 3.46). Likewise, β-lactam combinations with drugs interacting with a cellular respiration gene *ubiX*, a purine biosynthesis gene *purA* and a peptidoglycan recycling gene *mpl* were predicted to be synergistic (Fig. 3.46).

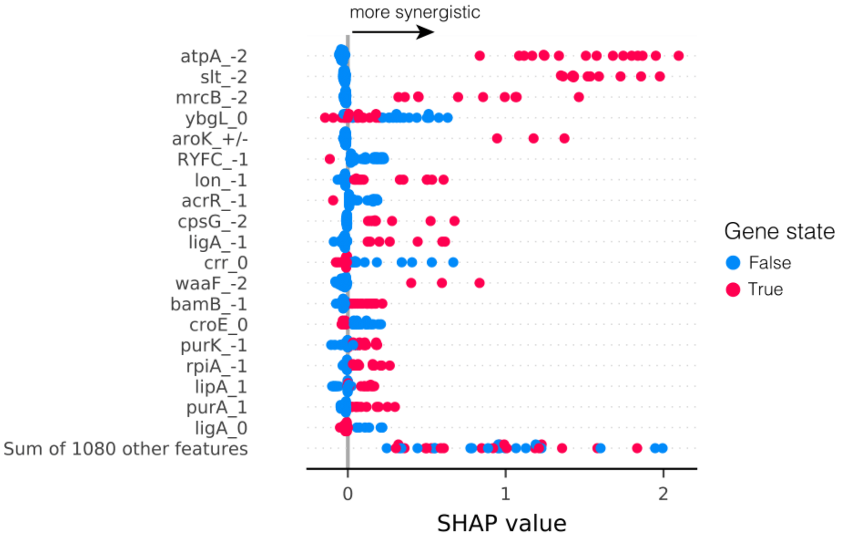


Fig. 3.44 Most important genes for β-lactam synergy predictions in *E. coli*. Positive SHAP values (x-axis) contribute to higher synergistic scores. Only synergistic combinations (points) with one or both drugs belonging to the β-lactam class are shown. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

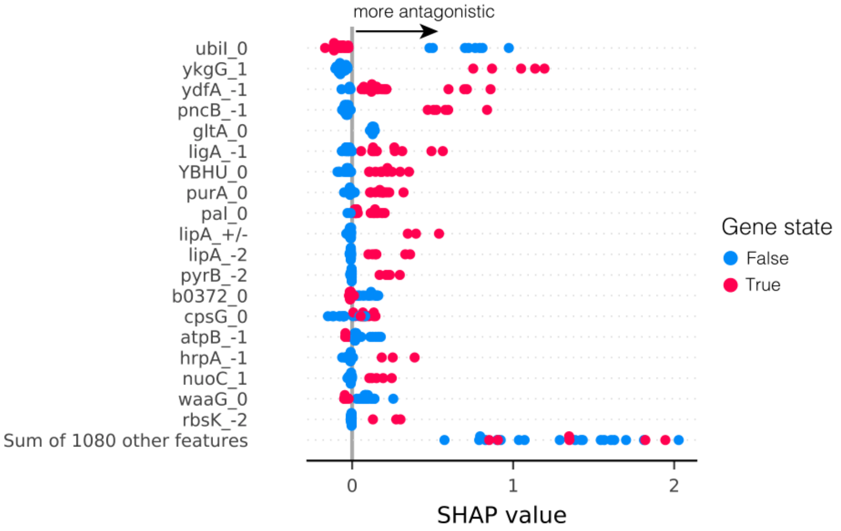


Fig. 3.45 Most important genes for prediction of antagonisms of DNA gyrase-targeting drugs in *E. coli*. Positive SHAP values (x-axis) contribute to higher antagonistic scores. Only antagonisms (points) with one or both drugs belonging to the DNA gyrase class are shown. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Nichols et al. (2011).

In *Salmonella*, antagonism predictions for drugs targeting DNA gyrase (Fig. 3.47) were driven by a putative mandelate racemase *STM14-4626*, a transcriptional regulator *soxS*, a phosphate transporter *pitA*, a nucleoside transporter *nupC* and a protein transporter *tatC*.

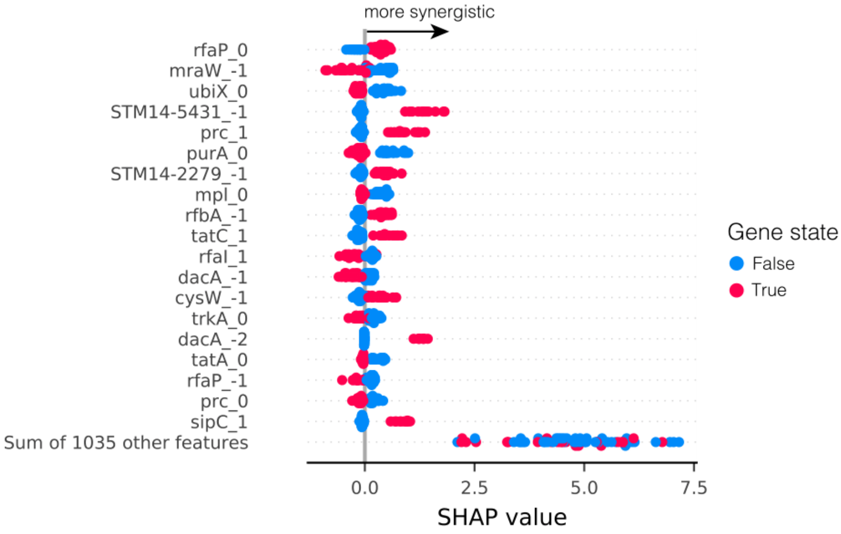


Fig. 3.46 Most important genes for β-lactam synergy predictions in *S. typhimurium*. Positive SHAP values (*x*-axis) contribute to higher synergistic scores. Only synergistic combinations (points) with one or both drugs belonging to the β-lactam class are shown. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Pfalz (2017)

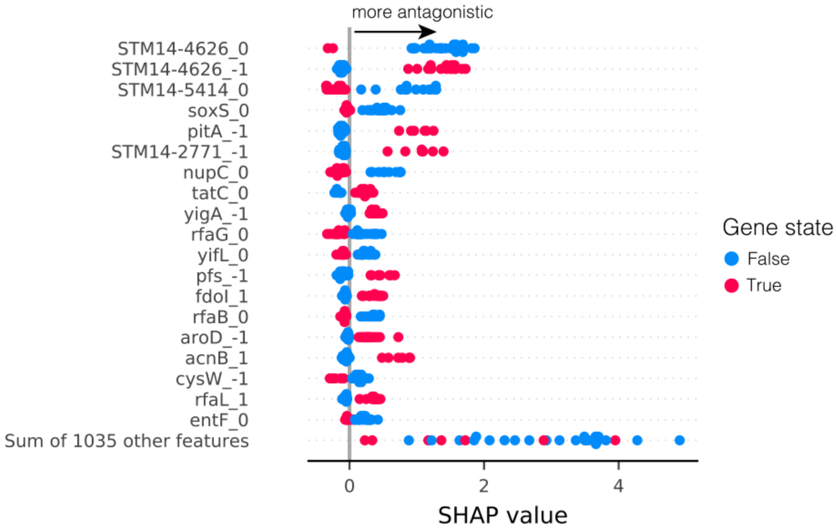


Fig. 3.47 Most important genes for prediction of antagonisms of DNA gyrase-targeting drugs in *S. typhimurium*. Positive SHAP values (*x*-axis) contribute to higher antagonistic scores. Only antagonisms (points) with one or both drugs belonging to the DNA gyrase class are shown. Drug interaction data: Brochado et al. (2018). Chemogenomic data: Pfalz (2017)

3.3.12 Directional network of drug interactions in Gram-negatives

Using single-drug and combination fitness data in 3 Gram-negative species (*E. coli*, *S. typhimurium*, *P. aeruginosa*), I inferred directional relationships in the drug interaction network (see Methods, Subsection 3.2.17). I generated a directional network of class-class interactions (Fig. 3.48) in Gram-negative species by aggregating compound classes (drug class annotation in Supplementary Table S4).

Aminoglycosides, a class of compounds relying on membrane potential for cell entry (Damper and Epstein, 1981), were weakened by multiple drug classes in antagonistic combinations (Fig. 3.48 and Supplementary Table S6). In particular, macrolides reduced the effect of aminoglycosides (Fig. 3.48) in 5 combinations. The antagonizing action of macrolides on aminoglycosides has been previously described in the literature (Nichols et al., 2017).

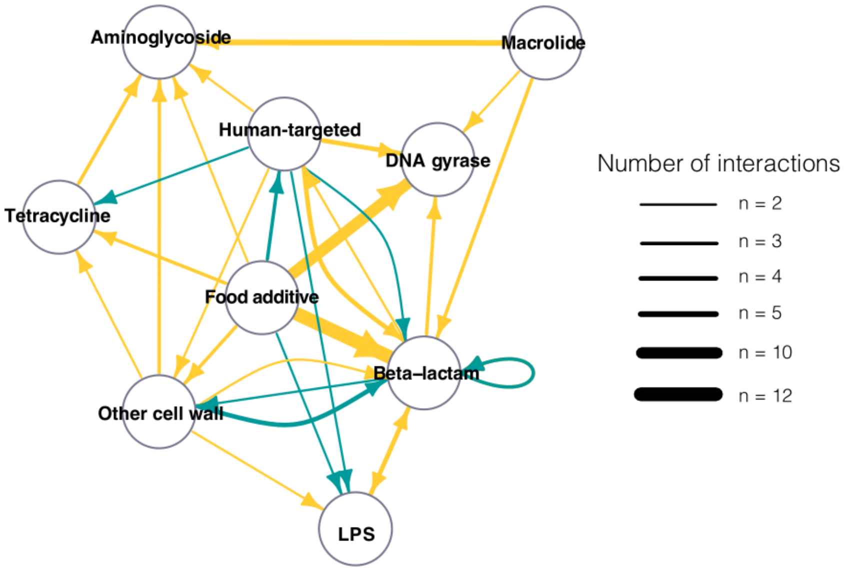


Fig. 3.48 Inferred directional network of drug interactions in Gram-negative species (*E. coli*, *S. typhimurium*, *P. aeruginosa*) aggregated by compound class. Edges indicate interactions between drug classes. Each arrow points to a drug class potentiated (for synergies) or weakened (for antagonisms) by the source node. Edge width is proportional to the number of class-class interactions. Edges with $n = 1$ were omitted. For drug class annotation refer to Supplementary Table S4. Drug interaction data source: Brochado et al. (2018).

Food additive class that includes vanillin, berberine, curcumin and caffeine (Supplementary Table S4) had the highest number of antagonistic interactions for which the direction could be inferred (Fig. 3.48 and Supplementary Table S6). The food additive class weakened the action of β-lactam antibiotics in 12 combinations and reduced the effect of DNA gyrase targeting compounds in 10 drug pairs (Fig. 3.48).

Directional synergies occurred between β-lactam antibiotics and other cell wall targeting compounds (Fig. 3.48 and Supplementary Table S6) with 2 combinations in which β-lactams potentiated other cell wall drugs and 4 synergies in which the direction was reversed. Within the β-lactam class, I observed the potentiation of one β-lactam antibiotic by another member of the same class in 3 combinations (Fig. 3.48 and Supplementary Table S6).

Interestingly, human-targeted compounds potentiated tetracyclines in 2 combinations and increased the activity of β-lactam antibiotics in 2 synergistic pairs (Fig. 3.48 and and Supplementary Table S6). However, human-targeted compounds also reduced the effect of aminoglycosides, other cell wall and DNA gyrase targeting drugs in several antagonistic combinations (Fig. 3.48).

3.4 Discussion

In this chapter I used chemogenomic data in *E. coli* and *S. typhimurium* to predict novel drug interactions (Subsection 3.3.10) and identify genes driving antagonism and synergy predictions (Subsection 3.3.11). In Subsection 3.3.8 I uncovered interaction-associated genes and highlighted key biological processes important for synergies and antagonisms in *E. coli* and *S. typhimurium*. The findings in Subsection 3.3.8 match the reported mechanisms in the literature (Chevereau and Bollenbach, 2015). In particular, drug efflux, ATP and lipopolysaccharide synthesis are known mechanisms of drug interactions that also appear as primary culprits in this analysis. The interaction-associated genes are likely to be implicated in a variety of drug interactions and these gene sets (Supplementary Table S5) can be used for targeted search of non-essential genes underlying antagonisms and synergies.

Using interaction-associated genes as predictors, I benchmarked 3 state-of-the-art machine learning algorithms (Subsection 3.3.10) and found that gradient boosting models (XGBoost) had the best drug interaction prediction performance in both *E. coli* and *S. typhimurium*.

There are a number of limitations when it comes to drug interaction prediction using chemogenomic profiles. Only non-essential genes are probed in gene deletion screens, which may mask some important information from the classifier. Additionally, combination chemogenomic profiles are “synthetic”, i.e. constructed from single-compound profiles and hence may not reflect real combination-gene states. Another problem is class imbalance due to the low occurrence of drug interactions. Supervised learning algorithms are known to perform poorly on imbalanced data (He and Garcia, 2009). I addressed the class imbalance problem by varying the class weight hyperparameter of machine learning models (Subsection 3.2.14).

In view of these caveats it is not surprising that the overall average precision across all cross-validation runs was about 0.45 for antagonisms and synergies in both *E. coli* and *S. typhimurium* (Subsection 3.3.10). For drug interaction prediction I used solely single-drug chemogenomic profiles to appraise the predictive power of chemical genetics data alone. Including compound physico-chemical properties may improve prediction accuracy.

In Subsection 3.3.11 I obtained feature importance for individual drug interaction predictions, making it possible to examine what genes are important for drug interactions occurring between particular compound classes. This more nuanced analysis may help advance our understanding as to which genetic features are essential for particular class-class interactions. This demonstrates that drug interaction mechanisms can be studied *in silico* using single-compound chemogenomic data, which is more tractable than testing drug combinations in genome-wide libraries. Mechanistic insights gained from such studies can improve the design of combinatorial drug therapies.

In most cases drug interactions exhibit inherent directionality, i.e. one drug modulates the efficacy of the other drug. Explicit directional modelling has been used to identify "perpertrator" and "victim" drugs. The general pharmacodynamic interaction model (GPDI) (Wicha et al., 2017) uncovered directional relationships in the yeast drug interaction network and found that the majority (67%) of drug interactions were unidirectional, with one drug potentiating or attenuating the action of the partner drug. In Subsection 3.3.12 I showed that the direction of drug interactions can be inferred from drug combination data.

The antagonistic interaction between macrolides and aminoglycosides is an example of an antagonism occurring due to reduced drug uptake. Macrolides alter the membrane potential which affects the import of aminoglycosides as the latter depend on proton motive force (PMF) for cellular uptake (Taber et al., 1987). Indeed in the directed graph of drug interactions (Subsection 3.3.12), macrolides weakened the action of aminoglycosides in 5 combinations. Interestingly, food additives such as vanillin and caffeine reduced the effect of many drugs in Gram-negative species (Subsection 3.3.12). Vanillin weakened the action of drugs in 13 antagonistic interactions (Supplementary Table S6). Vanillin is structurally similar to salicylate, a plant hormone known to induce drug extrusion (Sun et al., 2014), which largely explains its antagonistic behavior. Causal inference in drug interaction networks sheds light on which drugs increase drug uptake/extrusion and provides additional mechanistic clues that can be leveraged for improved design of combinatorial therapies.

Chapter 4

Discussion

4.1 Overcoming limitations of imaging-based assays

Image segmentation (Section 1.4) remains one of the most challenging steps in microscopy-based screens. Aside from standard cell-line-based systems, for which published workflows exist (Bray et al., 2016), establishing novel imaging-based screening techniques almost always requires *de-novo* analysis pipeline development. Even slight modifications in image acquisition settings can necessitate adjustments to previously used pipelines. There are several solutions to these challenges. First and foremost, it is important to encourage researchers to make their image analysis workflows public. Model zoos have already become commonplace in the deep learning community[1]. Leveraging transfer learning can expedite the analysis of image-based screens.

Another opportunity lies in unsupervised models handling images in domain-agnostic manner, which can be used for rapid data exploration. For instance, segmentation-free approaches do not require pixel-level labelling or as the name implies segmentation (Godinez et al., 2017). While these approaches have not come of age yet, there is a significant promise in these whole-image analysis tools (Godinez et al., 2017; Janssens et al., 2021).

Although thousands of morphological properties are computed in high-content screening assays, many of the image-based features are redundant due to high correlations among each other. Feature selection can be used to reduce redundancy and only choose the most informative features. While there are a number of feature selection methods (Subsection 3.1.4), most of these are *ad hoc* techniques with user-defined thresholds. Beyond filtering out noisy measurements based on replicate correlation, there is little consensus over what feature selection methods are best suited for a given problem (Caicedo et al., 2017).

---

[1] https://modelzoo.co

Another limitation of high-content screens is the lack of interpretability of certain morphological properties such as texture features or image moments (Subsection 2.2.4). To address this issue, one can find proxy features correlated with mathematical shape descriptors. Other approaches can leverage pixelwise feature importance measures such as saliency maps (Simonyan et al., 2014) or DeepLift (Shrikumar et al., 2017) used in deep learning.

Despite the tremendous progress in image analysis in recent years, segmentation of 3D confocal microscopy images remains complex. One of the issues is anisotropic resolution of fluorescence microscopy images (Weigert et al., 2017), in which axial resolution (z-axis) is often much lower than lateral resolution (xy-plane). In biomedical imaging, 3D convolutional neural networks have been used for the analysis of MRI and CT scans (Chen et al., 2018; Çiçek et al., 2016; Dolz et al., 2019). The adoption of 3D CNN models is still lagging in high-content screening applications. As organoid and organotypic models are becoming more popular (Rossi et al., 2018), analysis of 3D images will gain even more importance in the future (Boutros et al., 2015).

Lastly, machine-learning-based techniques are slowly taking over the field of bioimage analysis (Chandrasekaran et al., 2021; Moen et al., 2019). As of writing this chapter, tools for labelling ground truth remain somewhat limited in terms of installation and use especially for labelling 3D volumes or annotating videos. To facilitate AI-driven image analysis, next-generation labelling tools for high-dimensional image data should be developed.

4.2 Perspectives on combinatorial therapy design

Discovery of effective antibiotic combinations largely relies on high-throughput screening (Subsection 3.1.1), as mechanisms of most drug-drug interactions remain cryptic. Species and even strain specificity (Brochado et al., 2018) of drug-drug interactions makes it prerequisite to screen combinations in a variety of bacterial strains. At the same time species specificity of synergsitic interactions provides an opportunity to target pathogenic species without harming the microbiome (Cacace, 2021, Section 5.4). Studies investigating the effect of drug combinations on commensal organisms can provide valuable insights.

Drug-drug interactions can occur between co-administered antibiotic and non-antibiotic drugs. An appreciable number of drug-drug interactions were detected between human-targeted and antimicrobial drugs in *S. aureus* (Cacace et al., 2021). Previously, Brochado et al. (2018) reported numerous drug-drug interactions occuring between food additives and antibiotics. Systematic profiling of combinations of antibiotics with non-bactericidal compounds can shed light on this overlooked aspect.

To accelerate the discovery of effective drug combinations, it is extremely important to understand mechanisms of drug-drug interactions. Drug combinations can be used in conjunction with genetic perturbations to elucidate drug-drug interaction mechanisms (Subsection 3.1.2). The scale of these experiments, however, is daunting as drug combinations are screened in genome-wide mutant libraries. Therefore studying drug-drug interactions using single-drug chemogenomic profiles is an attractive alternative. In Subsection 3.3.8, I showed that it is possible to identify genes associated with drug-drug interactions, which narrows down the search for genes underlying these interactions. Furthermore, feature importance of individual predictions (Subsection 3.3.11) can be utilized to determine genes driving specific class-class interactions. This lays the foundation for mechanism-informed design of drug combinations.

In preceding studies researchers used single-drug features to predict drug-drug interactions. One of my primary objectives was to assess the predictive power of single-drug chemogenomic profiles. The prediction performance of machine learning models using chemogenomic data was modest, with average precision of about 0.45 (Subsection 3.3.10) for synergy and antagonism predictions in both studied species. However, precision of this order of magnitude is not uncommon in this domain (Bansal et al., 2014; Menden et al., 2019). In future studies, integrating physico-chemical and chemogenomic data could boost the classification performance.

References

Archer, K. J. and Kimes, R. V. (2008). Empirical characterization of random forest variable importance measures. *Comput. Stat. Data Anal.*, 52(4).

Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K. W., Schindelin, J., Cardona, A., and Seung, H. S. (2017). Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification. *Bioinformatics*.

Argelaguet, R., Velten, B., Arnol, D., Dietrich, S., Zenz, T., Marioni, J. C., Buettner, F., Huber, W., and Stegle, O. (2018). Multi-Omics Factor Analysis—a framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.*, 14(6).

Aslam, B., Wang, W., Arshad, M. I., Khurshid, M., Muzammil, S., Rasool, M. H., Nisar, M. A., Alvi, R. F., Aslam, M. A., Qamar, M. U., Salamat, M. K. F., and Baloch, Z. (2018). Antibiotic resistance: a rundown of a global crisis.

Atanasov, A. G., Zotchev, S. B., Dirsch, V. M., Orhan, I. E., Banach, M., Rollinger, J. M., Barreca, D., Weckwerth, W., Bauer, R., Bayer, E. A., Majeed, M., Bishayee, A., Bochkov, V., Bonn, G. K., Braidy, N., Bucar, F., Cifuentes, A., D'Onofrio, G., Bodkin, M., Diederich, M., Dinkova-Kostova, A. T., Efferth, T., El Bairi, K., Arkells, N., Fan, T. P., Fiebich, B. L., Freissmuth, M., Georgiev, M. I., Gibbons, S., Godfrey, K. M., Gruber, C. W., Heer, J., Huber, L. A., Ibanez, E., Kijjoa, A., Kiss, A. K., Lu, A., Macias, F. A., Miller, M. J., Mocan, A., Müller, R., Nicoletti, F., Perry, G., Pittalà, V., Rastrelli, L., Ristow, M., Russo, G. L., Silva, A. S., Schuster, D., Sheridan, H., Skalicka-Woźniak, K., Skaltsounis, L., Sobarzo-Sánchez, E., Bredt, D. S., Stuppner, H., Sureda, A., Tzvetkov, N. T., Vacca, R. A., Aggarwal, B. B., Battino, M., Giampieri, F., Wink, M., Wolfender, J. L., Xiao, J., Yeung, A. W. K., Lizard, G., Popp, M. A., Heinrich, M., Berindan-Neagoe, I., Stadler, M., Daglia, M., Verpoorte, R., and Supuran, C. T. (2021). Natural products in drug discovery: advances and opportunities.

Baddeley, A., Bárány, I., and Schneider, R. (2006). Spatial point processes and their applications. *Lect. Notes Math.*, 1892.

Baeuerle, P. A. and Reinhardt, C. (2009). Bispecific T-cell engaging antibodies for cancer therapy.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604).

Bansal, M., Yang, J., Karan, C., Menden, M. P., Costello, J. C., Tang, H., Xiao, G., Li, Y., Allen, J., Zhong, R., Chen, B., Kim, M., Wang, T., Heiser, L. M., Realubit, R., Mattioli, M., Alvarez, M. J., Shen, Y., Gallahan, D., Singer, D., Saez-Rodriguez, J., Xie, Y., Stolovitzky, G., Califano, A., Abbuehl, J. P., Altman, R. B., Balcome, S., Bell, A., Bender, A., Berger, B., Bernard, J., Bieberich, A. A., Borboudakis, G., Chan, C., Chen, T. H., Choi, J., Coelho, L. P., Creighton, C. J., Dampier, W., Davisson, V. J., Deshpande, R., Diao, L., Di Camillo, B., Dundar, M., Ertel, A., Goswami, C. P., Gottlieb, A., Gould, M. N., Goya, J., Grau, M., Gray, J. W., Hejase, H. A., Hoffmann, M. F., Homicsko, K., Homilius, M., Hwang, W., Ijzerman, A. P., Kallioniemi, O., Karacali, B., Kaski, S., Kim, J., Krishnan, A., Lee, J., Lee, Y. S., Lenselink, E. B., Lenz, P., Li, L., Li, J., Liang, H., Mpindi, J. P., Myers, C. L., Newton, M. A., Overington, J. P., Parkkinen, J., Prill, R. J., Peng, J., Pestell, R., Qiu, P., Rajwa, B., Sadanandam, A., Sambo, F., Sridhar, A., Sun, W., Toffolo, G. M., Tozeren, A., Troyanskaya, O. G., Tsamardinos, I., Van Vlijmen, H. W., Wang, W., Wegner, J. K., Wennerberg, K., Van Westen, G. J., Xia, T., Yang, Y., Yao, V., Yuan, Y., Zeng, H., Zhang, S., Zhao, J., and Zhou, J. (2014). A community computational challenge to predict the activity of pairs of compounds. *Nat. Biotechnol.*, 32(12).

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., Wilson, C. J., Lehár, J., Kryukov, G. V., Sonkin, D., Reddy, A., Liu, M., Murray, L., Berger, M. F., Monahan, J. E., Morais, P., Meltzer, J., Korejwa, A., Jané-Valbuena, J., Mapa, F. A., Thibault, J., Bric-Furlong, E., Raman, P., Shipway, A., Engels, I. H., Cheng, J., Yu, G. K., Yu, J., Aspesi, P., De Silva, M., Jagtap, K., Jones, M. D., Wang, L., Hatton, C., Palescandolo, E., Gupta, S., Mahan, S., Sougnez, C., Onofrio, R. C., Liefeld, T., MacConaill, L., Winckler, W., Reich, M., Li, N., Mesirov, J. P., Gabriel, S. B., Getz, G., Ardlie, K., Chan, V., Myer, V. E., Weber, B. L., Porter, J., Warmuth, M., Finan, P., Harris, J. L., Meyerson, M., Golub, T. R., Morrissey, M. P., Sellers, W. R., Schlegel, R., and Garraway, L. A. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, 483(7391).

Baym, M., Stone, L. K., and Kishony, R. (2016). Multidrug evolutionary strategies to reverse antibiotic resistance.

Bedard, P. L., Hansen, A. R., Ratain, M. J., and Siu, L. L. (2013). Tumour heterogeneity in the clinic.

Bendall, L. J., Daniel, A., Kortlepel, K., and Gottlieb, D. J. (1994). Bone marrow adherent layers inhibit apoptosis of acute myeloid leukemia cells. *Exp. Hematol.*

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B*, 57(1).

Bentéjac, C., Csörgő, A., and Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artif. Intell. Rev.*, 54(3).

Beucher, S. (1992). The Watershed Transformation Applied to Image Segmentation. In *Proc. 10th Pfefferkorn Conf. Signal Image Process. Microsc. Microanal.*

Bharati, M. H., Liu, J. J., and MacGregor, J. F. (2004). Image texture analysis: Methods and comparisons. *Chemom. Intell. Lab. Syst.*

Billard, C. (2013). BH3 mimetics: Status of the field and new developments.

Borten, M. A., Bajikar, S. S., Sasaki, N., Clevers, H., and Janes, K. A. (2018). Automated brightfield morphometry of 3D organoid populations by OrganoSeg. *Sci. Rep.*, 8(1).

Boutros, M. and Ahringer, J. (2008). The art and design of genetic screens: RNA interference.

Boutros, M., Heigwer, F., and Laufer, C. (2015). Microscopy-Based High-Content Screening.

Bray, M. A., Singh, S., Han, H., Davis, C. T., Borgeson, B., Hartland, C., Kost-Alimova, M., Gustafsdottir, S. M., Gibson, C. C., and Carpenter, A. E. (2016). Cell Painting, a high-content image-based assay for morphological profiling using multiplexed fluorescent dyes. *Nat. Protoc.*

Breiman, L. (2001). Random forests. *Mach. Learn.*, 45(1).

Breinig, M., Klein, F. A., Huber, W., and Boutros, M. (2015). A chemical–genetic interaction map of small molecules using high-throughput imaging in cancer cells. *Mol. Syst. Biol.*, 11(12).

Brochado, A. R., Telzerow, A., Bobonis, J., Banzhaf, M., Mateus, A., Selkrig, J., Huth, E., Bassler, S., Zamarreño Beas, J., Zietek, M., Ng, N., Foerster, S., Ezraty, B., Py, B., Barras, F., Savitski, M. M., Bork, P., Göttig, S., and Typas, A. (2018). Species-specific activity of antibacterial drug combinations. *Nature*.

Brodin, P. and Christophe, T. (2011). High-content screening in infectious diseases.

Broutier, L., Mastrogiovanni, G., Verstegen, M. M., Francies, H. E., Gavarró, L. M., Bradshaw, C. R., Allen, G. E., Arnes-Benito, R., Sidorova, O., Gaspersz, M. P., Georgakopoulos, N., Koo, B. K., Dietmann, S., Davies, S. E., Praseedom, R. K., Lieshout, R., IJzermans, J. N., Wigmore, S. J., Saeb-Parsy, K., Garnett, M. J., Van Der Laan, L. J., and Huch, M. (2017). Human primary liver cancer-derived organoid cultures for disease modeling and drug screening. *Nat. Med.*, 23(12).

Brown, E. D. and Wright, G. D. (2016). Antibacterial drug discovery in the resistance era.

Bungartz, K. D., Lalowski, K., and Elkin, S. K. (2018). Making the right calls in precision oncology.

Burger, J. A., Burger, M., and Kipps, T. J. (1999). Chronic lymphocytic leukemia B cells express functional CXCR4 chemokine receptors that mediate spontaneous migration beneath bone marrow stromal cells. *Blood*, 94(11).

Burger, J. A. and Kipps, T. J. (2002). Chemokine receptors and stromal cells in the homing and homeostasis of chronic lymphocytic leukemia B cells.

Burger, M., Hartmann, T., Krome, M., Rawluk, J., Tamamura, H., Fujii, N., Kipps, T. J., and Burger, J. A. (2005). Small peptide inhibitors of the CXCR4 chemokine receptor (CD184) antagonize the activation, migration, and antiapoptotic responses of CXCL12 in chronic lymphocytic leukemia B cells. *Blood*, 106(5).

Bush, K., Courvalin, P., Dantas, G., Davies, J., Eisenstein, B., Huovinen, P., Jacoby, G. A., Kishony, R., Kreiswirth, B. N., Kutter, E., Lerner, S. A., Levy, S., Lewis, K., Lomovskaya, O., Miller, J. H., Mobashery, S., Piddock, L. J., Projan, S., Thomas, C. M., Tomasz, A., Tulkens, P. M., Walsh, T. R., Watson, J. D., Witkowski, J., Witte, W., Wright, G., Yeh, P., and Zgurskaya, H. I. (2011). Tackling antibiotic resistance.

Cacace, E. (2021). High-throughput profiling of drug interactions in Gram-positive bacteria. *PhD thesis, University of Groningen*.

Cacace, E., Kim, V., Huber, W., and Typas, A. (2021). High-throughput profiling of drug interactions in Gram-positive bacteria. *Manuscript in preparation*.

Cacace, E., Kritikos, G., and Typas, A. (2017). Chemical genetics in drug discovery.

Caicedo, J. C., Cooper, S., Heigwer, F., Warchal, S., Qiu, P., Molnar, C., Vasilevich, A. S., Barry, J. D., Bansal, H. S., Kraus, O., Wawer, M., Paavolainen, L., Herrmann, M. D., Rohban, M., Hung, J., Hennig, H., Concannon, J., Smith, I., Clemons, P. A., Singh, S., Rees, P., Horvath, P., Linington, R. G., and Carpenter, A. E. (2017). Data-analysis strategies for image-based cell profiling. *Nat. Methods*.

Chait, R., Craney, A., and Kishony, R. (2007). Antibiotic interactions that select against resistance. *Nature*, 446(7136).

Chandrasekaran, S., Cokol-Cakmak, M., Sahin, N., Yilancioglu, K., Kazan, H., Collins, J. J., and Cokol, M. (2016). Chemogenomics and orthology-based design of antibiotic combination therapies. *Mol. Syst. Biol.*

Chandrasekaran, S. N., Ceulemans, H., Boyd, J. D., and Carpenter, A. E. (2021). Image-based profiling for drug discovery: due for a machine-learning upgrade?

Chen, H., Dou, Q., Yu, L., Qin, J., and Heng, P. A. (2018). VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images.

Chen, S. S., Chang, B. Y., Chang, S., Tong, T., Ham, S., Sherry, B., Burger, J. A., Rai, K. R., and Chiorazzi, N. (2016). BTK inhibition results in impaired CXCR4 chemokine receptor surface expression, signaling and function in chronic lymphocytic leukemia. *Leukemia*, 30(4).

Chen, T. and Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, volume 13-17-August-2016.

Chevereau, G. and Bollenbach, T. (2015). Systematic discovery of drug interaction mechanisms. *Mol. Syst. Biol.*

Chua, V., Orloff, M., Teh, J. L., Sugase, T., Liao, C., Purwin, T. J., Lam, B. Q., Terai, M., Ambrosini, G., Carvajal, R. D., Schwartz, G., Sato, T., and Aplin, A. E. (2019). Stromal fibroblast growth factor 2 reduces the efficacy of bromodomain inhibitors in uveal melanoma. *EMBO Mol. Med.*

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). 3D U-net: Learning dense volumetric segmentation from sparse annotation. In *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, volume 9901 LNCS.

Coelho, L. P. (2013). Mahotas: Open source software for scriptable computer vision. *J. Open Res. Softw.*

Cokol, M., Chua, H. N., Tasan, M., Mutlu, B., Weinstein, Z. B., Suzuki, Y., Nergiz, M. E., Costanzo, M., Baryshnikova, A., Giaever, G., Nislow, C., Myers, C. L., Andrews, B. J., Boone, C., and Roth, F. P. (2011). Systematic exploration of synergistic drug pairs. *Mol. Syst. Biol.*

Coscia, M., Pantaleoni, F., Riganti, C., Vitale, C., Rigoni, M., Peola, S., Castella, B., Foglietta, M., Griggio, V., Drandi, D., Ladetto, M., Bosia, A., Boccadoro, M., and Massaia, M. (2011). IGHV unmutated CLL B cells are more prone to spontaneous apoptosis and subject to environmental prosurvival signals than mutated CLL B cells. *Leukemia*, 25(5).

Costanzo, M., Baryshnikova, A., Bellay, J., Kim, Y., Spear, E. D., Sevier, C. S., Ding, H., Koh, J. L., Toufighi, K., Mostafavi, S., Prinz, J., St. Onge, R. P., Vandersluis, B., Makhnevych, T., Vizeacoumar, F. J., Alizadeh, S., Bahr, S., Brost, R. L., Chen, Y., Cokol, M., Deshpande, R., Li, Z., Lin, Z. Y., Liang, W., Marback, M., Paw, J., Luis, B. J. S., Shuteriqi, E., Tong, A. H. Y., Van Dyk, N., Wallace, I. M., Whitney, J. A., Weirauch, M. T., Zhong, G., Zhu, H., Houry, W. A., Brudno, M., Ragibizadeh, S., Papp, B., Pál, C., Roth, F. P., Giaever, G., Nislow, C., Troyanskaya, O. G., Bussey, H., Bader, G. D., Gingras, A. C., Morris, Q. D., Kim, P. M., Kaiser, C. A., Myers, C. L., Andrews, B. J., and Boone, C. (2010). The genetic landscape of a cell. *Science (80-.).*, 327(5964).

Curran, E. and Stock, W. (2019). Taking a “bite out of all”: Blinatumomab approval for MRD-positive ALL. *Blood*, 133(16).

Czerniecki, S. M., Cruz, N. M., Harder, J. L., Menon, R., Annis, J., Otto, E. A., Gulieva, R. E., Islas, L. V., Kim, Y. K., Tran, L. M., Martins, T. J., Pippin, J. W., Fu, H., Kretzler, M., Shankland, S. J., Himmelfarb, J., Moon, R. T., Paragas, N., and Freedman, B. S. (2018). High-Throughput Screening Enhances Kidney Organoid Differentiation from Human Pluripotent Stem Cells and Enables Automated Multidimensional Phenotyping. *Cell Stem Cell*, 22(6).

Damper, P. D. and Epstein, W. (1981). Role of the membrane potential in bacterial resistance to aminoglycoside antibiotics. *Antimicrob. Agents Chemother.*, 20(6).

Delbridge, A. R. and Strasser, A. (2015). The BCL-2 protein family, BH3-mimetics and cancer therapy.

Dent, P., Curiel, D. T., Fisher, P. B., and Grant, S. (2009). Synergistic combinations of signaling pathway inhibitors: Mechanisms for improved cancer therapy. *Drug Resist. Updat.*, 12(3).

Dietlein, F., Kalb, B., Jokic, M., Noll, E. M., Strong, A., Tharun, L., Ozretić, L., Künstlinger, H., Kambartel, K., Randerath, W. J., Jüngst, C., Schmitt, A., Torgovnick, A., Richters, A., Rauh, D., Siedek, F., Persigehl, T., Mauch, C., Bartkova, J., Bradley, A., Sprick, M. R., Trumpp, A., Rad, R., Saur, D., Bartek, J., Wolf, J., Büttner, R., Thomas, R. K., and Reinhardt, H. C. (2015). A Synergistic Interaction between Chk1- and MK2 Inhibitors in KRAS-Mutant Cancer. *Cell*, 162(1).

Dietrich, S., Oleś, M., Lu, J., Sellner, L., Anders, S., Velten, B., Wu, B., Hüllein, J., Liberio, M., Walther, T., Wagner, L., Rabe, S., Ghidelli-Disse, S., Bantscheff, M., Oleś, A., Słabicki, M., Mock, A., Oakes, C., Wang, S., Oppermann, S., Lukas, M., Kim, V., Sill, M., Benner, A., Jauch, A., Sutton, L., Young, E., Rosenquist, R., Liu, X., Jethwa, A., Lee, K., Lewis, J., Putzker, K., Lutz, C., Rossi, D., Mokhir, A., Oellerich, T., Zirlik, K., Herling, M., Nguyen-Khac, F., Plass, C., Andersson, E., Mustjoki, S., Von Kalle, C., Ho, A., Hensel, M., Dürig, J., Ringshausen, I., Zapatka, M., Huber, W., and Zenz, T. (2018). Drug-perturbation-based stratification of blood cancer. *J. Clin. Invest.*, 128(1).

Ding, L., Zhang, W., Yang, L., Pelicano, H., Zhou, K., Yin, R., Huang, R., and Zeng, J. (2018). Targeting the autophagy in bone marrow stromal cells overcomes resistance to vorinostat in chronic lymphocytic leukemia. *Onco. Targets. Ther.*, 11.

Ding, W., Nowakowski, G. S., Knox, T. R., Boysen, J. C., Maas, M. L., Schwager, S. M., Wu, W., Wellik, L. E., Dietz, A. B., Ghosh, A. K., Secreto, C. R., Medina, K. L., Shanafelt, T. D., Zent, C. S., Call, T. G., and Kay, N. E. (2009). Bi-directional activation between mesenchymal stem cells and CLL B-cells: Implication for CLL disease progression. *Br. J. Haematol.*, 147(4).

Döhner, H. and Gaidzik, V. I. (2011). Impact of genetic features on treatment decisions in AML.

Dolz, J., Gopinath, K., Yuan, J., Lombaert, H., Desrosiers, C., and Ben Ayed, I. (2019). HyperDense-Net: A Hyper-Densely Connected CNN for Multi-Modal Image Segmentation. *IEEE Trans. Med. Imaging*, 38(5).

Donnenberg, V. S., Corselli, M., Normolle, D. P., Meyer, E. M., and Donnenberg, A. D. (2018). Flow cytometric detection of most proteins in the cell surface proteome is unaffected by trypsin treatment. *Cytom. Part A*, 93(8).

Doroshow, D. B., Eder, J. P., and LoRusso, P. M. (2017). BET inhibitors: A novel epigenetic approach.

Duarte, D., Hawkins, E. D., and Lo Celso, C. (2018). The interplay of leukemia cells and the bone marrow microenvironment.

Dubois, N., Crompot, E., Meuleman, N., Bron, D., Lagneaux, L., and Stamatopoulos, B. (2020). Importance of Crosstalk Between Chronic Lymphocytic Leukemia Cells and the Stromal Microenvironment: Direct Contact, Soluble Factors, and Extracellular Vesicles.

Dumas, P. Y., Naudin, C., Martin-Lannerée, S., Izac, B., Casetti, L., Mansier, O., Rousseau, B., Artus, A., Dufossée, M., Giese, A., Dubus, P., Pigneux, A., Praloran, V., Bidet, A., Villacreces, A., Guitart, A., Milpied, N., Kosmider, O., Vigon, I., Desplat, V., Dusanter-Fourt, I., and Pasquet, J. M. (2019). Hematopoietic niche drives FLT3-ITD acute myeloid leukemia resistance to quizartinib via STAT5- And hypoxia-dependent upregulation of AXL. *Haematologica*, 104(10).

Eriksson, A., Österroos, A., Hassan, S., Gullbo, J., Rickardson, L., Jarvius, M., Nygren, P., Fryknäs, M., Höglund, M., and Larsson, R. (2015). Drug screen in patient cells suggests quinacrine to be repositioned for treatment of acute myeloid leukemia. *Blood Cancer J.*

Evenson, R. (1993). Patents, R&D, and invention potential: International evidence. *Am. Econ. Rev.*, 83(2).

Fischer, B., Sandmann, T., Horn, T., Billmann, M., Chaudhary, V., Huber, W., and Boutros, M. (2015). A map of directional genetic interactions in a metazoan cell. *Elife*.

Fittall, M. W. and Van Loo, P. (2019). Translating insights into tumor evolution to clinical practice: Promises and challenges.

Fong, C. Y., Gilan, O., Lam, E. Y., Rubin, A. F., Ftouni, S., Tyler, D., Stanley, K., Sinha, D., Yeh, P., Morison, J., Giotopoulos, G., Lugo, D., Jeffrey, P., Lee, S. C. W., Carpenter, C., Gregory, R., Ramsay, R. G., Lane, S. W., Abdel-Wahab, O., Kouzarides, T., Johnstone, R. W., Dawson, S. J., Huntly, B. J., Prinjha, R. K., Papenfuss, A. T., and Dawson, M. A. (2015). BET inhibitor resistance emerges from leukaemia stem cells. *Nature*.

Freund, Y. and Schapire, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *J. Comput. Syst. Sci.*, 55(1).

Friedman, A. A., Letai, A., Fisher, D. E., and Flaherty, K. T. (2015). Precision medicine for cancer with next-generation functional diagnostics.

Frieri, M., Kumar, K., and Boutin, A. (2017). Antibiotic resistance. *Journal of Infection and Public Health*, 10(4):369–378.

Fuchs, F., Pau, G., Kranz, D., Sklyar, O., Budjan, C., Steinbrink, S., Horn, T., Pedal, A., Huber, W., and Boutros, M. (2010). Clustering phenotype populations by genome-wide RNAi and multiparametric imaging. *Mol. Syst. Biol.*

Garnett, M. J., Edelman, E. J., Heidorn, S. J., Greenman, C. D., Dastur, A., Lau, K. W., Greninger, P., Thompson, I. R., Luo, X., Soares, J., Liu, Q., Iorio, F., Surdez, D., Chen, L., Milano, R. J., Bignell, G. R., Tam, A. T., Davies, H., Stevenson, J. A., Barthorpe, S., Lutz, S. R., Kogera, F., Lawrence, K., McLaren-Douglas, A., Mitropoulos, X., Mironenko, T., Thi, H., Richardson, L., Zhou, W., Jewitt, F., Zhang, T., O'Brien, P., Boisvert, J. L., Price, S., Hur, W., Yang, W., Deng, X., Butler, A., Choi, H. G., Chang, J. W., Baselga, J., Stamenkovic, I., Engelman, J. A., Sharma, S. V., Delattre, O., Saez-Rodriguez, J., Gray, N. S., Settleman, J., Futreal, P. A., Haber, D. A., Stratton, M. R., Ramaswamy, S., McDermott, U., and Benes, C. H. (2012). Systematic identification of genomic markers of drug sensitivity in cancer cells. *Nature*, 483(7391).

Ghatak, S., King, Z. A., Sastry, A., and Palsson, B. O. (2019). The y-ome defines the 35% of Escherichia coli genes that lack experimental evidence of function. *Nucleic Acids Res.*

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K. D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Güldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kötter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-Macdonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C. Y., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., and Johnston, M. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, 418(6896).

Godinez, W. J., Hossain, I., Lazic, S. E., Davies, J. W., and Zhang, X. (2017). A multi-scale convolutional neural network for phenotyping high-content cellular images. *Bioinformatics*.

Goebeler, M. E., Knop, S., Viardot, A., Kufer, P., Topp, M. S., Einsele, H., Noppeney, R., Hess, G., Kallert, S., Mackensen, A., Rupertus, K., Kanz, L., Libicher, M., Nagorsen, D., Zugmaier, G., Klinger, M., Wolf, A., Dorsch, B., Quednau, B. D., Schmidt, M., Scheele, J., Baeuerle, P. A., Leo, E., and Bargou, R. C. (2016). Bispecific T-cell engager (BiTE) antibody construct Blinatumomab for the treatment of Patients with relapsed/refractory non-Hodgkin lymphoma: Final results from a phase I study. *J. Clin. Oncol.*, 34(10).

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.

Greco, W. R., Bravo, G., and Parsons, J. C. (1995). The search for synergy: A critical review from a response surface perspective.

Gupta, S. V., Hertlein, E., Lu, Y., Sass, E. J., Lapalombella, R., Chen, T. L., Davis, M. E., Woyach, J. A., Lehman, A., Jarjoura, D., Byrd, J. C., and Lucas, D. M. (2013). The proteasome inhibitor carfilzomib functions independently of p53 to induce cytotoxicity and an atypical NF-κB response in chronic lymphocytic leukemia cells. *Clin. Cancer Res.*

Han, X., Dorsey-Oresto, A., Malik, M., Wang, J. Y., Drlica, K., Zhao, X., and Lu, T. (2010). Escherichia coli genes that reduce the lethal effects of stress. *BMC Microbiol.*, 10.

Haralick, R. M., Dinstein, I., and Shanmugam, K. (1973). Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.*

Harvey, A. L. (2008). Natural products in drug discovery.

Hasserjian, R. P. (2013). Acute myeloid leukemia: Advances in diagnosis and classification.

He, H. and Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.*

He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2020). Mask R-CNN. *IEEE Trans. Pattern Anal. Mach. Intell.*

Herbst, S. (2020). Systematic analysis of cell-intrinsic and extrinsic factors in chronic lymphocytic leukemia to understand functional consequences for drug response and clinical outcome. *PhD thesis, University of Heidelberg*.

Herbst, S. and Kim, V. (2021). Imaging-based stroma coculture model for drug response assays in hematological cancers. *Manuscript in preparation*.

Hernando-Amado, S., Sanz-García, F., and Martínez, J. L. (2020). Rapid and robust evolution of collateral sensitivity in Pseudomonas aeruginosa antibiotic-resistant mutants. *Sci. Adv.*, 6(32).

Holick, M. F. (2004). Sunlight and vitamin D for bone health and prevention of autoimmune diseases, cancers, and cardiovascular disease.

Hsieh, F. H. (2014). Primer to the Immune Response. *Ann. Allergy, Asthma Immunol.*, 113(3):333.

Hu, M. K. (1962). Visual Pattern Recognition by Moment Invariants. *IRE Trans. Inf. Theory*.

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S., and Murphy, K. (2017). Speed/accuracy trade-offs for modern convolutional object detectors. In *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*.

Huang, Z. and Leng, J. (2010). Analysis of Hu's moment invariants on image scaling and rotation. In *ICCET 2010 - 2010 Int. Conf. Comput. Eng. Technol. Proc.*

Huber, W., Von Heydebreck, A., Sültmann, H., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. In *Bioinformatics*.

Imamovic, L. and Sommer, M. O. (2013). Use of collateral sensitivity networks to design drug cycling protocols that avoid resistance development. *Sci. Transl. Med.*, 5(204).

Iorio, F., Knijnenburg, T. A., Vis, D. J., Bignell, G. R., Menden, M. P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., Cokelaer, T., Greninger, P., van Dyk, E., Chang, H., de Silva, H., Heyn, H., Deng, X., Egan, R. K., Liu, Q., Mironenko, T., Mitropoulos, X., Richardson, L., Wang, J., Zhang, T., Moran, S., Sayols, S., Soleimani, M., Tamborero, D., Lopez-Bigas, N., Ross-Macdonald, P., Esteller, M., Gray, N. S., Haber, D. A., Stratton, M. R., Benes, C. H., Wessels, L. F., Saez-Rodriguez, J., McDermott, U., and Garnett, M. J. (2016). A Landscape of Pharmacogenomic Interactions in Cancer. *Cell*, 166(3).

Janssens, R., Zhang, X., Kauffmann, A., de Weck, A., and Durand, E. Y. (2021). Fully unsupervised deep mode of action learning for phenotyping high-content cellular images. *Bioinformatics*.

Jones, G. M., Stalker, J., Humphray, S., West, A., Cox, T., Rogers, J., Dunham, I., and Prelich, G. (2008). A systematic library for comprehensive overexpression screens in Saccharomyces cerevisiae. *Nat. Methods*, 5(3).

Judd, S. E. and Tangpricha, V. (2009). Vitamin D deficiency and risk for cardiovascular disease. In *Am. J. Med. Sci.*, volume 338.

Kamdje, A. H., Bassi, G., Pacelli, L., Malpeli, G., Amati, E., Nichele, I., Pizzolo, G., and Krampera, M. (2012). Role of stromal cell-mediated Notch signaling in CLL resistance to chemotherapy. *Blood Cancer J.*, 2(5).

Kämpfer, A. A., Urbán, P., Gioria, S., Kanase, N., Stone, V., and Kinsner-Ovaskainen, A. (2017). Development of an in vitro co-culture model to mimic the human intestine in healthy and diseased state. *Toxicol. Vitr.*, 45.

Karjalainen, R., Pemovska, T., Popa, M., Liu, M., Javarappa, K. K., Majumder, M. M., Yadav, B., Tamborero, D., Tang, J., Bychkov, D., Kontro, M., Parsons, A., Suvela, M., Mayoral Safont, M., Porkka, K., Aittokallio, T., Kallioniemi, O., McCormack, E., Gjertsen, B. T., Wennerberg, K., Knowles, J., and Heckman, C. A. (2017). JAK1/2 and BCL2 inhibitors synergize to counteract bone marrow stromal cell–induced protection of AML. *Blood*.

Kay, N. E., Shanafelt, T. D., Strege, A. K., Lee, Y. K., Bone, N. D., and Raza, A. (2007). Bone biopsy derived marrow stromal elements rescue chronic lymphocytic leukemia B-cells from spontaneous and drug induced cell death and facilitates an "angiogenic switch". *Leuk. Res.*, 31(7).

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., and Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In *Adv. Neural Inf. Process. Syst.*, volume 2017-December.

Keating, G. M. (2010). Rituximab: A review of its use in chronic lymphocytic leukaemia, low-grade or follicular lymphoma and diffuse large b-cell lymphoma.

Kensert, A., Harrison, P. J., and Spjuth, O. (2019). Transfer Learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes. *SLAS Discov.*, 24(4).

Khotanzad, A. and Hong, Y. H. (1990). Invariant Image Recognition by Zernike Moments. *IEEE Trans. Pattern Anal. Mach. Intell.*

Kirkpatrick, C. J., Fuchs, S., and Unger, R. E. (2011). Co-culture systems for vascularization - Learning from nature.

Kirtonia, A., Pandya, G., Sethi, G., Pandey, A. K., Das, B. C., and Garg, M. (2020). A comprehensive review of genetic alterations and molecular targeted therapies for the implementation of personalized medicine in acute myeloid leukemia. 98(8).

Koenderink, J. J. and van Doorn, A. J. (1992). Surface shape and curvature scales. *Image Vis. Comput.*

Kong, H., Akakin, H. C., and Sarma, S. E. (2013). A generalized laplacian of gaussian filter for blob detection and its applications. *IEEE Trans. Cybern.*

Konopleva, M., Konoplev, S., Hu, W., Zaritskey, A. Y., Afanasiev, B. V., and Andreeff, M. (2002). Stromal cells prevent apoptosis of AML cells by up-regulation of anti-apoptotic proteins. *Leukemia*, 16(9).

Kraus, O. Z., Grys, B. T., Ba, J., Chong, Y., Frey, B. J., Boone, C., and Andrews, B. J. (2017). Automated analysis of high-content microscopy data with deep learning. *Mol. Syst. Biol.*, 13(4).

Kumar-Sinha, C. and Chinnaiyan, A. M. (2018). Precision oncology in the age of integrative genomics.

Kurtova, A. V., Balakrishnan, K., Chen, R., Ding, W., Schnabl, S., Quiroga, M. P., Sivina, M., Wierda, W. G., Estrov, Z., Keating, M. J., Shehata, M., Jäger, U., Gandhi, V., Kay, N. E., Plunkett, W., and Burger, J. A. (2009). Diverse marrow stromal cells protect CLL cells from spontaneous and drug-induced apoptosis: Development of a reliable and reproducible system to assess stromal cell adhesion-mediated drug resistance. *Blood.*

Lagneaux, L., Delforge, A., Bron, D., De Bruyn, C., and Stryckmans, P. (1998). Chronic lymphocytic leukemic B cells but not normal B cells are rescued from apoptosis by contact with normal bone marrow stromal cells. *Blood*, 91(7).

Landau, D. A., Carter, S. L., Stojanov, P., McKenna, A., Stevenson, K., Lawrence, M. S., Sougnez, C., Stewart, C., Sivachenko, A., Wang, L., Wan, Y., Zhang, W., Shukla, S. A., Vartanov, A., Fernandes, S. M., Saksena, G., Cibulskis, K., Tesar, B., Gabriel, S., Hacohen, N., Meyerson, M., Lander, E. S., Neuberg, D., Brown, J. R., Getz, G., and Wu, C. J. (2013). Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*, 152(4).

Langton, K. P., Henderson, P. J., and Herbert, R. B. (2005). Antibiotic resistance: Multidrug efflux proteins, a common transport mechanism?

Laufer, C., Fischer, B., Billmann, M., Huber, W., and Boutros, M. (2013). Mapping genetic interactions in human cancer cells with RNAi and multiparametric phenotyping. *Nat. Methods*.

Lausen, B. and Schumacher, M. (1992). Maximally Selected Rank Statistics. *Biometrics*, 48(1).

Lecun, J., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*.

Lehár, J., Stockwell, B. R., Giaever, G., and Nislow, C. (2008). Combination chemical genetics.

Letai, A. (2017). Functional precision cancer medicine-moving beyond pure genomics.

Levy, S. B. (2002). Active efflux, a common mechanism for biocide and antibiotic resistance.

Liberali, P., Snijder, B., and Pelkmans, L. (2014). A hierarchical map of regulatory genetic interactions in membrane trafficking. *Cell*, 157(6).

Liberali, P., Snijder, B., and Pelkmans, L. (2015). Single-cell and multivariate approaches in genetic perturbation screens.

Liebers, N., Bruch, P.-M., Gambietz, A., Giles, H., Lu, J., Knoll, M., Kolb, C., Dreger, P., Mueller-Tidow, C., Huber, W., Benner, A., Zenz, T., and Dietrich, S. (2019). Ex-Vivo Drug Response Profiling for Tailoring Treatment in Hematologic Malignancies: The Prospective Non-Interventional SMART-Trial. *Blood*, 134(Supplement_1).

Loo, L. H., Wu, L. F., and Altschuler, S. J. (2007). Image-based multivariate profiling of drug responses from single cells. *Nat. Methods*, 4(5).

Lukonin, I., Serra, D., Challet Meylan, L., Volkmann, K., Baaten, J., Zhao, R., Meeusen, S., Colman, K., Maurer, F., Stadler, M. B., Jenkins, J., and Liberali, P. (2020). Phenotypic landscape of intestinal organoid regeneration. *Nature*, 586(7828).

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., and Lee, S.-I. (2020). From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.*, 2(1).

Lundberg, S. M. and Lee, S. I. (2017). A unified approach to interpreting model predictions. In *Adv. Neural Inf. Process. Syst.*, volume 2017-December.

MacDonald, M. L., Lamerdin, J., Owens, S., Keon, B. H., Bilter, G. K., Shang, Z., Huang, Z., Yu, H., Dias, J., Minami, T., Michnick, S. W., and Westwick, J. K. (2006). Identifying off-target effects and hidden phenotypes of drugs in human cells. *Nat. Chem. Biol.*, 2(6).

Malani, D., Murumägi, A., Yadav, B., Pemovska, T., Mpindi, J. P., Kontro, M., Aittokallio, T., Heckman, C. A., Porkka, K., Wennerberg, K., Wolf, M., and Kallioniemi, O. (2014). AML Specific Targeted Drugs Identified By Drug Sensitivity and Resistance Testing: Comparison of Ex Vivo Patient Cells with in Vitro Cell Lines. *Blood*, 124(21).

Markham, P. N. and Neyfakh, A. A. (2001). Efflux-mediated drug resistance in Gram-positive bacteria.

Martin, R. G. and Rosner, J. L. (1995). Binding of purified multiple antibiotic-resistance repressor protein (MarR) to mar operator sequences. *Proc. Natl. Acad. Sci. U. S. A.*, 92(12).

Mason, D. J., Stott, I., Ashenden, S., Weinstein, Z. B., Karakoc, I., Meral, S., Kuru, N., Bender, A., and Cokol, M. (2017). Prediction of Antibiotic Interactions Using Descriptors Derived from Molecular Structure. *J. Med. Chem.*

Mattiazzi Usaj, M., Styles, E. B., Verster, A. J., Friesen, H., Boone, C., and Andrews, B. J. (2016). High-Content Screening for Quantitative Cell Biology.

Mauvezin, C. and Neufeld, T. P. (2015). Bafilomycin A1 disrupts autophagic flux by inhibiting both V-ATPase-dependent acidification and Ca-P60A/SERCA-dependent autophagosome-lysosome fusion. *Autophagy*, 11(8).

McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B. A., Karhohs, K. W., Doan, M., Ding, L., Rafelski, S. M., Thirstrup, D., Wiegraebe, W., Singh, S., Becker, T., Caicedo, J. C., and Carpenter, A. E. (2018). CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol.*

Meijering, E. (2020). A bird's-eye view of deep learning in bioimage analysis.

Mellman, I., Coukos, G., and Dranoff, G. (2011). Cancer immunotherapy comes of age.

Menden, M. P., Wang, D., Mason, M. J., Szalai, B., Bulusu, K. C., Guan, Y., Yu, T., Kang, J., Jeon, M., Wolfinger, R., Nguyen, T., Zaslavskiy, M., Abante, J., Y.Di, G., Fawell, S., Stolovitzky, G., Guinney, J., Dry, J. R., and Saez-Rodriguez, J. (2019). Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. *Nat. Commun.*, 10(1).

Meyer, M., Reimand, J., Lan, X., Head, R., Zhu, X., Kushida, M., Bayani, J., Pressey, J. C., Lionel, A. C., Clarke, I. D., Cusimano, M., Squire, J. A., Scherer, S. W., Bernstein, M., Woodin, M. A., Bader, G. D., and Dirks, P. B. (2015). Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity. *Proc. Natl. Acad. Sci. U. S. A.*, 112(3).

Miller, M. H., Feinstein, S. A., and Chow, R. T. (1987). Early effects of β-lactams on aminoglycoside uptake, bactericidal rates, and turbidimetrically measured growth inhibition in Pseudomonas aeruginosa. *Antimicrob. Agents Chemother.*, 31(1).

Moellering, R. C., Eliopoulos, G. M., and Allan, J. D. (1986). Beta-lactam/aminoglycoside combinations: Interactions and their mechanisms. *Am. J. Med.*, 80(5 C).

Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M., and Van Valen, D. (2019). Deep learning for cellular image analysis.

Moffat, J., Grueneberg, D. A., Yang, X., Kim, S. Y., Kloepfer, A. M., Hinkle, G., Piqani, B., Eisenhaure, T. M., Luo, B., Grenier, J. K., Carpenter, A. E., Foo, S. Y., Stewart, S. A., Stockwell, B. R., Hacohen, N., Hahn, W. C., Lander, E. S., Sabatini, D. M., and Root, D. E. (2006). A Lentiviral RNAi Library for Human and Mouse Genes Applied to an Arrayed Viral High-Content Screen. *Cell*, 124(6).

Mokhtari, R. B., Homayouni, T. S., Baluch, N., Morgatskaya, E., Kumar, S., Das, B., and Yeger, H. (2017). Combination therapy in combating cancer.

Molnar, C. (2019). *Interpretable Machine Learning*. https://christophm.github.io/interpretable-ml-book/.

Montresor, A., Toffali, L., Rigo, A., Ferrarini, I., Vinante, F., and Laudanna, C. (2018). CXCR4-and BCR-triggered integrin activation in B-cell chronic lymphocytic leukemia cells depends on JAK2-activated Bruton's tyrosine kinase. *Oncotarget*, 9(80).

Montserrat, E., Bauman, T., and Delgado, J. (2016). Present and future of personalized medicine in CLL.

Moreno, C. and Montserrat, E. (2008). New prognostic markers in chronic lymphocytic leukemia. *Blood Rev.*, 22(4).

Morid, M. A., Borjali, A., and Del Fiol, G. (2021). A scoping review of transfer learning research on medical image analysis using ImageNet.

Mott, B. T., Eastman, R. T., Guha, R., Sherlach, K. S., Siriwardana, A., Shinn, P., McKnight, C., Michael, S., Lacerda-Queiroz, N., Patel, P. R., Khine, P., Sun, H., Kasbekar, M., Aghdam, N., Fontaine, S. D., Liu, D., Mierzwa, T., Mathews-Griner, L. a., Ferrer, M., Renslo, A. R., Inglese, J., Yuan, J., Roepe, P. D., Su, X.-z., and Thomas, C. J. (2015). High-throughput matrix screening identifies synergistic and antagonistic antimalarial drug combinations. *Sci. Rep.*, 5(13891):doi: 10.1038/srep13891.

Mudry, M. E., Fortney, J. E., York, T., Hall, B. M., and Gibson, L. F. (2000). Stromal cells regulate survival of B-lineage leukemic cells during chemotherapy. *Blood*.

Natekin, A. and Knoll, A. (2013). Gradient boosting machines, a tutorial. *Front. Neurorobot.*, 7(DEC).

Nathan, C. and Goldberg, F. M. (2005). Outlook: The profit problem in antibiotic R&D. *Nat. Rev. Drug Discov.*, 4(11).

Neu, H. C. (1992). The crisis in antibiotic resistance. *Science (80-.).*

Nichols, D. P., Happoldt, C. L., Bratcher, P. E., Caceres, S. M., Chmiel, J. F., Malcolm, K. C., Saavedra, M. T., Saiman, L., Taylor-Cousar, J. L., and Nick, J. A. (2017). Impact of azithromycin on the clinical and antimicrobial effectiveness of tobramycin in the treatment of cystic fibrosis. *J. Cyst. Fibros.*, 16(3).

Nichols, R. J., Sen, S., Choo, Y. J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K. M., Lee, K. J., Wong, A., Shales, M., Lovett, S., Winkler, M. E., Krogan, N. J., Typas, A., and Gross, C. A. (2011). Phenotypic landscape of a bacterial cell. *Cell*.

Noguera, N. I., Breccia, M., Divona, M., Diverio, D., Costa, V., De Santis, S., Avvisati, G., Pinazzi, M. B., Petti, M. C., Mandelli, F., and Lo Coco, F. (2002). Alterations of the FLT3 gene in acute promyelocytic leukemia: Association with diagnostic characteristics and analysis of clinical outcome in patients treated with the Italian AIDA protocol. *Leukemia*, 16(11).

Norberg, M., Lindhagen, E., Kanduri, M., Rickardson, L., Sundström, C., Stamatopoulos, K., Rosenquist, R., and Åleskog, A. (2012). Screening for cytotoxic compounds in poor-prognostic chronic lymphocytic leukemia. *Anticancer Res.*, 32(8).

Otsu, N. (1979). THRESHOLD SELECTION METHOD FROM GRAY-LEVEL HISTOGRAMS. *IEEE Trans Syst Man Cybern.*

Pammolli, F., Magazzini, L., and Riccaboni, M. (2011). The productivity crisis in pharmaceutical R&D. *Nat. Rev. Drug Discov.*, 10(6).

Panayiotidis, P., Jones, D., Ganeshaguru, K., Foroni, L., and Hoffbrand, A. V. (1996). Human bone marrow stromal cells prevent apoptosis and support the survival of chronic lymphocytic leukaemia cells in vitro. *Br. J. Haematol.*

Paulsson, K., Forestier, E., Lilljebjörn, H., Heldrup, J., Behrendtz, M., Young, B. D., and Johansson, B. (2010). Genetic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Proc. Natl. Acad. Sci. U. S. A.*, 107(50).

Pemovska, T., Kontro, M., Yadav, B., Edgren, H., Eldfors, S., Szwajda, A., Almusa, H., Bespalov, M. M., Ellonen, P., Elonen, E., Gjertsen, B. T., Karjalainen, R., Kulesskiy, E., Lagström, S., Lehto, A., Lepistö, M., Lundán, T., Majumder, M. M., Marti, J. M., Mattila, P., Murumägi, A., Mustjoki, S., Palva, A., Parsons, A., Pirttinen, T., Rämet, M. E., Suvela, M., Turunen, L., Västrik, I., Wolf, M., Knowles, J., Aittokallio, T., Heckman, C. A., Porkka, K., Kallioniemi, O., and Wennerberg, K. (2013). Individualized systems medicine strategy to tailor treatments for patients with chemorefractory acute myeloid leukemia. *Cancer Discov.*, 3(12).

Perlman, Z. E., Slack, M. D., Feng, Y., Mitchison, T. J., Wu, L. F., and Altschuler, S. J. (2004). Multidimensional drug profiling by automated microscopy. *Science (80-.).*, 306(5699).

Peters, J. M., Colavin, A., Shi, H., Czarny, T. L., Larson, M. H., Wong, S., Hawkins, J. S., Lu, C. H., Koo, B. M., Marta, E., Shiver, A. L., Whitehead, E. H., Weissman, J. S., Brown, E. D., Qi, L. S., Huang, K. C., and Gross, C. A. (2016). A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. *Cell*, 165(6).

Petrou, M. and García Sevilla, P. (2006). *Image Processing: Dealing with Texture*.

Pfalz, B. (2017). Comparing bacterial gene networks based on high-throughput phenomics. *PhD thesis, University of Heidelberg*.

Piddock, L. J. (2012). The crisis of no new antibiotics-what is the way forward?

Pinheiro, P. O., Collobert, R., and Dollar, P. (2015). Learning to segment object candidates. In *Adv. Neural Inf. Process. Syst.*

Pinheiro, P. O., Lin, T. Y., Collobert, R., and Dollár, P. (2016). Learning to refine object segments. In *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics).*

Plackett, B. (2020). Why big pharma has abandoned antibiotics. *Nature*, 586(7830).

Plum, L. A. and Deluca, H. F. (2010). Vitamin D, disease and therapeutic opportunities.

Probst, P., Wright, M. N., and Boulesteix, A. L. (2019). Hyperparameters and tuning strategies for random forest.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A. (2018). Catboost: Unbiased boosting with categorical features. In *Adv. Neural Inf. Process. Syst.*, volume 2018-December.

Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., Norris, A., Sanseau, P., Cavalla, D., and Pirmohamed, M. (2018). Drug repurposing: Progress, challenges and recommendations.

Quintarelli, C., De Angelis, B., Errichiello, S., Caruso, S., Esposito, N., Colavita, I., Raia, M., Pagliuca, S., Pugliese, N., Risitano, A. M., Picardi, M., Luciano, L., Saglio, G., Martinelli, G., and Pane, F. (2014). Selective strong synergism of Ruxolitinib and second generation tyrosine kinase inhibitors to overcome bone marrow stroma related drug resistance in chronic myelogenous leukemia. *Leuk. Res.*

Ramón-García, S., Ng, C., Anderson, H., Chao, J. D., Zheng, X., Pfeifer, T., Av-Gay, Y., Roberge, M., and Thompson, C. J. (2011). Synergistic drug combinations for tuberculosis therapy identified by a novel high-throughput screen. *Antimicrob. Agents Chemother.*, 55(8).

Rao, R. D., Mladek, A. C., Lamont, J. D., Goble, J. M., Erlichman, C., James, C. D., and Sarkaria, J. N. (2005). Disruption of parallel and converging signaling pathways contributes to the synergistic antitumor effects of simultaneous mTOR and EGFR inhibition in GBM cells. *Neoplasia*, 7(10).

Rassenti, L. Z., Jain, S., Keating, M. J., Wierda, W. G., Grever, M. R., Byrd, J. C., Kay, N. E., Brown, J. R., Gribben, J. G., Neuberg, D. S., He, F., Greaves, A. W., Rai, K. R., and Kipps, T. J. (2008). Relative value of ZAP-70, CD38, and immunoglobulin mutation status in predicting aggressive disease in chronic lymphocytic leukemia. *Blood*, 112(5).

Ren, S., He, K., Girshick, R., and Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.*

Riches, J. C., O'Donovan, C. J., Kingdon, S. J., McClanahan, F., Clear, A. J., Neuberg, D. S., Werner, L., Croce, C. M., Ramsay, A. G., Rassenti, L. Z., Kipps, T. J., and Gribben, J. G. (2014). Trisomy 12 chronic lymphocytic leukemia cells exhibit upregulation of integrin signaling that is modulated by NOTCH1 mutations. *Blood*, 123(26).

Ridky, T. W., Chow, J. M., Wong, D. J., and Khavari, P. A. (2010). Invasive three-dimensional organotypic neoplasia from multiple normal human epithelia. *Nat. Med.*, 16(12).

Roemer, T., Davies, J., Giaever, G., and Nislow, C. (2012). Bugs, drugs and chemical genomics.

Roider, T., Brinkmann, B., and Kim, V. (2021). Ex vivo model of bispecific antibodies in nodal B cell lymphoma identifies a response-associated regulatory T cell phenotype and effective drug combinations. *Manuscript in preparation*.

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*.

Rossi, G., Manfrin, A., and Lutolf, M. P. (2018). Progress and potential in organoid research.

Rozovski, U., Hazan-Halevy, I., Keating, M. J., and Estrov, Z. (2014). Personalized medicine in CLL: Current status and future perspectives.

Sachs, N. and Clevers, H. (2014). Organoid cultures for the analysis of cancer phenotypes.

Sachs, N., de Ligt, J., Kopper, O., Gogola, E., Bounova, G., Weeber, F., Balgobind, A. V., Wind, K., Gracanin, A., Begthel, H., Korving, J., van Boxtel, R., Duarte, A. A., Lelieveld, D., van Hoeck, A., Ernst, R. F., Blokzijl, F., Nijman, I. J., Hoogstraat, M., van de Ven, M., Egan, D. A., Zinzalla, V., Moll, J., Boj, S. F., Voest, E. E., Wessels, L., van Diest, P. J., Rottenberg, S., Vries, R. G. J., Cuppen, E., and Clevers, H. (2018). A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. *Cell*, 172(1-2).

Saeys, Y., Inza, I., and Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics.

Sahu, S. K. and Anand, A. (2018). Drug-drug interaction extraction from biomedical texts using long short-term memory network. *J. Biomed. Inform.*, 86.

Santiago, M., Lee, W., Fayad, A. A., Coe, K. A., Rajagopal, M., Do, T., Hennessen, F., Srisuknimit, V., Müller, R., Meredith, T. C., and Walker, S. (2018). Genome-wide mutant profiling predicts the mechanism of a Lipid II binding antibiotic article. *Nat. Chem. Biol.*, 14(6).

Sargentini, N. J., Gularte, N. P., and Hudman, D. A. (2016). Screen for genes involved in radiation survival of Escherichia coli and construction of a reference database. *Mutat. Res. - Fundam. Mol. Mech. Mutagen.*, 793-794.

Scannell, J. W. and Bosley, J. (2016). When quality beats quantity: Decision theory, drug discovery, and the reproducibility crisis. *PLoS One*, 11(2).

Schütte, M., Risch, T., Abdavi-Azar, N., Boehnke, K., Schumacher, D., Keil, M., Yildiriman, R., Jandrasits, C., Borodina, T., Amstislavskiy, V., Worth, C. L., Schweiger, C., Liebs, S., Lange, M., Warnatz, H. J., Butcher, L. M., Barrett, J. E., Sultan, M., Wierling, C., Golob-Schwarzl, N., Lax, S., Uranitsch, S., Becker, M., Welte, Y., Regan, J. L., Silvestrov, M., Kehler, I., Fusi, A., Kessler, T., Herwig, R., Landegren, U., Wienke, D., Nilsson, M., Velasco, J. A., Garin-Chesa, P., Reinhard, C., Beck, S., Schäfer, R., Regenbrecht, C. R., Henderson, D., Lange, B., Haybaeck, J., Keilholz, U., Hoffmann, J., Lehrach, H., and Yaspo, M. L. (2017). Molecular dissection of colorectal cancer in pre-clinical models identifies biomarkers predicting sensitivity to EGFR inhibitors. *Nat. Commun.*, 8.

Senft, D., Leiserson, M. D., Ruppin, E., and Ronai, Z. A. (2017). Precision Oncology: The Road Ahead.

Severin, F., Frezzato, F., Visentin, A., Martini, V., Trimarco, V., Carraro, S., Tibaldi, E., Maria Brunati, A., Piazza, F., Semenzato, G., Facco, M., and Trentin, L. (2019). In chronic lymphocytic leukemia the JAK2/STAT3 pathway is constitutively activated and its inhibition leads to CLL cell death unaffected by the protective bone marrow microenvironment. *Cancers (Basel).*, 11(12).

Seyhan, A. A. (2019). Lost in translation: the valley of death across preclinical and clinical divide – identification of problems and overcoming obstacles. *Transl. Med. Commun.*, 4(1).

Shanafelt, T. D. (2009). Predicting clinical outcome in CLL: how and why.

Shanafelt, T. D., Drake, M. T., Maurer, M. J., Allmer, C., Rabe, K. G., Slager, S. L., Weiner, G. J., Call, T. G., Link, B. K., Zent, C. S., Kay, N. E., Hanson, C. A., Witzig, T. E., and Cerhan, J. R. (2011). Vitamin D insufficiency and prognosis in chronic lymphocytic leukemia. *Blood*.

Shapley, L. S. (1953). A Value for n-Person Games. In *Contributions to the Theory of Games (AM-28), Vol. II*.

Shi, J. and Vakoc, C. R. (2014). The Mechanisms behind the Therapeutic Activity of BET Bromodomain Inhibition.

Shrikumar, A., Greenside, P., and Kundaje, A. (2017). Learning important features through propagating activation differences. In *34th Int. Conf. Mach. Learn. ICML 2017*, volume 7.

Silver, L. L. and Bostian, K. A. (1993). Discovery and development of new antibiotics: The problem of antibiotic resistance.

Simonyan, K., Vedaldi, A., and Zisserman, A. (2014). Deep inside convolutional networks: Visualising image classification models and saliency maps. In *2nd Int. Conf. Learn. Represent. ICLR 2014 - Work. Track Proc.*

Sison, E. A. R., Magoon, D., Li, L., Annesley, C. E., Rau, R. E., Small, D., and Brown, P. (2014). Plerixafor as a chemosensitizing agent in pediatric acute lymphoblastic leukemia: Efficacy and potential mechanisms of resistance to CXCR4 inhibition. *Oncotarget*.

Snijder, B., Sacher, R., Rämö, P., Damm, E. M., Liberali, P., and Pelkmans, L. (2009). Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature*, 461(7263).

Snijder, B., Vladimer, G. I., Krall, N., Miura, K., Schmolke, A. S., Kornauth, C., Lopez de la Fuente, O., Choi, H. S., van der Kouwe, E., Gültekin, S., Kazianka, L., Bigenzahn, J. W., Hoermann, G., Prutsch, N., Merkel, O., Ringler, A., Sabler, M., Jeryczynski, G., Mayerhoefer, M. E., Simonitsch-Klupp, I., Ocko, K., Felberbauer, F., Müllauer, L., Prager,

G. W., Korkmaz, B., Kenner, L., Sperr, W. R., Kralovics, R., Gisslinger, H., Valent, P., Kubicek, S., Jäger, U., Staber, P. B., and Superti-Furga, G. (2017). Image-based ex-vivo drug screening for patients with aggressive haematological malignancies: interim results from a single-arm, open-label, pilot study. *Lancet Haematol.*

Sommer, C., Straehle, C., Kothe, U., and Hamprecht, F. A. (2011). Ilastik: Interactive learning and segmentation toolkit. In *Proc. - Int. Symp. Biomed. Imaging*.

Stamatopoulos, B., Meuleman, N., de Bruyn, C., Pieters, K., Mineur, P., le Roy, C., Saint-Georges, S., Varin-Blank, N., Cymbalista, F., Bron, D., and Lagneaux, L. (2012). AMD3100 disrupts the cross-talk between chronic lymphocytic leukemia cells and a mesenchymal stromal or nurse-like cell-based microenvironment: Pre-clinical evidence for its association with chronic lymphocytic leukemia treatments. *Haematologica*, 97(4).

Stengel, A., Kern, W., Zenger, M., Perglerová, K., Schnittger, S., Haferlach, T., and Haferlach, C. (2016). Genetic characterization of T-PLL reveals two major biologic subgroups and JAK3 mutations as prognostic marker. *Genes Chromosom. Cancer*, 55(1).

Sullivan, G. J., Delgado, N. N., Maharjan, R., and Cain, A. K. (2020). How antibiotics work together: molecular mechanisms behind combination therapy.

Summers, T. A. and Jaffe, E. S. (2011). Hairy cell leukemia diagnostic criteria and differential diagnosis. *Leuk. Lymphoma*, 52(SUPPL. 2).

Sun, J., Deng, Z., and Yan, A. (2014). Bacterial multidrug efflux pumps: Mechanisms, physiology and pharmacological exploitations.

Sun, W., Sanderson, P. E., and Zheng, W. (2016a). Drug combination therapy increases successful drug repositioning.

Sun, W., Weingarten, R. A., Xu, M., Southall, N., Dai, S., Shinn, P., Sanderson, P. E., Williamson, P. R., Frank, K. M., and Zheng, W. (2016b). Rapid antimicrobial susceptibility test for identification of new therapeutics and drug combinations against multidrug-resistant bacteria. *Emerg. Microbes Infect.*, 5(11).

Swords, R. T., Azzam, D., Al-Ali, H., Lohse, I., Volmar, C. H., Watts, J. M., Perez, A., Rodriguez, A., Vargas, F., Elias, R., Vega, F., Zelent, A., Brothers, S. P., Abbasi, T., Trent, J., Rangwala, S., Deutsch, Y., Conneally, E., Drusbosky, L., Cogle, C. R., and Wahlestedt, C. (2018). Ex-vivo sensitivity profiling to guide clinical decision making in acute myeloid leukemia: A pilot study. *Leuk. Res.*, 64.

Taber, H. W., Mueller, J. P., Miller, P. F., and Arrow, A. S. (1987). Bacterial uptake of aminoglycoside antibiotics.

Taigman, Y., Yang, M., Ranzato, M., and Wolf, L. (2014). DeepFace: Closing the gap to human-level performance in face verification. In *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*

Tang, J., Alelyani, S., and Liu, H. (2014). Feature selection for classification: A review. In *Data Classif. Algorithms Appl.*

Thimiri, D. B., Cremaschi, A., Skånland, S. S., Gade, A., Schjesvold, F. H., Tjønnfjord, G. E., A Munthe, L., and Tasken, K. (2018). In-Vitro Drug Sensitivity Screening in Chronic Lymphocytic Leukemia (CLL) Primary Patient Samples Identifies Drug Candidates for Precision Cancer Therapy. *Blood*, 132(Supplement 1).

Tissino, E., Benedetti, D., Herman, S. E., ten Hacken, E., Ahn, I. E., Chaffee, K. G., Rossi, F. M., Bo, M. D., Bulian, P., Bomben, R., Bayer, E., Härzschel, A., Gutjahr, J. C., Postorino, M., Santinelli, E., Ayed, A., Zaja, F., Chiarenza, A., Pozzato, G., Chigaev, A., Sklar, L. A., Burger, J. A., Ferrajoli, A., Shanafelt, T. D., Wiestner, A., Del Poeta, G., Hartmann, T. N., Gattei, V., and Zucchetto, A. (2018). Functional and clinical relevance of VLA-4 (CD49d/CD29) in ibrutinib-treated chronic lymphocytic leukemia. *J. Exp. Med.*, 215(2).

Toprak, E., Veres, A., Michel, J. B., Chait, R., Hartl, D. L., and Kishony, R. (2012). Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat. Genet.*, 44(1).

Tyers, M. and Wright, G. D. (2019). Drug combinations: a strategy to extend the life of antibiotics in the 21st century.

Tyner, J. W., Tognon, C. E., Bottomly, D., Wilmot, B., Kurtz, S. E., Savage, S. L., Long, N., Schultz, A. R., Traer, E., Abel, M., Agarwal, A., Blucher, A., Borate, U., Bryant, J., Burke, R., Carlos, A., Carpenter, R., Carroll, J., Chang, B. H., Coblentz, C., D'Almeida, A., Cook, R., Danilov, A., Dao, K. H. T., Degnin, M., Devine, D., Dibb, J., Edwards, D. K., Eide, C. A., English, I., Glover, J., Henson, R., Ho, H., Jemal, A., Johnson, K., Johnson, R., Junio, B., Kaempf, A., Leonard, J., Lin, C., Liu, S. Q., Lo, P., Loriaux, M. M., Luty, S., Macey, T., MacManiman, J., Martinez, J., Mori, M., Nelson, D., Nichols, C., Peters, J., Ramsdill, J., Rofelty, A., Schuff, R., Searles, R., Segerdell, E., Smith, R. L., Spurgeon, S. E., Sweeney, T., Thapa, A., Visser, C., Wagner, J., Watanabe-Smith, K., Werth, K., Wolf, J., White, L., Yates, A., Zhang, H., Cogle, C. R., Collins, R. H., Connolly, D. C., Deininger, M. W., Drusbosky, L., Hourigan, C. S., Jordan, C. T., Kropf, P., Lin, T. L., Martinez, M. E., Medeiros, B. C., Pallapati, R. R., Pollyea, D. A., Swords, R. T., Watts, J. M., Weir, S. J., Wiest, D. L., Winters, R. M., McWeeney, S. K., and Druker, B. J. (2018). Functional genomic landscape of acute myeloid leukaemia. *Nature*, 562(7728).

Tyner, J. W., Yang, W. F., Bankhead, A., Fan, G., Fletcher, L. B., Bryant, J., Glover, J. M., Chang, B. H., Spurgeon, S. E., Fleming, W. H., Kovacsovics, T., Gotlib, J. R., Oh, S. T., Deininger, M. W., Zwaan, C. M., Den Boer, M. L., Van Den Heuvel-Eibrink, M. M., O'Hare, T., Druker, B. J., and Loriaux, M. M. (2013). Kinase pathway dependence in primary human leukemias determined by rapid inhibitor screening. *Cancer Res.*, 73(1).

Uy, G. L., Rettig, M. P., Motabi, I. H., McFarland, K., Trinkaus, K. M., Hladnik, L. M., Kulkarni, S., Abboud, C. N., Cashen, A. F., Stockerl-Goldstein, K. E., Vij, R., Westervelt, P., and DiPersio, J. F. (2012). A phase 1/2 study of chemosensitization with the CXCR4 antagonist plerixafor in relapsed or refractory acute myeloid leukemia. *Blood*.

Van Der Maaten, L. and Hinton, G. (2008). Visualizing Data using t-SNE. *J. Mach. Learn. Res.*, 9:2579–2605.

van der Velden, D. L., van Herpen, C. M., van Laarhoven, H. W., Smit, E. F., Groen, H. J., Willems, S. M., Nederlof, P. M., Langenberg, M. H., Cuppen, E., Sleijfer, S., Steeghs, N., and Voest, E. E. (2017). Molecular Tumor Boards: Current practice and future needs. *Ann. Oncol.*, 28(12).

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., and the scikit-image contributors (2014). scikit-image: image processing in Python. *PeerJ*, 2:e453.

Ventola, C. L. (2015). The antibiotic resistance crisis: causes and threats. *P T J.*, 40(4).

Vianello, F., Villanova, F., Tisato, V., Lymperi, S., Ho, K. K., Gomes, A. R., Marin, D., Bonnet, D., Apperley, J., Lam, E. W., and Dazzi, F. (2010). Bone marrow mesenchymal stromal cells non-selectively protect chronic myeloid leukemia cells from imatinib-induced apoptosis via the CXCR4/CXCL12 axis. *Haematologica*, 95(7).

Weigert, M., Royer, L., Jug, F., and Myers, G. (2017). Isotropic reconstruction of 3D fluorescence microscopy images using convolutional neural networks. In *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, volume 10434 LNCS.

Weiner, L. M., Dhodapkar, M. V., and Ferrone, S. (2009). Monoclonal antibodies for cancer immunotherapy.

Wellek, S. (1993). A Log-Rank Test for Equivalence of Two Survivor Functions. *Biometrics*, 49(3).

Wicha, S. G., Chen, C., Clewe, O., and Simonsson, U. S. (2017). A general pharmacodynamic interaction model identifies perpetrators and victims in drug interactions. *Nat. Commun.*, 8(1).

Wiestner, A., Rosenwald, A., Barry, T. S., Wright, G., Davis, R. E., Henrickson, S. E., Zhao, H., Ibbotson, R. E., Orchard, J. A., Davis, Z., Stetler-Stevenson, M., Raffeld, M., Arthur, D. C., Marti, G. E., Wilson, W. H., Hamblin, T. J., Oscier, D. G., and Staudt, L. M. (2003). ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. *Blood*, 101(12).

Wildenhain, J., Spitzer, M., Dolma, S., Jarvik, N., White, R., Roy, M., Griffiths, E., Bellows, D. S., Wright, G. D., and Tyers, M. (2015). Prediction of Synergism from Chemical-Genetic Interactions by Machine Learning. *Cell Syst.*

Woo, S. U., Sangai, T., Akcakanat, A., Chen, H., Wei, C., and Meric-Bernstam, F. (2017). Vertical inhibition of the PI3K/Akt/mTOR pathway is synergistic in breast cancer. *Oncogenesis*, 6(10).

Wouters, O. J., McKee, M., and Luyten, J. (2020). Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018.

Wu, K., Otoo, E., and Shoshani, A. (2005). Optimizing connected component labeling algorithms. In *Med. Imaging 2005 Image Process.*, volume 5747.

Yasaka, K. and Abe, O. (2018). Deep learning and artificial intelligence in radiology: Current applications and future directions.

Yeh, P., Tschumi, A. I., and Kishony, R. (2006). Functional classification of drugs by properties of their pairwise interactions. *Nat. Genet.*, 38(4).

Yeh, P. J., Hegreness, M. J., Aiden, A. P., and Kishony, R. (2009). Drug interactions and the evolution of antibiotic resistance. *Nat. Rev. Microbiol.*, 7(6).

Zeng, Z., Shi, Y. X., Samudio, I. J., Wang, R. Y., Ling, X., Frolova, O., Levis, M., Rubin, J. B., Negrin, R. R., Estey, E. H., Konoplev, S., Andreeff, M., and Konopleva, M. (2009). Targeting the leukemia microenvironment by CXCR4 inhibition overcomes resistance to kinase inhibitors and chemotherapy in AML. *Blood*.

Zenz, T., Eichhorst, B., Busch, R., Denzel, T., Häbe, S., Winkler, D., Bühler, A., Edelmann, J., Bergmann, M., Hopfinger, G., Hensel, M., Hallek, M., Döhner, H., and Stilgenbauer, S. (2010a). TP53 mutation and survival in chronic lymphocytic leukemia. *J. Clin. Oncol.*, 28(29).

Zenz, T., Mertens, D., Küppers, R., Döhner, H., and Stilgenbauer, S. (2010b). From pathogenesis to treatment of chronic lymphocytic leukaemia.

Zhang, C. and Ma, Y. (2012). *Ensemble machine learning: Methods and applications*.

Zhang, S., Smartt, H., Holgate, S. T., and Roche, W. R. (1999). Growth factors secreted by bronchial epithelial cells control myofibroblast proliferation: An in vitro co-culture model of airway remodeling in asthma. *Lab. Investig.*, 79(4).

Zhang, X., Tu, H., Yang, Y., Jiang, X., Hu, X., Luo, Q., and Li, J. (2019). Bone marrow–derived mesenchymal stromal cells promote resistance to tyrosine kinase inhibitors in chronic myeloid leukemia via the IL-7/JAK1/STAT5 pathway. *J. Biol. Chem.*, 294(32).

Zhang, Y., Zheng, W., Lin, H., Wang, J., Yang, Z., and Dumontier, M. (2018). Drug-drug interaction extraction via hierarchical RNNs on sequence and shortest dependency paths. *Bioinformatics*, 34(5).

Zucchetto, A., Caldana, C., Benedetti, D., Tissino, E., Rossi, F. M., Hutterer, E., Pozzo, F., Bomben, R., Dal Bo, M., D'Arena, G., Zaja, F., Pozzato, G., Di Raimondo, F., Hartmann, T. N., Rossi, D., Gaidano, G., Del Poeta, G., and Gattei, V. (2013). CD49d is overexpressed by trisomy 12 chronic lymphocytic leukemia cells: Evidence for a methylation-dependent regulation mechanism. *Blood*, 122(19).

Supplement A

Supplementary figures

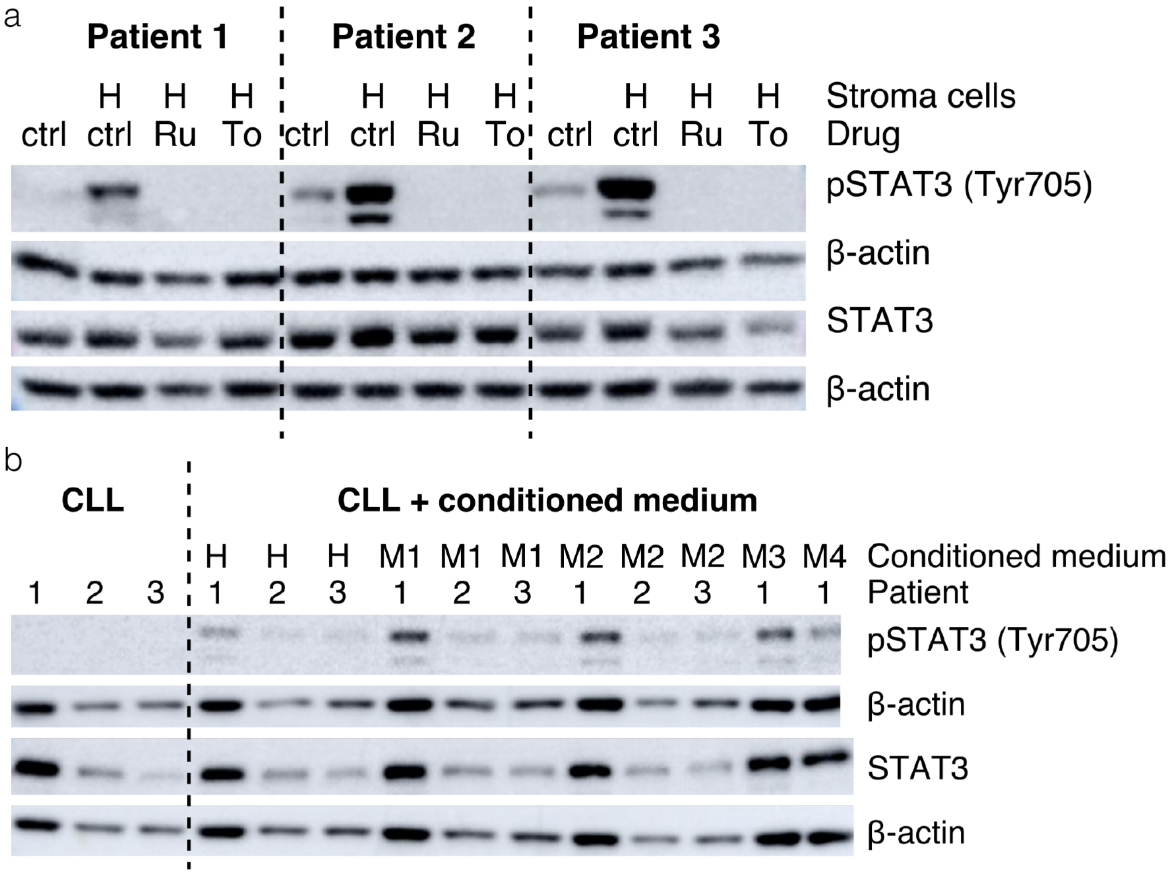


Fig. S1 JAK-STAT validations. Data produced by S. Herbst.

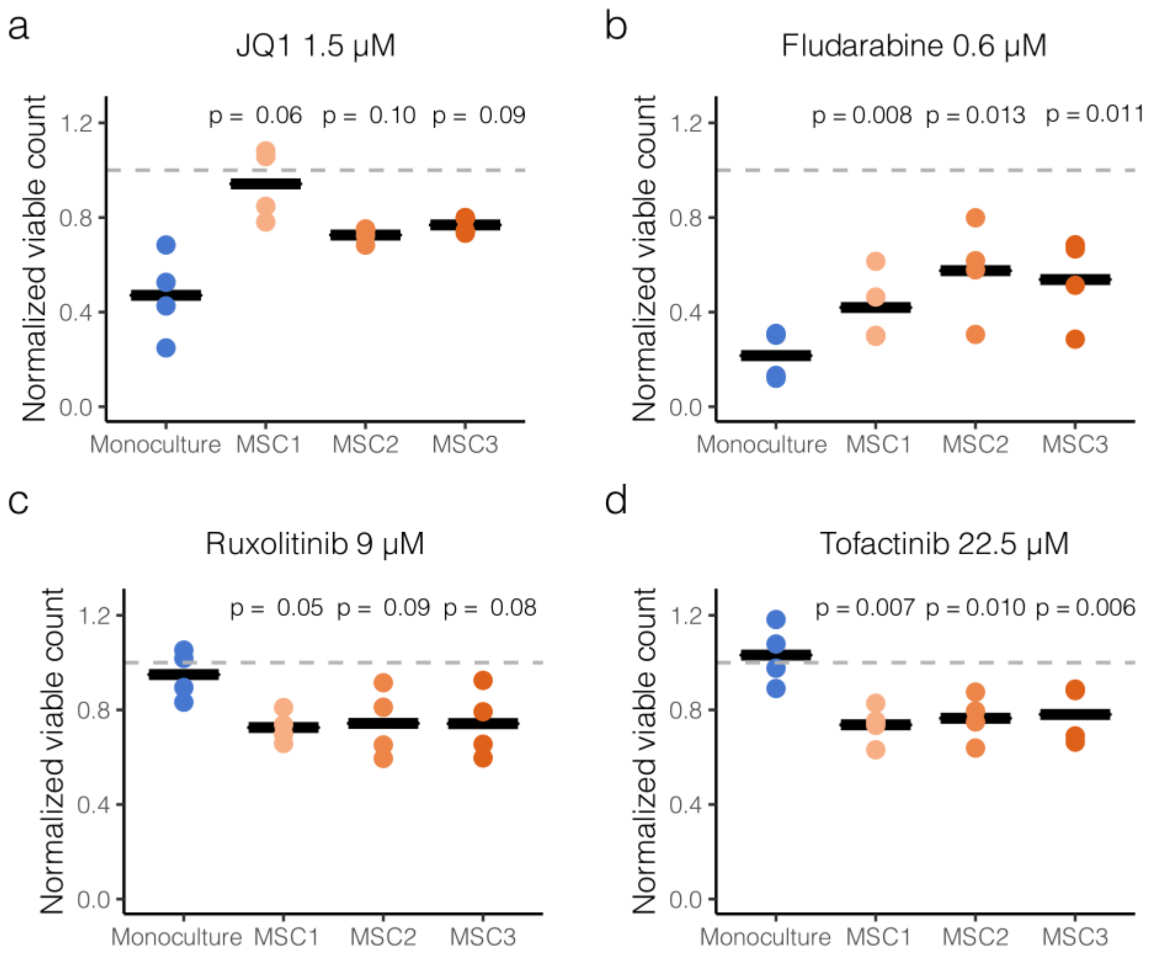


Fig. S2 MSC validations. Data produced jointly with S. Herbst.

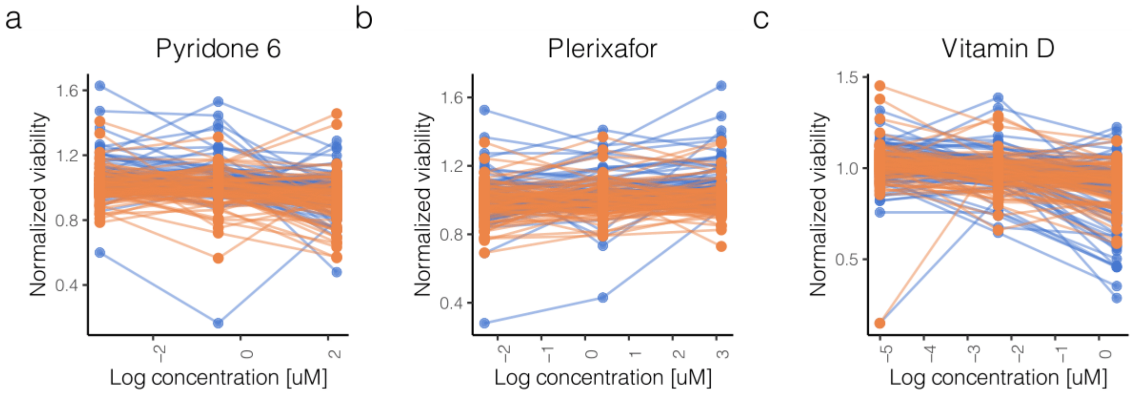


Fig. S3 Dose-response curves of a) pyridone 6 b) plerixafor and c) vitamin D in mono- and coculture. Data produced jointly with S. Herbst.

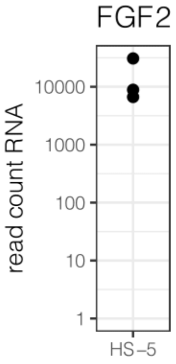


Fig. S4 FGF2 expression in HS-5 cells. Data produced by S. Herbst.

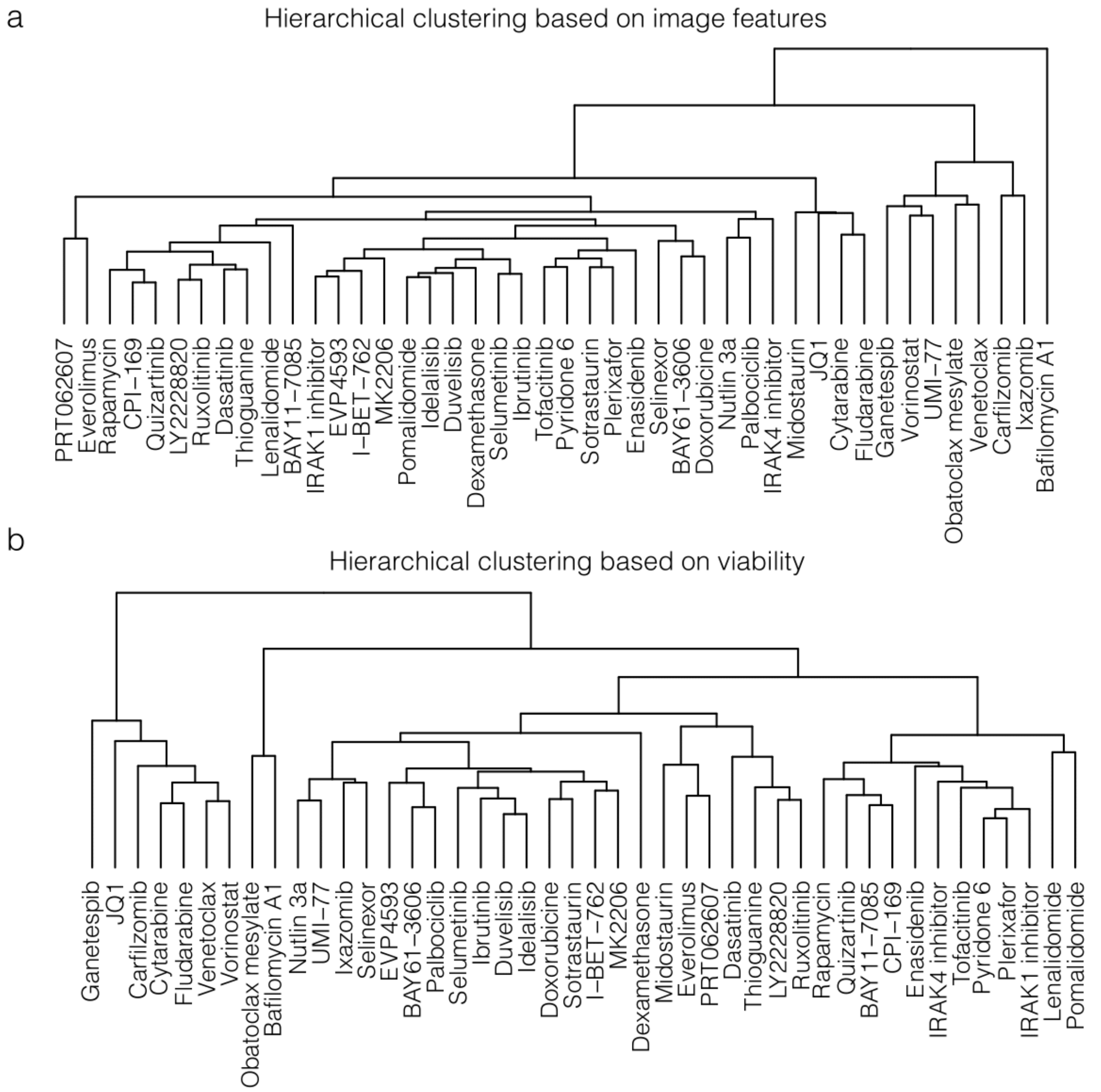


Fig. S5 Clustering based on a) image features b) viability alone. Data produced jointly with S. Herbst.

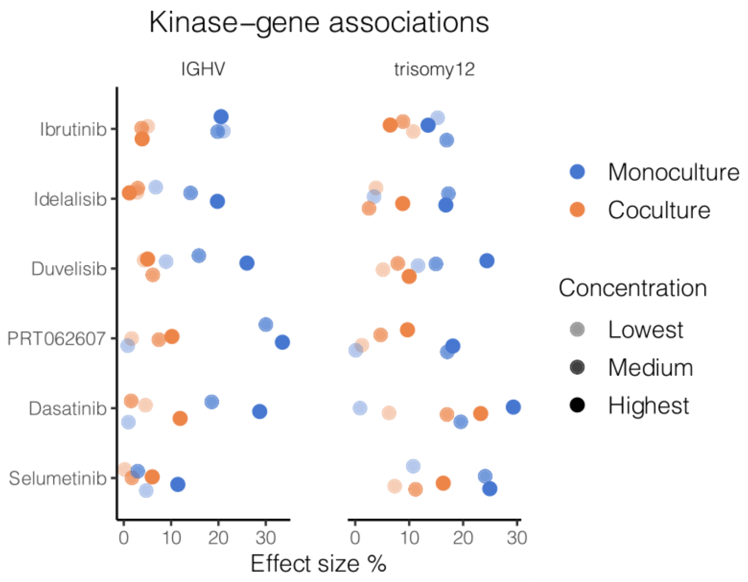


Fig. S6 Effect size of IGHV and trisomy12 associations with BCR inhibitor response in mono- and coculture. Data produced jointly with S. Herbst.

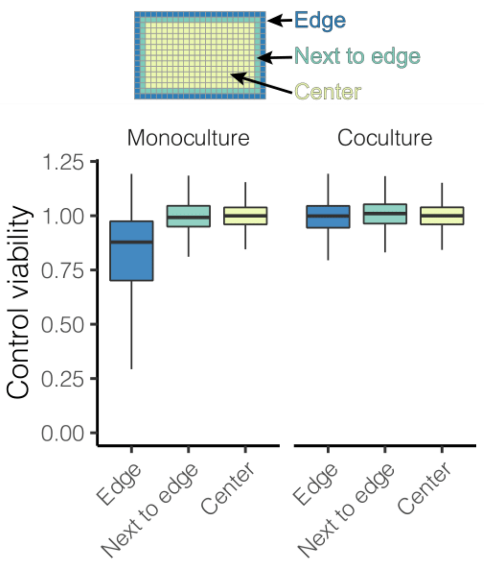


Fig. S7 Edge effect in mono- and coculture. Data produced by S. Herbst.

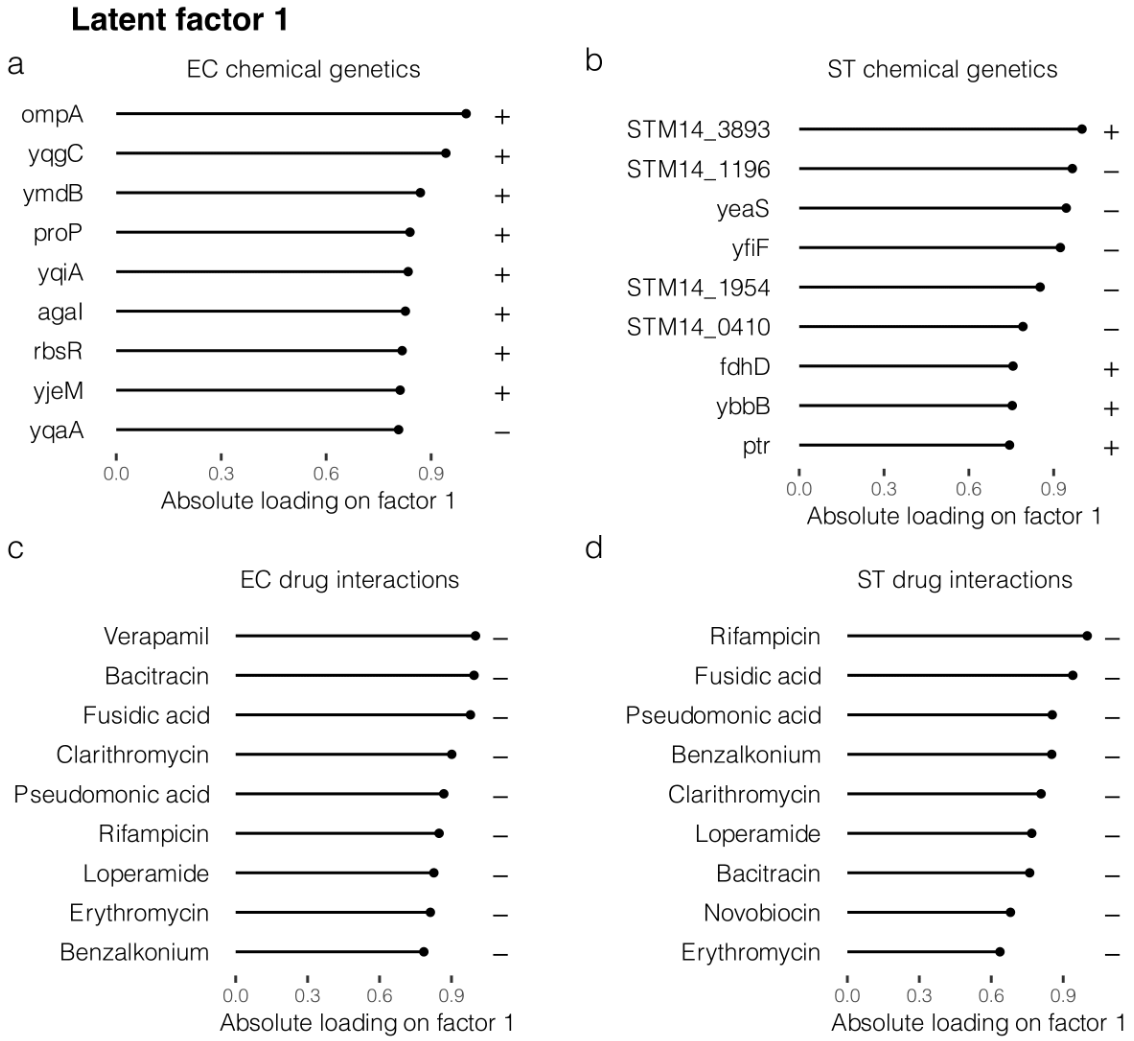


Fig. S8 Latent factor 1 loadings on different omic views. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

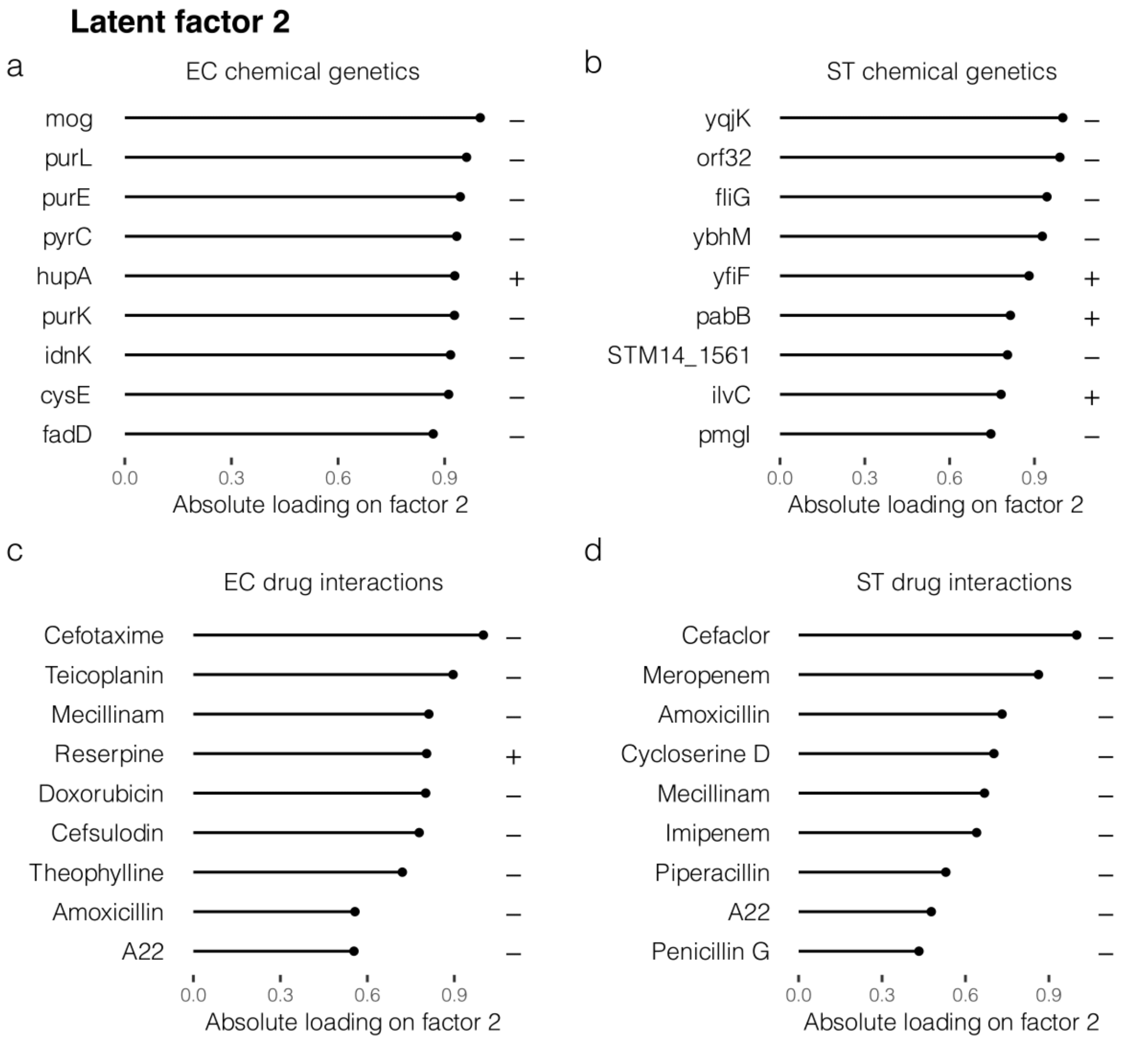


Fig. S9 Latent factor 2 loadings on different omic views. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

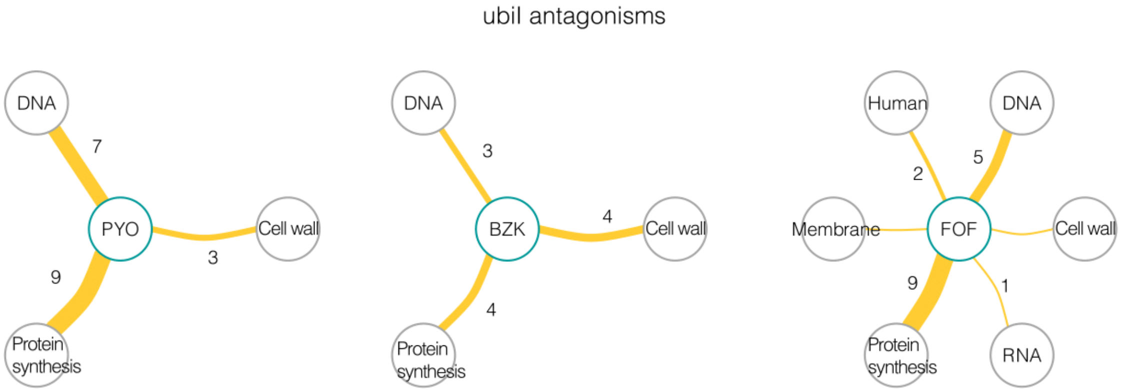


Fig. S10 Promiscuous antagonists pyocyanin (PYO), benzalkonium (BZK) and fosfomycin (FOF) all lose fitness in Δ*ubiI*. Edge width and label indicate the number of antagonisms occurring between the drug in the center and a compound class. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011).

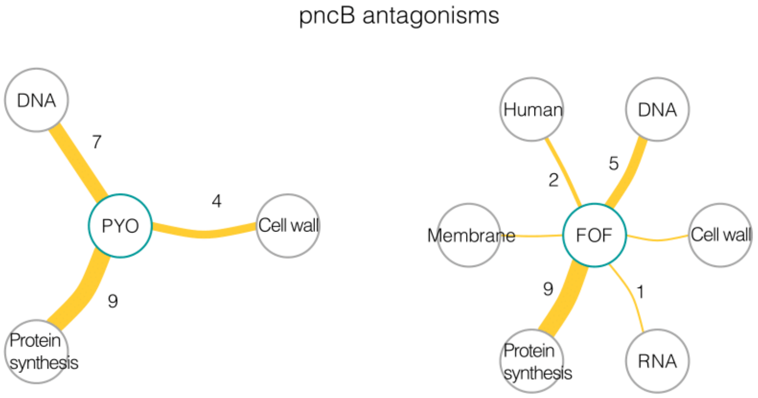


Fig. S11 Promiscuous antagonists pyocyanin (PYO) and fosfomycin (FOF) lose fitness in Δ*pncB*. Edge width and label indicate the number of antagonisms occurring between the drug in the center and a compound class. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011).

Supplement B

Supplementary tables

Table S1 Selected features used for compound profiling in non-CLL diseases (AML, HCL, MCL and T-PLL). Data produced jointly with S. Herbst.

Channel	Feature	Parameter/array index
Calcein	convex area	
Calcein	eccentricity	
Calcein	moments central	[2][2]
Calcein	moments hu	[1]
Calcein	weighted moments central	[2][2]
Hoechst	DifferenceVariance	$[d = 3][1]$
Hoechst	InfoMeas1	$[d = 5][3]$
Hoechst	mean intensity	
Hoechst	SumAverage	$[d = 3][0]$
Hoechst	SumVariance	$[d = 7][0]$
Hoechst	SumVariance	$[d = 3][0]$
Hoechst	Variance	$[d = 7][0]$
Hoechst	Variance	$[d = 3][0]$
Hoechst	Variance	$[d = 3][1]$
Lysosomal	area	
Lysosomal	contrast	[3]
Lysosomal	Contrast	$[d = 7][1]$
Lysosomal	Correlation	$[d = 7][3]$
Lysosomal	DifferenceEntropy	$[d = 3][2]$
Lysosomal	DifferenceEntropy	$[d = 7][2]$

Lysosomal	DifferenceEntropy	$[d = 5][2]$
Lysosomal	DifferenceEntropy	$[d = 7][1]$
Lysosomal	DifferenceVariance	$[d = 3][2]$
Lysosomal	DifferenceVariance	$[d = 7][1]$
Lysosomal	DifferenceVariance	$[d = 5][1]$
Lysosomal	DifferenceVariance	$[d = 3][3]$
Lysosomal	DifferenceVariance	$[d = 5][2]$
Lysosomal	dissimilarity	[1]
Lysosomal	dissimilarity	[3]
Lysosomal	dissimilarity	[5]
Lysosomal	dissimilarity	[10]
Lysosomal	InfoMeas1	$[d = 7][0]$
Lysosomal	InfoMeas1	$[d = 5][3]$
Lysosomal	InfoMeas2	$[d = 5][0]$
Lysosomal	InfoMeas2	$[d = 7][1]$
Lysosomal	InfoMeas2	$[d = 7][0]$
Lysosomal	InfoMeas2	$[d = 5][1]$
Lysosomal	mean intensity	
Lysosomal	SumAverage	$[d = 7][3]$
Lysosomal	SumVariance	$[d = 7][2]$
Lysosomal	SumVariance	$[d = 5][3]$
Lysosomal	SumVariance	$[d = 3][3]$
Lysosomal	Variance	$[d = 5][2]$
Lysosomal	zernike	$[r = 18][1]$

Table S2 Selected features used for compound profiling in CLL. Data produced jointly with S. Herbst.

Channel	Feature	Parameter/array index
Hoechst	Correlation	$[d=3][0]$
Hoechst	Correlation	$[d=7][3]$
Hoechst	eccentricity	
Hoechst	extent	
Hoechst	inertia tensor eigvals	[0]
Hoechst	inertia tensor eigvals	[1]
Hoechst	major axis length	
Hoechst	minor axis length	
Hoechst	moments hu	[1]
Hoechst	moments hu	[0]
Hoechst	moments hu	[3]
Hoechst	solidity	
Hoechst	weighted moments hu	[1]
Hoechst	zernike	$[r=15][1]$
Lysosomal	contrast	[3]
Lysosomal	contrast	[4]
Lysosomal	contrast	[0]
Lysosomal	Contrast	$[d=7][3]$
Lysosomal	Contrast	$[d=5][0]$
Lysosomal	Contrast	$[d=3][0]$
Lysosomal	DifferenceEntropy	$[d=3][0]$
Lysosomal	DifferenceEntropy	$[d=5][0]$
Lysosomal	dissimilarity	[3]
Lysosomal	dissimilarity	[7]
Lysosomal	dissimilarity	[1]
Lysosomal	dissimilarity	[11]
Lysosomal	dissimilarity	[0]
Lysosomal	inertia tensor eigvals	[0]
Lysosomal	major axis length	
Lysosomal	moments hu	[3]
Lysosomal	moments hu	[0]
Lysosomal	zernike	$[r=15][1]$

Lysosomal	zernike	$[r = 20]$[18]
Lysosomal	zernike	$[r = 18]$[1]
Lysosomal	zernike	$[r = 18]$[26]
Lysosomal	zernike	$[r = 20]$[48]
Lysosomal	zernike	$[r = 20]$[20]
Lysosomal	zernike	$[r = 18]$[4]
Lysosomal	zernike	$[r = 20]$[28]
Lysosomal	zernike	$[r = 20]$[35]
Lysosomal	zernike	$[r = 18]$[20]
Lysosomal	zernike	$[r = 20]$[22]

Table S3 Compounds used in the antibiotic combination screen in Gram-positive species. Compound class annotation generated jointly with E. Cacace.

Drug	Code	Subclass	Class
Amoxicillin	AMX	beta-lactams	cell wall
Amoxicillinclavulanic	AMXCLA	beta-lactams	cell wall
Cefepime	FEP	beta-lactams	cell wall
Cefotaxime	CTX	beta-lactams	cell wall
Cefuroxime	CXA	beta-lactams	cell wall
Cephalexin	LEX	beta-lactams	cell wall
Clavulanic	CLA	beta-lactams	cell wall
Imipenem	IPM	beta-lactams	cell wall
Oxacillin	OXA	beta-lactams	cell wall
PenicillinG	PEN	beta-lactams	cell wall
Daptomycin	DAP	glycopeptides and lipopeptides	cell wall
Oritavancin	ORI	glycopeptides and lipopeptides	cell wall
Teicoplanin	TEC	glycopeptides and lipopeptides	cell wall
Vancomycin	VAN	glycopeptides and lipopeptides	cell wall
Bacitracin	BAC	other cell wall	cell wall
CycloserineD	DCS	other cell wall	cell wall
Fosfomycin	FOF	other cell wall	cell wall
Nisin	NSN	other cell wall	cell wall
Metronidazole	MTR	DNA/oxidative stress	DNA
Nitrofurantoin	NIT	DNA/oxidative stress	DNA
Phleomycin	PLM	DNA/oxidative stress	DNA
Ciprofloxacin	CIP	fluoroquinolones	DNA
Levofloxacin	LCX	fluoroquinolones	DNA
Moxifloxacin	MXF	fluoroquinolones	DNA
Cotrimoxazole	SMXTMP	folate biosynthesis	DNA
Sulfamethoxazol	SMX	folate biosynthesis	DNA
Trimethoprim	TMP	folate biosynthesis	DNA
Doxorubicin	DXR	non-antibiotic	non-antibiotic
MitomycinC	MMC	non-antibiotic	non-antibiotic
Novobiocin	NVB	other DNA	DNA
Rifampicin	RIF	other DNA	DNA
Benzalkonium	BZK	membrane	membrane
EGCG	EGCG	membrane	membrane
Triclosan	TRI	membrane	membrane

Acetylsalicylicacid	ASA	non-antibiotic	non-antibiotic
Auranofin	AUR	non-antibiotic	non-antibiotic
Loperamide	LOP	non-antibiotic	non-antibiotic
Miconazole	MCZ	non-antibiotic	non-antibiotic
Procaine	PRC	non-antibiotic	non-antibiotic
Streptozotocin	STZ	non-antibiotic	non-antibiotic
CCCP	CCCP	oxidative stress/PMF	oxidative stress/PMF
Paraquat	PQ	oxidative stress/PMF	oxidative stress/PMF
Pyocyanin	PYO	oxidative stress/PMF	oxidative stress/PMF
ADEP	ADEP	protein degradation	protein degradation
U1	U1	protein degradation	protein degradation
Gentamicin	GEN	aminoglycosides	protein synthesis
Spectinomycin	SPT	aminoglycosides	protein synthesis
Streptomycin	STR	aminoglycosides	protein synthesis
Clindamycin	CLI	lincosamides	protein synthesis
Lincomycin	LIN	lincosamides	protein synthesis
Azithromycin	AZM	macrolides and ketolides	protein synthesis
Clarithromycin	CLR	macrolides and ketolides	protein synthesis
Erythromycin	ERI	macrolides and ketolides	protein synthesis
Telithromycin	TEL	macrolides and ketolides	protein synthesis
Chloramphenicol	CHL	other protein synthesis	protein synthesis
Fusidic acid	FUS	other protein synthesis	protein synthesis
Pseudomonic acid	MUP	other protein synthesis	protein synthesis
Puromycin	PUR	other protein synthesis	protein synthesis
Linezolid	LZD	oxazolidinones	protein synthesis
Tedizolid	TZD	oxazolidinones	protein synthesis
Pristinamycin	PRI	streptogramins	protein synthesis
Virginiamycin M1	VIR	streptogramins	protein synthesis
Clortetracycline	CTC	tetracyclines	protein synthesis
Doxycycline	DOX	tetracyclines	protein synthesis
Tigecycline	TGC	tetracyclines	protein synthesis

Table S4 Compounds used in the antibiotic combination screen in Gram-negative species (Brochado et al., 2018). Compound class annotation generated jointly with E. Cacace.

Drug	Code	Subclass	Class
A22	A22	other cell wall	cell wall
Amikacin	AMK	aminoglycoside	protein synthesis
Amoxicillin	AMX	beta-lactam	cell wall
Aztreonam	ATM	beta-lactam	cell wall
Azithromycin	AZM	macrolide	protein synthesis
Bacitracin	BAC	other cell wall	cell wall
Berberine	BBR	food additive	human
Bleomycin	BLM	multiple	DNA
Benzalkonium	BZK	human drug	human
CCCP	CCCP	PMF	PMF
Cefaclor	CEC	beta-lactam	cell wall
Cerulenin	CER	fatty acid biosynthesis	membrane
Cefsulodin	CFS	beta-lactam	cell wall
Chlorhexidine	CHG	LPS	membrane
CHIR-90	CHIR90	LPS	membrane
Chloramphenicol	CHL	chloramphenicol	protein synthesis
Ciprofloxacin	CIP	DNA gyrase	DNA
Clindamycin	CLI	lincosamide	protein synthesis
Clofazimine	CLZ	oxidative stress	DNA
Clarithromycin	CLR	macrolide	protein synthesis
Colistin	CST	LPS	membrane
Cefotaxime	CTX	beta-lactam	cell wall
Cycloserine D	DCS	other cell wall	cell wall
Doxycycline	DOX	tetracycline	protein synthesis
Doxorubicin	DXR	other DNA	DNA
EGCG	EGCG	fatty acid biosynthesis	membrane
Erythromycin	ERI	macrolide	protein synthesis
Fosfomycin	FOF	other cell wall	cell wall
Fusidic acid	FUS	other protein synthesis	protein synthesis
Gentamicin	GEN	aminoglycoside	protein synthesis
Imipenem	IPM	beta-lactam	cell wall
Levofloxacin	LCX	DNA gyrase	DNA
Cephalexin	LEX	beta-lactam	cell wall
Loperamide	LOP	human drug	human

Linezolid	LZD	oxazolidinone	protein synthesis
Mecillinam	MEC	beta-lactam	cell wall
Meropenem	MEM	beta-lactam	cell wall
Metformin	MTF	human drug	human
Minocycline	MIN	tetracycline	protein synthesis
Mitomycin C	MMC	other DNA	DNA
Metronidazole	MTR	other DNA	DNA
Pseudomonic acid	MUP	tRNA	protein synthesis
Nitrofurantoin	NIT	other DNA	DNA
Nonactin	NON	PMF	PMF
Nisin	NSN	other cell wall	cell wall
Novobiocin	NVB	DNA gyrase	DNA
Oxacillin	OXA	beta-lactam	cell wall
Penicillin G	PEN	beta-lactam	cell wall
Piperacillin	PIP	beta-lactam	cell wall
Phleomycin	PLM	other DNA	DNA
Polymyxin B	PMB	LPS	membrane
Paraquat	PQ	oxidative stress	oxidative stress
Procaine	PRC	human drug	human
Puromycin	PUR	tRNA	protein synthesis
Pyocyanin	PYO	oxidative stress	oxidative stress
Rifampicin	RIF	RNA polymerase	RNA
Reserpine	RSP	human drug	human
Sulfamonomethoxine	SMM	folic acid biosynthesis	DNA
Spiramycin	SPM	macrolide	protein synthesis
Spectinomycin	SPT	aminoglycoside	protein synthesis
Teicoplanin	TEC	other cell wall	cell wall
Tigecycline	TGC	tetracycline	protein synthesis
Trimethoprim	TMP	folic acid biosynthesis	DNA
Tobramycin	TOB	aminoglycoside	protein synthesis
Theophylline	TPH	food additive	human
Triclosan	TRI	fatty acid biosynthesis	membrane
Vanillin	VNL	food additive	human
Verapamil	VPM	human drug	human

Table S5 Top gene ontology (GO) terms enriched among the interaction-associated genes in *E. coli* and *S. typhimurium*. Drug interaction data source: Brochado et al. (2018). Chemogenomic data source: Nichols et al. (2011) and Pfalz (2017).

GO term	Genes	Associated
Organism: *E. coli*		
translation	adk, cysS, glyQ, infA pheS, prfC, rimK, rimP, rplI	Antagonism
aerobic respiration	erpA, nuoF, nuoG nuoM, nuoN	Antagonism
O antigen biosynthetic process	rfbD, rffG, rffA, rfbX	Antagonism
aromatic amino acid family biosynthetic process	aroA, aroB, aroC aroE, aroG, aroK	Antagonism
glycine catabolic process	gcvH, gcvP, gcvT, glyA	Antagonism
chorismate biosynthetic process	aroA, aroB, aroC aroE, aroG, aroK	Antagonism
lipopolysaccharide biosynthetic process	arnC, cpsG, etk, fepE galU, gmd, rfbD, rffG wbbI, wcaF, wcaJ, fcl galF, rffA, rfbX	Antagonism
’de novo’ IMP biosynthetic process	purB, purD, purK, purL	Synergy
proton transport	ampG, atpB, atpC atpD, atpE, atpF	Synergy
sulfate assimilation	cysC, cysD, cysJ, cysN	Synergy
hydrogen sulfide biosynthetic process	cysC, cysD, cysJ, cysN	Synergy
lipopolysaccharide biosynthetic process	arnC, cpsG, etk, fepE gmd, waaB, wcaF, wcaI fcl, rfe, rffC	Synergy
ATP synthesis coupled proton transport	atpA, atpB, atpC atpD, atpE, atpF	Synergy
Organism: *S. typhimurium*		
electron transport chain	STM14_0521, STM14_0523 STM14_1758, STM14_2857,	Antagonism

	STM14_2860, STM14_4850	
nonribosomal peptide biosynthetic process	STM14_0681, STM14_0694 STM14_0696	Antagonism
peptide metabolic process	STM14_0681, STM14_0694 STM14_0696, STM14_1573	Antagonism
generation of precursor metabolites and energy	STM14_0190, STM14_0521 STM14_0522, STM14_0523 STM14_1758, STM14_2857 STM14_2860, STM14_4464 STM14_4850, STM14_4885 STM14_5506	Antagonism
hydrogen ion transmembrane transport	STM14_0520, STM14_0521 STM14_0522, STM14_0523 STM14_4660, STM14_4960	Antagonism
chorismate biosynthetic process	STM14_0459, STM14_0884 STM14_1106, STM14_1649 STM14_4199	Antagonism
coenzyme biosynthetic process	STM14_2206, STM14_2883 STM14_4803, STM14_5090 STM14_930, STM14_931 STM14_932, STM14_934 STM14_987	Synergy
purine ribonucleotide biosynthetic process	STM14_0623, STM14_3064 STM14_3065, STM14_4739 STM14_5016, STM14_5248	Synergy
phosphate ion transport	STM14_4318, STM14_4649 STM14_4650	Synergy
’de novo’ UMP biosynthetic process	STM14_0077, STM14_1200 STM14_1332	Synergy
cofactor biosynthetic process	STM14_0519, STM14_0694 STM14_0696, STM14_2206 STM14_2883, STM14_4803 STM14_5090, STM14_930 STM14_931, STM14_932 STM14_934, STM14_987	Synergy

ribonucleoside monophosphate biosynthetic process	STM14_0077, STM14_0623 STM14_1200, STM14_1332 STM14_3064, STM14_3065 STM14_5016, STM14_5248	Synergy
prosthetic group metabolic process	STM14_4803, STM14_930 STM14_931, STM14_932 STM14_934, STM14_987	Synergy

Table S6 Inferred directional relationships in synergistic and antagonistic combinations in Gram-negative species (Brochado et al., 2018). For drug codes see Table S4.

drug1	acts on	drug2	type	class1	class2
SPT	attenuates	AMK	antagonism	aminoglycoside	aminoglycoside
SPT	attenuates	CPO	antagonism	aminoglycoside	antifungal
SPT	attenuates	CTX	antagonism	aminoglycoside	beta-lactam
SPT	attenuates	NVB	antagonism	aminoglycoside	DNA gyrase
SPT	attenuates	CHG	antagonism	aminoglycoside	LPS
SPT	attenuates	BLM	antagonism	aminoglycoside	multiple
SPT	attenuates	NIT	antagonism	aminoglycoside	other DNA
OXA	potentiates	TOB	synergy	beta-lactam	aminoglycoside
LEX	attenuates	PIP	antagonism	beta-lactam	beta-lactam
LEX	attenuates	CTX	antagonism	beta-lactam	beta-lactam
LEX	attenuates	CFS	antagonism	beta-lactam	beta-lactam
CTX	potentiates	MEM	synergy	beta-lactam	beta-lactam
PEN	potentiates	PIP	synergy	beta-lactam	beta-lactam
PEN	potentiates	MEC	synergy	beta-lactam	beta-lactam
LEX	potentiates	CHL	synergy	beta-lactam	chloramphenicol
PEN	attenuates	CIP	antagonism	beta-lactam	DNA gyrase
PIP	attenuates	MXF	antagonism	beta-lactam	DNA gyrase
LEX	attenuates	CIP	antagonism	beta-lactam	DNA gyrase
CFS	potentiates	NVB	synergy	beta-lactam	DNA gyrase
OXA	attenuates	BZK	antagonism	beta-lactam	human drug
PEN	attenuates	BZK	antagonism	beta-lactam	human drug
OXA	attenuates	PMB	antagonism	beta-lactam	LPS
AMX	attenuates	CST	antagonism	beta-lactam	LPS
AMX	attenuates	CHG	antagonism	beta-lactam	LPS
OXA	attenuates	CHG	antagonism	beta-lactam	LPS
PIP	potentiates	CHIR90	synergy	beta-lactam	LPS
CFS	attenuates	A22	antagonism	beta-lactam	other cell wall
PEN	potentiates	A22	synergy	beta-lactam	other cell wall
OXA	potentiates	A22	synergy	beta-lactam	other cell wall
CTX	attenuates	RIF	antagonism	beta-lactam	RNA polymerase
CHL	attenuates	TOB	antagonism	chloramphenicol	aminoglycoside
CHL	attenuates	IPM	antagonism	chloramphenicol	beta-lactam
CHL	attenuates	CST	antagonism	chloramphenicol	LPS
NVB	attenuates	CHIR90	antagonism	DNA gyrase	LPS
EGCG	potentiates	CPO	synergy	fatty acid biosynthesis	antifungal

EGCG	attenuates	MEC	antagonism	fatty acid biosynthesis	beta-lactam
CER	potentiates	PIP	synergy	fatty acid biosynthesis	beta-lactam
CER	attenuates	CHL	antagonism	fatty acid biosynthesis	chloramphenicol
CER	attenuates	NVB	antagonism	fatty acid biosynthesis	DNA gyrase
EGCG	attenuates	MXF	antagonism	fatty acid biosynthesis	DNA gyrase
EGCG	attenuates	NVB	antagonism	fatty acid biosynthesis	DNA gyrase
EGCG	attenuates	CIP	antagonism	fatty acid biosynthesis	DNA gyrase
CER	potentiates	STZ	synergy	fatty acid biosynthesis	human drug
CER	potentiates	PRC	synergy	fatty acid biosynthesis	human drug
CER	attenuates	CHIR90	antagonism	fatty acid biosynthesis	LPS
EGCG	attenuates	CHG	antagonism	fatty acid biosynthesis	LPS
TRI	potentiates	PMB	synergy	fatty acid biosynthesis	LPS
CER	attenuates	DOX	antagonism	fatty acid biosynthesis	tetracycline
CER	attenuates	MIN	antagonism	fatty acid biosynthesis	tetracycline
CER	attenuates	PUR	antagonism	fatty acid biosynthesis	tRNA
TMP	potentiates	AMK	synergy	folic acid biosynthesis	aminoglycoside
TMP	potentiates	CEC	synergy	folic acid biosynthesis	beta-lactam
TMP	attenuates	CIP	antagonism	folic acid biosynthesis	DNA gyrase
TMP	attenuates	BZK	antagonism	folic acid biosynthesis	human drug
TMP	potentiates	CLI	synergy	folic acid biosynthesis	lincosamide
TMP	attenuates	CHIR90	antagonism	folic acid biosynthesis	LPS
TMP	potentiates	PMB	synergy	folic acid biosynthesis	LPS
SMM	potentiates	BLM	synergy	folic acid biosynthesis	multiple
BBR	attenuates	AMK	antagonism	food additive	aminoglycoside
TPH	attenuates	AMK	antagonism	food additive	aminoglycoside
VNL	attenuates	AMX	antagonism	food additive	beta-lactam
CAF	attenuates	AMX	antagonism	food additive	beta-lactam
CAF	attenuates	CTX	antagonism	food additive	beta-lactam
CAF	attenuates	CEC	antagonism	food additive	beta-lactam
CCM	attenuates	MEC	antagonism	food additive	beta-lactam
CCM	attenuates	AMX	antagonism	food additive	beta-lactam
VNL	attenuates	CTX	antagonism	food additive	beta-lactam
VNL	attenuates	CEC	antagonism	food additive	beta-lactam
VNL	attenuates	ATM	antagonism	food additive	beta-lactam
VNL	attenuates	MEC	antagonism	food additive	beta-lactam
BBR	attenuates	CFS	antagonism	food additive	beta-lactam
TPH	attenuates	IPM	antagonism	food additive	beta-lactam
VNL	attenuates	CHL	antagonism	food additive	chloramphenicol
CAF	attenuates	CHL	antagonism	food additive	chloramphenicol

CAF	attenuates	CIP	antagonism	food additive	DNA gyrase
VNL	attenuates	LCX	antagonism	food additive	DNA gyrase
CCM	attenuates	LCX	antagonism	food additive	DNA gyrase
VNL	attenuates	CIP	antagonism	food additive	DNA gyrase
CCM	attenuates	NVB	antagonism	food additive	DNA gyrase
VNL	attenuates	NVB	antagonism	food additive	DNA gyrase
CAF	attenuates	NVB	antagonism	food additive	DNA gyrase
BBR	attenuates	NVB	antagonism	food additive	DNA gyrase
BBR	attenuates	CIP	antagonism	food additive	DNA gyrase
BBR	attenuates	LCX	antagonism	food additive	DNA gyrase
CCM	attenuates	TMP	antagonism	food additive	folic acid biosynthesis
CCM	attenuates	BZK	antagonism	food additive	human drug
BBR	potentiates	PRC	synergy	food additive	human drug
CCM	potentiates	PRC	synergy	food additive	human drug
VNL	potentiates	STZ	synergy	food additive	human drug
BBR	attenuates	CHG	antagonism	food additive	LPS
CCM	potentiates	PMB	synergy	food additive	LPS
CAF	potentiates	CHG	synergy	food additive	LPS
BBR	attenuates	CLR	antagonism	food additive	macrolide
CCM	attenuates	A22	antagonism	food additive	other cell wall
BBR	attenuates	FOF	antagonism	food additive	other cell wall
VNL	attenuates	A22	antagonism	food additive	other cell wall
CCM	attenuates	FUS	antagonism	food additive	other protein synthesis
VNL	attenuates	FUS	antagonism	food additive	other protein synthesis
CCM	potentiates	RIF	synergy	food additive	RNA polymerase
VNL	potentiates	RIF	synergy	food additive	RNA polymerase
CAF	attenuates	DOX	antagonism	food additive	tetracycline
VNL	attenuates	DOX	antagonism	food additive	tetracycline
TPH	attenuates	MIN	antagonism	food additive	tetracycline
CCM	potentiates	MIN	synergy	food additive	tetracycline
CAF	attenuates	PUR	antagonism	food additive	tRNA
VNL	attenuates	PUR	antagonism	food additive	tRNA
MTF	attenuates	AMK	antagonism	human drug	aminoglycoside
VPM	attenuates	GEN	antagonism	human drug	aminoglycoside
LOP	potentiates	GEN	synergy	human drug	aminoglycoside
PRC	attenuates	CEC	antagonism	human drug	beta-lactam
MTF	attenuates	CEC	antagonism	human drug	beta-lactam
LOP	attenuates	CFS	antagonism	human drug	beta-lactam
LOP	attenuates	IPM	antagonism	human drug	beta-lactam

VPM	potentiates	MEC	synergy	human drug	beta-lactam
RSP	potentiates	CEC	synergy	human drug	beta-lactam
STZ	potentiates	CHL	synergy	human drug	chloramphenicol
BZK	attenuates	CIP	antagonism	human drug	DNA gyrase
LOP	attenuates	NVB	antagonism	human drug	DNA gyrase
RSP	attenuates	NVB	antagonism	human drug	DNA gyrase
LOP	attenuates	CIP	antagonism	human drug	DNA gyrase
DIC	potentiates	CIP	synergy	human drug	DNA gyrase
DIC	attenuates	BZK	antagonism	human drug	human drug
RSP	potentiates	CHG	synergy	human drug	LPS
PRC	potentiates	PMB	synergy	human drug	LPS
PNF	attenuates	ERI	antagonism	human drug	macrolide
LOP	attenuates	BLM	antagonism	human drug	multiple
MTF	attenuates	BLM	antagonism	human drug	multiple
PRC	potentiates	BLM	synergy	human drug	multiple
MTF	attenuates	A22	antagonism	human drug	other cell wall
LOP	attenuates	A22	antagonism	human drug	other cell wall
RSP	potentiates	FOF	synergy	human drug	other cell wall
VPM	potentiates	MMC	synergy	human drug	other DNA
LOP	potentiates	MIN	synergy	human drug	tetracycline
LOP	potentiates	DOX	synergy	human drug	tetracycline
CLI	attenuates	TOB	antagonism	lincosamide	aminoglycoside
CLI	attenuates	ATM	antagonism	lincosamide	beta-lactam
CLI	potentiates	CHIR90	synergy	lincosamide	LPS
CLI	attenuates	BLM	antagonism	lincosamide	multiple
CST	attenuates	MEC	antagonism	LPS	beta-lactam
PMB	attenuates	MEC	antagonism	LPS	beta-lactam
PMB	attenuates	CHIR90	antagonism	LPS	LPS
PMB	attenuates	PLM	antagonism	LPS	other DNA
SPM	attenuates	AMK	antagonism	macrolide	aminoglycoside
SPM	attenuates	TOB	antagonism	macrolide	aminoglycoside
ERI	attenuates	AMK	antagonism	macrolide	aminoglycoside
CLR	attenuates	AMK	antagonism	macrolide	aminoglycoside
CLR	attenuates	TOB	antagonism	macrolide	aminoglycoside
SPM	attenuates	CEC	antagonism	macrolide	beta-lactam
ERI	attenuates	IPM	antagonism	macrolide	beta-lactam
AZM	attenuates	CFS	antagonism	macrolide	beta-lactam
CLR	attenuates	MXF	antagonism	macrolide	DNA gyrase
SPM	attenuates	CIP	antagonism	macrolide	DNA gyrase

CLR	potentiates	CHIR90	synergy	macrolide	LPS
SPM	attenuates	CLR	antagonism	macrolide	macrolide
ERI	attenuates	BLM	antagonism	macrolide	multiple
CLR	attenuates	BLM	antagonism	macrolide	multiple
CLR	attenuates	PLM	antagonism	macrolide	other DNA
CLR	attenuates	MMC	antagonism	macrolide	other DNA
DCS	attenuates	TOB	antagonism	other cell wall	aminoglycoside
DAP	attenuates	AMK	antagonism	other cell wall	aminoglycoside
NSN	attenuates	AMK	antagonism	other cell wall	aminoglycoside
BAC	potentiates	TOB	synergy	other cell wall	aminoglycoside
FOF	attenuates	CPO	antagonism	other cell wall	antifungal
DCS	attenuates	CTX	antagonism	other cell wall	beta-lactam
DAP	attenuates	CFS	antagonism	other cell wall	beta-lactam
BAC	potentiates	CEC	synergy	other cell wall	beta-lactam
BAC	potentiates	CTX	synergy	other cell wall	beta-lactam
FOF	potentiates	AMX	synergy	other cell wall	beta-lactam
BAC	potentiates	IPM	synergy	other cell wall	beta-lactam
DCS	attenuates	MXF	antagonism	other cell wall	DNA gyrase
DCS	attenuates	BZK	antagonism	other cell wall	human drug
FOF	attenuates	CHIR90	antagonism	other cell wall	LPS
DCS	attenuates	CHG	antagonism	other cell wall	LPS
A22	attenuates	MIN	antagonism	other cell wall	tetracycline
NSN	attenuates	MIN	antagonism	other cell wall	tetracycline
MMC	attenuates	GEN	antagonism	other DNA	aminoglycoside
MMC	attenuates	IPM	antagonism	other DNA	beta-lactam
MTR	attenuates	CTX	antagonism	other DNA	beta-lactam
MTR	attenuates	CFS	antagonism	other DNA	beta-lactam
PLM	attenuates	CFS	antagonism	other DNA	beta-lactam
DXR	potentiates	ATM	synergy	other DNA	beta-lactam
MTR	potentiates	ATM	synergy	other DNA	beta-lactam
NIT	attenuates	NVB	antagonism	other DNA	DNA gyrase
MMC	attenuates	LCX	antagonism	other DNA	DNA gyrase
DXR	potentiates	MXF	synergy	other DNA	DNA gyrase
NIT	attenuates	CHIR90	antagonism	other DNA	LPS
NIT	potentiates	BLM	synergy	other DNA	multiple
NIT	attenuates	A22	antagonism	other DNA	other cell wall
MTR	potentiates	MMC	synergy	other DNA	other DNA
DXR	potentiates	RIF	synergy	other DNA	RNA polymerase
FUS	potentiates	TOB	synergy	other protein synthesis	aminoglycoside

FUS	potentiates	CHL	synergy	other protein synthesis	chloramphenicol
FUS	attenuates	BZK	antagonism	other protein synthesis	human drug
FUS	potentiates	AZM	synergy	other protein synthesis	macrolide
LZD	attenuates	CFS	antagonism	oxazolidinone	beta-lactam
LZD	attenuates	BZK	antagonism	oxazolidinone	human drug
PQ	potentiates	TOB	synergy	oxidative stress	aminoglycoside
PQ	attenuates	ATM	antagonism	oxidative stress	beta-lactam
PQ	attenuates	PIP	antagonism	oxidative stress	beta-lactam
PQ	attenuates	MEC	antagonism	oxidative stress	beta-lactam
PQ	potentiates	IPM	synergy	oxidative stress	beta-lactam
PQ	potentiates	CEC	synergy	oxidative stress	beta-lactam
PQ	attenuates	CHL	antagonism	oxidative stress	chloramphenicol
CLZ	attenuates	CHL	antagonism	oxidative stress	chloramphenicol
PQ	attenuates	CIP	antagonism	oxidative stress	DNA gyrase
PQ	attenuates	NVB	antagonism	oxidative stress	DNA gyrase
PQ	potentiates	STZ	synergy	oxidative stress	human drug
PQ	attenuates	CHIR90	antagonism	oxidative stress	LPS
PQ	attenuates	AZM	antagonism	oxidative stress	macrolide
PQ	potentiates	DCS	synergy	oxidative stress	other cell wall
PQ	attenuates	MMC	antagonism	oxidative stress	other DNA
PQ	attenuates	PLM	antagonism	oxidative stress	other DNA
PQ	attenuates	FUS	antagonism	oxidative stress	other protein synthesis
PQ	attenuates	DOX	antagonism	oxidative stress	tetracycline
CLZ	attenuates	DOX	antagonism	oxidative stress	tetracycline
PQ	attenuates	PUR	antagonism	oxidative stress	tRNA
CCCP	potentiates	MEC	synergy	PMF	beta-lactam
CCCP	potentiates	CHL	synergy	PMF	chloramphenicol
CCCP	potentiates	CLR	synergy	PMF	macrolide
CCCP	potentiates	RIF	synergy	PMF	RNA polymerase
CCCP	potentiates	DOX	synergy	PMF	tetracycline
RIF	attenuates	IPM	antagonism	RNA polymerase	beta-lactam
TGC	attenuates	TOB	antagonism	tetracycline	aminoglycoside
TGC	attenuates	AMK	antagonism	tetracycline	aminoglycoside
DOX	attenuates	GEN	antagonism	tetracycline	aminoglycoside
DOX	attenuates	STZ	antagonism	tetracycline	human drug
TGC	potentiates	FOF	synergy	tetracycline	other cell wall
PUR	attenuates	CEC	antagonism	tRNA	beta-lactam
PUR	attenuates	RIF	antagonism	tRNA	RNA polymerase