Multi-task learning predicts drug combination synergy in cells and in the clinic

Coryandar Gilvary[1,2,3,4], Jonathan R Dry[5], Olivier Elemento[1, 2,3,4,6]

[1] HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Dept. of Physiology and Biophysics, Weill Cornell Medicine, New York, NY 10065, USA;
[2] Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY 10065, USA;
[3] Sandra and Edward Meyer Cancer Center, Weill Cornell Medicine, New York, NY 10065, USA;
[4] Tri-Institutional Training Program in Computational Biology and Medicine, New York, NY 10065, USA;
[5] Oncology, IMED Biotech Unit, AstraZeneca R&D Boston, Waltham MA, USA
[6] WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY 10065, USA

Correspondence: Olivier Elemento (ole2001@med.cornell.edu)

ABSTRACT

Combination therapies for various cancers have been shown to increase efficacy, lower toxicity, and circumvent resistance. However, despite the promise of combinatorial therapies, the biological mechanisms behind drug synergy have not been fully characterized, and the systematic testing of all possible synergistic therapies is experimentally infeasible due to the sheer volume of potential combinations. Here we apply a novel big data approach in the evaluation and prediction of drug synergy by using the recently released NCI-ALMANAC. We found that each traditional drug synergy metric (Bliss, Loewe, ZIP, HSA, ALMANAC Score) identified unique synergistic drug pairs with distinct underlying joint mechanisms of action. Leveraging these findings, we developed a suite of context specific drug synergy predictive models for each distinct synergy type and achieved significant predictive performance (AUC = 0.89-0.953). Furthermore, our models accurately identified clinically tested drug pairs and characterized the clinical relevance of each drug synergy metric, with Bliss Independence capturing clinically tested combinations best. Our findings demonstrate that drug synergy can be obtained from numerous unique joint mechanisms of action, captured by different synergy metrics. Additionally, we show that drug synergy, of all kinds, can be predicted with high degrees of accuracy with significant clinical potential. This breakthrough understanding of joint mechanisms of action will allow for the design of rational combinatorial therapeutics on a large scale, across various cancer types.

1

47 **INTRODUCTION**
48 Precision medicine has been instrumental in the creation of novel cancer therapeutics, however
49 many of these promising treatments have been limited by insufficient efficacy or acquired
50 resistance. Cancer cells have exhibited multiple types of resistance mechanisms, such as
51 redundant pathways[1] and pathway reactivation[2], among others[3-5]. Combination therapies have
52 been proposed as a general strategy to combat therapeutic resistance, as well as increase overall
53 efficacy[6]. Such is the case with the combination of MEK and ERK inhibition to overcome
54 MAPK pathway reactivation[7]. Overall, rational drug combination therapies have the potential to
55 create long-lasting therapeutic strategies against cancer and perhaps many other diseases.
56
57 While combination therapies have shown great potential, identifying effective drug pairings
58 presents a difficult challenge. Traditional methods for assessing drug efficacy of single drug
59 agents are not easily applied. With the number of approved or investigational drugs increasing
60 every year, the ability to pairwise test these agents across a wide breadth of disease models
61 becomes virtually impossible. This experimental infeasibility leads to the need for a
62 computational approach to predict drug synergy.
63
64 Quantitative models have been introduced to predict effective drug combinations, however they
65 tend to be limited in scope, either confined to certain drug or cancer types[8]. A recent DREAM
66 Challenge (dreamchallenges.org)[9] teamed up with AstraZeneca and called for models to predict
67 synergy across diverse drugs and cancer types[10, 11]. Despite >80 distinct computational methods
68 being submitted to harness biological knowledge and classify cells and drugs, with performance
69 matching the accuracy of biological replicates for many cases, there were still numerous drug
70 combinations that consistently performed poorly across all models.
71
72 The problem is confounded as there currently is no agreed upon gold standard to measure drug
73 synergy of clinical relevance *in vitro*, introducing variability across method development and
74 misclassification of valuable drug combinations. There are many different metrics used to
75 measure drug synergy that are all based on models with differing underlying assumptions.
76 Fourcquier and Guedi have presented a comprehensive and succinct description of many of these
77 models, how they differ and their practical limitations[12]. The most widely used of these methods
78 include Bliss Independence[13, 14], which assumes no interference/interaction between drugs, and
79 Loewe Additivity[14, 15], which is based on the dose equivalence principle. However, these models,
80 require large amounts of experimental data. Therefore, simpler models such as Highest Single
81 Agent (HSA), which calculates synergy as the excess over the maximum single agent response,
82 can become attractive options[14]. Newer methods such as the Zero Interaction Potency (ZIP)
83 model, a combination of Bliss and Loewe[16], have aimed to overcome the limitations of these past
84 models. Overall, each metric captures and measures different aspects of synergistic action.
85
86 The release of the NCI-ALMANAC[17], a publically available database of combination drug
87 efficacies, has presented a unique opportunity to develop models to predict drug synergy on a
88 large scale. Here, we describe our analysis of the mechanisms behind diverse types of drug
89 synergy and leverage this information to create a suite of machine learning models to predict
90 novel drug combinations and synergy scores from five different metrics, Bliss, Loewe, HSA, ZIP
91 and the ALMANAC Score (a modification of the Bliss model). Using this approach, we
92 identified the mechanisms and characteristics of the drug synergy identified by each metric and

93   highlight how we could efficiently pinpoint clinically relevant drug combinations. Overall this
94   approach has potential to accelerate the preclinical pipeline for pairwise drug combination
95   therapeutics and provides a conceptual framework for understanding the mechanistic basis of
96   drug synergy.
97
98   **RESULTS**
99   **Combination Efficacy Measures are Poorly Correlated**
100
101   Drug combination efficacy metrics have been crucial in the development of synergistic drug
102   treatments, however there has yet to be an in-depth evaluation of these metrics across diverse
103   drug and cancer types. The diversity of the publically available data set, NCI-ALMANAC,
104   provided the opportunity for us to test the different combination efficacy measures across various
105   cancer types and drug classes. We first interrogated the diversity of the NCI-ALMANAC dataset
106   and found that they tested over 100 drugs, which represented 12 distinct drug classes (**Figure
107   1A**), in pairwise drug efficacy screenings on 60 cell lines. The 60 cancer cell lines tested cover 9
108   diverse cancer types (**Figure 1B**). In addition to the experimental data, the NCI-ALMANAC
109   formulated their own drug synergy score (we will refer to this metric as the ALMANAC score)
110   and have released this score for each pairwise combination[17].
111
112   The disparity and diversity among combination efficacy metrics has long been remarked on[12]
113   and we verified this quantitatively by looking at the correlation between each synergy metric on
114   the NCI-ALMANAC drug pairs for each cell line (**Figure 1C, Supp Fig 1**). We found the
115   correlation between synergy measures using both Pearson and Spearmen, with the coefficients
116   ranging from 0.21-0.66 ($p < 0.001$) and 0.18-0.84 ($p < 0.001$), respectively. Only HSA and
117   Loewe had a Pearson and Spearman correlation above 0.5 (**Methods**). Importantly, we also
118   found that the majority of both the synergistic and antagonistic qualifications assigned to drug
119   pairs were unique to efficacy metrics (**Figure 1D**), even in the case of measures that were highly
120   correlated (proportion overlap 0.40 and 0.35 for HSA and Loewe, HSA and ZIP, respectively).
121   These results illustrate the distinctness between combination efficacy metrics across many drug
122   and cancer types.
123
124   **Synergistic Drug Pairs Share Distinct Attributes**
125
126   Drug synergy can arise due to a variety of diverse mechanisms which may present as distinct
127   patterns in *in vitro* assays[18, 19]. Due to the large discrepancy between combination efficacy
128   metrics we reasoned that each metric may be identifying different types of synergistic
129   combinations. Therefore, we looked to quantify which drug attributes were shared among all
130   synergistic drug pairs and which were metric specific. Since we have previously found drug
131   structure to effect pharmacological attributes such as toxicity and molecular targets[20-22], we
132   investigated if structure based similarity between drug pairs was indicative of a pair being
133   synergistic. Using chemical fingerprint similarity, we found that synergistic pairs were more
134   similar to each other than antagonistic and other non-synergistic drugs, across all metrics (KS
135   test, D-stat 0.23-0.307, p-value < 0.001, **Figure 2A**). Antagonistic and uncategorized, drug pairs
136   were associated with undistinguishable levels of structure similarity (KS Test, p-value > 0.05).
137   When we evaluated additional structure comparison measures, such as the similarity of hydrogen
138   atoms and bonds our findings remained consistent (**Supp Fig 2**), we found that regardless of

139 combination efficacy metric, drug synergy scores increased directly with drug structure
140 similarity. These results run counter to expectations that synergistic drugs target distinct
141 pathways.
142
143 We further investigated drug attributes that could characterize synergistic drug pairs and focused
144 on molecular targets of these drugs, which has also been shown to effect numerous other
145 pharmacological attributes in past research[20, 22]. We found that drug pairs sharing a higher
146 number of the same targets were not more synergistic or antagonistic than expected by chance
147 (KS test, D-stat < 0.1, p-value >0.05), for any metric (**Supp Fig 3A**). However, we found that
148 synergistic drugs tended to be found in the same pathway more than antagonistic drugs, as per
149 KEGG pathways, Reactome pathways or molecular function gene ontology (**Figure 2B, Supp**
150 **Fig 3B**). However, since each drug synergy metric is based on different underlying principles,
151 we hypothesized that there would be a large variation in how strongly drug synergy scores are
152 influenced by pathway similarity. We found that the Bliss model (which assumes independence
153 between drugs) tended to induce the lowest pathway similarity difference between synergistic
154 (strong pathway similarity) and antagonistic (low pathway similarity) drug pairs (KS test, D =
155 0.171, p <0.001, **Supp Fig 3B**). However, in a model that assumes a more additive effect, such
156 as Loewe, the separation greatly increases (KS test, D = 0.307, p <0.001, **Supp Fig 3B**).
157 Therefore, drug synergy metrics which incorporate drug pair interactions may be more strongly
158 affected when drugs target similar pathways.
159
160 **Synergy Metrics Identify Unique Synergistic Pathway Combinations**
161
162 Drug combination efficacy metrics principally vary in their intrinsic assumptions about drug
163 synergy. Previous work in cancer drug combinations has demonstrated that drug synergy can be
164 achieved through a variety of pathway mechanisms[10]; therefore each metric is most likely
165 identifying distinct pathway combinations. Using the KEGG database to identify pathways based
166 on the molecular targets of each drug, we evaluated whether the combination of targeting two
167 specific pathways was consistently synergistic or antagonistic for each metric. We specifically
168 evaluated the pathway combinations which were most variable among metrics (ie were
169 significantly enriched for synergy using some metrics and a loss of significance in other metrics)
170 (**Figure 2C, Methods**). With the identification of these top differential pathway combinations
171 we investigated the potential causes for the variability between metrics. Pathway combinations
172 such as 'Notch Signaling' with 'One Carbon Pool by Folate' (OR = 2.29, p <0.001) or 'Steroid
173 Hormone Biosynthesis' (Fisher's Exact OR = 2.63, p < 0.001) were significantly more likely to
174 be given high Bliss or ALMANAC scores. These pathways are distantly related, as in having no
175 proteins/compounds in common. Both Bliss Independence and the ALMANAC Score assume
176 little-to no interaction between drugs, therefore it would follow that these metrics are more likely
177 to capture drug combinations targeting distant pathways (**Supp Fig 4**). However, when
178 evaluating pathways that are more closely related (sharing proteins/compounds), such as 'Gap
179 Junction' with 'p53 Signaling' we found that the Bliss metric is likely to not find these drug
180 combinations synergistic (Fisher's Exact OR = 0.1326, p < 0.01). When evaluating drugs that
181 are specifically targeting the same pathway, in the case of the ErbB signaling pathway, the
182 Loewe score is significantly more likely to label these pairs as synergistic (Fisher's Exact OR =
183 2.61, p <0.001), whereas Bliss demonstrates the opposite pattern (Fisher's Exact OR = 0.686, p <
184 0.01). Past research has shown that dual targeting the ErbB pathway can be a promising clinical

4

185 treatments[23] within cancer. All of these pathway combinations highlight the different drug
186 combinations captured by the various combination efficacy metrics and are candidates for further
187 drug synergy development.
188
189 **Distinct Biological Mechanisms Lead to Differential Drug Synergy Scores**
190 To further characterize the mechanistic types of drug synergy each metric identifies we
191 categorized three types of pathway mechanisms at the network level based on Menden et al[10]
192 (**Methods**): same target, same pathway, and parallel pathway. We defined combinations with a
193 "same pathway" mechanism as those whose targets are up/downstream of each other and
194 "parallel pathway" mechanism combinations occurred when the drug targets were upregulated
195 or downregulated by the same gene. We found that the synergy score for all metrics increased
196 significantly for drug pairs that interacted in any of these pathway mechanisms by any
197 mechanism of action, compared to drug pairs that did not (Wilcox p-value <0.05, Figure 2D).
198 The greatest increase was seen in drug combinations with the same target, specifically when
199 using the ALMANAC score to measure synergy (Wilcox location shift = 5.27, p-value < 0.001,
200 Figure 2D). These results highlight the importance of having any type of interaction, supported
201 by our finding that synergistic drug combinations for all metrics had lower distance between
202 drug targets in a gene network when compared to antagonistic drug pairs (**Supp Fig 3C**), most
203 significantly HSA (D = 0.203, p < 0.001).
204
205 To better understand the pathway mechanisms of drug synergy, we assessed if both drugs were
206 inhibiting, activating or a mix of both (one inhibiting and one activating), using mechanism of
207 action data available for each drug on DrugBank[24]. Using our previously defined mechanisms
208 (same target, same pathway, parallel pathway), we first evaluated the 'same target' mechanism
209 and only looked at drug combinations that had at least one drug target in common. When at least
210 one of the drugs was activating its gene target, the ALMANAC Score was significantly higher
211 (Both drugs activate: Wilcox location shift = 5.27, p-value < 0.001; One drug activates: Wilcox
212 location shift = 4.32, p-value < 0.001, **Figure 2D, Supp Fig 5**). However, there was no
213 significant shift in synergy score if both drugs were inhibiting the same gene (Wilcox location
214 shift = 0.388, p-value =0.519). Therefore, the ALMANAC score shows no enrichment in finding
215 combinations of dual inhibitors. All other combination efficacy metrics retained a statistical
216 significant increase in synergy scores when both drugs inhibited the same target. Dual inhibitors
217 have been shown to be successful combination treatments within clinical trials[25, 26], which might
218 be missed if only using the ALMANAC scoring metric.
219
220 Besides targeting the same gene target, synergistic drug pairs will often target the same or
221 parallel pathways, therefore we wanted to quantify the effect on these pathway mechanisms as
222 well. Again when evaluating the ALMANAC score, synergy scores were significantly increased
223 when drug pairs were interacting in these pathway mechanisms, specifically when at least one
224 drug was an inhibitor and both drugs were targeting the same pathway (Both drugs inhibiting:
225 Wilcox location shift = 5.79 , p-value <0.001, One drug inhibiting: Wilcox location shift = 4.47,
226 p-value <0.001, **Supp Fig 6,7**) and parallel pathways (Wilcox location shift = 1.72, p-value =
227 0.002, One drug inhibiting: Wilcox location shift = 1.92, p-value <0.001, **Supp Fig 6,7**).
228 However, when both drugs activate the same/parallel pathway ALMANAC scores were
229 significantly lower than drug pairs not interacting in that manner (Same Pathway: Wilcox
230 location shift= -0.971, p-value = 0.00907, Parallel Pathway: Wilcox location shift = -2.61, p-

231 value < 0.001, **Supp Fig 6,7**). We found that synergistic scores were significantly higher for
232 drugs pairs that activated the same or parallel pathways in only two metrics, Loewe (Same
233 Pathway: Wilcox location shift= 1.66, p-value < 0.001, Parallel Pathway: Wilcox location shift =
234 1.47, p-value < 0.001, **Supp Fig 6,7**) and HSA (Same Pathway: Wilcox location shift= 0.48, p-
235 value < 0.001, Parallel Pathway: Wilcox location shift = 0.76, p-value < 0.001, **Supp Fig 6,7**)
236 (**Figure 2D**). Drugs may interact in unexpected ways when they activate the same or parallel
237 pathways, which is poorly modeled by the ALMANAC and Bliss metric. However, the Loewe
238 metric models additive effects of drug pairs and could better reflect cases where drugs interact.
239 These underlying model assumptions may allow for Loewe to identify dual pathway activating
240 drug combinations that may be missed by other common methods.
241
242 **Suite of Metric Specific Models Predicts Drug Synergy**
243
244 Since the synergistic drug combinations showed distinct characteristics when compared to
245 antagonistic or other drug pairs, we reasoned that using a computational approach we could build
246 a classification model to predict drug synergy or antagonism based on the similarity of various
247 pharmacological and genomic attributes. Due to the diverse nature of each combination efficacy
248 metric we chose to build a set of classification models, each fit with the synergistic/antagonism
249 labels found using a specific metric, to create a model toolbox. Additionally, to account for the
250 cell line specificity of drug synergy noted in past research[17] and found within our own data
251 (**Figure 1B**), we used a multi-task learning approach, which utilizes the strength of transfer
252 learning[27] while accounting for differences in synergy mechanisms between cell lines/cancer
253 types.
254
255 Our approach, a multi-task learning extreme randomized tree algorithm[28], was tested with 10
256 fold cross validation using the NCI-ALMANAC data and collected drug similarity features
257 (**Methods**). Each model, specific to one metric, had significant predictive power, with the area-
258 under-the –receiver-operator curve (AUROC) ranging from 0.89-0.95 (**Figure 3A**) and the area-
259 under-the-precision-recall curve (AUPRC) ranging from 0.51-0.71 (**Figure 3B**). These strong
260 AUPRCs demonstrate a fair trade-off between false positives and false negatives within our data
261 (**Supp Fig 8**). This is one of the highest reported accuracy measures for drug synergy prediction
262 models based on any of these combination efficacy metrics[10, 29].
263
264 To ensure our results were transferable to external data sets we investigated the AstraZeneca –
265 DREAM Challenge drug synergy publically available dataset[10, 11], which also contained a variety
266 of cancer and drug types. There were 19 new drug pairs tested in eight of the cancer cell lines our
267 model is trained for available in the DREAM challenge 1 training data. To ensure we were
268 appropriately comparing results we calculated four of the different combination efficacy metrics
269 (HSA, ZIP, Loewe, Bliss) from the experimental DREAM Challenge data. The metric used
270 within the original paper was most highly correlated with the HSA metric (cor = 0.961, p <
271 0.001, **Methods, Supp Fig 9**) and therefore we used the HSA metric to compare the predicted
272 results. We predicted the drug synergy for all DREAM challenge drug pairs using each model.
273 We found that the predictions were significantly correlated to the calculated AZ-DREAM score
274 (Pearson's Coefficient 0.313, p-value = 0.0493) used in the original research and the calculated
275 HSA (Pearson's Coefficient 0.363, p-value = 0.0213; **Figure 3C**). Although the other metrics do
276 not show a statistical significance between calculated and predicted score, this was expected.

277   This experiment was designed with the specific combination AZ-DREAM metric in mind and
278   therefore this was the cleanest data available and most accurate out-of-sample test set. This
279   external independent test set validates the use of our models on different drugs pairs in
280   independent experimental settings.
281
282   **Clinical Significance Differs based on Combination Efficacy Metric**
283
284   We further assessed our models by applying them to clinically tested pairwise drug combinations
285   found using the Drug Combination Data Base[30]. This database contains drug combinations for a
286   multitude of diseases at different points in the drug development phase, ranging from preclinical
287   to approved. We have limited our scope to drugs that have entered clinical trials for cancer. This
288   subset ensured a fair tradeoff between translatable relevancy and a large enough sample size.
289   Across all combination efficacy metrics our model assigned a significantly higher synergy score
290   to clinically tested drug pairs than randomly selected combinations (D = 0.242-0.297, p < 0.001)
291   (**Figure 4A**). While all metrics showed statistical significance, the ALMANAC Score and Bliss
292   metric demonstrated the best ability to distinguish clinically tested and randomly selected drug
293   pairs. This may be due to the popularity of Bliss scores in research, leading to more Bliss
294   synergistic pairs making their way to clinical trials. However, this could be due to the pathway
295   mechanisms enriched within Bliss/ALMANAC score synergistic pairs are more clinically
296   impactful. For example, the targeting of distinct pathways, which the Bliss metric favors, may
297   lead to lower toxicity, which can make those synergistic combinations a clinically viable
298   option[31].
299
300   In addition to accurately predicting synergistic combinations, we wanted to test if our models,
301   which predict cell-line specific drug synergy, are capable of properly distinguishing which
302   cancer type a combination will be clinically effective within. Across all metrics, predicted
303   synergy scores for clinically tested drug combinations were significantly higher in the cancer cell
304   lines matched with the clinically tested cancer type when compared to randomly selected cell
305   lines (p-value < 0.05, **Figure 4b**). The cancer type specificity was comparable for all metric
306   types, demonstrating the robustness of our models and their potential clinical impact.
307
308   <u>**DISCUSSION**</u>
309   Identifying rational, synergistic drug combinations has great potential to increase efficacy of
310   cancer treatments as well as combat therapeutic resistance, however experimental approaches to
311   pairwise test drug combinations are costly in both time and money. Additionally, there lacks a
312   true gold standard to measure drug synergy, due to the numerous different ways synergistic
313   action can be achieved between two drugs. We have proposed a suite of models to predict drug
314   synergy using the top drug combination efficacy metrics currently available, in addition to
315   identifying the specific types of synergy each metric is tuned to identify. When trained on NCI-
316   ALMANAC experimental combination data we achieved significant predictive power across all
317   metrics and showed the ability for these models to be applied to drug combination screenings
318   performed under different conditions, using different drugs and cancer cell lines. Additionally,
319   we have applied these models to drugs in the clinic and accurately distinguished between
320   clinically tested drug combinations and random pairs. Our metric specific suite of models also
321   enables us to identify which efficacy metrics have been most clinical successful and can inform
322   future experimental approaches.

7

323
324 Drug combinations have long been thought to be the answer to patient resistance and increasing
325 drug efficacy, with notable cases of success[6]. Researchers have realized the impossibility of
326 experimentally testing every pairwise drug combinations and therefore many computational
327 models have been created to answer this very issue[10, 11]. However, many current models were
328 created for either specific drug types or cancer types[8], which inherently limits the applicability of
329 these models. Recently there was a DREAM Challenge, partnered with AstraZeneca, to predict
330 drug synergy based on diverse drugs and cancer types[10, 11], however this challenge focused on
331 predicting drug synergy, calculated by only one metric. The metrics used to calculate drug
332 synergy have been shown to be discordant[12] and vary in many of the publically available dataset
333 and the work flows of pharmacological labs or companies. We focused our attention on
334 predicting a bevy of drug synergy types and furthering the understanding of the nuances between
335 drug synergy types.
336
337 While current large-scale experimental approaches have focused on pairwise drug combinations,
338 many drug cocktails (> two drugs) have shown to be promising as therapeutics[32, 33]. Many of the
339 metrics we have discussed can be applied to and measure the drug synergy of these drug
340 cocktails[14, 34]. Since our predictive models are based on drug similarity, this can be easily
341 extended to include more than only two drugs. As experimental methods become more
342 streamlined and less costly, we believe we could extend these models to the prediction of drug
343 synergy for drug cocktails.
344
345 However, drug combination efficacy among cell lines is only one piece of the necessary puzzle
346 to developing synergistic drug combinations. We have not addressed the ongoing issue of
347 toxicity as it pertains to drug combinations, which has led to promising candidates failing in
348 clinical settings[35]. Additionally, our model is not dose specific and therefore cannot be used to
349 identify optimal ratios or doses of drugs for treatments.
350
351 Overall, our suite of models has the potential to quickly predict synergist drug combinations to
352 rapidly speed up the preclinical pipeline for drug combination treatments. Additional, we have
353 outlined the differences between drug synergy metrics in terms of the type of synergistic action
354 they identify and their clinical significance. We hope this can be used to help guide future
355 experimental work to improve the development of pairwise combination treatments for cancer
356 treatments.
357
358 **METHODS**
359
360 ***Drug Synergy Cell Line Data***
361 The drug pair synergy data was downloaded via NCI-ALMANAC[17] and refined to include drug
362 pairs with enough publically available data. In total 3647 unique drug pairs in 60 cell lines were
363 analyzed. Using the raw data provided in NCI-ALMANAC, the R package SynergyFinder[36]
364 Version 1.6.1 was used to calculate the Bliss, ZIP, HSA and Loewe synergy score for each drug
365 pair. We categorized drug pairs as "synergistic" for each metric if their scores were within the top
366 5% and "antagonistic" if the scores were in the bottom 66.67%.
367
368 ***Pathway Analysis***

8

369 All known drug targets were collected from Drugbank[24] and matched to KEGG pathways via the
370 KEGGREST[37] R package using a custom R script. A fisher's exact test was used to find the Odd's
371 Ratio for targeted pathway combination likely to be marked as synergistic for each metric. The
372 most variable pathway combinations were found by identifying all combinations that had at least
373 one synergy metric with an Odd's Ratio lower confidence interval above 1.5 and an Odd's Ratio
374 higher confidence interval lower than 1.
375
376 ***Feature Collection***
377 *Compound Features*
378 For the 3,647 drug pairs, multiple compound similarity features were collected. Additionally, using
379 their known drug targets as listed in DrugBank[24], we collected drug target similarity features as
380 well. The feature, source and metric used to measure similarity is listed in **Supplementary Table**
381 **1**. The measures of similarity included but were not limited to Pearson Correlation, Jaccard Index
382 and Dice Similarity. In cases where there was insufficient or missing information, features were
383 imputed by using the median value for that feature in drug pairs with complete information.
384
385 *Network Features*
386 We curated a biological network that contains 22,399 protein-coding genes, 6,679 drugs, and 170
387 TFs. The protein-protein interactions represent established interaction[38-40], which include both
388 physical (protein-protein) and non-physical (phosphorylation, metabolic, signaling, and
389 regulatory) interactions. The drug-protein interactions were curated from several drug target
390 databases[40].
391
392 ***Predictive Model Suite***
393 Our predictive models were trained as binary classifiers using the features described above on the
394 NCI ALMANAC data, with synergistic and antagonistic drug pairs being our respective classes.
395 Every model included the same features, however the classes were determined by one of the five
396 drug synergy measures (HSA, Bliss, Loewe, ZIP, ALMANAC Score). Mulit-task extremely
397 randomized tree models, a decision tree model, was used after model selection and implemented
398 using the R statistical software with the extraTrees package[28], the cancer cell line was used as each
399 task. To evaluate predictive power 10-fold cross validation was used for each model. Down
400 sampling was the chosen sub-sampling approach applied to each fold to account for the class
401 imbalance between synergistic and antagonistic drug pairs.
402
403 ***Classification Evaluation***
404 For evaluating all the binary synergy classifications, receiver operating characteristic (ROC) and
405 precision-recall curve (PRC) curves were created in R using the pROC[41] and precrec[42] packages
406 respectively. Area-under-the-ROC curve (AUC) and area-under-the-PRC (AUPRC) scores were
407 used to evaluate model performance.
408
409 ***DREAM Challenge Validation Data***
410 Raw dose-response data from the DREAM-AZ Combination Prediction Challenge[10] was used as
411 an external dataset to test our models. We found 19 drug pairs, unseen by the models, available in
412 the Challenge 1 data set tested within cell lines our models were trained on. For these 19 pairs
413 features were collected in the same manner as described above and drug synergy scores for all
414 metrics, besides the ALAMANAC score, were calculated as well. The correlation between all

9

415 synergy scores were found using Pearson correlations. The synergy scores were predicted using
416 each model and then a Pearson correlation to the calculated scores were measured. Since the
417 calculated HSA score was most significantly correlated with the given DREAM challenge score,
418 the predicted HSA scores were used in the comparison to the DREAM challenge scores.
419
420 ***Clinical Trial Evaluation***
421 Current combinatorial therapeutics in clinical trials were found using the Drug Combination Data
422 Base[30]. Combination therapies were narrowed down to only pairwise combinations of drugs with
423 sufficient information to run the model and combinations being tested to treat cancer, which
424 resulted in 300 unique drug combinations within clinical trials. We created a set of ~250,000
425 random combinations by shuffling the drugs and removing any pairs that were in clinical trials.
426 We then collected all drug compound features as described above and ran them through our
427 model with all cancer cell lines our models have been trained on. Once predictive synergy scores
428 for every metric in all cell lines were obtained we used the mean of scores for each cancer type
429 (based on primary site) to determine the final score of clinically tested combinations. First we
430 tested the difference between clinically tested drug pairs and random drug pairs using
431 Kolmogorov–Smirnov tests. We used a bootstrapping approach to sample the non-clinically
432 tested drug pairs to get a confidence interval of the true model score distribution. Additionally,
433 we used a Kolmogorov–Smirnov test to determine the statistical significance between the scores
434 of our clinically tested drug pairs within the cancer type that is being tested and the drug pairs
435 synergy scores within cancer types that have not been clinically tested.
436
437 ***Code Availability***
438 Our suite of combination models will be available for download upon request to help facilitate
439 immediate impact within the research community.
440

458
459
460

REFERENCES

1. Thomson, S., Petti, F., Sujka-Kwok, I., Epstein, D. & Haley, J.D. Kinase switching in mesenchymal-like non-small cell lung cancer lines contributes to EGFR inhibitor resistance through pathway redundancy. *Clinical & experimental metastasis* **25**, 843-854 (2008).

2. Johannessen, C.M. et al. COT drives resistance to RAF inhibition through MAP kinase pathway reactivation. *Nature* **468**, 968 (2010).

3. Garraway, L.A. & Jänne, P.A. Circumventing cancer drug resistance in the era of personalized medicine. *Cancer discovery* (2012).

4. Duran, I. et al. Resistance to targeted therapies in renal cancer: the importance of changing the mechanism of action. *Targeted oncology*, 1-17 (2017).

5. Chong, C.R. & Jänne, P.A. The quest to overcome resistance to EGFR-targeted therapies in cancer. *Nature medicine* **19**, 1389-1400 (2013).

6. Al-Lazikani, B., Banerji, U. & Workman, P. Combinatorial drug therapy for cancer in the post-genomic era. *Nature biotechnology* **30**, 679 (2012).

7. Merchant, M. et al. Combined MEK and ERK inhibition overcomes therapy-mediated pathway reactivation in RAS mutant tumors. *PloS one* **12**, e0185862 (2017).

8. Gayvert, K.M. et al. A Computational Approach for Identifying Synergistic Drug Combinations. *PLoS computational biology* **13**, e1005308 (2017).

9. Saez-Rodriguez, J. et al. Crowdsourcing biomedical research: leveraging communities as innovation engines. *Nature Reviews Genetics* **17**, 470 (2016).

10. Menden, M.P. et al. Community assessment of cancer drug combination screens identifies strategies for synergy prediction| bioRxiv. *bioRxiv* (2017).

11. Menden, M.P. et al. A cancer pharmacogenomic screen powering crowd-sourced advancement of drug combination prediction. *bioRxiv*, 200451 (2018).

12. Foucquier, J. & Guedj, M. Analysis of drug combinations: current methodological landscape. *Pharmacology research & perspectives* **3** (2015).

13. Bliss, C. The toxicity of poisons applied jointly 1. *Annals of applied biology* **26**, 585-615 (1939).

14. Berenbaum, M. A method for testing for synergy with any number of agents. *Journal of Infectious Diseases* **137**, 122-130 (1978).

15. Loewe, S. & Muischnek, H. Über kombinationswirkungen. *Naunyn-Schmiedebergs Archiv für experimentelle Pathologie und Pharmakologie* **114**, 313-326 (1926).

16. Yadav, B., Wennerberg, K., Aittokallio, T. & Tang, J. Searching for drug synergy in complex dose–response landscapes using an interaction potency model. *Computational and structural biotechnology journal* **13**, 504-513 (2015).

17. Holbeck, S.L. et al. The National Cancer Institute ALMANAC: A Comprehensive Screening Resource for the Detection of Anticancer Drug Pairs with Enhanced Therapeutic Activity. *Cancer Research*, canres. 0489.2017 (2017).

18. Bulusu, K.C. et al. Modelling of compound combination effects and applications to efficacy and toxicity: state-of-the-art, challenges and perspectives. *Drug Discovery Today* **21**, 225-238 (2016).

11

504 19. Dry, J.R., Yang, M. & Saez-Rodriguez, J. Looking beyond the cancer cell for effective drug
505   combinations. *Genome medicine* **8**, 125 (2016).
506 20. Madhukar, N.S., Gayvert, K., Gilvary, C. & Elemento, O. A Machine Learning Approach
507   Predicts Tissue-Specific Drug Adverse Events. *bioRxiv*, 288332 (2018).
508 21. Madhukar, N.S. et al. A new big-data paradigm for target identification and drug
509   discovery. *bioRxiv*, 134973 (2017).
510 22. Gayvert, K.M., Madhukar, N.S. & Elemento, O. A data-driven approach to predicting
511   successes and failures of clinical trials. *Cell chemical biology* **23**, 1294-1301 (2016).
512 23. Zhang, H. et al. ErbB receptors: from oncogenes to targeted cancer therapies. *The
513   Journal of clinical investigation* **117**, 2051-2058 (2007).
514 24. Law, V. et al. DrugBank 4.0: shedding new light on drug metabolism. *Nucleic acids
515   research* **42**, D1091-D1097 (2013).
516 25. Janjigian, Y.Y. et al. Dual inhibition of EGFR with afatinib and cetuximab in kinase
517   inhibitor-resistant EGFR-mutant lung cancer with and without T790M mutations. *Cancer
518   discovery* (2014).
519 26. Fan, Q.-W. et al. A dual PI3 kinase/mTOR inhibitor reveals emergent efficacy in glioma.
520   *Cancer cell* **9**, 341-349 (2006).
521 27. Pan, S.J. & Yang, Q. A survey on transfer learning. *IEEE Transactions on knowledge and
522   data engineering* **22**, 1345-1359 (2010).
523 28. Simm, J., de Abril, I.M. & Sugiyama, M. Tree-based ensemble multi-task learning method
524   for classification and regression. *IEICE TRANSACTIONS on Information and Systems* **97**,
525   1677-1681 (2014).
526 29. Aben, N. et al. Identifying biomarkers of anti-cancer drug synergy using multi-task
527   learning. *bioRxiv*, 243568 (2018).
528 30. Liu, Y., Hu, B., Fu, C. & Chen, X. DCDB: drug combination database. *Bioinformatics* **26**,
529   587-588 (2009).
530 31. Mokhtari, R.B. et al. Combination therapy in combating cancer. *Oncotarget* **8**, 38022
531   (2017).
532 32. Zhu, B. et al. Assessment of cytochrome P450 activity by a five-drug cocktail approach.
533   *Clinical Pharmacology & Therapeutics* **70**, 455-461 (2001).
534 33. Nussinov, R., Tsai, C.-J. & Mattos, C. 'Pathway drug cocktail': targeting Ras signaling
535   based on structural pathways. *Trends in molecular medicine* **19**, 695-704 (2013).
536 34. Zimmer, A., Katzir, I., Dekel, E., Mayo, A.E. & Alon, U. Prediction of multidimensional
537   drug dose responses based on measurements of drug pairs. *Proceedings of the National
538   Academy of Sciences* **113**, 10442-10447 (2016).
539 35. Minor, D.R., Puzanov, I., Callahan, M.K., Hug, B.A. & Hoos, A. Severe gastrointestinal
540   toxicity with administration of trametinib in combination with dabrafenib and
541   ipilimumab. *Pigment cell & melanoma research* **28**, 611-612 (2015).
542 36. Ianevski, A., He, L., Aittokallio, T. & Tang, J. SynergyFinder: a web application for
543   analyzing drug combination dose–response matrix data. *Bioinformatics* **33**, 2413-2415
544   (2017).
545 37. Tenenbaum, D. (2015).
546 38. Das, J. & Yu, H. HINT: High-quality protein interactomes and their applications in
547   understanding human disease. *BMC systems biology* **6**, 92 (2012).

12

548   39.   Khurana, E., Fu, Y., Chen, J. & Gerstein, M. Interpretation of genomic variants using a
549   unified biological network approach. *PLoS computational biology* **9**, e1002886 (2013).
550   40.   Aksoy, B.A. et al. PiHelper: an open source framework for drug-target and antibody-
551   target data. *Bioinformatics* **29**, 2071-2072 (2013).
552   41.   Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC
553   curves. *BMC bioinformatics* **12**, 77 (2011).
554   42.   Saito, T. & Rehmsmeier, M. Precrec: fast and accurate precision–recall and ROC curve
555   calculations in R. *Bioinformatics* **33**, 145-147 (2017).
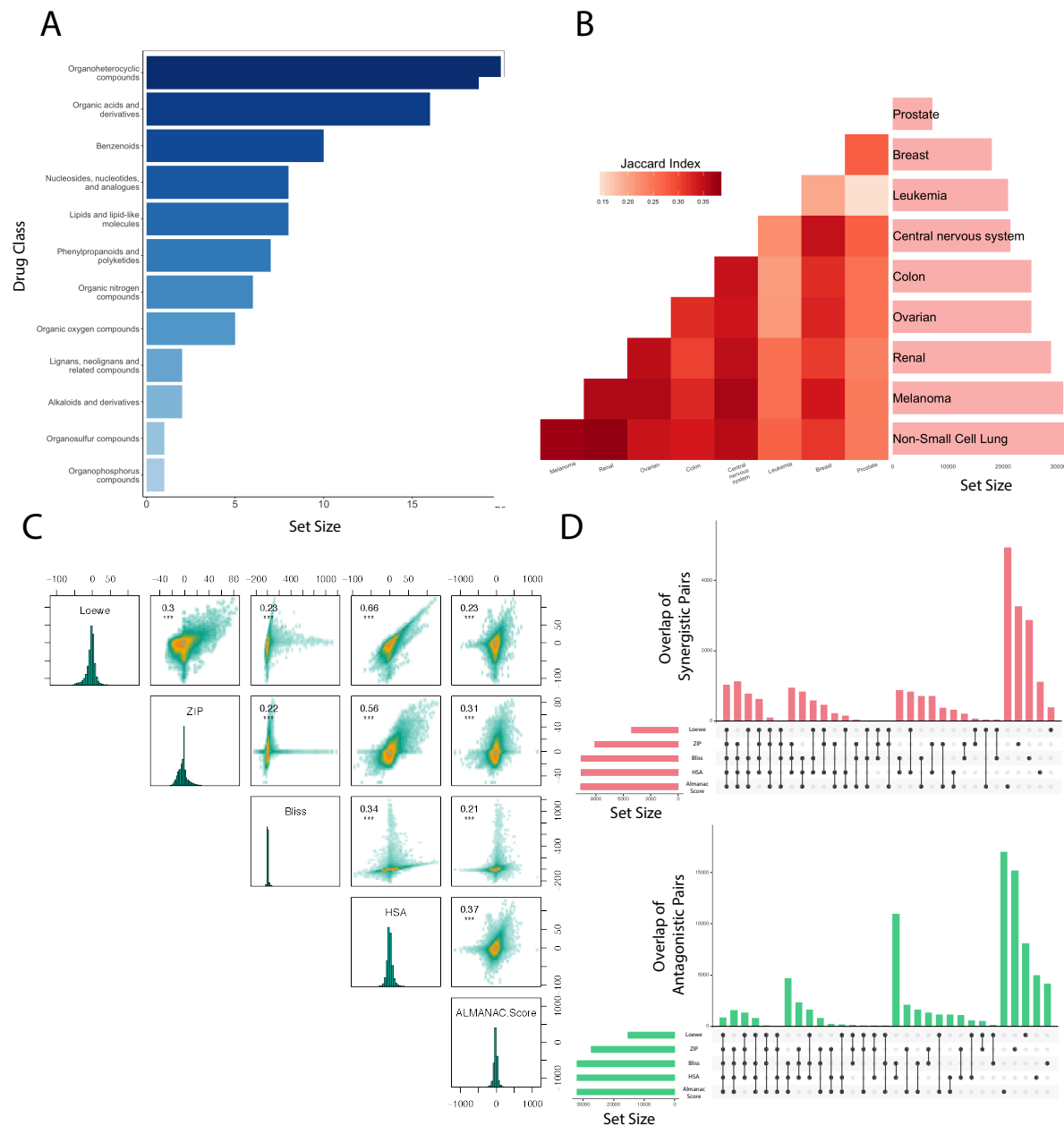556
557
558
559

Figure 1: Drug synergy metrics show distinct patterns across drug classes and cancer types. A) Distribution of drug classes for all the drugs tested within NCI-ALMANAC and used within this analysis. B) For each metric, the similarity of synergistic drug pairs between cancer types was measured using a Jaccard Index. The average of those 5 indices is shown above. The number of experiments done in cell lines of each primary site are also shown. C) The distribution of each synergy metric for all drug pairs and the correlation between them. D) The overlap between synergistic and antagonistic drug pairs in each tested cell line for each metric type.
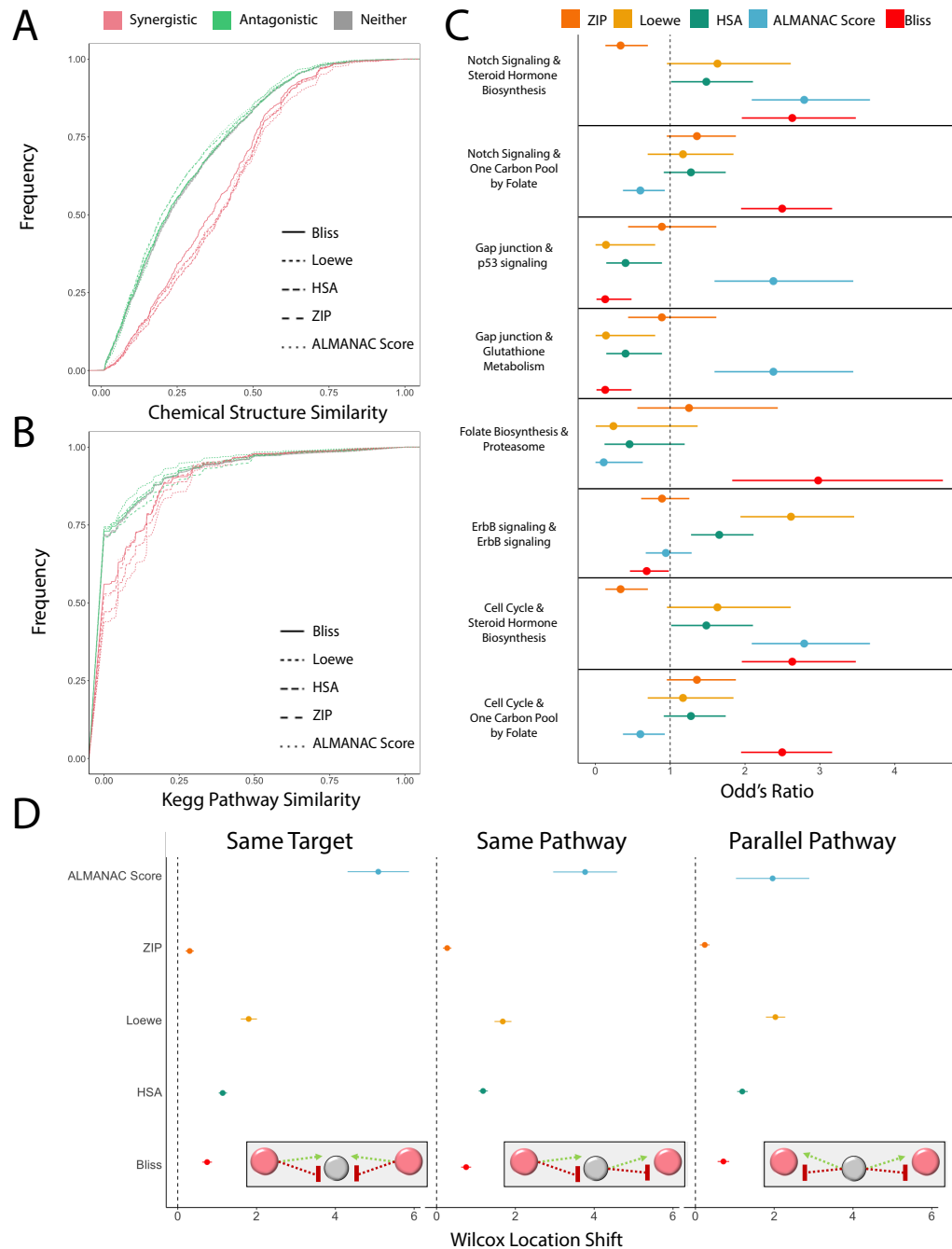
14

569
570 Figure 2: Distinct attributes of synergistic drug pairs as defined by each metric. The cumulative
571 distribution plot showing the proportional difference between antagonistic, synergistic, and all
572 other drug pairs at different similarity cut-offs for A) chemical structure and B) KEGG pathway
573 similarity. C) The odd's ratio plot, as measure by a Fisher's Exact test, for the enrichment of
574 pairs of targeted pathways in synergistic drug pairs. D) The Wilcoxon location shift for
575 synergistic drug pairs acting in one of three pathway mechanisms: same target, same pathway or
576 parallel pathway. The visual aids represent a simplified mechanistic overview, red indicating the
577 drug and grey indicating the targeted drug. Green arrows indicate a gene/drug that activates and
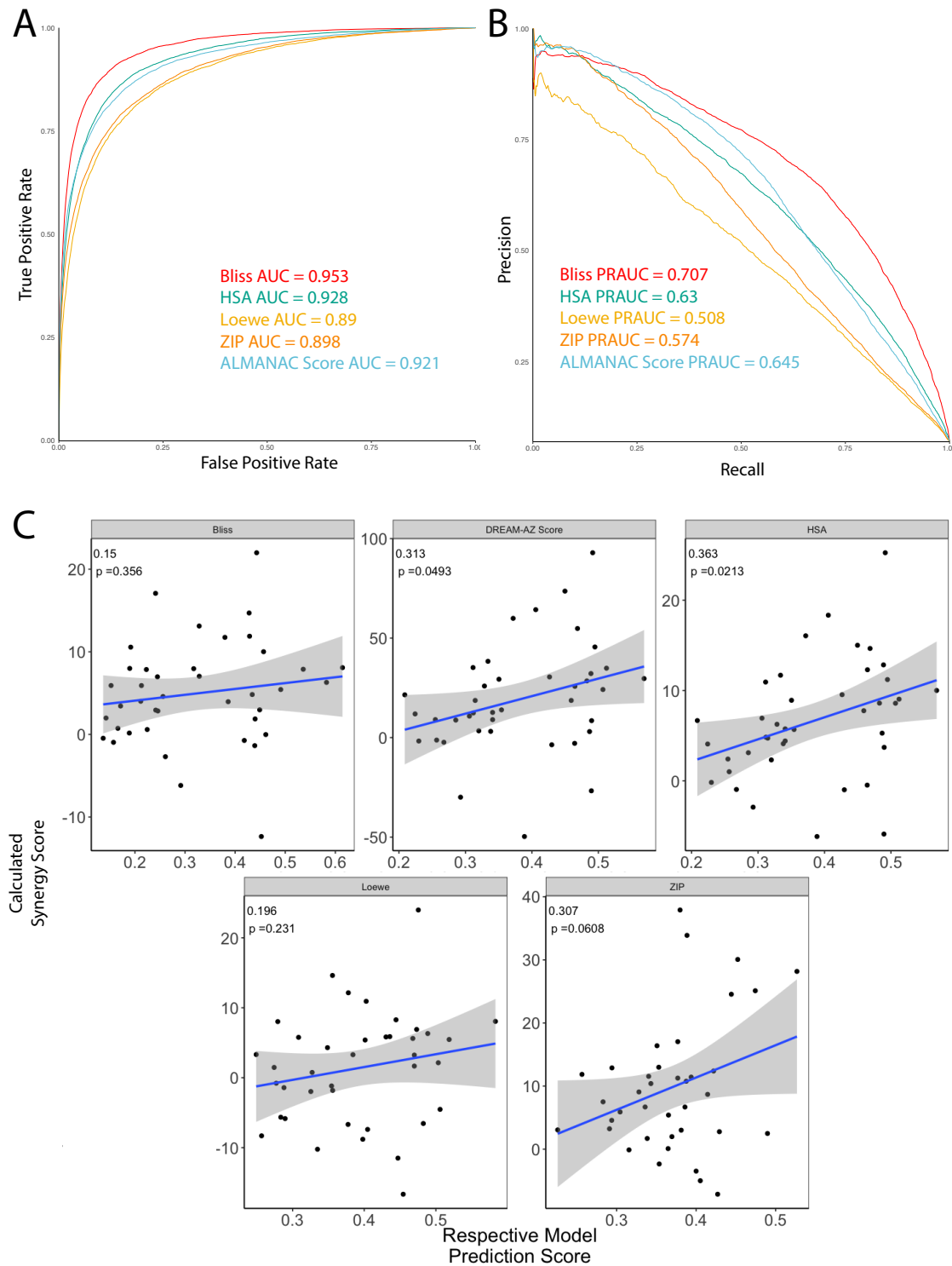578 red lines indicate a gene/drug that suppresses/inhibits.
579

15

Figure 3: The performance of the suite of predictive models for drug synergy across metrics. For each model, which predicts drug synergy as defined by each respective model, the performance was measured by A) the receiver-operating characteristic curve and B) the precision-recall curve. C) The correlation between predicted synergistic scores and the calculated scores from the DREAM –AZ challenge for all metrics.
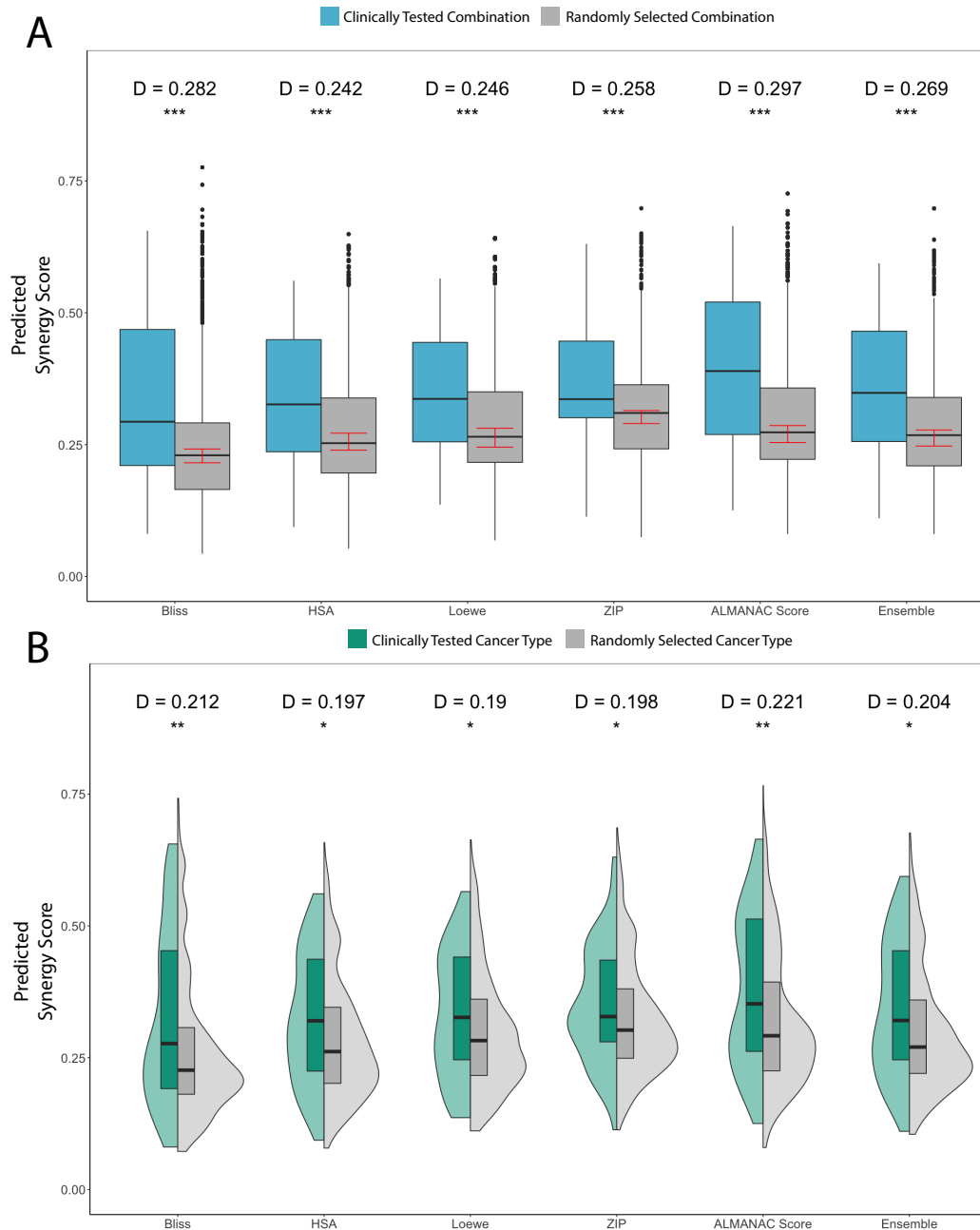
586
587 Figure 4: The clinical importance and performance of each drug synergy metric model. A)
588 Boxplots of the distribution of synergy scores, predicted by the suite of models, for drug pairs
589 that have been clinically tested and those that have not. The red bars show the confidence
590 interval for the median score of non-clinically tested pairs. B) Violin plots of the distribution of
591 predicted synergy score for clinically tested drug pairs within the cancer cell lines matched to the
592 clinically tested cancer types compared to those drug pairs in non-clinically tested cancer types.
593 KS tests were used to find the d-statistic and p-value.
594
595
596

597 **Supplementary Figures**

598 Figure S1: Spearman correlation between all synergy metrics. A) The distribution of each
599 synergy metric for all drug pairs and the spearman correlation between them.

600

601 Figure S2: Structural similarity across all drug pairs. The structural similarity for synergistic,
602 antagonistic and all other drug pairs as defined by A) Hydrogen Similarity and B) Bonds
603 Similarity, statistical significance found by Kolmogorov-Smirnov test.

604

605 Figure S3: Drug target similarity across all drug pairs. The similarity for synergistic, antagonistic
606 and all other drug pairs measured by A) Drug Target Similarity, B) KEGG/Reactome/GO
607 Pathway Similarity (based on drug targets) and C) Drug target network distance, statistical
608 significance found by Kolmogorov-Smirnov test.

609

610 Figure S4: Relatedness of KEGG pathway pairs of interest. The genes associated with the first
611 and second pathways listed shown in blue and gray, respectively. All overlapping genes are
612 shown in green.

613

614 Figure S5: The effect of drug pairs with the same target. For all metrics, the change in synergy
615 score when drug pairs A) targeted, B) activated or C) inhibited the same target, statistical
616 significance found by the Wilcoxon Test.

617

618 Figure S6: The effect of drug pairs with gene targets in parallel pathways. For all metrics, the
619 change in synergy score when drug pairs A) targeted, B) activated or C) inhibited parallel
620 pathways, statistical significance found by the Wilcoxon Test.

621

622 Figure S7: The effect of drug pairs with gene targets in the same pathway. For all metrics, the
623 change in synergy score when drug pairs A) targeted, B) activated or C) inhibited the same
624 pathway, statistical significance found by the Wilcoxon Test.

625

626 Figure S8: Model performance. For all models within our suite of drug synergy prediction
627 models the performance was measured when controlled for each "task", or cell line.

628

629 Figure S9: Correlation of AZ-DREAM score. The Pearson correlation between the provided
630 AstraZeneca-DREAM Challenge score and A) Loewe, B) Bliss, C) ZIP, and D) HSA.

631

632 **Supplementary Table**

633 Table S1: All features used within our extreme randomized model. Features are listed with their
634 source and the metric used to measure similarity.

635

636