A KINASE INHIBITION MAP APPROACH FOR TUMOR SENSITIVITY PREDICTION AND COMBINATION THERAPY DESIGN FOR TARGETED DRUGS

RANADIP PAL* NOAH BERLOW+

Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX, 79409, USA
*E-mail: *ranadip.pal@ttu.edu +noah.berlow@ttu.edu*

Drugs targeting specific kinases are becoming common in cancer research and are a basis for personalized cancer therapy. Some of these drugs have the capacity to target multiple kinases. Promiscuous kinase inhibitors can be effective but the "off-target" effects can bring in toxicity for the patient. Thus the success of targeted cancer therapies with nominal harmful side effects is dependent on administering a single or multiple combinations of kinase inhibitors that targets the minimum number of kinases required to inhibit the tumor pathways. This requires a framework to predict the tumor sensitivities of a drug or drug combination based on the knowledge of the kinase inhibitors of a drug. In this article, we present a novel approach to predict the tumor sensitivities of a drug based on the generation of deterministic and stochastic Kinase Inhibition Maps. We build sensitivity maps or truth tables for a cell line from experimentally generated tumor sensitivities to kinase inhibitor drugs and use them to predict the sensitivity of a new drug or drug combinations based on known kinase inhibitor targets. We test our algorithms on a dataset of a dog osteosarcoma cell line with 317 possible kinase inhibitor targets after application of 36 targeted drugs. Our proposed algorithms are able to predict the sensitivities with high accuracy based on the given kinase inhibitor targets.

Keywords: Targeted therapy design, Drug sensitivity prediction

1. Introduction

Drugs that target specific kinases are becoming common in cancer research (see[1–5]). Cancer-related kinases are the paradigm of molecularly-targeted therapies, and a cornerstone of Personalized Cancer Therapy. The kinome consists of greater than 500 diverse tyrosine- or serine-threonine kinases ([6]). Some of these drugs have the capacity to target multiple kinases ([7]) and effects of multiple kinase inhibition are still to be properly understood. A prediction problem that we often encounter is predicting the sensitivity of a new drug for cell lines when the kinases that are inhibited by the drug are known. To address this, we formulate the problem as a boolean logic problem and build Kinase inhibiting maps for each cell line based on prior data. The Kinase inhibiting maps are used to answer the prediction problem.

The structure of kinases have been well studied for increasing the specificity of targeted drugs (see[8]) or high throughput kinase profiling has been used as a platform for drug discovery ([9]). However, approaches for high throughput analysis of multiple kinase inhibitor data to predict the behavior of new drugs are lacking in the literature. Prediction of tumor sensitivity to drugs have been approached earlier as a classification problem using gene expression profiles in *Staunton et al.*[10] . In *Staunton et al.*,[10] gene expression profiles are used to predict the binarized efficacy of a drug over a cell line with the accuracy of the designed classifiers ranging from 64% to 92%. In our proposed approach, we use the kinase inhibitor profiles of drugs as opposed to gene expression profiles of cell lines. The kinase inhibitor profiles of a drug are not dependent on a cell line and thus for a new cell line corresponding to a patient, we dont

need to measure the gene expression profiles and just have to measure the tumor sensitivity after application of different drugs. Furthermore, our stochastic approach can predict the effectiveness of a drug on more than binary levels and produce higher accuracy results. In *Lee et al.*,[11] a co-expression extrapolation (COXEN) approach is used to predict the drug sensitivity in data points outside the training set with an accuracy of around 75% in predicting the binarized sensitivity. Our proposed approach is unique in not using gene expression biomarkers for prediction and using set theory and stochastic extensions to predict the non-binarized tumor sensitivity with high accuracy.

Drugs like Staurosporine have numerous kinase targets and are extremely effective in inhibiting the tumor but also have high toxicity ([7]). In fact, the toxicity of Staurosporine is similar to chemotherapy and thus working against the goal of targeted action on tumor cells without damaging normal cells and tissues. On the other hand, Imatinib (Gleevec) that has minimal kinase targets are being taken by leukemia patients for over a decade with no major side effects. Thus, it is extremely important that the off target effects of drugs are taken into consideration before prescribing a therapeutic regime for a tumor patient. Having a mathematical framework for the possible tumor pathways can allow us to generate the minimal set of kinase inhibition combinations that are required to block all the activated tumor pathways. A combination of drugs can be selected that just targets a minimal set of kinase inhibitors, with extremely few off target inhibitions. Majority of the current approaches for modeling genetic regulatory networks are not well suited for tackling this issue as the data requirements for model parameter estimation are significantly more in terms of number of samples and requirement of primarily time series data for estimation of the model parameters.[12] We propose a novel way to utilize steady-state experimental data on drug sensitivities over cell lines or animal models to generate robust maps for representing the tumor pathways. We exploit the fact that if a drug with inhibitor set S is effective, then any other drug with inhibitor set containing S will also be effective.

The paper is organized as follows: the mathematical formulation of the problem is provided in Section 2, algorithms to generate kinase inhibition maps are provided in Section 3, Section 4 covers the algorithm to predict the sensitivity of a new drug based on kinase inhibitor targets, Section 5 covers the probablistic maps; the results are presented in section 6 and finally section 7 considers the conclusions.

2. Mathematical Formulation

Let us a consider that we have drug sensitivity data for n cell lines after application of m drugs. The known multi kinase inhibiting sets for these drugs are denoted by S_1,S_2,...,S_m. The cardinality or number of elements of S_i is given by a_i for $i = 1, 2, \cdots, m$. Let S denote the union of $S_1, S_2, ..., S_m$ i.e. $S = \cup_{i=1}^{m} S_i$ and a denote the cardinality of S i.e. $a = |S|$. Thus there are a kinases that we are interested in and we will denote them as $k_1, k_2, \cdots, k_a$. The elements of sets S_i are denoted by $[e_{i,1}, e_{i,2}, \cdots, e_{i,a_i}]$ for $i = 1, 2, \cdots, m$. The $e_{i,j}$'s for $j = 1, .., a_i$ and $i = 1, 2, \cdots, m$ are elements from the set $[k_1, k_2, \cdots, k_a]$. For this scenario, the maximum possible number of distinct multiple kinase inhibitor activities exhibited by drugs are 2^a since each kinase can be either inhibited or not and there are a kinases, thus number of possibilities

are $(2 \times 2 \times \cdots \times 2)_{a \text{ times}} = 2^a$. To solve our prediction problem, we will construct Kinase Inhibiting Maps (KIMs) for each cell line similar to Karnaugh maps ([13]) to simplfy boolean expressions.

As an example, let us consider the abstract representation of a biological pathway shown in Fig. 1. We will consider that a drug works if it inhibits all the paths between X and Y.

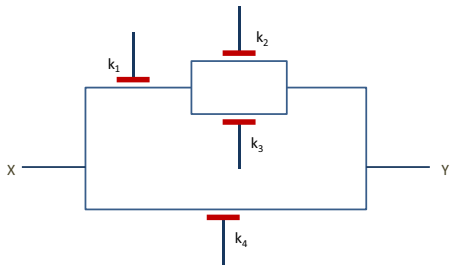

Fig. 1. Circuit 1

The circuit shown in Fig. 1 has three paths from X to Y through kinases $\{k_1, k_2\}$; $\{k_1, k_3\}$ and $\{k_4\}$. To block all the paths from X to Y, we have to inhibit all these parallel pathways. The first two pathways can be blocked by either inhibiting k_1 or inhibiting both k_2 and k_3. The third pathway has to be blocked by inhibiting k_4. If we represent this information as a map where a 1 indicates all pathways being blocked for that set of kinase inhibitions, then the resultant maping will be as shown in Figure 2. For example, a 1 in the 2nd row and 3rd column of the 4×4 matrix denotes that all the pathways for circuit 1 can be blocked by the combination $k_1 = 0, k_2 = 1, k_3 = 1, k_4 = 1$ i.e. any drug inhibiting k_2, k_3 and k_4 will have high sensitivity for circuit 1. Similarly, a 0 in the 2nd row and 2nd column denotes that inhibiting k_2 (i.e. k_2=1) and inhibiting k_4 (i.e. k_4=1) cannot produce high sensitivity for circuit 1.

$k_1\,k_2$ \ $k_3\,k_4$	**0 0**	**0 1**	**1 1**	**1 0**
0 0	0	0	0	0
0 1	0	0	1	0
1 1	0	1	1	0
1 0	0	1	1	0

Fig. 2. Kinase Inhibiting Map (KIM) corresponding to Circuit 1

The map shown in Table 2 will be termed as Kinase Inhibiting Map (KIM) and can be constructed from Circuit 1 by checking the output for all the 2^4 combinations of kinase activities. The boolean logic for the circuit can be constructed based on the fact that each series connection is like Boolean **OR** function and each parallel connection is like boolean **AND** function, thus circuit 1 refers to (k_1 **OR** (k_2 **AND** k_3)) **AND** (k_4).

3. Constructing kinase inhibition maps from drug targets and sensitivity data

For the construction of the kinase inhibition maps, the following two sets of rules relevant to our problem will assist in filling the entries of the map.

Rule 1: If S_i is the inhibiting set of kinases for drug i and the drug is successful in inhibiting the circuit, then any set B containing the set S_i (i.e. B is a superset of S_i, $B \supset S_i$) will also be successful in inhibiting the circuit.

Rule 2: If S_i is the inhibiting set of kinases for drug i and the drug is unsuccessful in inhibiting the circuit, then any set B that is the subset of set S_i (i.e. $B \subset S_i$) will also be unsuccessful in inhibiting the circuit.

Rule 1 essentially says that if inhibiting a number of kinases has blocked all the paths, then inhibiting more kinases will not open any path. For instance, if we consider circuit 1 and our experiments denote that the set $\{k_1, k_4\}$ is able to inhibit the circuit, then any superset of $\{k_1, k_4\}$ such as $\{k_1, k_4, k_2\}$ will also inhibit the circuit. The number of possible supersets of a set S_i containing a_i elements among possible a elements is 2^{a-a_i}. For circuit 1, if $S_i = \{k_1, k_4\}$, the number of possible supersets of S_i is $2^{4-2} = 4$ and they are $\{k_1, k_4\}, \{k_1, k_4, k_2\}, \{k_1, k_4, k_3\}$ and $\{k_1, k_4, k_2, k_3\}$. If we consider the kinase inhibiting map in Fig. 2, then knowing the information that $\{k_1, k_4\}$ inhibits the circuit, we can fill the entries of its superset as 1. The entries corresponding to its superset are $[k_1k_2k_3k_4] = \{1001, 1101, 1011, 1111\}$ which fills the KIM(3,2), KIM(3,3), KIM(4,2) and KIM(4,3) entries of the inhibition map. Here KIM(i, j) denotes the ith row and jth column entry of the Inhibition map. The kinase inhibition map is a matrix of size $2^{p_1} \times 2^{p_2}$ where $p_1 + p_2 = a$ and $p_1 = \lfloor a/2 \rfloor$ and $p_2 = a - p_1$.

Rule 2 captures the fact that if a set of kinase inhibitors is unsuccessful in blocking the paths of a circuit, then any reduced number of kinase inhibitors among the inhibiting kinases cannot block all the paths. For instance, in circuit 1, if our experiments denote that the set $\{k_1, k_2, k_3\}$ of kinase inhibition is not sucessful in blocking all the paths of the circuit, then any subset of $\{k_1, k_2, k_3\}$ such as $\{k_1, k_2\}$ will also be unsuccessful in blocking the paths of the circuit. The number of possible subsets of a set S_i containing a_i elements is 2^{a_i}. For circuit 1, if $S_i = \{k_1, k_2, k_3\}$, the number of possible subsets of S_i is $2^3 = 8$ and they are $\{k_1, k_2, k_3\}, \{k_1, k_2\}, \{k_1, k_3\}, \{k_2, k_3\}, \{k_1\}, \{k_2\}, \{k_3\}$ and $\{\}$. In the boolean logic, the subsets will be as follows $[k_1k_2k_3k_4] = \{1110, 1100, 1010, 0110, 1000, 0100, 0010, 0000\}$. Thus, if a drug with set S_i of inhibitors is unsuccessful, then we can mark the states in the kinase inhibition map that has zeros for the $S - S_i$ kinases as zeros. Thus for our example, $S - S_i = \{4\}$ and any state having zero for k_4 will be marked zero in the kinase inhibition map. So the first and fourth column of the kinase inhibition map will be marked 0 based on the experimental piece of information that the set $\{k_1, k_2, k_3\}$ is unsuccessful. This approach fills up a large number of entries of the kinase inhibition map based on limited experimental knowledge. For our example, the two experimental results that $\{k_1, k_4\}$ is successful and $\{k_1, k_2, k_3\}$ is not successful fills in 12 out of 16 entries in the kinase inhibition map. Higher number of entries are filled when a large set of kinase inhibitors is unsuccessful or a small set of kinase inhibitors is successful.

The algorithm for constructing the kinase inhibitor map is shown in Algorithm 3.1.

Algorithm 3.1 Algorithm for constructing Kinase Inhibitor Map KIM

```
for i = 1 to m do
  for j = 1 to n do
    if Drug is successful then
      {Use Rule 1 to fill up entries in the kinase inhibition map for cell line j (KIMj)}
      for all R such that R ⊃ S_i do
        KIMj (R) = 1;
      end for
    else
      {Use Rule 2 to fill up entries in the kinase inhibition map for cell line j (KIMj)}
      for all R such that R ⊂ S_i do
        KIMj (R) = 0;
      end for
    end if
  end for
end for
```

4. Predicting sensitivity of a new drug based on its set of kinase inhibitors

If the set of kinase inhibitors is known for a new drug, then we can check the kinase inhibition map entry for that set of inhibitors for each cell line and predict the outcome of the drug when applied to that cell line.

Example 4.1.

Let us consider that we have two cell lines whose abstract circuit representation of the pathways are shown in Figures 3 and 4. Based on experimental data on these cell lines, we can construct the Kinase Inhibitors Maps KIM1 and KIM2 as shown in Figures 5 and 6. Here $S = \{k_1, k_2, k_3, k_4\}$ and k_4 is not directly involved in pathway 1 whereas k_1 is not directly involved in pathway 2. If a new drug D_{m+1} has the kinase inhibition set $S_{m+1} = \{k_2, k_3, k_4\} = \{0111\}$, then based on the 2nd row, 3rd column entries of the inhibition maps IM1 and IM2, the drug will be ineffective for CP1 and effective for CP2.

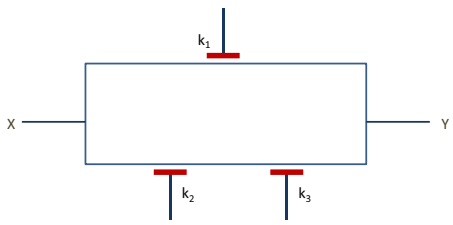

Fig. 3. CP1: Cellular pathway representation 1

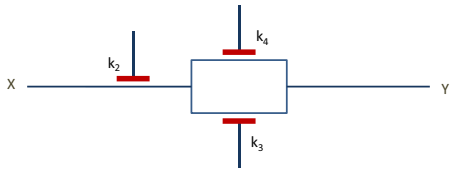

Fig. 4. CP2: Cellular pathway representation 2

$k_1\,k_2$ \ $k_3\,k_4$	0 0	0 1	1 1	1 0
0 0	0	0	0	0
0 1	0	0	0	0
1 1	1	1	1	1
1 0	0	0	1	1

Fig. 5. KIM1: Kinase Inhibition map corresponding to Cellular pathway representation 1 shown in Fig. 3

$k_1\,k_2$ \ $k_3\,k_4$	0 0	0 1	1 1	1 0
0 0	0	0	1	0
0 1	1	1	1	1
1 1	1	1	1	1
1 0	0	0	1	0

Fig. 6. KIM2: Kinase Inhibition map corresponding to Cellular pathway representation 2 shown in Fig. 4

5. Stochastic Extension

The approach described in the previous sections depend on the accuracy of the experimental data and accurate modeling assumptions. When there is a mismatch in the biological pathway and its abstract modeling and there are noises in the extraction of the data, we have to consider modifications of our algorithm to make it robust to uncertainties. Currently, the entries of the kinase inhibition map are updated based on the latest experimental data and in the process, the inferences of previous experimental data can be lost. For instance, let us consider the pathway representation of Figure 1 and our first experiment states that a drug combination of $\{k_1, k_2, k_3\}$ is ineffective, then the third row fourth column entry is zero. If now, an error in the data extraction of a new experiment states that $\{k_1, k_3\}$ is effective, then the third row fourth column entry will be erroneously changed to 1. Furthermore, the sensitivity to drugs will not necessarily be binary (effective or ineffective), thus our modified algorithm will also take into account the percentage of effectiveness of a drug. The effectiveness of a drug is usually measured by their IC_{50} values which denotes the amount of drug required to inhibit

the activity of a process by half. So if the levels of drugs used for our experiment is from 1 unit to U units (say 1 to 10,000 nano-molar units), we will use the following approach to get the scaled sensitivity y of the drug with IC_{50} value x:

$$y = \frac{U - x}{U} \tag{1}$$

Note that equation 1 is one possible approach to map the IC_{50} values to numbers between 0 and 1 and there can be many other ways for this maping. To arrive at the new stochastic Kinase Inhibition Map, we will initially consider two matrices IM^1 and IM^0 denoting the most effective and ineffective combinations respectively. Then, the two matrices are combined in a proportional ratio to arrive at the final Probabilistic Kinase Inhibition Map ($PKIM$).

Algorithm 5.1 Algorithm for constructing Probabilistic Kinase Inhibitor Map PKIM

{Generate Kinase Inhibition Maps IM^1 and IM^0}
for $j = 1$ to n **do**
  Initialize $IM_j^1 = IM_j^0 = \mathbf{0}$
  for $i = 1$ to m **do**
    {Use Rule 1 to fill up entries in IM^1} for cell line j
    for all R such that $R \supset S_i$ **do**
      IM_j^1 (R) = IM_j^1 (R) + y_i;
    end for
    {Use Rule 2 to fill up entries in IM^0} for cell line j
    for all R such that $R \subset S_i$ **do**
      IM_j^0 (R) = IM_j^0 (R) + 1 − y_i;
    end for
  end for
end for
{Generate Probabilistic Kinase Inhibition Map $PKIM$ }
for all $CL_j \in \{CL_1, \cdots, CL_n\}$ **do**
  Initialize $PKIM_j = \mathbf{0}$
  for all $v \in V$ **do**
    $PKIM_j$ (v) = $\frac{IM_j^1(v)}{IM_j^1(v)+IM_j^0(v)}$;
  end for
end for

return $PKIM_1, PKIM2, ..., PKIM_n$

6. Results

To illustrate the effectiveness of the proposed algorithms, we consider experimental data on a canine osteosarcoma cell line *CanOS1224* that was treated with 36 targeted cancer drugs and the tumor sensitivities measured. The experiments were conducted in the labs of Charles

Keller and Brian Druker in Oregon Health and Science University, Portland, USA. The cell line *CanOS1224* was derived from an actual canine with osteosarcoma who is being treated in OHSU as part of a clinical trial. Thus the data reflects a new cell line without much prior biological knowledge specific to that cell line and success of tumor sensitivity prediction on such a cell line holds promise for personalized therapy for a new patient coming to the clinic. The canine osteosarcoma primary cell culture was plated in 96 well plates at a seeding density of 2000 cells per well over graded concentrations of 36 small-molecule kinase inhibitors. Each inhibitor was plated individually at four concentrations predicted to bracket the IC_{50} for that drug. Cells were cultured in RPMI 1640 supplemented with 2mM glutamine, 2mM sodium pyruvate, 2mM HEPES, 1% penicillin streptomycin, and 10% fetal bovine serum for 72 hours. At the end of the 72 hour incubation, cell viability was assessed using the MTS assay. All values were normalized to the mean of seven wells on each plate containing no drug. The IC_{50} for each drug was then determined by identification of the two concentrations bracketing 50% cell viability and application of the following formula: [((A-50)/(A-B))*(Dose B - Dose A)]+ Dose A where cell viability value above 50% = A (drug dose for this value is Dose A) and cell viability value below 50% = B (drug dose for this value is Dose B).

The drug targets (kinases inhibited by a drug) are obtained from the supplementary tables of[7] and[14] based on experimental quantitative dissociation constant (k_d) values for each drug across 317 kinase assays. We considered a drug to be inhibiting the kinase if the k_d value is less than 10 percent of the maximum. The sensitivities (in terms of IC_{50} values) of the 36 drugs are experimentally generated and converted to numbers between 0 and 1 with 2 significant digits using Eq. 1 where U is 10000nM. Using all the 317 kinases to build the kinase inhibition map has the following problems: (a) the size of the Kinase Inhibition Map with $2^{317} = 2.67 \times 10^{95}$ entries will be computationally intractable; (b) data on 36 drugs is not rich enough to fill majority of the entries and keeping too many kinases will lead to over-fitting and (c) a number of kinases in the non-activated pathways of the cell line are not necessary to predict the tumor sensitivity to drugs. Thus, we used sequential feature selection to narrow down our number of Kinases.The cost function considered was mean absolute error in predicting the tumor sensitivities. We initially started with the kinase EGFR and sequentially added other kinases to it with 8 being the maximum number of kinases to be selected. For a set of kinases at any stage of feature selection, we used leave-one-out error estimation technique to calculate the mean absolute error. The 8 kinases selected through sequential selection approach are PIM1, PIK3CA, MRCKA, EPHA3, MAP4K5, MET, ACVR1B and EGFR. The data is shown in Figure 7. The first column shows the 36 drugs and the next 8 columns are the scaled dissociation constants for the 8 selected Kinases. The scaled dissociation constants ($0 \leq k_d \leq 1$) are obtained from[7] and[14] where a high value reflects that the Kinase is inhibited by the drug. To binarize the inhibition of kinases by a drug, we considered a kinase to be inhibited by the drug if the scaled k_d value is ≥ 0.9. The scaled tumor sensitivities are shown in column 10 in Figure 7 where a value of 1 reflects that the tumor is highly sensitive to the drug. We considered the probabilistic Kinase Inhibition Map method and used leave one out cross validation to measure the effectiveness of the PKIM technique. For leave one out error, we used the data on 35 drugs to build the PKIM and used the generated PKIM to predict the

sensitivity for the 36th drug. This is done for all the 36 drugs and the average error is 0.067 which is really low for leave one out validation. Note that the maximum error can be 1 and a fixed prediction of 0.5 will result in an error of 0.437. The leave one out errors for each drug is included in the 11th column of Figure 7.

If we use re-substitution error estimate, then the average error decreases to 0.018. The re-substitution errors for each drug is shown in Column 12 of Figure 7. The decrease in error for re-substitution is expected as we are using the same data for training and testing. For example, we should note that the leave one out error for the drug PI-103 is 0.903. This is because the kinase signature of this drug is not available in any other drug i.e. no other drug targets the PIK3CA pathway as this drug does. Targeting the PIK3CA pathway for this cell line is able to reduce the tumor and this knowledge can only be gained from the data from the drug PI-103 and none of the other 35 drugs considered here. However, when we use the sensitivity of drug PI-103 in training the PKIM, we can predict the sensitivity for the drug PI-103 perfectly from the PKIM as shown in the case of re-substitution error. To be able to accurately predict the sensitivity of a new drug, it is imperative that our training data contain sets of kinase inhibitors that target most of the important pathways such as MAPK, JAK-STAT, mTOR, ERBB etc. The important pathway information can be obtained from public databases such as *KEGG*(http : //www.genome.jp/kegg/) or *pathway commons* (http : // www.pathwaycommons.org). Also drugs targeting both individual and combination of pathways can provide complementary information in estimating the entries of the kinase inhibition map.

The probabilistic KIM that was generated using the experimental data from the 36 drugs is shown in Figure 8. The PKIM was then used to build a circuit representation for cell line CanOS1224. The shaded areas in Figure 8 were considered to be 1 for generating the circuit shown in Figure 9. The circuit representation is easier to visualize and to study effective drug combinations, however the probabilistic Kinase Inhibition Map in Figure 8 should be referred to get the actual prediction. To arrive at a simpler circuit representation, some entries in Figure 8 is considered effective even when they have a low value. For instance, the entry for PIM1 being inhibited alone is 0.26 and still we considered it as effective for drawing our circuit. The PKIM or the circuit representation can be extremely useful in designing combination therapy for diseases. For instance, if our goal is to avoid resistance to drugs evolving in a patient, then we should try to block more than one series of kinase sets that each individually can block the tumor. For instance, in Figure 9, if we block PIK3CA, MRCKA and PIM1 then we will have two independent sets of Kinases that block the tumor. If the cancer becomes resistant to inhibition of PIM1, then the other set of Inhibitors MRCKA and PIK3CA will be the next line of defense against the tumor. From the data in figure 7, Flavopiridol and PI-103 can be used to target the three Kinases PIK3CA, MRCKA and PIM1. If our goal is to maximize chances of success and minimize off target effects, then we will pick the drug or drug combinations that inhibit minimum number of kinases but blocks all the paths of the circuit. To have a measure of toxicity of a drug, we should consider the total number of kinases inhibited by the drug among the 317 kinases tested and not just the 8 kinases considered here.

	Kinases								Sensitivities	Leave one out error	Re-substitution error
drugs	EGFR	PIM1	MRCKA	ACVR1B	PIK3CA	EPHA3	MAP4K5	MET	CanOS		
ABT-869	0	0	0	0	0	0.65	0.91	0.87	0.31	0.00791	0.004051
AMG-706	1	0	0	0	0	0	0	0	0	0	0
AST-487	0.95	0	0	0	0	0.99	0.96	0.56	0.51	0.328909	0.208381
AZD-1152HQPA	0.95	0	0	0	0	0	0.98	0	0	0	0
BIRB-796	0.3	0	0.06	0	0	0.91	0	0	0	0	0
BMS-387032/SNS-	0	0	0	0	0	0	0	0	0	0	0
CHIR-258/TKI-258	0	0	0	0	0	0	0.94	0	0	0	0
CHIR-265/RAF-265	0	0	0	0	0	0	0.92	0	0	0	0
CI-1033	1	0	0	0	0	0.79	0.74	0.44	0.26	0.063798	0.040419
CP-690550	0	0	0	0	0	0	0	0	0	0	0
Dasatinib	0.99	0	0.8	0.97	0	1	1	0	0.95	0.050409	0.027331
EKB-569	1	0	0	0	0	0.6	1	0.38	0	0	0
Erlotinib	1	0	0	0	0	0.76	0	0.62	0	0	0
Flavopiridol	0	0.94	0.05	0	0	0.67	0	0	0.94	0.052976	0.011501
Gefitinib	1	0	0	0	0	0	0	0	0	0	0
GW-2580	0	0	0	0	0	0	0	0	0	0	0
GW-786034	0	0	0	0	0	0	0.7	0	0	0	0
Imatinib	0	0	0	0	0	0	0	0	0	0	0
JNJ-7706621	0	0	0	0	0	0	0.79	0.94	0	0	0
Lapatinib	1	0	0	0	0	0	0	0	0	0	0
LY-333531	0	0.97	0	0	0	0	0	0	0.28	0.2766	0.021387
MLN-518	0.96	0	0	0	0	0	0.92	0	0	0	0
MLN-8054	0	0	0	0	0	0.79	0	0	0	0	0
PI-103	0	0	0.82	0	1	0	0	0	0.91	0.9064	0
PKC-412	0.87	0.94	0	0	0	0	0.89	0.93	0.97	0.025076	0.005444
PTK-787	0	0	0	0	0	0	0	0	0	0	0
SB-202190	0.74	0	0	0.91	0	0	0	0	0	0	0
SB-203580	0.83	0	0.38	0.7	0	0	0	0	0	0	0
SB-431542	0	0	0	0.98	0	0	0	0	0	0	0
Sorafenib	0	0	0	0	0	0.81	0.84	0	0.05	0.0489	0.039753
Staurosporine	0.96	1	0.99	0.93	0	1	1	0.98	1	0.0007	0.000581
SU-14813	0	0	0	0	0	0	0.97	0	0	0	0
Sunitinib	0	0	0	0	0	0.79	1	0	0	0.011252	0.009147
VX-680/MK-0457	0	0	0	0	0	0.85	0.99	0.93	0.26	0.074105	0.037949
VX-745	0	0	0	0	0	0	0	0	0	0	0
ZD-6474	1	0	0.74	0	0	0.8	0.95	0.43	0.37	0.568434	0.263415

Fig. 7. Experimental data and predicton errors using probabilistic KIMs.

7. Conclusions

In this article, we presented an approach to generate abstract representation of cancer pathways that can assist in predicting the sensitivities of a new drug given the kinase inhibitors of the drug. We also extended our algorithm to the probabilistic case to tackle latent kinases and noisy data and to predict non-binarized tumor sensitivities. The algorithms were validated on data obtained after application of targeted cancer drugs on a canine cancer cell line. The approach presented here is a novel way to analyze tumor sensitivity data based on Boolean logic and set theory. We expect that improved results can be obtained when the training set

KINASES																	
		MET	MAP4K5 MET	MAP4K5	EPHA3 MAP4K5	EPHA3 MAP4K5 MET	EPHA3 MET	EPHA3	PIK3CA EPHA3	PIK3CA EPHA3 MET	PIK3CA EPHA3 MAP4K5 MET	PIK3CA EPHA3 MAP4K5	PIK3CA MAP4K5	PIK3CA MAP4K5 MET	PIK3CA MET	PIK3CA
	0	0	0	0	0.01	0.3	0	0	0	0	1	1	0	0	0	0
ACVR1B	0	0	0	0	0.49	1	0	0	0	0	1	1	0	0	0	0
MRCKA ACVR1B	0	0	0	0	0.49	1	0	0	1	1	1	1	1	1	1	1
MRCKA	0	0	0	0	0.07	1	0	0	1	1	1	1	1	1	1	0.91
PIM1 MRCKA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PIM1 MRCKA ACVR1B	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PIM1 ACVR1B	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PIM1	0.26	0.91	0.91	0.91	1	1	1	0.96	1	1	1	1	1	1	1	1
EGFR PIM1	0.91	0.91	0.98	0.91	1	1	1	1	1	1	1	1	1	1	1	1
EGFR PIM1 ACVR1B	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
EGFR PIM1 MRCKA ACVR1B	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
EGFR PIM1 MRCKA	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
EGFR MRCKA	0	0	0	0	0.64	1	0	0	1	1	1	1	1	1	1	1
EGFR MRCKA ACVR1B	0	0	0	0	0.98	1	0	0	1	1	1	1	1	1	1	1
EGFR ACVR1B	0	0	0	0	0.94	1	0	0	0	0	1	1	0	0	0	0
EGFR	0	0	0	0	0.3	1	0	0	0	0	1	1	0	0	0	0

Fig. 8. Probabilistic Kinase Inhibition Map for CanOS1224. The first row and First column denotes the Kinases being inhibited.

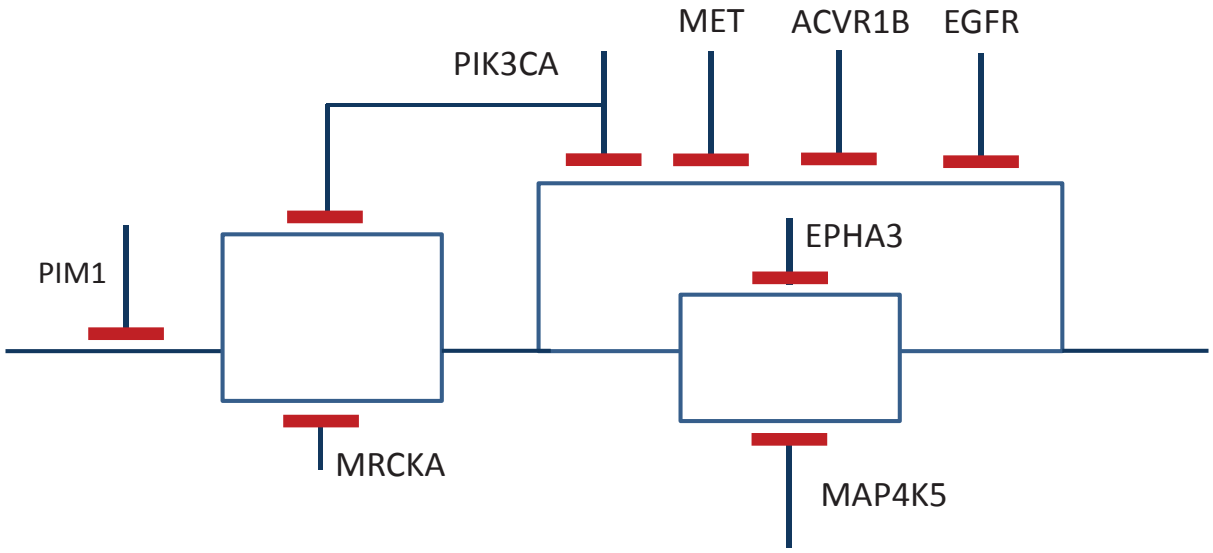

Fig. 9. Circuit representation for CanOS1224 based on PKIM in Figure 8.

will include more drugs with various combination of inhibiting kinases.

8. Acknowledgments

This research was supported by NSF grant CCF 0953366. We would like to thank Charles Keller, Jinu Abraham, Nicolle Hofmann, Jeffrey W. Tyner, Marc M. Loriaux and Brian J. Druker from Oregon Health & Science University and Bernard Seguin from Oregon State University for generating and providing the biological data.

References

1. C. Sawyers, *Nature* **432**, 294 (2004).
2. M. R. Green, *New England Journal of Medicine* **350:21**, 2191 (2004).
3. B. J. Druker, *Oncologist* **9**, 357 (2004).
4. A. Hopkins, J. Mason and J. Overington, *Current Opinion in Structural Biology* **16**, 127 (2006).
5. Z. A. Knight and K. M. Shokat, *Chemistry & Biology* **12**, 621 (2005).
6. G. Manning, D. B. Whyte, R. Martinez, T. Hunter and S. Sudarsanam, *Science* **298**, 1912 (2002).
7. M. W. Karaman, S. Herrgard, D. K. Treiber, P. Gallant, C. E. Atteridge, B. T. Campbell, K. W. Chan, P. Ciceri, M. I. Davis, P. T. Edeen, R. Faraoni, M. Floyd, J. P. Hunt, D. J. Lockhart, Z. V. Milanov, M. J. Morrison, G. Pallares, H. K. Patel, S. Pritchard, L. M. Wodicka and P. P. Zarrinkar, *Nature biotechnology* **26**, 127 (January 2008).
8. Chen, Jianping, Zhang and Xi, *Bioinformatics* **23**, 563 (2007).
9. D. Goldstein and et al., *Nat. Rev. Drug Discov.* **7**, 391 (2008).
10. J. E. Staunton, D. K. Slonim, H. A. Coller, P. Tamayo, M. J. Angelo, J. Park, U. Scherf, J. K. Lee, W. O. Reinhold, J. N. Weinstein, J. P. Mesirov, E. S. Lander and T. R. Golub, *Proceedings of The National Academy of Sciences* **98**, 10787 (2001).
11. J. K. Lee, D. M. Havaleshko, H. Cho, J. N. Weinstein, E. P. Kaldjian, J. Karpovich, A. Grimshaw and D. Theodorescu, *Proceedings of the National Academy of Sciences* **104**, 13086 (August 2007).
12. Z. Szallasi, J. Stelling and V. Periwal, *System Modeling in Cell Biology from Concepts to Nuts and Bolts* (MIT Press, Cambridge, MA, 2006).
13. M. Karnaugh, *Trans. AIEE. pt. I* **72**, 593 (1953).
14. P. P. Zarrinkar, R. N. Gunawardane, M. D. Cramer, M. F. Gardner, D. Brigham, B. Belli, M. W. Karaman, K. W. Pratz, G. Pallares, Q. Chao, K. G. Sprankle, H. K. Patel, M. Levis, R. C. Armstrong, J. James and S. S. Bhagwat, *Blood* **114**, 2984 (2009).