Human Genomics

K-Map: connecting kinases with therapeutics for drug repurposing and development

Jihye Kim[1], Minjae Yoo[1], Jaewoo Kang[2] and Aik Choon Tan[1*]

Abstract

Protein kinases play important roles in regulating signal transduction in eukaryotic cells. Due to evolutionary conserved binding sites in the catalytic domain of the kinases, most inhibitors that target these sites promiscuously inhibit multiple kinases. Quantitative analysis can reveal complex and unexpected interactions between protein kinases and kinase inhibitors, providing opportunities for identifying multi-targeted inhibitors of specific diverse kinases for drug repurposing and development. We have developed K-Map—a novel and user-friendly web-based program that systematically connects a set of query kinases to kinase inhibitors based on quantitative profiles of the kinase inhibitor activities. Users can use K-Map to find kinase inhibitors for a set of query kinases (obtained from high-throughput 'omics' experiments) or to reveal new interactions between kinases and kinase inhibitors for rational drug combination studies.

Availability and implementation: K-Map has been implemented in python scripting language and the website is freely available at: http://tanlab.ucdenver.edu/kMap.

Introduction

Protein kinases represent one of the largest 'druggable' and well-studied families in the human genome [1]. This class of proteins (kinome) plays a key role as regulators and transducers of signaling in eukaryotic cells. There is an estimated >500 members of the human kinome which can be classified into seven different kinase families based on their conserved catalytic domain sequences [2]. Kinases are relatively easy to target with small molecules and have been extensively studied at the biochemical, structural, and physiological levels. In cancer cells, some kinases are mutated and acquire oncogenic properties to drive tumorgenesis. Small molecules that inhibit these oncogenic kinases can effectively kill cancer cells, as demonstrated by the success story of imatinib (Gleevec®, Novartis, Basel, Switzerland) in inhibiting the activity of *BCR-ABL* in chronic myelogenous leukemia [3]. Imatinib also inhibits *KIT* and *PDGFRA*, which are commonly dysregulated in gastrointestinal stromal tumors [4]. The imatinib example illustrates that small-molecule kinase inhibitors interact with multiple protein kinase family members (*BCR-ABL*,

KIT, PDGFRA), and understanding these complex interactions between kinases and inhibitors could be useful for drug repurposing and development. These complex interactions could only be revealed by systematic interrogation of the small molecules across a large panel of kinases using quantitative assays (kinase activity profiles). Here, we have developed K-Map—a novel and user-friendly web-based program that systematically connects a set of query kinases to kinase inhibitors based on quantitative profiles of the kinase inhibitor activities. K-Map is motivated by the 'connectivity map' concept [5] where gene expression changes could be used as the 'universal language' to connect between biological systems, genes, and drugs. Instead of gene expression signatures, we used the kinase activity profiles as the 'language' for connecting kinases and small molecules in K-Map to reveal the complex interactions of kinases and inhibitors.

K-Map methods and features

Quantitative kinase inhibitor selectivity data sources

Two recently published comprehensive analyses of kinase inhibitor selectivity [6,7] were used to construct the K-Map reference database (kinase activity profiles). The first study systematically interrogates 178 commercially available inhibitors against a panel of 300 protein kinases using a radiometric phospho-transfer method to assess the percent

* Correspondence: aikchoon.tan@ucdenver.edu
[1]Translational Bioinformatics and Cancer Systems Biology Laboratory, Division of Medical Oncology, Department of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA
Full list of author information is available at the end of the article

kinase inhibition (IC_{50}) [6]. The second study measures inhibitor selectivity and potency of 72 inhibitors across 442 kinases using direct binding affinities between inhibitors and kinases (K_d) [7]. These kinase activity profiles were converted into rank-ordered lists according to their inhibitions and potencies against the kinases and used as the K-Map reference profiles for matching query kinases. For each study, the kinase activity profiles for individual drugs were converted into rank-ordered lists according to their inhibitions and potencies against the kinases. As a result, we generated two K-Map reference databases from these two studies: one for IC_{50} and the other one are for K_d. Both databases will be used to connect the query kinases and return the drugs in K-Map.

Pattern matching strategy

We implemented the K-Map pattern matching strategy based on the Kolmogorov-Smirnov (KS) statistics. The KS test is a nonparametric, rank-based pattern matching approach implemented in the connectivity map [5]. The query is a list of kinases, and the goal of the algorithm is to correlate kinase inhibitor that enriches the same kinases based on kinase inhibition profiles. For every inhibitor in the reference database, the KS statistic is computed and a connectivity score is defined.

Similar to the connectivity map approach [5], to compute the connectivity score, let N be the number of kinases in the reference database and M be the number of query kinases. For every drug in the reference database, we can compute the rank-ordered list R for all kinases $(1, 2, ..., N)$ based on the drug inhibitions and potencies against the kinases. For a list of query kinases of j, where $j = 1, 2, ..., M$, compute the following two values for each drug i in the reference database:

$$a = \max_{j=1}^{M} \left[\frac{j}{M} - \frac{R(j)}{N} \right]$$

and

$$b = \max_{j=1}^{M} \left[\frac{R(j)}{N} - \frac{(j-1)}{M} \right]$$

Let KS^i be the KS score for drug i,

$$KS^i = \begin{cases} a, & \text{if } a > b; \\ -b, & \text{if } b > a. \end{cases}$$

Finally, to compute the connectivity score (S^i) for drug i in a reference database, let $P = \max(KS^i)$ and $Q = \min(KS^i)$,

$$S^i = \begin{cases} \dfrac{KS^i}{P}, & \text{if } KS^i > 0; \\ -\left(\dfrac{KS^i}{Q}\right), & \text{if } KS^i < 0. \end{cases}$$

The connectivity score (S^i) for every drug is reported as the 'Score' in the results page. A positive score represents that the inhibitor has a similar rank order as the query kinases, indicating that the inhibitor is more specific in inhibiting the query kinases. A negative score represents that the inhibitor has a reverse rank order as the query, hence not specific in inhibiting the query kinases. Connectivity scores for each inhibitor were normalized to yield a score ranging 0 to 1, and inhibitors were ranked based on this normalized score. We also computed the running sum of the connectivity score for each inhibitor. The maximum value of the running sum is equivalent to the connectivity score of each inhibitor. Since the query kinases are unitless, K-Map can be applied to any technology platform.

Computing the permutation *p* value

To estimate the p value for each drug i, we perform a permutation test by randomly selecting the number of M instances from the rank-ordered drug i kinase profiles. Let $t = 1, 2, ..., T$ trials; the same procedure as computing the KS^i for drug i is performed T times and is denoted as KS_t^i. Let KS_0^i denote the actual KS^i for drug i. Count the number of times f where

$$\left[\left| KS_t^i \right| \geq \left| KS_0^i \right| \right]_{t=1}^{T}$$

is true. The frequency of this event (f/T) is estimated as a (two-sided) p value. This procedure is similar to the implementation of permutation test by the connectivity map [5]. The p value reported in the results page of K-Map is computed by 500 permutations.

Query features

K-Map implements three query functions: users can (1) directly enter query kinases in the query text box or upload a list of query kinases (Figure 1 (A)) in the K-Map tab, (2) select kinases from the kinase family (Figure 1 (B)) in the K-Map (by family) tab, or (3) query a set of kinases involved in certain biological processes according to Gene Ontology (Figure 1 (C)) in the K-Map (by GO) tab. Users also need to define which database they would like their query kinases to connect with (IC_{50} or K_d). All inhibitors available in the K-Map could be browsed under the Drug Info tab (Figure 1 (D)). Under the Download tab, users can search and download kinase-inhibitor relationships. The Help tab provides user guide to query and navigate the K-Map. The user manual for K-Map is available at http://tanlab.ucdenver.edu/kMap.

Connectivity results and linking features

The output of K-Map is a rank-ordered list of inhibitors based on the normalized connectivity scores, accompanied by p values and running sum plots. The 2D drug structure is viewable by scrolling through the drug name. Kinase
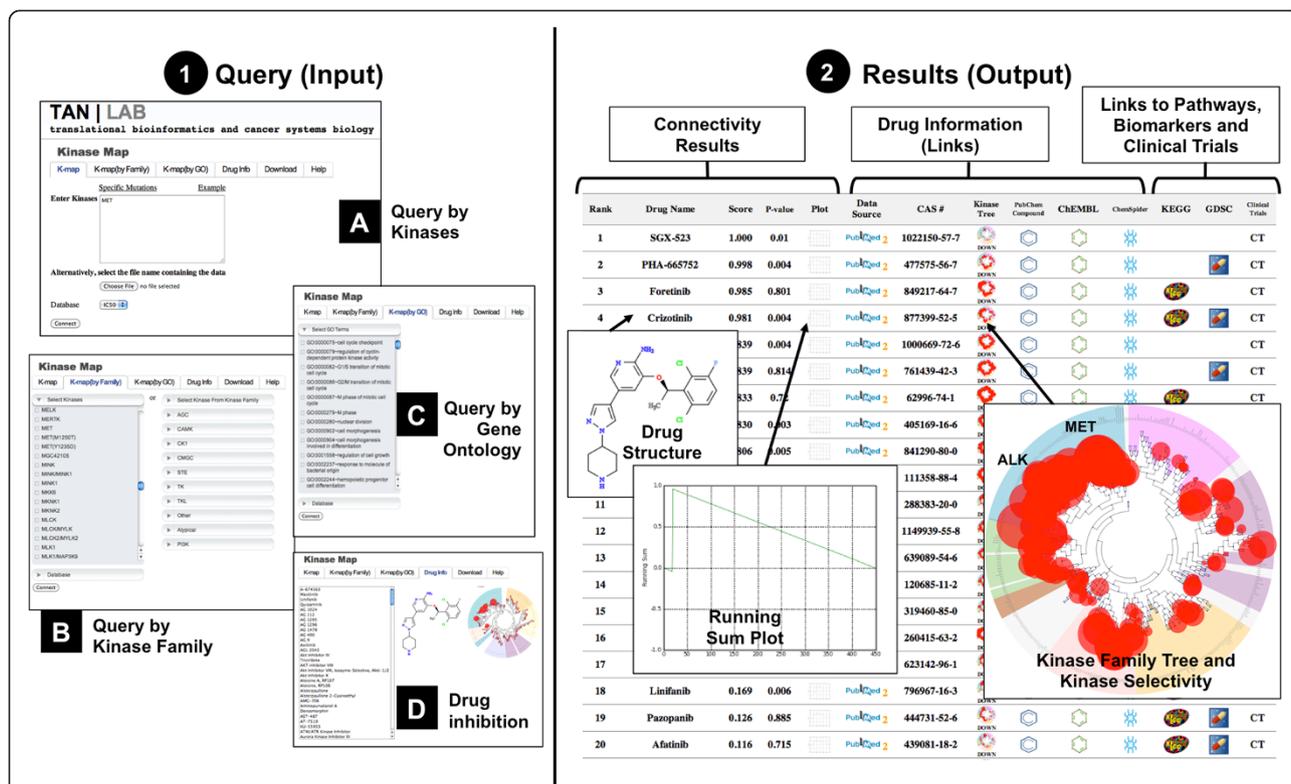
Figure 1 Query and results of the K-Map. Left, query features. K-Map could be queried by: **(A)** direct input or directly uploading a list of kinases in the K-Map tab, **(B)** by selecting kinases from the kinase family, or **(C)** querying by Gene Ontology Biological Processes. All drug information is available in the Drug Info tab **(D)** and could be downloaded in the Download tab. Right, connectivity results. Query kinases were connected to the inhibitors and sorted by normalized connectivity scores. Link-out features include PubMed, PubChem, ChEMBL, and ChemSpider for drug sources and information. K-Map also provides links to pathway, biomarker, and clinical trial information for each drug via KEGG, GDSC, and ClinicalTrials.gov databases, respectively.
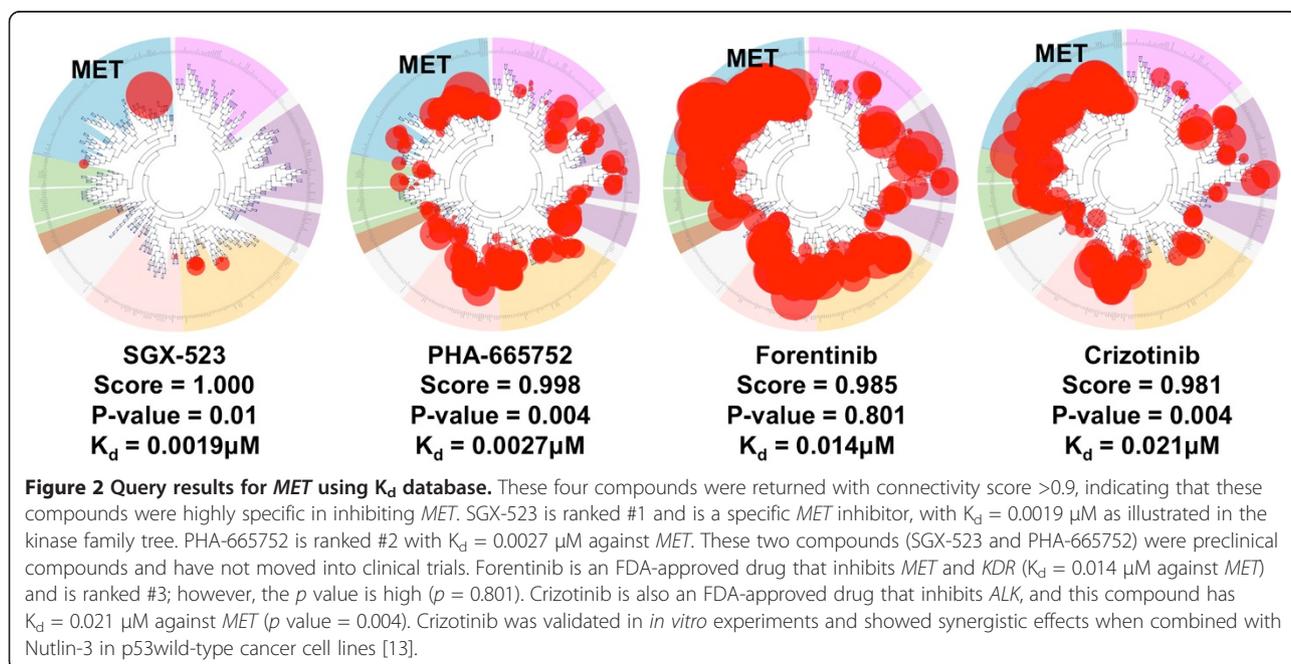
Figure 2 Query results for *MET* using K_d database. These four compounds were returned with connectivity score >0.9, indicating that these compounds were highly specific in inhibiting *MET*. SGX-523 is ranked #1 and is a specific *MET* inhibitor, with K_d = 0.0019 μM as illustrated in the kinase family tree. PHA-665752 is ranked #2 with K_d = 0.0027 μM against *MET*. These two compounds (SGX-523 and PHA-665752) were preclinical compounds and have not moved into clinical trials. Forentinib is an FDA-approved drug that inhibits *MET* and *KDR* (K_d = 0.014 μM against *MET*) and is ranked #3; however, the *p* value is high (p = 0.801). Crizotinib is also an FDA-approved drug that inhibits *ALK*, and this compound has K_d = 0.021 μM against *MET* (p value = 0.004). Crizotinib was validated in *in vitro* experiments and showed synergistic effects when combined with Nutlin-3 in p53wild-type cancer cell lines [13].

inhibitor specificity within the kinase family tree is generated under KinaseTree column where the red circles indicate degrees of inhibition. Linking features are available for data source of the kinase inhibition assay (via PubMed) and three major chemical databases (PubChem [8], ChEMBL [9], and ChemSpider (www.chemspider.com)). Additional links to drug pathway and drug biomarkers are available through the Kyoto Encyclopedia of Genes and Genomes (KEGG) [10] and Genomics of Drug Sensitivity in Cancer (GDSC) [11] databases, respectively. K-Map also provides link-out to ClinicalTrials.gov for ongoing or completed clinical trials of these inhibitors in various diseases. We plan to update the K-Map database every quarter to keep up with the new data and link-out information.

Implementation

K-Map is implemented in python (v2.6) and CGI script. The kinase family tree map and 2D drug structure are generated by the E.T.E. software (v2.0) and Open Babel (v2.3.1) [12], respectively. The K-Map website is freely available at: http://tanlab.ucdenver.edu/kMap.

K-Map application: case study

We have recently performed a genome-wide functional genetic screen to identify synthetic lethality genes for Nutlin-3 (p53 inhibitor) in p53 wild-type cancer cell lines [13]. From this screening, we identified *MET* as a synthetic lethal gene with Nutlin-3 in killing cancer cell. By querying *MET* in the K-Map using K_d database (Figure 1 (A)), four compounds were returned with connectivity score >0.9 (Figure 1,right side). All four compounds are specific in inhibiting *MET* with $K_d \leq 0.025$ μM (Figure 2). Interestingly, crizotinib, a recently FDA-approved *ALK* inhibitor is ranked #4 (*p* value = 0.004). As we demonstrated, treating p53 wild-type cancer cells with Nutlin-3 and crizotinib inhibits proliferation and enhances cell killing in *in vitro* experiments [13]. This supports the finding that the K-Map could reveal new inhibition of kinase inhibitor (Figure 1, kinase family tree indicates that crizotinib shows the highest selectivity in inhibiting *ALK* and *MET*).

Summary

K-Map is a novel and user-friendly web-based tool for connecting kinases with drugs based on quantitative profiles of the kinase inhibitor activities. Many kinase inhibitors could promiscuously inhibit multiple kinases due to conserved sequence similarity among kinase family members; we have exploited these complex and unexpected interactions between kinases and inhibitors as opportunities for drug repurposing and development. Users can use K-Map to search kinase inhibitors for a set of query kinases (obtained from high-throughput 'omics' experiments) or to reveal new interactions between kinases and kinase inhibitors for rational combination studies. In the future, we plan to extend K-Map by including more kinase inhibitor profiles. In summary, we believe that K-Map will be a valuable bioinformatics tool in connecting altered/mutated genes identified by next-generation sequencing with therapeutics, accelerating the process of personalized medicine.

Author details
[1]Translational Bioinformatics and Cancer Systems Biology Laboratory, Division of Medical Oncology, Department of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA. [2]Data Mining and Information Systems Laboratory, Department of Computer Science and Engineering, Korea University, Seoul 136-713, South Korea.

References
1. Cohen P: Protein kinases—the major drug targets of the twenty-first century. *Nat Rev Drug Discov* 2002, **1**(4):309–315.
2. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S: The protein kinase complement of the human genome. *Science* 2002, **298**(5600):1912–1934.
3. Druker BJ, Talpaz M, Resta DJ, Peng B, Buchdunger E, Ford JM, Lydon NB, Kantarjian H, Capdeville R, Ohno-Jones S, Sawyers CL: Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. *N Engl J Med* 2001, **344**:1031–1037.
4. Demetri GD, von Mehren M, Blanke CD, Van den Abbeele AD, Eisenberg B, Roberts PJ, Heinrich MC, Tuveson DA, Singer S, Janicek M, Fletcher JA, Silverman SG, Silberman SL, Capdeville R, Kiese B, Peng B, Dimitrijevic S, Druker BJ, Corless C, Fletcher CD, Joensuu H: Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. *N Engl J Med* 2002, **347**:472–480.
5. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, Reich M, Hieronymus H, Wei G, Armstrong SA, Haggarty SJ, Clemons PA, Wei R, Carr SA, Lander ES, Golub TR: The connectivity map: using gene-expression signatures to connect small molecules, genes and disease. *Science* 2006, **131**(5795):1929–1935.
6. Anastassiadis T, Deacon SW, Devarajan K, Ma H, Peterson JR: Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. *Nat Biotechnol* 2011, **29**(11):1039–1045.
7. Davis MI, Hunt JP, Herrgard S, Ciceri P, Wodicka LM, Pallares G, Hocker M, Treiber DK, Zarrinkar PP: Comprehensive analysis of kinase inhibitor selectivity. *Nat Biotechnol* 2011, **29**(11):1046–1051.
8. Bolton E, Wang Y, Thiessen PA, Bryant SH: PubChem: Integrated platform of small molecules and biological activities. In *Annual Reports in Computational Chemistry*. 4th edition. Washington, DC: American Chemical Society; 2008.
9. Gaulton A, Bellis LJ, Bento AP, Chambers J, Davies M, Hersey A, Light Y, McGlinchey S, Michalovich D, Al-Lazikani B, Overington JP: ChEMBL: a large-scale bioactivity database for chemical biology and drug discovery. *Nucleic Acids Res* 2011, **40**(Database issue):D1100–1107.

10. Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M: **KEGG for integration and interpretation of large-scale molecular datasets.** *Nucleic Acids Res* 2012, **40:**D109–D114.

11. Yang M, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, Bindal N, Beare D, Smith JA, Thompson IR, Ramaswamy S, Futreal PA, Haber DA, Stratton MR, Benes C, McDermott U, Garnett MJ: **Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells.** *Nucleic Acids Res* 2013, **41**(Database issue):D955–961.

12. O'boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR: **Open Babel: an open chemical toolbox.** *J Cheminform* 2011, **3:**33.

13. Sullivan KD, Padilla-Just N, Henry RE, Porter CC, Kim J, Tentler JJ, Eckhardt SG, Tan AC, DeGregori J, Espinosa JM: **ATM and MET kinases are synthetic lethal with nongenotoxic activation of p53.** *Nat Chem Biol* 2012, **8**(7):646–654.