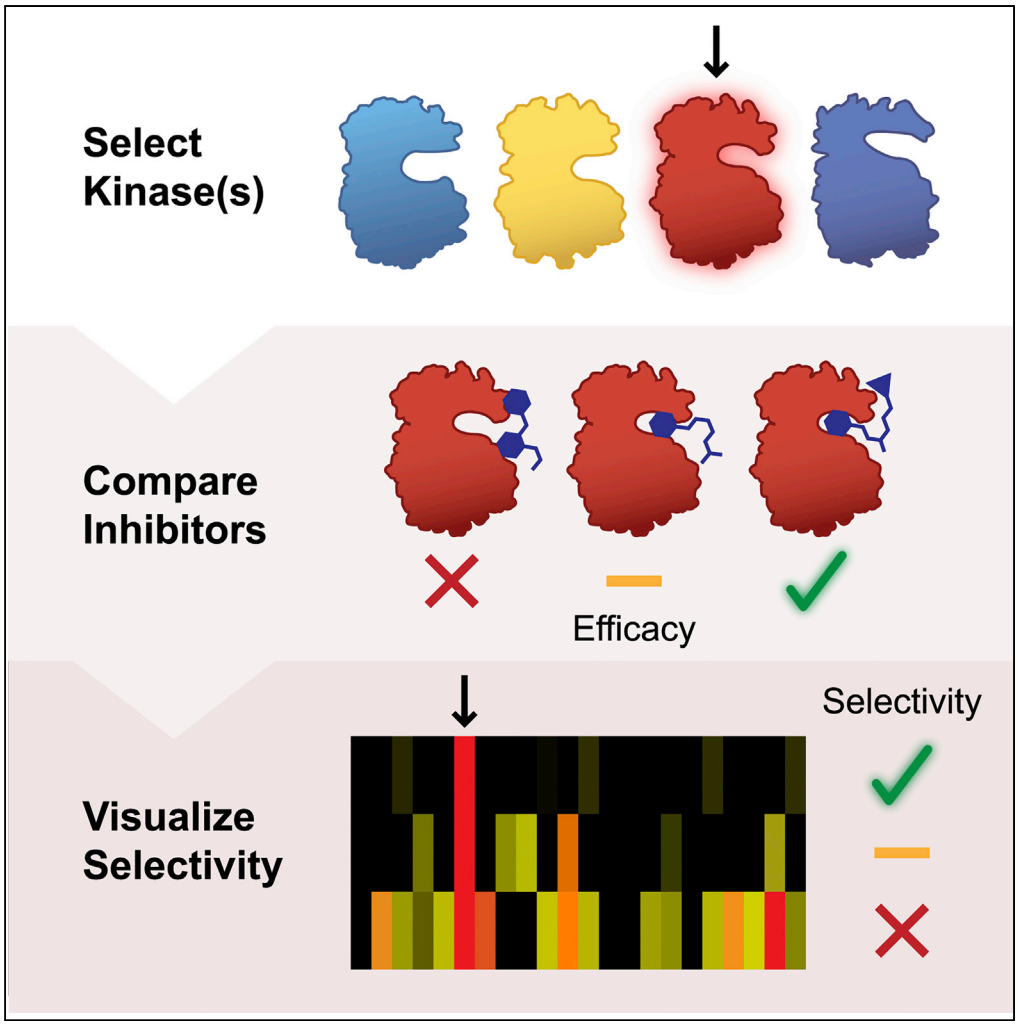
iScience

CellPress

Article

KInhibition: A Kinase Inhibitor Selection Portal



Thomas Bello, Taranjit S. Gujral

tgujral@fredhutch.org

HIGHLIGHTS

An easy-to-use tool that compiles datasets from kinase inhibitor screens

A KInhibition Selectivity Score quantifies each compound's selectivity

Users can select kinases, view compounds and selectivity, and download the results

KInhibition is broadly applicable across life science disciplines

Bello & Gujral, iScience *8*, 49–53
October 26, 2018
https://doi.org/10.1016/j.isci.2018.09.009

iScience

CellPress

Article

KInhibition: A Kinase Inhibitor Selection Portal

Thomas Bello[1,2] and Taranjit S. Gujral[1,2,3,*]

SUMMARY

Protein kinases constitute a large class of signaling molecules frequently targeted in research and clinical uses. However, kinase inhibitors are notoriously non-specific, making it difficult to select an appropriate inhibitor for a given kinase. Available data from large-scale kinase inhibitor screens are often difficult to query. Here, we present KInhibition (https://kinhibition.fredhutch.org), an online portal that allows users to search publicly available datasets to find selective inhibitors for a chosen kinase or group of kinases. Compounds are sorted by a KInhibition Selectivity Score, calculated based on compounds' activity against the selected kinase(s) versus activity against all other kinases for which that compound has been profiled. The current version allows users to query four datasets, with a framework that can easily accommodate additional datasets. KInhibition represents a powerful platform through which researchers from broad areas of biology, chemistry, and pharmacology can easily interrogate large datasets to help guide their selection of kinase inhibitors.

INTRODUCTION

A fundamental aspect of cell biology is the study of cellular signaling, the process by which cells sense their surroundings, respond to environmental cues, and transfer information (Downward, 2001). Both clinicians and researchers rely on the ability to selectively perturb the function of specific signaling molecules, often by using small molecule inhibitors (Jin et al., 2014). In particular, kinases represent a large class of proteins that are key mediators of signaling pathways and important targets for research and therapy (Wu et al., 2015). In the last 10 years alone, there have been more than 1.5 million publications on kinases, and countless small-molecule inhibition studies spanning the majority of the >500 kinases in the human kinome (Wu et al., 2015), underscoring the central role of kinase signaling and inhibition in molecular and cellular biology.

Despite the enormity of research that has been done on kinase signaling, there remains a confounding challenge when attempting to selectively inhibit a desired molecule. Mainly due to the high structural similarity among kinases, nearly all available small-molecule kinase inhibitors exhibit some promiscuity, causing undesired "off-target" effects (Anastassiadis et al., 2011; Bain et al., 2007; Davis et al., 2011; Karaman et al., 2008). Even many compounds described as being "specific" or "selective" have confounding off-target effects, making the selection, use, and analysis of the appropriate kinase inhibitor difficult. A number of large-scale kinase inhibitor screens have been undertaken in an attempt to quantify these effects (Anastassiadis et al., 2011; Dranchak et al., 2013; Duong-Ly et al., 2016; Gao et al., 2013; Klaeger et al., 2017; Koleti et al., 2017), and many of these results are publicly available. However, these data are decentralized and difficult to query, with results often being spread across multiple data files that must be downloaded and opened individually. Furthermore, there are multiple methods employed for representing or quantifying "selectivity" (Anastassiadis et al., 2011; Cheng et al., 2010; Davis et al., 2011; Graczyk, 2007; Karaman et al., 2008; Klaeger et al., 2017; Uitdehaag and Zaman, 2011), which may yield conflicting results. Thus, the ultimate challenge of identifying the right kinase inhibitor for a biological task remains unresolved. Here we present KInhibition, a powerful platform tool that allows researchers to search through multiple kinase inhibitor screens, visualize the relevant data, and choose the most selective and appropriate kinase inhibitor for the task at hand. We anticipate that KInhibition will be adopted by the broader research community equivalent to RNAi target sequence or CRISPR guide RNA selection tools.

RESULTS

The Theory behind KInhibition

KInhibition is a platform tool designed to answer the question "Which compound should be used to inhibit a chosen kinase or pathway?" The first, but often the most critical, step in answering this question is to

[1]Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

[2]Department of Molecular and Cellular Biology, University of Washington, Seattle, WA 98195-7275, USA

[3]Lead Contact

*Correspondence: tgujral@fredhutch.org

https://doi.org/10.1016/j.isci.2018.09.009

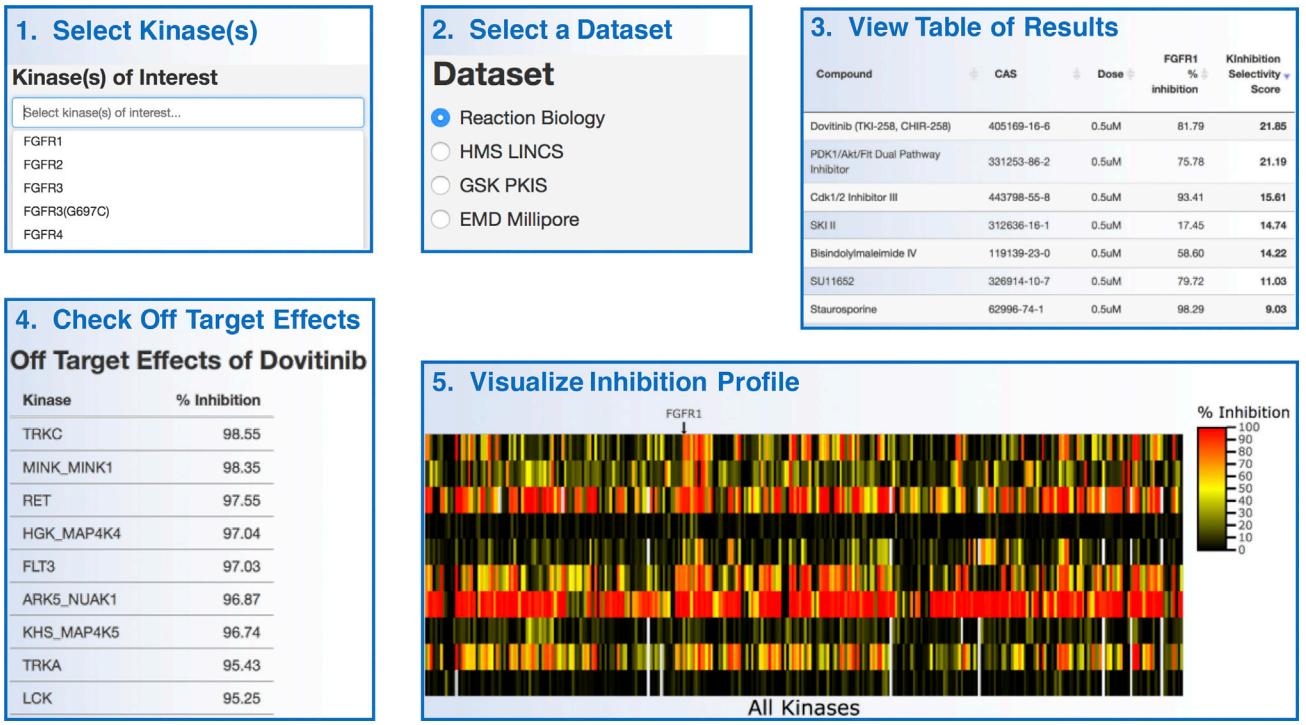


Figure 1. Using KInhibition

Users first select a kinase or group of kinases that they wish to inhibit, and then select a dataset in which to search for compounds. The table of compounds updates and sorts compounds based on the calculated KInhibition Selectivity Score. Clicking on a row opens a second table below, which displays off-target effects of that compound. The Heatmap tab displays the full inhibition profiles across all kinases for the compounds displayed in the first page of the Table of Results. The table, and consequently the heatmap, can be reordered by clicking on any of the column headers.

locate and format the relevant data from the publicly available kinase inhibitor screens. These datasets may be initially found in somewhat intractable formats, but nearly all of them can be summarized by a matrix of drug-target interactions, with rows representing the compounds tested, columns representing the kinases screened, and entries being that compound's effect on that kinase. The main requirement for a dataset to be included in KInhibition is for it to be formatted in this manner, making the addition of future datasets to this platform a relatively trivial task.

A unified data format allows us to address the more nuanced issue of how to quantify the "selectivity" of a given inhibitor. Although there have been many proposed metrics for selectivity (Anastassiadis et al., 2011; Cheng et al., 2010; Graczyk, 2007; Klaeger et al., 2017; Uitdehaag and Zaman, 2011), very few are computationally robust enough to apply across datasets. We developed a "KInhibition Selectivity Score," which quantifies the selectivity of a compound based on its on-target inhibition ("inhibition score") and its off-target effects ("inhibition penalty"). The inhibition score is simply the inhibition of the selected kinase, or a geometric mean if multiple kinases are chosen. The inhibition penalty is further divided into two sub-penalties. The first quantifies the broad inhibition activity of a compound, and will therefore best account for the extreme case in which a compound inhibits nearly every kinase tested, but to a small degree (e.g., 10% of control) compared with the intended target. The second sub-penalty specifically quantifies off-target effects that are close in magnitude to the inhibition of the intended target. This accounts for the other extreme case, in which a compound inhibits only a few (e.g., 2–10) off-target kinases, but with a magnitude comparable to or greater than the intended target. Both these extremes represent distinct types of off-target effects that must be considered when choosing the appropriate kinase inhibitor for an experiment.

The KInhibition Selectivity Score has numerous advantages that give the user the most relevant information and the most control over the decision of which compound to use. First, it is designed to work with percent-of-control data, rather than binding coefficients (K_i) or IC_{50} values, like previously reported partition indices

Dataset	Compounds Tested	Compound + Dose Combinations	Kinases Screened	Pairwise Coverage	Reference
Reaction Biology[a]	178	178	300	98.9%	Anastassiadis et al., 2011
HMS LINCS	121	134	471	88.5%	Koleti et al., 2017
GSK PKIS	367	734	224	99.9%	Dranchak et al., 2013
EMD Millipore	128	255	234	100%	Gao et al., 2013

Table 1. Informational Summary of Datasets Currently Included in KInhibition

[a]An updated version of the Reaction Biology dataset will be included in this portal as soon as it is publicly available. Currently, the updated version contains 385 compounds, 427 compound + dose combinations, 298 kinases screened, and 99.5% pairwise coverage.

or entropy-based scores (Uitdehaag and Zaman, 2011). This allows it to be used for screens performed even at a single dose, treating additional doses as separate compounds or entries in the matrix. Second, this score quantifies selectivity for a user-defined set of up to 10 "on-target" kinases, rather than simply basing all calculations on the most inhibited kinase for each compound. Third, as previously mentioned, this score accounts for both the number and the magnitude of off-target effects better than previous scores (such as the Gini coefficient), allowing researchers to select the inhibitor most suited to their needs. Finally, this scoring metric does not rely on any hard-coded values, arbitrary thresholds, or data binning, giving it an advantage over S(x) scores or Ambit scores (Cheng et al., 2010).

Using KInhibition

The KInhibition app is run entirely in-browser and does not require the user to upload or download any data or files. It can be found at https://kinhibition.fredhutch.org/. Upon loading the app webpage, users select a kinase or a group of kinases they wish to inhibit (Figure 1). Kinase names are standardized across all datasets, as described toward the bottom of the "Datasets" tab. After selecting the kinase(s), the "Table of Results" tab will automatically update to list the inhibitors in the first available dataset, sorted based on their KInhibition Selectivity Score for the chosen kinase(s). A set of radio buttons will also appear to allow the user to choose between all the datasets that include their chosen kinase(s). The data from each dataset are kept separate and must be searched one at a time, as each screen is done using different experimental conditions, kinase panels, and compound doses, and thus results may not be comparable enough to simply merge the datasets. Details about each dataset can be found in the Datasets tab (Table 1). The Table of Results can be searched using the search box in the top right, can display more or fewer compounds per page using the drop-down menu in the top left, and can be sorted based on the values in any of the columns by clicking on the column header. Furthermore, clicking on any row generates a new table below this one, which lists the significant off-targets of that compound. "Significant off-targets" are defined as kinases inhibited at least half the amount of the chosen kinase (or half of the geometric mean for multiple chosen kinases).

Finally, after exploring the Table of Results, users can click on the "Heatmap" tab to load a heatmap of the inhibition profiles for the compounds displayed in the Table of Results. The compounds included in the heatmap will mirror those in the first page of the Table of Results, including any changes to the number or order of compounds in this table. Inhibition is represented as a color spectrum from black (no inhibition) to yellow (moderate inhibition) to red (maximal inhibition) (Figure S1). Users can mouse over the heatmap to view the details of any individual point, or click and drag to selectively zoom in on a portion (double-clicking zooms out to the full heatmap). This heatmap, along with all the previously mentioned tables, can be downloaded using buttons below each element.

DISCUSSION

Kinases remain one of the few classes of biomolecules whose function (as opposed to simply concentration or abundance) can be easily detected, quantified, and perturbed. Kinases therefore sit at a critical intersection between basic research and clinical applications, making data on kinase inhibitors a lucrative asset in both academia and life science industries. With the above functionality, KInhibition fills a much needed role in modern cell biology by allowing researchers to make data-driven decisions regarding kinase inhibitors. By following the aforementioned steps, researchers can easily find and choose the most selective and

appropriate compound for their particular target or pathway. Due to the robust and careful design of the app, this platform can be easily updated and expanded to include additional datasets and information as they become available. We therefore expect this portal to see broad use and adoption akin to other selection tools.

Limitations of This Study

The KInhibition platform, and the associated KInhibition Selectivity Score were designed to best leverage the currently available data. However, it should be noted that the KInhibition Selectivity Score and all other metrics listed in this tool are based only on a single dose of the compound used. The efficacy, selectivity, and off-target effects of a given compound depend heavily on the concentration used in the actual experiment, as well as on the biological context in which the compound is applied. Thus, the information presented in the Table of Results (i.e., Percent Inhibition) may not directly translate to a cellular or *in vivo* context. Therefore, the goal of this portal, and the datasets included in it, is to obtain a qualitative assessment of selectivity, using the quantitative metrics presented as data-driven guidelines for decision making in the context of past experience and other pharmacologic properties of the compounds in question (i.e., solubility, bioavailability, and metabolism).

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND SOFTWARE AVAILABILITY

The app portal can be accessed at https://kinhibition.fredhutch.org. The source code and all other files can be found at the Github repository listed in the Key Resources Table.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods and one figure and can be found with this article online at https://doi.org/10.1016/j.isci.2018.09.009.

ACKNOWLEDGMENTS

This study was supported by grants from the NIH (K22CA201229) and the Pacific Northwest Prostate Cancer SPORE (P50CA097186-14). T.B. is the recipient of the Fred Hutchinson Cancer Research Center Interdisciplinary Dual Mentor Fellowship in Cancer Research. We thank Dirk Peterson, Robert McDermott, and the Fred Hutch Scientific Computing Core for their help and support throughout this project, and Anna Russel for graphic design assistance.

AUTHOR CONTRIBUTIONS

T.S.G. and T.B. conceived the study. T.B. performed all the analysis and developed Kinhibition application. T.S.G. and T.B. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 8, 2018
Revised: August 21, 2018
Accepted: September 10, 2018
Published: October 26, 2018

REFERENCES

Anastassiadis, T., Deacon, S.W., Devarajan, K., Ma, H., and Peterson, J.R. (2011). Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. Nat. Biotechnol. *29*, 1039–1045.

Bain, J., Plater, L., Elliott, M., Shpiro, N., Hastie, C.J., Mclauchlan, H., Klevernic, I., Arthur, J.S.C., Alessi, D.R., and Cohen, P. (2007). The selectivity of protein kinase inhibitors: a further update. Biochem. J. *408*, 297–315.

Cheng, A.C., Eksterowicz, J., Geuns-Meyer, S., and Sun, Y. (2010). Analysis of kinase inhibitor selectivity using a thermodynamics-based partition index. J. Med. Chem. *53*, 4502–4510.

Davis, M.I., Hunt, J.P., Herrgard, S., Ciceri, P., Wodicka, L.M., Pallares, G., Hocker, M., Treiber, D.K., and Zarrinkar, P.P. (2011). Comprehensive analysis of kinase inhibitor selectivity. Nat. Biotechnol. *29*, 1046–1051.

Downward, J. (2001). The ins and outs of signalling. Nature *411*, 759–762.

Dranchak, P., MacArthur, R., Guha, R., Zuercher, W.J., Drewry, D.H., Auld, D.S., and Inglese, J. (2013). Profile of the GSK published protein kinase inhibitor set across ATP-dependent and-independent luciferases: implications for reporter-gene assays. PLoS One *8*, e57888.

Duong-Ly, K.C., Devarajan, K., Liang, S., Horiuchi, K.Y., Wang, Y., Ma, H., and Peterson, J.R. (2016). Kinase inhibitor profiling reveals unexpected opportunities to inhibit disease-associated mutant kinases. Cell Rep. *14*, 772–781.

Gao, Y., Davies, S.P., Augustin, M., Woodward, A., Patel, U.A., Kovelman, R., and Harvey, K.J. (2013). A broad activity screen in support of a chemogenomic map for kinase signalling research and drug discovery. Biochem. J. *451*, 313–328.

Graczyk, P.P. (2007). Gini coefficient: a new way to express selectivity of kinase inhibitors against a family of kinases. J. Med. Chem. *50*, 5773–5779.

Jin, L., Wang, W., and Fang, G. (2014). Targeting protein-protein interaction by small molecules. Annu. Rev. Pharmacol. Toxicol. *54*, 435–456.

Karaman, M.W., Herrgard, S., Treiber, D.K., Gallant, P., Atteridge, C.E., Campbell, B.T., Chan, K.W., Ciceri, P., Davis, M.I., Edeen, P.T., et al. (2008). A quantitative analysis of kinase inhibitor selectivity. Nat. Biotechnol. *26*, 127–132.

Klaeger, S., Heinzlmeir, S., Wilhelm, M., Polzer, H., Vick, B., Koenig, P.-A., Reinecke, M., Ruprecht, B., Petzoldt, S., Meng, C., et al. (2017). The target landscape of clinical kinase drugs. Science *358*, eaan4368.

Koleti, A., Terryn, R., Stathias, V., Chung, C., Cooper, D.J., Turner, J.P., Vidović, D., Forlin, M., Kelley, T.T., D'Urso, A., et al. (2017). Data portal for the library of integrated network-based cellular signatures (LINCS) program: integrated access to diverse large-scale cellular perturbation response data. Nucleic Acids Res. *46* (D1), D558–D566.

Uitdehaag, J.C., and Zaman, G.J. (2011). A theoretical entropy score as a single value to express inhibitor selectivity. BMC Bioinformatics *12*, 94.

Wu, P., Nielsen, T.E., and Clausen, M.H. (2015). FDA-approved small-molecule kinase inhibitors. Trends Pharmacol. Sci. *36*, 422–439.

ISCI, Volume *8*

Supplemental Information

KInhibition: A Kinase Inhibitor Selection Portal

Thomas Bello and Taranjit S. Gujral

SUPPLEMENTAL FIGURES:

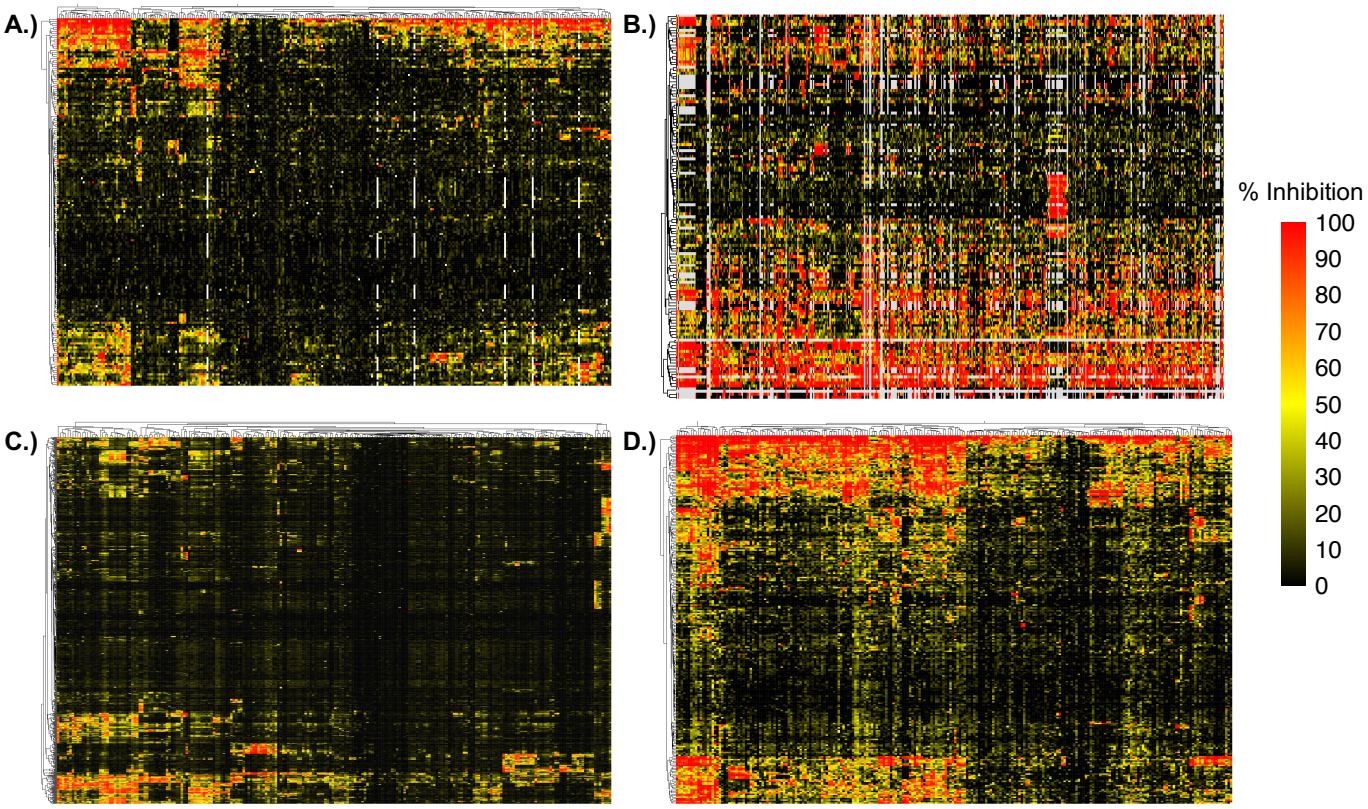


Figure S1. Heatmaps showing the inhibition profiles of all the compounds in each dataset, related to Table 1. Rows represent the compounds tested, columns represent the kinases screened. White boxes denote missing data. A, Reaction Biology dataset; B, HMS LINCS dataset; C, GSK PKIS dataset; D, EMD Millipore dataset.

TEST PROCEDURES AND VIGNETTE:

Finding a selective inhibitor of EGFR and ErbB family kinases, related to Figure 1.

To begin, navigate to https://kinhibition.fredhutch.org in a modern desktop browser (while mobile browsers may load the page, some functionality may be lost). Chrome, Firefox, and Safari have all been tested for compatibility. On loading, the sidebar on the left will have two entry fields, “Kinase(s) of Interest” and “Dataset”, as well as brief instructions for use. The first step is to select a kinase or group of kinases to inhibit. In this vignette, we will search for inhibitors of EGFR and other ErbB family members. Click on the kinase selection field (labeled with the placeholder text “Select Kinase(s) of interest…”). This will open another drop-down menu with all of the kinases profiled in the chosen dataset. Select “EGFR” by either scrolling down to find it, or by typing the first few letters into the field to filter the results, then clicking on “EGFR” in the drop-down menu. Note that you must click the option even if you have typed the entire name into the field in order for the tool to register your selection. Once EGFR is selected, the Table of Results to the right will automatically populate with the top ten selective inhibitors from the first available dataset (in this case, “Reaction Biology”), initially sorted by KInhibition Selectivity Score (KISS). Additionally, a set of radio buttons will appear to allow you to select any dataset that includes the selected kinase(s) (descriptions of each dataset can be found by navigating to the “Datasets” tab). For now, let us keep “Reaction Biology” selected. If it has changed, click on the “Reaction Biology” button.

In the Table of Results, we see that the top result is “EGFR Inhibitor”, with a KISS of 98.72, followed by PD 174265 with a KISS of 88.37. To the right side of the table, we notice that PD 174265 has one off-target effect that is greater than half of its inhibition of EGFR, while EGFR inhibitor has none. Note that you may need to scroll to the right to see the whole table, depending on the size of your browser window. To check this off target effect, click on the second row of the Table of Results. This will open a second table below the “Table of Results” that displays the off-target results of the selected compound (you may need to scroll down to see this table). Here, we see that PD 174265 also has significant inhibition on ErbB2 that is considered “off-target” and is lowering its KISS. To get a better idea of the full inhibition profiles of these compounds, click on the “Heatmap” tab towards the top of the screen. Note that the heatmap may take a few seconds to load, and may require scrolling to the right to see the entire image. From this heatmap, we can see that EGFR Inhibitor is, indeed, highly selective for EGFR at 0.5 μM. The white tiles represent areas of missing data, so we can see that EGFR Inhibitor was not profiled for 6 kinases in this dataset. If we hover over these tiles, a pop-up window reveals that these kinases are (from left to right) EPHA6, MYO3B, PRKACG, STK38, TLK2, and ULK3.

From the heatmap, we can see that Lapatinib (row 3) also has high inhibition of EGFR, but also inhibits two other kinases. Hovering the mouse over these bright red tiles reveals that these are ErbB2 and ErbB4. To modify our search to

include these as on-target inhibitions, first navigate back to the “Table of Results” tab. Now, click on the “Kinase(s) of Interest” field on the left again and, without removing EGFR, add “ERBB2” and “ERBB4” the same way as before (either by scrolling to find them, or typing the first few characters, then clicking on the desired kinase). With these selected, we see that Lapatinib is now the top result with a KISS of 96.31. Use the drop-down menu to the top-left of the table to change “Show 10 entries” to “Show 20 entries”. Towards the bottom of this table, we see that “EGFR Inhibitor” now has a KISS of 10.58 due to its low inhibition of ErbB2 and ErbB4. If we desire an inhibitor that has higher inhibition of ErbB4 specifically, we can click on the column name “ERBB4 % Inhibition” twice to reorder the compounds based on their inhibition of ErbB4. The top result, Bosutinib, does not appear very selective (KISS = 27.42). However, the second result is a compound named “EGFR/ErbB2/ErbB4 Inhibitor” (KISS = 87.50). Once again, we can click on this row to generate a table of the 7 notable off-target effects.

To get a well-rounded view of which inhibitor to use, it is a good idea to search multiple datasets. With the same group of kinases selected (EGFR, ERBB2, ERBB4), click on the “HMS LINCS” button under the “Datasets” header on the left. When the table loads, we see that, in this dataset, Lapatinib (10μM) has a KISS of 88.37. Clicking on the second row to display the off-target effects of Lapatinib (10μM) shows that this dataset also includes many mutant forms of EGFR that are significantly inhibited by Lapatinib. Since these are counted as off-targets, they will lower the KISS. If we wish to get a better idea of Lapatinib’s score, we can add each of these mutant forms to our “Kinase(s) of Interest). If we wish to search for a particular subset of compounds, we can type “AZ” in the search field to the top right of the table to restrict the compounds displayed to only those with “AZ” in their names. After doing this, clicking on the “Heatmap” tab will display a heatmap of only these compounds. Finally, all of these outputs (Table of Results, off-target effect table, and heatmap) can be downloaded by clicking on the buttons below each object.

TRANSPARENT METHODS:

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
App script	This paper	https://github.com/FredHutch/KInhibition-public (app.R)
Reaction Biology Formatted Dataset	(Anastassiadis et al., 2011)	https://github.com/FredHutch/KInhibition-public (rbio_old_dataset.csv)
HMS LINCS Formatted Dataset	(Koleti et al., 2017)	https://github.com/FredHutch/KInhibition-public (LINCS_dataset.csv)
GSK PKIS Formatted Dataset	(Dranchak et al., 2013)	https://github.com/FredHutch/KInhibition-public (PKIS_dataset.csv)
EMD Millipore Formatted Dataset	(Gao et al., 2013)	https://github.com/FredHutch/KInhibition-public (EMD_dataset.csv)
Software and Algorithms		
R 3.3.0	(R Core Team, 2016)	https://www.r-project.org/
Shiny	(Chang et al., 2017)	https://cran.r-project.org/package=shiny
Shiny Semantic	(Stachura, 2018)	https://cran.r-project.org/package=shiny.semantic
DT: ‘DataTables’ Wrapper	(Xie, 2018)	https://cran.r-project.org/package=DT
dplyr	(Wickham et al., 2017)	https://cran.r-project.org/package=dplyr
reshape2	(Wickham, 2007)	http://www.jstatsoft.org/v21/i12/
ggplot2	(Wickham, 2009)	http://ggplot2.org
webshot	(Chang, 2017)	https://cran.r-project.org/package=webshot
HTMLwidgets	(Vaidyanathan et al., 2018)	https://cran.r-project.org/package=htmlwidgets
plotly	(Sievert et al., 2017)	https://cran.r-project.org/package=plotly

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Taran Gujral (tgujral@fredhutch.org).

METHOD DETAILS

Calculating the KInhibition Selectivity Score

The KInhibition Selectivity Score, $KISS$, quantifies the selectivity of a given compound screened at a specific dose against a large panel of kinases. The data from such screens is treated as an $m \times n$ drug-target interaction matrix, where m is the number of compounds (with different doses of the same compound being treated as separate compounds for this purpose) and n is the number of kinases screened. All entries in this matrix fall in the range $[0, 100]$ and represent the inhibition of the kinase by that compound. This is "percent of control" data, with 0 being no inhibition (activity of the kinase equal to or greater than the control), and 100 being complete inhibition (no detected kinase activity in the presence of that compound).

Before computing the score, each row in the drug-target interaction matrix is scaled so that the maximum of every row is exactly 100. This decouples compounds' selectivity from their efficacy to avoid unnecessarily penalizing compounds that are highly selective, but display low absolute inhibition at the dose tested. Efficacy at the tested dose can still be determined in the final results table, which displays the unscaled inhibition values.

The following calculations are computed row-wise on the scaled drug-target matrix, so that each compound has an assigned selectivity score based solely on the properties of that compound, independent of the other compounds (rows) in the matrix. The non-missing elements of each row, $z_1 \dots z_n$, can be partitioned into on-target effects, $x_1 \dots x_k$, and off-target effects, $y_1 \dots y_{n-k}$. First, the Inhibition Score, IS, is computed from the k chosen on-target inhibitions $x_1 \dots x_k$ as a geometric mean:

$$IS = \sqrt[k]{\prod_{i=1}^{k} x_i}$$

We used a geometric mean to best represent compounds that inhibit only some of the chosen on-target kinases, as a geometric mean is always less than or equal to an arithmetic mean. This value thus lies in the interval of the scaled data:

$$0 \leq IS \leq 100$$

The Inhibition Penalty, IP, is then computed in two parts. The first part, IP_1, is computed as an arithmetic mean of the off-target effects $y_1 \dots y_{n-k}$:

$$IP_1 = \left(\frac{1}{n-k}\right) \sum_{i=1}^{n-k} y_i$$

This portion of the penalty represents the baseline inhibition across all of the off-target kinases, which is useful for penalizing compounds with extremely broad activity. However, it poorly accounts for compounds that have a small number (i.e. <10) of very large (near 100% inhibition) off-target effects compared to the on-target effects. To better capture this aspect, IP_2 is computed as a ratio of normalized variances between the off-target effects $y_1 \dots y_{n-k}$ and the total effects $z_1 \dots z_n$ as follows. First, the variances are computed as

$$s_y^2 = \left(\frac{1}{n-k-1}\right) \sum_{i=1}^{n-k} (y_i - \bar{y})$$

$$s_z^2 = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} (z_i - \bar{z})$$

where $\bar{y} = \left(\frac{1}{n-k}\right) \sum_{i=1}^{n-k} y_i$ and $\bar{z} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} z_i$ are the arithmetic means of the off-target and total inhibitions, respectively. When $k \ll n$, we expect these variances to be approximately equal for all but the most selective compounds, whose total variances will instead be explained largely by the difference between the on- and off-target effects. An approximate lower bound for the off-target variance can be computed using a modified version of the von Szokefalvi Nagy Inequality:

$$s_{lo}^2 = \frac{\max(y_i)^2}{2n}$$

The variances can then be centered using this lower bound in order to shift IP_2 to have a minimum closer to 0, and IP_2 is calculated as the ratio between the two shifted variances:

$$IP_2 = \frac{s_y^2 - s_{lo}^2}{s_z^2 - s_{lo}^2}$$

While the modifications to the lower bound (namely, using n rather than $n-k$ in the denominator) mean that IP_2 can fall outside the interval of $[0, 1]$, in practice this only occurs for the most selective compounds, and then only by a very small margin. We then combine the two penalties in an empirically determined manner:

$$IP = \frac{IP_1 + 100(IP_2)^5}{2}$$

With the rare exception mentioned above, both terms in the numerator of this equation lie in the interval $[0, 100]$, and therefore their arithmetic mean IP does as well. Finally, $KISS$ is calculated as a different between IS and IP

$$KISS = IS - IP$$

Based on the bounds mentioned above, $KISS$ generally lies in the interval $[-100, 100]$, with 100 representing near perfect selectivity (i.e. only on-target inhibition), and -100 representing the opposite (i.e. only off-target inhibition).

QUANTIFICATION AND STATISTICAL ANALYSIS

All software development, calculations, and analyses were carried out using R 3.3.0 (https://www.r-project.org/). All packages used can be found in the Key Resources Table.

DATA SOFTWARE AND AVAILABILITY

The app portal can be accessed at https://kinhibition.fredhutch.org. The source code and all other files can be found at the Github repository listed in the Key Resources Table.

SUPPLEMENTAL REFERENCES:

Anastassiadis, T., Deacon, S.W., Devarajan, K., Ma, H., and Peterson, J.R. (2011). Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity. Nat. Biotechnol. *29*, 1039–1045.

Chang, W. (2017). webshot: Take Screenshots of Web Pages.

Chang, W., Cheng, J., Allaire, J.J., Xie, Y., and McPherson, J. (2017). shiny: Web Application Framework for R.

Dranchak, P., MacArthur, R., Guha, R., Zuercher, W.J., Drewry, D.H., Auld, D.S., and Inglese, J. (2013). Profile of the GSK Published Protein Kinase Inhibitor Set Across ATP-Dependent and-Independent Luciferases: Implications for Reporter-Gene Assays. PLoS One *8*, e57888.

Gao, Y., Davies, S.P., Augustin, M., Woodward, A., Patel, U.A., Kovelman, R., and Harvey, K.J. (2013). A broad activity screen in support of a chemogenomic map for kinase signalling research and drug discovery. Biochem. J. *451*, 313–328.

Koleti, A., Terryn, R., Stathias, V., Chung, C., Cooper, D.J., Turner, J.P., Vidović, D., Forlin, M., Kelley, T.T., D'Urso, A., et al. (2017). Data Portal for the Library of Integrated Network-based Cellular Signatures (LINCS) program: integrated access to diverse large-scale cellular perturbation response data. Nucleic Acids Res.

R Core Team (2016). R: A Language and Environment for Statistical Computing.

Sievert, C., Parmer, C., Hocking, T., Chamberlain, S., Ram, K., Corvellec, M., and Despouy, P. (2017). plotly: Create Interactive Web Graphics via "plotly.js."

Stachura, F. (2018). shiny.semantic: Semantic UI Support for Shiny.

Vaidyanathan, R., Xie, Y., Allaire, J.J., Cheng, J., and Russell, K. (2018). htmlwidgets: HTML Widgets for R.

Wickham, H. (2007). Reshaping Data with the {reshape} Package. J. Stat. Softw. *21*, 1–20.

Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag New York).

Wickham, H., Francois, R., Henry, L., and Muller, K. (2017). dplyr: A Grammar of Data Manipulation.

Xie, Y. (2018). DT: A Wrapper of the JavaScript Library "DataTables."