Assessing Darkness of the Human Kinome from a Medicinal Chemistry Perspective

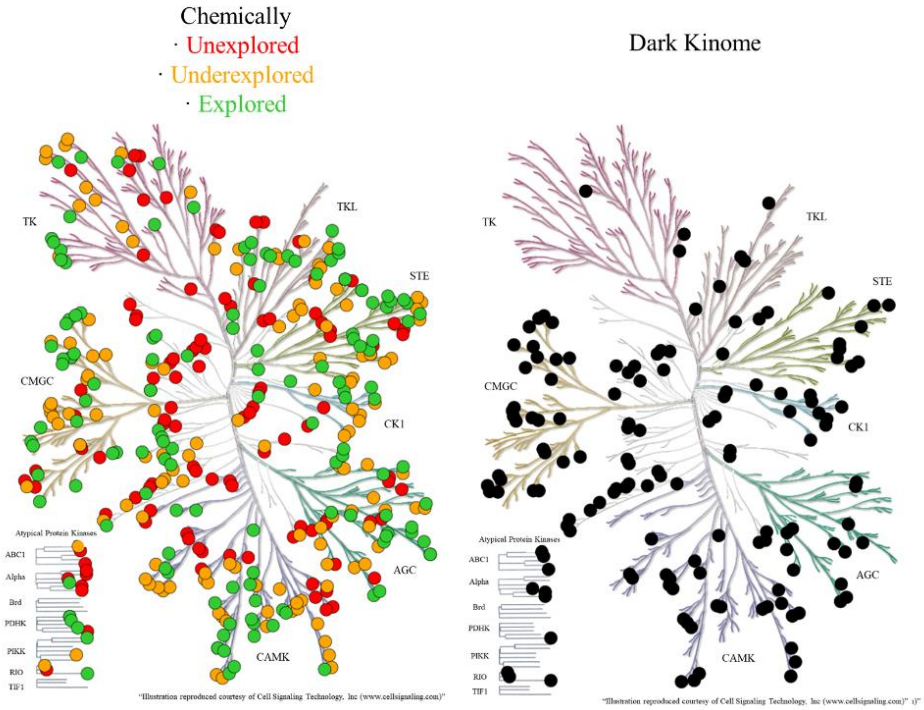
Selina Voßen,[1] Elena Xerxa,[1,2] and Jürgen Bajorath[1-3]*

[§]B-IT, Department of Life Science Informatics and Data Science; [#]Lamarr Institute for Machine Learning and Artificial Intelligence; [€]LIMES Institute, Program Unit Chemical Biology and Medicinal Chemistry, Rheinische Friedrich-Wilhelms-Universität, Friedrich-Hirzebruch-Allee 5/6, D-53115 Bonn, Germany.

Corresponding Author

*Phone: 49-228-7369-100, E-mail: bajorath@bit.uni-bonn.de.

Graphical Abstract



Abstract

In drug discovery, human protein kinases (PKs) represent one of the major target classes, due to their central role in cellular signaling, implication in various diseases as a consequence of deregulated signaling, and their notable druggability. Individual PKs and their disease biology have been explored to different degrees, giving rise to heterogeneous functional knowledge and disease associations across the human kinome. The U.S. National Institutes of Health previously designated 162 understudied ("dark") human PKs and lipid kinases, due to the lack of functional annotations and high-quality molecular probes for functional investigations. Given large volumes of available PK inhibitors (PKIs) and activity data, we have systematically analyzed the distribution of PKIs and associated data at different confidence levels across the human kinome and distinguished between chemically explored, underexplored, and unexplored PKs. The analysis provides a medicinal chemistry-centric view of PK exploration and further extends prior assessment of the dark kinome.

Introduction

The human kinome contains 518 wild-type protein kinases (PKs) that are organized into different groups.[1] PKs catalyze the adenosine triphosphate (ATP)-dependent phosphorylation of tyrosine or serine/threonine residues in PKs or other proteins and play a key role in cellular signaling. The transduction of signals into cells is mediated by receptor-associated PKs that initiate phosphorylation cascades involving downstream PKs and other signaling proteins.[2] Uncontrolled PK activity and abnormal signaling have been implicated in the development and progression of many diseases, making PKs prime targets in drug discovery.[2,3] In addition to insights from studying PK disease biology, the early-on recognized druggability of the PK catalytic domain further increased the interest in PKs as therapeutic targets and favored small molecular PK inhibitors (PKIs) as drug candidates.[4] For targeting the catalytic domain, PKIs with varying modes of action were discovered including different types of competitive ATP/active site-directed, allosteric, or covalent inhibitors.[5-8] The therapeutic potential of PKIs was first explored and exploited in oncology[9-11] and then in other therapeutic areas[11-13] including neurodegenerative disorders, metabolic diseases, and immunology.[14-16] In 2001, with imatinib, the first PKI was approved for the treatment of chronic leukemia by drugs by the U.S. Food and Drug Administration (FDA).[9] Beginning with imatinib, 80 PKIs have thus far obtained FDA approval[17] and ~200 PKIs are at different stages of clinical development. In addition to clinical compounds, very large numbers of PKIs from medicinal chemistry have become publicly available,[18] which also reflects the intensity of PK drug discovery efforts.

In drug discovery, biological evidence of disease implication of potential targets and availability of relevant active compounds typically determine target priorities. Even if discovered serendipitously, promising targets are then often pursued simultaneously in different pharma

environments, spurred on by "me-tooism", as exemplified by the dominant early interest in tyrosine PKs as targets for cancer treatment.[4,10] As a consequence of propagating target preferences, however they might be arrived at, and varying degrees of exploration, there generally is considerable heterogeneity in therapeutic target classes, that is, some targets are intensely investigated whereas others are only little explored. For instance, in 2010, literature analysis of PKs implicated in cancer revealed that ~25% of the human kinome lacked functional annotations and ~50% was only preliminarily studied.[19] Given this imbalance oberved for PKs and other major pharmaceutical target classes, the U.S. National Institutes of Health (NIH) launched the Illuminating the Druggable Genome (IDG) project aiming to further explore and characterize underinvestigated targets in major druggable gene families including PKs, G protein coupled receptors (GPCRs), and ion channels.[20] As a part of this NIH-funded initiative, the Kinase Data and Resource Generating Center (Kinase DRCG) was formed to investigate understudied kinases.[21] In 2019, the NIH reported a list of 162 human kinases (including PKs and lipid kinases) representing targets for the Kinase DRCG.[21,22] Criteria for understudied kinases included the lack of literature citations, functional, regulatory, and signaling pathway information as well as the lack of monoclonal antibodies as detection reagents and active compounds as molecular probes (tool compounds) for functional studies (as further discussed below).[21,23] These understudied kinases were designated "dark kinases" or the "dark kinome" by the IDG Knowledge Management Center and the Dark Kinase Knowledgebase (DKK).[24,25] The mission of the Kinase DRGC has been -and continues to be- the generation of functional annotations and molecular probes for the designated dark kinases. Therefore, the research consortium developed assays for quantifying kinase gene expression, identifying interaction networks, and detecting compounds with kinase activity in living cells, including molecular probes.[21] The resulting data are compiled, organized, and made

available in the DKK repository.[25] These Kinase DRCG efforts to illuminate dark kinases have been complemented by other studies including the development of a statistical scoring scheme based on differential kinase gene expression and clinical parameters to prioritize understudied kinases as potential cancer targets,[26] a kinase ontology for the integration and interactive analysis of functional data,[27] and multi-level kinase citation analysis,[28] also including database searches for molecular probes.[28]

While most of the IDG activities and related studies have concentrated on elucidating cellular functions of dark PKs using biological approaches, molecular probes represent the chemical component of the methodological portfolio. In chemical biology, probe compounds were originally introduced to provide an alternative to biological knockout experiments and models for specifically interfering with molecular functions of target proteins and studying functional consequences.[29-31] Compared to genetic approaches that deprive cells of individual target proteins, molecular probes have the advantage of temporal interference and assessment of dose-response behavior. However, to ensure non-ambiguous functional analysis, molecular probes should specifically inhibit their targets and not elicit any secondary effects. Accordingly, stringent criteria are applied to candidate compounds to qualify for inclusion in public probe repositories such as the molecular probes collection of the Structural Genomics Consortium (SGC)[32,33] or the Chemical Probes Portal (CPP).[34] Molecular probe criteria partly differ from and might go beyond early-phase selection criteria for drug candidates. For instance, probes should be potent (nanomolar) inhibitors of a given target, with at least 30-fold selectivity over targets from the same family, and display dose-dependent on-target activity in cells (at no more than 1 μM concentration).[32] In addition, structurally analogous inactive (negative control) compounds should be available. Given these requirements, qualifying molecular probes are typically difficult to obtain. For example, a

recent search in SGC and CPP identified at least one qualifying probe compound for only 129 PKs of the human kinome.[28] Furthermore, Kinase DRCG efforts over the past years yielded high-quality tool compounds for only 44 of the 162 understudied kinases.[21] Thus, given the sparseness of chemical probes, they can complement functional illumination of PKs using biological approaches, but are insufficient for chemically-oriented assessment of PK darkness.

Herein, we report a systematic analysis of the chemical exploration of the human kinome based on large volumes of compound activity data from various public sources at different levels of curation, without pre-conceived notion of previously designated understudied PKs. The results complement functional characterization of the dark kinome with medicinal chemistry-centric analysis of PK exploration.

Results and Discussion

Study concept

Ascertainment of the dark kinome was primarily based thus far on lack of functional information and molecular probes as chemical tools for functional analysis. We have reasoned that the assessment of "chemical exploration" of PKs should complement and further extend functional investigations, especially for drug discovery. Taking into consideration that designated small molecular probes meeting the desired potency, selectivity, and negative control requirements are currently only available for ~25% of the human kinome, probes are clearly insufficient as a measure of chemical exploration of PKs. However, large volumes of public PKIs and their activity data are available to determine the distribution of active compounds across the human kinome and differentiate PKs according to their degree of chemical exploration, that is, varying availability of qualifying PKIs. Therefore, we have curated available PKIs including clinical compounds at different data confidence levels (distinguishing between low- and/or high-confidence activity data), formulated criteria for assessing chemical exploration, characterized PKIs including promiscuous compounds available for each wild-type PK, and differentiated chemically explored, underexplored, and unexplored PKs. Our analysis was conducted without any pre-conceived notion of dark PKs, providing an independent medicinal chemistry-centric view of PK exploration, as reported in the following.

Data sources, retrieval, and curation

In a previous survey of public compound repositories,[35] the vast majority of PKIs and activity data was found in ChEMBL[36] and BindingDB,[37] which were used herein as major sources for PKI

curation. Additional specialized database included in our analysis are reported in the Methods section. Prior to data curation, in ChEMBL, a total of 197,115 PKIs and 696,081 activity records were available and in BindingDB, 245,137 and 420,164 PKIs and activity records, respectively. Consistent with previous findings,[35,38] there was substantial overlap between the two databases, with 129,244 shared PKIs (41%) and 243,136 (28%) activity records. Furthermore, PKIs and activity data were reported in ChEMBL and BindingDB for 445 human PKs and only in ChEMBL or Binding DB for 17 PKs and one PK, respectively, hence for 463 PKs combining these resources. In addition, we detected PKI activity data for an additional PK in the chemoproteomics data set of Reinecke et al.[46], thus amounting to a total of 464 human PKs.

Data curation based on the criteria for PKIs with available high- or low-confidence activity data (see Methods) resulted in a high-confidence data set containing 199,428 PKIs with activity against a total of 446 PKs and a low-confidence data set of 287,564 PKIs with activity against 459 PKs. Taken together, activity data in the high- and low-confidence sets covered a total of 459 PKs (89% of the human kinome). In the low-confidence data set, approximate measurements were included and records of activity or inactivity (such as ">10,000 nM") were distinguished (see Methods).

Classification criteria

Our analysis aimed to distinguish between three categories of PKs including *chemically explored, underexplored,* and *unexplored* PKs defined based upon the following criteria.

PKs were classified as chemically unexplored if *(i)* no PKIs were available in the low- and high-confidence data sets and no PKIs were reported in the Reinecke data set[39] or if *(ii)* PKs were exclusively annotated with promiscuous PKIs with reported activity against at least 10 PKs (see Methods).

PKs for which at least one approved PKI drug or clinical compound was available were classified as chemically explored. The availability of at least one drug or PKI in clinical trials was considered a sufficient indicator of chemical exploration, taking into consideration that the development of clinical compounds requires extensive compound optimization efforts and that corresponding data obtained on the way to clinical candidates are often kept proprietary. Notably, for most PKs with clinical compound(s), varying (and often large) numbers of other qualifying PKIs were available. In addition, based on statistical assessment (see below), PKs for which at least 30 unique PKIs with high-confidence data were available, were also classified as chemically explored. PKs having no PKIs with high-confidence activity data were classified as chemically underexplored.

For the remaining PKs, for which qualifying PKIs with high- and low-confidence activity data were available, promiscuous PKIs were also omitted from further consideration. These PKs were classified as chemically explored or underexplored "on a sliding scale" based on available compound numbers and data quality, as detailed below.

Chemically unexplored protein kinases

For 52 PKs, no compound records were available, for two PKs, all activity records were incomplete, and for four others, no qualifying PKIs remained in the low- or high-confidence data sets after initial data curation. Furthermore, 62 PKs were only annotated with (one to 348) promiscuous PKIs. Thus, on the basis of these criteria, a total of 120 PKs were classified as chemically unexplored. The different subsets of chemically unexplored PKs are reported in **Supplementary Table S1**.

Protein kinases with clinical compounds

We next searched for PKs for which approved drugs and/or other clinical PKs were reported in ChEMBL or other specialized databases (see Methods) and identified a total of 131 "clinical PKs" that were classified as chemically explored, as rationalized above. Clinical PKs are reported in **Supplementary Table S2**. In addition, **Table 1** reports the top-10 clinical PKs together with the numbers of associated drugs, clinical PKIs, and other PKIs.

Table 1. Top-10 clinical protein kinases.

PK	UniProt ID	Approved Drugs	Clinical PKIs	Low-confidence PKIs	High-confidence PKIs
Vascular endothelial growth factor receptor 2	P35968	12	41	15,970	9783
Mast/stem cell growth factor receptor Kit	P10721	12	17	5004	2593
Epidermal growth factor receptor	P00533	9	34	15,655	9673
Platelet-derived growth factor receptor alpha	P16234	8	17	2604	916
Tyrosine-protein kinase ABL1	P00519	8	12	8031	2861
Receptor-type tyrosine-protein kinase FLT3	P36888	7	22	7409	3906
Proto-oncogene tyrosine-protein kinase receptor Ret	P07949	7	4	5208	3413
Receptor tyrosine-protein kinase erbB-2	P04626	6	16	6060	3221
Fibroblast growth factor receptor 2	P21802	6	13	5315	2466
ALK tyrosine kinase receptor	Q9UM73	6	4	4144	2260

Reported are the top-10 clinical PKs based on number of approved drugs. In addition, for each PK, the number of PKIs in clinical trials and other PKIs in the low- and high-confidence data set is reported. PKs with the same number of approved drugs are ranked based on PKIs in clinical trials.

With eight to 12 approved drugs per PK, highest ranked are four growth factor receptor variants, which are primary cancer targets, followed by ABL1 and FLT3. Overall, receptor Tyr PKs implicated in cancer currently are the most established PK drug targets.

Kinases with varying degrees of chemical exploration

For clinical PKs, all associated PKIs exclusively reported to be active against these PKs were removed from the high- and low-confidence data sets for further analysis. In addition, after determining PKs only annotated with promiscuous compounds, the promiscuous PKIs were removed from the data sets. The substantially reduced versions of the high- and low-confidence data sets included 40,895 PKIs with activity against a total of 264 PKs and 63,436 PKIs with activity against 267 PKs, respectively, which provided the basis of our subsequent analysis of other chemically explored and underexplored PKs.

For three PKs included in the low-confidence set, no high-confidence PKI data were available. Thus, these PKs were regarded as chemically underexplored. Next, for the 264 PKs in the high-confidence data set, the distribution of high-confidence PKIs was determined (**Figure 1**).

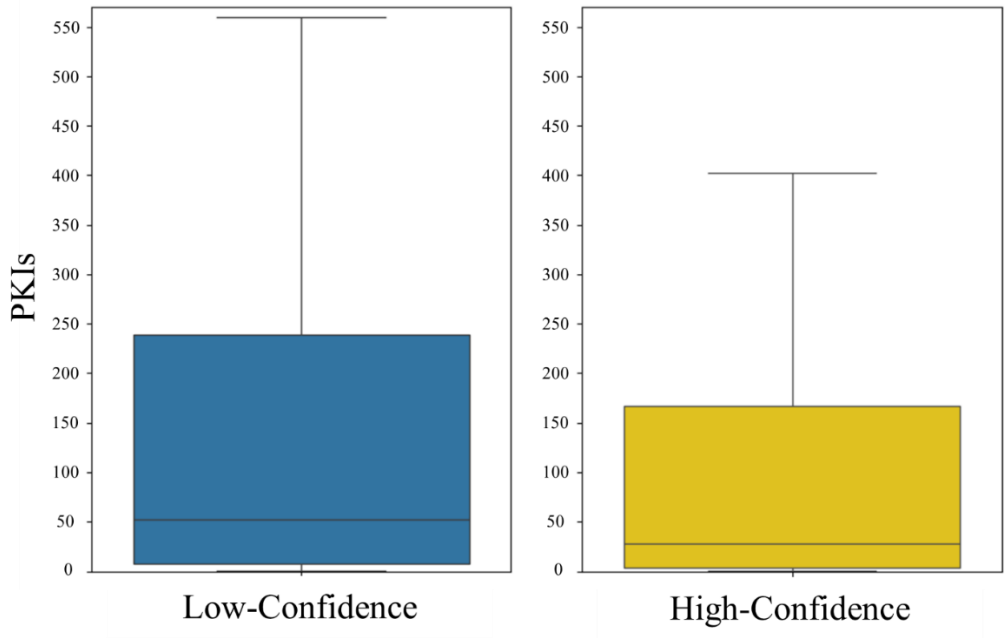


Figure 1. Distribution of high-confidence protein kinase inhibitors.

Boxplots show the distribution of PKIs from the low- (left) and high- confidence set (right). The horizontal lines represent the median values and separate the upper and lower quartile. The whiskers mark the value range of the data within 1.5 times the interquartile range. Statistical outliers have been removed for clarity and are reported in Table 2 and Supplementary Table S3.

The distributions of the low- and high-confidence had median values of 52 and 28 PKIs per PK, respectively, and the upper quartiles of the distribution covered broad value ranges. Given the distribution of the high-confidence set, we classified 129 PKs as chemically explored, for which at least 30 unique PKIs with high-confidence activity data were available (that is, a number of PKIs slightly above the mean). For these 129 PKIs, a median (mean) value of 244 (575) low-confidence PKIs were also available.

Table 2 reports the top-10 chemically most explored PKs from the high-confidence set, all of which were statistical outliers of the distribution in **Figure 1**. For each of these PKs, thousands of low- and high-confidence PKIs were available. Different from the leading clinical PKs, Ser/Thr PKs such as pim-1 and pim-2 proto-oncogene and different MAP PKs represented the chemically by far most explored PKs, which were not among the leading clinical PKs.

Table 2. Top-10 chemically explored PKs with largest number of PKIs.

PK	UniProt ID	Low-confidence PKIs	High-confidence PKIs
Serine/threonine-protein kinase pim-1	P11309	5502	4335
Leucine-rich repeat serine/threonine-protein kinase 2	Q5S007	4523	3360
Mitogen-activated protein kinase kinase kinase kinase 1	Q92918	5379	3122
Serine/threonine-protein kinase pim-2	Q9P1W9	3318	2700
Serine/threonine-protein kinase TBK1	Q9UHD2	1742	1354
Epithelial discoidin domain-containing receptor 1	Q08345	1368	1350
MAP kinase-interacting serine/threonine-protein kinase 2	Q9HBH9	1745	1292
Tyrosine-protein kinase receptor TYRO3	Q06418	1565	1268
Casein kinase I isoform delta	P48730	2784	1251
Dual specificity tyrosine-phosphorylation-regulated kinase 1A	Q13627	3996	1155

Reported are the top-10 chemically most explored PKs based on largest numbers of PKIs with high-confidence activity data.

Chemically explored protein kinases

The union of the 131 clinical PKs and the other 129 PKs with at least 30 high-confidence PKIs represented chemically explored PKs. For the 135 PKs remaining in the high-confidence data set, the number of high- and low-confidence PKIs were compared, as shown in **Figure 2**.

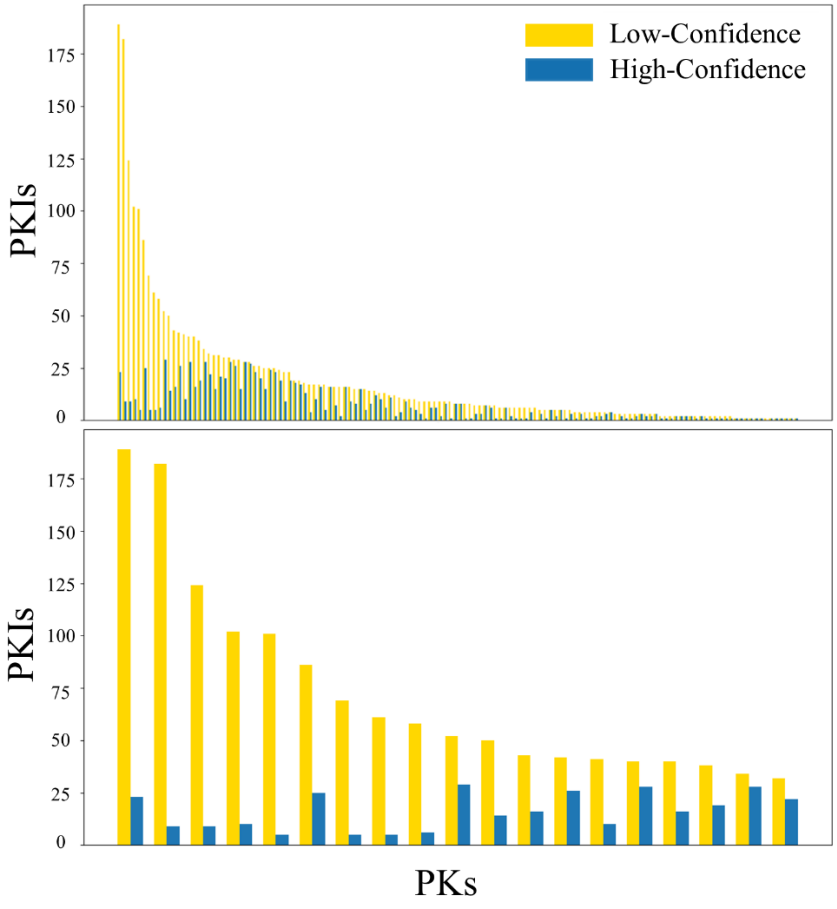


Figure 2. Protein kinase subset with high- and low-confidence inhibitors. For PKs with at least one but fewer than 30 high-confidence PKIs, the histogram representation compares the number of available high- and low-confidence PKIs. At the top, the global distribution is displayed after removal of the first two PKs with largest numbers of low-confidence PKIs (for better visibility). From the left to the right, PKs were ordered according to decreasing numbers of low-confidence PKIs. At the bottom, only PKs number 3-20 of the distribution are shown on a different scale after removal of the first two PKs with largest numbers of low-confidence PKIs.

The distribution showed a sharp decline in the number of low-confidence PKIs for the remaining subset of PKs with less than 30 high-confidence PKIs. The first two PKs, 5'-AMP-activated protein kinase catalytic subunit alpha-2 and cyclin-dependent kinase 19, dominated the distribution. For 5'-AMP-activated protein kinase catalytic subunit alpha-2, 758 low- and 22 high-confidence PKIs were available and for cyclin-dependent kinase 19, 339 and 25, respectively. Given the large number of low-confidence PKIs that were available for a limited number of PKs with varying numbers of less than 30 high-confidence PKIs, we classified eight additional PKs (i.e., PK 1-8 of the distribution in **Figure 2**, top) as chemically explored. For the first seven PKs, more than 100 low-confidence were available. For the PK at rank eight, calcium/calmodulin-dependent protein kinase type II subunit alpha, which we included as the last PK in the chemically explored category, 86 low- and 25 high-confidence PKIs were available.

Thus, based on our analysis, we identified a total of 268 chemically explored PKs including the 131 clinical PKs (**Supplementary Table S2**) and 137 (129 plus eight) others, reported in **Supplementary Table S3**.

Underexplored protein kinases

The remaining 127 PKs with low- and high-confidence PKIs were classified as chemically underexplored. In addition, there were three PKs for which only one to six (non-promiscuous) low- and no high-confidence PKIs were available, hence yielding a total of 130 chemically underexplored PKs, reported in **Supplementary Table S4**.

Protein kinase classification

Figure 3 summarizes the PK classification and shows the distribution across the human kinome. Chemically unexplored, underexplored and explored PKs were widely distributed over all PK groups. Chemically unexplored PKs were also found within the most intensely investigated groups of Tyr and Ser/Thr PKs (group TK and CMGC, respectively). **Figure 4** shows exemplary chemically explored and underexplored PKs from two other PK groups.

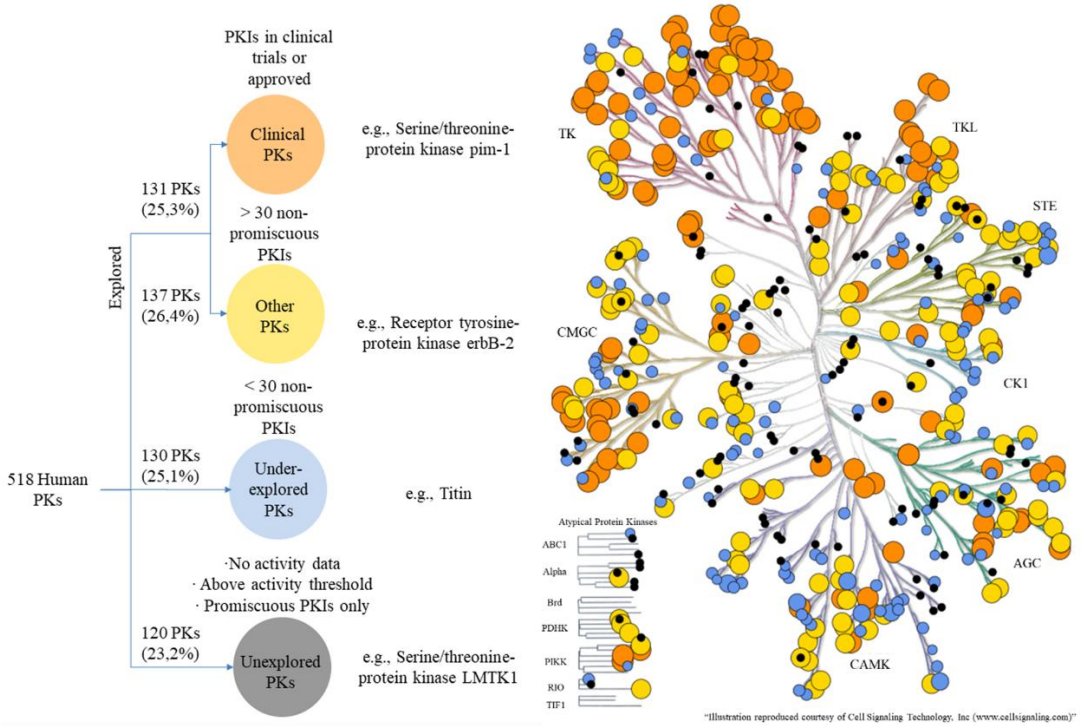


Figure 3. Protein kinase classification. On the left, our PKI data-based PK classification is summarized, yielding 120 chemically unexplored, 130 underexplored, and 268 explored PKs. On the right, the classification is mapped onto a phylogenetic tree presentation of the human kinome (different groups are labeled using standard abbreviations). Each dot represents a PK that are classified using the following color code: clinical PKs, orange; other chemically explored PKs, yellow; underexplored, blue; unexplored, black. For chemically explored and underexplored PKs, dots are scaled in size according to the total number of available PKIs. Chemically unexplored PKs are consistently represented using smallest black dots.

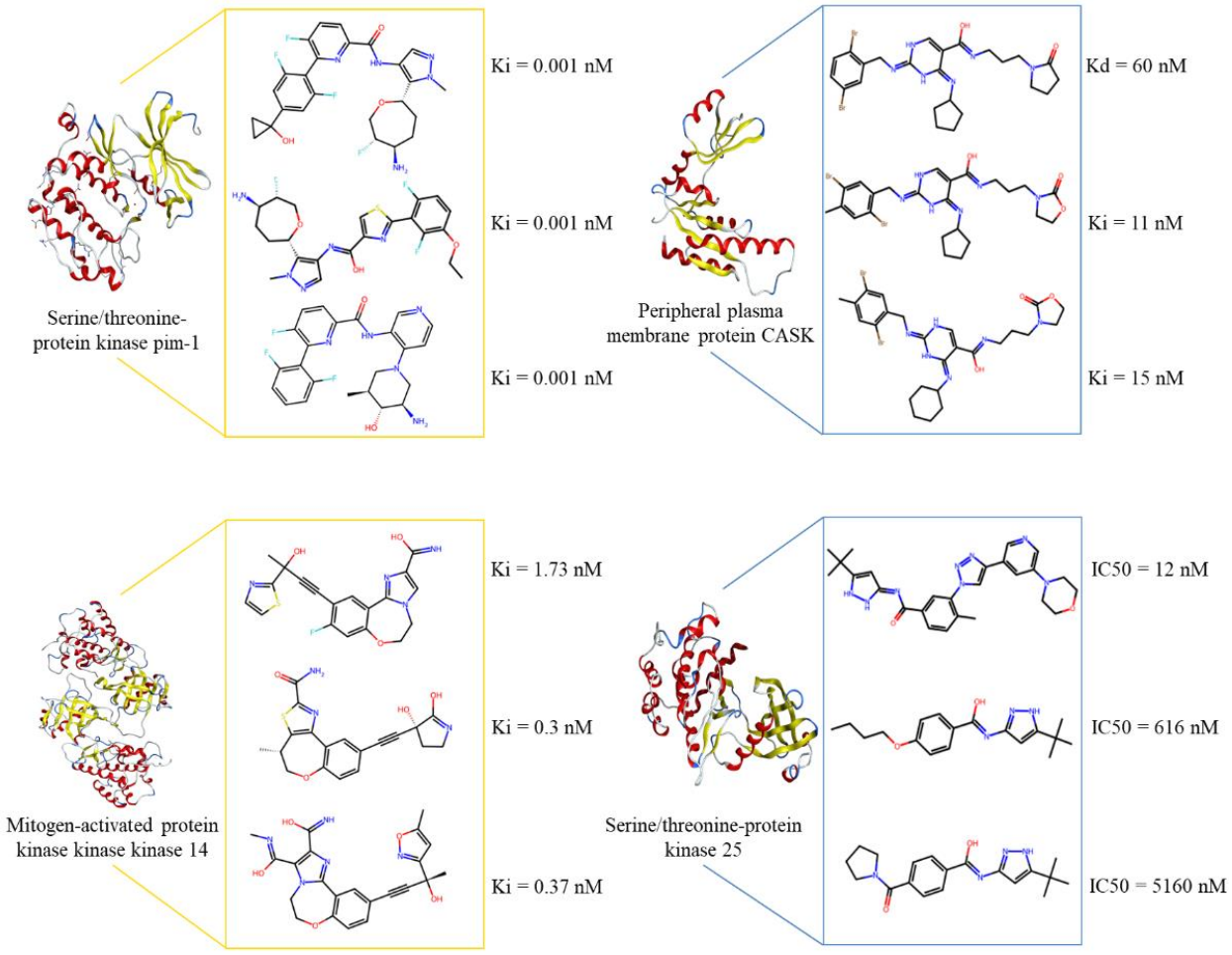


Figure 4. Exemplary protein kinases with different degrees of chemical exploration. Shown are exemplary PKs (X-ray structure renderings) and their PKIs from two different groups (CAMK, top; STE, bottom). On the left and right, chemically explored and underexplored PKs are shown, respectively. For chemically explored and underexplored PKs, exemplary PKIs from the high- and low-confidence data set are shown, respectively. PKs are identified using their UniProt IDs: P11309, serine/threonine-protein kinase pim-1; O14936, peripheral plasma membrane protein CASK; Q99558, Mitogen-activated protein kinase kinase kinase 14; O00506, serine/threonine-protein kinase 25. Only non-promiscuous compounds are shown.

Chemically classified versus dark protein kinases

The 162 understudied/dark kinases published by the NIH[21,22] were found to contain six lipid kinases and one pseudokinase. The remaining 155 dark PKs were compared to chemically unexplored and underexplored PKs classified in our analysis (**Figure 5**). A total 120 chemically unexplored (62) and underexplored (58) PKs overlapped with the 155 dark PKs designated by the NIH. Hence, for these PKs reported in **Supplementary Table S5,** both functional information and compound data were limited (or lacking).

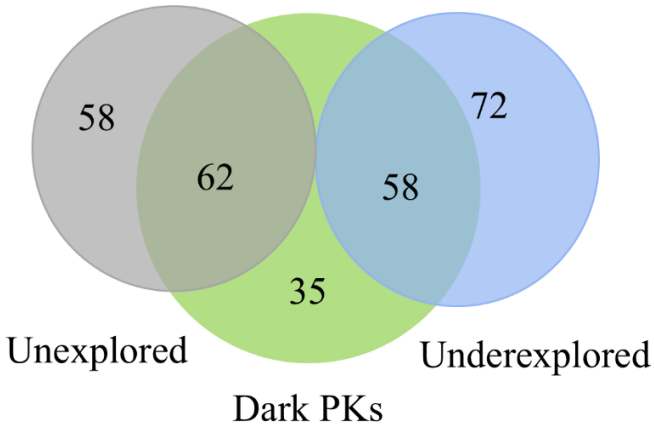


Figure 5. Chemically un- and underexplored versus dark protein kinases. The Venn diagram compares chemically un- or underexplored with dark PKs.

However, all of the remaining 35 designated dark PKs were classified as chemically explored in our analysis, including six clinical PKs, as reported in **Supplementary Table S6**. Most of these PKs were associated with large numbers of PKIs. For 15 PKs, 100 or more (up to 1292) high-confidence PKIs (and, in each case, larger numbers of low-confidence PKIs) were available. Among these were, for instance, MAP kinase-interacting serine/threonine-protein kinase 1 with one clinical PKI, 1134 high-, and 1669 low-confidence PKIs or protein kinase C theta type with one approved drug, three other clinical PKIs, 812 high-, and 1210 low-confidence PKIs.

Figure 6 shows that these chemically explored PKs were widely distributed across the human kinome, essentially covering all PK groups. Notably, 10 chemically explored Ser/Thr PKs of the CMGC group for which (with one exception) large numbers of PKIs were available, were previously classified as understudied. Taken together, the 120 overlapping chemically un- or underexplored and understudied PKs and the 35 chemically explored PKs previously categorized as understudied provide a further refined view of the remaining dark human kinome.

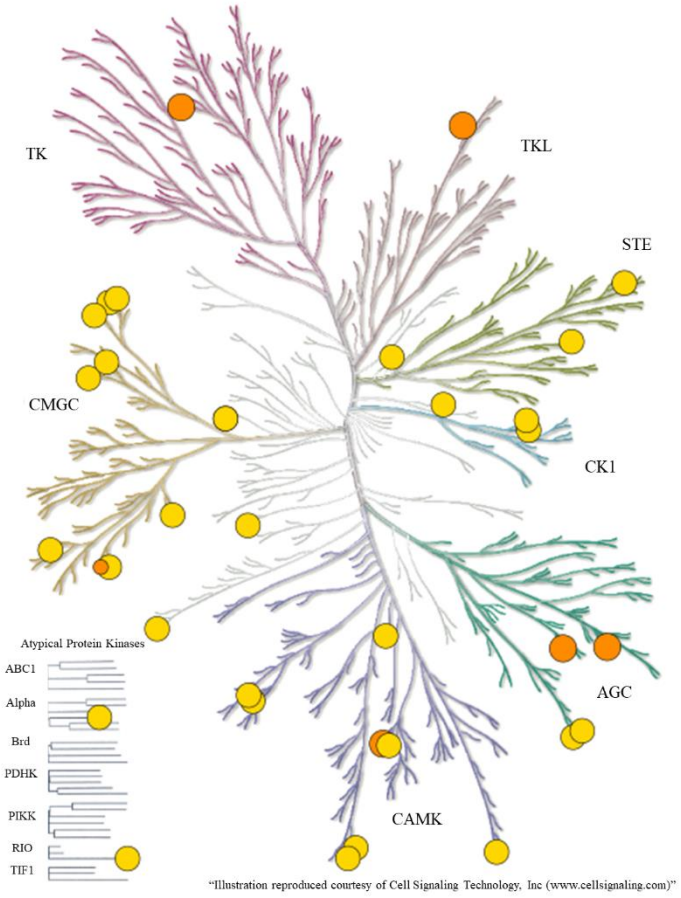


Figure 6. Distribution of chemically explored protein kinases categorized as understudied. The 35 PKs previously categorized as understudied are mapped on a phylogenetic tree of the human kinome. The representation is according to Figure 3.

Concluding Remarks

Although PKs are among the most heavily investigated pharmaceutical target proteins, the 80 FDA-approved PKI drugs are directed against only 22 primary PK targets.[17] Given the history of PK drug discovery, anti-cancer PKIs clearly dominate. The 80 approved drugs include 69 anti-cancer agents compared to six PKI drugs for the treatment of inflammatory diseases. Forty-three PKI drugs are active against receptor Tyr PKs, which are eminent cancer targets, followed by 20 with activity against non-receptor Tyr PKs, and 13 with activity against Ser/Thr PKs.[17] However, while clinical compounds currently are available for more than 100 PKs, as discussed above, PKI drugs have thus only been approved for a very small fraction of the human kinome. Accordingly, given that only ~25% of the human kinome is charted with clinical compounds at present, PK targets remain of high interest in different therapeutic areas, and the quest for new insights into PK disease biology and for new chemical matter continues.

In this context, the IDG initiative of the NIH including the Kinase DRCG[21,22] has been a milestone event for the field, focusing on PKs with little or no functional annotations. As an initial part of this endeavor, 162 understudied kinases including 155 PKs were identified as primary targets for further functional evaluation and the development of chemical probes, representing the often cited dark kinome, considered as a source of potential drug targets with new disease implications.

Our current analysis attempted to provide a complementary medicinal chemistry-centric view of PK exploration, concentrating on available PKIs including clinical compounds and activity data at different confidence levels. To ensure rigor and comprehensiveness of the analysis, extensive data curation across different compound repositories was mandatory. This made it possible to organize (and rank) wilde-type PKs of the human kinome based on different degrees of chemical

exploration, also taking into consideration that especially for PKs with clinical compounds, PKI data are often kept proprietary. To provide guidelines for PK classification, we distinguished between chemically unexplored, underexplored, and explored Pks based on the criteria specified above. Of course, chemically underexplored and explored PKs represent a continuum with respect to PKI volumes and confidence levels, rather than discrete states. However, they can be differentiated based on chemcial and statistical considerations. These criteria are not written in stone and can be adjusted for specific applications, and the PK rankings reported herein can also be viewed as a continuum.

Given the absence of any qualifying PKIs or presence of exclusive activity against promiscuous (broad spectrum) PKIs, we classified 120 PKs as chemically unexplored. Furthermore, based on available PKIs including clinical compounds and varying data confidence levels, we differentiated between 130 chemically underexplored and 268 explored PKs including 131 clinical PKs. Hence, nearly half of the human kinome currently remains chemically un- or underexplored.

Following our analysis and PK classification, we then compared the 155 previously designated (functionally) dark PKs and 250 chemically un- or underexplored PKs reported herein. With 120 shared PKs, there was large overlap between these independently derived PK subsets; an encouraging finding. If target functions remain unclear, extensive medicinal chemistry efforts for target intervention are very unlikely. However, we also found that 58 unexplored and 72 underexplored PKs were not included in the collection of understudied/dark PKs, leaving considerable room for further chemical exploration of these candidate PKs, if prioritzed in light of disease biology. Moreover, 35 dark PKs were classified as chemically explored and in many cases, these PKs were highly explored, including clinical compounds. This noteworthy discrepancy might at least in part be attributable to progress made by the Kinase DRCG over the past four

years, mirrored by increasing medicinal chemistry efforts on selected PKs with new and interesting functional annotations. On the other hand, these observations also point at the complementary nature of biological and medicinal chemistry-centric evaluation of PK exploration, adding a new dimension to the further evaluation of the dark kinome.

Methods

Data retrieval and curation

A comprehensive list wild-type PKs constituting the human kinome was compiled by integrating information from Manning et al.[1] and UniProt[39]. A total of 517 UniProt target identifiers (IDs) were retrieved (with one ID covering two PKs reported by Manning et al.) and used to extract PK information from different databases. ChEMBL[36] (version 34) and BindingDB[37] (accessed May 2024) were the main sources for PKI data. For retrieval of PKIs and activity data, UniProt PK IDs were mapped to corresponding human ChEMBL IDs. Target and compound information extracted from ChEMBL included PK gene name, standard activity measurement, target type, standard relation, standard units, activity value, and Simplified Molecular Input Line System (SMILES)[41] and InChI strings of PKIs. Activity records were not considered if any of the information above was missing. Duplicate entries were removed. For BindingDB, additional data processing was required to align with the data structure of ChEMBL. For example, target type information was not explicitly provided but inferred from the number of targeted protein chains (e.g., 1 corresponds to single target). For each PK, unique PKIs and activity data from ChEMBL and BindingDB were combined. In addition, for the ~1000 PKIs investigated by Reinecke et al., PK activity annotations were analyzed to potentially identify PKs not covered by ChEMBL/BindingDB assay data. Ultimately, for all qualifying PKIs, IUPAC International Chemical Identifier (InChI)[41] strings were subjected to chemical standardization including salt removal, canonicalization, and neutralization. The resulting canonical InChI served as unique compound identifiers for subsequent data processing.

Protein kinase inhibitor sets with high- or low-confidence activity data

High-confidence PKI activity data were selected based on previously reported selection criteria.[18] Accordingly, standard activity measurements (K_d, K_i or IC_{50}) with numerically specified values, "nM" units, standard relation "=", and "single protein" target type were required. If multiple values of the same activity measurement type were available, they were averaged if all values fell into one order of magnitude; otherwise, the measurements were discarded. If multiple measurement types were available, the final activity annotation was selected according to the priority order (K_i> K_d> IC_{50}). For activity, a threshold of at least 50,000 nM or lower was applied (that is, a record > 50,000 nM was classified as inactive and disregarded).

Low-confidence PKI data included records with quantitative measurements that did not meet all criteria required for high-confidence data as well as approximate measurements with standard relation "<", "<=", ">" or >=. For quantitative measurements, multiple values of the same type were averaged (or discarded) and as a threshold for activity, a numerically specified (K_i, K_d, IC_{50}, or EC50) value of at least 50,000 nM or lower was required, consistent with high-confidence data. For records with % inhibition measurements, values of >=50% were classified as active.

If multiple measurement types were available, the final activity annotation was selected according to the priority order (K_i> K_d> IC_{50}>EC50). If multiple approximate measurements of the same type were available, values with operators ("<", "<=") were separated from those with (">", ">=") if applicable and the lowest and highest value was selected, respectively. If numerically specified and approximate activity values were available for a PKI, numerically specified values were retained. Approximate annotations were classified as active if they met threshold values of <= 10,000 nM for (K_i, K_d, IC_{50}, or EC50) or >=50% inhibition.

Notably, the most frequently detected approximate activity annotation was "> 10,000 nM" for different measurement types, which was classified as an indicator of inactivity in a PK assay.

Promiscuous protein kinase inhibitors

Promiscuous (multi-PK) inhibitors were identified based on low-confidence and high-confidence data by counting the total number of PK annotations. If a PKI was reported to be active against at least 10 PKs it was classified as a promiscuous PKI.

Kinome representations

Annotated phylogenetic tree representations of the human kinome were generated with KinMap.[42] In display items and the test, PKs were designated using standard UniProt abbreviations.[39]

Clinical protein kinases

PKs with PKIs approved as drugs or in clinical trials were classified as clinical PKs. Therefore, ChEMBL and the Pharos[23] database and different drug repositories including DrugBank,[43] DrugCentral,[44] and GuidetoPHARMACOLOGY[45] were searched for clinical PKIs. For a given PK, at least one clinical PKI was required to be reported in each database to obtain clinical PK status. PKs with clinical PKIs were classified as chemically explored PKs.

Acknowledgment

The authors thank Martin Vogt for support and helpful discussions.

Abbreviations: ATP, adenosine triphosphate; DKK, Dark Kinase Knowledgebase; FDA, Food and Drug Administration; GPCRs, G protein-coupled receptors; ID, identifier; IDG, Illuminating the Druggable Genome; InChI, IUPAC International Chemical Identifier; NIH, National Institutes of Health; PD, promiscuity degree; PK, protein kinase; PKI, protein kinase inhibitor; CPP, Chemical Probes Portal; SGC, Structural Genomics Consortium; SMILES, Simplified Molecular Input Line System; TCI, target-compound interaction.

References

1. Manning, G.; Whyte, D.; Martinez, R.; Hunter, T.; Sudarsanam, S. The Protein Kinase Complement of the Human Genome. *Science* **2002**, *298*, 1912-1934.

2. Ferguson, F. M.; Gray, N. S. Kinase Inhibitors: The Road Ahead. *Nat. Rev. Drug Discovery* **2018**, *17*, 353-376.

3. Attwood, M. M.; Fabbro, D.; Sokolov, A.V.; Knapp, S.; Schiöth, H. B. Trends in Kinase Drug Discovery: Targets, Indications and Inhibitor Design. *Nat. Rev. Drug Discov.* **2021**, *20*, 839-861.

4. Levitzki, A.; Gazit, A. Tyrosine Kinase Inhibition: An Approach to Drug Development. *Science* **1995**, *267*, 1782-1788.

5. Gavrin, L. K.; Saiah, E. Approaches to Discover Non-ATP Site Kinase Inhibitors. *Med. Chem. Comm.* **2013**, *4*, 41–51.

6. Müller, S.; Chaikuad, A.; Gray, N. S.; Knapp, S. The Ins and Outs of Selective Kinase Inhibitor Development. *Nat. Chem. Biol.* **2015**, *11*, 818– 821,

7. Lu, X.; Smaill, J. B.; Ding, K. New Promise and Opportunities for Allosteric Kinase Inhibitors. *Angew. Chem. Int. Ed.* **2020**, *59*,13764-13776.

8. Abdeldayem, A.; Raouf, Y. S.; Constantinescu, S. N.; Moriggl, R.; Gunning, P. T. Advances in Covalent Kinase Inhibition. *Chem. Soc. Rev.* **2020**, *49*, 2617-2687.

9. Deininger, M.; Buchdunger, E.; Druker, B. J. The Development of Imatinib as a Therapeutic Agent for Chronic Myeloid Leukemia. *Blood* **2005**, *105*, 2640-2653.

10. Baselga, J. Targeting Tyrosine Kinases in Cancer: The Second Wave. *Science* **2006**, *312*, 1175-1178.

11. Cohen, P.; Cross, D.; Jänne, P. A. Kinase Drug Discovery 20 Years After Imatinib: Progress and Future Directions. *Nat. Rev. Drug Discov.* **2021**, *20*, 551-569.

12. Xie, Z.; Yang, X.; Duan, Y.; Han, J.; Liao, C. Small-Molecule Kinase Inhibitors for the Treatment of Nononcologic Diseases. *J. Med. Chem.* **2021**, *64*, 1283-1345.

13. Ayala-Aguilera, C. C.; Valero, T.; Lorente-Macias, A.; Baillache, D. J.; Croke, S.; Unciti-Broceta, A. Small Molecule Kinase Inhibitor Drugs (1995–2021): Medical Indication, Pharmacology, and Synthesis. *J. Med. Chem.* **2022**, *65,* 1047-1131.

14. Cuny, G. D. Kinase Inhibitors as Potential Therapeutics for Acute and Chronic Neurodegenerative Conditions. *Curr. Pharm. Des.* **2009**, *15*, 3919-3939.

15. Nikolic, I.; Leiva, M.; Sabio, G. The Role of Stress Kinases in Metabolic Disease. *Nat. Rev. Endocrinol.* **2020**, *16*, 697-716.

16. Zarrin, A. A.; Bao, K.; Lupardus, P.; Vucic, D. Kinase Inhibition in Autoimmunity and Inflammation. *Nat. Rev. Drug Discov*. **2021**, *20*, 39-63.

17. Roskoski Jr., R. Properties of FDA-Approved Small Molecule Protein Kinase Inhibitors: A 2024 Update. *Pharmacol. Res.* **2024**, *200*, 107059.

18. Xerxa, E.; Miljković, F.; Bajorath, J. Data-Driven Global Assessment of Protein Kinase Inhibitors with Emphasis on Covalent Compounds. *J. Med. Chem.* **2023**, *66*, 7657-7665.

19. Fedorov, O.; Müller, S.; Knapp, S. The (Un)Targeted Cancer Kinome. *Nat. Chem. Biol.* **2010**, *6*, 166-169.

20. Rodgers, G.; Austin, C.; Anderson, J.; Pawlyk, A.; Colvis, C.; Margolis, R.; Baker, J. Glimmers in Illuminating the Druggable Genome. *Nat. Rev. Drug Discov.* **2018**, *17*, 301-302.

21. Gomez, S. M.; Axtman, A. D.; Willson, T. M.; Major, M. B.; Townsend, R. R.; Sorger, P. K.; Johnson, G. L. Illuminating Function of the Understudied Druggable Kinome. *Drug Discovery Today* **2024**, *29*, 103881.

22. Illuminating the Druggable Genome, Understudied Proteins, Eligible Kinases. https://commonfund.nih.gov/idg/understudiedproteins (accessed June 30 2024).

23. Nguyen, D. T.; Mathias S.; Bologa C.; Brunak S.; Fernandez N.; Gaulton A.; Hersey A.; Holmes J.; Jensen L. J.; Karlsson, A. Pharos: Collating Protein Information to Shed Light on the Druggable Genome. *Nucl. Acids Res.* **2017**, *45*, D995-D1002.

24. Oprea, T.I.; Bologa, C.; Holmes, J.; Mathias, S.; Metzger, V. T.; Waller, A.; Yang, J. J.; Leach, A. R.; Jensen, L. J.; Kelleher, K. J.; Sheils, T. K.; Mathé, E.; Avram, S.; Edwards, J. S. Overview of the Knowledge Management Center for Illuminating the Druggable Genome. *Drug Discovery Today* **2024**, *29*, 103882.

25. Berginski, M. E.; Moret, N.; Liu, C.; Goldfarb, D.; Sorger, P. K.; Gomez, S. M. The Dark Kinase Knowledgebase: An Online Compendium of Knowledge and Experimental Results of Understudied Kinases. *Nucleic Acids Res*. **2021**, *49*, D529-D535.

26. Essegian, D.; Khurana, R.; Stathias, V.; Schürer, S. C. The Clinical Kinase Index: A Method to Prioritize Understudied Kinases as Drug Targets for the Treatment of Cancer. *Cell. Rep. Med.* **2020**, *1*, 100128.

27. Soleymani, S.; Gravel, N.; Huang, L. C.; Yeung, W.; Bozorgi, E.; Bendzunas, N. G.; Kochut, K. J.; Kannan, N. Dark Kinase Annotation, Mining, and Visualization Using the Protein Kinase Ontology. *PeerJ.* **2023**, 1, e16087.

28. Anderson, B.; Rosston, P.; Ong, H. W.; Hossain, M. A.; Davis-Gilbert, Z. W.; Drewry, D. H. How Many Kinases Are Druggable? A Review of Our Current Understanding. *Biochem. J.* **2023**, *480*, 1331-1363.

29. Frye, S. V. The Art of the Chemical Probe. *Nat. Chem. Biol.* **2010**, *6*, 159-161.

30. Bunnage, M. E.; Checkler, E. L. P.; Jones, L. H. Target Validation Using Chemical Probes. *Nat. Chem. Biol.* **2013**, *9*, 195-199.

31. Schwarz, D. M.; Gestwicki, J. E. Revisiting the "Art of the Chemical Probe". *ACS Chem. Biol.* **2018**, *13*, 1109-1110.

32. Drewry, D. H.; Wells, C. I.; Andrews, D. M.; Angell, R.; Al-Ali, H.; Axtman, A. D.; Capuzzi, S. J.; Elkins, J. M.; Ettmayer, P.; Frederiksen, M.; Gileadi, O.; Gray, N.; Hooper, A.; Knapp, S.; Laufer, S.; Luecking, U.; Michaelides, M.; Müller, S.; Muratov, E.; Denny, R. A.; Saikatendu, K. S.; Treiber, D. K.; Zuercher, W. J.; Willson, T. M. Progress Towards a Public Chemogenomic Set for Protein Kinases and a Call for Contributions. *PLoS One*. **2017**, *12*, e0181585.

33. Structural Genomics Consortium Chemical Probes. https://www.thesgc.org/chemical-probes (accessed July 05 2024).

34. Antolin, A. A.; Sanfelice, D.; Crisp, A.; Villasclaras Fernandez, E.; Mica, I. L.; Chen, Y.; Collins, I.; Edwards, A.; Müller, S.; Al-Lazikani, B.; Workman, P. The Chemical Probes Portal: An Expert Review-Based Public Resource to Empower Chemical Probe Assessment, Selection and Use. *Nucl. Acids Res.* **2023**, *51*, D1492-D1502.

35. Miljković, F.; Bajorath, J. Computational Analysis of Kinase Inhibitors Identifies Promiscuity Cliffs across the Human Kinome. *ACS Omega* **2018**, *3*, 17295-17308.

36. Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A Large-Scale Bioactivity Database for Drug Discovery. *Nucleic Acids Res.* **2012**, *40*, D1100-1107.

37. Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K**.** BindingDB: A Web-Accessible Database of Experimentally Determined Protein–Ligand Binding Affinities. *Nucleic Acids Res.* **2007**, *35*, D198-D201.

38. Yonchev, D.; Dimova, D.; Stumpfe, D.; Vogt, M.; Bajorath, J. Redundancy in Two Major Compound Databases. *Drug Discovery Today* **2018**, *23*, 1183-1186.

39. Reinecke, M.; Brear, P.; Vornholz, L.; Berger, B. T.; Seefried, F.; Wilhelm, S.; Samaras, P.; Gyenis, L.; Litchfield, D. W.; Médard, G.; Müller, S.; Ruland, J.; Hyvönen, M.; Wilhelm, M.; Kuster, B. Chemical Proteomics Reveals the Target Landscape of 1,000 kinase Inhibitors. *Nat. Chem. Biol.* **2024**, *20*, 577-585.

40. UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **2023**, *51*, D523-D531.

41. Weininger, D. SMILES, a Chemical Language and Information System. 1. Introduction to Methodology and Encoding Rules. *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31-36.

42. Heller, S. R.; McNaught, A.; Pletnev, I.; Stein, S.; Tchekhovskoi, D. InChI, the IUPAC International Chemical Identifier. *J. Cheminform.* **2015**, *7*, 23.

43. Eid, S.; Turk, S.; Volkamer, A.; Rippmann, F.; Fulle, S. KinMap: A Web-Based Tool for Interactive Navigation through Human Kinome Data. *BMC Bioinf.* **2017**, *18*, 16.

44. Wishart, D. S.; Knox, C.; Guo, A. C.; Shrivastava, S.; Hassanali, M.; Stothard, P.; Chang, Z.; Woolsey, J. DrugBank: A Comprehensive Resource for In Silico Drug Discovery and Exploration. *Nucleic Acids Res.* **2006**, *34*, D668-D672.

45. Avram, S.; Wilson, T. B.; Curpan, R., Halip, L.; Borota, A.; Bora, A.; Bologa, C. G.; Holmes, J., Knockel, J.; Yang, J. J.; Oprea, T. I. (2023). DrugCentral 2023 Extends Human Clinical Data and Integrates Veterinary Drugs. *Nucleic Acids Res.* **2023**, *51*, D1276–D1287.

46. Harding, S. D.; Armstrong, J. F.; Faccenda, E.; Southan, C.; Alexander, S. P. H.; Davenport, A. P.; Spedding, M.; Davies, J. A. The IUPHAR/BPS Guide to PHARMACOLOGY in 2024. *Nucleic Acids Res.* **2024**, *52*, D1438-D1449.

.

Supporting Information

Table S1. Chemically unexplored protein kinases.

UniProt ID	Protein kinase name
Q9UJY1	Heat shock protein beta-8
Q38SD2	Leucine-rich repeat serine/threonine-protein kinase 1
Q96LW2	Ribosomal protein S6 kinase-related protein
Q7Z7A4	PX domain-containing protein kinase-like protein
Q504Y2	Extracellular tyrosine-protein kinase PKDCC
Q9NSY0	Nuclear receptor-binding protein 2
Q6IBS0	Twinfilin-2
Q6J9G0	Tyrosine-protein kinase STYK1
Q8NCB2	CaM kinase-like vesicle-associated protein
Q8TAS1	Serine/threonine-protein kinase Kist
Q96QS6	Serine/threonine-protein kinase H2
Q9UHY1	Nuclear receptor-binding protein
Q8N2I9	Serine/threonine-protein kinase 40
Q9H5K3	Protein O-mannose kinase
O60229	Kalirin
Q15772	Striated muscle preferentially expressed protein kinase

Q496M5	Inactive serine/threonine-protein kinase PLK5
Q86TB3	Alpha-protein kinase 2
Q12792	Twinfilin-1
Q7Z2Y5	Nik-related protein kinase
Q96L96	Alpha-protein kinase 3
Q6ZS72	Protein PEAK3
P0C263	Serine/threonine-protein kinase SBK2
Q01973	Inactive tyrosine-protein kinase transmembrane receptor ROR1
Q9Y6S9	Ribosomal protein S6 kinase-like 1
O43930	Putative serine/threonine-protein kinase PRKY
Q96RU8	Tribbles homolog 1
Q02846	Retinal guanylyl cyclase 1
P51841	Retinal guanylyl cyclase 2
Q13308	Inactive tyrosine-protein kinase 7
Q6A1A2	Putative 3-phosphoinositide-dependent protein kinase 2
Q5JZY3	Ephrin type-A receptor 10
Q8NE28	Serine/threonine kinase-like domain-containing protein STKLD1
Q96C45	Serine/threonine-protein kinase ULK4
Q14296	Fas-activated serine/threonine kinase
Q86YV5	Inactive tyrosine-protein kinase PRAG1
Q6P3W7	SCY1-like protein 2

Q8N165	Serine/threonine-protein kinase PDIK1L
Q6ZMQ8	Serine/threonine-protein kinase LMTK1
P34925	Tyrosine-protein kinase RYK
Q96KG9	N-terminal kinase-like protein
Q8TEA7	TBC domain-containing protein kinase-like protein
Q8IWU2	Serine/threonine-protein kinase LMTK2
Q96RU7	Tribbles homolog 3
Q7Z695	Uncharacterized aarF domain-containing protein kinase 2
A0A0B4J2F2	Putative serine/threonine-protein kinase SIK1B
Q16671	Anti-Muellerian hormone type-2 receptor
Q8IWB6	Inactive serine/threonine-protein kinase TEX14
Q8IZE3	Protein-associating with the carboxyl-terminal domain of ezrin
Q5VST9	Obscurin
Q96QP1	Alpha-protein kinase 1
P11801	Serine/threonine-protein kinase H1
Q01974	Tyrosine-protein kinase transmembrane receptor ROR2
O14874	Branched-chain alpha-ketoacid dehydrogenase kinase
P25092	Guanylyl cyclase C
O75962	Triple functional domain protein
P00540	Proto-oncogene serine/threonine-protein kinase mos
Q9C0K7	STE20-related kinase adapter protein beta

Q7RTN6	STE20-related kinase adapter protein alpha
Q9Y6R4	Mitogen-activated protein kinase kinase kinase 4
Q96Q40	Cyclin-dependent kinase 15
Q5TCX8	Mitogen-activated protein kinase kinase kinase 21
Q9Y2H9	Microtubule-associated serine/threonine-protein kinase 1
Q00537	Cyclin-dependent kinase 17
Q8NFD2	Ankyrin repeat and protein kinase domain-containing protein 1
Q96D53	Atypical kinase COQ8B, mitochondrial
Q9Y2H1	Serine/threonine-protein kinase 38-like
P51956	Serine/threonine-protein kinase Nek3
P21127	Cyclin-dependent kinase 11B
Q86SG6	Serine/threonine-protein kinase Nek8
P20794	Serine/threonine-protein kinase MAK
Q86UX6	Serine/threonine-protein kinase 32C
O60566	Mitotic checkpoint serine/threonine-protein kinase BUB1 beta
P57078	Receptor-interacting serine/threonine-protein kinase 4
Q52WX2	Serine/threonine-protein kinase SBK1
Q15208	Serine/threonine-protein kinase 38
P57058	Hormonally up-regulated neu tumor-associated kinase
Q9C098	Serine/threonine-protein kinase DCLK3
Q6ZN16	Mitogen-activated protein kinase kinase kinase 15

Q07002	Cyclin-dependent kinase 18
O95747	Serine/threonine-protein kinase OSR1
Q96S38	Ribosomal protein S6 kinase delta-1
Q8TDR2	Serine/threonine-protein kinase 35
Q6DT37	Serine/threonine-protein kinase MRCK gamma
Q9Y2U5	Mitogen-activated protein kinase kinase kinase 2
Q9UL54	Serine/threonine-protein kinase TAO2
Q9BXM7	Serine/threonine-protein kinase PINK1, mitochondrial
Q8IVT5	Kinase suppressor of Ras 1
Q8TD19	Serine/threonine-protein kinase Nek9
Q9BXU1	Serine/threonine-protein kinase 31
Q15831	Serine/threonine-protein kinase STK11
Q99759	Mitogen-activated protein kinase kinase kinase 3
Q96S44	EKC/KEOPS complex subunit TP53RK
Q8WXR4	Myosin-IIIb
Q96GX5	Serine/threonine-protein kinase greatwall
Q9BXA6	Testis-specific serine/threonine-protein kinase 6
P0C264	Uncharacterized serine/threonine-protein kinase SBK3
Q9BRS2	Serine/threonine-protein kinase RIO1
Q6P5Z2	Serine/threonine-protein kinase N3
Q9UF33	Ephrin type-A receptor 6

Q8NG66	Serine/threonine-protein kinase Nek11
Q8IV63	Serine/threonine-protein kinase VRK3
Q15131	Cyclin-dependent kinase 10
Q58A45	PAN2-PAN3 deadenylation complex subunit PAN3
Q8WU08	Serine/threonine-protein kinase 32A
Q8N568	Serine/threonine-protein kinase DCLK2
Q9UKI8	Serine/threonine-protein kinase tousled-like 1
Q6XUX3	Dual serine/threonine and tyrosine protein kinase
P31152	Mitogen-activated protein kinase 4
P14616	Insulin receptor-related protein
Q9NRP7	Serine/threonine-protein kinase 36
Q96RG2	PAS domain-containing serine/threonine-protein kinase
Q3MIX3	Uncharacterized aarF domain-containing protein kinase 5
Q8NEV4	Myosin-IIIa
Q96QT4	Transient receptor potential cation channel subfamily M member 7
Q8IZL9	Cyclin-dependent kinase 20
Q9BX84	Transient receptor potential cation channel subfamily M member 6
Q9UPZ9	Serine/threonine-protein kinase ICK
Q6SA08	Testis-specific serine/threonine-protein kinase 4

Different subsets of chemically unexplored PKs are color-coded as follows: black, PKs without PKI records; blue, PKs with incomplete/inconclusive PKI activity records; red, PKs with no qualifying low- and high-confidence PKIs following data curation; green, PKs exclusively annotated with promiscuous PKIs.

Table S2. Clinical protein kinases.

UniProt ID	Protein kinase name
P25092	Guanylyl cyclase C
P35916	Vascular endothelial growth factor receptor 3
P16591	Tyrosine-protein kinase Fer
Q08881	Tyrosine-protein kinase ITK/TSK
Q96GD4	Aurora kinase B
P24941	Cyclin-dependent kinase 2
P21860	Receptor tyrosine-protein kinase erbB-3
P00519	Tyrosine-protein kinase ABL1
Q8IW41	MAP kinase-activated protein kinase 5
Q16584	Mitogen-activated protein kinase kinase kinase 11

Q02763	Angiopoietin-1 receptor
P45983	Mitogen-activated protein kinase 8
P49840	Glycogen synthase kinase-3 alpha
Q14289	Protein-tyrosine kinase 2-beta
P12931	Proto-oncogene tyrosine-protein kinase Src
P29597	Non-receptor tyrosine-protein kinase TYK2
Q15759	Mitogen-activated protein kinase 11
Q04759	Protein kinase C theta type
Q07912	Activated CDC42 kinase 1
Q9Y243	RAC-gamma serine/threonine-protein kinase
O96017	Serine/threonine-protein kinase Chk2
P49137	MAP kinase-activated protein kinase 2
P04629	High affinity nerve growth factor receptor
O15530	3-phosphoinositide-dependent protein kinase 1
P49841	Glycogen synthase kinase-3 beta
P42338	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit beta isoform

P50750	Cyclin-dependent kinase 9
P42345	Serine/threonine-protein kinase mTOR
Q00535	Cyclin-dependent kinase 5
P52333	Tyrosine-protein kinase JAK3
P16234	Platelet-derived growth factor receptor alpha
P37023	Serine/threonine-protein kinase receptor R3
Q06187	Tyrosine-protein kinase BTK
O75116	Rho-associated protein kinase 2
O43283	Mitogen-activated protein kinase kinase kinase 13
Q13882	Protein-tyrosine kinase 6
P50613	Cyclin-dependent kinase 7
Q12852	Mitogen-activated protein kinase kinase kinase 12
Q9H3Y6	Tyrosine-protein kinase Srms
P17948	Vascular endothelial growth factor receptor 1
O96013	Serine/threonine-protein kinase PAK 4
P07947	Tyrosine-protein kinase Yes
P10721	Mast/stem cell growth factor receptor Kit

Q15418	Ribosomal protein S6 kinase alpha-1
P21802	Fibroblast growth factor receptor 2
P00533	Epidermal growth factor receptor
P41279	Mitogen-activated protein kinase kinase kinase 8
O60674	Tyrosine-protein kinase JAK2
Q00534	Cyclin-dependent kinase 6
P36507	Dual specificity mitogen-activated protein kinase kinase 2
P54760	Ephrin type-B receptor 4
P07332	Tyrosine-protein kinase Fes/Fps
P78527	DNA-dependent protein kinase catalytic subunit
O14965	Aurora kinase A
P36897	TGF-beta receptor type-1
P08069	Insulin-like growth factor 1 receptor
P42336	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform
P04049	RAF proto-oncogene serine/threonine-protein kinase
P42685	Tyrosine-protein kinase FRK

Q9NWZ3	Interleukin-1 receptor-associated kinase 4
Q13546	Receptor-interacting serine/threonine-protein kinase 1
P48736	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit gamma isoform
P31749	RAC-alpha serine/threonine-protein kinase
Q02779	Mitogen-activated protein kinase kinase kinase 10
P06493	Cyclin-dependent kinase 1
P45984	Mitogen-activated protein kinase 9
P29376	Leukocyte tyrosine kinase receptor
P22612	cAMP-dependent protein kinase catalytic subunit gamma
Q9NYV4	Cyclin-dependent kinase 12
Q9UM73	ALK tyrosine kinase receptor
Q15118	[Pyruvate dehydrogenase
Q12866	Tyrosine-protein kinase Mer
Q05655	Protein kinase C delta type
Q13535	Serine/threonine-protein kinase ATR
P06239	Tyrosine-protein kinase Lck

P15056	Serine/threonine-protein kinase B-raf
P04626	Receptor tyrosine-protein kinase erbB-2
Q04912	Macrophage-stimulating protein receptor
P07333	Macrophage colony-stimulating factor 1 receptor
P27361	Mitogen-activated protein kinase 3
P08922	Proto-oncogene tyrosine-protein kinase ROS
Q9UQB9	Aurora kinase C
Q16620	BDNF/NT-3 growth factors receptor
P53779	Mitogen-activated protein kinase 10
P68400	Casein kinase II subunit alpha
Q15303	Receptor tyrosine-protein kinase erbB-4
Q16539	Mitogen-activated protein kinase 14
Q16832	Discoidin domain-containing receptor 2
P11362	Fibroblast growth factor receptor 1
P36888	Receptor-type tyrosine-protein kinase FLT3
P08631	Tyrosine-protein kinase HCK
Q86V86	Serine/threonine-protein kinase pim-3

Q13464	Rho-associated protein kinase 1
P23443	Ribosomal protein S6 kinase beta-1
P23458	Tyrosine-protein kinase JAK1
O14757	Serine/threonine-protein kinase Chk1
O14920	Inhibitor of nuclear factor kappa-B kinase subunit beta
P22455	Fibroblast growth factor receptor 4
P29317	Ephrin type-A receptor 2
O00311	Cell division cycle 7-related protein kinase
Q9BUB5	MAP kinase-interacting serine/threonine-protein kinase 1
P07949	Proto-oncogene tyrosine-protein kinase receptor Ret
P30530	Tyrosine-protein kinase receptor UFO
P07948	Tyrosine-protein kinase Lyn
P35968	Vascular endothelial growth factor receptor 2
O00444	Serine/threonine-protein kinase PLK4
P06213	Insulin receptor
Q99683	Mitogen-activated protein kinase kinase kinase 5
P06241	Tyrosine-protein kinase Fyn

Q13233	Mitogen-activated protein kinase kinase kinase 1
Q16288	NT-3 growth factor receptor
P10398	Serine/threonine-protein kinase A-Raf
P28482	Mitogen-activated protein kinase 1
P30291	Wee1-like protein kinase
O00329	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit delta isoform
P05771	Protein kinase C beta type
P80192	Mitogen-activated protein kinase kinase kinase 9
P43405	Tyrosine-protein kinase SYK
Q05397	Focal adhesion kinase 1
P08581	Hepatocyte growth factor receptor
P09619	Platelet-derived growth factor receptor beta
P16066	Atrial natriuretic peptide receptor 1
P31751	RAC-beta serine/threonine-protein kinase
P11802	Cyclin-dependent kinase 4
P33981	Dual specificity protein kinase TTK

Q13131	5'-AMP-activated protein kinase catalytic subunit alpha-1
P20594	Atrial natriuretic peptide receptor 2
Q02750	Dual specificity mitogen-activated protein kinase kinase 1
P53350	Serine/threonine-protein kinase PLK1
P22607	Fibroblast growth factor receptor 3
P49674	Casein kinase I isoform epsilon

Table S3. Other chemically explored protein kinases.

UniProt ID	**Protein kinase name**
Q16654	[Pyruvate dehydrogenase
Q5TCY1	Tau-tubulin kinase 1
Q13164	Mitogen-activated protein kinase 7
Q13627	Dual specificity tyrosine-phosphorylation-regulated kinase 1A
O14733	Dual specificity mitogen-activated protein kinase kinase 7
Q15349	Ribosomal protein S6 kinase alpha-2
Q16513	Serine/threonine-protein kinase N2

Q16659	Mitogen-activated protein kinase 6
P48730	Casein kinase I isoform delta
Q15746	Myosin light chain kinase, smooth muscle
Q13418	Integrin-linked protein kinase
O00238	Bone morphogenetic protein receptor type-1B
P41240	Tyrosine-protein kinase CSK
Q59H18	Serine/threonine-protein kinase TNNI3K
P51955	Serine/threonine-protein kinase Nek2
Q15120	[Pyruvate dehydrogenase
Q99640	Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase
O15264	Mitogen-activated protein kinase 13
Q9UHD2	Serine/threonine-protein kinase TBK1
Q9Y4K4	Mitogen-activated protein kinase kinase kinase kinase 5
Q12851	Mitogen-activated protein kinase kinase kinase kinase 2
P57059	Serine/threonine-protein kinase SIK1
Q13470	Non-receptor tyrosine-protein kinase TNK1
Q9Y2K2	Serine/threonine-protein kinase SIK3
P19784	Casein kinase II subunit alpha'
P49760	Dual specificity protein kinase CLK2
O60285	NUAK family SNF1-like kinase 1

P43250	G protein-coupled receptor kinase 6
O95835	Serine/threonine-protein kinase LATS1
Q7L7X3	Serine/threonine-protein kinase TAO1
Q13153	Serine/threonine-protein kinase PAK 1
O43318	Mitogen-activated protein kinase kinase kinase 7
Q15139	Serine/threonine-protein kinase D1
P17612	cAMP-dependent protein kinase catalytic subunit alpha
Q13315	Serine-protein kinase ATM
Q9BWU1	Cyclin-dependent kinase 19
O43781	Dual specificity tyrosine-phosphorylation-regulated kinase 3
P45985	Dual specificity mitogen-activated protein kinase kinase 4
Q9BZL6	Serine/threonine-protein kinase D2
O75582	Ribosomal protein S6 kinase alpha-5
O00418	Eukaryotic elongation factor 2 kinase
Q96PY6	Serine/threonine-protein kinase Nek1
Q9Y572	Receptor-interacting serine/threonine-protein kinase 3
P53778	Mitogen-activated protein kinase 12
Q9NZJ5	Eukaryotic translation initiation factor 2-alpha kinase 3
P54646	5'-AMP-activated protein kinase catalytic subunit alpha-2
Q9BQI3	Eukaryotic translation initiation factor 2-alpha kinase 1
O75460	Serine/threonine-protein kinase/endoribonuclease IRE1

P19525	Interferon-induced, double-stranded RNA-activated protein kinase
P51451	Tyrosine-protein kinase Blk
Q9BYT3	Serine/threonine-protein kinase 33
P37173	TGF-beta receptor type-2
Q8NEV1	Casein kinase II subunit alpha 3
P27448	MAP/microtubule affinity-regulating kinase 3
O14976	Cyclin-G-associated kinase
P51812	Ribosomal protein S6 kinase alpha-3
O94804	Serine/threonine-protein kinase 10
O15111	Inhibitor of nuclear factor kappa-B kinase subunit alpha
Q9NYL2	Mitogen-activated protein kinase kinase kinase 20
P48729	Casein kinase I isoform alpha
O43683	Mitotic checkpoint serine/threonine-protein kinase BUB1
O94768	Serine/threonine-protein kinase 17B
P49336	Cyclin-dependent kinase 8
P53667	LIM domain kinase 1
Q9H2K8	Serine/threonine-protein kinase TAO3
Q13557	Calcium/calmodulin-dependent protein kinase type II subunit delta
O00141	Serine/threonine-protein kinase Sgk1
Q06418	Tyrosine-protein kinase receptor TYRO3
O94921	Cyclin-dependent kinase 14

Q96RR4	Calcium/calmodulin-dependent protein kinase kinase 2
Q6IQ55	Tau-tubulin kinase 2
P43403	Tyrosine-protein kinase ZAP-70
P25098	Beta-adrenergic receptor kinase 1
Q00526	Cyclin-dependent kinase 3
O43353	Receptor-interacting serine/threonine-protein kinase 2
Q9UKE5	TRAF2 and NCK-interacting protein kinase
O95819	Mitogen-activated protein kinase kinase kinase kinase 4
Q9P1W9	Serine/threonine-protein kinase pim-2
P42681	Tyrosine-protein kinase TXK
O75385	Serine/threonine-protein kinase ULK1
Q04771	Activin receptor type-1
Q9UK32	Ribosomal protein S6 kinase alpha-6
Q14004	Cyclin-dependent kinase 13
Q99558	Mitogen-activated protein kinase kinase kinase 14
Q96L34	MAP/microtubule affinity-regulating kinase 4
P36896	Activin receptor type-1B
Q13043	Serine/threonine-protein kinase 4
Q9HAZ1	Dual specificity protein kinase CLK4
P41743	Protein kinase C iota type
Q9Y5S2	Serine/threonine-protein kinase MRCK beta

Q9UQ07	MAPK/MAK/MRK overlapping kinase
Q9HBH9	MAP kinase-interacting serine/threonine-protein kinase 2
Q05513	Protein kinase C zeta type
O00443	Phosphatidylinositol 4-phosphate 3-kinase C2 domain-containing subunit alpha
Q9H2G2	STE20-like serine/threonine-protein kinase
Q96SB4	SRSF protein kinase 1
P49759	Dual specificity protein kinase CLK1
Q13188	Serine/threonine-protein kinase 3
P49761	Dual specificity protein kinase CLK3
Q8NB16	Mixed lineage kinase domain-like protein
Q96KB5	Lymphokine-activated killer T-cell-originated protein kinase
P24723	Protein kinase C eta type
Q2M2I8	AP2-associated protein kinase 1
P17252	Protein kinase C alpha type
Q5VT25	Serine/threonine-protein kinase MRCK alpha
Q8IVH8	Mitogen-activated protein kinase kinase kinase kinase 3
P36894	Bone morphogenetic protein receptor type-1A
Q15119	Pyruvate dehydrogenase
Q08345	Epithelial discoidin domain-containing receptor 1
P53671	LIM domain kinase 2

P51813	Cytoplasmic tyrosine-protein kinase BMX
Q13976	cGMP-dependent protein kinase 1
P11309	Serine/threonine-protein kinase pim-1
Q9NYY3	Serine/threonine-protein kinase PLK2
Q9UQM7	Calcium/calmodulin-dependent protein kinase type II subunit alpha
Q8IU85	Calcium/calmodulin-dependent protein kinase type 1D
P54764	Ephrin type-A receptor 4
P34947	G protein-coupled receptor kinase 5
Q9Y463	Dual specificity tyrosine-phosphorylation-regulated kinase 1B
P42680	Tyrosine-protein kinase Tec
Q5S007	Leucine-rich repeat serine/threonine-protein kinase 2
Q9H2X6	Homeodomain-interacting protein kinase 2
Q14680	Maternal embryonic leucine zipper kinase
Q92918	Mitogen-activated protein kinase kinase kinase kinase 1
Q02156	Protein kinase C epsilon type
P54753	Ephrin type-B receptor 3
Q9BVS4	Serine/threonine-protein kinase RIO2
Q9UEE5	Serine/threonine-protein kinase 17A
P51617	Interleukin-1 receptor-associated kinase 1
P42684	Tyrosine-protein kinase ABL2
Q9H0K1	Serine/threonine-protein kinase SIK2

P05129	Protein kinase C gamma type
Q9P2K8	eIF-2-alpha kinase GCN2
Q9H4B4	Serine/threonine-protein kinase PLK3
Q8TF76	Serine/threonine-protein kinase haspin
Q92630	Dual specificity tyrosine-phosphorylation-regulated kinase 2
Q14164	Inhibitor of nuclear factor kappa-B kinase subunit epsilon

The union of clinical PKs in Table S2 and other PKs reported in Table S3 represents chemically explored PKs.

Table S4. Chemically underexplored protein kinases.

UniProt ID	Protein kinase name
Q6P2M8	Calcium/calmodulin-dependent protein kinase type 1B
P35626	G protein-coupled receptor kinase 3
Q9NQU5	Serine/threonine-protein kinase PAK 6
Q13163	Dual specificity mitogen-activated protein kinase kinase 5
P53355	Death-associated protein kinase 1
Q9H4A3	Serine/threonine-protein kinase WNK1
Q8TDC3	Serine/threonine-protein kinase BRSK1
Q9BYP7	Serine/threonine-protein kinase WNK3
Q13237	cGMP-dependent protein kinase 2
Q6VAB6	Kinase suppressor of Ras 2
O15075	Serine/threonine-protein kinase DCLK1
Q8NE63	Homeodomain-interacting protein kinase 4
P32298	G protein-coupled receptor kinase 4
Q9H422	Homeodomain-interacting protein kinase 3
Q13554	Calcium/calmodulin-dependent protein kinase type II subunit beta
Q96BR1	Serine/threonine-protein kinase Sgk3
Q05823	2-5A-dependent ribonuclease
Q00532	Cyclin-dependent kinase-like 1

P35590	Tyrosine-protein kinase receptor Tie-1
Q9Y3S1	Serine/threonine-protein kinase WNK2
O43293	Death-associated protein kinase 3
Q8IYT8	Serine/threonine-protein kinase ULK2
Q13705	Activin receptor type-2B
Q9NRH2	SNF-related serine/threonine-protein kinase
P29323	Ephrin type-B receptor 2
Q15569	Dual specificity testis-specific protein kinase 1
P0C1S8	Wee1-like protein kinase 2
Q16566	Calcium/calmodulin-dependent protein kinase type IV
Q96PF2	Testis-specific serine/threonine-protein kinase 2
Q15835	Rhodopsin kinase GRK1
O15146	Muscle, skeletal receptor tyrosine-protein kinase
P27037	Activin receptor type-2A
Q96PN8	Testis-specific serine/threonine-protein kinase 3
Q9H093	NUAK family SNF1-like kinase 2
Q96Q15	Serine/threonine-protein kinase SMG1
Q8IWQ3	Serine/threonine-protein kinase BRSK2
Q9Y616	Interleukin-1 receptor-associated kinase 3
Q99986	Serine/threonine-protein kinase VRK1
Q32MK0	Myosin light chain kinase 3

Q9H1R3	Myosin light chain kinase 2, skeletal/cardiac muscle
O15021	Microtubule-associated serine/threonine-protein kinase 4
Q96S53	Dual specificity testis-specific protein kinase 2
P46734	Dual specificity mitogen-activated protein kinase kinase 3
Q96Q04	Serine/threonine-protein kinase LMTK3
Q9UQ88	Cyclin-dependent kinase 11A
Q9UBS0	Ribosomal protein S6 kinase beta-2
Q15375	Ephrin type-A receptor 7
Q9NR20	Dual specificity tyrosine-phosphorylation-regulated kinase 4
Q9Y6E0	Serine/threonine-protein kinase 24
Q8IY84	Serine/threonine-protein kinase NIM1
Q16644	MAP kinase-activated protein kinase 3
Q16816	Phosphorylase b kinase gamma catalytic chain, skeletal muscle/heart isoform
Q7KZI7	Serine/threonine-protein kinase MARK2
Q86Z02	Homeodomain-interacting protein kinase 1
Q9UBE8	Serine/threonine-protein kinase NLK
O14936	Peripheral plasma membrane protein CASK
Q13523	Serine/threonine-protein kinase PRP4 homolog
Q9H792	Inactive tyrosine-protein kinase PEAK1
Q86Y07	Serine/threonine-protein kinase VRK2

P15735	Phosphorylase b kinase gamma catalytic chain, liver/testis isoform
O75914	Serine/threonine-protein kinase PAK 3
O76039	Cyclin-dependent kinase-like 5
Q9P0L2	Serine/threonine-protein kinase MARK1
Q8TD08	Mitogen-activated protein kinase 15
O15197	Ephrin type-B receptor 6
Q09013	Myotonin-protein kinase
Q8IVW4	Cyclin-dependent kinase-like 3
Q99570	Phosphoinositide 3-kinase regulatory subunit 4
O14578	Citron Rho-interacting kinase
Q92519	Tribbles homolog 2
Q86TW2	AarF domain-containing protein kinase 1
P51957	Serine/threonine-protein kinase Nek4
P29322	Ephrin type-A receptor 8
Q9NRM7	Serine/threonine-protein kinase LATS2
Q92772	Cyclin-dependent kinase-like 2
P42679	Megakaryocyte-associated tyrosine-protein kinase
P54762	Ephrin type-B receptor 1
P09769	Tyrosine-protein kinase Fgr
Q8TDX7	Serine/threonine-protein kinase Nek7
P78362	SRSF protein kinase 2

Q9Y6M4	Casein kinase I isoform gamma-3
O95382	Mitogen-activated protein kinase kinase kinase 6
Q8NI60	Atypical kinase COQ8A, mitochondrial
P22694	cAMP-dependent protein kinase catalytic subunit beta
Q9BXA7	Testis-specific serine/threonine-protein kinase 1
Q76MJ5	Serine/threonine-protein kinase/endoribonuclease IRE2
Q9UPE1	SRSF protein kinase 3
Q9NY57	Serine/threonine-protein kinase 32B
Q9P286	Serine/threonine-protein kinase PAK 5
P78368	Casein kinase I isoform gamma-2
Q9P289	Serine/threonine-protein kinase 26
Q96NX5	Calcium/calmodulin-dependent protein kinase type 1G
P51817	cAMP-dependent protein kinase catalytic subunit PRKX
O60307	Microtubule-associated serine/threonine-protein kinase 3
Q14012	Calcium/calmodulin-dependent protein kinase type 1
Q96J92	Serine/threonine-protein kinase WNK4
Q9HCP0	Casein kinase I isoform gamma-1
O94806	Serine/threonine-protein kinase D3
Q8N4C8	Misshapen-like kinase 1
O43187	Interleukin-1 receptor-associated kinase-like 2
Q9HBY8	Serine/threonine-protein kinase Sgk2

Q6ZWH5	Serine/threonine-protein kinase Nek10
Q8N5S9	Calcium/calmodulin-dependent protein kinase kinase 1
Q16512	Serine/threonine-protein kinase N1
Q6P3R8	Serine/threonine-protein kinase Nek5
P54756	Ephrin type-A receptor 5
Q8N752	Casein kinase I isoform alpha-like
Q8WTQ7	Rhodopsin kinase GRK7
Q13555	Calcium/calmodulin-dependent protein kinase type II subunit gamma
Q8NER5	Activin receptor type-1C
Q9HC98	Serine/threonine-protein kinase Nek6
Q6PHR2	Serine/threonine-protein kinase ULK3
Q56UN5	Mitogen-activated protein kinase kinase kinase 19
Q9UEW8	STE20/SPS1-related proline-alanine-rich protein kinase
Q13873	Bone morphogenetic protein receptor type-2
Q5MAI5	Cyclin-dependent kinase-like 4
P52564	Dual specificity mitogen-activated protein kinase kinase 6
Q86UE8	Serine/threonine-protein kinase tousled-like 2
P21709	Ephrin type-A receptor 1
Q6P0Q8	Microtubule-associated serine/threonine-protein kinase 2
Q13177	Serine/threonine-protein kinase PAK 2

O75716	Serine/threonine-protein kinase 16
O00506	Serine/threonine-protein kinase 25
Q86YV6	Myosin light chain kinase family member 4
Q9NSY1	BMP-2-inducible protein kinase
Q00536	Cyclin-dependent kinase 16
O75676	Ribosomal protein S6 kinase alpha-4
P29320	Ephrin type-A receptor 3
Q9UIK4	Death-associated protein kinase 2
O14730	Serine/threonine-protein kinase RIO3

Table S5. Chemically un- or underexplored protein kinases categorized as understudied.

UniProt ID	Protein kinase name
P21127	Cyclin-dependent kinase 11B
O14874	Branched-chain alpha-ketoacid dehydrogenase kinase
Q6SA08	Testis-specific serine/threonine-protein kinase 4
Q58A45	PAN2-PAN3 deadenylation complex subunit PAN3
Q6DT37	Serine/threonine-protein kinase MRCK gamma
Q9NSY0	Nuclear receptor-binding protein 2
Q8IWU2	Serine/threonine-protein kinase LMTK2
Q86TB3	Alpha-protein kinase 2
Q9NRP7	Serine/threonine-protein kinase 36
Q7Z695	Uncharacterized aarF domain-containing protein kinase 2
Q96C45	Serine/threonine-protein kinase ULK4
Q8NE28	Serine/threonine kinase-like domain-containing protein STKLD1
Q6XUX3	Dual serine/threonine and tyrosine protein kinase
Q9Y6S9	Ribosomal protein S6 kinase-like 1
Q8IZE3	Protein-associating with the carboxyl-terminal domain of ezrin
Q96QS6	Serine/threonine-protein kinase H2
Q9BXA6	Testis-specific serine/threonine-protein kinase 6
P31152	Mitogen-activated protein kinase 4

Q9UKI8	Serine/threonine-protein kinase tousled-like 1
Q8N2I9	Serine/threonine-protein kinase 40
Q8IZL9	Cyclin-dependent kinase 20
Q8N165	Serine/threonine-protein kinase PDIK1L
Q96S38	Ribosomal protein S6 kinase delta-1
P0C263	Serine/threonine-protein kinase SBK2
Q96KG9	N-terminal kinase-like protein
Q38SD2	Leucine-rich repeat serine/threonine-protein kinase 1
Q9BRS2	Serine/threonine-protein kinase RIO1
Q8NCB2	CaM kinase-like vesicle-associated protein
Q6P3W7	SCY1-like protein 2
Q96D53	Atypical kinase COQ8B, mitochondrial
Q8IV63	Serine/threonine-protein kinase VRK3
Q5TCX8	Mitogen-activated protein kinase kinase kinase 21
Q96Q40	Cyclin-dependent kinase 15
Q6P5Z2	Serine/threonine-protein kinase N3
Q9Y2H1	Serine/threonine-protein kinase 38-like
P0C264	Uncharacterized serine/threonine-protein kinase SBK3
Q6ZN16	Mitogen-activated protein kinase kinase kinase 15
Q7Z2Y5	Nik-related protein kinase
Q8TD19	Serine/threonine-protein kinase Nek9

Q96L96	Alpha-protein kinase 3
P51956	Serine/threonine-protein kinase Nek3
Q86YV5	Inactive tyrosine-protein kinase PRAG1
Q9UPZ9	Serine/threonine-protein kinase ICK
Q5VST9	Obscurin
Q15131	Cyclin-dependent kinase 10
Q96LW2	Ribosomal protein S6 kinase-related protein
Q9UL54	Serine/threonine-protein kinase TAO2
Q7Z7A4	PX domain-containing protein kinase-like protein
Q52WX2	Serine/threonine-protein kinase SBK1
Q96S44	EKC/KEOPS complex subunit TP53RK
P11801	Serine/threonine-protein kinase H1
Q8TEA7	TBC domain-containing protein kinase-like protein
Q3MIX3	Uncharacterized aarF domain-containing protein kinase 5
Q07002	Cyclin-dependent kinase 18
Q9C098	Serine/threonine-protein kinase DCLK3
Q8NG66	Serine/threonine-protein kinase Nek11
Q9H5K3	Protein O-mannose kinase
Q86SG6	Serine/threonine-protein kinase Nek8
Q00537	Cyclin-dependent kinase 17
Q86UX6	Serine/threonine-protein kinase 32C

Q9BXU1	Serine/threonine-protein kinase 31
Q8WU08	Serine/threonine-protein kinase 32A
Q6P2M8	Calcium/calmodulin-dependent protein kinase type 1B
Q9NQU5	Serine/threonine-protein kinase PAK 6
Q8TDC3	Serine/threonine-protein kinase BRSK1
Q9BYP7	Serine/threonine-protein kinase WNK3
Q8NE63	Homeodomain-interacting protein kinase 4
Q9H422	Homeodomain-interacting protein kinase 3
Q00532	Cyclin-dependent kinase-like 1
Q05823	2-5A-dependent ribonuclease
Q9Y3S1	Serine/threonine-protein kinase WNK2
Q15569	Dual specificity testis-specific protein kinase 1
P0C1S8	Wee1-like protein kinase 2
Q96PF2	Testis-specific serine/threonine-protein kinase 2
Q96PN8	Testis-specific serine/threonine-protein kinase 3
Q9H093	NUAK family SNF1-like kinase 2
Q8IWQ3	Serine/threonine-protein kinase BRSK2
O15021	Microtubule-associated serine/threonine-protein kinase 4
Q96S53	Dual specificity testis-specific protein kinase 2
Q96Q04	Serine/threonine-protein kinase LMTK3
Q9UQ88	Cyclin-dependent kinase 11A

Q9NR20	Dual specificity tyrosine-phosphorylation-regulated kinase 4
Q8IY84	Serine/threonine-protein kinase NIM1
Q9Y6E0	Serine/threonine-protein kinase 24
Q16816	Phosphorylase b kinase gamma catalytic chain, skeletal muscle/heart isoform
Q86Z02	Homeodomain-interacting protein kinase 1
Q13523	Serine/threonine-protein kinase PRP4 homolog
O76039	Cyclin-dependent kinase-like 5
P15735	Phosphorylase b kinase gamma catalytic chain, liver/testis isoform
O75914	Serine/threonine-protein kinase PAK 3
Q86Y07	Serine/threonine-protein kinase VRK2
Q9P0L2	Serine/threonine-protein kinase MARK1
Q8TD08	Mitogen-activated protein kinase 15
Q8IVW4	Cyclin-dependent kinase-like 3
Q86TW2	AarF domain-containing protein kinase 1
P51957	Serine/threonine-protein kinase Nek4
Q92772	Cyclin-dependent kinase-like 2
Q8TDX7	Serine/threonine-protein kinase Nek7
Q9Y6M4	Casein kinase I isoform gamma-3
Q8NI60	Atypical kinase COQ8A, mitochondrial
P22694	cAMP-dependent protein kinase catalytic subunit beta

Q9BXA7	Testis-specific serine/threonine-protein kinase 1
Q76MJ5	Serine/threonine-protein kinase/endoribonuclease IRE2
Q9UPE1	SRSF protein kinase 3
Q9NY57	Serine/threonine-protein kinase 32B
Q9P286	Serine/threonine-protein kinase PAK 5
P78368	Casein kinase I isoform gamma-2
Q96NX5	Calcium/calmodulin-dependent protein kinase type 1G
O60307	Microtubule-associated serine/threonine-protein kinase 3
Q9HCP0	Casein kinase I isoform gamma-1
Q6ZWH5	Serine/threonine-protein kinase Nek10
Q8N5S9	Calcium/calmodulin-dependent protein kinase kinase 1
Q6P3R8	Serine/threonine-protein kinase Nek5
Q8N752	Casein kinase I isoform alpha-like
Q9HC98	Serine/threonine-protein kinase Nek6
Q5MAI5	Cyclin-dependent kinase-like 4
Q86UE8	Serine/threonine-protein kinase tousled-like 2
Q6P0Q8	Microtubule-associated serine/threonine-protein kinase 2
Q00536	Cyclin-dependent kinase 16
O14730	Serine/threonine-protein kinase RIO3

Chemically un- or underexplored PKs previously categorized as understudied are reported in black and blue, respectively.

Table S6. Chemically explored protein kinases categorized as understudied.

UniProt ID	Protein kinase name	Approved drugs	Clinical PKIs	Low-confidence PKIs	High-confidence PKIs
Q9HBH9	MAP kinase-interacting serine/threonine-protein kinase 2			1745	1292
Q9BUB5	MAP kinase-interacting serine/threonine-protein kinase 1		1	1669	1134
Q99558	Mitogen-activated protein kinase kinase kinase 14			1190	1062
Q04759	Protein kinase C theta type	1	3	1210	812
Q9Y463	Dual specificity tyrosine-phosphorylation-regulated kinase 1B			662	527
Q92630	Dual specificity tyrosine-phosphorylation-regulated kinase 2			412	307
Q9HAZ1	Dual specificity protein kinase CLK4			925	263
Q96PY6	Serine/threonine-protein kinase Nek1			300	244

P19784	Casein kinase II subunit alpha			697	175
Q9NYV4	Cyclin-dependent kinase 12		4	995	147
Q8IU85	Calcium/calmodulin-dependent protein kinase type 1D			134	130
O00418	Eukaryotic elongation factor 2 kinase			127	116
O94768	Serine/threonine-protein kinase 17B			179	115
O43781	Dual specificity tyrosine-phosphorylation-regulated kinase 3			147	114
Q9UEE5	Serine/threonine-protein kinase 17A			159	100
Q9Y5S2	Serine/threonine-protein kinase MRCK beta			85	85
Q13188	Serine/threonine-protein kinase 3			184	81
Q96RR4	Calcium/calmodulin-dependent protein kinase kinase 2			90	75
P27448	MAP/microtubule affinity-regulating kinase 3			87	70
Q5VT25	Serine/threonine-protein kinase MRCK alpha			76	62
Q99640	Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase			65	61

Q7L7X3	Serine/threonine-protein kinase TAO1			108	59
Q5TCY1	Tau-tubulin kinase 1			57	57
Q6IQ55	Tau-tubulin kinase 2			51	51
Q14004	Cyclin-dependent kinase 13			103	47
P49761	Dual specificity protein kinase CLK3			740	40
Q9BVS4	Serine/threonine-protein kinase RIO2			59	32
Q96L34	MAP/microtubule affinity-regulating kinase 4			51	30
Q9BWU1	Cyclin-dependent kinase 19			339	25
Q9BYT3	Serine/threonine-protein kinase 33			189	23
Q02779	Mitogen-activated protein kinase kinase kinase 10		1	19	19
P29376	Leukocyte tyrosine kinase receptor	4		18	11
O94921	Cyclin-dependent kinase 14			102	10
Q8NEV1	Casein kinase II subunit alpha 3			182	9
P22612	cAMP-dependent protein kinase catalytic subunit gamma		1	252	6

Shown are chemically explored PKs previously categorized as understudied. The PKs are ranked by the number of available high-confidence PKIs. Clinical PKs are colored in orange.