

# High-Throughput Kinase Activity Mapping (HT-KAM) system: analysis of phospho-catalytic profiles

Jean-Philippe Coppé (✉ [jean-philippe.coppe@ucsf.edu](mailto:jean-philippe.coppe@ucsf.edu))

Kinome Activity Mapping

Christina Yau

Kinome Activity Mapping

Denise M. Wolf

Kinome Activity Mapping

---

## Method Article

**Keywords:** kinase enzymes, phospho-catalytic activity screening, peptide sensors, combinatorial profiling, cell extracts, tissue biospecimen extracts, functional proteomics, signaling networks

**Posted Date:** June 12th, 2019

**DOI:** <https://doi.org/10.1038/protex.2019.030>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

Phosphorylation networks intimately regulate mechanisms of response to therapies. Identifying the phospho-catalytic activities of kinases in biological samples remains a challenge. Here, we introduce a high-throughput system to detect kinases' enzymatic activity using their biological peptide targets as phospho-sensors. Libraries of peptides operate as specific, distinct combinatorial peptide sets that simultaneously distinguish and measure the activity of a multiplicity of kinase enzymes. Our strategy provides access to a vast, untapped resource of meaningful measurements, whether readouts are interpreted irrespective of which enzymes phosphorylate which probes, or analyzed to convert global phospho-signatures into functional profiles of kinase activities. The procedure described in this Protocol Exchange chapter focuses on detailing the statistical and computational analysis steps that allow deconvoluting peptide phosphorylation profiles into kinase activity signatures. This is related to the Nature Cell Biology manuscript NCB-C36710, titled: "Mapping phospho-catalytic dependencies of therapy-resistant tumors reveals actionable vulnerabilities".

## Introduction

A key to successful therapy is the identification of critical aberrant signaling networks whose inhibition would result in system failure of diseased cells. We designed a protein enzyme activity screening system that relies on peptides as surrogate sensors of the phospho-catalytic functionality of kinases. The technology is a modular biochemical assay platform that users can adapt to their needs (e.g. probe libraries or assay conditions; see details in the related Nature Protocol Exchange chapter titled: "High-Throughput Kinase Activity Mapping (HT-KAM) system: biochemical assay"), and for which we developed a number of computational and statistical steps that can further help scientists make best use of the experimental output of such screen. Below, we describe methods to analyze phospho-catalytic signatures established from high throughput ATP-consumption measurements. We provide examples using a proof-of-concept 228-peptide library to explore the actionable phosphorylation signatures of tumor cells or patient tissues as test scenarios. Please note the following. First, in the current protocol, we focus on the method to analyze results being generated from the ATP-consumption profiles measured in presence of peptides. Second, the reason why this protocol is separate from the biochemical assay protocol it is associated with, is because we know that other readouts than ATP-consumption could be generated using the combinatorial peptide-library approach we designed, yet the readouts of the such profiles could still be interpreted/analyzed using the methods/step-by-step-process/logics described in the below analytical protocol. Third, we provide details of the 228-peptide library we had originally used as proof-of-concept so that users can then repeat the analytical/computational processes we describe in detail in the PROCEDURE section further below (e.g. users may use more, or less, or different peptides to build their sensor libraries yet they can use similar analytical/computational logics as those described below).

## Reagents

Samples and assay reagents • Described in the related Nature Protocol Exchange chapter titled: “High-Throughput Kinase Activity Mapping (HT-KAM) system: biochemical assay”. Peptide library and 384-well assay plates • The 228-peptide library included: o 151 biological peptides o 14 generic positive control peptides o 63 reference peptides that include 27 mutated (Tyrosine (Y) / Serine (S) / Threonine (T) → Glycine (G)) and 31 pre-phosphorylated (Y / S / T → pY / pS / pT) peptides, and 5 random peptide sequences o Biological peptides correspond to phosphorylatable amino acid regions of substrate protein identified from literature and curated in resources such as PhosphoAtlas27 (<http://cancer.ucsf.edu/phosphoatlas>; US20120296880). o Each generic positive control (CON+) peptide corresponds to a kinase activity reporting probe commonly used in single-peptide assays as available/advertised from literature/manufacturers, and may correspond to a commonly known ‘consensus’ amino acid sequence. o Peptide library can be built/designed using pre-defined knowledge of available from literature: • For instance, for biological peptide targets of kinases, there are ~2,600 different biological peptide sequences that are on average 51% unique to each human kinase, which users can find by referring to PhosphoAtlas27 (<http://cancer.ucsf.edu/phosphoatlas>), which also contains an additional set of >2,800 peptides that related to all known cancer-mutated phosphorylatable peptide regions of kinase substrate proteins. • As well, for generic CON+ peptides, there are ~160 different generic CON+ peptide sequences that are on average <0.1% unique to each human kinase and are currently used, available, advertised and commonly used for pharmacological screens (or potentially other applications), which users can find by referring to catalogs from SignalChem, Promega, AnaSpec, ReactionBiology, PamGen, KINOMEScan/DiscoverX, KiNativ/ActivX, PhosphoNet/Kinexus, JPT, PerkinElmer, ThermoFisher, and many more. • Each 384-well assay plate also includes: o 14 peptide-free wells (i.e. all reagents including ATP and sample but without peptide). o all other controls for such assay: ATP standard (serial dilutions), background (all reagents but without any ATP and sample and peptide), sample-free ATP-loading baseline (all reagents including ATP but without sample and peptide).

## Equipment

• Computer / computer software: XLS (versions 14.0 and 16.0), R (version 3.5.0), Prism (version 6.0e), MATLAB (version 8.5), SIGMAPLOT (version 12.5.0.38), cBioPortal, HTseq-count (version 0.10.0), DESeq2 (version DESeq2\_1.18.1). Versions of software are provided as examples.

## Procedure

PART1. Data normalization. Normalization methods to transform raw ATP-consumption measurements into interpretable peptide phosphorylation profiles are needed to be able to compare profiles of different samples. 1. For biochemical samples (e.g. recombinant kinases +/- compounds): a. The average value of ATP-consumption across the 228-peptide sensors for each experimental run can be systematically used for internal normalization of each experimental run. b. Alternatively, to further analyze or cross-validate the output of results, users can use other normalization schemes relying on specific peptide sensor subsets, for instance: i. 63 reference peptides, or ii. 16 Y/S/T-free peptides, or iii. 5 random

peptides. c. Activity per-peptide can then be calculated as the difference in ATP-consumption between individual peptide-derived values and the internal mean (whichever one listed/chosen from points a)-b) above). i. Users may also consider using alternatives comparisons, such as fold change versus to the baseline of their choice. d. Peptide-specific activity values can then be averaged across independent repeats to establish the activity signature of each sample/recombinant across all peptide sensors.

2. For cell or tissue samples: a. The average value of ATP consumption across the 228 peptides and 14 data-points from cell/tissue extract alone (i.e. established from 14 peptide-free control wells per 384-well plate) was systematically used for internal normalization of each experimental run. (This normalization is somewhat comparable to (1) how western blots are normalized to total protein amount and/or a set of 'stable' proteins, and (2) how TaqMan/RT-PCR/microarrays are normalized to the expression levels of a set of house-keeping genes and/or across all readouts.) b. Alternatively, to further analyze or cross-validate the output of results, users can use other normalization schemes relying on specific peptide sensor subsets, for instance: i. 14 peptide-free control wells (i.e. cell or tissue extract alone), or ii. 16 Y/S/T-free peptides, or iii. 63 reference peptides. c. Activity per-peptide can then be calculated as the difference in ATP-consumption between individual peptide-derived values and the internal mean (whichever one listed/chosen from points a)-b) above). i. Users may also consider using alternatives comparisons, such as fold change versus to the baseline of their choice. d. Peptide-specific activity values can then be averaged across independent repeats to establish the activity signature of each sample/recombinant across all peptide sensors. The results of these different normalization schemes for either biochemical samples, or cell extracts, or tumor tissue extracts, are then subjected to statistical and comparative analyses described below. Note that, in the case of biospecimen tissues, another useful dataset can be the 'un-normalized' (i.e. 'raw') ATP-consumption profiles measured across wells/peptides, and can be directly used to compare individual samples. Note that, as a mean to control for the quality of the output of the assay, but also in order to identify which peptide sensors out of the compendium of peptides composing a library can best report on any recombinant kinase of interest, users can systematically calculate Z-factor profiles. Indeed, comparing the dynamic range to data variation of 'positive' versus 'negative' controls (i.e. Z-factor or Z') is a standard method in the field to evaluate the performance of an enzymatic assay. Comparing Z' outputs depending on peptides can be considered as a measure of fitness of a probe in a kinase assay. Z' is calculated as  $Z' = 1 - \frac{3 * (\text{StDev Pos} + \text{StDev Neg})}{|\text{Ave Pos} - \text{Ave Neg}|}$ , where Neg are ATP consumption values measured in absence of any peptide (e.g. measured in the 14 peptide-free wells), and where Pos are ATP consumption values measured in presence of a peptide probe. Such peptide probe is usually a commonly used generic CON+ peptide, but can also correspond to other peptide probes included in our assay, such as either best activity-reporting peptide among any other/non-advertised generic CON+ for a tested kinase, or best activity-reporting peptide among biological peptides.

PART2. Compare peptide-phosphorylation signatures between samples. Once data have been normalized (see above), results can be (i) either interpreted irrespective of which enzymes phosphorylate which probes (as if peptide-phosphorylation profiles were agnostic readouts of overall phospho-catalytic activities; current PART2), or (ii) analyzed to convert global phospho-signatures into functional profiles of kinase activities (considering that peptide sensors are related to a kinase enzyme that phosphorylate the residue/region of a given substrate protein in

biological settings; i.e. PART3 further below). Users may find advantages using either of these analyses depending on their questions/topics. In the PART2 and PART3 below we provide a non-exhaustive list of statistical tools to interpret profiles. (These are examples of methods that may guide users in their analysis of data, but computational procedures and statistical processes are necessarily user-dependent and should be further tailored to users' samples and hypotheses.) Note that all statistical analyses described below can be ran with any of the normalized datasets (four different options for biochemical samples, and four different options for biological samples, respectively detailed in PART1 in items 1.a-b) and 2.a-b) above).

1. Analysis of peptide-phosphorylation profiles measured with recombinant kinases (+/- inhibitors):
  - a. Apply unsupervised or semi-supervised hierarchical clustering of kinases' phospho-catalytic activity signatures monitored across all-228 peptides, in order to group phosphorylation activity signatures based on their similarities/differences per peptide probe and per kinase (e.g. Euclidean distance or (Absolute) Correlation (centered or uncentered), and Ward linkage or complete or average linkage).
  - b. Apply Pearson- or Spearman- correlation to highlight the functional relationships of kinase enzymes.
  - c. Examine how including a multiplicity of peptide sensors impacts the sensitivity and specificity of the assay for predicting the identity of an individual kinase, by computing Area Under the Curve (AUC) from repeated iteration of random peptide sampling:
    - i. To do so, use random peptide sampling of combinations of up to 50 (or 100 or more) peptides out of all (228) peptides, using Diagonal Linear Discriminant Analysis (DLDA) class predictors, where AUC values reflect the performance of any given assay for predicting the identity of a kinase family by comparing all its kinases' 228-peptide phospho-signatures, versus the 228-peptide phospho-signatures of all other tested kinases, when relying on one or multiple peptide sensors (random sampling per combination of any set of n-peptides can be ran for 1,000 iterations but this may depend on users' preference and peptide library size).
    - ii. For each kinase or kinase family, Receiver Operating Characteristic (ROC) curves and AUC values can then be computed from kinases' phospho-catalytic activity profiles measured with particular sets of peptides, for instance a kinase's set of biological peptides, or a kinase's set of generic CON+ peptides, or all-random peptides (for negative control for instance), or any subset of interest to the user, in order to provide a comparison of the sensitivity/specificity.
  - d. Identify combinatorial peptide sets that best differentiate a kinase from others by comparing all phospho-catalytic profiles of kinases using a dual significance threshold ( $p < 0.05$  for False Discovery Rate (FDR/BH)-corrected t-test and Wilcoxon rank sum test):
    - i. All 228-peptide activity profiles from a given kinase (or kinase family) are compared to the 228-peptide profiles of all other tested recombinant kinases, so that all peptides associated with differential activity values (up or down) passing a significant  $p < 0.05$  threshold for both FDR-corrected t-test and Wilcoxon rank sum test, can then be selected.
      1. It is best to consider comparing the 228-peptide activity signatures of an individual kinase (or all kinases belonging to a given family) for which users have generated enough independent experimental repeats (e.g.  $n \geq 6$ ), although this is ultimately a user-defined threshold that depends on how robust differences in peptide phosphorylations need to be.
      2. Note that the selected, most differential peptides can be associated with either low, or high, or 'average' phospho-catalytic activities specific to a kinase or kinase family as long as they significantly contrast with activities observed across all other kinases. Following this principle, a peptide can be found as part of the differential signature of multiple kinase families at once owing activity levels and significances

that are specific to the differential signature of its given kinase family versus all other kinases. 3. Note that the other outcome of this process is that, the activities from the selected, most significantly differential peptides specific to a kinase family follow a trend that may vary from one individual kinase to another within that family (and/or between experimental read outs). Some individual kinases may even cluster away from the majority of the other kinase family members. Such outcome underlines the functional precision of combinatorial measurements provided by the HT-KAM strategy toward the systematic identification of specific enzymatic activity features unique to most kinases within a kinase family, yet remaining capable of functionally distinguishing some sub-family members. 4. To confirm the validity of the differential peptide signature, the analysis can be complemented using Monte Carlo cross validation to further estimate the performance accuracy of the predictive calculation outlined above. ii. Once peptide sensors have been selected from step i) above, sub-activity heatmaps can then be generated to display the distinct peptide subsets identified as functional predictors of the differential activity signature of the kinase (or kinase family) of interest. Users may elect to further apply unsupervised clustering or other grouping/ranking statistical tools to highlight the relationships between the selected set of peptides most differentially associated with low/high activity of a kinase or a kinase family, and their ability to distinguish all other kinases tested by the user. iii. Peptides composing the differential signatures can be further classified as 'predicted' or not ('other'), where 'predicted' defines a peptide previously identified in literatures as a target of a given kinase. It is anticipated that many of the peptides would fall under the class 'other' since many peptides included in the differential signature of a kinase/kinase family would necessarily match peptides that in fact match other kinases/kinase families and thus allow to distinguish the kinase/kinase family of interest versus all others. iv. Peptides composing the differential signatures can be further used to generate ROC curves and AUC values calculated from differential peptide set per kinase family signature. (Please refer to section 1.c) above, and run the same method but using the specific set of defined 'differential peptide subset' to compare the peptide-phosphorylation profile of a kinase/kinase family of interest versus all other kinases.) e. As an alternative method to find peptide sensors that are most significantly associated with high/low activity per kinase, users can test each kinase independently of all others, and use initially-separate-but-eventually-converging computational methods to compare levels of ATP consumption per individual peptide to the pool of (63) reference peptides. i. In the first approach, the average 228-activity data points from all experimental repeats is used in a Kalmagorov-Smirnov (KS) test comparing each 165-non-reference peptides (i.e. 151 biological peptides, and 14 positive control peptides) to the 63-reference peptides (p values with or without BH correction controlling for false-discovery rate). In parallel, the mean and standard deviation (SD) of the 63-reference peptides is computed to then identify which peptides among the 165-non-reference peptides display activity signals >2 fold Standard Deviation from the mean (>highest 2.5% of reference). ii. In the second approach, all experimental replicates (instead of averaging them as in the first method) are used in either a linear additive model ( $\lambda$ ) with BH corrected p-values from each 165-non-reference peptide versus 63-reference peptide (BH.p. $\lambda$ <0.05 threshold), or an ANOVA model with BH corrected replicate error. iii. The overlapping results of the computational processes and statistical cut-offs resulting from these two separate computational methods identify the most significantly and stringently selected high –and low– activities per peptide per kinase (i.e. robust

sensors of kinases' catalytic activities). f. As another method to find (and/or validate) peptide sensors that are most significantly associated with high activity per kinase, users can measure the ATP-consumption profiles in presence of increasing concentration of a kinase-specific inhibitor. The underlying postulate is that, when the activity of a kinase is measured in presence of an inhibitor that should inhibit its activity, then any peptide associated with a significant decrease in activity of this kinase may be considered as a suitable sensor to detect the activity of this kinase. Specifically, for any given kinase, the output of such approach can be plotted and interpreted as follow: i. Calculate the Pearson correlation coefficient between drug concentration and ATP consumption for each peptide, as a mean to evaluate the levels of inhibition (to plot on the y-axis). ii. Calculate the activity level per peptide in the untreated control setting (to plot on the x-axis). iii. Calculate the correlation between data points ( $R^2$  (Fisher (inhibition), activity) along with significance of the association (p-val)), and identify which peptides report on (i) higher kinase activity levels (i.e. dots located toward the right-end of the x-axis, indicating highest ATP consumption), and (ii) also exhibit greater activity inhibition in presence of increasing concentrations of the kinase-specific inhibitor (i.e. dots located toward the bottom-end of the y-axis, indicating strongest negative correlation and thus strongest inhibition). iv. These results also can help assess the utility / reliability / quality of biological peptide targets of kinases as sensors. v. This method and its logics can also be used for many other purposes such as identifying how specific a kinase-inhibitor is, or identify additional targets of a drug. g. Compare activity profiles of Tyrosine Kinases or Serine/Threonine Kinases measured in presence of any of their predicted Y- or S/T- containing biological peptides, versus any Y- or S/T- free biological or reference peptides. This can help provide a control readout of the specific activity of kinases for their reporting sensor probes.

2. Analysis of peptide-phosphorylation profiles measured in presence of protein extracts from cell lines (+/- inhibitors) or tissues (such as tumors): a. Please note that statistical methods described for biochemical compositions (see section 1.a), 1.b), 1.c), 1.d), 1.e), 1.f), and 1.g) above) can also be used to explore the peptide phosphorylation profiles of biological samples (e.g. define the sensitivity/specificity, or the differentiability, of peptide sensor-derived signatures that best predict the identity of an individual biological sample based on the levels of peptide phosphorylation measured across/between samples). i. Below we detail some additional/complementary methods – which can also/conversely/obviously be applied to study biochemical samples if relevant to users' questions. b. Apply unsupervised or semi-supervised hierarchical clustering using peptide-phosphorylation activity profiles monitored across all-228 peptides, in order to group cell or tissue extracts based on the similarities/differences of their respective phosphorylation activity profiles (e.g. HT-KAM-generated profiles can be clustered using Euclidean distance and ward linkage.) i. Note that this simple step can already have a direct diagnostic value when testing tissue samples from patients, for instance revealing sub-signatures that match patients outcome / survival, or recurrence, or therapeutic resistance / response. c. Apply principal component analysis (PCA) to investigate the potential association between a variable of interest and the principal components (PCs) defining the phosphorylation signatures of different biological samples. i. Use linear regression, overall fit of univariate model  $PC(i)$  for variable  $(j)$  to plot graph displaying the relationship between PCs of peptide-phosphorylation activity profiles and a biological or technical variable, along with the related significance. ii. For technical variables: This method can be an effective

way to assess whether replicate runs from the same sample are significantly similar or not, or whether days at which assays are significantly similar or not, which can be used to assess or show the level of performance and reproducibility of the HT-KAM screening system (i.e. experimental procedure, instrumentation, data analysis, among many technical variables users may want to question). From a technical standpoint, this analysis can also be used to assess the various outputs that can be obtained using different normalization methods, or comparison methods.

iii. For biological variables: This method can be an effective way to assess whether biological or clinical characteristics such as drug-resistance or survival outcome (or any phenotypic or molecular or medical characteristic is of interest to the user) are associated or not with PC signatures of peptide phosphorylation signature of samples (cells or tissues, including tumors).

d. To find peptides that qualify as best predictors of a biological variable of interest, peptides can be selected based on whether the phosphorylation activities they report on, concurrently pass both FDR-adjusted two-sided Student t-test and Wilcoxon rank sum test  $p < 0.05$ .

i. Such rigorous dual significance threshold selection can identify a subset of peptides as the most significantly differentially phosphorylated peptides associated with the biological variable of interest (e.g. drug-resistance or survival outcome).

1. the phospho-fingerprints of tumors were highly robust signatures that strongly associated with outcome

ii. This method can also allow users to identify which kind of peptides (e.g. biological peptide sequences) they included in their assay, are most associated with the biological variable of interest, and in which proportion such peptides display higher/lower phosphorylation activity profiles. (This can be useful to then define which kinases are more or less active in their samples of interest, as explained in the PART3 below, and based on the connectivity between biological or CON+ peptides and their respective kinase enzymes.)

iii. Other thresholds may be used; e.g. users may derive peptide-phosphorylation signatures by selecting and displaying peptides that match the top-10% or top-25% most-differential (up and down) activities.

e. If users want to assess differences in levels of peptide phosphorylation between samples representing different conditions of interest (e.g. treated or not with an inhibitor, or treated or not with a combination of inhibitors, or drug-sensitive vs. drug-resistant tumors), users can simply calculate the average differences (or fold) in phosphorylation activities per peptide (and for each of the 228 peptides) between treated samples versus control untreated counterparts.

PART3. Calculate and compare kinase activity signatures between biological samples. Since a biological peptide or a generic CON+ peptide are –by definition– related to a kinase enzyme that phosphorylates them (see the REAGENTS section describing how peptide libraries are design and where peptide sequences come from), then peptide-phosphorylation profiles can be deconvoluted and transformed into individual kinase's phosphorylation activities. This is the logical premise for using biological peptides of kinases as specific discriminators of kinases respective identity and activity, and for using biological peptide libraries as combinatorial sensors of enzymes' activity to convert complex peptide-phosphorylation profiles into enzyme activity signatures (which can thus be simultaneously and directly measure at once in biological samples). As such, complex peptide-phosphorylation profiles can be systematically analyzed using computational methods and statistical tools to: (1) establish the phospho-catalytic activity of many kinases at once, (2) derive the global kinase activity signatures of each biological sample, (3) analyze and compare kinase activity signatures between biological samples. Below we provide a simple way to estimate the activity levels of kinases derived from phosphorylation

activity levels measured with multiple ( $n \geq 4$ ) biological peptides related to each kinase. 1. Average peptide-phosphorylation levels for all biological peptides corresponding to a given kinase. a. Note that we decided to only derive kinase activity levels for kinases with  $n \geq 4$  different biological peptides. i. Users may however decide to extend kinase signatures by calculating kinase activities for kinases with  $\geq 3$  (or less) different biological peptides, or conversely, narrow down signatures to kinases with  $\geq 5$  (or more) different biological peptides. ii. We chose  $n \geq 4$  for the following reasons: 1/ this reduces chances for peptides to be shared between kinases (such effect can be estimated when using CON+ peptides which are commonly shared between many kinases); 2/ this allows to –for example– rationally elude cross-reaction effects from parallel feedback loops from drug treatments, or to provide stronger statistical analysis to compare kinase activity profiles (within a sample or between samples). 2. Users can then apply the statistical tests/computational procedures described in PART2 above to compare and identify kinase enzymes whose catalytically activities are more or less active in cell/tissue samples. Below are some examples as well as other complementary analyses: a. Apply unsupervised hierarchical clustering, or semi-supervised hierarchical clustering, or principal component analysis, using kinase activity signatures of samples to investigate potential association between kinase(s) and variables/samples of interest. b. Once kinases of interest have been identified, the individual heatmaps of biological peptide phosphorylation activity signatures per kinase can be displayed along with the significance per peptide. c. Additional analysis (or validation) can be calculated as follow: i. Apply an enrichment analysis EASE – Fisher one-sided test to select the most differentially phosphorylated peptides associated with a sample (versus other samples) out of all (228) peptide sensors, to then identify which kinases' biological peptides are most represented within that sub-peptide-phosphorylation-profile. ii. Apply FDR-corrected one-sided or two-sided Student t-test using all (unselected) biological peptides per kinase, and comparing all experimental runs between different sample groups, to then identify which kinases' biological peptide phosphorylation sub-signatures are most systematically significantly upregulated/downregulated. 3. These analyses can also be applied to study profiles derived from generic CON+ peptides, and compare the output of generic CON+ peptides versus biological peptides.

## Timing

The amount of time entirely depends on the kind and number of questions and methods to be used, as well as on the computational/statistical background/skills of users/research teams. Note that we are actively working on building fully automated programs to efficiently generate data.

## Troubleshooting

- Ensure that your kinase-peptide list is as accurate as possible, and that not mistake while deconvoluting peptide phosphorylation profiles into kinase activity signatures cause peptide-kinase connectivity errors that may skew results.
- Users should choose the statistical method/computational process that best matches their specific question (i.e. none of the above methods provides the same output/interpretation of results).
- With regard to PART3 above, additional methods can be used to refine results by correcting

for the effects of peptides shared between kinases \ (we are currently developing the processes to automatize the inclusion of these methods and assess their usability).

## References

1. Olow, A., et al. An Atlas of the Human Kinome Reveals the Mutational Landscape Underlying Dysregulated Phosphorylation Cascades in Cancer. *Cancer Res* 76, 1733-1745 \ (2016).