Phosphoproteomics data classify hematological cancer cell lines according to tumor type and sensitivity to kinase inhibitors

Pedro Casado1, María P Alcolea1, Francesco Iorio2,3, Juan-Carlos Rodríguez-Prados1, Bart Vanhaesebroeck4, Julio Saez-Rodriguez2, Simon Joel5 and Pedro R Cutillas1,6*

Abstract

Background: Tumor classification based on their predicted responses to kinase inhibitors is a major goal for advancing targeted personalized therapies. Here, we used a phosphoproteomic approach to investigate biological heterogeneity across hematological cancer cell lines including acute myeloid leukemia, lymphoma, and multiple myeloma.

Results: Mass spectrometry was used to quantify 2,000 phosphorylation sites across three acute myeloid leukemia, three lymphoma, and three multiple myeloma cell lines in six biological replicates. The intensities of the phosphorylation sites grouped these cancer cell lines according to their tumor type. In addition, a phosphoproteomic analysis of seven acute myeloid leukemia cell lines revealed a battery of phosphorylation sites whose combined intensities correlated with the growth-inhibitory responses to three kinase inhibitors with remarkable correlation coefficients and fold changes (> 100 between the most resistant and sensitive cells). Modeling based on regression analysis indicated that a subset of phosphorylation sites could be used to predict response to the tested drugs. Quantitative analysis of phosphorylation motifs indicated that resistant and sensitive cells differed in their patterns of kinase activities, but, interestingly, phosphorylations correlating with responses were not on members of the pathway being targeted; instead, these mainly were on parallel kinase pathways.

Conclusion: This study reveals that the information on kinase activation encoded in phosphoproteomics data correlates remarkably well with the phenotypic responses of cancer cells to compounds that target kinase signaling and could be useful for the identification of novel markers of resistance or sensitivity to drugs that target the signaling network.

Background

Hematologic malignancies are a group of neoplastic diseases that originate from the transformation of bone marrow-derived cells. This group, which includes leukemias, lymphomas, and myelomas, is extraordinarily heterogeneous, which reflects the complexity of normal hematopoiesis and the immune system [1]. Although gene expression signatures can be used to classify malignancies into subgroups [2-4], a system-level understanding of the biochemical pathways (both signaling and metabolic) responsible for tumor phenotypes requires knowledge of signaling pathway activity, information that cannot be provided by measuring mRNA or protein expression alone [5,6], as enzyme expression does not necessarily correlate with pathway activity [7].

Essentially all cancers are driven by deregulation of protein kinase cascades downstream of growth factor, antigen, and G protein-coupled receptors [8]. Consequently, several kinase inhibitors that block cell transduction pathways overactive in cancer are already in the clinic while others are undergoing pre-clinical or clinical development. However, although clinical impact is observed in some patients, many patients do not respond to these therapies or subsequently develop resistance [9,10]. The use of predictive biomarkers, or ‘companion diagnostics’, is therefore important in individualizing such targeted agents [11]. While the activity of the target
kinase can in some instances predict response [12], this is not always the case, as the activity of parallel pathways in the network can contribute to resistance [13,14]. It could therefore be envisaged that the analysis of kinase signaling without a preconception of the pathways that may be active could be advantageous in predicting responses to kinase inhibitors.

Phosphorylation is a posttranslational modification regulated by the activity of kinases and phosphatases. By definition, each phosphorylation site is the result of a kinase/phosphatase reaction pair. Changes in phosphorylation status can alter many aspects of protein biology, including their localization, protein-protein interactions, stability, and enzymatic activity [15]. Although the information coded by phosphorylation patterns has not been completely deciphered, many phosphorylation sites can be associated with the activity of a specific protein kinase and thereby classified into signaling pathways [16-18]. Thus, global analysis of protein phosphorylation using quantitative techniques may in principle be translated into knowledge of the activation status of signaling pathways. This information, in turn, could be used to rationalize how the wiring of the kinase network contributes to the phenotypic characteristics of different tumors, such as aggressiveness, metastatic potential, and sensitivity to therapy.

The application of new proteomic techniques for phosphopeptide quantification is contributing to an improved understanding of cancer cell biology [19-23]. Several techniques for quantitative proteomics have been developed; these can be divided into those that require labeling of proteins with stable isotopes (for example, SILAC and iTRAQ) and those that do not require labeling [24,25]. Approaches based on labeling techniques usually detect a larger number of phosphopeptides than those based on label-free approaches because labeling techniques are compatible with extensive fractionation prior to mass spectrometry analysis. However, because of the time-consuming nature of such analyses, studies based on labeling techniques normally compare a small number of samples with no (or very few) biological replicates, a feature that limit the statistical significance of the results. Therefore there is a trade-off between the number of peptide/proteins identified and samples that can be compared in a study. Label-free approaches are preferred when the aim is to compare large sample numbers and replicates [26,27] even though the penetrance of the approach may not be as large as when using techniques that allow extensive fractionation before mass spectrometry analysis.

In the current study, label-free mass spectrometry (MS) was first used to analyze the phosphoproteomes of nine different hematological cancer cell lines. Unsupervised analysis of the data based on principal component analysis (PCA) and hierarchical clustering classified these cell lines according to their pathological origin, namely acute myeloid leukemia (AML), lymphoma, or multiple myeloma. Through a lasso linear regression analysis we assessed the potential of this data in predicting the level of sensitivity of seven AML to three kinase inhibitors: PI-103 (a PI3K/mTOR inhibitor), MEK-i, and JAK-i. Finally we identified phosphopeptides whose intensities across the cell lines correlated with the sensitivity to three inhibitors. Our results revealed a battery of phosphorylation sites whose intensities strongly correlated with responses of our AML panel to the compounds (with \( R > 0.9 \) and \( > 100\)-fold difference between most sensitive and most resistant cell line). These data therefore indicate that MS-based quantitative phosphoproteomics has the potential to classify cell lines into distinct subgroups according to their pathological origin and to their sensitivity to drugs that target kinase signaling. Phosphorylation sites that correlated with resistance were enriched in basic and proline-directed motifs, whereas those that correlated with sensitivity were mainly acidic or hydrophobic containing, thus suggesting that the relative activities of basophilic, proline directed, and acidophilic kinases may determine responses to the inhibitors. Therefore, the results of our study indicate that unbiased profiling of phosphorylation has the potential to stratify AML cells based on their responses to signaling inhibitors because these responses may be dependent on the combination of pathways (both target and parallel) active in cells, rather than on the activity of the target kinase/pathway only.

**Results**

**Overview of protein phosphorylation in hematological cancer cells**

A group of nine different hematological cell lines, consisting of three acute myeloid leukemia (AML), three lymphoma, and three multiple myeloma cell lines (Additional file 1, Table S1), was selected for analysis using a quantitative LC-MS/MS phosphoproteomics workflow summarized in Additional file 2, Figure S1A. This approach to quantify phosphorylation, which involves comparing peak intensities of phosphopeptides calculated as the height and areas of ions extracted from aligned chromatograms, has been independently validated by immunoblotting in our previous work that showed that this methodology can be used to quantify phosphopeptide levels with good precision and accuracy [16,28]. The technique is similar to that used in other laboratories [29-31].

Three biological replicates were analyzed on two separate occasions to give a total of six replicates per cell line, requiring 54 LC-MS/MS runs. The approach led to the identification of 2,050 phosphopeptides in 1,664
proteins (Additional file 3, Dataset 1). Given that phosphopeptides can contain more than one phosphorylation site, in total we identified 2,434 unique phosphorylation sites. Of these, the precise position of the modification was ambiguous for 738 sites. The remaining 1,696 sites (70% of total) were classified according to the amino acid bearing the site of phosphorylation. Additional file 2, Figure S1B shows that 1,254 (74%), 344 (20%), and 98 (6%) of phosphorylation sites were on serine, threonine, and tyrosine, respectively. These results are comparable with the distribution of phosphorylation sites in HeLa carcinoma cells and in platelets [32,33].

In order to further assess the nature of the phosphoproteomes identified in this study, we compared the assigned gene ontologies (GO) of the phosphoproteins with those proteins present in the whole human proteome (obtained from the SwissProt database, Additional file 4, Figure S2). Each protein was classified according to the three different domains included in the gene ontology project [34], namely (1) cellular component which describes the subcellular localization of the identified protein (or its extracellular environment); (2) biological process which indicates the biological function to which the gene product contributes; and (3) molecular function which describes the elemental activities of a gene product at a molecular level [34]. The distributions of these domains in our phosphoprotein database were compared to those in the SwissProt database (which lists all the known gene products and can thus be considered to represent the whole human proteome). The data show that phosphoproteins were on the whole not biased towards any GO, in that the subcellular distribution of phosphoproteins identified were similar to those reported in the SwissProt database (Additional file 4, Figure S2A). The minor differences in the proportion of membrane proteins could be explained by the difficulty in solubilizing membrane proteins with buffers compatible with the rest of the workflow. We also observed that phosphoproteins located in the cytosol/cytoplasm and those with roles in translation, cell cycle regulation, and proliferation were well represented relative to the whole proteome (Additional file 4, Figure S2A and S2B). Other well represented phosphoproteins include those with protein kinase activity and other enzymes (Additional file 4, Figure S2C).

Phosphoproteomics classified hematological cancer cells according to pathology

We next asked whether quantitative phosphoproteomics could be used to classify hematological cell lines according to their pathological origin. Normalized peptide intensities were used to calculate the overall fold difference (relative to the mean value across all samples) for each phosphopeptide identified and quantified in our study. These data were then subjected to principal component analysis (PCA). When analyzing the nine cell lines together, inspection of the PCA outputs (Figure 1A) revealed that principal component 1 (PC1) produced two clearly separated groups, one containing AML cell lines and another containing lymphoma and multiple myeloma cell lines, while principal component 2 (PC2) separated cells derived from lymphoma from those that originated from multiple myeloma (Figure 1A). We also analyzed the phosphorylation data of each cell line in three separate PCA plots corresponding to the three different cancer types. Only data on the cell lines of each different malignancy were included in each particular PCA plot. These data, summarized in Figure 1B, revealed that cell lines within a given disease could also be separated based on their phosphoprotein content, indicating global differences in phosphorylation between cells of the same pathology with the exception of RL and DoHH2 lymphoma cell lines which could not be separated by PCA (Figure 1B).

The ability of phosphoproteomics to classify hematological cell lines was also assessed by unsupervised hierarchical clustering. This analysis also classified the samples into three main groups corresponding to AML, lymphoma, and multiple myeloma, and clustered replicates for each cell line within those groups. Interestingly lymphoma and multiple myeloma cells were clustered together, consistent with AML cells being the most differentiated group (Figure 1C). This is consistent with the data obtained by PCA (Figure 1A). Taken together, the data in Figure 1 indicate that although there are differences in global phosphoprotein abundance within cell lines of the same pathology, the differences in the phosphoproteomes of hematological cancer cells are greater across pathology groups than within a given disease.

Stringent statistical analysis (as indicated in materials and methods) identified 609 phosphopeptides differentially regulated in the AML cell lines, of which 544 showed increased phosphorylation and 65 decreased phosphorylation, 46 were different in lymphoma cells (30 increased and 16 decreased) and 53 in multiple myeloma (27 increased and 26 decreased) (Figure 1D). Representative examples of phosphopeptides increased or decreased in AML, lymphoma, and multiple myeloma cells are shown in Additional file 5, Figure S3. Taken together, these data show that label-free quantitative phosphoproteomic data can be used to reproducibly classify hematological cancer cell lines according to their pathological origin.

Phosphorylation patterns in hematological cancer cells associated with sensitivity/resistance to kinase inhibitors

In order to assess whether phosphoproteomics could also be used to classify cells according to their responses to kinase inhibitor treatment, we compared phosphorylation patterns with the sensitivity of a panel of seven
AML cell lines (Additional file 1, Table S2) to a PI3K/mTOR inhibitor (PI-103), a MEK-inhibitor (MEK-i), and a JAK-inhibitor (JAK-i). Cell lines were ranked according to the percentage reduction in MTS signals when incubated with 1 μM PI-103, 1 μM JAK-i, or 10 μM MEK-i (Figure 2). MTS measures mitochondrial redox activity, a parameter that often correlates with cell viability. In our assay a reduction in MTS signal as a function of compound treatment may be taken to indicate a reduction in the number of viable cells relative to the

Figure 1 Phosphoproteomics classify hematological cell lines according to their tissue origin. Phosphopeptides identified by LC-MS/MS with Mascot and quantified with Pescal in nine hematological cell lines were analyzed by two different clustering tools. (a) Principal component analysis (PCA) of all the cell lines based on the analysis of the most intense 1,500 phosphopeptides. (b) Independent PCA plots of cells belonging to the same pathological group. (c) Unsupervised hierarchical clustering using Kendall’s tau coefficient to measure association between peptides and arrays and complete linkage clustering to calculate distances between clusters; AML samples are highlighted in red, lymphoma in blue, and multiple myeloma in green. Each replicate is shown separately and the cell line number is followed by the order in which the samples were analyzed by LC-MS/MS. (d) Summary of phosphopeptides differentially regulated in AML, lymphoma, and multiple myeloma cell lines with two-fold difference over mean expression and \( P < 0.05 \) after t-test and FDR multiple test correction.

AML 609
LPH 46
MM 53

Total 107

AML 544 65
LPH 30 16
MM 27 26

AML (b)
control as a combined action of the drugs in inhibiting proliferation and inducing cell death.

We then performed phosphoproteomics analysis on these cell lines (the data are shown in Additional file 6, Dataset 2). These experiments were performed in three independent cell cultures per cell line and further data analysis was based on phosphopeptide intensities from individual replicates rather than the averages values of the three replicates.

Assessing the potential of phosphoproteomics data in predicting drug responses

We assessed the ability of phosphoproteomics data in predicting the drug sensitivity profiles of our panel of cell lines against the three tested inhibitors through a ‘leave-one-cell-line-out’ (LOCLO) approach based on lasso regression [55], detailed in the methods section. For each of the three tested drugs, we trained seven different models by leaving out the samples corresponding to each of the seven cell lines in turn, and composing each time a test set with them. The trained models were finally used to make predictions on the test set.

Each training phase was composed by a three-fold cross validation estimation of the lasso shrinkage parameter (involving 18 of the corresponding training set only) followed by an optimization phase of the regressor coefficients (on the training set) and was repeated 20 times. Finally an average trained model was assembled.

Figure 2 Responses of hematological cell lines to kinase inhibitors. AML cells were exposed to 1 μM PI-103, 10 μM MEK-i, or 1 μM JAK-i for 72 h and viability measured by MTS and expressed as percentage to DMSO (control) treated cells. Cell lines were then ranked according to their responses to the inhibitors. Values are mean ± SD (n = 4).
across these 20 iterations and used to predict drug-responses for the corresponding test set.

A scatter plot of the predicted viability scores for each of the test set versus the observed ones is depicted in Figure 3A. To make results comparable in this plot, test set predictions were normalized together with the corresponding training set predictions and for the same reason, all the observed viability scores were normalized drug-wisely. A remarkably high and statistically significant values of Pearson correlation coefficient (R) were measured between observed/predicted viability scores drug-wisely. The inserts report average non-null coefficients and average non-null coefficient rate in the final descriptive models for PI-103 (b), JAK-i (c), and MEK-i (d) for phosphopeptides whose rate is > 50% (that is, sufficiently stably included in the optimal models).

In Figure 3B, C, and D) we reported excerpts of the final descriptive models including only regressors whose inclusion frequency is > 50% (sufficiently stably included). Taken together these results suggest that phosphoproteomics data have the potential to predict responses to the analyzed compounds in cell line models and that a relatively small core-set of phosphopeptides is strongly associated to viability scores.

Identification of subsets of phosphorylation sites that correlate with sensitivity

To complement the analysis based on LASSO, we also performed a correlation study to identify further phosphopeptides associated with responses to the three inhibitors. Figure 4A illustrates phosphopeptide intensities showing stronger associations with resistance or sensitivity to PI-103. We observed that the
correlations with responses were not always linear but instead often followed an exponential trend. This is illustrated in Figure 4B and 4C (left graphs) for phosphopeptides derived from NBEAL1 pT225 and DBNL pT295 which showed the greater correlation with resistance or sensitivity, respectively. Expressing normalized peak heights as a function of cell viability in semi-logarithmic scale produced a linear relationship between viability and phosphopeptide intensities (Figure 4B and 4C, right graphs) with R = 0.78 for NBEAL and R = -0.74 for DBNL phosphopeptides, consistent with the existence of an exponential relationship between viability and phosphopeptide intensities for these peptides.
We also investigated whether a combination of several phosphopeptides could be used to find associations between our panel of AML cells and their response to PI-103. As Figure 4D shows, the added intensities of the 10 phosphopeptides with greater correlation with the responses to PI-103 (denoted as group 1 in Figure 4D) showed a very strong correlation (R = 0.94) with the viability of cells exposed to PI-103. As with the behavior of single phosphopeptides, the correlation was more linear in semi-log than in linear scale (Figure 4D). The sum of the phosphopeptides with greater correlation with sensitivity (denoted as group 2 in Figure 4E) also showed a linear relationship with cell viability in semi-log scale (R = -0.94, Figure 4E). Remarkably, the ratio of phosphopeptide intensities that correlate with resistance (group 1) to those that correlate with sensitivity (group 2) showed a Pearson score of R = 0.99 with viability (Figure 4F). The differences in combined phosphopeptide intensities between the most resistant and most sensitive cell lines was 150-fold for the group 1:group 2 ratio (Figure 4F) compared to just 23-fold for group 1 (Figure 4D), indicating that the ratio of phosphopeptide intensities that correlate with sensitivity relative to those that correlate with resistance could be a strong classifier for the stratification of AML cells based on the likelihood that they respond to PI-103.

Phosphopeptide intensities were also correlated with the responses to MEK-i and JAK-i. Phosphopeptides showing the stronger association with resistance or sensitivity to MEK-i are shown in Figure 5A. The added intensities of phosphopeptides that correlated with resistance (group 1) or sensitivity (group 2) are shown in Figure 5B and 5C, respectively. As with the analysis of phosphopeptides correlating with responses to PI-103, the ratio of group 1 to group 2 showed a greater linear correlation (R = 0.95) and fold difference (284) between the most sensitive and resistant cells to MEK-i (Figure 5D) than when considering either group alone (Figures 5B and 5C). Similar results were obtained when the phosphoproteomes of our AML cell line panel were correlated with resistance/sensitivity to JAK-i. Figure 6A shows the phosphopeptides with best correlation with the response to this compound. While the combined intensities of the

---

**Figure 5** Correlating responses to MEK-i with protein phosphorylation data. The normalized intensities of phosphopeptides quantified across seven AML cells were correlated with the responses to the inhibition of proliferation of cells exposed to 10 μM MEK-i, as measured by cell viability. (a) Phosphopeptides showing greater correlation with resistance or sensitivity to MEK-i. (b) Correlation of the sum of the intensities showing greater correlation with resistance. (c) Correlation of the sum of the intensities showing greater correlation with sensitivity. (d) Correlation of the ratio between phosphopeptide intensities showing greater correlation to resistance relative to those showing greater correlation with sensitivity. In each case, R refers to the Pearson correlation coefficient of the log2 transformed data; fold refers to the division of the intensities of the peptide group increased in the most resistant cell line by that of the most sensitive cell line; open squares are individual data points of three biological replicates; blue diamonds are the averages of the three replicates.
phosphopeptides that correlated with resistance (group 1, Figure 6B) or sensitivity (group 2, Figure 6C) showed good correlation with viability (R = 0.89 and -0.95, respectively), the ratio of group 1 to group 2 phosphopeptide intensities showed the greater fold difference between the most resistant and sensitive cell lines to JAK-i, as measured by cell viability. (a) Phosphopeptides showing greater correlation with resistance or sensitivity to JAK-i. (b) Correlation of the sum of the intensities showing greater correlation with resistance. (c) Correlation of the sum of the intensities showing greater correlation with sensitivity. (d) Correlation of the ratio between phosphopeptide intensities showing greater correlation to resistance relative to those showing greater correlation with sensitivity. In each case, R refers to the Pearson correlation coefficient of the log2 transformed data, fold refers to the division of the intensities of the peptide group increased in the most resistant cell line by that of the most sensitive cell line; open squares are individual data points of three biological replicates; blue diamonds are the averages of the three replicates.

To investigate whether phosphopeptides associated with resistance/sensitivity in AML cells (Figures 4 to 6) would also be associated with responses in other cell types, we assessed whether these potential markers of sensitivity would also correlate in lymphoma and multiple myeloma cells shown in Figure 1. This analysis was undertook with the view that not all sites associated with responses in AML may be associated with the same parameters in lymphoma or multiple myeloma because of the marked differences in phosphorylation across these malignancies (Figure 1). First, sensitivity of the lymphoma and multiple myeloma cells shown in Figure 1 to PI103, JAK-i, and MEK-i was determined using MTS (Additional file 9, Figure S5A). These values of sensitivity were then correlated with the phosphorylation patterns of the corresponding peptides previously found associated to the inhibitors in the AML cells. Although not all markers of sensitivity found for AML were also associated with responses in lymphoma or multiple myeloma, we found that the phosphorylation signals of PML p-S506 and MARCKS p-S170 also positively correlate, with viability after treatment with PI-103 in our set of lymphoma and multiple myeloma cells (R = 0.72 and 0.59, respectively) (Additional file 9, Figure S5B). Similarly, phosphorylation signals on KRT25 p-S7 and SKIV2L p-T186 positively correlated with viability after treatment with JAK-i (R = 0.63 and 0.78, respectively) while the phosphorylation of HNRNPU p-272 negative correlated with the viability after treatment with MEK-i (R = 0.97 and 0.82) (Additional file 9, Figure S5C and S5D). Thus these data (Additional file 9, Figure S5) show that phosphorylation markers may be able to predict responses to kinase inhibitors across different diseases despite their marked differences in basal phosphorylation (Figure 1).
Pathways and ontologies associated with responses to kinase inhibitors

In order to investigate pathways represented in phosphopeptide sets associated with the responses of our AML cell line panel to kinase inhibitors, we performed a bioinformatics analysis using DAVID pathway analysis tools [35]. This analysis was based on phosphopeptides correlating with responses with R > 0.45 or R < -0.45. Phosphopeptides that correlated with the responses to MEK-i and JAK-i were present in proteins with diverse functions, including mRNA splicing, transcription and nuclear proteins. PKC pathway members were significantly enriched (P < 0.001) in the dataset of phosphorylation sites that correlated with the resistance to PI-103 (Additional file 10, Figure S6A). Ten phosphorylation sites on kinases, six of which are Ser/Thr protein kinases, were also found to correlate with resistance (Additional file 10, Figure S6B). Phosphorylation sites on proteins involved in transcription were well represented in the dataset that correlated with sensitivity to PI-103 (Additional file 10, Figure S6C, although the enrichment was not statistically significant, P = 0.2) or with resistance (Additional file 10, Figure S6D).

To assess whether phosphorylation sites on known members of PI3K, MEK, and JAK pathways would correlate with the responses of our AML cell line panel to the kinase inhibitors tested, we also specifically considered phosphorylation sites that, based on the literature, are known to be downstream of these kinases. Phosphorylation sites on 4EBP1 (gene name EIF4EBP1) or Ribosomal S6 (gene name RPS6), which are known to be downstream of PI3K/mTOR(Ref [36]), did not correlate with the responses of our cancer cell panel to the PI3K/mTOR inhibitor PI-103 (Figure 7A). In contrast, it was interesting to observe that phosphorylation sites on two phosphopeptides derived from MARCKS, which are substrates of PKCs (Ref [37]) and which thus provide a measure of PKC activities, correlated (R = 0.74 and 0.78) with the resistance of our AML panel to PI-103 (Figure 7B). These results are in line with those from PI-103 (Additional file 10, Figure S6C, although the enrichment was not statistically significant, P = 0.2) or with resistance (Additional file 10, Figure S6D).

Figure 7 Correlation of known PI3K, JAK and MEK pathway markers with responses to kinase inhibitors. (a) Correlation of phosphorylated peptides containing phosphorylation sites known to be downstream of PI3K/mTOR with sensitivity to PI-103. (b) Correlation of phosphorylated peptides on MARCKS, known to contain sites phosphorylated by PKCs, with responses to PI-103. (c) Correlation of phosphorylation sites on STAT5A, a known JAK site, and ribosomal S6 (RPS6 gene product) with sensitivity to JAK-i. (d) Correlation of several sites on stathmin (STMN1 gene product) with the sensitivity to MEK-i. Open squares: individual data points; filled diamonds: averages of individual data points.
Phosphoprotein concentrations in cells can be a consequence of the expression level of the phosphoprotein bearing the site of modification as well as of the activity of the kinases and phosphatases acting on these sites. To investigate whether the phosphorylation sites that correlated with resistance or sensitivity in this study were the result of differential kinase activities across the AML cell line panel or these were just a reflection of differences in the expression level of the phosphoprotein gene expression, we performed a further bioinformatics analysis of the data aimed at assessing common phosphorylation motifs in phosphopeptides correlating with the responses to the inhibitors. In this, phosphorylation motifs were obtained from the data using motif-X [40] and from the literature [41]. The normalized intensities of phosphorylation sites bearing these motifs were then averaged and correlations with the viability of cells to each inhibitor recalculated for each of the averaged motif intensities. This analysis was based on phosphopeptides with $R > 0.45$ or $R < -0.45$.

Figure 8 shows that phosphorylation motifs that correlated with resistance were rich in basic residues and/or had a proline at the carboxyl terminus of the phosphorylated residue. In contrast, phosphorylation motifs that correlated with sensitivity to the inhibitors were predominantly on acidic motifs for PI-103 and JAK-i (Figure 8A and 8C) or they contained a hydrophobic residue at the N terminus for JAK-i and MEK-i (Figure 8B and 8C). These data suggest that basophilic kinases (which include the AGC family protein kinases such as S6K, PKC, PKA, and PKB) and proline directed kinases (which include mTOR, CDKs, and MAP kinases) may be more active in resistant cells (Figure 8), whereas acidophilic kinases (such as casein kinases) and hydrophobic amino acid directed kinases (which include AMPK) may be more active in sensitive cells. These data are consistent with the correlation of PKC activity with resistance to PI-103 inferred from pathway and substrate phosphorylation analyses (Figure 8 and Additional file 9, Figure S5) in that the averaged intensities of 13 peptides containing the preferred motif for PKC (xRxSx) and of eight peptides containing the related xKxSx motif also correlated with the resistance to PI-103 (Figure 8A). Also consistent with the data shown in Figure 7C, suggesting a correlation between S6K activity and the resistance to JAK-i, the combined intensities of 43 phosphopeptides in the context of the xRxxSx motif also correlated with the resistance to this compound (Figure 8C). Of these, eight and nine phosphopeptides were in the context of the RxRxxSx and KxRxSx motifs, respectively, which are the preferred recognition motifs of S6K and related upstream kinases such as PKB/Akt. Overall, the data in Figure 8 indicate that the correlation of phosphopeptide intensities with resistance and sensitivity to kinase inhibitors is most likely due to differential kinase pathway activation in these cells, rather than just reflecting differences in the expression levels of the phosphoproteins bearing the sites of modification.

**Discussion**

Tumor classification informs therapeutic strategies and provides information on prognosis [42]. A classification based on the likely response to a particular therapy may be of particular importance in the use of agents that target protein and lipid kinases in signaling pathways [43]. These enzymes are drug targets deregulated in essentially all cancers [8] but the lack of robust and...
Phosphorylation motifs that correlate with the responses to kinase inhibitors. Phosphopeptides showing correlation to the responses to kinase inhibitions were grouped based on these containing specific phosphorylation motifs. The intensities of phosphopeptides in these groups were then averaged and correlated with viability of cells after exposure to PI-103, MEK-i, or JAK-i. Motifs were then ranked according to their correlation to the responses to PI-103 (a), MEK-i (b), or JAK-i (c). Motifs were color-coded based on their type: R, Pearson correlation to semi-log transformed data, n, number of phosphopeptides matched to the named motif.

Figure 8 Phosphorylation motifs that correlate with the responses to kinase inhibitors. Phosphopeptides showing correlation to the responses to kinase inhibitions were grouped based on these containing specific phosphorylation motifs. The intensities of phosphopeptides in these groups were then averaged and correlated with viability of cells after exposure to PI-103, MEK-i, or JAK-i. Motifs were then ranked according to their correlation to the responses to PI-103 (a), MEK-i (b), or JAK-i (c). Motifs were color-coded based on their type: R, Pearson correlation to semi-log transformed data, n, number of phosphopeptides matched to the named motif.

In order to investigate whether phosphoproteomics may be used to classify hematological cancers based on their phenotype, and to provide proof-of-principle of the approach, we analyzed the most abundant phosphopeptides of nine hematological cancer cell lines. Qualitative analysis of the data based on Gene Ontologies revealed that our analysis was broad and included phosphopeptides from all types of proteins, including those in cell membranes, cytosol, and subcellular compartments. While membrane proteins were slightly less represented than expected, proteins involved in cell cycle regulation, proliferation, and translation, and protein kinases were particularly well represented in our phosphoprotein dataset.

As for the quantitative analyses, our data showed that cancer cells of different phenotypes and origin had markedly different patterns of phosphorylation, thus indicating that phosphorylation could be used to classify these cells according to phenotype. PCA and unsupervised hierarchical clustering analysis classified cell lines...
according to their pathology (AML, lymphoma, and multiple myeloma), thus indicating that the phosphoproteomes of cells of the same pathology are more similar than those derived from different diseases. Furthermore, PC1 in the PCA and unsupervised hierarchical clustering clearly separated AML cells that have a myeloid lineage from the other two diseases that have a lymphoid lineage.

We evaluated the potential ability of phosphoproteomics data in predicting drug response through a robust cross-validation framework based on Lasso regression analysis, showing that the phosphorylation levels of a relatively small subset of phosphopeptides was informative with regard to treatment outcomes (Figure 3). In addition, using a complementary approach based on a correlation study, we identified phosphopeptide sets whose combined intensities correlated with the responses to the three compounds with remarkable correlation coefficients (R = 0.99, 0.95, and 0.94 for the responses to PI-103, MEK-i, and JAK-i, respectively). In addition to observing strong correlations with cell viability, we also found that the magnitude of changes between the phosphopeptide signatures of the most resistant and most sensitive cells were greater, or close, to 100-fold in each case (Figures 4F, 5D, and 6D). These large differences in phosphopeptide intensities are noteworthy and important as ideal predictors of response need to show robust differences and steep slopes in the regression function in addition to having a good correlation. Therefore, our proof-of-concept study indicates that phosphoproteomics may represent a general tool to identify biomarkers of response to kinase inhibitors and to classify cancers based on their likelihood that these respond to targeted therapies. Although some peptides that predicted responses to kinase inhibitors in AML were also associated with responses to the same inhibitors in lymphoma and multiple myeloma cells, there were subsets of phosphopeptides that predicted responses in AML cells only. This could be attributed to the marked differences in the patterns of basal phosphorylation between lymphoma, multiple myeloma, and AML cells (Figure 1) which may be a reflection of differences in the activity of kinase signaling across these diseases.

The correlation between PI3K pathway activity markers and sensitivity to PI3K compounds has been found to be weak in recent studies [49,50]. For example, no correlation between activating mutations on PIK3CA (the gene coding for the PI3K catalytic alpha isoform) with responses to PI3K inhibitors were found in cell lines [49], while a recently reported clinical trial found that just 30% of breast and gynecologic patients harboring PIK3CA mutations responded to PI3K therapies [50]. Dan et al. found a weak correlation between responses to PI3K inhibitors and phosphorylation of Akt (R < 0.45 and fold < 10 [49]), which is a marker of PI3K activity. Although the primary aim of this work was not to identify mechanisms of resistance, it was interesting to observe that, consistent with the data in Dan et al. [49] and Alcolea et al. [28], the levels of PI3K pathway activity markers did not correlate with the responses of cancer cells to the PI3K inhibitor (Figure 7A). Similarly, phosphorylation of STAT5A, a marker of JAK activity, did not correlate with the responses to JAK-i (Figure 7C). Interestingly, the phosphorylation sites that correlated with resistance were on pathways parallel to those being targeted; that is, on sites phosphorylated by PKC for the PI3K inhibitor (Figure 7B), thus raising the hypothesis that pathways parallel to those being targeted were contributing to resistance. The finding that in PKC and PI3K inhibitors had additive effects in decreasing the viability resistant AML cells (Additional file 10, Figure S6) is consistent with this possibility.

Conclusion

Our data therefore suggest that the activity of pathways parallel to those being targeted by the specific inhibitors might contribute to intrinsic resistance to kinase inhibitors, as observed in cells that acquired resistance as a result to chronic exposure to compounds that target signaling nodes [13,14,28,51,52].

Our analysis identified phosphorylation motifs that on the whole correlated with resistance or sensitivity to kinase inhibitors (Figure 8) indicating that different kinases may be activated in resistant and sensitive cells. Thus, it may be proposed that, in this study, unbiased profiling of phosphorylation stratified AML cells based on their responses to signaling inhibitors (Figures 3, 4, 5, 6, 7, 8) because these responses may be dependent on a combination of active pathways (both target and parallel) in cells, rather than on the activity of the target kinase/pathway only. Therefore, the ability to quantitatively analyze protein phosphorylation without a preconception of the pathways that may be implicated in conferring sensitivity may represent an important advance for therapies that target kinase signaling.

Materials and methods

Cell culture

Cell lines were obtained as indicated in Additional file 1, Table S1 and Additional file 1, Table S2. All cell lines were grown in RPMI-1640 media supplemented with 10% FBS and 100 units/mL penicillin/streptomycin at 37°C in a humidified atmosphere of 5% CO2. For experiments, 10 x 10^6 cells at 0.5 x 10^6 cells/mL were seeded 24 h before harvesting for each replicate.

Viability assay

Cell lines were seeded in 96-well plates and treated 24 h later with 1 μM PI-103 (Calbiochem cat # 528100),
10 µM MEK inhibitor I (referred as MEK-i hereafter, Calbiochem cat # 444937), or 1 µM JAK inhibitor (referred as JAK-i hereafter, Calbiochem cat # 420099) for a further 72 h. Cell viability was assessed by MTS assay (CellTiter 96® AQueousOne Solution Cell Proliferation assay, Promega Corporation, Madison, WI, USA).

Cell lysis, digestion, and solid-phase extraction
Cells were harvested by centrifugation at 300 × g for 5 min and washed twice with ice cold Dulbecco’s Phosphate Buffered Saline (DPBS) supplemented with 1 mM Na3VO4 and 1 mM NaF. Cell pellets were lyzed with 1 mL denaturing buffer (8 M urea in 20 mM HEPES pH 8.0) supplemented with phosphatase inhibitors (1 mM Na3VO4, 1 mM NaF, 1 mM β-glycerol phosphate, 1.25 mM sodium pyrophosphate) and further homogenized by sonication (three pulses of 15 s). Cell debris was removed by centrifugation (20,000 g for 5 min at 5°C), protein levels in the supernatant were quantified by Bradford analysis and an aliquot containing 500 µg of protein was diluted to a final volume of 1 mL in denaturing buffer. Cysteines were reduced and alkylated by sequential incubation with 4.1 mM dithiothreitol and 5 mM MEK inhibitor I (referred as JAK-i hereafter, Calbiochem cat # 420099), or 1 µM JAK inhibitor (referred as JAK-i hereafter, Calbiochem cat # 420099) for a further 72 h. Cell viability was assessed by MTS assay (CellTiter 96® AQueousOne Solution Cell Proliferation assay, Promega Corporation, Madison, WI, USA).

Immobilized metal ion affinity chromatography (IMAC)
Phosphopeptides were enriched by IMAC using a modified protocol previously described [24]. In brief, Sep-Pak eluents were incubated for 1 h at room temperature with 300 µL of equilibrated Fe(III) coated sepharose beads used as 50% slurry in 50:50:0.1 acetonitrile:water:TFA. For equilibration, beads were washed twice with 200 mM EDTA, three times with 50:50:0.1 acetonitrile:water:TFA, incubated twice for 5 min at RT with 100 mM FeCl3 (same volume as beads) and washed six times with 50% ACN/0.1% TFA. Unbound peptides were discarded and beads were sequentially washed with 300 µL of 50% acetonitrile/0.1% TFA and 300 µL of 50% acetonitrile/1% TFA. Phosphopeptides were recovered by incubating the beads twice with 300 µL 50% acetonitrile/1.5% ammonia water pH 11.0 for 1 min, after which the recovered solutions were pooled, acidified (by addition of 10% formic acid), dried in a speedvac, and stored at -80°C until analysis.

Identification and quantitation of phosphopeptides
For phosphopeptide identification, mascot Distiller 2.3.2 was used to smoothen and centroid the MS/MS data. The processed files were searched against the human sequence library in the SwissProt database (version 2010_03 containing 23,000 entries [53]) using the Mascot search engine [54]. Searches were automated with Mascot Daemon (v2.2.2; Matrix Science, London, UK). The search parameters included the following parameters: trypsin as digestion enzyme with two missed cleavages allowed, carbamidomethyl (C) as fixed modification, and Pyro-glu (N-term), Oxidation (M), and Phospho (STY) as variable modifications. A mass tolerance of ± 7 ppm for the precursor ion and ± 800 mzm for fragment ions was allowed. Hits were considered significant when they had an Expectation value of < 0.05 (as returned by Mascot). False discovery rates were approximately 2% as determined by decoy database searches. Results from Mascot
searches were deposited in to the ProteomeXchange Consortium [55] via the PRIDE partner repository [56] with the dataset identifier PXD000217 and DOI 10.6019/PXD000217. Sites of modification were reported when the Mascot delta score [57] was > 10; otherwise site assignment was deemed ambiguous. Phosphopeptides are reported in the results as gene name followed by phosphorylation site within the protein sequence and charge of the measured ion.

Phosphopeptide quantification was performed as described before [16,28]. Briefly, Pescal was used to align the chromatograms of all samples to be analyzed, to construct extracted ion chromatograms (XICs) for the first three isotopes of each phosphopeptide ion selected for quantitation, to pick the peaks to be quantified, and to measure their peak height and area. Windows of molecular mass and retention time were 7 ppm and 1.5 min, respectively. The resulting quantitative data were parsed into Excel files for further normalization and statistical analysis. Phosphopeptide intensities were normalized to the total chromatogram intensity and to the mean value across samples.

PCA, clustering, and statistical analysis of MS data
To avoid interference from low intensity phosphopeptides that have the lowest quantification quality [16], a peak intensity cut-off value was established. Thus, only the phosphopeptides with a maximum intensity across all samples above that cut-off were included in further data analysis by principal component analysis (PCA) and hierarchical clustering, both performed using R software (v. 2.12.2). For this analysis, mass intensity data were log transformed, peptides and arrays were normalized and mean centered. Kendall’s tau coefficient was used to measure the association between peptides and arrays. Finally complete linkage clustering was applied to calculate distances between clusters.

For stringent identification of differences across samples, phosphopeptides were considered differentially phosphorylated between cell types when the Bonferroni-corrected ANOVA and Tukey Pvalues were < 0.01 and the fold difference was > 2. To correlate peptide phosphorylation with drug sensitivity, phosphopeptide intensities derived from each cell line replicate were correlated with their viability after drug treatment at either 1 μM or 10 μM relative to DMSO-treated control. The correlation between relative phosphopeptide intensity and viability was calculated using Pearson’s coefficient (R) of either linear or logarithm transformed phosphorylation data.

LASSO regression analysis
Least absolute shrinkage and selection operator (Lasso) regression is a multivariate variable selection technique with a penalization approach that controls the number of regressors to be included in the optimized model [58]. This is achieved by including a ‘shrinkage’ parameter λ in the cost function.

For each of the three drugs we considered a n x p matrix X (where n = number of biological samples (21 in this study) and p = number of quantified phosphopeptides (2,135 in this study)) containing the column-wise normalized (μ = 0, δ = 1) intensities of all the phosphopeptides measured across the seven cell lines (three measurements for each cell line).

We assessed the potential of these data in predicting the level of resistance/sensitivity of the seven cell lines to three kinase inhibitors through a ‘leave-one-cell-line-out’ approach based on lasso linear regression. With this method, seven different models were trained for each drug, leaving out the three samples corresponding to each of the seven cell lines, in turn. Then the trained models were used to make predictions on the samples that were left out in the corresponding training phase. The predictive ability was evaluated in terms of correlation between predicted and observed percentage of surviving cells when compared to the negative control, after the drug treatment.

For a given drug D, and a cell line C left out from the training phase, let Y be the vector containing the 18 viability scores (respectively for the remaining 18 samples) and λ a non-negative number. We solved through lasso regression the problem:

\[ \min_{\beta_0, B} \left\{ \frac{1}{2n} \sum_{i=1}^{n} (y_i - \beta_0 - x_i^T B)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \]  

where n is the number of observations (that is, the 18 samples from measurements on the remaining six cell lines, in triplicate) y_i is the viability score of sample i following treatment with D; x_i is the row vector containing the normalized intensities of the p phosphopeptides when measured in the i-th sample; \( \beta_0 \) and B are a scalar and a p-vector, respectively. B contains the coefficients of the regressors (that is, all the phosphopeptides) to be optimized. As λ increases, the number of non-zero components (hence phosphopeptides with non-null coefficient in the model) decreases.

We determined the optimal value for the λ parameter with a three-fold cross-validation on the remaining 18 samples and solved equation (1) for vector B without considering the samples of the left out cell line. In order to reduce the instability of the final models across the three-fold cross-validation used to determine λ, these two final steps were repeated 20 times (for each left-out cell line) and the entries of the resulting B vector averaged across these 20 iterations, ending up in the final average model M_{D, C} (that is, final model for drug D, leaving out the cell line C samples). The frequency of
observing a non-null coefficient for each regressor across the 20 iterations (quantifying how much the corresponding phosphopeptide is stably included in the optimal models) was also computed and reported in the final results. The viability of each left-out cell line $C$ was finally predicted through the corresponding $M_{D, C}$.

In order to make the values predicted through $M_{D, C}$ on the left-out samples across the seven different cell lines $C$ and the three drugs $D$ comparable to each other, these values were normalized ($\mu = 0$, $\delta = 1$) together with the predictions of $M_{D, C}$ on the corresponding training set. For the same reason, to produce the scatter plot in Figure 3, all the observed viability were normalized ($\mu = 0$, $\delta = 1$) drug-wise.

To produce a final descriptive model $M_{D^*}$ of response to drug $D$, the coefficients of all the phosphopeptides (and their non-null coefficient frequencies) were averaged across the seven corresponding $M_{D, C}$.

Phosphopeptides whose average non-null coefficient frequency is $> 50\%$ in these final descriptive models are those reported in the insets of Figure 3.

Bioinformatics

Proteins containing phosphopeptides that significantly correlated with phenotypes were used for gene ontology (GO) and pathway enrichment analysis using either an in-house script that matched ontologies listed in SwissProt to each gene product or by David analysis tools [35]. As for phosphorylation motifs analysis, polypeptide sequences were obtained from each phosphopeptide in the dataset by leaving the phosphorylated residue in the center of a sequence that was flanked by seven amino acids on each side. In cases where the phosphorylated residue in the original phosphopeptide had less than seven amino acids at either terminus, these were extended by blasting them against the SwissProt database. Phosphorylation motifs were obtained from MotifX [40] and from the literature [41] to assemble a total of 108 different motifs. Because no differences between the rates at which Ser/Thr kinases phosphorylate Ser and Thr residues have been reported, no distinction was made between p-Ser and p-Thr containing motifs. Peptides phosphorylated at tyrosines were grouped in a single motif. Polypeptide sequences in the dataset were matched to these phosphorylation motifs and the average of the normalized and log-transformed intensities of all the phosphopeptides containing each of the predefined phosphorylation motifs were then averaged and correlated to sensitivity. A script in VBA was written to automate the implementation of these algorithms.

Western blot

AML cell lines were seeded at $5 \times 10^5$ cells/mL. Cells were harvested by centrifugation at $300 \times g$ for 5 min, washed twice with ice cold Dulbecco’s Phosphate Buffered Saline (DPBS), supplemented with $1 \text{mM Na}_2\text{VO}_4$ and $1 \text{mM NaF}$. Cell pellets were lysed with lysis buffer (50 mM tris-HCL pH 7.4, 150 NaCl, 1 mM EDTA, 1% Triton X-100). Protein concentration was calculated using Bradford and 50 $\mu$g of protein were run in 10% SDS-PAGE. Proteins were transferred to PVDF membranes that were block with TBS-Tween (0.1%) supplemented with 5% skim milk. Then membranes were incubated o/n with primary antibody and secondary antibody for 1 h. Protein bands were detected using ECL. Primary antibodies pMARKS S153/156 (Cell Signaling; Cat.2741) and Vinculin (Cell Signaling; Cat. 4650) were used at 1:1,000 and 1:10,000 dilution, respectively. Secondary antibodies were used at 1:5,000 dilution.

Data availability

Mass spectrometry data have been deposited to the ProteomeXchange with identifier PXD000217.

Additional material

Additional file 1: Table S1 - Hematological cell lines used to compare phosphoproteomes of different hematological cancers.
Additional file 2: Table S2 - AML cell lines used to correlate sensitivity to kinase inhibitors with phosphoproteomics data.
Additional file 3: Dataset 1 - Identities and quantitative values of all phosphopeptides identified in AML, lymphoma, and multiple myeloma cell lines.
Additional file 4: Figure S2 - Protein classes represented in the phosphoproteomes of hematological cancer cell lines.
Additional file 5: Figure S3 - Representative examples of phosphopeptides differentially regulated in AML, lymphoma, and multiple myeloma cell lines.
Additional file 6: Dataset 2 - Correlation of phosphoprotein data with responses to kinase inhibitors in AML.
Additional file 7: Figure S4 - Scatter plots between predicted/observed viability scores for individual drugs with cell lines identifiers, correlations scores, and P values.
Additional file 8: Dataset 3 - Final descriptive models of drug responses as resulting from the lasso regression analysis. Listed are predictive phosphopeptides together with their average coefficients and inclusion frequency.
Additional file 9: Figure S5 - Association between the markers of sensitivity to kinase inhibitors found for AML cells with the sensitivity to the same inhibitors in lymphoma and multiple myeloma cells.
Additional file 10: Figure S6 - Pathway analysis of phosphopeptides that correlate with the responses to PI-103.
Additional file 11: Figure S7 - An inhibitor of PKC reduced the viability of AML cells resistant to PI-103 inhibition and had an additive effect with PI-103.

Abbreviations

AML: acute myeloid leukemia; GO: gene ontology; IMAC: immobilized metal affinity chromatography; PCA: principal component analysis; XIC: extracted ion chromatogram.
Authors’ contributions
PC performed research, analyzed data, and wrote the paper; MPA performed research; FL analyzed data and wrote the paper; JCRP analyzed data; BV contributed reagents and edited the paper; JSR analyzed data; SJ contributed reagents and edited paper; PRC designed research, analyzed data, and wrote the paper.
BV is an Advisor to GSK (Stevenage, UK) and Actiomics (London, UK). PRC is an Advisor to Actiomics (London, UK).
All authors read and approved the final manuscript.

Acknowledgements
We thank Alex Montoya for technical assistance and the PRIDE team for depositing the mass spectrometry data in the ProteomeXchange consortium. This work was supported by the MRC (G0700755/82351), BBSRC (BB/G015023/1), and the Barts and London Charity (297/298 and 297/997). FL acknowledges funding from the EMBL-EBI Wellcome Trust Sanger Institute Post-Doctoral (ESPOD) programme.

Author details
1Analytical Signalling Group, Centre for Cell Signalling, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London EC1B 6QQ, UK. 2European Bioinformatics Institute, EMBL-EBI, Wellcome Trust Genome Campus - Cambridge CB10 1SD, UK. 3Cancer Genome Project, Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus - Cambridge CB10 1SD, UK. 4Cell Signalling Group, Centre for Cell Signalling, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London EC1B 6QQ, UK. 5Centre for Haematology-Oncology, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London EC1B 6QQ, UK. 6Current address: MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London, W12 ONN, UK.

Received: 1 December 2012 Revised: 19 April 2013 Accepted: 29 April 2013 Published: 29 April 2013

References
1. Harris NL, Jaffe ES, Diebold J, Flandrin G, Muller-Hermelink HK, Vardiman J, Lister TA, Bloomfield CD. The World Health Organization classification of hematological malignancies report of the Clinical Advisory Committee Meeting, Airlie House, Virginia, November 1997. Mod Pathol 2000, 13:195-207.
2. Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Bandlow B,hev M, Gunsalus FC, Feinman RS, Feinberg AP, et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat Genet 2002, 30:41–47.
3. Levene AP, Morgan GJ, Davies FE. The use of genetic microarray analysis to classify and predict prognosis in haematological malignancies. Clin Lab Haematol 2003, 25:209-220.
4. Golub TR, Slonim DK, Tamayo P, Huard P, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: discovery and class prediction by gene expression monitoring. Science 1999, 286:531-537.
5. Pascal LE, True LD, Campbell DS, Deutsch EW, Risk M, Coleman IM, Eichner LJ, Nelson PS, Liu AF. Correlation of mRNA and protein levels: cell type-specific gene expression of cluster designation antigens in the prostate. BMC Genomics 2008, 9:246.
6. Greenbaum D, Colangelo C, Williams K, Gerstein M. Comparing protein abundance and mRNA expression levels on a genomic scale. Genome Biol 2003, 4:117.
7. Cascante M, Boros LG, Comin-Anduix B, de Atraiu P, Centelles JJ, Lee PW. Metabolic control analysis in drug discovery and disease. Molbiotechnol 2002, 20:243-249.
8. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011, 144:646-674.
9. O’Hare T, Zabirskie MS, Ede CA, Agarwal A, Adrian LT, You H, Corbin AS, Yang F, Press RD, Rivera VM, Toplin J, Wong S, Deininger MW, Druker BJ. The BCR-ABL35INS insertion/truncation mutant is kinase-inactive and does not contribute to tyrosine kinase inhibitor resistance in chronic myeloid leukemia. Blood 2011, 118:5250-5254.
10. Boulus N, Mulder HL, Calabrese CR, Morrison JB, Rehg J, Relling MV, Sherr CJ, Williams RT. Chemotherapeutic agents circumvent emergence of dasatinib-resistant BCR-ABL kinase mutations in a precise mouse model of Philadelphia chromosome-positive acute lymphoblastic leukemia. Blood 2011, 117:3585-3595.
11. Keroff GU, Sigman CC. Cancer biomarkers: selecting the right drug for the right patient. NatRev Drug Discov 2012, 11:201-214.
12. Drucker BJ. Perspectives on the development of imatinib and the future of cancer research. NatMed 2009, 15:1149-1152.
13. Serra V, Scaltriti M, Prudkin L, Eichhorn PJ, Ibrahim YH, Chandarlapaty S, Markman B, Rodriguez O, Guzman M, Rodriguez S, Gil M, Russillo M, Pana JL, Singh S, Arribas J, Rosen N, Basilio J. PI3K inhibition results in enhanced HER signaling and acquired ERK dependency in HER2-overexpressing breast cancer. Oncogene 2011, 30:2547-2557.
14. Liu P, Cheng H, Santiago S, Rauder M, Zhang F, Isabella A, Yang J, Semaan DJ, Chen C, Fox EA, Gray NS, Monahan J, Schlegel R, Beroukhim R, Mills GB, Zhao J. Oncogenic PI3KCA-driven mammary tumors frequently recur via PI3K pathway-dependent and PI3K pathway-independent mechanisms. NatMed 2011, 17:1116-1120.
15. Manning G, Plovman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. Trends Biochem Sci 2002, 27:514-520.
16. Casado P, Cutillas PR. A self-validating quantitative mass spectrometry method for assessing the high-content phosphoproteomic experiments. MolCellProteomics 2011, 10M1100503709.
17. Smerer R, Jensen LJ, Oltersdorff G, van Koot M, Jorensen C, Miron IM, Diella F, Golovil K, Taylor L, Elder K, Metalnikov P, Nguyen Y, Pasculesu A, Jin J, Park JS, Samson LD, Woodgett JR, Russell RB, Bork P, Yaffe MB, Pawson T. Systematic discovery of in vivo phosphorylation networks. Cell 2007, 129:1415-1426.
18. Choudhary C, Mann M. Decoding signalling networks by mass spectrometry-based proteomics. Nat Rev MolCell Biol 2010, 11:427-439.
19. Hus P, Kang SA, Ramieseder J, Zhang Y, Ottna KA, Lim D, Peterson TR, Choi Y, Gray NS, Yaffe MB, Marto JA, Sabatini DM: The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. Science 2011, 332:1317-1322.
20. Goia A, Leroy C, Drullion C, Lagarde V, Etienne G, Dulaucq S, Lippert E, Roche S, Mahon FX, Pasque J-M. Quantitative phosphoproteomics revealed interplay between Syk and Lyn in the resistance to nilotinib in chronic myeloid leukemia cells. Blood 2011, 118:2211-2221.
21. Rubbi L, Titz B, Brown L, Galvan E, Komisiopoulou E, Chen SS, Low T, Tahmasian M, Skaggs B, Muschen M, Pellegrini M, Graeber TG. Global phosphoproteomics reveals crosstalk between Bcr-Abl and negative feedback mechanisms controlling Src signaling. Sci Signal 2011, 4ra18.
22. Choudhary C, Olsen J, Brandts C, Cox J, Reddy PN, Bohmer FD, Gerke V, Schmidt-Arras DE, Berdel WE, Muller-Tidow C, Mann M, Serve H. Methylated activation of oncogenic RTKs switches downstream signalling outcomes. Mol Cell 2009, 36:326-339.
23. Kummer M, Kaminiski M, Ziedler A, Oppermann F, Blencowe S, Max S, Muller S, Trebbe A, Godl K, Schaab C. Phosphosignature predicts dasatinib response in non-small cell lung cancer. MolCellProteomics 2012, 11:651-668.
24. Alcolea MP, Kleiner O, Cutillas PR. Absolute quantification of microbial proteomes at different treated conditions identifies markers of kinase pathway activation and protein kinase enzymatic activities. MolCellProteomics 2011, 10M1100069202.
25. Neilson KA, Ali NA, Muralidharan S, Mizraei M, Mariani M, Assadourian G, Lee A, van Sluyter SC, Haynes PA. Less label, more free: approaches in label-free quantitative mass spectrometry. Proteomics 2011, 11:535-553.
26. Schmidt A, Beck M, Malmstrom J, Lam H, Claassen M, Campbell D, Schmidt-Arras DE, Berdel WE, Muller-Tidow C, Mann M, Serve H. Methylated activation of oncogenic RTKs switches downstream signalling outcomes. Mol Cell 2009, 36:326-339.
27. Alcolea MP, Kleiner O, Cutillas PR. Absolute quantification of microbial proteomes at different treated conditions identifies markers of kinase pathway activation and protein kinase enzymatic activities. MolCellProteomics 2011, 10M1100069202.
28. Alcolea MP, Casado P, Druker BJ. Global phosphoproteomics reveals crosstalk between Bcr-Abl and negative feedback mechanisms controlling Src signaling. Sci Signal 2011, 4ra18.
29. Choudhary C, Olsen J, Brandts C, Cox J, Reddy PN, Bohmer FD, Gerke V, Schmidt-Arras DE, Berdel WE, Muller-Tidow C, Mann M, Serve H. Methylated activation of oncogenic RTKs switches downstream signalling outcomes. Mol Cell 2009, 36:326-339.
30. Alcolea MP, Kleiner O, Cutillas PR. Absolute quantification of microbial proteomes at different treated conditions identifies markers of kinase pathway activation and protein kinase enzymatic activities. MolCellProteomics 2011, 10M1100069202.

Casado et al. Genome Biology 2013, 14:R37
http://genomebiology.com/2013/14/4/R37
Page 18 of 18

30. Stulemeijer IJ, Joosten MH, Jensen ON. Quantitative phosphoproteomics of tomato mounting a hypersensitive response reveals a switch suppression of biosynthetic activity and a differential role for hsp90 isoforms. J Proteome Res 2009, 8:1168-1182.

31. Bodenmiller B, Wanka S, Kraft C, Urban J, Campbell D, Pedrioli PG, Gerin B, Picotti P, Lam H, Witek O, Brunning M, ROCHE, Rimmer B, Lobb RJ, Schlapbach R, Colman-Lerner A, Nolan GP, Niesvzhietski AI, Peter M, Loevith R, von Mering C, Aebersold R. Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. Sci Signal 2010, 3.194.

32. Olsen JV, Blagoev B, Gnad F, Maier J, Macek B, Kumar C, Mortensen P, Mann M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. Cell 2007, 127:635-648.

33. Qureshi AH, Chaoy, Maiguel D, Faridi MH, Barth CJ, Salem SM, Singhal M, Stobb D, Krastins B, Ogihara M, Zaki MI, Gupta V. Proteomic and phosphoproteomic profile of human platelets in platelets, resting state: insights into integrin signaling. PLoS One 2009, 4:e6727.

34. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000, 25:25-29.

35. Huang da W, Sherman BT, Tan Q, Kanes J, Liu D, Bryant D, Bryant D, Guo Y, Stephens R, Brusniak MY, Llampich RA. DAVID Bioinformatics Resources: expanded annotation database and novel tools to better extract biology from large gene lists. Nucleic Acids Res 2007, 35:W169-175.

36. Fingar DC, Salama S, Tsou C, Harlow E, Blenis J. Mammalian cell size is controlled by mTOR and its downstream targets S6K1 and 4EBP1/eur4E. Genes Dev 2002, 16:1472-1487.

37. Ramidson JJ. MARCKS: a case of molecular exaptation? Int J Biochem Cell Biol 2000, 32:475-479.

38. Beretta L, Dubois MF, Sobel A, Curmi PA, Mikoshiba K, Bibb JA. Phosphorylation of the tubulin-binding protein, stathmin, by Cdk5 and MAP kinases in the brain. J Neurochem 2006, 99:237-250.

39. Schwartz D, Gygi SP. An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. Nucleic Acids Res 2005, 23:1591-1598.

40. Miller ML, Jensen LJ, Diel F, Jorgensen C, Tinti M, Li L, Hsiung M, Parker SA, Bordeje X, Sichett-Pontier T, Olovemsky M, PasculoIcu A, Alexander J, Knapp S, Blom B, Ibor P, Li S, Cesareni G, Pawson T, Turk B, Yaffe MB, Brunak S, Linding R. Linear motif atlas for phosphorylation-dependent signalling. Science Signaling 2008, 1.56.

41. Vardiman JW, Thele J, Arber DA, Bunning RD, Borowitz MJ, Porwit A, Harris NL, Le Beau MM, Hellstrom-Lindberg E, Tefferi A, Bloomfield CD. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood 2009, 114:937-951.

42. Sawyers CL. The cancer biomarker problem. Nature 2008, 452:548-552.

43. Gonzalez-Angulo AM, Hennenessy BT, Meric-Bernstam F, Sahin A, Liu W, Lu Z, Casey MS, Myers S, Steer S, Dang C, Broadus S, Lucht A, Aparicio S, Brown P, Pusztai L, Symmans WF, Aisner J, Ovegaard J, Borresen-Dale AL, Hortobsay GN, Coombers KR, Mills GB. Functional proteomics can define prognosis and predict pathologic complete response in patients with breast cancer. Clin Proteomics 2011, 8:11.

44. Lee JM, Kohn EC. Proteomics as a guiding tool for more effective personalized therapy. Ann Oncol 2010., Suppl 7:v205-210.

45. Deeb SJ, D’Souza RC, Cox J, Schmidt-Supprian M, Mann M. Super-SILAC Allows Classification of Distinctive Large B-cell Lymphoma Subtypes by Their Protein Expression Profiles. Mol Cell Proteomics 2012, 11:77-89.

46. Hennenessy BT, Smith DL, Ram PT, Lu Y, Mills GB. Exploiting the PI3K/AKT pathway for cancer drug discovery. Nat Rev Drug Discov 2005, 4:988-1004.

47. Roberts PJ, Der CJ. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. Oncogene 2007, 26:3291-3310.

48. Dan S, Okamura M, Seki M, Yamazaki K, Sugita H, Okui M, Mukai Y, Nishimura H, Asaka R, Nomura K, Ishikawa Y, Yamori T. Correlating phosphatidylinositol 3-kinase inhibitor efficacy with signaling pathway status: in silico and biological evaluations. Cancer Res 2010, 70:4962-4994.

50. Janku F, Wheler JJ, Westin SN, Moulder SL, Naing A, Timberland AM, Fu S, Falchook GS, Hong DS, Garrido-Laguna I, Luthra R, Lee JJ, Lu KH, Kurzrock R. PI3K/AKT/mTOR Inhibitors in Patients With Breast and Gynecologic Malignancies Harboring PRKCAKA2 Mutations. Clin Oncol 2012, 10:777-782.

51. Pilotto O, Wright B, Brown P, Kn RT, Lewis M, Small D. Prolonged exposure to FLT3 inhibitors leads to resistance via activation of parallel signaling pathways. Blood 2007, 109:1643-1652.

52. Balmanno K, Kell SD, Gillings AS, Hayat S, Cook SJ. Intrinsic resistance to the MEK1/2 inhibitor AZD6244 (ARRY-142886) is associated with weak ERK1/2 signalling and/or strong PI3K signalling in colorectal cancer cell lines. Br J Cancer 2009, 125:2332-2341.

53. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 1999, 20;3551-3567.

54. http://expasy.org/spivot/

55. Vazianio JA, Cote RG, Cordas A, Dianes JA, Fabregat A, Foster JM, Giss J, Alp E, Birim M, Contell J, O’Kelly G, Schoenegger A, Ovleko R, Perez-Riverol Y, Reissinger F, Rios D, Wang R, Herrajak H. The PIProteomics IDentifications (PRIDE) database and associated tools: status in 2013. Nucleic Acids Res 2013, 41:D1063-1069.

56. http://proteomexchange.org.

57. Savitski MM, Lemeer S, Boesche M, Lang M, Mathieson T, Bantscheff M, Kuster B. Confident phophorylation site localization using the Mascot Delta Score. Mol Cell Proteomics 2010, 10:M7003380.

58. Tibshirani R. The lasso method for variable selection in the Cox model. StatMed 1996, 13:385-395.

59.http://www.biomedcentral.com/submit

Submit your next manuscript to BioMed Central and take full advantage of:

• Convenient online submission
• Thorough peer review
• No space constraints or color figure charges
• Immediate publication on acceptance
• Inclusion in PubMed, CAS, Scopus and Google Scholar
• Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit