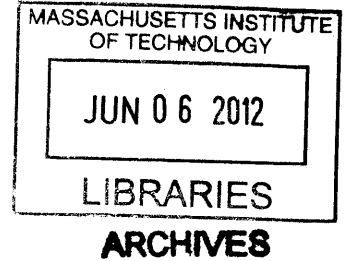


**A Systems Approach to Uncovering the Adaptive Response of Cancer to Targeted Therapies**

by

Adrian Joseph Randall

B.S., University of North Carolina (2009)



Submitted to the Computational and Systems Biology Program  
In Partial Fulfillment of the Requirements for the Degree of

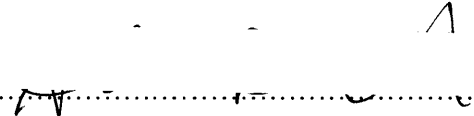
Master of Science

at the

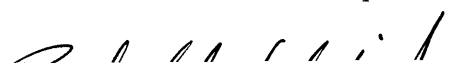
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

© 2012 Massachusetts Institute of Technology. All rights reserved.

Signature of Author..... 

Computational and Systems Biology Graduate Program  
May 25<sup>th</sup>, 2012



Certified by.....

Forest M. White  
Associate Professor of Biological Engineering  
Thesis Supervisor



Accepted by.....

Chris Burge  
Professor of Biology and Biological Engineering  
CSB Graduate Program Director



## Contents

Abstract.....	3
Acknowledgements.....	4
List of Figures.....	5
Chapter 1 – Tumor Adaptive Response to Chemotherapy.....	6
1.1 Tyrosine Kinases and their Inhibitors.....	6
1.2 Phosphorylation and Studying Network Dynamics in NSCLC.....	9
1.3 Cancer Response to Kinase Inhibition.....	13
1.4 Profiling the phosphorylation network of EGFR driven NSCLC.....	14
1.5 Response of NSCLC to EC5 leads to a potential boost in proliferation.....	19
1.6 Conclusions and looking forward.....	21
Chapter 2 – Long-term Resistance to Targeted Therapy.....	23
2.1 Mechanisms of long-term resistance to Kinase Inhibitors.....	23
2.2 Development of TKI resistance <i>in vitro</i> in NSCLC.....	23
2.2 EGFR can compensate for loss of EML4-ALK Signaling.....	25
2.3 Conclusions and looking forward.....	28
Chapter 3 – Using 14-3-3 Protein to Interrogate Phosphoserine and Phosphothreonine Signaling.....	29
3.1 Phosphoserine and phosphothreonine Signaling – the bigger picture.....	29
3.2 Use of 14-3-3 to interrogate phosphoserine and phosphothreonine signaling.....	29
3.3 14-3-3 pulldown of phosphorylated peptides.....	30
3.4 14-3-3 Pulldown isolates signaling protein and allows for studying signaling dynamics.....	32
3.5 Conclusions and future directions.....	33
Chapter 4 – Computational Tools for Analyzing Mass Spectrometry Data.....	34
4.1 Computational tools to process peptide data from a mass spectrometer.....	34
4.2 Generation of scripts for the automation of routine steps in data formatting and cleanup.....	34
4.3 Design of scripts to automate cleanup, averaging, and normalization of mass spectrometry data.....	35
4.4 Conclusions and looking forward.....	40
Conclusions and Looking Forward.....	41
Supplemental Information.....	44
Maps of 14-3-3 Vectors.....	44
Sample Data from a 14-3-3 Pulldown of Serum Stimulated HCC827 Cells.....	46
References.....	47

## **Abstract**

Tyrosine kinase inhibitors have significant promise in the fight to develop agents that can target cancer in a tumor-specific manner. A number of drugs have been and are currently in development to inhibit specific kinases that can mediate uncontrolled proliferation; however, an unfortunate eventuality for most patients receiving these treatments is the development of resistance that renders these drugs almost completely ineffective. While a number of mechanisms can evolve within a tumor to mitigate effects of kinase inhibitors, we sought to uncover what changes are occurring in the tyrosine phosphorylation network at both short timescales (minutes to 72 hours) and long timescales (120 hours+) that can be playing a role in helping a tumor become resistant to driver-kinase inhibition. It is our hypothesis that specific feedback networks are able to detect and overcome driver kinase inhibition through activation of potential other pathways, which can go on to mediate a longer term resistance phenotype.

In order to probe dynamics in the tyrosine phosphorylation network, we employed mass spectrometry to analyze peptides derived from six non-small cell lung cancer cell lines that we classify as either EGFR+ or EML4-ALK+. From both mass spectrometry data and growth assays, we identified an unintuitive boost in signaling and growth in response to low inhibitor concentrations, suggestive of a cellular mechanism that is adaptive to driver kinase inhibition. Studies of EML4-ALK driven H3122 cells showed that this short-term response is not the same as the known long-term resistance mechanism to ALK inhibition, leading support to the notion that the short-term “adaptive response” may be a novel type of mechanism to aid tumor adaptation to targeted therapies. In an effort to better probe signaling events occurring downstream of the phosphotyrosine network, a new pull down technique for mass spectrometry using 14-3-3 protein against phosphoserine and phosphothreonine peptides is described. The results of these studies open up many potential avenues for further exploration into the immediate and long-term signaling response of cancer to targeted therapies.

Thesis Supervisor: Forest M. White

Title: Associate Professor

## Acknowledgements

There are a number of people to whom I'd like to express my most sincere gratitude. I'd like to thank the CSB program for allowing me the opportunity to study and perform research at MIT, in particular Chris Burge for helping, advising, and supporting me during my time here. I'm grateful to Forest White for the opportunity to work in his lab and exposing me to the world of cancer signaling and mass spectrometry. I'd especially like to thank Dan Sears for being a wonderful colleague and friend this past year, with whom most of this work was completed. I look forward to seeing the wonderful and creative research I'm sure you'll do in your life. I only regret that we never got to see this work through together, but I wish you all the best in life and research.

I'd like to thank my undergraduate research advisor, Jack Griffith, for seeing potential in me and supporting me in my quest to do what I can to help improve the world. Without your guidance and mentorship, I wouldn't have dreamed of having the opportunity to study at MIT, and for seeing more in me than I ever saw in myself I'm eternally grateful.

I'd like to thank my friends for enriching my life in a very meaningful way. For motivating me, believing in me, supporting me, and for all the happy memories together and those yet to come, I'm forever grateful, especially to W. Anthony Calore, Tim Y. Chi, Juan Camilo Mendez Guzman, and Jorge Izquierdo. I look forward to all the amazing things I'm confident you all will do in life, and I hope to be able to help you in all of your pursuits.

A special thanks to my best friend, Lindsay Aldridge. Thank you for being supportive, encouraging, patient, loving, and when I need to hear it most, honest. I wish to always be there for you.

Lastly, I'd like to dedicate this thesis to my mother, Dr. Elizabeth Randall, and my sister, Dr. Eveleen Randall. We've been through hell and high water, and there are very few people who could have predicted that we would make this far together. I can't think of a better pair of people with whom I would have rather navigated life and its interesting turns together. Your endless selflessness and patience with me is worthy of the highest recognition. Thank you.

## List of Figures

Figure 1: Sensitivities of NSCLC cell lines to kinase inhibitors. ....	14
Figure 2: Measurement of basal levels of phosphorylation in untreated, serum-depleted NSCLC cell lines.....	15
Figure 3: Response of EGFR-driven cell lines to gefitinib 72 hours post EC95 treatment .....	16
Figure 4: Response of HCC827 to varying amount of kinase inhibitor.....	18
Figure 5: Cell counts relative to untreated 48 hours post treatment with Gefitinib and PF-1066.....	19
Figure 6: Mass spectrometry analysis of H3122 cells after short-term treatment with PF-1066.....	20
Figure 7: Growth experiment of EGFR+ NSCLC cell lines in gefitinib to measure timescales of resistance.....	24
Figure 8: Light microscopy images of live H3122 cells that are sensitive and resistant to ALK inhibitor PF-1066.....	26
Figure 9: Sensitivity of H3122 to various combined TKI treatments.....	26
Figure 10: A phosphoproteomic comparison of sensitive and PF-1066 resistant H3122 cells.....	27
Figure 11: Schematic of the E. coli expression vector for adding a single biotin group N-terminal to a protein of interest.....	30
Figure 12: Coomassie stain and Western Blot of 14-3-3 lysate and streptavidin beads confirming expression and purification.....	32
Figure 13: Dynamics from HGF-stimulated H1975 cells showing expected patterns in phosphorylation sites with known functions.....	32
Figure 14: A sample image of the GUI highlighting the number of options for data cleanup and processing .....	35
Figure 15: Sample data output from processing with the mzxmlread MATLAB function combined with in-house scripts to reference and combine peptide sequence information inside an HTML output file from MASCOT.....	36
Figure 16: Sample output of the cleanupTXT.m script as described above .....	38
Figure 17: Sample of iTRAQ data of 9 scans that are not averaged with using a 100% threshold. ....	38

## Chapter 1 – Tumor Adaptive Response to Chemotherapy

### 1.1 Tyrosine Kinases and their Inhibitors

Landmark work in the 1950s by Rita Levi-Montalcini and Stanley Cohen led to the discovery of growth factors and generated the idea of specific proteins (termed hormones then) that could drive growth of transformed cells in a growth factor-specific manner.[1-3] Continued research in the 1960s by Stanley Cohen led to the discovery of another growth factor, epidermal growth factor (EGF), termed for its ability to drive cellular proliferation in epithelial cells. [4] It was not, however, until the late 1970s that a receptor to epidermal growth factor, EGFR, was discovered, and that upon cellular treatment with EGF that cellular uptake of  $^{32}\text{P}$  increased, leading to the notion that growth factors lead to kinase activity, possibly by EGFR itself, and that growth factors acting through receptors can induce downstream phosphorylation events. [5-7] Concurrent to this finding was the discovery of kinases that phosphorylate tyrosine (tyrosine kinases) and that kinases have the ability to transform cells into a malignant state, further driving home the notion that growth factors, kinases, and phosphorylation events drive proliferation in both normal and abnormal contexts.[8, 9]

In a normal context, tyrosine kinases are an integral component of the cellular signaling network that drive phenotypes like growth, proliferation, migration, differentiation, and even cell death. After research uncovered that a tyrosine kinase, SRC, can transform a cell into a malignant phenotype, followed by the discovery that epithelial cancers can highly overexpress another tyrosine kinase EGFR led to intense speculation that like SRC, it too could be an important oncogenic driver. [10-13] In 1983, in a landmark study by John Mendelsohn and colleagues, it was shown that monoclonal antibodies engineered to target EGFR inhibit proliferation, leading to the later development of cetuximab (Erbiximab) as a therapy for colorectal and head and neck cancers, currently on the market today. [14] Around the same time, researchers at Genentech also developed a monoclonal antibody against the related HER2 receptor tyrosine kinase and saw similar antiproliferative effects, leading to the first approved kinase-targeted therapy, trastuzumab (Herceptin) for HER2+ breast cancers. [15]

In addition to targeting receptor tyrosine kinases (RTKs), a variety of therapies have been developed to target cytosolic kinases as well. For example, in 1960, Peter Nowell and David Hungerford noted that in cancer cells derived from leukemia patients that chromosome 22 was smaller than normal, and follow-up work identified that the so-called “Philadelphia Chromosome” contained a genetic translocation in the two genes BCR and ABL and that the gene fusion product was constitutively active and could induce CML in mice. [16, 17] Through high-throughput screening of drug compounds, imatinib (Gleevec) was discovered and shown to inhibit the BCR-ABL fusion product and slow growth and even induce apoptosis in CML cells. [18] With the paradigm of kinase inhibition leading to cancer cell death in place, a rigorous search for both compounds and drug targets ensued, leading to the discovery a number of both kinases and inhibitors that have had success in the clinic.

Lung cancer is the leading cause of cancer deaths worldwide, primarily due to smoking, and because of the difficulty of diagnosis it is normally at an advanced, metastatic stage before treatment and thus has a comparatively poor prognosis. [19] While more kinases have been uncovered over the years as cancers drivers in lung cancer, a significant number of cancers with no known oncogenic drivers remains. Genetic mutations found in signaling pathway genes involved with proliferation in NSCLC (with corresponding frequency) include mutations in KRAS (10-30%), BRAF (2%), EGFR (10-40%), HER2 (4%), ALK (7%), MET (14%), and P13KCA (2%). [19] Even with knowledge of potential drivers, it is estimated that roughly half of non-small cell lung cancer cases are driven through unknown mechanisms, and due to the difficulty of diagnosis as well as the lack of effective targeted treatments, five-year survival rates in patients are only on the order of roughly 15%. [20] In order to address this problem, we selected a study system of six non-small cell lung cancer (NSCLC) cell lines to probe kinase-driven events that could be responsible for abnormal cell growth. Of the six cell lines, four we classified as EGFR+ and two we classified EML4-ALK+.

“Activating” mutations in EGFR that act to destabilize the inactive, non-ligand bound form of the receptor have been isolated in clinical samples, and provide a rational basis for treatment using targeted

inhibitors. [21, 22] The activating mutations L858R and delE746-A750 comprise almost 90% of mutations in EGFR found in NSCLC, and these mutations confer sensitivity to EGFR inhibitors erlotinib (Tarceva) and gefitinib (Iressa).[23, 24] Interestingly, these mutations are strongly correlated with non-smoker lung cancer patients, for reasons that as of yet appear to be unknown. [25] However, treatment of lung cancer with gefitinib in patients that have been preselected for activating mutations in EGFR leads to an improvement in median progression-free survival to only 10.8 months, compared to 5.4 months with carboplatin and paclitaxel, lending credence to the idea that identifying driver kinases may only be a small part of the clinical solution for treating NSCLC. [26]

In addition to mutations in EGFR, roughly 7% of lung cancers contain genetic fusions of echinoderm microtubule-associated protein-like 4 (EML4) and anaplastic lymphoma kinase (ALK) genes, leading to constitutive growth activation and cellular transformation. [27] This fusion gene was discovered through a creative study involving creation and isolation of a cDNA library from a single patient sample. DNA was transfected and screened for transformation ability; from a particular colony of transformed cells, and one clone was sequenced and found to align to both the EML4 and ALK genes, suggesting the transforming agent was a fusion product of the two genes. [27] While a provocative finding, fusions between ALK and nucleophosmin (NPM) have previously been noted in non-Hodgkin's lymphoma, giving precedence that an ALK translocation event could be possible [28] The exact mechanism of how a genetic fusion can be possible at such a specific locus appears to remain unclear. However, from biochemical studies, constitutive activation of the ALK kinase is thought to be due to induced oligomerization due to the basic domain of EML4. [27] An inhibitor of the ALK kinase, TAE684, was discovered and found to prevent growth of an EML4-ALK+ tumor model in mice, providing the basis for clinical develop of this and other ALK inhibitors. [29] Interestingly, though, in the same study it was found that coactivation of HER2 and EGFR genes can compensate for inhibition of EML4-ALK, where cells containing HER2 and EGFR overexpression could withstand ALK inhibition. An EGFR/HER2 inhibitor combined with ALK inhibition, however, lead to loss of phosphorylation on



ERK and AKT and subsequent apoptosis, suggesting that these “compensatory” kinases could provide the basis for targeted cocktail therapies.

## 1.2 Phosphorylation and Studying Network Dynamics in NSCLC

In order to study the phosphorylation dynamics in NSCLC, mass spectrometry was employed to analyze and quantify the phosphoproteome of lung cancer cell lines. Liquid chromatography mass spectrometry (LC-MS/MS) has been developed for the study of proteins with post-translational modifications, and the advent of faster and more sensitive analyzers has allowed for detection and quantification of hundreds to thousands of unique peptides and phosphorylation sites. [30-33] Basic workflow for all mass spectrometry experiments described in this thesis includes the isolation and cleanup of peptides from crude cell lysates followed by liquid chromatography over an acetonitrile gradient and MS/MS sequencing and quantification.

At the beginning of this project, the question of timescales of adaptation and selection in cancer in response to treatment was considered in order to attempt to better understand how resistance develops in NSCLC. It has previously been shown that long-term adaptation to kinase inhibition can occur in cancer cell lines, and following a similar pattern, kinase inhibitor resistance has been noted clinically. Until more recently, however those mechanisms have remained unclear. Further complicating matters, studies have shown that the cells that develop resistance to TKI treatment tend to display heterogeneity in their resistance mechanism. For example, in a previous study of EGFR-driven NSCLC H1650 cells, cancer cells were made resistant to the EGFR inhibitor erlotinib (Tarceva), and of the screened drug resistance clones, only 13% had a mesenchymal-like morphology and molecular characteristics of having undergone an epithelial-mesenchymal transition (EMT) that is characteristic of EGFR-inhibitor resistant cancers, leading to a question of what occurs in the other 87%. [34, 35] Therefore, given this heterogeneity in response *in vitro*, it seems possible that tumor bulk could respond in a similar fashion, and if so, would make downstream treatment significantly more challenging. [36] To address this concern, we sought to uncover how cells responded to kinase inhibition at shorter time scales (hours to

days, i.e. before they develop resistance) with the goal of potentially mitigating a heterogeneous long-term response that may hamper clinical outcomes down the line in patients.

Before performing mass spectrometry, the inhibitors gefitinib (SelleckChem, Houston, TX, USA) and PF-1066 (ChemieTek, Indianapolis, IN, USA) were applied to the EGFR and EML4-ALK cell lines, respectively, at a range of 3nM to 10uM to determine sensitivity of these cells to kinase inhibition. Cell lines at least one passage post-thawing were cultured in RPMI-1640 supplemented with 10% FBS and 1x Streptomycin-Penicillin antibiotic solution (referred to in this thesis as RF-10). Three thousand cells were counted using a Cellometer Vision Analysis System (Nexcelom, Lawrence, MA, USA), and placed into a 96-well plate and allowed to attach overnight. The next day, the media was removed and replaced with RF-10 and the kinase inhibitors were applied. After 72 hours in the presence of inhibitors, cell viability was analyzed by adding the CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS) (Promega, Madison, WI, USA) colorimetric assay per standard protocols and measured absorbance at wavelength 490nm using an Infinite® M1000 PRO plate reader (Tecan, Männedorf, Switzerland) allows for the approximation of cell numbers remaining. Effective concentration (EC) values were calculated from dose-response curve fits made using GraphPad Prism software and GraphPad's online ECAAnything calculator (GraphPad Software, La Jolla, CA, USA).

To prepare protein samples, cell lines were cultured *in vitro* in RPMI-1640 including 10% FBS and 1x Strep-Pen antibiotic solution. Cells were plated at a lower density (1.5 million cells per 15 cm plate, or roughly 10,000 cells/cm<sup>2</sup>) to maintain consistency with MTS assays used previously for finding EC values. Cells were left overnight for 24 hours to allow them to attach to the culture dish and kinase inhibitors were added with fresh RF-10 media. In order to profile responses to match MTS experiments, EGFR cell lines were treated with gefitinib at their corresponding EC95 values for 72 hours, and then washed with cold PBS and lysed using cold 8M urea containing 1mM sodium orthovanadate to inhibit phosphatases. Bulk protein from cell lysates was quantified using the BCA assay kit per standard protocols (Pierce Biotechnology, Rockford, IL, USA) and an equal amount of protein from each condition

was reduced with 10mM DTT for 1 hour at 56°C and alkylated with 55mM iodoacetamide for 1 hour under constant rotation. Each tube was then diluted with 12mL of 100mM ammonium acetate buffer (pH 8.9) and digested with trypsin (Promega, Madison, WI, USA) overnight at room temperature at a 1:100 enzyme:protein mass ratio. Peptides samples were acidified with 1mL of glacial acetic acid and desalted with 10mL of 0.1% acetic acid on a C18 Sep-Pak Plus cartridge (Waters, Milford, MA, USA), followed by elution with 10mL of a 0.1% acetic acid/25% acetonitrile solution and lyophilized to dryness.

Prior to mass spectrometry analysis, samples for studying dynamics were iTRAQ labeled to allow for relative quantification of identical peptides across differing conditions. iTRAQ labeling involves the attachment of an isotopic label to primary amines to “tag” peptides from a specific condition as a basis for measuring relative changes in peptide levels across different samples. [37] For 8-plex iTRAQ experiments, 400ug of lyophilized peptide was dissolved in 30 uL of a 0.5 M Triethylammonium bicarbonate (TEAB) dissolution buffer. iTRAQ reagents (Applied Biosystems, Foster City, CA, USA) were thawed and 70 uL of isopropanol was added to each tube. After vortexing each tube for 1 minute, the iTRAQ reagents were individually added to lyophilized peptide, and the subsequent mixture was vortexed and spun down and allowed to sit at room temperature for 2 hours. Each vial was then spun down and dried to ~30 uL using a speed-vac to stop the labeling reaction, and then each tube was combined into one tube, followed by two washing steps with 25% acetonitrile in 0.1% acetic acid to wash iTRAQ-labeled peptides bound to the walls of each tube. After this washing step, the sample was then dried down in a speed-vac again, however this time to complete dryness.

In order to enrich for peptides with tyrosine phosphorylation, an immunoprecipitation (IP) using monoclonal antibodies against phosphotyrosine were attached to a bead resin. Twelve micrograms of three anti-phosphotyrosine antibodies, 4G10 (EMD Millipore, Billerica, MA, USA), PT-66 (Sigma-Aldrich, St. Louis, MO, USA) and P-Tyr-100 (Cell Signaling Technology, Danvers, MA, USA) were incubated for 8 hours with 70 uL of washed Protein G slurry (EMD Millipore, Billerica, MA, USA) at 4°C under constant rotation. Once the iTRAQ-labeled pellet was completely dried down, it was

resuspended in 150 uL of iTRAQ IP buffer (100 mM TRIS, 100 mM NaCl, 1% NP-40, pH 7.4), 300 uL of HPLC-grade water (Sigma-Aldrich, St. Louis, MO, USA) and the pH was adjusted to 7.4 using TRIS buffer. Resuspended peptides were added to the Protein G slurry loaded with anti-phosphotyrosine antibodies and incubated under constant rotation at 4<sup>o</sup>C overnight.

After overnight incubation, beads were gently spun down and washed with 450 uL of IP buffer (100 mM TRIS, 100 mM NaCl, 0.3% NP-40, pH 7.4) once for five minutes, followed by three washes with 450 uL of rinse buffer (100 mM TRIS, 100 mM NaCl, pH 7.4). Peptides were eluted from the antibody in 60 uL of 100 mM glycine pH 2 for 30 minutes. In order to further enrich and boost the signal of phosphorylated peptides, immobilized-metal affinity chromatography (IMAC) was performed on capillary tubing and high pressure bombs. [32] The elution from the peptide IP was applied to a 10 cm, 200 um inner diameter capillary column packed with Poros MC 20 um beads loaded with >10 column volumes of 100 mM FeCl<sub>3</sub> (Sigma-Aldrich, St. Louis, MO, USA) at a 1 uL/minute flow rate. The column was rinsed with an Organic Rinse solution (25% acetonitrile, 1% acetic acid, 100 mM NaCl) for 10 minutes at a 10 uL/minute flow rate, acidified with 0.1% acetic acid for 10 minutes at a 10 uL/minute flow rate, and then eluted onto a 10 cm, 100 um inner-diameter precolumn packed with 5 um C18 beads (Waters, Milford, MA, USA) with 40 uL of 250 mM sodium phosphate pH 8.0 buffer at a flow rate of 1 uL/minute. The precolumn with loaded peptides was attached to an analytical column packed with 10 cm of 5 um C18 beads with a sub-micron tip to ensure nanoliter flow. An acetonitrile:acetic acid HPLC gradient was applied to the columns as previously described. [38, 39] Sequencing and quantification of peptides was performed by electrospray ionization followed by tandem MS on an LTQ-Orbitrap instrument (Thermo Fisher Scientific, Waltham, MA, USA).

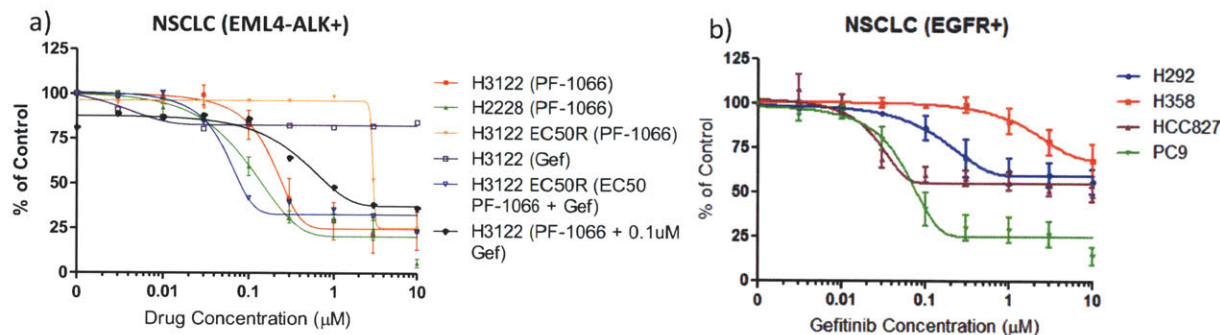
MS/MS spectra were extracted from Thermo RAW files through generation of an MGF file using MSQuant software as well as custom MATLAB (Mathworks, Natick, MA, USA) scripts for extracting peptide quantification information. Peptide sequence identification was extracted through MASCOT software using the Human-2009 protein database (Matrix Bioscience, Boston, MA, USA). [40] To correct

for differences in peptide loading, all data was normalized to the supernatant containing non-phosphorylated peptides from the phosphotyrosine IP at a dilution of 1:1000; the mean and median iTRAQ values were calculated for each channel and fold-changes were applied such that each peptide in the 113 (or 114 in 4-plex experiments) channel had a normalized fold change equal to 1. Peptides with MASCOT scores less than 25 containing no phosphotyrosine or having a missing iTRAQ channel were removed, and peptides with multiple spectra were averaged together using custom MATLAB scripts described in Chapter 4.

For short-term adaptive responses in the tyrosine phosphorylation network, 1.5 million cells were plated and allowed to attach and deplete serum overnight for 24 hours, except inhibitors were added at 10, 100, and 1000 nM for only 10 and 60 minutes followed by lysis and protein collection using the protocol described above. In order to measure proliferation to correlate phosphorylation data from short-term experiments, 100,000 cells were plated in a 6-well dish and left 24 hours to settle in a similar fashion. Inhibitors were added (without fresh media) at concentrations of 10, 100, and 1000 nM and allowed to incubate for 48 hours. Before harvesting cells, they were washed once with trypsin-EDTA, and then 500 uL of trypsin-EDTA was applied. The cells were incubated with trypsin for 15 minutes (or until they all detached) at 37°C, and then 500uL of RF-10 media was applied and mixed. The 1 mL sample was then counted on a ViCell instrument (Beckman Coulter, Brea, CA, USA) using the default method.

### **1.3 Cancer Response to Kinase Inhibition**

As stated previously, the model system chosen for studying changes in the signaling network in response to kinase inhibitor treatment entails six non-small cell lung cancer (NSCLC) cell lines *in vitro*.



**Figure 1:** Sensitivities of NSCLC cell lines to kinase inhibitors (a) Crizotinib (PF-1066) and (b) Iressa (Gefitinib) as measured by MTT assay 72 hours post-treatment. H3122 EC50R cells are resistant to its EC50 of PF-1066.

Four cell lines, PC9, HCC827, H292, and H358 were studied as models for EGFR+ cancers, while two other cell lines, H2228 and H3122 contain the EML4-ALK gene fusion previously described. Prior to mass spectrometry experiments, the sensitivity of the six cell lines to their respective inhibitors was established (Figure 1). In the two EML4-ALK cell lines studied, both demonstrate sensitivity to the ALK inhibitor PF-1066, establishing that constitutive ALK kinase activity leads to cell growth (Figure 1a, red and green lines). Of the EGFR cell lines, the HCC827 and PC9 cells were found to be the most sensitive, while the H292 and H358 cell lines display little sensitivity to gefitinib (Figure 1b). The two cell lines most sensitive to gefitinib, HCC827 and PC9, both contain the activating delE746-A750 mutation in EGFR, which is predictive of sensitivity to gefitinib. [41, 42] In addition, as expected the KRAS-driven H358 cell line was largely insensitive to gefitinib, since the constitutive G12V mutation can activate downstream survival and proliferation independently of EGFR. [43] Lastly, the H292 cell line that has been previously shown to have wildtype EGFR and KRAS demonstrated lower sensitivity to gefitinib. [23, 44]

#### 1.4 Profiling the phosphorylation network of EGFR driven NSCLC

While the presence of a known driver mutation can be predictive of the sensitivity of cell lines to inhibitors designed to target it, the MTS assay experiments from Figure 1 suggest that cells can demonstrate a varied response, leading to the question of how cells adapt at both very short-term (<1

hour) and medium-term (72 hours) timescales. In order to address the question of adaptation at 72 hours, 1.5 million cells of the four EGFR cell lines were plated in 15 cm plates and analyzed for both basal levels of signaling as well as treatment with an EC95 concentration of gefitinib to see how the phosphorylation network changes at that time point (Figures 2 and 3).

In measurements of basal levels of phosphorylation in untreated, serum-depleted cells, not surprisingly, the two most sensitive cell lines, HCC827 and PC9, contained the highest amounts of phosphorylation on EGFR, indicative of its status as the predominant driver gene (Figure 2). Interestingly, phosphorylation of Met in the HCC827 cell line was also very high compared to other cell lines studied. It has previously been shown that Met can drive resistance to EGFR inhibitors and that a small fraction of cells contain duplications in the Met gene suggesting that selection of Met overexpressing cells is a

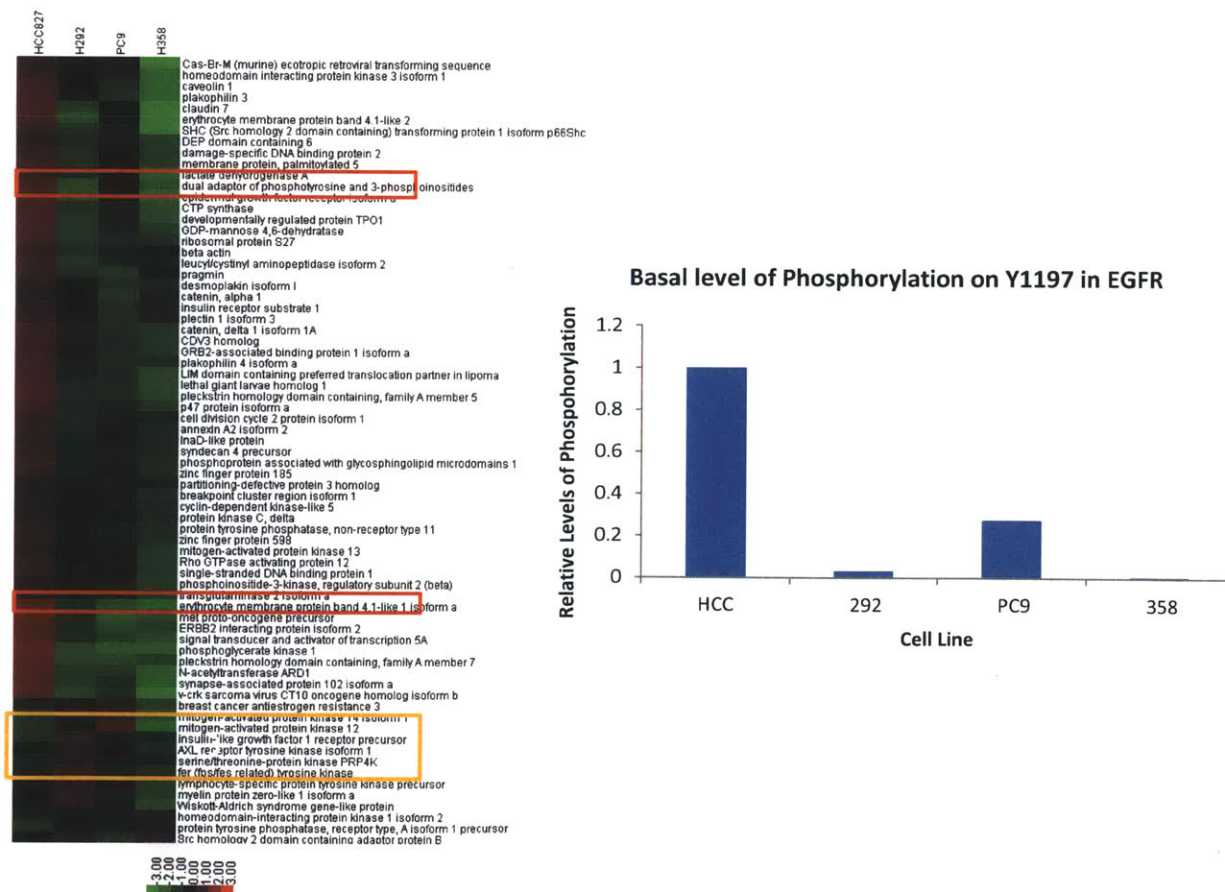
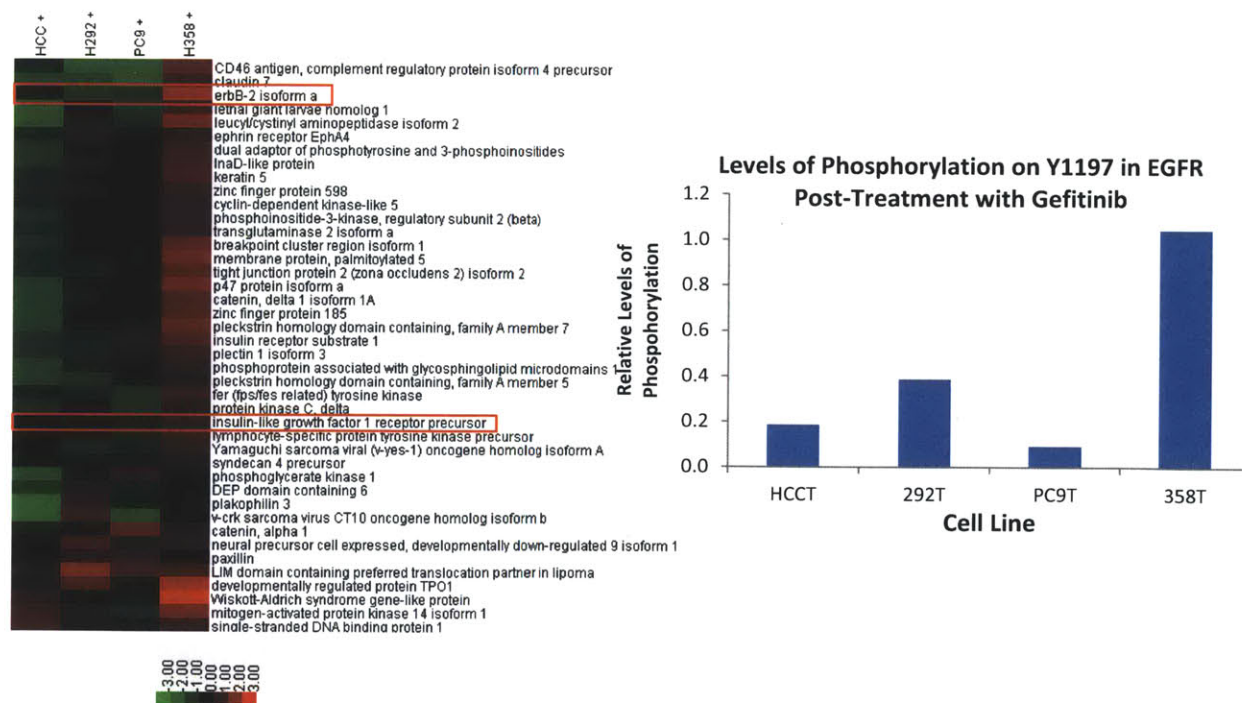


Figure 2: a) Measurement of basal levels of phosphorylation in untreated, serum-depleted NSCLC cell lines. Data is mean-centered, log 2 transformed and clustered using a Euclidean distance metric. b) Levels of phosphorylation on Y1197 in EGFR correlates strongly with sensitivity to gefitinib. Heatmap courtesy of Dan Sears.

mechanism for how a tumor can adapt to gefitinib. [42, 45] Other data of interest includes the high levels of phosphorylation in the receptors IGF1R and AXL, as well as a number of src-family kinases in the H292 cell line, suggesting other potential drug targets (highlighted in the yellow box in figure 2 to the left). While not approved for clinical use yet, drugs targeting the IGF1R and AXL receptors are currently being explored for treating lung as well as other types of cancer. [46, 47]

In cells treated with their corresponding EC95 concentration of gefitinib, levels of phosphorylation on the Y1197 site on EGFR decreased in the HCC827, H292, and PC9 cell lines 72 hours after treatment with EC95 gefitinib. Interestingly, however, levels of phosphorylation in H358 cells did not decrease (Figure 3). It has previously been shown that KRAS can drive expression of the EGFR ligands TGF $\alpha$ , amphiregulin, and heparin-binding EGF-like growth factor, which could explain the robust levels of phosphorylation seen in the KRAS mutant H358 cells even in the presence of gefitinib. [48, 49] In addition, phosphorylation of other receptor tyrosine kinases IGF1R and HER2 increased in these cells

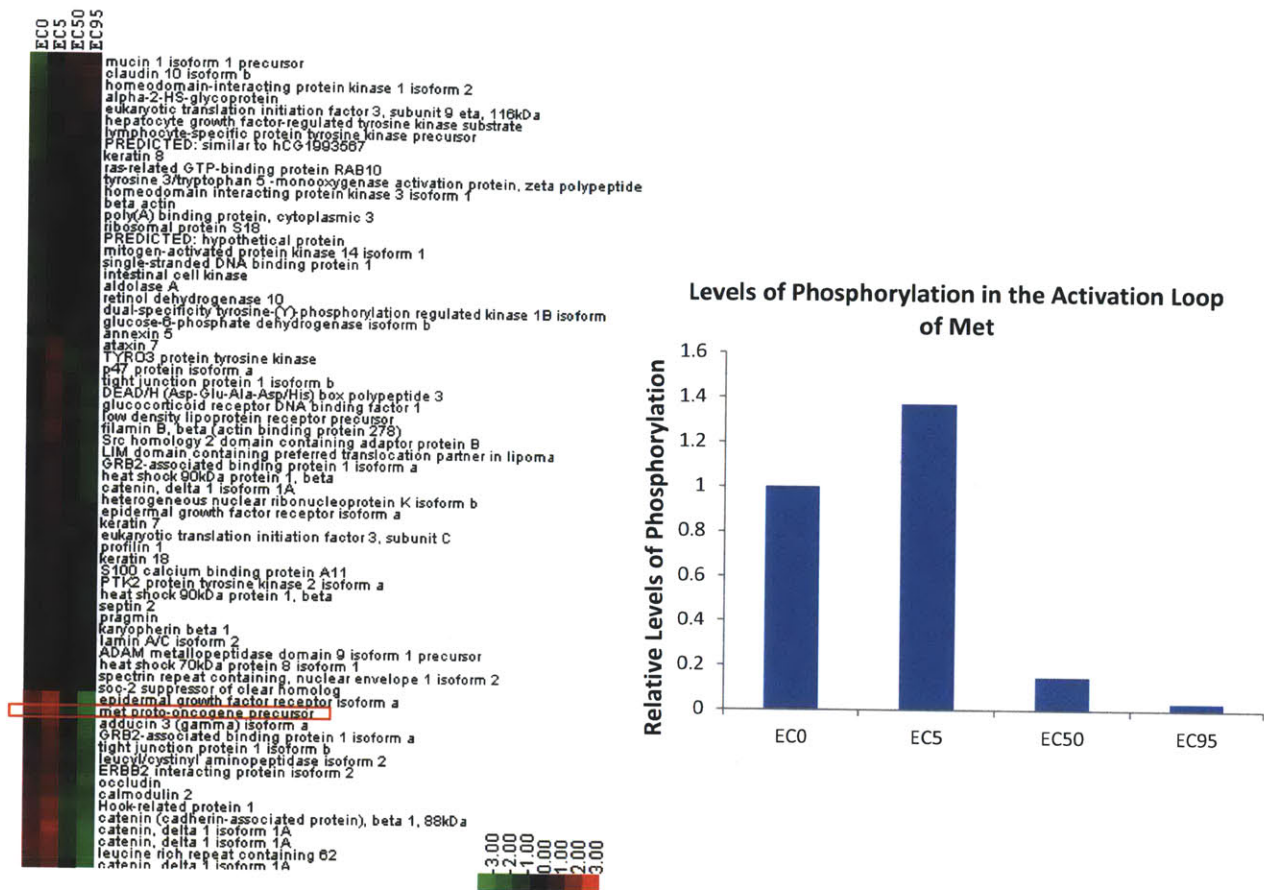


**Figure 3: a) Response of EGFR-driven cell lines to gefitinib 72 hours post EC95 treatment. While the most sensitive HCC827 cell line shows little to no increases in phosphorylation, the most resistant H358 shows significant increases in phosphorylation demonstrating feedback and activation of potential survival pathways. Highlighted is the upregulation of HER2 and IGF1R phosphorylation in H358 cells in response to gefitinib. Data is mean-centered, log<sub>2</sub> transformed and clustered with a Euclidean distance metric (scale shown below heatmap). b) Level of phosphorylation in Y1197 of EGFR showing the activity of gefitinib 72 hours on cells. Heatmap courtesy of Dan Sears.**



as well, potentially through similar mechanisms, suggesting for other means that KRAS-driven tumors can subvert EGFR inhibition (Figure 3, highlighted in the red box).

In addition to profiling how different cell lines respond to treatment at 72 hours at high concentration of inhibitor, we sought to understand in a specific cell line how varying the concentration with which we treat the cancer cell lines would affect the tyrosine phosphorylation network. Clinically, this may be important due to the fact that some cancers display poor vasculature and drugs may be unable to penetrate deep inside a tumor. While we can understand a potential mechanism of resistance due to high drug concentrations *in vitro*, where drug distribution is not as significant of a concern, there remains the question of just what effect low drug penetrance has on a tumor. [50] With this in mind, we treated the HCC827 cell line with gefitinib at its EC5, EC50, and EC95 concentrations and noted a significant difference in response of the drug (Figure 4). For example, with treatment of EC5 of gefitinib (second column of the heatmap), somewhat surprisingly, boosts in the levels of phosphorylation can be seen on a number of tyrosine kinases, including Met, which is known to mediate long-term resistance to gefitinib in this cell line.



**Figure 4:** HCC827 in response to varying doses of gefitinib 72 hours post-treatment, selecting for sites whose phosphorylation increases at least 1.3-fold at EC5. At EC5, significant increases in phosphorylation are seen, indicating possible mechanisms of adaptation, in particular in the activation loop of the receptor tyrosine kinase MET (highlighted in red), a known mediator of gefitinib-resistance. Data is mean-centered, log 2 transformed and clustered using a Euclidean distance metric. Graph to the right highlights the increase in phosphorylation in the activation loop of Met.

If targeted therapies can have the physiological effect of activation rather than inhibition of pathways at low concentrations due to feedback mechanisms that are still unclear, it stands to reason that parts of a tumor that are only exposed to limiting amounts of kinase inhibitor can be stimulated to grow by activation of survival pathways that mediate survival and eventual resistance. While no known publications to date explore this idea, there is precedence that initial stimulation of alternate survival pathways can mediate long-term resistance. Using the same HCC827 cell line, and with the knowledge that Met amplification leads to eventual resistance to gefitinib, it has been shown that a transient treatment with hepatocyte growth factor (HGF), which stimulates Met, can accelerate Met amplification in cells treated with gefitinib, thereby leading to long-term resistance. [45] While it remains unclear in our

data what is responsible for this boost in signaling at low levels of gefitinib, it may be possible that this response could have clinical implications and warrants further investigation.

### 1.5 Response of NSCLC to EC50 leads to a potential boost in proliferation

To examine if the boost in signaling could lead to increased proliferation, growth experiments using cell counting were performed at a range of drug concentrations from 3 to 100 nM of gefitinib in HCC827 cells and 10 to 1000 nM of PF-1066 in H3122 cells. In order to maintain consistent cell density with MTS and mass spectrometry experiments, 90,000 cells were plated in each well of a 6-well plate in RF-10 media and left overnight for 24 hours to settle and deplete serum. To diminish effects of fresh serum containing growth factors, inhibitors were added without added fresh media/serum, and cells were counted 48 hours later using a Vi-cell instrument as described earlier. While no boost in proliferation could be seen with the HCC827 cells at low concentration of gefitinib, interestingly, H3122 cells treated with PF-1066 exhibited a dramatic boost in proliferation (Figure 5).

While examining this data, a question of timescales that are driving this proliferation event became apparent, in particular, the question of if this is an adaptation-like event versus selection of cells

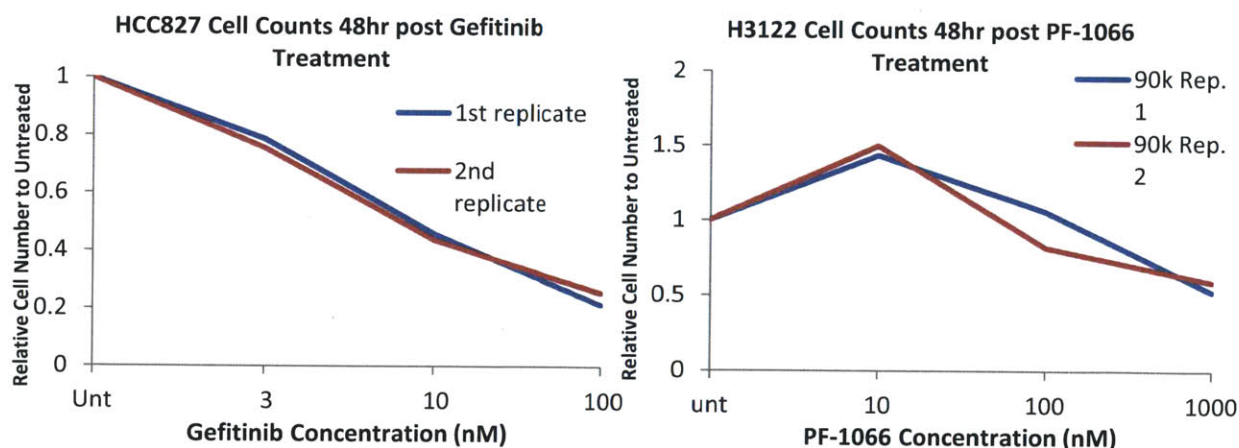


Figure 5: Cell counts relative to untreated 48 hours post treatment with Gefitinib and PF-1066 at the concentrations indicated. While HCC827 exhibited no discernible boost in proliferation, H3122 cells had a dramatic boost of proliferation at 10 nM of PF-1066.

that are insensitive to treatment. Therefore, we sought to better understand how inhibitors affect the

tyrosine phosphorylation network at shorter time scales (less than one hour). An 8-plex mass spectrometry experiment was set up where 1.5 million H3122 cells were plated in a 15 cm dish, allowed to attach overnight for 24 hours, and then 10-1000 nM of PF-1066 were added to the cells; four of the dishes were lysed after 10 minutes and the other four were lysed at 60 minutes of treatment. While phosphorylation of peptides on EML4 decreased as would be expected with treatment of an ALK inhibitor, treatment also caused dramatic increases of phosphorylation in other pathways (Figure 6).

In this dataset, of particular interest is the increase in phosphorylation at Y505 on lymphocyte-



Figure 6: Mass spectrometry analysis of H3122 cells after short-term treatment with PF-1066 reveals both up- and downregulation in a number of different phosphorylation sites. As expected, phosphorylation of EML4 decreases as a function of drug concentration. Data is mean-centered, log 2 based and clustered using a Euclidean distance metric.

specific protein tyrosine kinase as a function of drug concentration up to 1000 nM at the 10 minute time point. Phosphorylation at this site is mediated by C-src tyrosine kinase (CSK), which when it becomes activated phosphorylates a number of SRC-family kinases at inhibitory sites, thereby deactivating them. While CSK has traditionally been thought of as an anti-oncogene due to its inhibitory activities, for reasons that are not yet known, it has been found to be overexpressed in epithelial cancers, even cancers with high SRC activity. [51-54] Mediators of activation on CSK are relatively unknown, however a phosphoproteomic analysis of leukemia driven by another cytosolic kinase fusion, BCR-ABL, found that inhibiting ABL with imatinib (Gleevec) lead to rapid decreases in phosphorylation in SRC-family kinases (SFKs), mediated by CSK, suggesting its involvement may be key to cancer adaptation to inhibition. [55]

## **1.6 Conclusions and looking forward**

The mass spectrometry and phenotypic data thus far indicate that cells are rewiring their signaling networks in response to kinase inhibition, and in particular with the EGFR driven cell lines, quantitative analysis provides a list of potential drug targets for further study. For example, at low concentration (EC5) treatment with gefitinib, a boost in signaling was seen in EGFR, Met, and AXL receptor tyrosine kinases, suggesting potential pathways that cells can respond to loss of driver kinase activity (Figure 4). While the phenotype of this signaling boost remains unclear for EGFR driven cell lines, in the EML4-ALK driven H3122 cell line, treatment with EC5 concentration of PF-1066 leads to a significant (~50%) increase in proliferation 48 hours later compared to the untreated case. The mechanism of this increase remains unclear, but the existence of the phenotype suggests that targeted inhibition of driver kinases with low levels of inhibitor may, perhaps unintuitively, cause increases in proliferation rather than cell death. Clinically, the issue of drug penetrance has gained traction in the research community, and a number of potential mechanisms to alleviating poor bioavailability of chemotherapy are currently being pursued and follow-up work in analyzing cellular response at EC5 concentration may provide explanations of responses seen clinically. [50]

A number of papers suggesting the presence of feedback mechanisms in response to targeted inhibition that can allow for survival and proliferation of cancer cells have arisen, which strongly supports the importance of continued work in the area to uncover either new mechanisms or how mechanisms currently known may play a role in resistance and clinical treatment. [56] For example, in V600E mutant BRAF-driven colorectal cancer (CRC), treatment with a BRAF inhibitor PLX4032 (vemurafenib) has not been as effective as it has been in melanoma containing the same activating BRAF mutation.[57] With this in mind, a study was done to try to examine why this is the case; using an shRNA library to test for what genes may be mediating resistance to BRAF inhibition, researchers noted that knockdown of EGFR lead to synergy with PLX4032 suggesting that EGFR could be playing a role in these CRCs. In order to probe a potential feedback between BRAF and EGFR, in CRC cells containing the activating V600E mutation yet poor sensitivity to PLX4032, they found a strong increase in phosphorylation in EGFR after 6 hours of treatment with 1 uM of inhibitor. From data supported by the shRNA screen, adding an inhibitor to EGFR restored expected sensitivity to BRAF inhibition, leading to antitumor activity *in vivo*.

A number of other similar studies have been published demonstrating the presence of other feedback mechanisms in cancer in response to targeted treatment. [58-62] To date, no known clinical trials are in existence that are designed around targeted a kinase driver as well as a feedback mechanism in parallel; however, as more molecular mechanisms are uncovered, and the importance of these mechanisms are further established *in vivo*, it is conceivable that treatments with this paradigm in mind will begin to appear in the clinic. A potential problem that can arise, of course, is the increased amount of side-effects due to inhibition of the targets present in others tissue that are responsible for normal cellular function, and these effects may amplified as the number of drugs in a given cocktail is increased. Therefore, with this and other reasons in mind, research on delivery mechanisms for better targeting personalized drug cocktails to patient tumors is being aggressively pursued. [63] Further elucidation of signaling networks via mass spectrometry as well as perhaps the combined use of knockdown-libraries for hypothesis generation will likely lead to novel targets of adaptation and eventual resistance.

## **Chapter 2 – Long-term Resistance to Targeted Therapy**

### **2.1 Mechanisms of long-term resistance to Kinase Inhibitors**

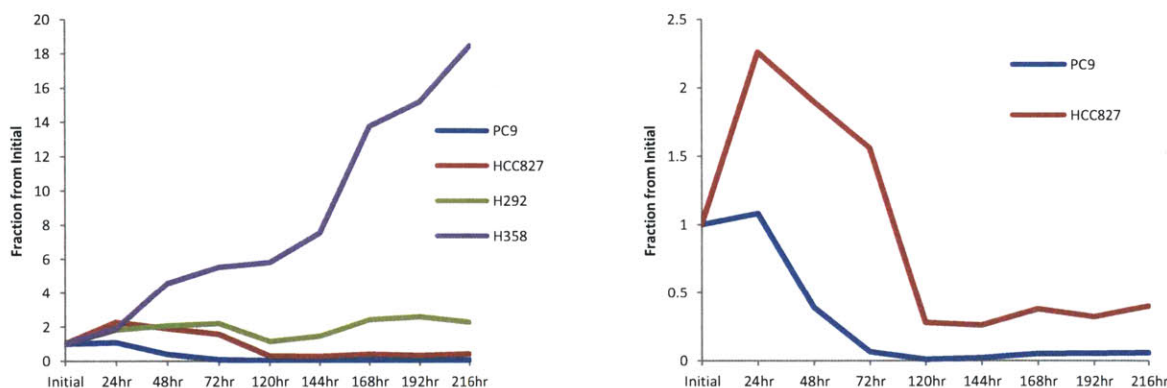
An unfortunate inevitability in most patients treated with kinase inhibitors is the eventual development of tumors that are refractory to treatment. Options that are available to patients once they have developed resistance are few and often times leads to eventual mortality. Therefore, there exists a strong need for the development of either more effective front-line therapies that are designed to prevent known mechanisms of resistance, or alternatively, more second- and even third-line therapies that target the tumors in the presence of known mechanisms of resistance.

A number of mechanisms that tumors develop to evade therapy have been previously uncovered and it is likely a number of other mechanisms have yet to be discovered. As a small example of the types of resistance mechanisms known to exist, tumors have been shown to upregulate alternative survival signaling pathways, prevent apoptosis through expression of antiapoptotic proteins, increase activity/expression of membrane “pumps” to actively remove drugs from cells, and even evolve gatekeeper mutations to block drug target sites themselves. [64] In particular for our studies, we sought to use mass spectrometry-based phosphoproteomics to uncover which alternative kinases can be upregulated in cancer cell lines that can mediate growth signaling in the presence of inhibitors to the known driver kinases.

### **2.2 Development of TKI resistance *in vitro* in NSCLC**

A distinction to the “adaptive response” as referred to in Chapter 1 and a longer-term response is the time frame by which the mechanism of resistance develops. For example, early in the project we sought to understand time scales it takes for cells to develop resistance. As a model system, we again chose the non-small cell lung cancer cell lines driven by the kinases EGFR and EML4-ALK described in Chapter 1. To better understand these timescales, 100,000 (or 1,000,000 in the case with PC9 cell since they are hypersensitive to EGFR inhibition) NSCLC cell lines were plated in full serum, and treated with

1 $\mu$ M of gefitinib with media and inhibitor being replaced every 2-3 days to test how long cell death takes as well as eventual resistance develops, (Figure 7).



**Figure 7:** Growth of EGFR+ NSCLC cell lines in the presence of 1 $\mu$ M Gefitinib, highlighting the timescales of cell death as well as ultimate resistance. Note that H292 and H358 are relatively insensitive to gefitinib and thus have little growth inhibition. The graph to the right highlights the sensitive HCC827 and PC9 cell lines, whose growth begins at ~120-144 hours post treatment.

Cells were trypsinized and resuspended in full RPMI-1640 with 10% FBS (RF-10) and quantified using a Vi-CELL counter every 24 hours. As stated earlier, the H358 cell line’s growth is not inhibited in the present of gefitinib (Figure 7, purple line), due to the presence of the constitutively activating G12V mutation on KRAS making the cells’ growth independent to EGFR activity. [43] However, in the most sensitive HCC827 and PC9 cell lines, growth in the presence of greater than EC95 amounts of gefitinib appears to begin between 120-144 hours, or roughly 5-6 days, suggesting that cancers can evolve within a few cell cycles to inhibition of driver kinases (Figure 7).

Whether this process occurs by adaptation or selection, however, remains unclear. In EGFR-driven cancer NSCLC, it has previously been shown that upregulation of Met kinase can drive proliferation in the presence of gefitinib, and perhaps more interestingly, that a small fraction of NSCLC cells in a gefitinib-sensitive population appear to naturally contain gene amplifications of Met, which are

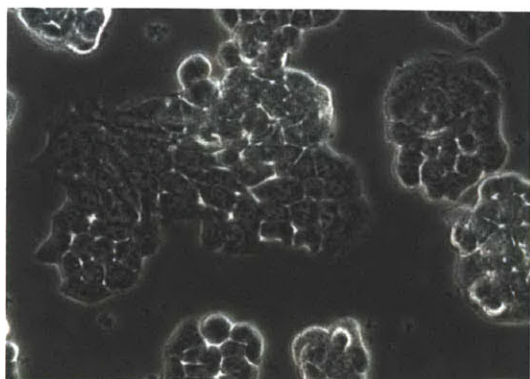


later selected for after treatment with gefitinib. [45, 65] More recently, studies have shown that this genetic heterogeneity can affect clinical outcomes in other types of cancer. For example, in glioblastoma multiforme (GBM), tumors that had a clonal origin and an overexpression of EGFR appear to contain parts which do not express EGFR at all, and in those cells, increased expression of other RTKs like PDGFR and Met was noted. [66, 67] The presence of these other receptor tyrosine kinases appears to be a stable co-existence, and cell lines derived from such tumors require inhibition of each of the overexpressed RTKs to suppress growth. While these studies are suggestive of selective processes being significant for cancer resistance, another study showed that “crosstalk”, or co-activation of Met by an EGFR variant, EGFRvIII, was found to be significant for cancer growth. Co-treatment with gefitinib and a Met inhibitor had a synergistic effect, supporting the idea of networks being adaptable. [68, 69] Further supporting this notion of targeting multiple relevant kinases in cancer, dasatinib (Sprycel) has been quite successful clinically, hypothesized due to its ability to inhibit over 100 kinases, including a number of known kinases that can mediate cancer growth, like ABL and SRC-family kinases. [70-72] Therefore, while the question of adaptation and selection remains unclear and will require further experimental work to track cells and signaling events with more granularity over the course of treatment, these studies together strongly support the paradigm of clinically targeting multiple kinases together that can either be compensatory or present in a stable, co-existing population in cancer.

## **2.2 EGFR can compensate for loss of EML4-ALK Signaling**

In order to better understand the development of long-term resistance to kinase inhibition in NSCLC, both EGFR and EML4-ALK positive cell lines were made resistant by culturing the cells in the presence of EC50 concentrations of gefitinib and PF-1066, respectively. While an initial period of cell death was noted, over time cells began to proliferate despite the presence of drugs targeting the driver kinase. The EML4-ALK+ H3122 cells displayed a dramatic change in morphology as it developed resistance to PF-1066 (Figure 8 for images). An MTS assay per standard protocol confirmed increased resistance to PF-1066 (Figure 9, orange line). To identify kinase drivers that could be responsible for

H3122 (PF-1066 sensitive)

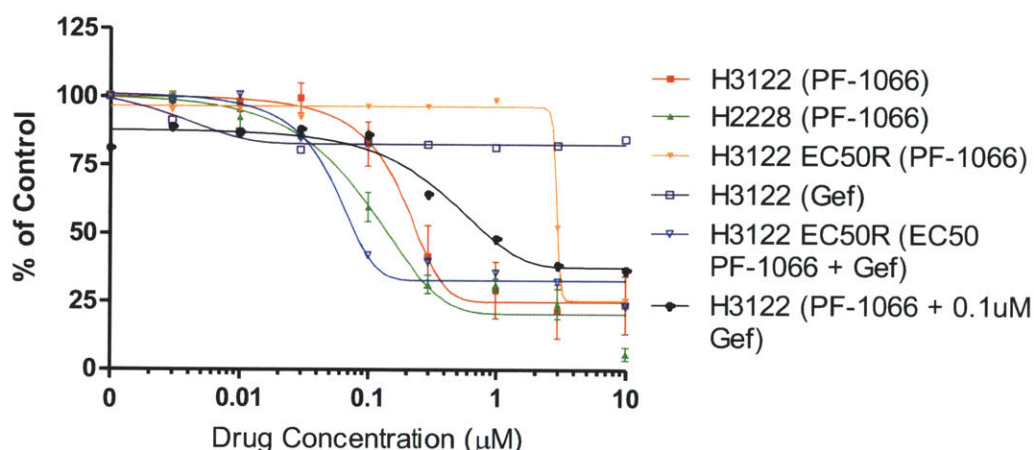


H3122 (PF-1066 resistant)



**Figure 8: Light microscopy images of live H3122 cells that are sensitive and resistant to ALK inhibitor PF-1066. Of note is in the resistant cell line a "flattening out" of the cell-cell contacts causing a dramatic change in morphology. 40x magnification.**

resistance and development of an epithelial morphology, an 8-plex phosphoproteomics experiment containing four channels of sensitive and four channels of resistant H3122 cells was set up. 1.5 million

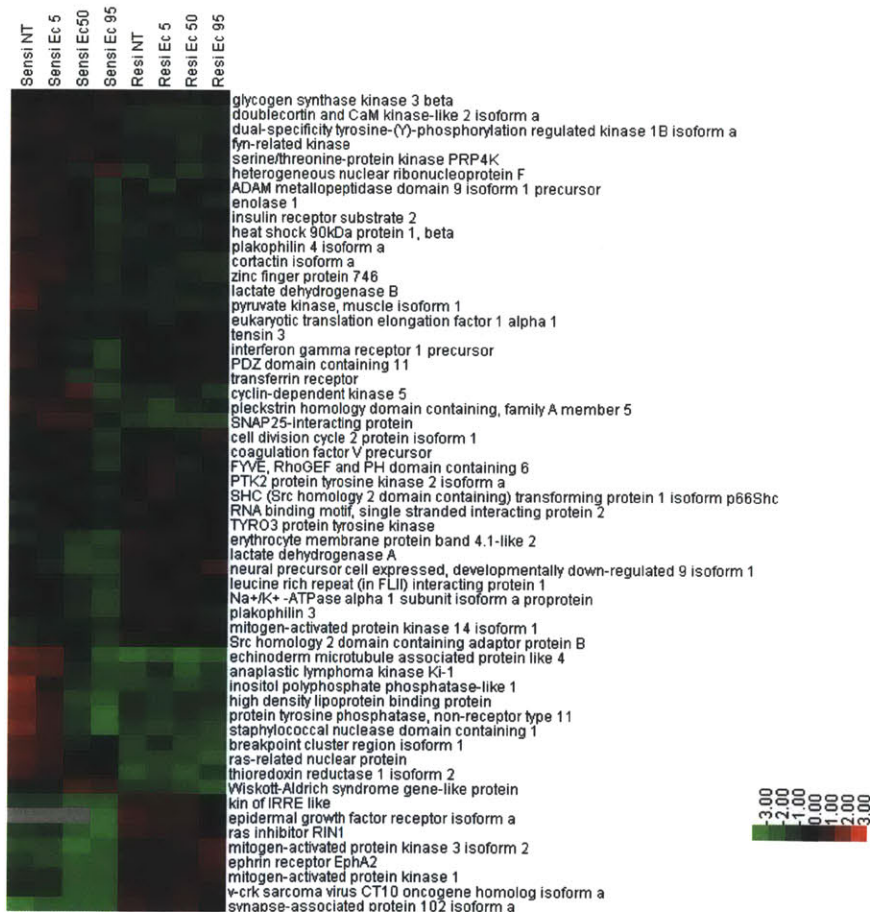


**Figure 9: Sensitivity of H3122 cells to various TKI treatment. While sensitive H3122 cells show no sensitivity to gefitinib, once the cells become resistant to PF-1066, they demonstrate an increase in sensitivity, indicative of EGFR becoming the driver kinase in these cells. Interestingly, treatment of sensitive cells with PF-1066 and gefitinib leads to no synergy, suggestion that reliance on EGFR is a longterm response in these cells.**

cells were plated in full RF-10 media and allowed to settle for 24 hours as in previous experiments, and PF-1066 was added at EC5, EC50, and EC95 concentrations for 72 hours, followed by lysis, phosphopeptide enrichment, and mass spectrometry.

While treatment with PF-1066 demonstrated a decrease in phosphorylation of ALK and EML4 as a function of drug concentration, in resistant cells both untreated and treated with PF-1066, EML4 and

ALK showed decreased phosphorylation while the phosphorylation of EGFR increased dramatically (Figure 10). Based on this data, it was hypothesized that EGFR could be driving growth in these resistant cells. In order to confirm the importance of EGFR in these cells, they were treated with a dose-response of gefitinib and showed a dramatic increase in gefitinib sensitivity (Figure 9, blue line). Interestingly, co-treatment of sensitive cells with both PF-1066 and gefitinib lead to no synergy (black line, Figure 9), suggestive that reliance on EGFR is a process that takes longer than 72 hours. Concurrently to this work, was been published that not only is EGFR a mechanism of resistance to ALK inhibition, but also that high expression of EGFR in EML4-ALK positive cells is a marker for decreased initial sensitivity to ALK-inhibitors, presumably due to its ability to compensate for loss of ALK signaling. [73]



**Figure 10: A phosphoproteomic comparison of sensitive and PF-1066 resistant H3122 cells, highlighting a decrease in EML4-ALK phosphorylation and a dramatic increase in EGFR phosphorylation. Data is log2 transformed, mean-centered, and clustered using a Euclidean metric. Heatmap courtesy of Dan Sears.**

## 2.3 Conclusions and looking forward

In the course of performing this work it became clear that there is a strong need for the development of both diagnostic and prognostic tools for both 1) uncovering drivers and 2) predicting future drivers in response to TKI therapy, and in particular also highlights the utility of mass spectrometry to identify driver kinases that are upregulated as they evolve resistance. Interestingly, work with evolved resistance in H3122 to ALK-inhibition suggests that the short-term response that allows for survival of cells may not match the long-term response, pointing to the need to understand how resistance develops with more time granularity. In addition to studying signaling dynamics at more time points, applying phosphoproteomic studies to more cell lines with different drivers could potentially uncover novel mechanisms that would aid in clinical treatment of cancer. Ultimately, in case factors of 2D culture affect development of resistance, presumably due to altered cell-cell contact signaling, expansion of studies to *in vivo* xenograft models of resistant tumors as well use of 3D cell culture techniques may prove to be more realistic in rationalizing phenotypes seen clinically in patients. Nonetheless, this work points to a direction of mass spectrometry studies performed as described as a tool for uncovering other mechanisms of altered kinase regulation in cancers that evolve resistance to TKI treatment.

## **Chapter 3 – Using 14-3-3 to Interrogate Phosphoserine and Phosphothreonine**

### **Signaling**

#### **3.1 Phosphoserine and phosphothreonine Signaling – the bigger picture**

While significant insight can be derived from studying phosphotyrosine networks, the eventual phenotype of a cell may be due to downstream signaling networks that are predominantly phosphoserine and phosphothreonine mediated. For example, Akt, an important kinase involved in proliferation, transcription, and migration, phosphorylates serine and threonine residues and acts as a conduit between tyrosine networks and downstream serine-threonine pathways. [74, 75] While a phosphorylated tyrosine residue provides a large enough epitope for efficient antibody recognition and pull down as required for our mass spectrometry studies, phosphorylated serine and threonine residues are not as bulky and antibody generation to these epitopes remains challenging. Further complicating matters, serine and threonine phosphorylation is extremely abundant inside the cell, and selectively enriching for sites that are physiologically important remains largely impossible. With this in mind, we sought to develop a means of enriching for phosphorylated serine and threonine residues that are meaningful within the context of combining phosphotyrosine data with phenotypic data to offer a potentially more global view of cellular signaling. Given the complexity of cellular networks in terms of feedback and feedforward loops, it is our hypothesis that functional characterization of phosphoserine and phosphothreonine networks will provide significantly better understanding of how perturbations in signaling networks can affect cellular outcomes and may open up potential opportunities for drug targeting down the line.

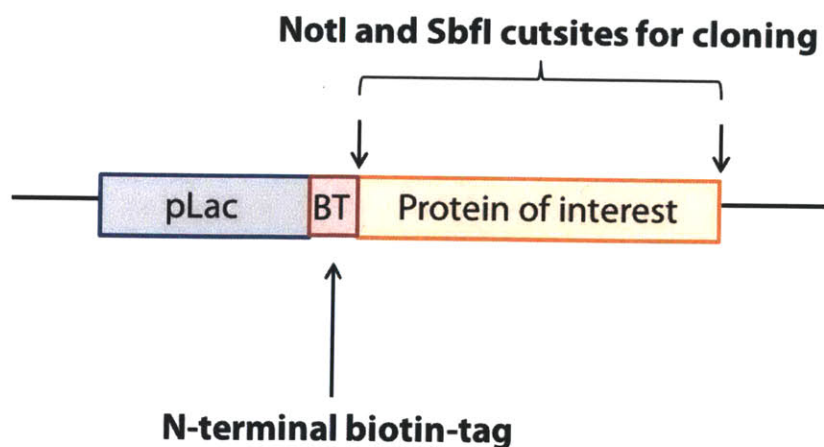
#### **3.2 Use of 14-3-3 to interrogate phosphoserine and phosphothreonine signaling**

In order to isolate serine/threonine phosphorylation sites that are of importance for the overall signaling network, we used full length 14-3-3 protein to pull down phosphorylated peptides. Via recognition of specific serine/threonine phosphorylated sites within the Mode 1 (RXX(pS/pT)XP) and Mode 2 (LX(R/K)SX(pS/pT)XP) motif, 14-3-3 functions as an important scaffolding protein to help

regulate signal transduction pathways involved in a variety of cellular functions like apoptosis, DNA-damage response, growth, and proliferation. [76-78] Of importance to our goal, 14-3-3 isoforms bind to a significant number of proteins in the cell, thereby maximizing the potential number of sites we can study via peptide pull down and mass spectrometry. For example, a protein array study identified 1,752 unique proteins that are pulled down by the 14-3-3 epsilon isoform. [79] In addition, binding of 14-3-3 proteins has been shown to be phosphorylation-specific with a decrease in  $K_d$  by at least 100-fold between a phosphorylated and unphosphorylated form of a given peptides containing a 14-3-3 motif, suggesting a 14-3-3 peptide pull down approach could work to isolate phospho-specific peptides. [80] With these facts in mind, we sought to perform mass spectrometry experiments as described in Chapter 1, however replacing the antibody pulldown step with a 14-3-3 pulldown step instead to isolate phosphorylated peptides containing Mode 1 and Mode 2 binding motifs.

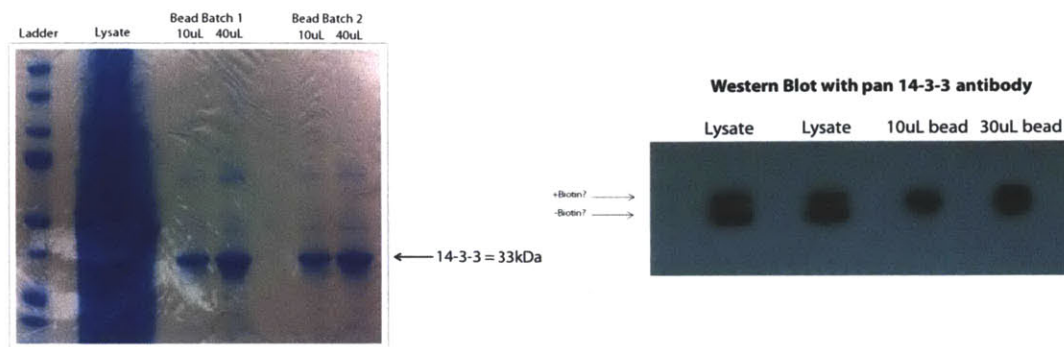
### 3.3 14-3-3 pulldown of phosphorylated peptides

An *E. coli* expression vector was constructed using a modified “Avi-tag” biotin ligase system that was a gift from the Niles lab at MIT. The birA biotin ligase adds a biotin group to the lysine residue of the recognition sequence, MASSLRQILDSQKMEWRSNAGGS, in order to provide an attachment site to the streptavidin protein, and to biotin-tag a protein of interest, this recognition sequence is added N-



**Figure 11:** Schematic of the *E. coli* expression vector for adding a single biotin group N-terminal to a protein of interest (see text for details)

terminal to a gene and inserted into a standard *E. coli* expression vector. [81] Eight-base NotI and SbfI restriction enzyme sites were added after the biotin-ligase recognition sequence and immediately after the stop codon for ease of shuffling 14-3-3 isoforms or any other protein of interest (see Figure 11 for the schematic as well as Supplemental Information for vector maps). A birA-expression vector pACYC177\_birA (KanR) and a 14-3-3 expression vector pET16Trx\_NB-14-3-3 (AmpR) were co-transformed into *E. coli* BL21(DE3) cells and selected using half antibiotic concentrations (50 ug/mL ampicillin, 25 ug/mL kanamycin), as the expression plasmids are low-copy number and growth can be significantly inhibited at higher antibiotic concentrations. Colonies were picked and grown for 8 hours in 2 mL of LB supplemented with antibiotics, after which 500 uL of culture were added to 250 mL of LB containing antibiotics as well as 0.1 mM IPTG and 1X biotin and grown for 24 hours at 25°C. Cells were harvested by spinning at 4,000g for 10 minutes at 4°C, and resuspended in a 25 mL 1x TBS buffer containing 1x protease inhibitor cocktail, 1 kunitz/mL DNase, 1 mg/mL lysozyme, and 0.1% NP-40. Under constant stirring, the suspension was incubated at room temperature for 2-3 hours to lyse cells; lysis was further enhanced by several rounds of sonication at 4°C. The lysate was cleared of cell debris by spinning at 12,000g for 30 minutes at 4°C, and the supernatant was applied to 1 mL of packed streptavidin-agarose beads (EMD Millipore, Billerica, MA, USA). Lysate was mixed with beads for 2 hours at 4°C, followed by 4 washes with 1x TBS. To confirm isolation of protein, 10 and 40 uL of slurry was run on a SDS-polyacrylamide gel, followed by transfer and a Western Blot per standard protocols against a pan-14-3-3 antibody (Santa Cruz Biotechnology, Santa Cruz, CA, USA) at a 1:10000 dilution (Figure 12).



**Figure 12:** Coomassie stain and Western Blot of 14-3-3 lysate and streptavidin beads confirming expression and purification. The two gels were from different purifications of the same bacterial stock. The two bands on lysate wells of the Western Blot are potentially from the presence of both biotinylated and unbiotinylated 14-3-3 protein.

The mass spectrometry protocol was not altered, except peptides were pulled down using biotin-tagged 14-3-3. Peptides were eluted using 100 mM glycine pH 2 and IMAC was performed to further enrich for phosphopeptides. A sample list of peptides seen in a run of 2 mg of unlabeled peptide from 30 minute serum-stimulated HCC827 cells is shown in the Supplementary Information.

### 3.4 14-3-3 Pulldown isolates signaling protein and allows for studying signaling dynamics

Through 14-3-3 pulldowns, peptides were isolated containing Mode 1 and Mode 2 motifs that correspond to known 14-3-3 binders. In addition to unlabeled peptides, iTRAQ-labeled peptides can be pulled down using this technique to identify dynamics in response to perturbation. In HGF stimulated H1975 cells, expected dynamics in a number of signaling proteins can be seen (Figure 13).

	H1975	+	+	+	+
	HGF (10 ng/ml)		+		
	HGF (50 ng/ml)			+	
	HGF (500 ng/ml)				+
Protein	Sequence	114	115	116	117
aldehyde dehydrogenase 3 family, member A1	K.KMIAETSSGGVAANDVIVHITLHSLPFGGVGNSGMGSYHGK.K + 2 Phospho	1.0	1.1	1.6	1.3
ATPase, class VI, type 11B	R.SWSASDPFYTNDR.S + Phospho (ST)	1.0	1.7	2.0	1.9
BCL2-antagonist of cell death protein	R.SRSAPPNLWAAQR.Y + Phospho (ST)	1.0	2.6	2.1	3.0
forkhead box O3A	R.SCTWPLQRPELQASPAKPSGETAADMIPPEEEDDEDEDGGGR.A + Phospho	1.0	1.9	1.5	1.8
forkhead box O4	R.SCTWPLRPEIANQPSEPPEVEPDLGEK.V + Phospho (ST)	1.0	1.5	1.3	1.6
GTPase activating protein and VPS9 domains 1	R.RPMSDPSWNR.R + Phospho (ST)	1.0	1.1	1.5	1.3
histone deacetylase 5 isoform 1	R.TQSSPLPQSPQALQQLVMQQHQHFLEK.Q + Phospho (ST)	1.0	1.1	0.9	1.0
Ia related protein isoform 1	K.GLSASLPDLSENWIEVK.K + Phospho (ST)	1.0	1.2	1.4	1.6
slingshot 2	R.TTNPFYNTM.- + Phospho (ST)	1.0	0.9	0.7	0.6
spectrin, alpha, non-erythrocytic 1 (alpha-fodrin)	R.WRSLQQLAEER.S + Phospho (ST)	1.0	1.1	1.1	1.1
v-raf murine sarcoma viral oncogene homolog B1	R.DRSSAPNVHINTIEPVNIDDLIR.D + Phospho (ST)	1.0	0.8	1.6	1.1
WW domain containing transcription regulator 1	R.SHSSPASLQLGTGAGAAGSPAQQHAHLR.Q + 2 Phospho (ST)	1.0	1.2	1.3	0.8

**Figure 13:** Dynamics from HGF-stimulated H1975 cells showing expected patterns in phosphorylation sites with known functions. Data courtesy of Dan Sears



For example, phosphorylation of the pro-apoptotic BCL2-antagonist of cell death (BAD) protein at serine-99 as seen in Figure 13 has been shown to be pro-survival, which would be consistent with treatment of a growth factor and subsequent activation of pro-survival kinases like RAF. [82, 83] In addition to the examples listed above, other phosphorylation sites noted on important kinases in signaling pathways observed by this method include ARAF, BRAF, CRAF, and the key regulatory Serine-473 site on AKT. However, a notable shortcoming of the method thus far has been the significant contamination of acidic, non-14-3-3 specific peptides. Work is continuing to improve binding, washing, and elution of specific peptides to improve coverage of the phosphoproteome that is bound by 14-3-3.

### **3.5 Conclusions and future directions**

While this technique is still in development, in particular to decrease the amount of acidic contaminating peptides, we have been able to isolate known 14-3-3 binding pairs at specific key regulatory sites. Future work to improve the coverage will better be able to elucidate how perturbations in phosphotyrosine signaling feeds forward, as well as how perturbation of phosphoserine and phosphothreonine signaling can feed back upstream. In particular to our work in the area of cancer adaptation and resistance, a well-established 14-3-3 pull down protocol would allow the interrogation of known driver kinases like BRAF and Akt, as well as identification of potential key nodes that may not be apparent from the tyrosine signaling network alone. For example, a number of signaling events related to regulation of apoptosis, like sites on BAD and Caspase 9, are serine and threonine phosphorylation events, and given the altered regulation of apoptosis in cancer, information about these dynamics may be useful toward deriving a better understanding of tumor biology. Therefore, in the absence of an antibody that is able to effectively pull down peptides relevant to 14-3-3-mediated signaling, improvements of this technique could prove to be valuable in providing potentially clinically relevant information from tumor cells.

## **Chapter 4 – Computational Tools for Analyzing Mass Spectrometry Data**

### **4.1 Computational tools to process peptide data from a mass spectrometer**

As the sensitivity and speed of electronics and sensors inside mass spectrometers improves, more and more proteomic data will be generated. Manual processing of data will likely eventually prove to be an impossible task, yet given the amount of hypothesis one can derive from high-throughput proteomic data, there is a strong need for tools that can reliably extract meaningful information that is useful to help scientists guide experimental design. An ideal computational tool will save time for researchers without sacrificing quality or integrity of data, as a major concern in proteomics is the quality of data that is being deposited and used as rationale for project development. [84] Manual validation of spectra data is currently the best means of ensuring that sequencing and quantitative information from database searching software like MASCOT is correct. Nonetheless, scripts that automate steps currently done by hand that require limited to no judgment, like data formatting, routine calculation, and basic cleanup, can remove cumbersome steps that can take on the order of hours to get data from common mass spectrometry formats (RAW, MGF, mzXML, etc) into data processing software like MATLAB and Excel.

### **4.2 Generation of scripts for the automation of routine steps in data formatting and cleanup**

Scripts were written to automate a number of the manual calculation and formatting steps that are done after the creation of HTML files with peptide information from the MASCOT server. The goal of these scripts is to streamline and ideally remove some of the manual steps required to generate Excel files with easily readable and usable data from raw mass spectrometry files. In order to improve accessibility to those with minimal computation expertise, a graphical user interface (GUI) was written to help with option selection, as well as providing an intuitive means for a user to customize how scripts are executed through passing in parameters (Figure 14). In addition, the MATLAB software environment allows a number of complicated data operations and contains a rich library of both built-in and user community-

written functions for downstream data analysis. Lastly, a long-term goal that was in mind during development would be for the software to be as extensible as possible for running as many of the routine computational analyses that are performed in a proteomic/mass spectrometry lab, including the ability to easily add new or update existing functionalities as desired.

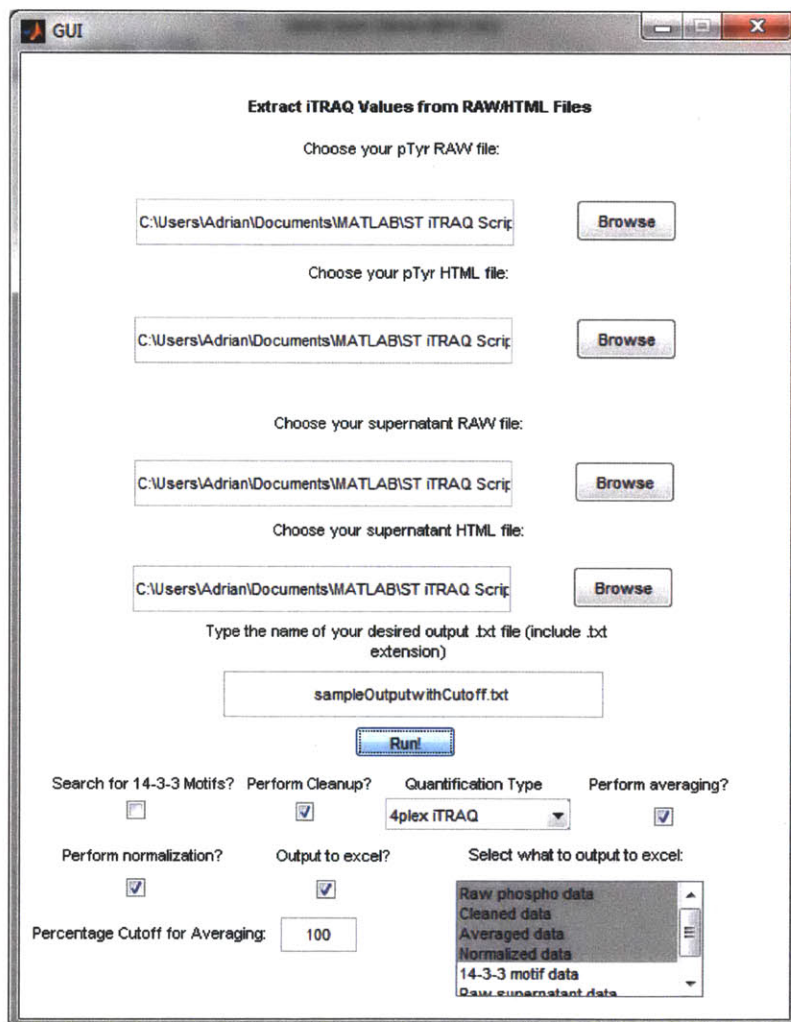


Figure 14: A sample image of the GUI highlighting the number of options for data cleanup and processing

#### 4.3 Design of scripts to automate cleanup, averaging, and normalization of mass spectrometry data

Previous scripts written in the lab to process RAW files from a Thermo Fisher Orbitrap-LTQ and MASCOT were used as a starting template to build a GUI and additional scripts for added function. At a high level, the scripts work sequentially to process both data from a phosphotyrosine as well as a supernatant run in parallel (for loading control) as desired. The scripts as published here are able to work to clean up quantification data, average peptide values, normalize to a supernatant run of non-phosphorylation peptides, and lastly search for and mark peptides containing 14-3-3 motifs. Each algorithm for the listed functionalities is described at a high level in the following subsections.

### 4.3.1 Scripts for cleanup of peptide and iTRAQ information

Prior to analysis with Excel, MATLAB, or comparable software, data from a mass spectrometry run must be formatted into a MATLAB-friendly format. Using the `mzxmlread` function found in the Bioinformatics Toolbox, a structure of spectra scans as well as a list of peaks can be generated for extracting iTRAQ data as well as peptide M/Z information for validating sequencing and the presence of contaminating peaks. This information can be combined with sequencing information from MASCOT to generate a list of peptides and as their corresponding iTRAQ values. A sample output of the data at this point is illustrated below in Figure 15.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
Scan	Protein	Sequence	GI	Elution	MZ	Charge	Score	113	114	115	116	117	118	119	121	
2	27415	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (ST)	8.9E+07	109.833	1411.16	2	91	0	165.5	283.61	164.332	154.599	307.356	0	393.06
3	29210	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	116.065	941.111	3	69	0	0	0	0	0	0	0	0
4	29211	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	116.067	941.111	3	51	10349	18104.2	18172	20357.5	21571	11867.6	16338.2	18710.7
5	24301	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (ST)	8.9E+07	98.482	941.112	3	39	377.043	574.988	583.438	508.21	817.617	368.89	752.367	280.146
6	27825	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	111.273	706.086	4	18	5231.97	2856.4	7652.79	5090.89	6658	7088.88	7132.37	2859.74
7	28371	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	113.16	706.086	4	32	0	8383.05	6449.79	0	0	8696.3	3901.6	0
8	27952	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (ST)	8.9E+07	111.712	1411.17	2	21	0	0	0	0	0	0	0	0
9	27953	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (ST)	8.9E+07	111.715	1411.17	2	82	719.269	1029.32	1246.97	977.096	1473.84	1045.82	1145.79	1017.9
10	28118	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	112.285	941.113	3	58	0	0	0	0	0	0	0	0
11	28119	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	112.287	941.113	3	59	96932.9	111659	131658	143364	161679	134715	101754	147158
12	28664	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	114.173	941.113	3	56	0	0	0	0	0	0	0	0
13	28665	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	114.175	941.113	3	67	28201.8	37438.7	46356.2	39832.5	60265.5	44436	36668.6	36031.3
14	26480	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	106.492	941.113	3	59	0	0	0	0	0	0	0	0
15	26481	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	106.493	941.113	3	42	1983.61	2655.68	3480.73	3146.81	3358.29	3013.14	2573.86	3878.15
16	27572	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	110.387	941.113	3	89	0	0	0	0	0	0	0	0
17	27573	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	110.388	941.113	3	57	58168.6	68284	96648.9	86324.5	92694.1	92515.9	60418.8	87936.3
18	27026	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	108.462	941.113	3	52	0	0	0	0	0	0	0	0
19	27027	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	108.463	941.113	3	66	6784.29	11714.1	13845.5	16646.9	21299	15510.2	10180.1	19155.8
20	25937	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	104.528	941.114	3	37	696.35	1058.09	1060.32	850.734	1061.35	1158.37	946.882	1005.03
21	25936	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (ST)	8.9E+07	104.525	941.114	3	20	0	0	0	0	0	0	0	0
22	24849	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	100.517	941.114	3	42	330.124	540.135	712.152	440.684	771.314	460.204	389.757	671.471
23	23765	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	96.47	941.115	3	30	379.577	485.637	226.172	401.354	469.533	476.175	398.369	487.661
24	28526	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (ST)	8.9E+07	113.697	1411.17	2	105	258.541	233.919	319.712	306.838	360.191	401.558	0	355.025
25	44114	lymphocyte-specific protein tyrosine kinase	R.SVLEDFFTATEGYQPQP.- + Phospho (Y)	1.1E+08	164.758	1221.07	2	36	0	0	0	0	0	0	0	0
26	43469	lymphocyte-specific protein tyrosine kinase	R.SVLEDFFTATEGYQPQP.- + Phospho (ST)	1.1E+08	162.363	814.38	3	36	0	0	0	0	0	0	0	0

Figure 15: Sample data output from processing with the `mzxmlread` MATLAB function combined with in-house scripts to reference and combine peptide sequence information inside an HTML output file from MASCOT

From visual inspection of Figure 15, the data contains a number of potential errors that need to be examined and cleaned. For example, in the columns containing iTRAQ data (columns I to P), a number of zeroes are present indicating the HCD on the instrument could not detect an iTRAQ data for a particular scan. In addition, in column H, MASCOT outputs a MudPIT score based on estimating the likelihood that multiple matches to a protein in the database is random in order to estimate the confidence of a given sequence. As a rule of thumb (with exceptions), a peptide with a score of 25 is likely to be real; shorter peptide reads can still be real with a score less than 25, however, care must be taken to manually validate the hit and determine if the sequence read is real. Reading down column C, there are multiple scans for a given peptide, each with their own iTRAQ data, as well as peptides not containing phosphorylated tyrosine.

From this raw data output shown and described above, a MATLAB structure containing the text and numerical is passed into a cleanupTXT.m function which will perform six operations:

- Delete a row of data if there is no phosphorylation group in the peptide
- Delete a row of data if there is no tyrosine amino acid in the peptide
- Delete a row of data if there are any zeroes in any iTRAQ channel. While very low (<1000) values in a iTRAQ channel may not be real, this must be determined by the experimenter via manual inspection of a given scan
- Delete a row of data if the MudPIT score is below 20
- Delete a row if MASCOT outputs more than 3 phosphorylation sites in the peptide
- After cleaning up scans based on the preceding rules, correct iTRAQ values for contamination from other channels due to isotopic impurities

After these operations are completed, this data will be output, and the sample from above is shown below in Figure 16.

Scan	Protein	Sequence	GI	Elution	MZ	Charge	Score	I113	I114	I115	I116	I117	I118	I119	I121	
2	29211	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	116.067	941.111	3	51	10956.3	18297	17585.8	19996.2	21686.6	11024.7	15676.3	20166.5
3	24501	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (ST)	8.9E+07	98.482	941.112	3	39	400.075	577.027	569.15	480.076	838.345	321.178	775.973	296.853
4	27953	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (ST)	8.9E+07	111.715	1411.117	2	82	763.97	1025.06	1234.98	920.398	1487.66	1003.04	1120.2	1094.59
5	28119	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	112.287	941.113	3	59	103245	109792	128916	140219	160434	133168	98678.6	158886
6	28665	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	114.175	941.113	3	67	29985.8	37102.7	45828.3	37761	60722.1	43462.6	35321.3	38786.5
7	26481	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	106.493	941.113	3	42	2108.89	2629.14	3452.78	3052.97	3314.83	2968.9	2364.16	4188.22
8	27573	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	110.388	941.113	3	57	61945.9	66877.4	96292.9	83662.8	90993.7	92643.3	55301.6	94951
9	27027	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	108.463	941.113	3	66	7184.48	11789.5	13476.2	16251.9	21347.6	15347.7	8947.78	20713.1
10	25937	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	104.528	941.114	3	37	738.926	1061.83	1038.36	812.938	1043.08	1150.53	905.049	1082.64
11	24849	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	100.517	941.114	3	42	349.931	540.371	712.506	402.148	786.168	445.229	350.808	725.709
12	23795	serine/threonine-protein kinase PRP4K	K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y)	8.9E+07	96.47	941.115	3	30	403.708	488.246	195.917	400.282	461.967	471.44	375.413	525.916
18	48470	lumiperoxidase-specific protein tyrosine kinase 2	K.LCDFGATGCVADNDITPYLVS.R.F + Phospho (Y)	1.1E+08	162.567	814.48	3	50	1147.47	1764.17	2699.48	5809.14	1903.71	1802.97	1247.61	2122.60

Figure 16: Sample output of the cleanupTXT.m script as described above

### 4.3.2 Script for averaging iTRAQ data from multiple scans

From Figure 16 above, for the given tryptic peptide K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y), a number of scans contain different iTRAQ information. In order to get an idea of the trend in quantification of the data, averaging can be performed of the iTRAQ data. However, straight averaging may not be appropriate if iTRAQ information is showing different trends for a given peptide; this may be due to contamination from other peptides if the chromatography is poor, and thus must be accounted for. Therefore, as a user input, a percent threshold option is added which, if inputted, will look at the trends in the data and determine if any of the dynamics for a given peptide exceeds the user-defined threshold. For example, with a 100% threshold option added, data for K.LCDFGSASHVADNDITPYLVS.R.F + Phospho (Y) is not averaged, as the software will detect that dynamics are not consistent and require manual inspection. Depending on the user's tolerance, this number can be increased or decreased as

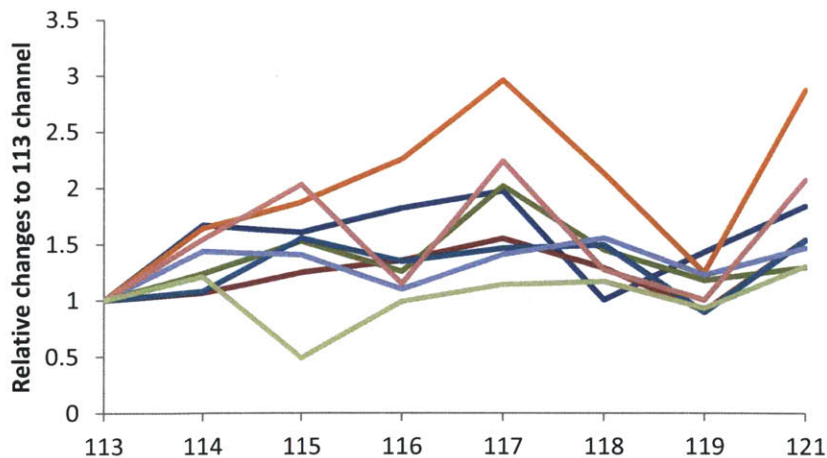


Figure 17: Sample of iTRAQ data of 9 scans that are not averaged with using a 100% threshold.

needed. If the script finds that a given peptide has dynamics that surpass the threshold, it will output an asterisk in an added column to either the tab-delimited text file or Excel file to alert the researcher that the dynamics in question for a given peptide need to be manually inspected for contamination from other peptides or some other problem.

#### ***4.3.3 Script for normalizing iTRAQ data from non-phosphorylated supernatant peptides***

In order to obtain more accurate quantification, the iTRAQ data must be normalized to a supernatant of non-phosphorylated peptides to account for discrepancies in peptide loading. Supernatant data is converted to a tab-delimited text file as described earlier in Chapter 4.3.1; however, the cleanup step only entails deleting rows if a given iTRAQ channel is below 1000 as well as any peptides with a MudPIT score below 25. The iTRAQ data of this cleaned set is corrected for contamination from other iTRAQ channels, and then the mean iTRAQ value for a given channel is calculated and used to create an estimated ratio for the loading control. Other methods exist for this normalization step, and a user can choose to edit the script as desired. The ratio for loading control is applied to each channel in the cleaned dataset can be output to both a tab-delimited text file as well as an Excel file as desired.

#### ***4.3.4 Script for finding 14-3-3 motifs in tryptic peptides***

To aid in studies using 14-3-3 protein to pull-down phosphoserine and phosphothreonine peptides, a motif searcher was added to highlight peptides that match the expected 14-3-3 protein is known to have two binding modes: Mode 1, RXX(pS/pT)XP, and Mode 2, LX(R/K)SX(pS/pT)XP, where X stands for any amino acid. [85] The script searches peptide by peptide and if it finds a sequence that matches either motif, it will mark the line with an asterisk that is stored in an added column.

#### **4.4 Conclusions and looking forward**

In the preceding chapter, a set of scripts for cleaning and formatting data from raw mass spectrometry files was presented that can save hours of manual work. The scripts are not meant to replace manual validation; however, rather they aim to save time by removing repetitive manual calculations, as well as alerting a researcher when dynamics in a peptide do not match across scans. Eventually as the size of data files increases and manual validation becomes increasingly difficult, more complex computational tools will need to be written for both processing and validating proteomic data, for example, matching a peak list with expected fragmentation patterns to determine if contaminating spectra are present. In the meantime, the scripts as presented have been designed and written with extensibility in mind such that new functionalities could be added at will to automate other parts of high-throughput proteomic data analysis.



## Conclusions and Looking Forward

In this thesis I've discussed projects aimed at exploring three different aspects of cancer biology from a signaling perspective: 1) how signaling networks adapt and remodel themselves to drug-induced perturbation, 2) how cancer cells can evolve long-term mechanisms of resistance to targeted therapies, and 3) the development of a novel technique to better probe signaling networks downstream of the phosphotyrosine network that may play significant roles in signal transduction.

Life is an especially robust process and signaling pathways appear to be selected for with redundancy in mind. [86, 87] Currently we seek to inhibit what we view of as main drivers of cancer growth, as well as potentially other kinases that can mediate survival of a cancer cell population long-term (i.e. Met compensating for EGFR-inhibition, or EGFR compensating for EML4-ALK inhibition). Furthermore, the heterogeneity of a tumor may be great enough that the number of inhibitors needed to eradicate a cancer's growth will be too high from a side effects standpoint. Perhaps with the further development of nanoparticles and other means of targeted drug delivery, these effects will be mitigated, but it remains unclear what their potential is in the future.

Looking forward, I envision the further development of more complex, perhaps even "computable" therapies that rather than seeking to inhibit cancer, they stimulate means for cells to apoptose in a cancer-specific manner. In this context, I view the complexity of signaling circuitry almost as a "waterfall" of growth signals, where cells can dynamically rewire themselves to activate different pathways that can mediate their survival. In the same sense as building a dam to block water, adding an inhibitor to a cancerous cell may just lead to an alternate "path" being explored. The effects of the drug may just be cytostatic rather than cytotoxic, so perhaps looking more in the direction of conditional therapies that can activate apoptosis inside of a target cell of interest may prove to be more effective clinically. [88] This idea has been explored by many and numerous studies have been published in the area of computable therapies for cancer. [89-95] Cancer has a number of immutable targets associated with it at both the genomic and proteomic level that are sufficient at differentiating it from other cells in

the body. For example, it has been estimated that a tumor can contain on the order of 20,000 unique mutations in more than 130 coding exons, and besides a myriad of genetic mutations within a cell, there will also exist another myriad of mutations within a context of a tumor (i.e. mutations unique to different sub-clones of a tumor). [96] While the immune system in a normal context should be able to target cells that have developed mutations in coding regions through MHC-peptide recognition by T-cells, cancers evolve to adapt in many clever ways, such as secreting cytokines to suppress the immune system, or altering their apoptotic pathways to become resistant to secreted proteins of the immune system. [97] With this in mind, given the sheer amount of defects present, it seems eradicating cancer may potentially be an insurmountable problem from an inhibitory perspective, but from an activation perspective, it may prove to be a more reasonable direction for therapies in the future as single markers can be sufficient to delineate a cancerous cell from a non-cancerous cell.

Until these more complex treatments can be administered clinically, there is a significant need to discover novel “druggable” pathways, as well as develop tools to best predict *a priori* patient sensitivities to available treatments. From a practical, economic perspective, with new drug costing on the order of \$100,000 per year, identifying patients likely to benefit versus those unlikely to benefit would be a means of decreasing unnecessary financial and treatment burden both to the patient as well as society. The benefit that patients derive from drugs that are currently approved for targeted treatment of cancers can vary greatly; some even with appropriate genetic preselection (for example, activating mutations in EGFR or ALK translocations) do not respond to treatments predicted to succeed based on what we currently know, so there is a strong need for diagnostic tools as well as corresponding computational analyses that can connect genetic, proteomic, and phenotypic data with clinical outcomes.

With this in mind, there appears to be significant clinical value in technologies that are able to extract as much information as possible from tumors with very high granularity, ideally at the single-cell level. Heterogeneity is a natural problem in tumors, given the different clonal expansions that can occur both with and without treatment over the course of cancer development. If heterogeneity is amplified due

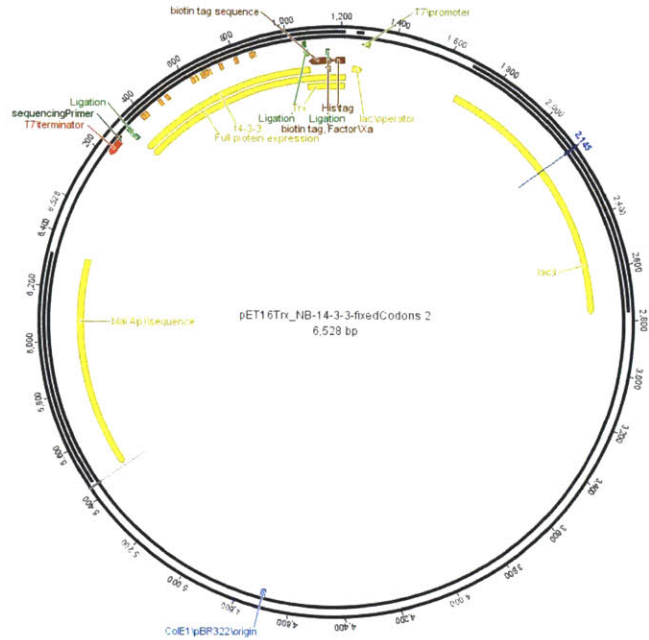
to cells developing resistance to inhibition, the problem will likely increase even more in complexity, as multiple new clonal populations can arise in this resistance developing process. Therefore, it seems an ideal solution would be the development of diagnostic tools that can gather as much relevant information about a tumor as possible, ultimately even at the single-cell level where resistance can develop, and use this data to construct predictive models for clinical outcomes. [45] Fortunately, a host of technologies are currently in development to uncover signaling dynamics at the single-cell level, which could aid in potentially uncovering novel targets that are masked by bulk studies like those performed in this thesis. [72, 98] Better identification of critical signaling events at the single-cell level through genomic or proteomic techniques may prove someday to provide significant value to patients in the clinic through better personalized treatments aimed at mitigating driver growth signals as well as targets that can drive resistance. With this in mind, it is my most sincere hope that we're in the midst of truly exciting developments in our understanding of tumor biology, and that the discoveries we're making now will prove to be very fruitful for many years to come for those suffering from such a terrible disease.

“If those committed to the quest fail, they will be forgiven. When lost, they will find another way. The moral imperative of humanism is the endeavor alone, whether successful or not, provided the effort is honorable and failure memorable. The ancient Greeks expressed the idea in a myth of vaulting ambition. Daedalus escapes from Crete with his son Icarus on wings he has fashioned from feathers and wax. Ignoring the warnings of his father, Icarus flies toward the sun, whereupon his wings come apart and he falls into the sea. That is the end of Icarus in the myth. But we are left to wonder: Was he just a foolish boy? Did he pay the price for hubris, for pride in sight of the gods? I like to think that on the contrary his daring represents a saving human grace. And so the great astrophysicist Subrahmanyan Chandrasekhar could pay tribute to the spirit of his mentor, Sir Arthur Eddington, by saying: Let us see how high we can fly before the sun melts the wax in our wings.”

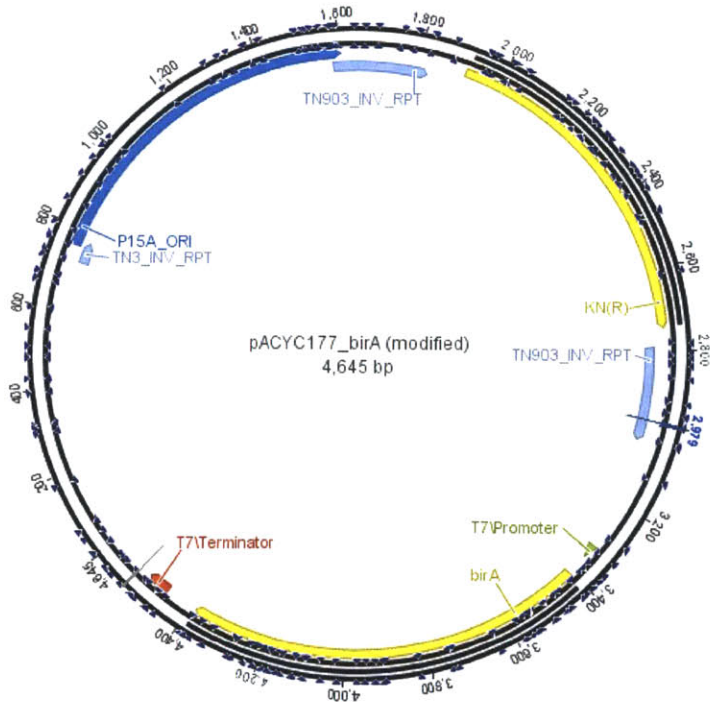
-- Edward O. Wilson

## Supplemental Information

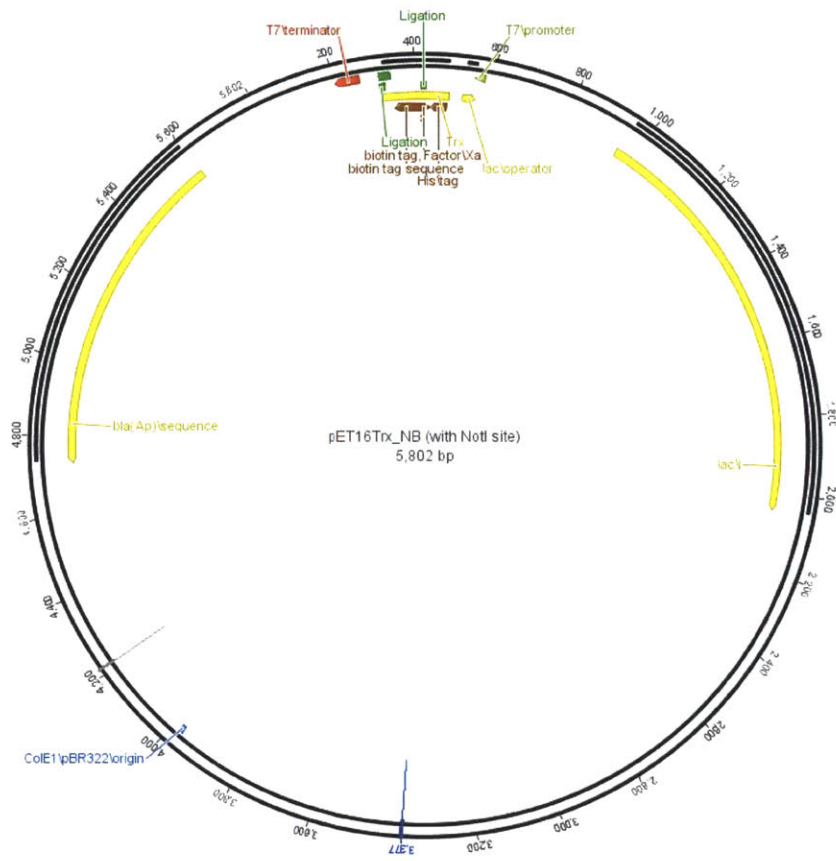
### Maps of 14-3-3 Vectors pNB-14-3-3zeta plasmid:



### pACYC177\_birA:



pNB\_NotI-SbfI (for cloning):



## Sample Data from a 14-3-3 Pull-down of Serum Stimulated HCC827 Cells

Protein	Sequence
WD repeat domain 44 protein	K.SVRDEVFHTDQDDPSSSDDEGMPYTRPVK.F + 3 Phospho (ST)
CLIP-associating protein 2	R.YESYGMHSDDDANS DASSAC SERSYSSRNGSIPTYMR.Q + 2 Oxidation (M); 2 Phospho (ST)
DEAD box polypeptide 42 protein	R.QQFH SKPVDS DSDDDPLEAFMAEVEDQAAR.D + 2 Phospho (ST)
dedicator of cytokinesis 8	K.MQVTMSLASLVGR.A + Oxidation (M); 3 Phospho (ST)
eukaryotic translation initiation factor 2B, subunit 5 epsilon, 82kDa	R.FIQWLKEAEEESSEDD.- + 2 Phospho (ST)
BCL2-associated transcription factor 1 isoform 1	K.NTPSQHSHSIQHSPER.S + Phospho (ST)
high mobility group AT-hook 1 isoform b	K.KLEKEEEEGISQESSEEEQ.- + 2 Phospho (ST)
heterogeneous nuclear ribonucleoprotein U isoform b	R.AKSPQPVEEEDHFDDTVVCLDTYNCDLHFK.I + Phospho (ST)
insulin-like growth factor 2 receptor precursor	K.LVSFHDDSDDELLHI.- + Phospho (ST)
kinesin family member 13B	R.RSISPNVNR.L + Phospho (ST)
DNA topoisomerase II, beta isozyme	K.YTFDFSEEEEDDADDDDDNNDLEELKVK.A + Phospho (ST)
cingulin	R.SHSQASLAGPGVPDPSNRSNSMLELAPK.V + 2 Phospho (ST)
cleavage and polyadenylation specific factor 2	K.EADIDSSDESIEEDIDQPSAHK.T + 3 Phospho (ST)
cleavage and polyadenylation specific factor 2	K.KLEQSKEADIDSSDESIEEDIDQPSAHK.T + 3 Phospho (ST)
microfibrillar-associated protein 1	K.RPDYAPMESSDEEDEFQFIKK.A + 2 Phospho (ST)
solute carrier family 4 (anion exchanger), member 1, adaptor protein	K.NWEDEDFYDSDDDTFLDR.T + Phospho (ST)
splicing factor 3B subunit 2	K.GFEEHKDSDDSSDDEQEKKPEAPK.L + 3 Phospho (ST)
v-raf murine sarcoma 3611 viral oncogene homolog	R.SASEPSLHR.T + Phospho (ST)
Yes-associated protein 1, 65 kD	R.AHSSPASLQLGAVSPGTLTPTGVVSGPAATPTAQHLR.Q + Phospho (ST)
BCL2-antagonist of cell death protein	R.SRSAPPNLWAAQR.Y + Phospho (ST)
tight junction protein 2 (zona occludens 2) isoform 2	R.RHQYSYDYHSSSEK.L + Phospho (ST)
ataxin 1	K.RRWSAPESR.K + Phospho (ST)

## References

1. Levi-Montalcini, R. and V. Hamburger, *A diffusible agent of mouse sarcoma, producing hyperplasia of sympathetic ganglia and hyperneurotization of viscera in the chick embryo.* Journal of Experimental Zoology, 1953. 123(2): p. 233-287.
2. Levi-Montalcini, R., *EFFECTS OF MOUSE TUMOR TRANSPLANTATION ON THE NERVOUS SYSTEM.* Annals of the New York Academy of Sciences, 1952. 55(2): p. 330-344.
3. Cohen, S. and R. Levi-Montalcini, *Purification and Properties of a Nerve Growth-promoting Factor Isolated from Mouse Sarcoma 180* Cancer Research 1957 17 (1 ): p. 15-20
4. Cohen, S., *The stimulation of epidermal proliferation by a specific protein (EGF).* Developmental Biology, 1965. 12(3): p. 394-407.
5. CARPENTER, G., L. KING, and S. COHEN, *Epidermal growth factor stimulates phosphorylation in membrane preparations in vitro.* 1978. 276(5686): p. 409-410.
6. Ushiro, H. and S. Cohen, *Identification of phosphotyrosine as a product of epidermal growth factor-activated protein kinase in A-431 cell membranes.* . Journal of Biological Chemistry 1980 255 (18 ): p. 8363-5
7. Hunter, T. and J.A. Cooper, *Epidermal growth factor induces rapid tyrosine phosphorylation of proteins in A431 human tumor cells.* Cell, 1981. 24(3): p. 741-752.
8. Hunter, T. and B.M. Sefton, *Transforming gene product of Rous sarcoma virus phosphorylates tyrosine* Proceedings of the National Academy of Sciences 1980 77 (3 ): p. 1311-1315
9. Levinson, A.D., et al., *Evidence that the transforming gene of avian sarcoma virus encodes a protein kinase associated with a phosphoprotein.* Cell, 1978. 15(2): p. 561-572.
10. Gusterson, B., et al., *Evidence for increased epidermal growth factor receptors in human sarcomas.* Int J Cancer, 1985. 36(6): p. 689-93.
11. Wong, A.J., et al., *Increased expression of the epidermal growth factor receptor gene in malignant gliomas is invariably associated with gene amplification* Proceedings of the National Academy of Sciences 1987 84 (19 ): p. 6899-6903
12. Cowley, G.P., J.A. Smith, and B.A. Gusterson, *Increased EGF receptors on human squamous carcinoma cell lines.* Br J Cancer, 1986. 53(2): p. 223-9.
13. King, C.R., M.H. Kraus, and S.A. Aaronson, *Amplification of a novel v-erbB-related gene in a human mammary carcinoma.* Science, 1985. 229(4717): p. 974-6.
14. Kawamoto, T., et al., *Growth stimulation of A431 cells by epidermal growth factor: identification of high-affinity receptors for epidermal growth factor by an anti-receptor monoclonal antibody* Proceedings of the National Academy of Sciences 1983 80 (5 ): p. 1337-1341

15. Hudziak, R.M., et al., *p185HER2 monoclonal antibody has antiproliferative effects in vitro and sensitizes human breast tumor cells to tumor necrosis factor*. Mol Cell Biol, 1989. 9(3): p. 1165-72.
16. *National Academy of Sciences*. Science 1960 132 (3438): p. 1488-1501
17. Daley, G.Q., R.A. Van Etten, and D. Baltimore, *Induction of chronic myelogenous leukemia in mice by the P210bcr/abl gene of the Philadelphia chromosome*. Science, 1990. 247(4944): p. 824-30.
18. Druker, B.J., et al., *Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells*. Nat Med, 1996. 2(5): p. 561-6.
19. Herbst, R.S., J.V. Heymach, and S.M. Lippman, *Lung Cancer*. New England Journal of Medicine, 2008. 359(13): p. 1367-1380.
20. Pao, W. and N. Girard, *New driver mutations in non-small-cell lung cancer*. The Lancet Oncology, 2011. 12(2): p. 175-180.
21. Yun, C.-H., et al., *Structures of Lung Cancer-Derived EGFR Mutants and Inhibitor Complexes: Mechanism of Activation and Insights into Differential Inhibitor Sensitivity*. Cancer Cell, 2007. 11(3): p. 217-227.
22. Pao, W. and J. Chmielecki, *Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer*. Nat Rev Cancer. 10(11): p. 760-74.
23. Lynch, T.J., et al., *Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-small-Cell Lung Cancer to Gefitinib*. New England Journal of Medicine, 2004. 350(21): p. 2129-2139.
24. Paez, J.G., et al., *EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy*. Science, 2004. 304(5676): p. 1497-500.
25. Pao, W., et al., *EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib*. Proc Natl Acad Sci U S A, 2004. 101(36): p. 13306-11.
26. Maemondo, M., et al., *Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR*. N Engl J Med. 362(25): p. 2380-8.
27. Soda, M., et al., *Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer*. Nature, 2007. 448(7153): p. 561-6.
28. Morris, S.W., et al., *Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma*. Science, 1994. 263(5151): p. 1281-4.
29. Koivunen, J.P., et al., *EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer*. Clin Cancer Res, 2008. 14(13): p. 4275-83.



30. Chi, A., et al., *Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry*. Proc Natl Acad Sci U S A, 2007. 104(7): p. 2193-8.
31. Biemann, K. and H.A. Scoble, *Characterization by tandem mass spectrometry of structural modifications in proteins*. Science, 1987. 237(4818): p. 992-8.
32. Ficarro, S.B., et al., *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae*. Nat Biotechnol, 2002. 20(3): p. 301-5.
33. Molina, H., et al., *Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry*. Proc Natl Acad Sci U S A, 2007. 104(7): p. 2199-204.
34. Yao, Z., et al., *TGF-beta IL-6 axis mediates selective and adaptive mechanisms of resistance to molecular targeted therapy in lung cancer*. Proc Natl Acad Sci U S A. 107(35): p. 15535-40.
35. Frederick, B.A., et al., *Epithelial to mesenchymal transition predicts gefitinib resistance in cell lines of head and neck squamous cell carcinoma and non-small cell lung carcinoma*. Molecular Cancer Therapeutics 2007 6 (6): p. 1683-1691
36. Qi, J., et al., *Multiple mutations and bypass mechanisms can contribute to development of acquired resistance to MET inhibitors*. Cancer Res. 71(3): p. 1081-91.
37. Gygi, S.P., et al., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags*. 1999. 17(10): p. 994-999.
38. Joughin, B.A., et al., *An integrated comparative phosphoproteomic and bioinformatic approach reveals a novel class of MPM-2 motifs upregulated in EGFRvIII-expressing glioblastoma cells*. Mol Biosyst, 2009. 5(1): p. 59-67.
39. Zhang, Y., et al., *Time-resolved Mass Spectrometry of Tyrosine Phosphorylation Sites in the Epidermal Growth Factor Receptor Signaling Network Reveals Dynamic Modules*. Molecular & Cellular Proteomics 2005 4 (9): p. 1240-1250
40. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data*. Electrophoresis, 1999. 20(18): p. 3551-67.
41. Noro, R., et al., *Gefitinib (IRESSA) sensitive lung cancer cell lines show phosphorylation of Akt without ligand stimulation*. BMC Cancer, 2006. 6: p. 277.
42. Engelman, J.A., et al., *MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling*. Science, 2007. 316(5827): p. 1039-43.
43. Sunaga, N., et al., *Knockdown of Oncogenic KRAS in Non-Small Cell Lung Cancers Suppresses Tumor Growth and Sensitizes Tumor Cells to Targeted Therapy*. Molecular Cancer Therapeutics 2011 10 (2): p. 336-346
44. Giovannetti, E., et al., *Molecular mechanisms underlying the synergistic interaction of erlotinib, an epidermal growth factor receptor tyrosine kinase inhibitor, with the multitargeted antifolate pemetrexed in non-small-cell lung cancer cells*. Mol Pharmacol, 2008. 73(4): p. 1290-300.

45. Turke, A.B., et al., *Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC*. *Cancer Cell*. 17(1): p. 77-88.
46. Ramalingam, S.S., et al., *Randomized phase II study of erlotinib in combination with placebo or R1507, a monoclonal antibody to insulin-like growth factor-1 receptor, for advanced-stage non-small-cell lung cancer*. *J Clin Oncol*. 29(34): p. 4574-80.
47. Li, Y., et al., *Axl as a potential therapeutic target in cancer: role of Axl in tumor growth, metastasis and angiogenesis*. *Oncogene*, 2009. 28(39): p. 3442-55.
48. Gangarosa, L.M., et al., *A Raf-independent Epidermal Growth Factor Receptor Autocrine Loop Is Necessary for Ras Transformation of Rat Intestinal Epithelial Cells*. *Journal of Biological Chemistry* 1997 272 (30): p. 18926-18931
49. Harris, R.C., E. Chung, and R.J. Coffey, *EGF receptor ligands*. *Exp Cell Res*, 2003. 284(1): p. 2-13.
50. Minchinton, A.I. and I.F. Tannock, *Drug penetration in solid tumours*. 2006. 6(8): p. 583-592.
51. Armstrong, E., et al., *The c-src tyrosine kinase (CSK) gene, a potential antioncogene, localizes to human chromosome region 15q23---q25*. *Cytogenet Cell Genet*, 1992. 60(2): p. 119-20.
52. Nada, S., et al., *Constitutive activation of Src family kinases in mouse embryos that lack Csk*. *Cell*, 1993. 73(6): p. 1125-35.
53. Suzuki, K., et al., *Down-regulation of the tumor suppressor C-terminal Src kinase (Csk)-binding protein (Cbp)/PAG1 is mediated by epigenetic histone modifications via the mitogen-activated protein kinase (MAPK)/phosphatidylinositol 3-kinase (PI3K) pathway*. *J Biol Chem*. 286(18): p. 15698-706.
54. Bénistant, C., et al., *The COOH-Terminal Src Kinase Csk Is a Tumor Antigen in Human Carcinoma* *Cancer Research* 2001 61 (4): p. 1415-1420
55. Rubbi, L., et al., *Global Phosphoproteomics Reveals Crosstalk Between Bcr-Abl and Negative Feedback Mechanisms Controlling Src Signaling*. *Sci. Signal.*, 2011. 4(166): p. ra18-.
56. Huang, P.H., *Phosphoproteomic studies of receptor tyrosine kinases: future perspectives*. *Mol Biosyst*. 8(4): p. 1100-7.
57. Prahallad, A., et al., *Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR*. 2012. 483(7388): p. 100-103.
58. Serra, V., et al., *PI3K inhibition results in enhanced HER signaling and acquired ERK dependency in HER2-overexpressing breast cancer*. 2011. 30(22): p. 2547-2557.
59. Amin, D.N., et al., *Resiliency and vulnerability in the HER2-HER3 tumorigenic driver*. *Sci Transl Med*. 2(16): p. 16ra7.
60. Sergina, N.V., et al., *Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3*. *Nature*, 2007. 445(7126): p. 437-41.

61. Turke, A.B., et al., *MEK inhibition leads to PI3K/AKT activation by relieving a negative feedback on ERBB receptors*. *Cancer Res*.
62. Yonesaka, K., et al., *Activation of ERBB2 signaling causes resistance to the EGFR-directed therapeutic antibody cetuximab*. *Sci Transl Med*. 3(99): p. 99ra86.
63. Chiu, G.N., et al., *Lipid-based nanoparticulate systems for the delivery of anti-cancer drug cocktails: Implications on pharmacokinetics and drug toxicities*. *Curr Drug Metab*, 2009. 10(8): p. 861-74.
64. Sierra, J.R., V. Cepero, and S. Giordano, *Molecular mechanisms of acquired resistance to tyrosine kinase targeted therapy*. *Mol Cancer*. 9: p. 75.
65. Engelman, J.A., et al., *MET Amplification Leads to Gefitinib Resistance in Lung Cancer by Activating ERBB3 Signaling*. *Science* 2007 316 (5827): p. 1039-1043
66. Szerlip, N.J., et al., *Intratatumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response*. *Proceedings of the National Academy of Sciences* 2012 109 (8): p. 3041-3046
67. Snuderl, M., et al., *Mosaic Amplification of Multiple Receptor Tyrosine Kinase Genes in Glioblastoma*. *Cancer cell*, 2011. 20(6): p. 810-817.
68. Huang, P.H., et al., *Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma*. *Proc Natl Acad Sci U S A*, 2007. 104(31): p. 12867-72.
69. Huang, P.H., A.M. Xu, and F.M. White, *Oncogenic EGFR signaling networks in glioma*. *Sci Signal*, 2009. 2(87): p. re6.
70. Karaman, M.W., et al., *A quantitative analysis of kinase inhibitor selectivity*. *Nat Biotechnol*, 2008. 26(1): p. 127-32.
71. Lindauer, M. and A. Hochhaus, *Dasatinib*. *Recent Results Cancer Res*. 184: p. 83-102.
72. Bendall, S.C., et al., *Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum*. *Science*. 332(6030): p. 687-96.
73. Sasaki, T., et al., *A novel ALK secondary mutation and EGFR signaling cause resistance to ALK kinase inhibitors*. *Cancer Res*. 71(18): p. 6051-60.
74. Bellacosa, A., et al., *A retroviral oncogene, akt, encoding a serine-threonine kinase containing an SH2-like region*. *Science*, 1991. 254(5029): p. 274-7.
75. Burgering, B.M. and P.J. Coffey, *Protein kinase B (c-Akt) in phosphatidylinositol-3-OH kinase signal transduction*. *Nature*, 1995. 376(6541): p. 599-602.
76. Bridges, D. and G.B.G. Moorhead, *14-3-3 Proteins: A Number of Functions for a Numbered Protein*. *Sci. STKE*, 2005. 2005(296): p. re10-

77. Muslin, A.J., et al., *Interaction of 14-3-3 with signaling proteins is mediated by the recognition of phosphoserine*. Cell, 1996. 84(6): p. 889-97.
78. B. Yaffe, M. and L. C. Cantley, *Signal transduction: Grabbing phosphoproteins*. 1999. 402(6757): p. 30-31.
79. Satoh, J., Y. Nanri, and T. Yamamura, *Rapid identification of 14-3-3-binding proteins by protein microarray analysis*. J Neurosci Methods, 2006. 152(1-2): p. 278-88.
80. Yaffe, M.B., et al., *The structural basis for 14-3-3:phosphopeptide binding specificity*. Cell, 1997. 91(7): p. 961-71.
81. Beckett, D., E. Kovaleva, and P.J. Schatz, *A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation*. Protein Sci, 1999. 8(4): p. 921-9.
82. Polzien, L., et al., *Identification of novel in vivo phosphorylation sites of the human proapoptotic protein BAD: pore-forming activity of BAD is regulated by phosphorylation*. J Biol Chem, 2009. 284(41): p. 28004-20.
83. Hekman, M., et al., *Dynamic changes in C-Raf phosphorylation and 14-3-3 protein binding in response to growth factor stimulation: differential roles of 14-3-3 protein binding sites*. J Biol Chem, 2004. 279(14): p. 14074-86.
84. White, F.M., *The potential cost of high-throughput proteomics*. Sci Signal. 4(160): p. pe8.
85. Johnson, C., et al., *Bioinformatic and experimental survey of 14-3-3-binding sites*. Biochem J. 427(1): p. 69-78.
86. Kitano, H., *Biological robustness*. Nat Rev Genet, 2004. 5(11): p. 826-37.
87. Kitano, H., *Cancer as a robust system: implications for anticancer therapy*. 2004. 4(3): p. 227-235.
88. Kreuzaler, P. and C.J. Watson, *Killing a cancer: what are the alternatives?* 2012. advance online publication.
89. Venkataraman, S., et al., *Selective cell death mediated by small conditional RNAs*. Proc Natl Acad Sci U S A. 107(39): p. 16777-82.
90. Xie, Z., et al., *Multi-input RNAi-based logic circuit for identification of specific cancer cells*. Science. 333(6047): p. 1307-11.
91. Varshavsky, A., *Targeting the absence: homozygous DNA deletions as immutable signposts for cancer therapy*. Proc Natl Acad Sci U S A, 2007. 104(38): p. 14935-40.
92. Chen, Y.Y. and C.D. Smolke, *From DNA to targeted therapeutics: bringing synthetic biology to the clinic*. Sci Transl Med. 3(106): p. 106ps42.
93. Culler, S.J., K.G. Hoff, and C.D. Smolke, *Reprogramming cellular behavior with RNA controllers responsive to endogenous proteins*. Science. 330(6008): p. 1251-5.

94. Win, M.N. and C.D. Smolke, *Higher-order cellular information processing with synthetic RNA devices*. Science, 2008. 322(5900): p. 456-60.
95. von Maltzahn, G., et al., *Nanoparticles that communicate in vivo to amplify tumour targeting*. Nat Mater. 10(7): p. 545-52.
96. Pleasance, E.D., et al., *A small-cell lung cancer genome with complex signatures of tobacco exposure*. 2010. 463(7278): p. 184-190.
97. Hanahan, D. and R.A. Weinberg, *Hallmarks of Cancer: The Next Generation*. Cell, 2011. 144(5): p. 646-674.
98. Shi, Q., et al., *Single-cell proteomic chip for profiling intracellular signaling pathways in single tumor cells*. Proceedings of the National Academy of Sciences 2012 109 (2 ): p. 419-424