Title
Multi-omic Analysis of B-Cell Lymphoma Reveals Novel Mechanisms of Chemotherapeutic Drug Resistance

Author
Flinders, Colin

Publication Date
2014

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Mulit-omic Analysis of B-Cell Lymphoma Reveals Novel Mechanisms of Chemotherapeutic Drug Resistance

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biological Chemistry

by

Colin Flinders

2014

ABSTRACT OF THE DISSERTATION

Mulit-omic Analysis of B-Cell Lymphoma Reveals Novel Mechanisms of Chemotherapeutic Drug Resistance

by

Colin Flinders

Doctor of Philosophy Biological Chemistry

University of California, Los Angeles, 2014

Professor Joseph Loo, Chair

The genetic origins of chemotherapy resistance are well established, however the role of the epigenome and post-transcriptional regulation in drug resistance is less well understood. To investigate mechanisms of drug resistance we performed a systematic genetic, epigenetic, transcriptomic and proteomic analysis of a mafosfamide sensitive and resistant murine lymphoma cell line, along with a series of resistant lines derived by drug dose escalation. Our data suggest that acquired resistance could not be explained by genetic alterations. By integrating our transcriptional profiles with transcription factor binding data we hypothesize that the resistance was associated with changes in the activity of the polycomb repressive complex (Prc2) as well as the transcription factor *E2a*. We verified that the resistant cells had distinct H3K27me3 and DNA methylation profiles, compared to the parental lines, and differentially expressed genes were enriched for targets of *E2a*. In addition, the resistant lines appear to de-differentiate to a less mature state along the B cell maturation axis. Overall, we propose that resistant lines are transformed by an *E2a* and Pcr2 driven cellular program that leads to a less mature B cell state in which the apoptotic cascade induced by mafosfamide treatment is attenuated.

Furthermore, combined transcriptomic and proteomic data analysis elucidated mechanisms of resistance involving the ubiquitination activating enzyme Uba1 which were not revealed by analysis of either transcriptomic or proteomic data alone.

The dissertation of Colin Flinders is approved by.

David Eisenburg

Kathrin Plath

Matteo Pellegrini

Parag Mallick

Joseph Loo, Committee Chair

University of California, Los Angeles

2014

This dissertation is dedicated to Kathleen Simpson.

Table of Contents and Lists

List of Figures and Tables

Acknowledgements

I would like to acknowledge the following people for their support and mentorship: Shannon Mumenthaler, Jonathan Katz, Greg Payne, and David Agus. I would also like to acknowledge Larry Lam who did the statistical analysis for Chapter 2, Liudmilla Rubbi who constructed the whole genome sequencing libraries, Roberto Ferrari who performed the ChIP-Seq experiments, Sorel Fitz-Gibbon who did the mapping and base calling, Pao-Yang Chen who did the initial entropy measurements, Michael Thompson who analyzed the Cancer Genome Atlas data, Heather Christofk who performed the metabolomics studies, David Agus who supervised the project, Daniel Ruderman who supervised the project, Parag Mallick who supervised the project, and Matteo Pellegrini who supervised the project for their contributions to Chapter 2 of this dissertation. Additionally, I would like to acknowledge Dror Berel who did the statistical analysis for Chapter, Ahyoung Joo who operated the mass spectrometers, Ruth Alvarez who operated the mass spectrometers, Dan Ruderman who supervised the project, Jonathan Katz who supervised the project, David Agus who supervised the project, and Parag Mallick who supervised the project for their contributions to Chapter 3 of this dissertation. Finally, I would like to acknowledge the National Cancer Institute for funding this work.

Vita

Education

University of California, Berkeley

B.A., Molecular and Cellular Biology, May 2006

Research Experience

Graduate student

University of California, Los Angeles

Los Angeles Ca, 9/2008-12/2014

- Developed *in vitro* model of epigenetic resistance in Burkitt's lymphoma.
- Worked in multidisciplinary teams to develop computational cell based models to changes in drug, microenvironment, and protein translational response.

Staff Research Associate I

Sangamo BioScience, Inc.

Richmond Ca, 8/2006-5/2008

- Created knockout as well as knockin cell lines.
- Characterized clonal populations of CD8+ cell lines for lack of gluccocortiod receptor and presence of chimeric T-cell receptor.

Posters & Presentations

Flinders, C., Poultney, C., Greenfield, A., Katz, J., Bonneau, R., Agus, D., Mallick, M. (2011). Proteomic and transcriptomic network regulatory analysis of chemotherapeutic resistant and sensitive murine B cell lymphoma. Poster presented at the second annual meeting of the National Cancer Institute Physical Sciences in Oncology, San Diego, Ca.

Flinders, C., Poultney, C., Greenfield, A., Katz, J., Bonneau, R., Agus, D., Mallick, P. (2012). Interrogation of proteomic and transciptomic temporal dynamics in response to DNA damaging agents. Poster presented at the third annual meeting of the National Cancer Institute Physical Sciences in Oncology, Tampa Bay, Fl.

Flinders, C., Larry, L., Rubbi, L., Agus, D., Christofk, H., Ruderman, D., Mallick, P., Pellegrini, P. (2013). Tumor evolution via alteration in DNA methylation. Poster presented at the fourth annual meeting of the National Cancer Institute Physical Sciences in Oncology, Scottsdale, Az.

Flinders, C., Larry, L., Rubbi, L., Ferrari, R., Fitz-Gibbon, S., Chen, P.Y., Christofk, H., Agus, D., Ruderman, D., Mallick, P., Pellegrini, P. (2014). Epigenetic Drug Resistance in B Cell Lymphoma. Poster presented at the fifth annual meeting of the National Cancer Institute Physical Sciences in Oncology, National Institute of Health, Bethesda, Md.

"Epigenetic Drug Resistance in B Cell Lymphoma" (2014) Talk presented at the fifth annual meeting of the National Cancer Institute Physical Sciences in Oncology, National Institute of Health, Bethesda, Md.

Chapter 1: Drug Resistance in Burkitt's Lymphoma

Introduction

Since the discovery of the first oncogene mutation in human cancer, research into the development of cancer and therapeutic resistance has primarily focused on mutations of the genome. Mutations can provide a survival advantage such as the nullification of the target of drug action, inactivation of apoptotic pathways, over expression of DNA repair genes, or the activation of proliferative pathways. Whereas traditional chemotherapy non-specifically kills rapidly dividing cells more modern targeted therapies leverage knowledge of these specific mutations to specifically target only those cells that contain a specific unique mutation or gene product. Despite the success of both traditional and targeted therapies, resistance is a major cause for the failure of treatment and as such understanding the mechanisms of acquired drug resistance will have important implications in the development of new therapies and drug regimens.

Despite extensive investigation, genetic mutations are unable to explain all cases of acquired resistance. Non-genetic mechanism of resistance have been suggested as an alternate to explain those cases of resistance that arise rapidly, reverse with drug holiday, lack detectable mutations, or where there exits heterogeneity in drug responsiveness amongst cells sharing the same genetic profile. Epigenetics, the heritable change in gene expression that is independent of alterations in the underlying DNA sequence is capable of explaining these cases of resistance. Epigenetic mechanisms of drug resistance can involve the aberrant loss or acquisition of DNA methylation or through the chemical modification of histone tails. Alteration of either DNA methylation patterns or histone modifications can lead to changes in higher order higher order chromatin structure resulting in either the compaction (heterochormatin) or decompression (euchromatin) of a section of the chromosome. Such structural

changes lead to the physical occlusion of DNA binding factors and preventing transcription in the case of compaction or allowing for an open confirmation which would permit their binding and allow for transcription. Furthermore, both types of modifications can act as binding sites for specific activator or repressors of gene expression (such as bromodomain or methyl-CpG-binding domain containing proteins).

During carcinogenesis alterations in the epigenome can lead to the transcriptional repression or activation of targets of drug action, DNA repair pathways, apoptosis, or the activation of cell proliferative pathways. Epigenetic alterations occur early in tumorigenesis and as a result tumors show a high degree of heterogeneity at the epigenetic level[1,2]. This epigenetic heterogeneity has important clinical implications as it can be a source for therapeutic selection. Furthermore, alterations of the epigenome can result in the ectopic expression of genes involved in differentiation and stem cell maintenance leading to the reprogramming of cancer cells towards a more stem cell like state.

B-cell Lymphoma

Lymphomas affect an estimated 450,000 new cases every year with 225,000 deaths annually[3]. Increased precision of histological and molecular classification techniques has led to the identification of over 60 lymphoma subtypes[4]. Treatment outcome is widely varied between these subtypes with some indolent types such as follicular and marginal zone lymphoma remaining incurable chronic diseases that require repeated rounds of treatment to more aggressive subtypes such as diffuse large B-cell lymphoma (DLBCL) and Burkitt's lymphoma (BL) that require intensive high dose therapy. B-cell lymphomas are a group heterogeneous malignancies that arise from mature naive B-cells that have successfully undergone rearrangement of the B-cell receptor. In normal B-cell development antigen activated B-cells migrate to lymphoid structures known as germinal centers (GC) where they undergo

stomatic hypermutation and class-switch recombination. B-cell lymphoma classification is based on the rationale that malignant B-cells are arrested at the stage of development from which they originate[5].

Burkitt's Lymphoma accounts for 1-2% of all cases of adult lymphoma in Westerns Europe and the United States[6]. BL is characterized by a C-myc translocation leading to over expression of C-Myc mRNA and protein[7]. Approximately, 80% of BL cases harbor a t(8;14) translocation, which places C-myc under the control of the IgH enhancer elements, with the other 20% of cases having translocation that place c-myc adjacent to either the κ or λ light chain loci and enhancer elements[7]. The structure of the IgH c-myc translocation suggests that this event occurs during class-switch recombination during normal B-cell development[8]. C-myc is a helix-loop-helix leucine zipper transcription factor involved in multiple pathways important in cancer including cell cycle regulation, apoptosis, cell growth, cell adhesion and differentiation[9].

Despite the greater understanding of molecular differences between various lymphoma subtypes the treatment paradigm for the vast majority lymphomas and in particular aggressive B-cell lymphomas has remained unchanged. The most commonly used therapeutic regime is the combinatorial treatment known as R-CHOP (rituximab, cyclophosphamide, hydroxydaunorubicin, oncovin, and prednisone)[6,10]. In BL, high intensity short duration R-CHOP regimens result in complete response (CR) rates of 47-100% in adults with most attaining CR within 4 to 6 weeks of the start of treatment[11,12]. However, only 47%-86% of individuals with CR maintain these results for at least a year following therapy[6]. Furthermore, for those individuals with only a partial response or who relapse an optimal salvage therapy remains unknown[6]. Even those patients who obtain long term remission of the disease often suffer from lasting toxicity induced health problems, predisposition to secondary malignancies, and an overall lower quality of life[13].

Drug Resistance in Cancer

Cancer cells are capable of becoming resistant to therapy in several different ways, some of which act broadly to enable the cancer to become resistant to multiple drug types, where as others are more specific and affect sensitivity to only one drug or class of drugs. Resistance can arise during the course of treatment (acquired resistance) or resistant cells can be present before the incitation of therapy (innate resistance). Regardless of how or when resistance arises, it ultimately results in a population of cells which are selected for and enriched during the course of treatment. Overtime, these cells may come to constitute the bulk of the tumor at which point therapy becomes ineffective and unless alternative therapies can be found the disease will progress.

Largely, resistance arises either through the over expression of pro-survival pathways or the down regulation of apoptotic pathways. Over expression or activating mutations in genes in pro-survival pathways such as the PI3K/Akt/mTor pathway can lead to cell survival. In the PI3K/Akt/mTor pathway for example, activating mutations in PI3KCA leads to hyperphosphorylation of phosphatidylinositol ultimately leading to the activation of Akt and mTor. Akt and mTor are then able to activate multiple anti-apoptotic molecules such as Bad, caspase 9 and various members of the forkhead transcription factor family[14,15]. Alternatively, over expression of drug efflux pumps, specifically the ATP-binding cassette (ABC) transporter family can lead to resistance against a broad swath of anti-cancer compounds. By up regulating ABC transporter genes cells are capable of lowering intracellular drug concentration, thus reducing the effect the drug has[16]. Over expression of ABC transporter proteins can be inherent to specific tissues from which the cancer arises, however over expression can also be induced by chemotherapy exposure[17]. For example, over expression of ABCB1 is associated with treatment failure in a wide range of cancers including kidney, colon, lung, as well as leukemias and lymphomas[18]. Cancer stem cells (discussed below) are inherently more resistant to therapy have been shown to express higher levels of drug eflux proteins[19].

Inactivating mutations in tumor suppressor genes can broadly affect drug resistance by disabling cell cycle checkpoints or inhibiting apoptosis. One of the most extensively studies examples is the tumor suppressor Tp53 which affects multiple survival pathway, including cell cycle regulation, senescence, DNA damage and repair[20]. Disruption of Tp53-dependent apoptosis either through inactivating mutations in Tp53 itself or of components of the Tp53 pathway are essential for the development of lymphoproliferative diseases[21,22]. As such the Tp53 status is an important prognostic indicator in lymphoid malignancies where it is associated with unfavorable outcome[23] . Although the frequency of Tp53 mutations across all lymphoid malignancies is low, over 30% of BL samples possess mutated Tp53[24].

Besides these general mechanisms of drug resistance, cells are capable of undergoing alterations that are specific to individual drugs or families of drugs. Targeted therapy is more specific than traditional chemotherapy in that it is directed at specific genetic legions that are altered in cancer. The first clinically successful targeted therapy is the kinase imatinib which inhibits the Bcr-Abl fusion gene that is a hallmark of chronic myeloid leukemia (CML)[25]. Resistance to targeted therapies is often the result of either point mutations which prevent drug binding or activation of pathway components downstream of the target protein. In CML patients, resistance to imatinib is most often the result of the former with point mutations in the ATP binding domain of Bcr-Abl preventing imatinib binding[25].

B-cell lymphomas are a highly heterogeneous malignancies and as such the mechanisms of resistance are varied. In BL, mutations in the hematopoietic differentiation transcription factors ID3 and TCF3 (E2a) result in PI3K/Akt/mTOR pathway dependency[24,26] and mutations in this pathway leaad to worse prognosis[27]. As mentioned above mutations in p53 or p53 pathway components are essential to disease development and progression. The location of disruption of the pathway can have important clinical implications[21,28]. Besides these intrinsic forms of resistance, B-cell lymphomas also experience

acquired resistance. Up regulation of ABC family members or drug metabolism genes such as Adh1 frequently occur[29]. Similarly, down regulation of rituximab target CD20 has been noted in individuals who relapse following R-CHOP[30].

Epigenetics and Tumor Heterogeneity

Tumor heterogeneity

Cellular heterogeneity is a well characterized attribute of cancer tissue. Tumor heterogeneity may be the result of genetic drift during early stages of cancer progression which leads to several clonal subpopulations. Deep-sequencing studies have revealed genetic heterogeneity through the analysis of copy number variations at the level of single cell sequencing and of genetic mutations from different tumor sites and metastases from an individual patients[31-33]. Identified genetic changes can be traced along a phylogenetic tree to develop a history of the appearance of specific mutations. These clonal subpopulations could then go through evolutionary selection as the result of nutrient deficiencies that arise as the initial tumor grows or as the result of therapeutic intervention[34]. The resultant heterogeneity will have therapeutic consequences as genetically distinct clonal subpopulations may have varying responses to targeted therapies[31,35]. Whether resistance arises as the result of de novo mutations during treatment or are present in the tumor at the time of initial treatment remains open for debate[36]. These two models are not mutually exclusive and which model may dominate in a particular case may be dependent on tumor type.

While it is clear that certain epigenetic modifications occur early in tumor progression and that they can be selected for, the lack of longitudinal epigenetic studies in cancer have hindered a better understanding of the evolutionary dynamics of these marks during the progression and treatment of the disease. Such studies are necessary in order to better understand epigenetic heterogeneity and how it can arise. Currently, epigenetic heterogeneity is thought to arise in one of two ways. In deterministic

variability, epigenetic differences represent distinct cell type specific alterations that lead to differentiation hierarchies. In stochastic variability, cell to cell variability within differentiation status results in heterogeneity within a population of cells at the same developmental stage. An epigenetic alteration in a progenitor cell population can lead to the clonal expansion of a distinct sub-population within a larger polyclonal population. These cells can then undergo further genetic or epigenetic alterations that can result in cycles of selection and clonal expansion. Ultimately, this can lead to increased plasticity amongst various tumor cell subpopulations which allows for greater survival. All tumor types show widespread aberrant epigenetic modification, which is capable of generating diversity in tumor populations that can be a source of drug resistant subclones[2,37].

While heterogeneity in cancer cell populations exists at both the genetic and epigenetic level, it is far greater at the epigenetic level. The frequency of gene promoter hypermethylation is several times higher than the frequency of nonsynonymous gene mutations in cancer, with a median of 250-800 versus 150-170 respectively[38]. Additionally, the fidelity at which CpG methylation is copied during cell division is estimated to be $2x10^{-5}$ per CpG per division compared to the mutation rate in cancer cells of 10^{-10} mutations per nucleotide base pair per division[39]. This results in a far higher level of heterogeneity at the DNA methylation level in tumors than is found in mutations. The resulting heterogeneity has clinical implications as well. It has been shown that in individuals with diffuse large B-cell lymphoma (DLBCL) DNA methylation heterogeneity is correlated with poor prognosis[40].

Epigenetic modifications in cancer

DNA methylation

DNA methylation is perhaps the most well studied chromatin modification. In mammals, DNA methylation primarily occurs on the cytosine residues of CpG dinucleotides by the action of DNA methyltransferases (DNMTs) forming the more hydrophobic 5-methly-cytosine. Much of the DNA

methylation of the genome occurs at regions enriched for CpG dinucleotides known as CpG islands which were first described in 1982 by Wolf and Migeon when they observed that methylated regions on the inactive X chromosome were associated with silencing of housekeeping genes[41]. DNA methylation acts through affecting higher order chromatin structure leading to a more condensed chromatin state and as a binding site for chromatin binding proteins. DNA methylation plays a vital role in gene expression during embryogenesis and development, where methylation of CpG islands in the promoter regions of genes leads to gene silencing. Conversely, methylation can also increase gene expression when it occurs in gene bodies where it is thought to facilitate transcription elongation. Besides its role in gene expression, DNA methylation is also critical in X chromosome inactivation, imprinting, and chromosomal stability through the methylation of repetitive DNA regions.

Cells from neoplastic tissue show aberrant DNA methylation patterns and increased heterogeneity compared to normal matched controls. Generally speaking, tumors are hypomethylated which leads to increased variable gene expression in differentially methylated regions[1] and increased chromosomal instability potentially contributing to gene copy number changes seen in cancer[42]. Simiarly, site specific hypomethylation has been observed as well. For example, hypomethylation of the drug eflux pump ABCB1 has been observed during treatment in AML patients and correlates with increased ABCB1 mRNA within 4 hours of the start of chemotherapy [43]. In addition to hypomethylation, there are examples of gene specific hypermethyation in cancer. Hypermethylation typically occurs in the promoter regions of tumor suppression genes such as RB1[44,45] or the CDKN2A locus. Hypermethylation has been shown to inactivate several important tumor suppressor genes in Burkitt's Lymphoma including Klf4[46], p16Ink4a, p15Ink4b[47], PUMA[48] and Bcl2l11[49]. Hypermethylation of tumor suppressor genes appears to occur after genes have already been repressed by histone modifications indicating a link between histone modification mediated gene silencing and DNA methylation[50,51]. Similarly, hypermethylation of CpG islands in the promoter of mismatch repair (MMR) enzyme MGMT results in its

silencing in numerous tumors including lymphoma. In glioblastoma, the methylation status of MGMT is used as a biomarker since hypermethylation of MGMT correlates with better response to the alkylating chemotherapy temozolomide[52]. MGMT repairs temozolomide induced O^6-methylguanine converting it back to guanine and thus preventing cell death. Hypomethylation of MGMT results in increased expression of MGMT and thus increased ability to repair temozolomide induced damage.

DNA methylation can be inhibited through the use of cytosine analog 5'-aza-cytosine (5-Aza) whose incorporation into DNA leads to the reduction of DNA methylation. When incorporated into DNA 5-Aza will irreversibly bind DNMTs thereby preventing them from methylating cytosine residues. Sub-lethal doses of 5-Aza lead to widespread de-methylation of the genome resulting in the re-expression of genes actively repressed by hypermethylation. Treatment of c-MYC driven lymphomas with 5-Aza has been shown to be able to induce expression of ID2 a negative regulator of c-MYC causing the reduction in both c-MYC RNA and protein levels and leading to the inhibition of proliferation[53]. Currently 5-Aza is approved for the treatment of several myelodysplastic syndromes including chronic myelomonocytic leukemia.

Histone Modifications

The packaging of DNA into nucleosomes by histones allows for the dynamic regulation of gene activity through the extensive modification of histone tails. Histone tails are capable of undergoing extensive post-translational modification including acetylation, methylation, phosphorylation, ubiquination and sumoylation. Beyond potentially acting as binding sites for other chromatin modifying proteins as well as activators and repressors of gene transcription, these modifications directly cause changes in the chromatin structure. Histone acetyltransferases (HATs) acetylate lysine residues thereby neutralizing the positive charge of lysine and resulting in a weaker interaction between nucleosomes and DNA allowing for greater accessibility to transcription factors. Histone deacetyltransferases (HDACs)

catalyze the opposite reaction reestablishing the positive charge resulting in a tightly packed chromatin structure that physically occludes transcription factors from binding the underlying DNA sequence. Histone methylation can occur at both lysine and arginine residues and results in an increase in the hydrophobicity of the histone tails. Whereas lysine acetylation is exclusively an activating modification, histone methylation can serve both as an activation as well as a repressive mark depending on the residue and the amount of methylation (mono-, di-, or tri-). For example, monomethylation of histone H3 lysine 9 (H3K9) and trimethylation of H3K4 leads to activation, while trimethylation of H3K9 and H3K27 lead to transcriptional repression. As mentioned previously, histones can exist in a bivalant state containing both activating as well as repressive marks leading to low levels of gene transcription, while keeping genes poised for activation.

Histone modifying proteins as well as the altered histone marks themselves have been shown to play a role both in the development of tumors as well as the acquisition of resistance. For example, over expression of the H3K27 methyltransferase EZH2 is implicated in tumorogensis and its experession correlates with poor prognosis in multiple cancers[54-56]. Additionally, activating mutations in EZH2 have been found in primary tumor samples taken from patients with non-Hodgkin's lymphoma[57]. These mutations change the substrate specificity of the enzyme allowing for the more favorable addition of second and third methyl groups to H3K27[58]. Additionally, ER positive breast cancer patients who replase early in response to aromatase inhibitors have an enrichment for H3K27me3 marks[59]. Similarly, loss of H3K18 acetylation is correlated with tumor grade and poor prognosis in breast, lung, prostate, pancreatic and kidney cancer[60-62]. Mutations in epigenetic modifiers frequently occur in cancers that either replase or are resistant to therapy such as the histone acetyltransferase CBP in acute lymphoblastic leukemia[63].

Given the role that altered histone modifications play in the development of cancer and their potential role in drug resistance several small molecule inhibitors have been developed. Currently two HDAC inhibitors (HDACi), vorinostat and romidepsin are approved for the treatment of advanced coetaneous T-cell lymphoma. The use of HDACi's to treat multiple other cancer types such as diffuse large B-cell lymphoma and non-small cell lung carcinomas that are resistant to EGF TKIs has shown them to have anti-cancer activity and to be well tolerated[64,65]. Similarly, allosteric inhibitors to the histone methyltransferase EZH2 are currently in the preclinical development phase for the treatment of non-Hodgkin's lymphoma. Another inhibitor currently being developed targets the histone methyltransferase DOT1L which catalyzes the methylation of H3K79. Increased DOT1L activity is found in some mixed lineage leukemias (MLLs) which contain rearrangements of the MLL gene. Inhibition of DOTL1 with shRNA or small molecules inhibits proliferation of MLL rearranged leukemia cells in vitro and extends survival of MLL tumor xenografts[66].

Investigating various mechanisms of drug resistance in Burkitt's lymphoma

Mutations in Ezh2 have been shown to exist in various forms of lymphoma including BL. However, it is less clear how alterations in the Ezh2 mark H3K27me3 affect drug sensitivity and whether such changes can occur independently to activating mutations in Ezh2. Chapter 2 of this dissertation examines how H3K27me3 occupancy changes during adaptation to the DNA alkylating drug mafosfamide. Our data indicates that alterations in gene expression which accompany increased drug tolerance show increased H3K27me3 occupancy. Furthermore, those genes with increased H3K27me3 occupancy are associated with the master B-cell regulator E2a, which shows a similar pattern of expression as does H3k27me3.

The studies conducted in Chapter 3 concern the combined analysis of omic data (transcriptomic and proteomic) and their application in investigating drug resistance. By studying gene and protein

expression changes in response to mafosfamide over time in both resistant and sensitive Burkitt's lymphoma cell lines we elucidate the role that Uba1 plays in drug resistance. Our data shows that in a Tp53 null Burkitt's lymphoma cell line increased protein expression of Uba1 and that the inhibition of Uba1 shows a synergistic effect with mafosfamide.

References

1. Hansen KD, Timp W, Bravo HC, et al. Increased methylation variation in epigenetic domains across cancer types. *Nat Genet*. 2011;43(8):768-775.
2. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nat Rev Cancer*. 2011;11(10):726-734.
3. Intlekofer AM, Younes A. Precision therapy for lymphoma-current state and future directions. *Nat Rev Clin Oncol*. 2014;11(10):585-596.
4. Campo E, Swerdlow SH, Harris NL, Pileri S, Stein H, Jaffe ES. The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood*. 2011;117(19):5019-5032.
5. Kuppers R. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer*. 2005;5(4):251-262.
6. Blum KA, Lozanski G, Byrd JC. Adult Burkitt leukemia and lymphoma. *Blood*. 2004;104(10):3009-3020.
7. Hecht JL, Aster JC. Molecular biology of Burkitt's lymphoma. *J Clin Oncol*. 2000;18(21):3707-3721.
8. Kuppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene*. 2001;20(40):5580-5594.
9. Sewastianik T, Prochorec-Sobieszek M, Chapuy B, Juszczynski P. MYC deregulation in lymphoid tumors: molecular mechanisms, clinical consequences and therapeutic implications. *Biochim Biophys Acta*. 2014;1846(2):457-467.
10. Roschewski M, Staudt LM, Wilson WH. Diffuse large B-cell lymphoma-treatment approaches in the molecular era. *Nat Rev Clin Oncol*. 2014;11(1):12-23.
11. Bernstein JI, Coleman CN, Strickler JG, Dorfman RF, Rosenberg SA. Combined modality therapy for adults with small noncleaved cell lymphoma (Burkitt's and non-Burkitt's types). *J Clin Oncol*. 1986;4(6):847-858.
12. Thomas DA, Cortes J, O'Brien S, et al. Hyper-CVAD program in Burkitt's-type adult acute lymphoblastic leukemia. *J Clin Oncol*. 1999;17(8):2461-2470.
13. Ng AK, LaCasce A, Travis LB. Long-term complications of lymphoma and its treatment. *J Clin Oncol*. 2011;29(14):1885-1892.
14. De Luca A, Maiello MR, D'Alessio A, Pergameno M, Normanno N. The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches. *Expert Opin Ther Targets*. 2012;16 Suppl 2:S17-27.
15. Hafsi S, Pezzino FM, Candido S, et al. Gene alterations in the PI3K/PTEN/AKT pathway as a mechanism of drug-resistance (review). *Int J Oncol*. 2012;40(3):639-644.
16. Gottesman MM, Fojo T, Bates SE. Multidrug resistance in cancer: role of ATP-dependent transporters. *Nat Rev Cancer*. 2002;2(1):48-58.

17. Chen KG, Sikic BI. Molecular pathways: regulation and therapeutic implications of multidrug resistance. *Clin Cancer Res*. 2012;18(7):1863-1869.
18. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer*. 2013;13(10):714-726.
19. Shervington A, Lu C. Expression of multidrug resistance genes in normal and cancer stem cells. *Cancer Invest*. 2008;26(5):535-542.
20. Xu-Monette ZY, Medeiros LJ, Li Y, et al. Dysfunction of the TP53 tumor suppressor gene in lymphoid malignancies. *Blood*. 2012;119(16):3668-3683.
21. Schmitt CA, Fridman JS, Yang M, Baranov E, Hoffman RM, Lowe SW. Dissecting p53 tumor suppressor functions in vivo. *Cancer Cell*. 2002;1(3):289-298.
22. Peller S, Rotter V. TP53 in hematological cancer: low incidence of mutations with significant clinical relevance. *Hum Mutat*. 2003;21(3):277-284.
23. Sturm I, Bosanquet AG, Hermann S, Guner D, Dorken B, Daniel PT. Mutation of p53 and consecutive selective drug resistance in B-CLL occurs as a consequence of prior DNA-damaging chemotherapy. *Cell Death Differ*. 2003;10(4):477-484.
24. Schmitz R, Young RM, Ceribelli M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature*. 2012;490(7418):116-120.
25. O'Hare T, Walters DK, Stoffregen EP, et al. In vitro activity of Bcr-Abl inhibitors AMN107 and BMS-354825 against clinically relevant imatinib-resistant Abl kinase domain mutants. *Cancer Res*. 2005;65(11):4500-4505.
26. Love C, Sun Z, Jima D, et al. The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet*. 2012;44(12):1321-1325.
27. Maxwell SA, Mousavi-Fard S. Non-Hodgkin's B-cell lymphoma: advances in molecular strategies targeting drug resistance. *Exp Biol Med (Maywood)*. 2013;238(9):971-990.
28. Schmitt CA, McCurrach ME, de Stanchina E, Wallace-Brodeur RR, Lowe SW. INK4a/ARF mutations accelerate lymphomagenesis and promote chemoresistance by disabling p53. *Genes Dev*. 1999;13(20):2670-2677.
29. Ghetie MA, Ghetie V, Vitetta ES. Anti-CD19 antibodies inhibit the function of the P-gp pump in multidrug-resistant B lymphoma cells. *Clin Cancer Res*. 1999;5(12):3920-3927.
30. Duman BB, Sahin B, Ergin M, Guvenc B. Loss of CD20 antigen expression after rituximab therapy of CD20 positive B cell lymphoma (diffuse large B cell extranodal marginal zone lymphoma combination): a case report and review of the literature. *Med Oncol*. 2012;29(2):1223-1226.
31. Anderson K, Lutz C, van Delft FW, et al. Genetic variegation of clonal architecture and propagating cells in leukaemia. *Nature*. 2011;469(7330):356-361.
32. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011;472(7341):90-94.
33. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. 2012;366(10):883-892.
34. Mullighan CG, Phillips LA, Su X, et al. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science*. 2008;322(5906):1377-1380.
35. Yang J, Luo H, Li Y, et al. Intratumoral heterogeneity determines discordant results of diagnostic tests for human epidermal growth factor receptor (HER) 2 in gastric cancer specimens. *Cell Biochem Biophys*. 2012;62(1):221-228.
36. Sierra JR, Cepero V, Giordano S. Molecular mechanisms of acquired resistance to tyrosine kinase targeted therapy. *Mol Cancer*. 2010;9:75.
37. Glasspool RM, Teodoridis JM, Brown R. Epigenetics as a mechanism driving polygenic clinical drug resistance. *Br J Cancer*. 2006;94(8):1087-1092.

38. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546-1558.
39. Yatabe Y, Tavare S, Shibata D. Investigating stem cells in human colon by using methylation patterns. *Proc Natl Acad Sci U S A*. 2001;98(19):10839-10844.
40. De S, Shaknovich R, Riester M, et al. Aberration in DNA methylation in B-cell lymphomas has a complex origin and increases with disease severity. *PLoS Genet*. 2013;9(1):e1003137.
41. Wolf SF, Migeon BR. Studies of X chromosome DNA methylation in normal human cells. *Nature*. 1982;295(5851):667-671.
42. De S, Michor F. DNA secondary structures and epigenetic determinants of cancer genome evolution. *Nat Struct Mol Biol*. 2011;18(8):950-955.
43. Hu XF, Slater A, Kantharidis P, et al. Altered multidrug resistance phenotype caused by anthracycline analogues and cytosine arabinoside in myeloid leukemia. *Blood*. 1999;93(12):4086-4095.
44. Sakai T, Toguchida J, Ohtani N, Yandell DW, Rapaport JM, Dryja TP. Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet*. 1991;48(5):880-888.
45. Greger V, Passarge E, Hopping W, Messmer E, Horsthemke B. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. *Hum Genet*. 1989;83(2):155-158.
46. Guan H, Xie L, Leithauser F, et al. KLF4 is a tumor suppressor in B-cell non-Hodgkin lymphoma and in classic Hodgkin lymphoma. *Blood*. 2010;116(9):1469-1478.
47. Lindstrom MS, Wiman KG. Role of genetic and epigenetic changes in Burkitt lymphoma. *Semin Cancer Biol*. 2002;12(5):381-387.
48. Garrison SP, Jeffers JR, Yang C, et al. Selection against PUMA gene expression in Myc-driven B-cell lymphomagenesis. *Mol Cell Biol*. 2008;28(17):5391-5402.
49. Richter-Larrea JA, Robles EF, Fresquet V, et al. Reversion of epigenetically mediated BIM silencing overcomes chemoresistance in Burkitt lymphoma. *Blood*. 2010;116(14):2531-2542.
50. Bachman KE, Park BH, Rhee I, et al. Histone modifications and silencing prior to DNA methylation of a tumor suppressor gene. *Cancer Cell*. 2003;3(1):89-95.
51. Sproul D, Kitchen RR, Nestor CE, et al. Tissue of origin determines cancer-associated CpG island promoter hypermethylation patterns. *Genome Biol*. 2012;13(10):R84.
52. Hegi ME, Diserens AC, Gorlia T, et al. MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med*. 2005;352(10):997-1003.
53. Guan H, Xie L, Klapproth K, Weitzer CD, Wirth T, Ushmorov A. Decitabine represses translocated MYC oncogene in Burkitt lymphoma. *J Pathol*. 2013;229(5):775-783.
54. Varambally S, Dhanasekaran SM, Zhou M, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature*. 2002;419(6907):624-629.
55. Wagener N, Macher-Goeppinger S, Pritsch M, et al. Enhancer of zeste homolog 2 (EZH2) expression is an independent prognostic factor in renal cell carcinoma. *BMC Cancer*. 2010;10:524.
56. Takawa M, Masuda K, Kunizaki M, et al. Validation of the histone methyltransferase EZH2 as a therapeutic target for various types of human cancer and as a prognostic marker. *Cancer Sci*. 2011;102(7):1298-1305.
57. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011;476(7360):298-303.
58. Sneeringer CJ, Scott MP, Kuntz KW, et al. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc Natl Acad Sci U S A*. 2010;107(49):20980-20985.
59. Jansen MP, Knijnenburg T, Reijm EA, et al. Hallmarks of aromatase inhibitor drug resistance revealed by epigenetic profiling in breast cancer. *Cancer Res*. 2013;73(22):6632-6641.

60. Elsheikh SE, Green AR, Rakha EA, et al. Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. *Cancer Res*. 2009;69(9):3802-3809.
61. Manuyakorn A, Paulus R, Farrell J, et al. Cellular histone modification patterns predict prognosis and treatment response in resectable pancreatic adenocarcinoma: results from RTOG 9704. *J Clin Oncol*. 2010;28(8):1358-1365.
62. Seligson DB, Horvath S, McBrian MA, et al. Global levels of histone modifications predict prognosis in different cancers. *Am J Pathol*. 2009;174(5):1619-1628.
63. Mullighan CG, Zhang J, Kasper LH, et al. CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature*. 2011;471(7337):235-239.
64. Nolan L, Johnson PW, Ganesan A, Packham G, Crabb SJ. Will histone deacetylase inhibitors require combination with other agents to fulfil their therapeutic potential? *Br J Cancer*. 2008;99(5):689-694.
65. Sharma SV, Lee DY, Li B, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*. 2010;141(1):69-80.
66. Daigle SR, Olhava EJ, Therkelsen CA, et al. Potent inhibition of DOT1L as treatment of MLL-fusion leukemia. *Blood*. 2013;122(6):1017-1025.

Chapter 2: Epigenetic Changes Mediated by Polycomb Repressive Complex 2 and E2a are Associated with Drug Resistance in a Mouse Model of Lymphoma

Introduction

Evolutionary processes are a central component of cancer initiation and progression. During treatment, cancers often acquire resistance, which ultimately leads to poor clinical outcome. A number of explanations for resistance have been proposed, including the presence of cancer stem cells[1], and of mutations conferring drug resistance[2]. Here we examine drug resistance mechanisms in a model of Burkitt's lymphoma. Though extremely aggressive[3], Burkitt's lymphoma has a high cure rate with complete remission among 90-95% of children receiving the standard of care therapy, a combinatorial treatment of rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone. In adults, treatment isn't as successful[4], in part due to acquired resistance. In both children and in adults, salvage treatment has a poor success rate; only one third of children and very few adults have positive outcomes from salvage therapy. Uncovering novel mechanisms of resistance may lead to new approaches to more effective treatment strategies.

Currently, the mechanisms underlying acquisition of resistance in Burkitt's lymphoma are only partially understood. Numerous genetic mechanisms have been hypothesized including up regulation of drug efflux genes such as the ATP-binding cassette (ABC) transporter family, cyclophosphamide inactivation through aldheyde dehydrogenase up regulation, increased expression of DNA repair enzymes, or deregulation of apoptosis through the loss of *Tp53*[5]. However, genetic mutations are unable to explain cases of acquired resistance that arise rapidly, or reverse in response to a drug holiday[6,7]. Alterations in histone modifications and DNA methylation that lead to an altered transcriptional program have also been proposed as epigenetic mechanisms for acquired drug resistance

in B cell lymphoma and other types of cancer[8,9]. Recent work has shown that treatment with the DNA methylation inhibitor 5-azacytidine is capable of reactivating *Id2*, a repressor of translocated Myc under control of the Eμ promoter in an *in vitro* model of Burkitt's lymphoma resulting in the inhibition of proliferation[10]. Similarly, pretreatment of lymphoma cell lines with the histone deacetylase (HDAC) inhibitor suberoylanilide hydroxamic acid (SAHA) has been able to re-sensitize lymphoma cell lines to various therapeutic agents and resulted in better treatment outcome *in vitro* [11].

Furthermore, studies of clinical specimens revealed that tumors are both genetically and epigenetically heterogeneous[8,12]. While the role of genetic heterogeneity within tumors and its effect on treatment response and outcome has been extensively studied, less is known about the role that epigenetic heterogeneity plays in disease progression and clinical outcome. Previous studies have shown that drug treatment can generate a selective pressure upon heterogeneous populations, leading to the enrichment of specific genetically distinct subpopulations[13,14]. These subpopulations can ultimately become the dominant population of the tumor, resulting in resistance to the therapeutic agent. It is possible that similar mechanisms of selection may act at the epigenetic level. Recent research in prostate cancer has documented the inherent heterogeneity of DNA methylation in patient tumor samples[15], though the selection of epigenetically distinct subpopulations has yet to be shown.

Beyond evident genetic alterations leading to resistance, there is also accumulating evidence that cell differentiation may impact chemosensitivity. Loss of differentiation and subsequent acquisition of a stem-like phenotype enables cancer cells to survive treatment through an increase in DNA damage response and alterations in cell cycle progression[16]. There is also evidence to suggest that epigenetic alterations can lead to transcriptional programs that resemble those of less differentiated cell types[17]. Perhaps the best-studied example of this phenomenon is the epithelial to mesenchymal transition (EMT). EMT is characterized by dysregulation of the *TGF-β* signaling pathway and down regulation of E-

cadherin[18] and has been observed in a variety of cancers where it is associated with a stem cell phenotype, metastasis, and multi-drug resistance[19,20]. Similarly, previous work has shown that hematopoietic stem cells are more radioresistant than their more differentiated progeny (Meijne et al. 1991).

Through a dose escalation experiment we investigate the mechanism by which *Eμ-Myc Cdkn2a*$^{-/-}$ Non-Hodgkin's B cell lymphoma cells acquire resistance to the chemotherapeutic agent mafosfamide, a DNA alkylating agent that is similar to the active form of cyclophosphamide. By sequencing the genomes of both the parental and resistant cell lines we were unable to identify genetic changes that might underlie the acquired resistance. In contrast, we hypothesize that the acquired resistance is mediated by Polycomb Repressive Complex 2 associated alterations in histone H3 lysine 27 trimethylation (H3K27me3) and DNA methylation, and transcriptional changes mediated by *E2a*, a master regulator of B cell development. Overall, our results indicate that the *Eμ-Myc Cdkn2a*$^{-/-}$ lymphoma cell line may become resistant to mafosfamide through the epigenetic reactivation of developmental pathways leading to a less mature state of B cell development.

Methods

Creation of resistant Eμ-Myc Cdkn2a-/- lines

Eμ-Myc Cdkn2a-/- parental lymphoma lines were generated from a C57BL/6J mouse as described in Schmitt, *et. al.* 1999[21] (**Supplementary Figure 1A**). Lymphoma cells from these mice were cultured *in vitro* to generate the cell line. Resistant strains were generated from this parental line via dose escalation of mafosfamide (Cell Signaling Technology, Danvers, Ma) over a 34-day peroid.

Cell Viability and Cell Cycle Analysis

Cell viability was measured using the Perkin-Elmer Operetta platform, and 2.5µM Draq5 (Abcam, Cambridge, Ma) for nuclear detection and 5 µg/mL of propidum iodide (Sigma-Aldrich, St. Louis, Mo) to detect dead cells. For cell cycle analysis, cells were fixed in EtOH and placed in solution with propidium iodide and measured on a Beckman LSRII and gated for G0/G1, S, and M phase cells.

Genome Sequencing

To identify variations between the sequences of the parental and resistant lines, all cell lines were sequenced to a minimum of 8x average coverage (after QC) using Illumina HiSeq 2000 sequencers. The reads were aligned to the mm9 (MGSCv37) *Mus musculus* reference genome using BWA version 0.6.2-r126 (backtrack)[22], with default parameters. Duplicate reads were removed using PICARD Version: 1.85(1345), default parameters (See supplemental methods).

Microarray

Oligonucleotide microarray analysis was carried out using GeneChip Mouse Gene ST 1.0 (Affymetrix, Santa Clara, Ca). The resulting data is publically available via Gene Expression Omnibus (GEO) Accession GSE60342. The Affymetrix Mouse Gene ST 1.0 Array expression measures were quantified and processed with robust multi-array average using the justRMA function of the 1.40.0 affy R package (RMA)[23]. Expression values were $\log_2$ transformed for further downstream analysis. Probe sets were annotated using the Affymetrix MoGene-1_0-st-v1.na33.2.mm9.probeset.csv file. We selected the top 1000 probe sets ranked by their covariance to identify up regulated genes across the time course or resistant cell lines (See supplemental methods).

Transcription Factor Analysis

Targets for 64 murine transcription factors were identified from ChIPBase (http://deepbase.sysu.edu.cn/chipbase, downloaded August 1, 2013)[24] and limited to genes with binding events within 5kb of transcriptional start sites (See Supplemental Table S2 for list of mouse cell lines used). To identify potential upstream regulators of lists of genes, we identified the overlap of these lists with transcription factor targets and used a one-sided Fisher's exact test to determine significance.

ChIP-Seq

Chromatin was immunoprecipitated as described previously[25]. Briefly, cells were grown to 50% confluency and formaldehyde was added for 10 min at room temperature. 100 µl of the lysate (5×10^6 cells) were used for each immunoprecipitation with anti-H3K27me3 (Active Motive cat # 39155). Libraries were sequenced using Illumina HIseq-2000 to obtain 50-bp-long reads.

Peaks were called by comparing counts in the immunoprecipitated libraries with input libraries in windows tiling the genome, using Poisson statistics as previously described[25]. Combinatorial clustering of data was achieved by determining significant enrichment for the histone mark in each condition within 5 Kb spanning the TSS of mouse promoters (at least 3 50 bp bin with $p < 1.0e-6$). A binary distribution was created based on a promoter being enriched (1) or not enriched (0) and a combinatorial matrix was created with all possible combinations across all conditions. H3K27me3 data was plotted based on the combinatorial clustering and visualized by Cluster 3.0-generated CDT file loaded on Java-Tree view to produce a heat map.

Bisulfite Sequencing

Reduced Representation Bisulfite (RRBS) libraries were generated following the standard RRBS protocol[26]. The genome was digested with the methylation-insensitive restriction enzyme MspI, and fragments from 100 to 300 bases were selected. The fragments were ligated with Illumina adaptors,

denatured and treated with sodium bisulfite. The libraries were sequenced using Illumina Hiseq 2000 sequencers. The reads were aligned using the bisulfite aligner, BS Seeker2[27], to determine which fragment they uniquely mapped to allowing for 3 mismatches to the reference genome (mouse mm9).

DNA Methylation Analysis

To identify DNA methylation changes that increased or decreased significantly with resistance, we computed RRBS fragment CpG methylation levels and calculated the covariance between the fragment methylation score and sample order (ordered from least to most resistant) (See supplemental methods).

Principal Component Analysis

Principal Component Analysis of expression profiles was performed by applying the R *prcomp* function with the scaled option to the expression microarray values of the resistant cell lines and B cells at different stages of development (NCBI GEO GSE15907) (See supplemental methods for more details on samples used).

Results

Generation of resistant lines and assessment of cell cycle characteristics.

To investigate the factors driving acquisition of resistance to chemotherapy, we selected a widely used cell line derived from a *Eμ-Myc* mouse model of Burkitt's lymphoma[21] which expresses the c-myc oncogene under control of an immunoglobulin heavy chain enhancer, thereby restricting its expression to B cell lineage cells. In addition, this line has a deletion in *Cdkn2a* that recapitulates a common mutation seen in human tumors[28]. *Eμ-Myc Cdkn2a*$^{-/-}$ mice experience accelerated lymphomagenesis and tumors that are highly invasive with apoptotic defects[21]. We refer to this line as

our 'parental' line. Resistant lines were generated by gradually exposing the parental line to increasing amounts of mafosfamide in cell culture. (**Figure 1A & B**). The cell cycle progression of our parental and resistant lines in the absence of mafosfamide was indistinguishable (**Supplementary Table S1**). Upon treatment with mafosfamide there was a decrease in the percentage of apoptotic cells in the resistant lines compared to the parental (**Figure 1C**). Among the non-apoptotic fraction of the population there was a similar decrease in percentage of cells in G1 and a concurrent increase in G2 percentage across all lines. As all the lines showed approximately the same amount of cell cycle delay upon treatment in the non-apoptotic fraction, it does not appear that the resistance phenotype is the result of increased cell cycle delay in the resistant lines.

Resistance is unlikely to have arisen by genetic mechanisms.

Given the pronounced difference in drug sensitivity between the parental and resistant cell lines, we initially hypothesized that resistance could be mediated by genetic variants that were initially present in a small number of cells and then came to dominate the population over the course of the dose escalation. To investigate this we attempted to identify genetic alterations in the resistant lines by searching for single nucleotide variants (SNVs) which showed and an increase in allele frequency in each successively more resistant line. We were unable to find any high confidence SNVs that displayed this pattern (see Supplementary Results). Similarly, no large structural variations or potential copy number variations were found to convincingly vary between the lines. Taken together, this data suggest that genetic alterations do not contribute to the acquired resistance phenotype.

Analysis of upstream regulators of gene expression point to Prc2 and E2a.

To gain insights into possible resistance mechanisms, we first sought to characterize the transcriptome changes of the parental line in the presence of mafosfamide over a 48-hour period

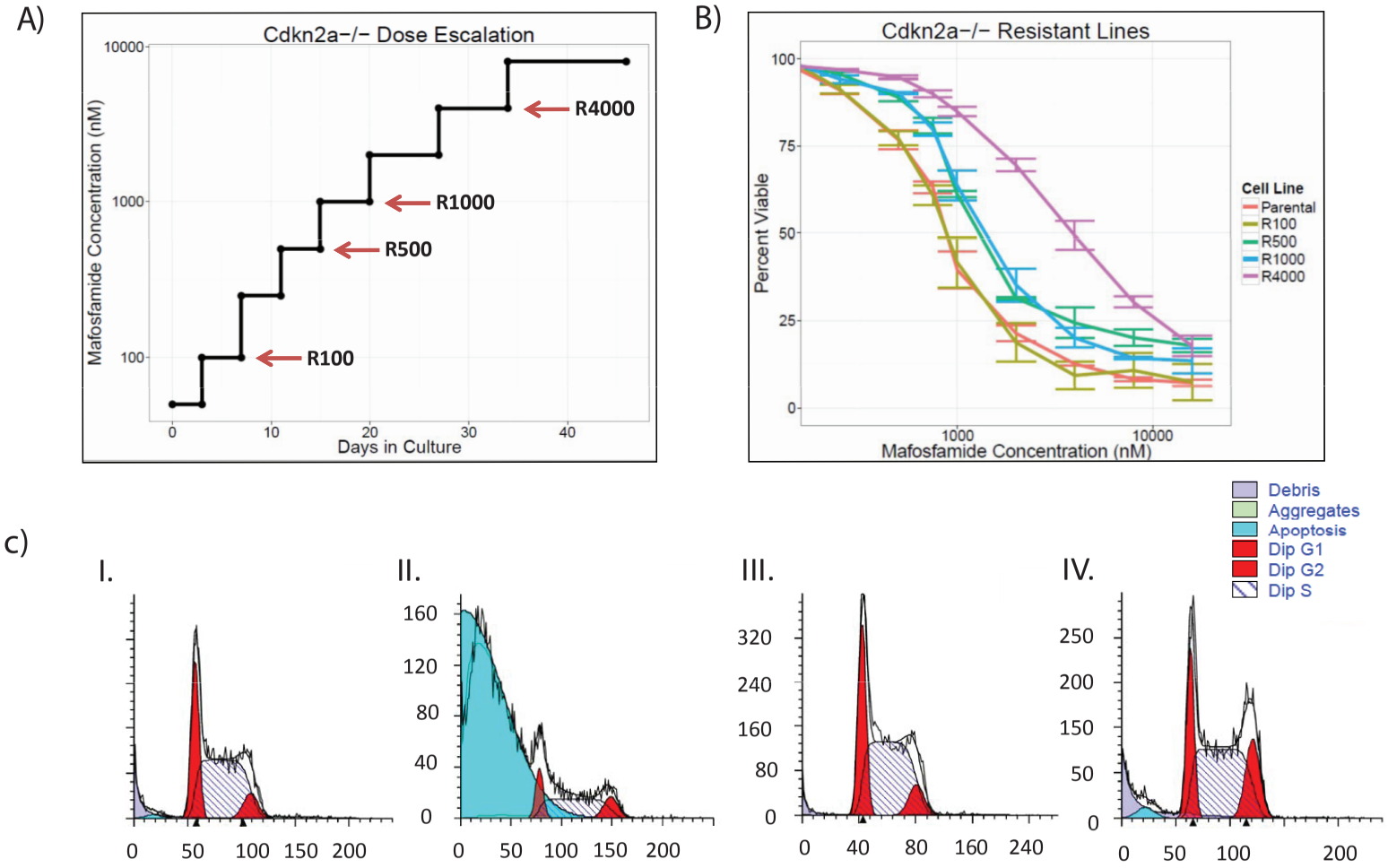


Figure 1. Dose escalation of Eμ-Myc Cdkn2a -/- cells and resistance to mafosfamide.

A) Eμ-Myc Cdkn2a-/- cells were cultured with increasing doses of mafosfamide. After 10, 18, 23, and 37 days, surviving drug resistant cells were extracted at 100, 500, 1000, and 4000 nM of mafosfamide. B) The % viability for the resistant cells was measured with 24hr exposure to drug at 100, 500, 750 and 1000nM of drug. C) I. FACS quantification of the parental cells, II. parental cells after treatment, III. untreated 4000 resistant cells, and IV. treated 4000 resistant cells.

(**Figure 2A**). Initially we focused on genes known to be involved in either drug metabolism or transport of mustard alkylating agents and those involved in DNA repair[29,30]. Overall, we found that the expression of these genes did not significantly increase in our dose escalation lines, suggesting that they also are unlikely to play a significant role in mediating resistance (see Supplementary Materials for details).

We identified genes that show significantly increased or decreased expression over time by computing the covariance of the top and bottom 1000 probe sets with respect to time (**Figure 2A**). In order to elucidate a common mechanism responsible for the transcriptional changes we observed, we sought to identify upstream regulators of the co-regulated genes using a mouse chromatin immunoprecipitation followed by sequencing (ChIP-seq) database on 64 transcription factors and their targets[24]. Each transcription factor was tested for association with the probe sets identified to have consistent changes in expression. Targets of transcription factors involved in hematopoietic stem cell development including members of Prc2, *Suz12* and *Pcl2*, are enriched in the high covariance gene list ($p < 1e-12$), while *E2a* transcription factor target genes are highly enriched in the low covariance list ($p < 1e-18$) (**Figure 2B)**[31-34].

Resistant cell lines show altered H3K27me3 occupancy.

As Prc2 complex is involved in the trimethylation of lysine 27 of Histone H3 (H3K27me3), we performed ChIP-seq for H3K27me3 in all lines. We found rapid and widespread acquisition of H3K27me3 at the transcriptional start sites (TSS) of genes after only a short time in co-culture with mafosfamide (**Figure 3A**). Upon further co-culture, H3K27me3 is globally reduced at the TSS but is retained and even further enriched at specific genomic loci. As most H3K27me3 peaks are near the TSS, we analyzed promoters by means of combinatorial clustering and identified three major groups of genes, one of which showed a high level of H3k27me3 throughout the time course (K1), another with rapid and widespread H3K27me3 acquisition followed by a gradual decrease (K2), and one group with

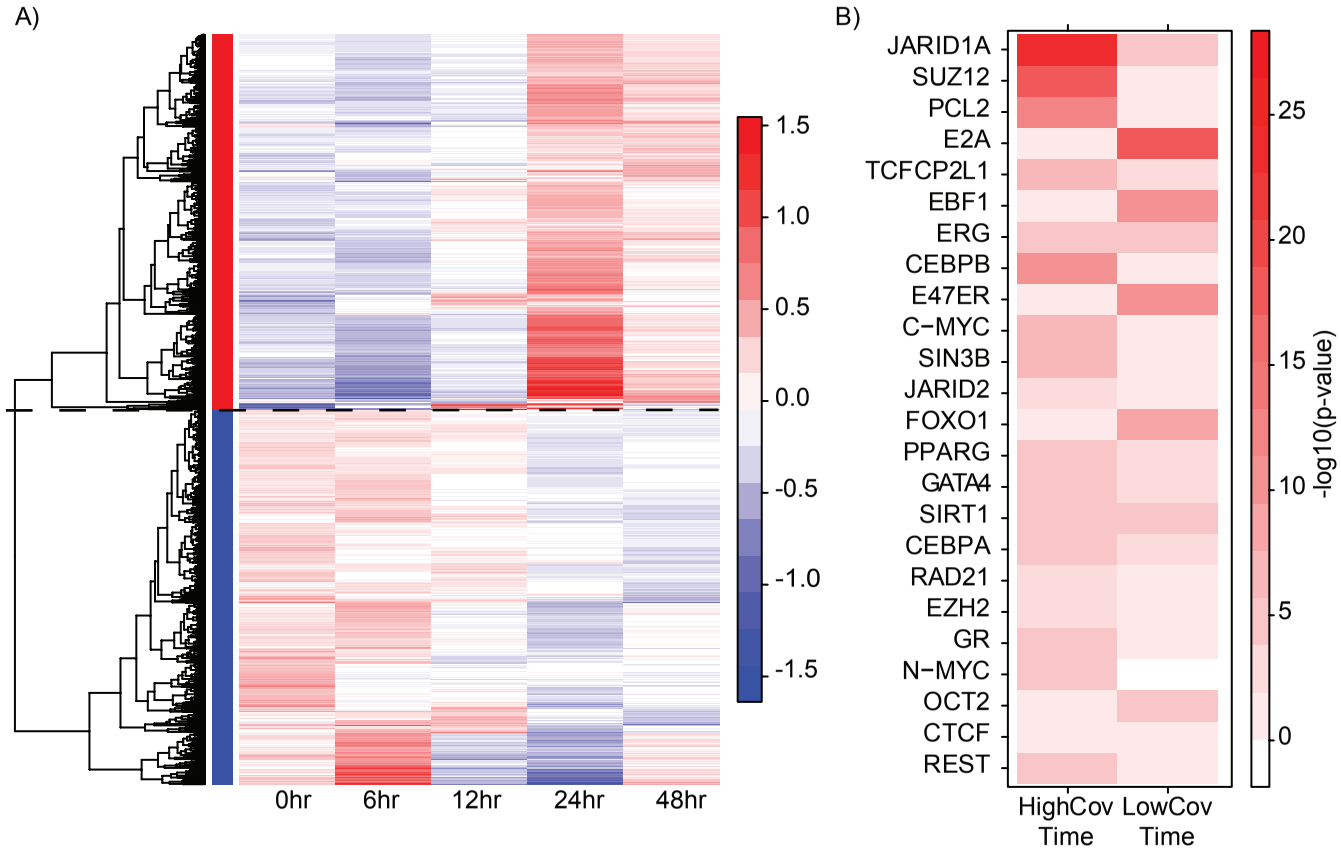


Figure 2. Microarray expression profile of the parental Eμ-Myc Cdkn2a -/- time course.
Microarray expression was performed on parental cells after 0, 6, 12, 24 and 48 hours of exposure to mafosfamide. A) Heatmap of the probe sets with high covariance with the treatment order in the red cluster along with the probe sets with low covariance with the treatment in the green cluster in the parental cell line time course (n=3). B) Genes with extreme covariance were tested for overlap with the targets of 64 different transcription factors. The heatmap shows –log(fisher test p-value) between the gene list and 23 of the transcription factors with significant overlap. Genes with positive covariance are enriched for targets of Prc2 subunits, Suz12 and Pcl2. Genes with negative covariance are enriched for E2a target genes.

little to no H3k27me3 (K3). Functional analysis of these groups showed enrichment for developmentally-regulated genes in K1 and cell cycle-regulated genes in K2 (**Figure 3B & C**). Measuring the level of H3K27me3 across all *E2a* bound genes revealed that the average profile follows a similar pattern of change to that of cluster K2 (**Figure 3D**). Furthermore, when we looked for enrichment of *E2a* bound genes within clusters K1, K2, and K3 we found a significant enrichment for them in cluster K2, and a significant depletion in K1 (**Figure 3E**).

Alteration of DNA methylation in Prc2 target genes in resistant cell lines.

To gain further insights into epigenetic changes that occur during the acquisition of drug resistance, we measured DNA methylation using Reduced Representation Bisulfite Sequencing (RRBS) of the parental and resistant lines. Focusing our analysis on the top and bottom 1000 genes (**Figure 4A)** with the highest and lowest covariance with respect to resistance revealed an enriched for Pcr2 and *E2a* (**Figure 4B**). Consistent with the previous results, we found that these fragments are also proximal to genes that are enriched for binding by Prc2 subunits *Suz12*, *Pcl2*, *Ezh2*, and *Jarid2* (**Figure 4C**). Furthermore the level of DNA methylation across ChIP-seq H3K27me3 clusters showed that genes in K2 increased in DNA methylation between the parental and resistant cell lines (**Figure 4E**) compared to those from K1 and K3 (**Figures 4D and F**). On a more global scale entropy analysis, a measure of the variability of methylation patterns, of RRBS fragments show that those loci with increased CpG methylation entropy show an increase in H3K27me3 occupancy suggesting that fragments undergoing changes in their methylation status are being bound by *Prc2* (**Supplementary Figure 5)**.

Gene expression changes in mafosfamide treated resistant cell lines.

For each resistant line we measured the changes in gene expression that occur upon treatment, and calculated the covariance of the expression response to mafosfamide with respect to increasing resistance. Comparing these covariance measures to those computed for the parental line revealed a

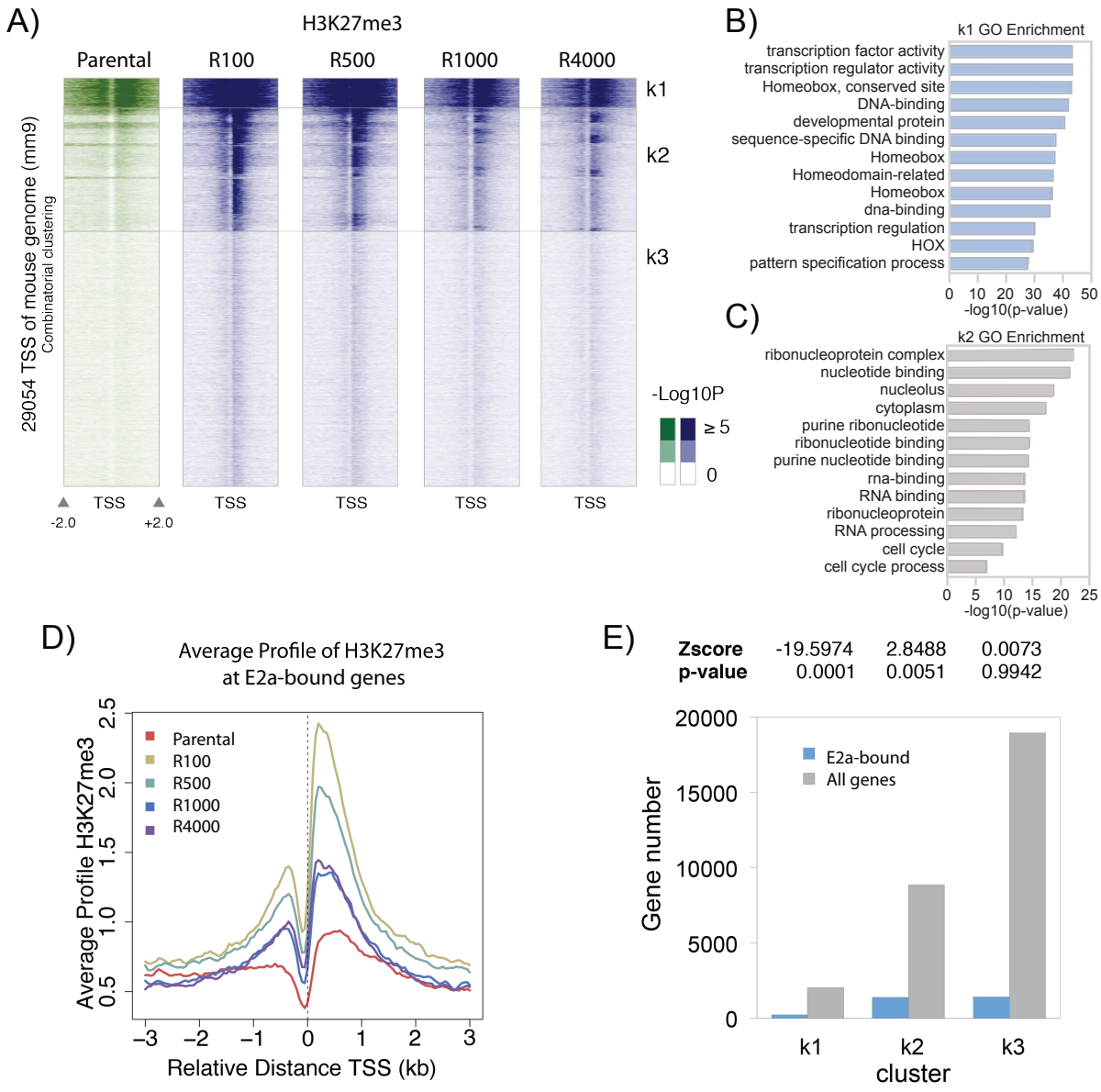


Figure 3. H3K27me3 ChIP-Seq.

ChIP-seq against histone H3 Lysine 27 tri-methylation. A) Clusters of promoters based on their H3K27me3 levels across the parental and resistant lines (n=2). B) Group K1 shows enrichment for developmental genes. C) Group K2 shows enrichment for cell cycle regulated genes. D) Average H2K27me3 levels across E2a bound genes. E) Analysis of E2a bound genes in the three clusters.

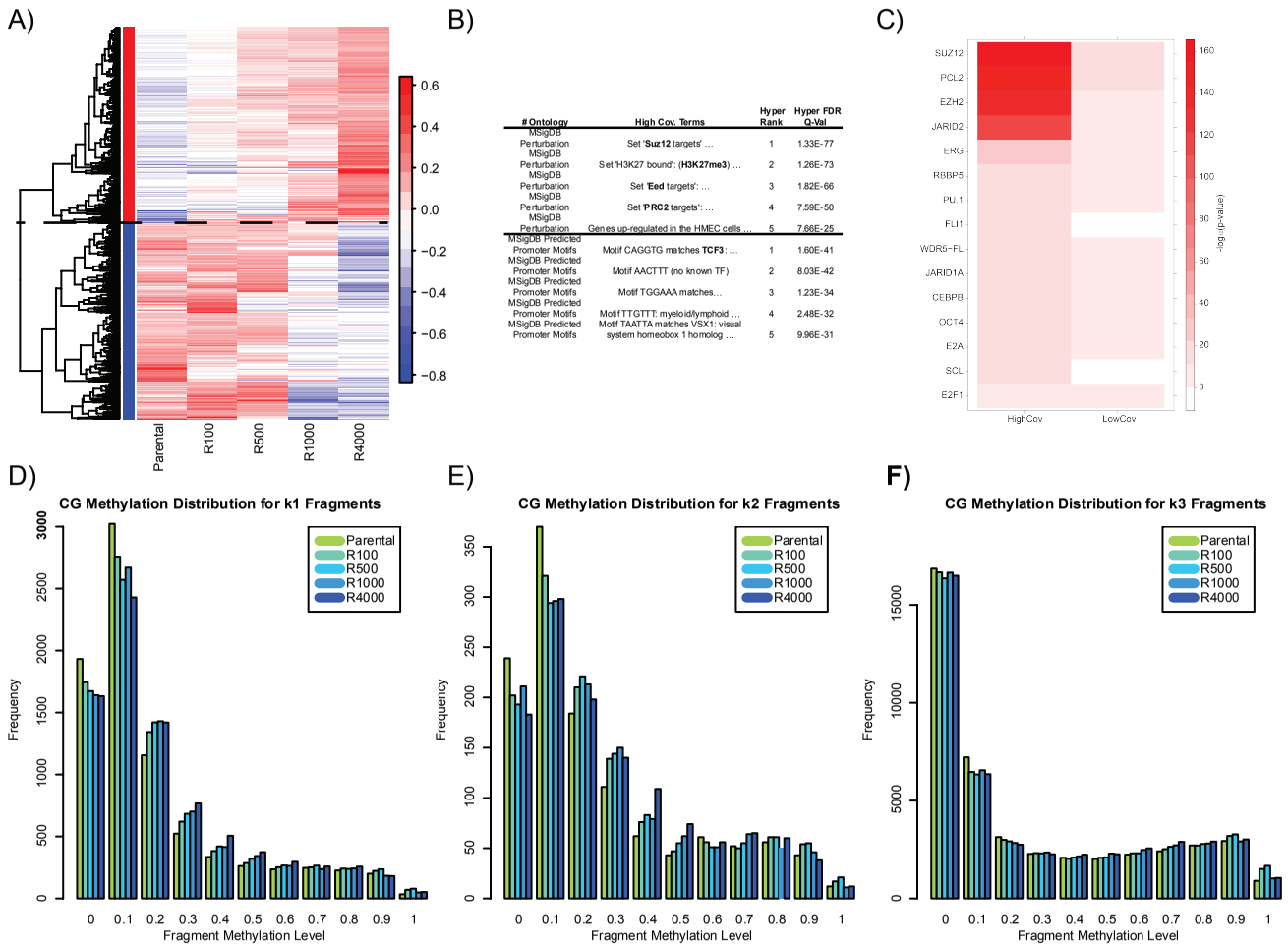


Figure 4. High DNA fragment methylation analysis.

A) Heatmap of the top and bottom 1000 RRBS fragments based on covariance. B) Ontology terms with significant overlap with the high covariance RRBS fragments with an adjusted FDR < 0.05 generated by GREAT. C) A heatmap of the –log10(P) values for the overlap of high and low covariance genes with transcription factor targets. D, E, F) Histogram of the methylation entropy levels for H3K27me3 ChIP-Seq clusters K1, K2, and K3.

significant inverse correlation between the two covariance measures with a Pearson coefficient of -0.646 ($p < 1e-23$), suggesting that many genes that are induced by mafosfamide in the parental line are repressed in the resistant ones (**Figure 5A**). For example Rras2, an oncogene associated with B cell proliferation, is strongly induced by mafosfamide in the parental line, but its expression decreased in the resistant ones (**Figures 5B and C**).

The upstream regulatory analysis of the covarying probe sets also revealed opposite trends between the parental and resistant lines (**Figure 5D**). Functional analysis of gene expression data from the mafaosfamide treated parental line showed an enrichment for target genes of Prc2 component *Suz12* in the Gene Set Enrichment Analysis (GSEA) (**Figure 5E**). Conversely, the same functional analysis of gene expression in response to treatment for the resistant cell lines showed depletion for the same gene set (**Figure 5F**).

Principal Component Analysis (PCA) of basal gene expression indicates alterations in B Cell maturation.

The observation that changes in gene expression, DNA methylation and H3K27me3 in the resistant lines is associated with *E2a* bound genes, a master regulator of B cell development, lead us to hypothesize that gradual incremental dose escalation results in epigenetic changes associated with B cell maturation. PCA of gene expression data obtained from the untreated dose escalated cell lines combined with expression data of B cells at different stages of development gathered from the Immgen dataset[35] showed that the second principal component (PC2) captures the developmental state of B cells (**Figure 6A**), with less differentiated progenitor cells toward the negative direction of the PC.2 axis and the more mature states in the positive. Based on this interpretation of PC.2, we find that our four resistant lines are ordered by their differentiation, with the least resistant line being more differentiated than the most resistant line (from right to left). Furthermore, when the Immgen gene expression data

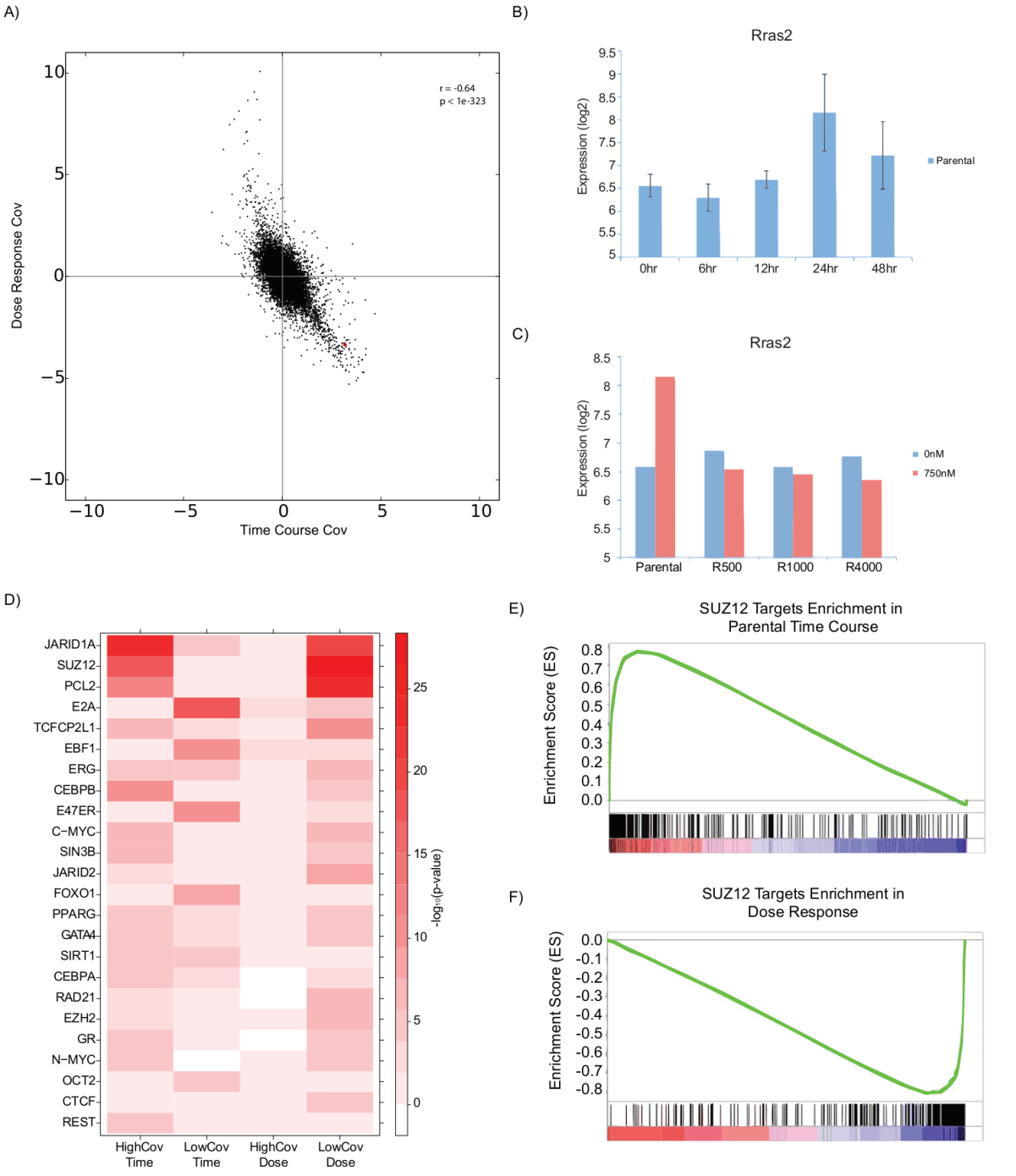


Figure 5. Comparison of resistant and parental gene expression.
A) A scatter plot of the covariance of the dose response of the resistant cells versus the covariance of the parental time course. B) Expression of Rras2 during treatment over time in the parental line. C) Comparison of Rras2 expression in response to mafosfamide treatment between the parental and resistant lines (n=3, mean ± SEM). D) Heatmap of the –log10(P) values of the overlap between TF targets and high and low covariance gene lists. E & F) Functional analysis of gene expression data using GSEA shows an enrichment for Suz12 target genes in the treated parental line, but a depletion in the treated resistant lines respectively.

set is analyzed, those genes that had a positive correlation with development (i.e. an increase in expression during differentiation) showed greater significance in enrichment for targets of *E2a* (**Figure 6B),** supporting its role as a master regulator of B cell development.

Human Diffuse Large B-Cell Lymphoma

Based on the above finding we sought to determine the methylation status of *E2a* bound genes in a clinically relevant dataset. We obtained DNA methylation data from diffuse large B-cell lymphoma patients collected from the Cancer Genome Atlas project. Of 19 total samples we were able to obtain, 14 corresponded to patients who were tumor free while 5 experienced disease progression. For all methylation sites, we computed the Kolmogorov-Smirnov statistic and an associated p-value between these two groups. From 482,421 total CG sites on the Illumina 450K array, 5541 had a KS test with $p < 0.05$. We linked CpG sites to *E2a* (*TCF3*) binding regions if both the CpG and binding region occurred in the gene or within 10 kilobases of the genes transcription start site (TSS). We then identified the intersection of genes that had both evidence of *TCF3* binding and at least 2 CG sites with significant methylation changes between the two patient populations. From a total of 17,744 genes used in the linking calculations, we derived Table 1 (p-value ≈ 0), which shows a strong enrichment of *TCF3* genes associated with differentially methylated sites. This analysis suggests that *E2a* (*TCF3*) bound genes are differentially methylated and are associated with treatment failure in a lymphoma subtype.

Discussion

Through a series of drug escalation co-cultures of *Eμ-Myc Cdkn2a*-/- cells with mafosfamide we developed a series of cell lines that showed increasing resistance to this cyclophosphamide analog. We hypothesized that if a mutation provided the cells a proliferative or survival advantage, then the frequency of that mutation would have increased over time in culture. However, we observed no high

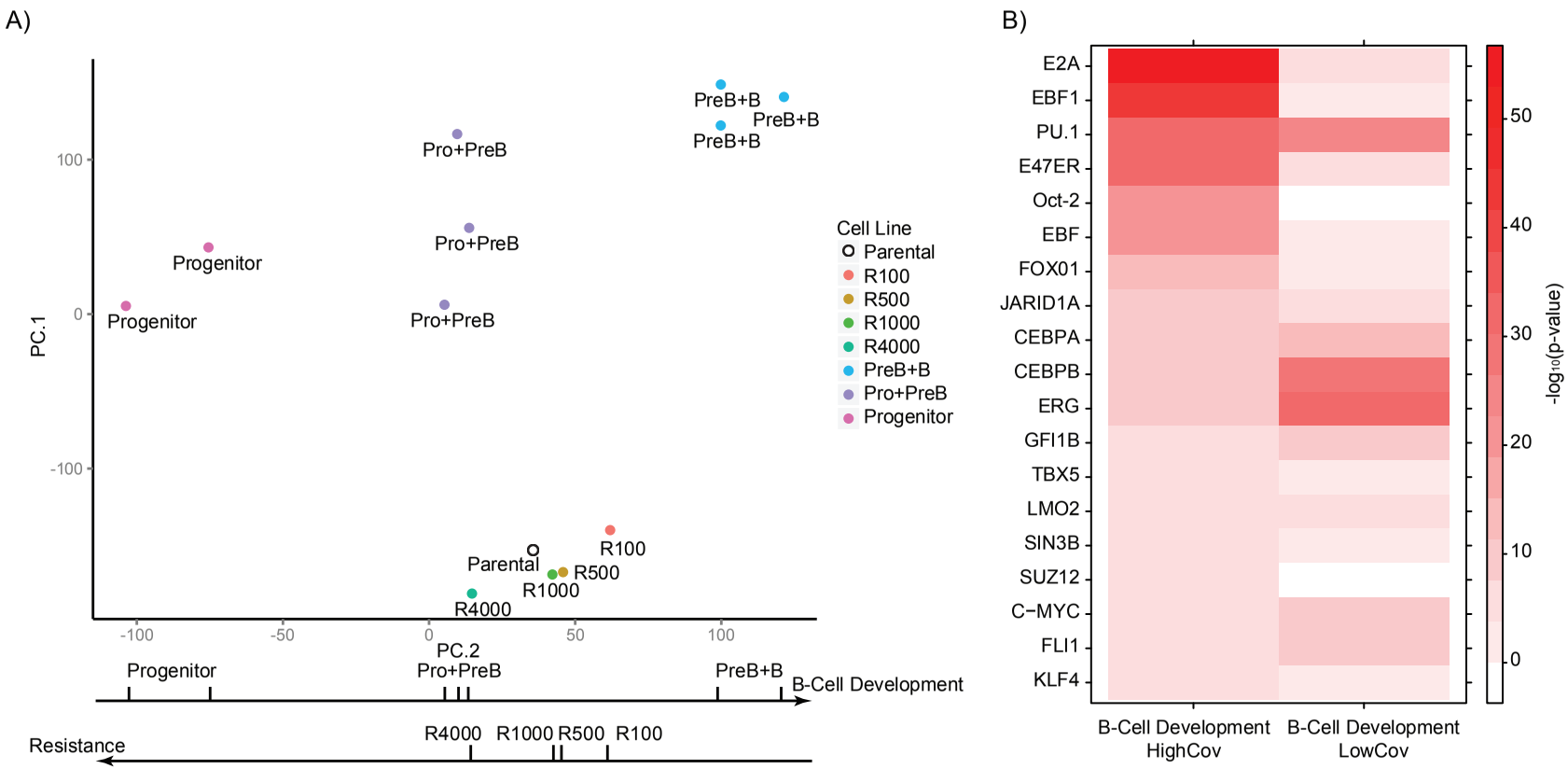


Figure 6. B cell development principal component analysis.
A) Scatter plot of the Principal Components of the Immunological Genome Project samples and the resistant lines. B) Heatmap of the –log10(p) values for the overlap of high and low covariance genes with TF targets.

	TCF3 Target	Not TCF3 Target
Differential Methylation	208	80
No Differential Methylation	5086	12370

Table 1: TCF3 (E2a) target gene methylation.
DNA methylation data obtained from TCGA shows an increase in differential methylation in TCF3 (E2a) target genes in patients with diffuse large B-cell lymphoma who showed stable or progressive disease compared to individuals who had complete response.

confidence mutations that displayed this pattern. Based on these results we concluded that the observed resistance results from epigenetic changes.

Reasoning that the gene expression response to mafosfamide might identify altered processes in the resistant cells we identified alterations in transcript levels that co-varied with resistance. Upstream regulatory analysis of these co-varying genes suggested that they are significantly enriched for binding of Prc2 as well as *E2a*. Prc2 plays a critical role in B cell development by repressing genes necessary for hematopoietic differentiation through the addition of tri-methylation to lysine 27 of histone H3. During hematopoietic stem cell differentiation this repressive mark is lost allowing for the expression of genes that commit hematopoietic stem cells to differentiation[31,32,36]. Numerous studies have shown that over expression of or activating mutations in Prc2 components, particularly Ezh2 contribute to proliferation and lymphomagenesis in diffuse large B cell lymphoma (DLBCL)[37,38]. The pattern we observe of rapid H3K27 methylation followed by demethylation, is in agreement with what has been reported in patients where loss of H3K27me3 is a poor predictor of outcome [39].

To measure the activity of Prc2, we performed ChIP-seq on H3K27me3 in the parental and resistant cell lines. A large fraction of promoters in the resistant lines showed rapid and widespread H3K27me3 acquisition followed by a gradual reduction. These results suggest that Prc2 is rapidly activated by the DNA damage induced by mafosfamide, which is consistent with previous literature suggesting the Prc2 is targeted to sites of DNA damage[40,41]. Furthermore, these results suggest that as cells become adapted to higher doses of mafosfamide, this initial response is attenuated, and the most resistant line returns to a more basal state of H3K27me3.

However, the observation that H3K27me3 is rapidly gained but then lost, suggests that it may not be the primary mechanism leading to resistance. The removal of H3K27me3 at many loci may lead to its replacement with DNA methylation, a more permanent mark. Previous work has shown a possible

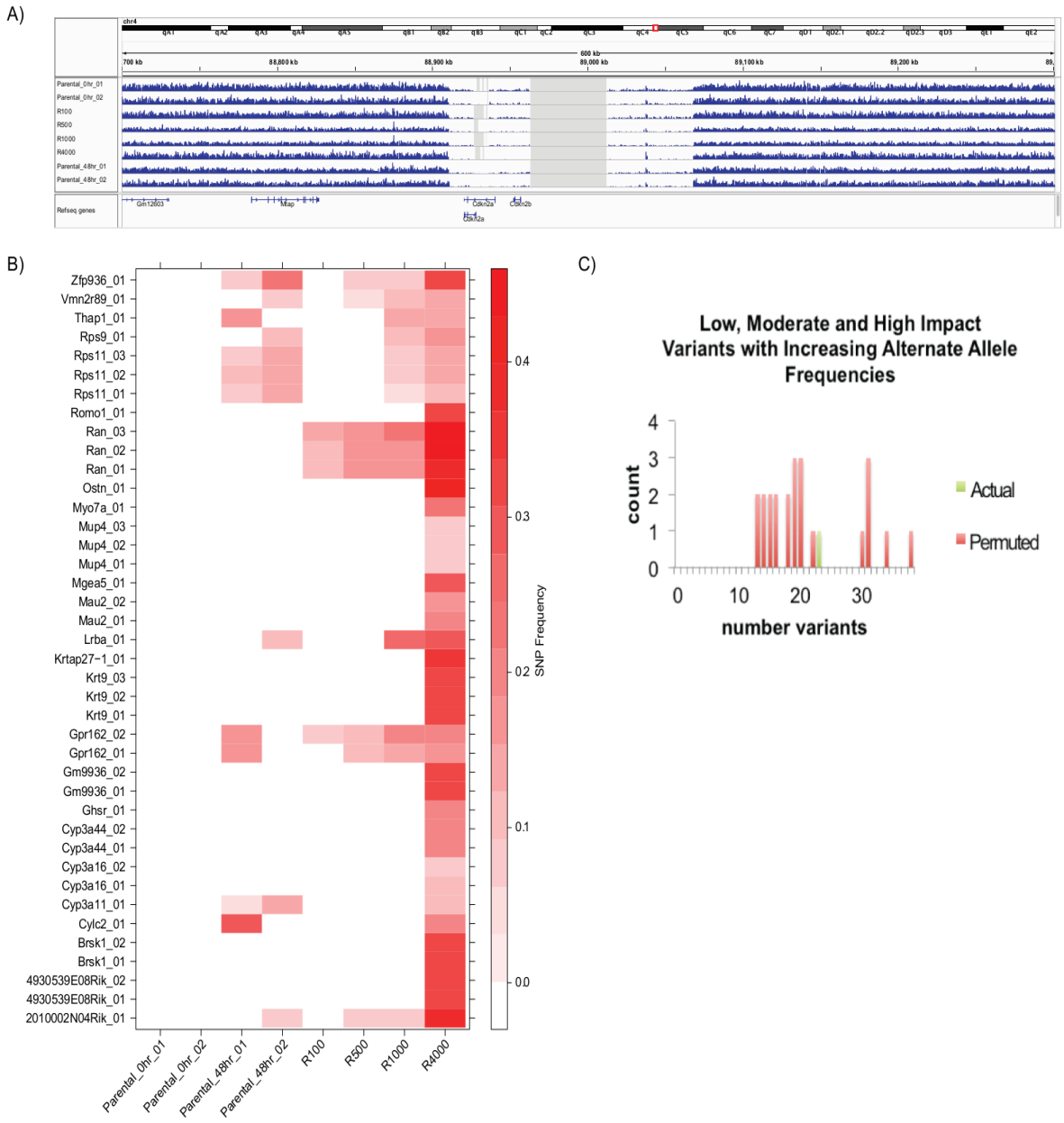
connection between DNA methylation, gene expression, and the Prc2 complex. It has been suggested that histone methylation by Prc2 may recruit *Dnmt3l*, an inactive homolog of DNA methyltransferases, resulting in the inhibition of CpG methylation[33]. Thus the gradual removal of H3K27me3 could lead to the replacement of *Dnmt3l* with its active counterparts, *Dnmt3a* and *Dnmt3b*, leading to an increase in DNA methylation. Because of this potential coupling of Prc2 with DNA methytransferases, we sought to characterize changes in DNA methylation using RRBS. In agreement with this model, we observed that genes that gained H3K27me3 early, and then lost it (cluster K2), showed a gradual increase in DNA methylation. Moreover, fragments whose methylation entropy increases also show the same pattern of increase and then decrease of H3k27me3. Thus, we conclude that the addition of H3k27me3 to these loci perturbs DNA methylation, and increases the diversity of methylation patterns found there. One possible explanation for this effect, is that H3k27me3 often affects DNA methylation on only one of the two chromosomes, in a stochastic fashion, generating allelic methylation, a state of inherently high entropy.

Clustering of the H3K27me3 ChIP-seq data also showed that the K2 cluster is strongly enriched for *E2a* target genes and that *E2a* is also strongly associated with those genes that gain and subsequently lose H3K27me3, while gradually gaining DNA methyaltion. *E2a* encodes two proteins, E12 and E47 that are known to be master regulators involved in the process of B cell lineage commitment[34]. PCA analysis of gene expression from the resistant lines and from various stages of B cell development indicated that the expression of *E2a* target genes is strongly correlated with B cell maturation states. Early in B cell development the *E2a* locus is repressed but becomes transcriptionally active during B cell commitment. *E2a*'s central role in Burkitt's lymphoma pathogenesis is made evident by the fact that it is the fifth most mutated gene in Burkitt's lymphoma with all mutations affecting its DNA binding domain[42]. Depletion of *E2a* using siRNA has been shown to lead to lower *Cdkna1* (*p21*) accumulation and higher PUMA expression, leading to impairment of cell cycle arrest and increased *Tp53*-dependent

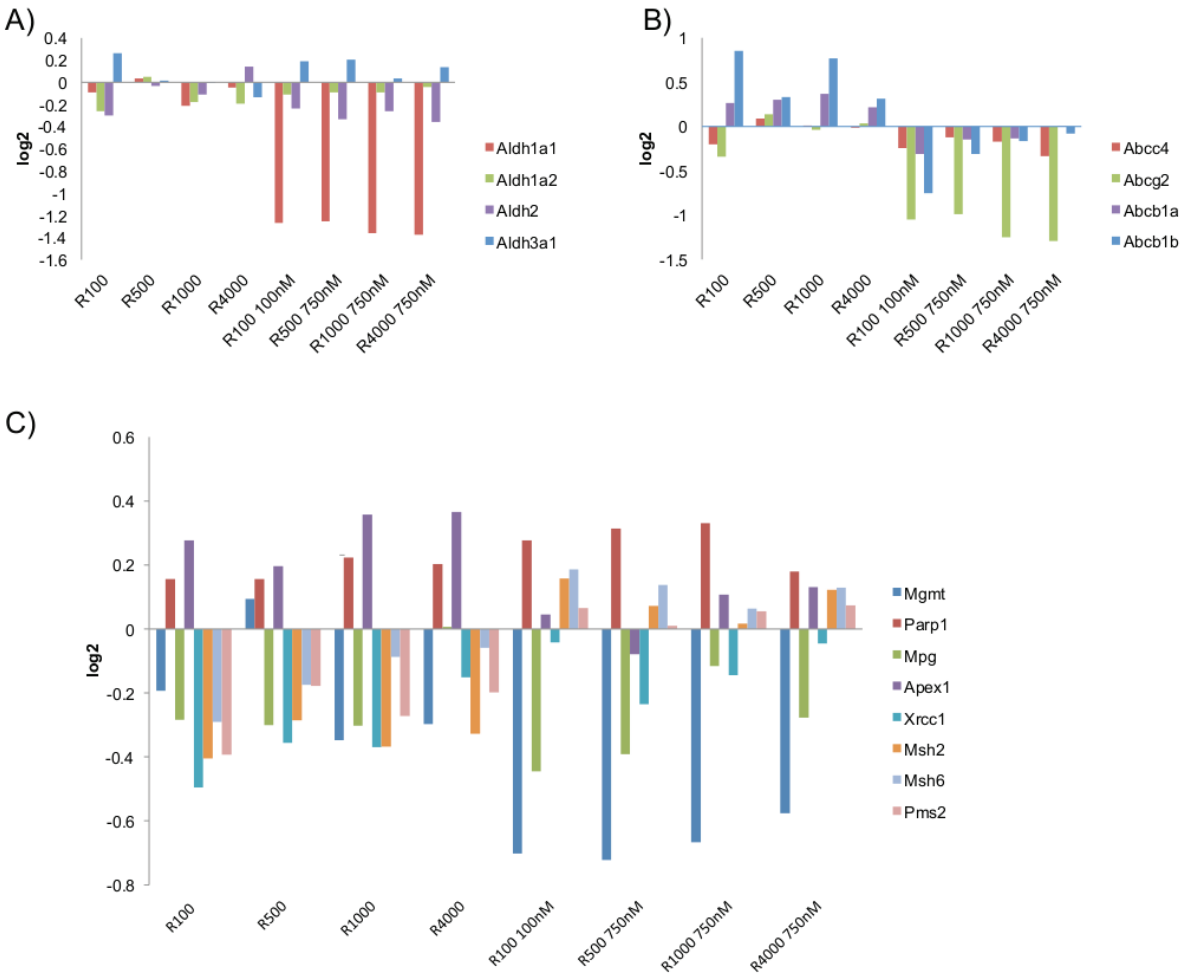
apoptosis[43]. Our data suggests that *E2a* may be playing a critical role in mediating the resistance phenotype, as it appears to be a key regulator of the mafosfamide response in the parental line.

Taken together, our results suggest that Prc2 and *E2a* activity may be responding to mafosfamide to induce epigenetic changes in the transcription of key hematopoietic developmental and pluripotency genes. We propose that upon treatment with mafosfamide, our lines undergo rapid and widespread acquisition of the repressive mark H3K27me3 and that this methylation is coupled to changes in the activity of *E2a*, leading to changes in transcription of genes involved in B cell development. As the cells become more resistant, the H3K27me3 response is attenuated, and replaced by the more permanent mark, 5-methylcytosine. We further speculate that these changes lead the *Eμ-Myc* cell lines to revert to an earlier developmental state that suppresses the apoptotic response upon exposure to mafosfamide. This reversion is supported by our integrative analysis of Immgen data that captures a principal component of the B cell developmental axis. As our cells become more resistant to mafosfamide, they appear to move backwards along this axis, suggesting they are becoming more stem like as they become more resistant. The diminished apoptotic response of stem cells compared to differentiated cells has been described in many other systems, and may thus represent a critical mechanism for the acquisition of drug resistance[44,45].

Finally, we have shown that similar mechanisms may also be at play in human lymphomas. Our analysis of diffuse large B-cell lymphoma suggests that there are epigenetic differences between patients that remain disease free versus those that have disease progression. Moreover, these sites of differential methylation are significantly associated with genes that are bound by Tcf3. Thus our murine model may be capturing epigenetic resistance mechanisms that are also present in human disease, suggesting that epigenetic plasticity may impair therapeutic regimens.

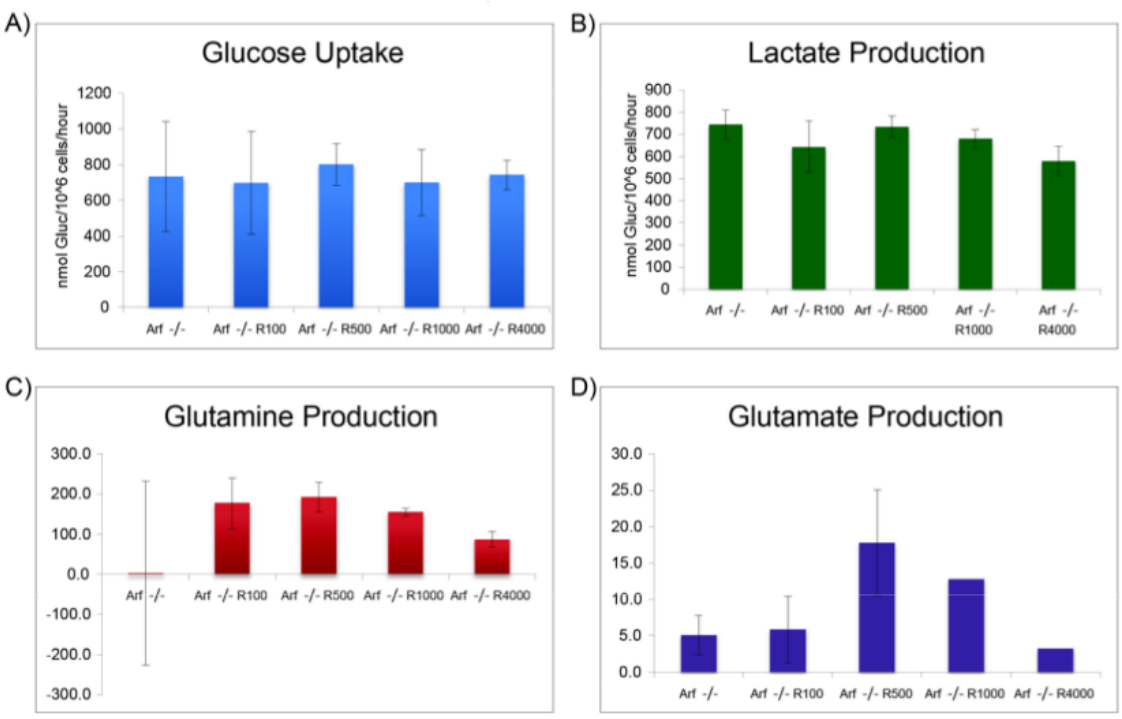


Supplemental Figure 1: Whole genome sequencing of parental and resistant lines. A) Cdkn2a deletion in parental and resistant lines. B) Single nucleotide variant (SNV) allele frequency in parental line in the presence and absence of treatment, and resistant lines. C) Stastical anlaysis of SNV allele frequency.

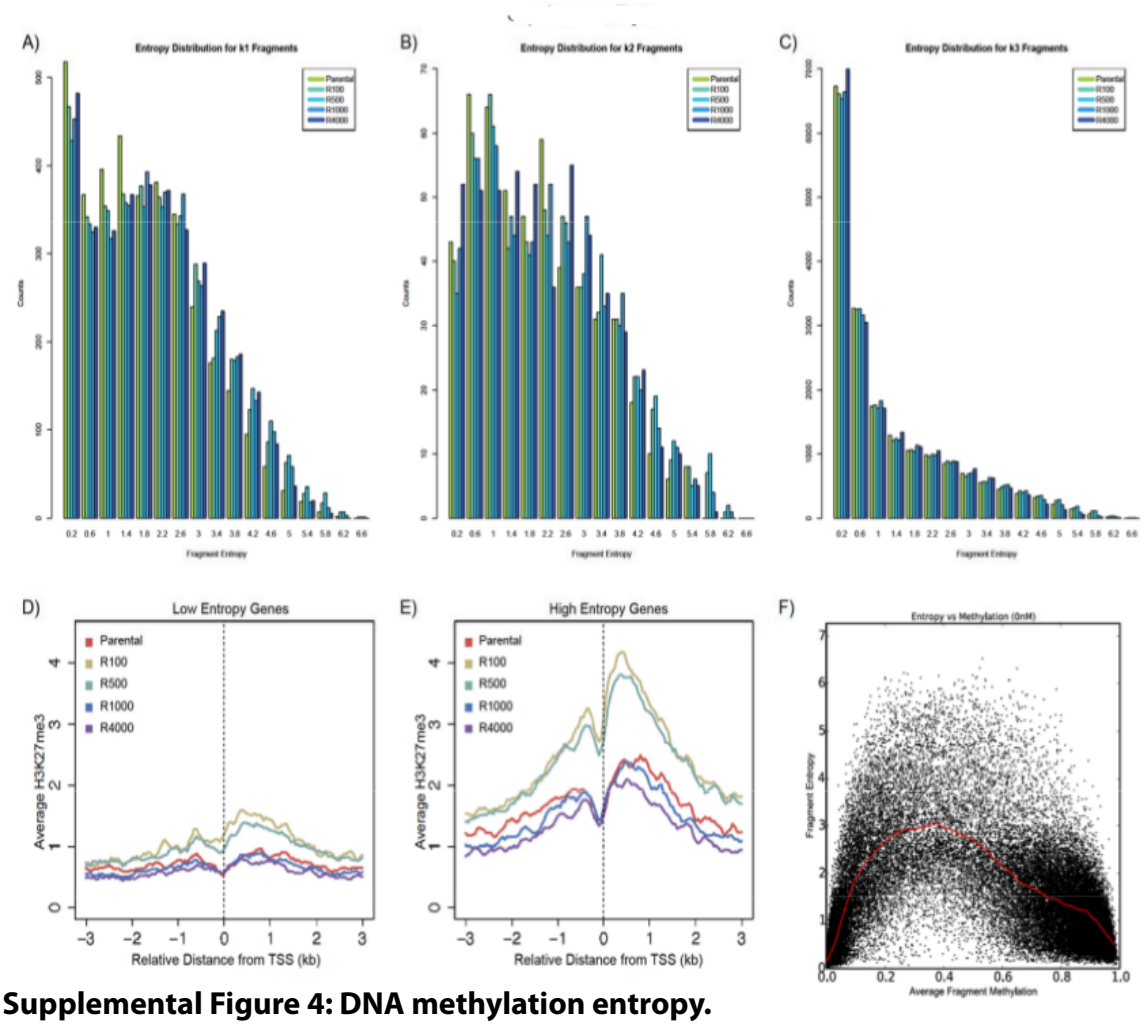


Supplemental Figure 2: Gene expression of genes known to be involved in drug resistance

A) Gene expression of genes known to be involved in mafosfamide metabolism, expression relative to parental line. B) Gene expression of genes known to be involved in mafosfamide export, expression relative to parental line. C) Gene expression of genes known to be involved in DNA repair and apoptosis, expression relative to parental line.

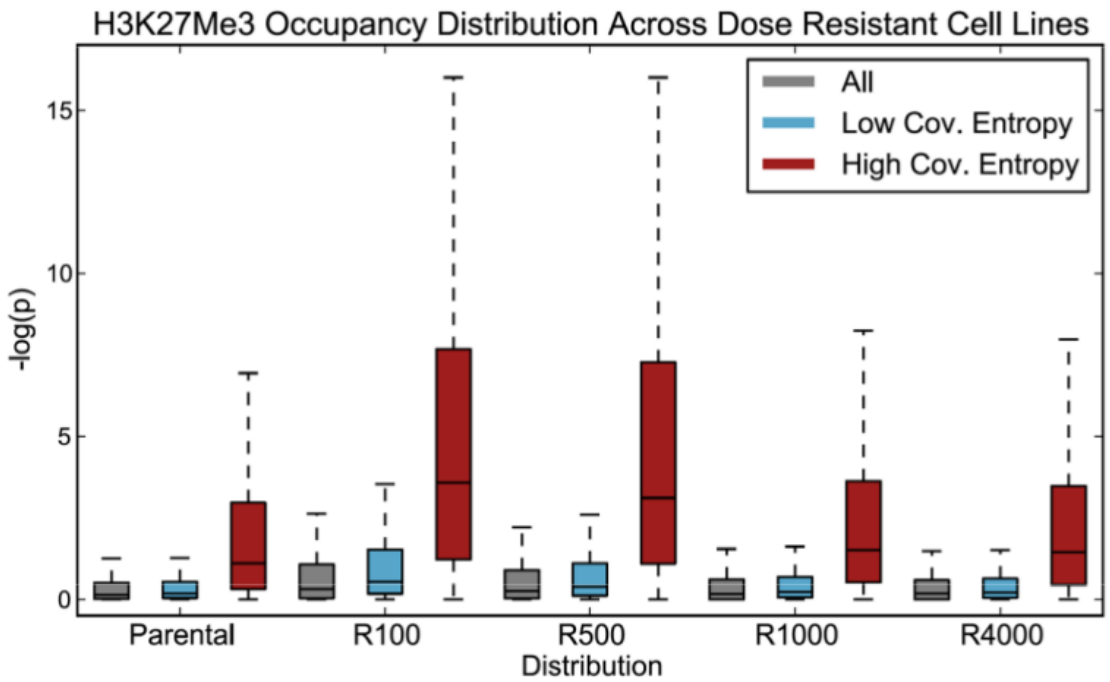


Supplemental Figure 3: Metabolomic analysis of parental and resistant lines. A) Glucose uptake in parental and resistant cell lines. B) Lactate production in parental and resistant cell lines. C) Glutamine production in parental and resistant cell lines. D) Glutamate production in parental and resistant cell lines.



Supplemental Figure 4: DNA methylation entropy.

A) DNA entropy distribution of K1 ChIP-seq cluster fragments. B) DNA entropy distribution of K2 ChIP-seq cluster fragments. C) DNA entropy distribution of K3 ChIP-seq cluster fragments. D) Average H3K27me3 level of low entropy methylation fragments. E) Average H3K27me3 level of high entropy methylation fragments. F) Entropy versus DNA methylation level.



Supplemental Figure 5: H3K27me3 occupancy across DNA methylation fragment entropy.

Significance of H3K27me3 occupancy of high and low DNA methylation fragment entropy levels across parental and resistant lines.

Cell Line	Treatment	Non-Apoptotic			% Apoptotic
		% G1	% G2	% S	
Eu-Myc Parental	0nM	29.22	8.86	61.92	1.36
Eu-Myc Parental	750nM	21.63	17.57	69.79	81.47
Eu-Myc R500	0nM	27.49	8.25	64.26	0.78
Eu-Myc R500	750nM	23.39	15.72	60.89	46.5
Eu-Myc R1000	0nM	29.62	7.2	63.18	1.03
Eu-Myc R1000	750nM	22.83	16.55	60.55	36.3
Eu-Myc R4000	0nM	29.93	9	61.07	0.1
Eu-Myc R4000	750nM	22.06	19.27	58.67	3.26

Supplmental Table 1:
Cell cycle analysis of parental and resistant lines in the presence and absence of mafosfamide.

References

1. Weigert O, Weinstock DM. The evolving contribution of hematopoietic progenitor cells to lymphomagenesis. *Blood*. 2012;120(13):2553-2561.

2. Mumenthaler SM, Foo J, Leder K, et al. Evolutionary modeling of combination treatment strategies to overcome resistance to tyrosine kinase inhibitors in non-small cell lung cancer. *Mol Pharm*. 2011;8(6):2069-2079.

3. Choi MK, Jun HJ, Lee SY, et al. Treatment outcome of adult patients with Burkitt lymphoma: results using the LMB protocol in Korea. *Ann Hematol*. 2009;88(11):1099-1106.

4. Richter-Larrea JA, Robles EF, Fresquet V, et al. Reversion of epigenetically mediated BIM silencing overcomes chemoresistance in Burkitt lymphoma. *Blood*. 2010;116(14):2531-2542.

5. Holohan C, Van Schaeybroeck S, Longley DB, Johnston PG. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer*. 2013;13(10):714-726.

6. Wilting RH, Dannenberg JH. Epigenetic mechanisms in tumorigenesis, tumor cell heterogeneity and drug resistance. *Drug Resist Updat*. 2012;15(1-2):21-38.

7. Knoechel B, Roderick JE, Williamson KE, et al. An epigenetic mechanism of resistance to targeted therapy in T cell acute lymphoblastic leukemia. *Nat Genet*. 2014;46(4):364-370.

8. Chambwe N, Kormaksson M, Geng H, et al. Variability in DNA methylation defines novel epigenetic subgroups of DLBCL associated with different clinical outcomes. *Blood*. 2014.

9. Hiraga J, Tomita A, Sugimoto T, et al. Down-regulation of CD20 expression in B-cell lymphoma cells after treatment with rituximab-containing combination chemotherapies: its prevalence and clinical significance. *Blood*. 2009;113(20):4885-4893.

10. Guan H, Xie L, Klapproth K, Weitzer CD, Wirth T, Ushmorov A. Decitabine represses translocated MYC oncogene in Burkitt lymphoma. *J Pathol*. 2013;229(5):775-783.

11. Valdez BC, Nieto Y, Murray D, et al. Epigenetic modifiers enhance the synergistic cytotoxicity of combined nucleoside analog-DNA alkylating agents in lymphoma cell lines. *Exp Hematol*. 2012;40(10):800-810.

12. Murawski N, Pfreundschuh M. New drugs for aggressive B-cell and T-cell lymphomas. *Lancet Oncol*. 2010;11(11):1074-1085.

13. Sharma SV, Lee DY, Li B, et al. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell*. 2010;141(1):69-80.

14. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*. 2012;481(7382):506-510.

15. Aryee MJ, Liu W, Engelmann JC, et al. DNA methylation alterations exhibit intraindividual stability and interindividual heterogeneity in prostate cancer metastases. *Sci Transl Med*. 2013;5(169):169ra110.

16. Blanpain C, Mohrin M, Sotiropoulou PA, Passegue E. DNA-damage response in tissue-specific and cancer stem cells. *Cell Stem Cell*. 2011;8(1):16-29.

17. Marconett CN, Zhou B, Rieger ME, et al. Integrated transcriptomic and epigenomic analysis of primary human lung epithelial cell differentiation. *PLoS Genet*. 2013;9(6):e1003513.

18. Massague J. TGFbeta signalling in context. *Nat Rev Mol Cell Biol*. 2012;13(10):616-630.

19. Arumugam T, Ramachandran V, Fournier KF, et al. Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer. *Cancer Res*. 2009;69(14):5820-5828.

20. Roxanis I. Occurrence and significance of epithelial-mesenchymal transition in breast cancer. *J Clin Pathol*. 2013;66(6):517-521.

21. Schmitt CA, McCurrach ME, de Stanchina E, Wallace-Brodeur RR, Lowe SW. INK4a/ARF mutations accelerate lymphomagenesis and promote chemoresistance by disabling p53. *Genes Dev*. 1999;13(20):2670-2677.

22. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.

23. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 2003;19(2):185-193.

24. Yang JH, Li JH, Jiang S, Zhou H, Qu LH. ChIPBase: a database for decoding the transcriptional regulation of long non-coding RNA and microRNA genes from ChIP-Seq data. *Nucleic Acids Res*. 2013;41(Database issue):D177-187.

25. Ferrari R, Su T, Li B, et al. Reorganization of the host epigenome by a viral oncogene. *Genome Res*. 2012;22(7):1212-1221.

26. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res*. 2005;33(18):5868-5877.

27. Guo W, Fiziev P, Yan W, et al. BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics*. 2013;14(1):774.

28. Maggi LB, Jr., Winkeler CL, Miceli AP, et al. ARF tumor suppression in the nucleolus. *Biochim Biophys Acta*. 2014;1842(6):831-839.

29. Panasci L, Xu ZY, Bello V, Aloyz R. The role of DNA repair in nitrogen mustard drug resistance. *Anticancer Drugs*. 2002;13(3):211-220.

30. Sarkaria JN, Kitange GJ, James CD, et al. Mechanisms of chemoresistance to alkylating agents in malignant glioma. *Clin Cancer Res*. 2008;14(10):2900-2908.

31. Aloia L, Di Stefano B, Di Croce L. Polycomb complexes in stem cells and embryonic development. *Development*. 2013;140(12):2525-2534.

32. Majewski IJ, Ritchie ME, Phipson B, et al. Opposing roles of polycomb repressive complexes in hematopoietic stem and progenitor cells. *Blood*. 2010;116(5):731-739.

33. Neri F, Krepelova A, Incarnato D, et al. Dnmt3L antagonizes DNA methylation at bivalent promoters and favors DNA methylation at gene bodies in ESCs. *Cell*. 2013;155(1):121-134.

34. Murre C. Regulation and function of the E2A proteins in B cell development. *Adv Exp Med Biol*. 2007;596:1-7.

35. Heng TS, Painter MW. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9(10):1091-1094.

36. Majewski IJ, Blewitt ME, de Graaf CA, et al. Polycomb repressive complex 2 (PRC2) restricts hematopoietic stem cell activity. *PLoS Biol*. 2008;6(4):e93.

37. Velichutina I, Shaknovich R, Geng H, et al. EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis. *Blood*. 2010;116(24):5247-5255.

38. Yap DB, Chu J, Berg T, et al. Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood*. 2011;117(8):2451-2459.

39. Wei Y, Xia W, Zhang Z, et al. Loss of trimethylation at lysine 27 of histone H3 is a predictor of poor outcome in breast, ovarian, and pancreatic cancers. *Mol Carcinog*. 2008;47(9):701-706.

40. Campbell S, Ismail IH, Young LC, Poirier GG, Hendzel MJ. Polycomb repressive complex 2 contributes to DNA double-strand break repair. *Cell Cycle*. 2013;12(16):2675-2683.

41. Chou DM, Adamson B, Dephoure NE, et al. A chromatin localization screen reveals poly (ADP ribose)-regulated recruitment of the repressive polycomb and NuRD complexes to sites of DNA damage. *Proc Natl Acad Sci U S A*. 2010;107(43):18475-18480.

42. Schmitz R, Young RM, Ceribelli M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature*. 2012;490(7418):116-120.

43. Andrysik Z, Kim J, Tan AC, Espinosa JM. A genetic screen identifies TCF3/E2A and TRIAP1 as pathway-specific regulators of the cellular response to p53 activation. *Cell Rep*. 2013;3(5):1346-1354.

44. Bakker ST, Passegue E. Resilient and resourceful: genome maintenance strategies in hematopoietic stem cells. *Exp Hematol*. 2013;41(11):915-923.

45. Zhou BB, Zhang H, Damelin M, Geles KG, Grindley JC, Dirks PB. Tumour-initiating cells: challenges and opportunities for anticancer drug discovery. *Nat Rev Drug Discov*. 2009;8(10):806-823.

Chapter 3: Combined Transcriptomic-Proteomic Analysis of Drug Response Reveals a Role for Uba1 in Drug Resistance

Introduction

To date, transcriptomic analysis has been used to define gene expression signatures that differentiate between cancer subpopulations and guide the design of personalized treatment regimen[1,2]. However, changes in gene expression are not always capable of discerning two cancer subtypes from one another, or even more importantly discerning two cancer cells from the same tumor that have differential response to treatment[3,4]. Since proteins are the effectors of genes and are ultimately responsible for a given phenotype protein expression signatures can similarly differentiate between cancer cell subpopulations. The proteome is a dynamic system, whose expression levels are determined by a myriad of factors including transcription and translation rates as well as protein degradation[5]. Due to this extensive post-transcriptional regulation gene expression shows poor correlation with protein expression[5]. While some genes are controlled primarily at the level of transcription and can be explained by simple mathematical equations, many are regulated in a more complex fashion which makes them less amenable to this type of analysis[6]. As such it is important to understand at which stage of regulation differences in mRNA and protein levels arise since it would be indicative of post-transcriptional regulation.

Eighty percent of all BL cases in humans harbor t(8;14) translocation which drives c-Myc expression and this translocation is a hallmark of Burkitt's lymphoma[7]. Myc is a transcription

factor that controls a wide spectrum of pathways involved in carcinogenesis and drug resistance including the control of proliferation, metabolism and apoptosis by binding to the enhancers of active genes amplifying their expression[8,9]. Myc is also one of the most frequently deregulated oncogenes[9]. Myc binds to the enhancers of active genes and amplifies their expression resulting in elevated levels of cellular RNA in c-Myc over expressing cells. Additionally, approximately 13% of Myc bound promoters have a direct role in ribosome biogenesis and protein synthesis including translation initiating and elongation factors[10]. Transgenic mice over expressing Myc in B lymphocytes have been shown to have an increase in total protein synthesis [11]. Myc over expressing lymphomas become oncogene addicted and inhibition of Myc through the induction of dominant-negative, recombinant Myc has lead to significant inhibition in cancer progression with limited toxicities[12]. Similarly, silencing of Myc with shRNA resulted in suppressed proliferation and increased apoptosis[13]. Pharmacological inhibition of myc represents a rational approach to therapy in Myc over expressing cancers. However, due to the nature of its protein binding domains which are large and do not contain clefts or pockets that could bind small molecules development of small molecule inhibitors of Myc is difficult[9]. One potential way side step Myc inhibition would be to inhibit pathways downstream of Myc potentially phenocopying Myc inhibition itself.

To do so, we sought to characterize the proteome and transcriptome dynamics in response to DNA damage in a mouse BL cell line that contains a translocation of c-Myc placing it under the control of the IgH enhancer (Eμ-Myc). Additionally, the two lines being tested contain mutations in either Cdkn2a or Tp53 leading to the inactivation Myc over expressed induced apoptosis. Both lines give rise to lymphomas that form rapidly and display apoptotic

defects *in vivo* [14]. BL is treated with high intensity, short duration regimens of a combination of cytotoxic agents including the monoclonal antibody rituximab, cyclophosphamide, hydroxydaunorubicin (doxorubicin), oncovin, and prednisone (R-CHOP) and is able to achieve complete response rates between 70% and 90% of patients, with long term event-free survival rates between 45% and 97%[15]. Due to the intense nature of the therapeutic regime in BL treatment drug toxicity is a major cause of mortality and morbidity. In one study, only 32% of patients older than 50 years of age were able to complete 6 to 7 treatment cycles compared with 79% of younger patients, with increases in mortality, disease progression and toxicity rates in those individuals older than 50 years[16].

Resistance to therapy is a major problem for the successful treatment of BL since a large percentage of those individuals who do not achieve remission with the standard therapy regime are unlikely to do so with salvage treatment[17,18]. Tp53 is one of the most frequently mutated genes in Burkitt's lymphoma[18]. Inactivation of the Tp53 pathway is frequently disrupted in Myc over expressing malginacies[18,19] since over expression of Myc triggers Tp53-dependent apoptosis[20,21]. Tp53 is a tumor suppressor gene that integrates signals from various stressor indicators such as DNA damage and nutrient stress[22]. Normally present in the cytosol, upon phosphorylation p53 translocates to the nucleus in response to DNA damage[22]. Exposure to DNA-alkylating agents is suggested to be responsible for the development of mutation in Tp53 and resistance to chemotherapy[23]. However, despite Tp53's role in DNA damage response and maintaining genome integrity, the prognostic value of Tp53 is of limited use in several malignancies including lymphoma[24].

Methods

Cell Culture

Eμ-Myc Cdkn2a-/- and *Tp53-/-* lymphoma lines were generated from a C57BL/6J mouse as described in Schmitt, *et. al.* 1999[14]. Lymphoma cells from these mice were cultured in 50% DMEM (Invitrogen, Waltham, MA, USA) and 50% IMEM (Invitrogen, Waltham, MA, USA), with 10% FBS, 1% Pen-Strep, 4mM L-Glutamine and 50μM β-mercaptoethanol. *Eμ-Myc* lines were cultured on a feeder layer of irradiated *Cdkn2a-/-* mouse embryonic fibroblasts.

Cell Viability Measurements

Cell viability was measured using the Perkin-Elmer Operetta platform, and 2.5μM Draq5 (Abcam, Cambridge, Ma) for nuclear detection and 5 μg/mL of propidum iodide (Sigma-Aldrich, St. Louis, Mo) to detect dead cells. Viability was measured after 24 hours in the prescence of either mafosfamide (Abcam, Cambridge, England), Pyr-41 (Cell Signaling Technologies, Danvers, MA, USA), Bortezomib (Selleck Chemicals, Houston, Tx, USA), or Chloroquine (Sigma Aldirch, St. Louis, MO, USA).

Microarray

Oligonucleotide microarray analysis was carried out using GeneChip Mouse Gene ST 1.0 (Affymetrix, Santa Clara, Ca). The Affymetrix Mouse Gene ST 1.0 Array expression measures were quantified and processed with robust multi-array average using the justRMA function of the 1.40.0 affy R package (RMA)[25]. Expression values were $\log_2$ transformed for further downstream analysis. Probe sets were annotated using the Affymetrix MoGene-1_0-st-

v1.na33.2.mm9.probeset.csv file. We selected the top 1000 probe sets ranked by their covariance to identify up regulated genes across the time course or resistant cell lines (See supplemental methods).

LC-MS/MS and Protein Identification

Samples were lysed in RIPA buffer with HALT protease inhibitor (Invitrogen, Waltham, MA, USA) and lysates were cleared by centrifugation at 14,000g for 10 minutes at 4°C. Protein concentration was measured using the bicinchonimic acid assay (BCA). 100µg of protein was then reduced, alkylated, trypsin digested and isobarically labeled using Tandem Mass Tags (TMTs) (Thermo Scientific, Waltham, MA, USA). according to the Thermo Scientific protocol[26]. Samples were then pooled, lyophilized, and resuspended in 50µl 10mM KH2PO4, 25% Acn, pH (SCX Buffer A). Samples were then separated using a 100 x 2.1mm 5um 200Å polySULFOETHYL A strong cation exchange column (PolyLC, Columbia, MD, USA) with a flow rate of 200µl/minute with a gradient of 10mM KH2PO4, 500mM KCl, 25% Acn, pH 2.85 (Buffer B) of 100% for 30 minutes, followed by a linear gradient to 50% B in 50 minutes, followed by a linear gradient to 100% B 10 minutes. Fractions were collected every minute and poolled into 26 total fractions which were lyophilized and resuspended in 200µl of 95% H2O, 5% acetonitrile, 0.1% formic acid (Buffer A). 5µl of sample were injected on to an Agilent Zorbax SB C18 150 x 0.3 mm 5µm 200Å column. The samples were eluted with a linear gradient from 5% Buffer A to 65% Buffer B over 60 minutes, followed by a linear gradient from 65% B to 95% B 5 minutes.

An Orbitrap XL (Thermo Fisher Scientific) mass spectrometer was used to collect peptide masses. The mass spectrometer was operated in positive ion mode with an acquisition time of 180

minutes. A hybrid data acquisition method was used where both a CID and HCD scan were acquired for the six most abundant precursor ions, excluding singly charged ions. Monoisotopic precursor selection and FT master scan preview mode were enabled. For the CID scan, normalized collision energy (CE) was set to 35, activation time was 10 ms, and activation *Q* was 0.25. For the HCD scan, CE was set to 45, and HCD spectra were recorded at a resolution of 7500 (at 400 *m*/*z*) starting at 100 *m*/*z*.

Proteome Discoverer 1.2 was used for the analysis of mass spectrometry data. Ion trap CID spectra were searched with the SEQUEST version implemented in Proteome Discoverer 1.2. Settings for peptide identification were: precursor mass tolerance, 8 ppm; fragment mass tolerance, 0.6 Da; trypsin allowing max. two missed cleavage sites; static modifications, TMT 6plex (peptide N-terminus and Lys) and methylthione modification of Cys residues; dynamic modifications, phosphorylation (Tyr, Ser), oxidation (Met), deamidation (Asn, Gln) and TMT 6plex (Tyr) with a maximum of 4 modifications per peptide.

Analysis of Variance

Differentially expressed genes across omic, cell type, and time were selected by calculating their statistical significance using an ANOVA model, with an error term controlling for the hierarchy of the experiment design. We specified the hierarchical variance structure (across technical replicates and time points) by including an *Error* term in the model. The 50 most significant genes were selected as those with the smallest (most significant) p-value of the interaction coefficient between omic, cell type, and time. These genes were then used to perform group set analysis.

Group set analysis

Functional categorization and classification of differentially expressed genes was performed according to functional annotation data, including the Gene Ontology annotation database (Biological Process), based on group set (enrichment) analysis, as implemented in the R/topGO package. 1,246 GO terms that contained at least 5 of the ~600 genes available in our measurements were ranked according to their strength of association with the selected genes using Fisher's one sided test.

SDS-PAGE and Western Blots

Protein samples were prepared by mixing the protein solution with sample buffer (Life Technologies) followed by denaturation of the sample at 90 °C for 10 minutes. Protein samples were loaded on the precast 4-20% polyacrylamide gel (Criterion TGX Precast Gels, BioRad, Hercules, CA, USA) next to 15 µl of the SeeBlue Plus2 pre-stained standard (Life Technologies). Protein electrophoresis was performed in running buffer (100mM HEPES, 100mM Tris-base, 1% SDS) using 120 V at room temperature. Proteins were transferred to nitrocellulose membrane using the Trans-Blot Turbo Transfer System (BioRad, Hercules, CA, USA) using the standard protocol.

After protein transfer, the nitrocellulose membranes were incubated for 1 hour at room temperature 5% milk PBS-T blocking buffer. Membranes were incubated with primary antibodies diluted in blocking buffer overnight at 4 °C. Membranes were then washed with PBS-T 3 times for 10 minutes. Membranes were incubated with HRP-conjugated secondary antibodies (anti-mouse or anti-rabbit, Life Technologies) for 1 hour at room temperate followed by washing with PBS-T 3 times for 10 minutes. Proteins were detected by incubating the membrane for 5 minutes in SuperSignal West Pico Chemiluminescent substrate (Thermo

Scientific). Primary antibodies used were: α-Psmb1 (Abcam), α-Tubulin (Life Tecnologies), α-Ubiquitin (Abcam), α-Uba1 (Cell Signaling Technologies), and α-Fkbp3 (Abcam).

Results

Temporal transcriptomic and proteomic sample generation

Eμ-Myc Tp53-/- cells are significantly more resistant to chemotherapies including mafosfamide (a derivative of cyclophosphamide) compared to Cdkn2a-/- cells *in vitro* **(Figure 1A)**. Cdkn2a-/- and Tp53-/- Eμ-Myc lymphoma cells were cultured in the presence of the mafosfamide for 48 hours and samples from each line were collected at the time of drug addition (0 hrs) as well as 6, 12, 24 and 48 hours post drug addition (**Figure 1B**). Each sample was split and half the sample was used for transcriptomic analysis **(see Methods)** and the other half was used for quantitative proteomic analysis **(see Methods)(Figure 1B).** For isobaric peptide labeling, 100ng of protein from each time point was trypsinized and labeled with a specific TMT tag **(Figure 1B)**. The samples from each cell line time course study were mixed and separated into 20 fractions using strong cation exchange (SCX) liquid chromatography. SCX fractions were further separated using reverse-phase liquid chromatography **(Figure 1B)** in line to the tandem mass spectrometer to obtain spectra for peptide sequence determination.

After performing the time series experiments in biological triplicate for each cell line, a total of 2694 proteins were identified by at least one peptide with high confidence of correct sequence assignment. The three Cdkn2a-/- biological triplicates resulted in a total of 2,137 proteins identified, with 1065 proteins common amongst all three replicates **(Supplmental Figure 1A)**, while the three Tp53-/- biological replicates yeilded 1,595 proteins with 675

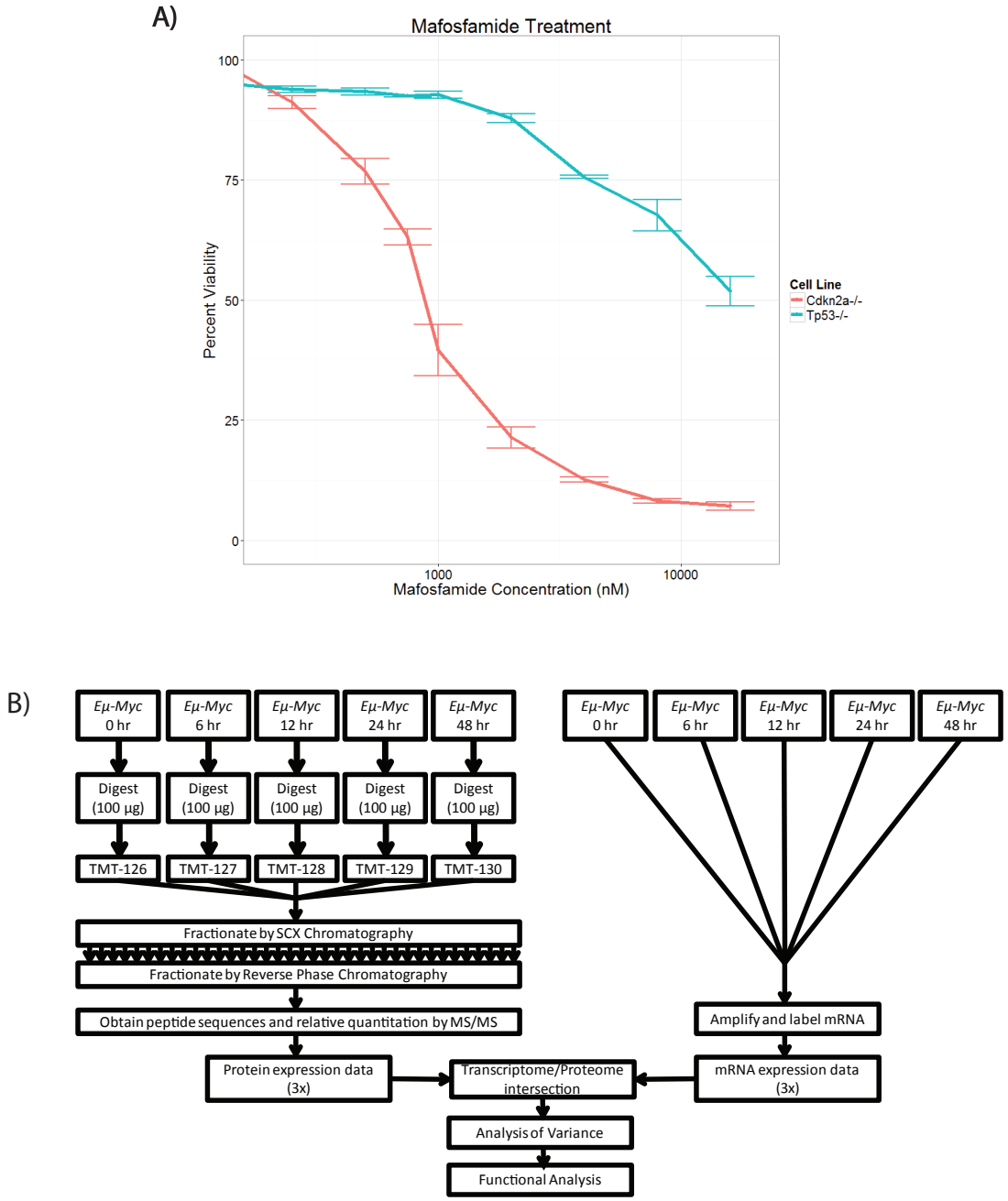


Figure 1: EC-50 of Eμ-Myc cell lines and experimental desgin.
A) EC-50 of Eμ-Myc Cdkn2a-/- and Tp53-/- lines after 24 hours of mafosfamide treatment. B) Sample generation for each line was performed as shown above in biological triplicate.

proteins in common **(Supplemental Figure 1B)**. Of those proteins identified, only those present in all three replicates of both experiments were kept for further analysis resulting in 607 proteins identified across all sample sets **(Supplemental Figure 1C).**

Analysis of variance (ANOVA) of either the transcriptomic or proteomic data resulted in different lists of significant genes and the enrichment of non-overlapping gene ontology (GO) processes. For the transcriptomic analysis, all 21720 transcripts identified in common across all samples were ranked by their interaction significance using a two-way ANOVA whose factors were time since treatment (continuous) and the cell type (categorical) **(Supplemental Table 1)**. A multi-level noise model was employed including strata for technical replicate and sample splitting by time since treatment. GO term group set analysis for the top 50 significant genes showed enrichment for biological processes related to cellular response to DNA damage and cell cycle checkpoint control **(Supplemental Table 2)**. For the proteomic analysis, the 607 proteins identified in common across all samples were analyzed using the same two-way ANOVA resulting in a different list of top 50 most significant proteins **(Supplemental Table 3)**. Furthermore, GO term group set analysis showed enrichment for metabolic processes such as gylcolysis and carbohydrate catabolic processes **(Supplemental Table 4)**. Since Myc is known to be involved in the regulation of both cell cycle checkpoint and proliferation as well as promote the transcription of genes involved in glycolysis the differing results of the two independent analyses suggests that their exists differential regulation of the transcriptome and proteome between the two lines centered around Myc controlled genes. Based on these differing results we sought to investigate potential mechanisms of regulation that would give rise to such differences.

Transcriptomic and proteomic analysis

Clustering of the transcriptomic and proteomic data with respect to cell line and time showed an inverse in the regulation of the transcriptome and proteome over time **(Figure 2A)**. Overall protein expression appeared to increase in the Tp53-/- line with a corresponding decrease in mRNA expression particularly at the 48 hour time point. This trend is reversed in the Cdkn2a-/- line which displays decreasing protein expression, but increasing mRNA expression. The overall correlation between gene and protein expression after 48 hours of drug exposure within these 594 mRNA-protein pairs was low ($R^2 \approx 0.0$) suggesting that there exists extensive post-transcriptional regulation occurring during drug exposure **(Figure 2B & C).**

Combined transcriptomic-proteomic ANOVA resulted in pathways with overall higher significance involved in post-transcriptional regulation compared to either transcriptomic or proteomic data alone. To identify genes and pathways involved in post-transcriptional regulation we first reduced the 607 proteins identified in all replicates to those that had corresponding mRNA values yielding 594 mRNA-protein expression pairs. We then performed a triple interaction ANOVA similar to the one described above, but with the additional factor of "omic" state. mRNA- protein pairs where ranked and selected based on their interaction significance **(Table 1)**. GO term group set analysis of the 169 significant mRNA-protein pairs identified by the triple interaction ANOVA differed greatly from the individual analyses with enrichment for completely different GO terms and overall higher significance. Specifically, the combined analysis showed enrichment for terms related to the regulation of protein ubiquitination or the proteasome, which did not appear in either the mRNA or protein analysis

Gene Name	p-value	Gene Name	p-value
Psmb1	1.36E-07	Suclg1	0.00074
Aldh9a1	1.70E-06	Hprt	0.00081
Fkbp3	2.41E-06	Atic	0.00082
Uba1	4.41E-06	Mdh1	0.00086
Cycs	9.30E-06	Psma3	0.00103
Hist1h1a	4.40E-05	Psmb7	0.00104
Etfb	4.60E-05	Ptma	0.00112
Tagln2	7.23E-05	Bub3	0.0012
Lmnb1	7.95E-05	Psmd6	0.00133
Ndufa4	9.99E-05	Wars	0.00148
Psma6	0.0001	Psmb4	0.0016
Ywhaq	0.00012	Snrpd1	0.00166
Ywhae	0.00012	Psma1	0.00171
Rpl34	0.00015	Rnaseh2b	0.00205
Eif5b	0.00017	Eef1b2	0.00211
Adk	0.00018	Pa2g4	0.00218
Psme1	0.00023	Ide	0.00232
Hdgf	0.00025	Psma5	0.00232
Srsf3	0.00028	Mcm7	0.00254
Actn4	0.00028	Pfn1	0.00275
Atad3a	0.00037	Xpo1	0.00307
Psmc2	0.00039	Ncbp1	0.00333
Alb	0.00048	Ywhab	0.00336
Dars	0.00057	Strap	0.00356
Iars	0.00062	Eif3g	0.00365

Table 1: Top 50 mRNA-protein pairs by interaction significance.
Top 50 genes by interaction significance from the three-way ANOVA for mRNA-protein expression pairs in Tp53-/- and Cdkn2a-/- lines after 48 hours of mafosfamide treatment.

(Table 2). One such gene, E1 ubiquitin ligase Uba1 showed increasing protein expression over time in the Tp53-/- line compared to the Cdkn2a-/- line despite decreasing levels of Uba1 transcript levels (**Figure 2D)**. Ubiquitination is a multistep enzymatic pathway whereby ubiquitin is conjugated to target proteins via Uba1, an E1 ATP-dependent ubiquitin ligase[27]. Protein degradation via the ubiquitin-proteasome pathway is the major method by which cells remove excess or damaged proteins[28]. Pharmacological inhibition of the proteasome using clinically approved inhibitors has been shown to be cytotoxic to malignant cells and improve patient outcome in certain hematological malignancies[29]. However, the consequences of inhibiting the ubiquitin-proteasome pathway upstream of protein degradation is less well understood. Based off of our initial data showing increased Uba1 protein expression, but decreasing mRNA levels over time in the Tp53-/- cell line and the opposite trend in the Cdkn2a-/- cell line we set out to investigate the effect of Uba1 inhibition on drug sensitivity.

Uba1 inhibition

Using sub-lethal concentrations of Uba1 inhibitor Pyr-41 in combination with mafosfamide significantly reduced the LC50 of the Tp53-/- line compared to that of the Cdkn2a-/- line **(Figure 3A)(Supplemental Figure 2A)**. At lower concentrations of mafosfamide, the addition of Pyr-41 displayed an additive effect marginally increasing cell death in both lines. However, at the highest concentration of mafosfamide tested the percent of viable cells decreased from ~52% to less than 30% suggesting that the effects of both Uba1 inhibition and DNA damage induced cell death becomes synergistic. While the Cdkn2a-/- cell line also displayed increased cell death with the addition of Pyr-41 compared to mafosfamide alone, at

higher mafosfamide concentrations the combinatorial effect was less pronounced than in the Tp53-/- line. Furthermore, Uba1 inhibition in the presence of mafosfamide resulted in a decrease of Uba1 expression in the Tp53-/- line without changing the pattern of expression seen in the Cdkn2a-/- line under mafosfmide treatment alone **(Figure 3B)**. The changes in Uba1 expression under combinatorial treatment suggests that inhibition of Uba1 results in decreased Uba1 protein expression. Both lines had substantially less protein ubiquitination with almost no detectable ubiquitination by 48 hours **(Figure 3B)**.

Inhibition of pathways downstream of Uba1 including the proteasome were shown preferentially affect the Tp53-/- line, but do not appear to be synergistic when combined with mafosfamide. Since Uba1 mediated protein ubiquitination is one of the earliest steps in the ubiquitin-protease pathway we asked whether inhibition of downstream components of the pathway would have a similar result to Uba1 inhibition. Uba1 is involved in the proteasome as well as the autophagy pathways, both of which have been previously been shown to contribute directly to drug resistance. As such, both pathways are currently under investigation as pharmacological targets for the treatment of various malignancies. Bortezomib is one such small molecule that inhibits the proteasome and has been approved for use in the clinic for the treatment of hematological malignancies[30]. While bortezomib resulted in cell death in the nano-molar range and preferentially killed Tp53-/- cells compared to Cdkn2a-/- cells, combinatorial treatment with sub-lethal doses of bortezomib and mafosfamide did not synergistically increase in cell death in either cell line **(Figure 3C)(Supplemental Figure 2B)**. Cholorquine inhibits **a**utophagy induced protein and organelle turnover and has been implicated in drug resistance in Tp53-/- Burkitt's lymphoma[31]. However, in our model neither

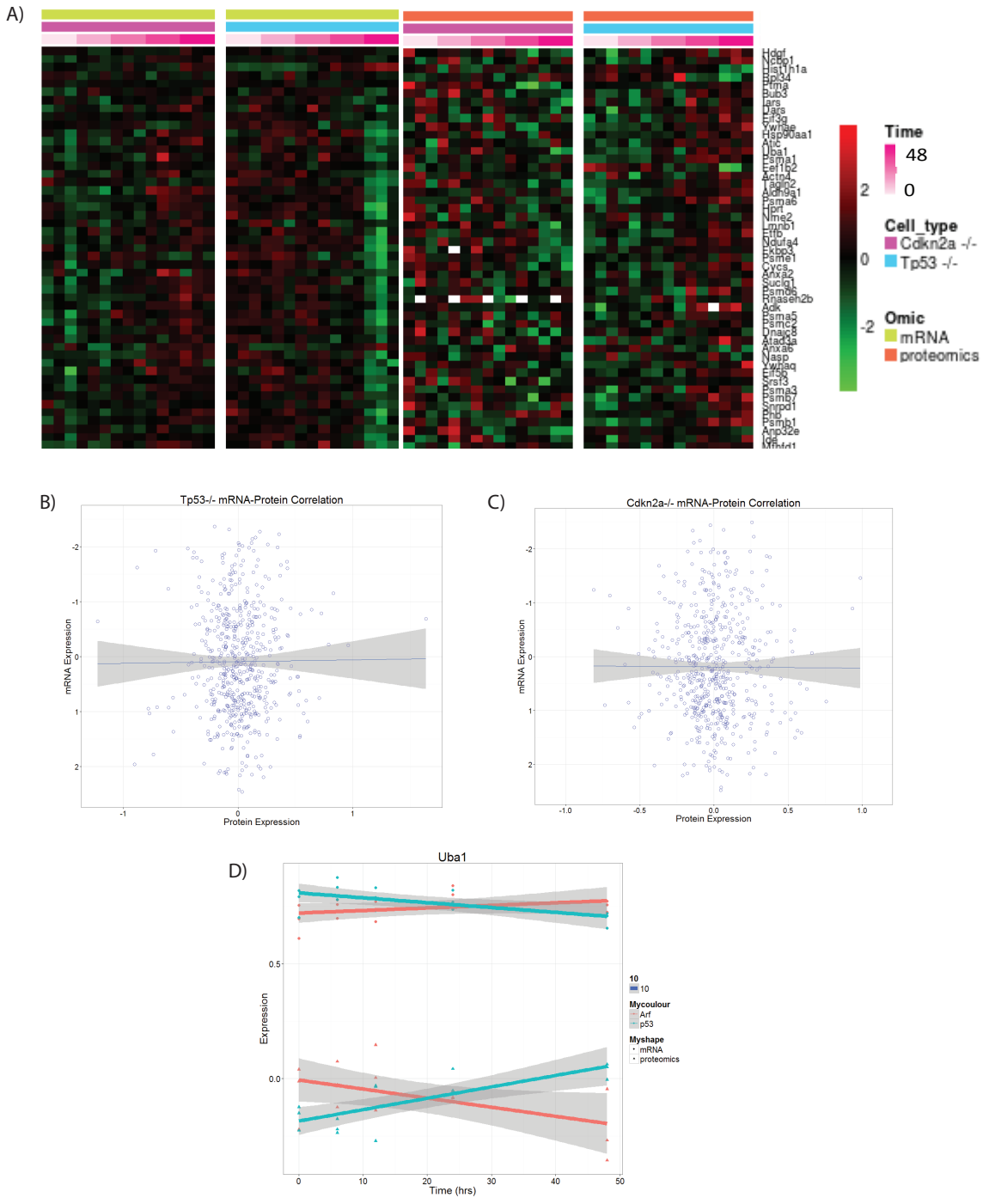


Figure 2: mRNA and protein expression under mafosfamide treatment.
A) (From left to right) Cdkn2a-/- and Tp53-/- mRNA expression respectively over 48 hours. Cdkn2a-/- and Tp53-/- protein expression respectively over 48 hours. B) Tp53-/- mRNA-protein correlation after 48 hours of mafosfamide treatment. C) Cdk2na-/- mRNA-protein correlation after 48 hours of mafosfamide treatment. D) Uba1 mRNA and protein expression over time in both Tp53-/- and Cdkn2a-/- cells.

GO ID	GO Biological Process Term	Annotated	Significant	Expected	Rank in Classic Fisher	Classic Fisher
GO:0051436	negative regulation of **ubiquitin**-protein **ligase** activity involved in mitotic cell cycle	27	10	2.4	1	**3.30E-05**
GO:0051352	negative regulation of **ligase** activity	27	10	2.4	2	**3.30E-05**
GO:0051444	negative regulation of **ubiquitin-protein ligase** activity	27	10	2.4	3	**3.30E-05**
GO:0031145	anaphase-promoting complex-dependent proteasomal **ubiquitin-dependent** protein catabolic process	28	10	2.49	4	**4.80E-05**
GO:0007093	mitotic cell cycle checkpoint	28	10	2.49	5	**4.80E-05**
GO:0009308	amine metabolic process	28	10	2.49	6	**4.80E-05**
GO:0031397	negative regulation of protein **ubiquitination**	28	10	2.49	7	**4.80E-05**
GO:0051439	regulation of ubiquitin-protein **ligase** activity involved in mitotic cell cycle	28	10	2.49	8	**4.80E-05**
GO:0051340	regulation of **ligase** activity	29	10	2.58	9	**6.90E-05**
GO:0051438	regulation of ubiquitin-protein **ligase** activity	29	10	2.58	10	**6.90E-05**
GO:0000075	cell cycle checkpoint	36	11	3.2	11	**1.00E-04**
GO:0006521	regulation of cellular amino acid metabolic process	25	9	2.22	12	**0.00011**
GO:0033238	regulation of cellular amine metabolic process	25	9	2.22	13	**0.00011**
GO:0031396	regulation of protein **ubiquitination**	31	10	2.75	14	**0.00013**
GO:0006977	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	26	9	2.31	15	**0.00016**
GO:0044106	cellular amine metabolic process	26	9	2.31	16	**0.00016**
GO:0031571	mitotic cell cycle G1/S transition DNA damage checkpoint	26	9	2.31	17	**0.00016**
GO:0031575	mitotic cell cycle G1/S transition checkpoint	26	9	2.31	18	**0.00016**
GO:0071779	G1/S transition checkpoint	26	9	2.31	19	**0.00016**
GO:0072395	signal transduction involved in cell cycle checkpoint	26	9	2.31	20	**0.00016**
GO:0072401	signal transduction involved in DNA integrity checkpoint	26	9	2.31	21	**0.00016**
GO:0072404	signal transduction involved in G1/S transition checkpoint	26	9	2.31	22	**0.00016**
GO:0072413	signal transduction involved in mitotic cell cycle checkpoint	26	9	2.31	23	**0.00016**
GO:0072422	signal transduction involved in DNA damage checkpoint	26	9	2.31	24	**0.00016**
GO:0072431	signal transduction involved in mitotic cell cycle G1/S transition DNA damage checkpoint	26	9	2.31	25	**0.00016**
GO:0072474	signal transduction involved in mitotic cell cycle G1/S checkpoint	26	9	2.31	26	**0.00016**
GO:2000045	regulation of G1/S transition of mitotic cell cycle	26	9	2.31	27	**0.00016**
GO:0000216	M/G1 transition of mitotic cell cycle	32	10	2.84	29	**0.00018**
GO:0007346	regulation of mitotic cell cycle	32	10	2.84	30	**0.00018**

Table 2: Group set anlaysis of GO terms from mRNA-protein three way ANOVA.
Top 30 GO terms for biological processes enriched for from the 169 mRNA-protein pairs that showed signficant interaction analysis from the three way ANOVA for Cdkn2a-/- and Tp53-/- mafosfamide treated cells over 48 hours.

cell lines showed an increase in cell death when mafosfamide was combined with chloroquine **(Figure 3C)(Supplemental Figure 2C)**. Since neither combinatorial proteasome and autophagy inhibitors resulted in increased cell death when combined with mafosfamide, Uba1 inhibition appears be acting independently of either pathway to result in increased drug tolerance in the Tp53-/- cell line. Uba1 is known to be involved in DNA repair and it must be this function of Uba1 acting independently of either autophagy or the proteasome to increase drug resistance[27]. This suggests that targeting Uba1 may be a clinically relevant target independent of drugs targeting either the proteasome or autophagy.

Discussion

Developing an understanding of global proteome changes is important in identifying pathways that are of biological significance. Extensive work using mRNA platforms have done just that at the transcriptomic level. However, given the poor correlation between mRNA and protein level it is important to develop independent proteomic signatures. Similarly, developing such proteomic signatures and comparing them to the transcriptomic signatures should elucidate genes and pathways that are regulated at the translational level. Interestingly, in our data protein levels in both lines fall within a much narrower range than those of the corresponding mRNA levels suggesting that the control over protein expression is more tightly controlled than that of mRNA. Current methods to de-convolute mRNA and protein dynamics rely on mathematical models that are only able to explain a small portion of mRNA-protein expression dynamics and do not elucidate molecular mechanisms through which they might act.

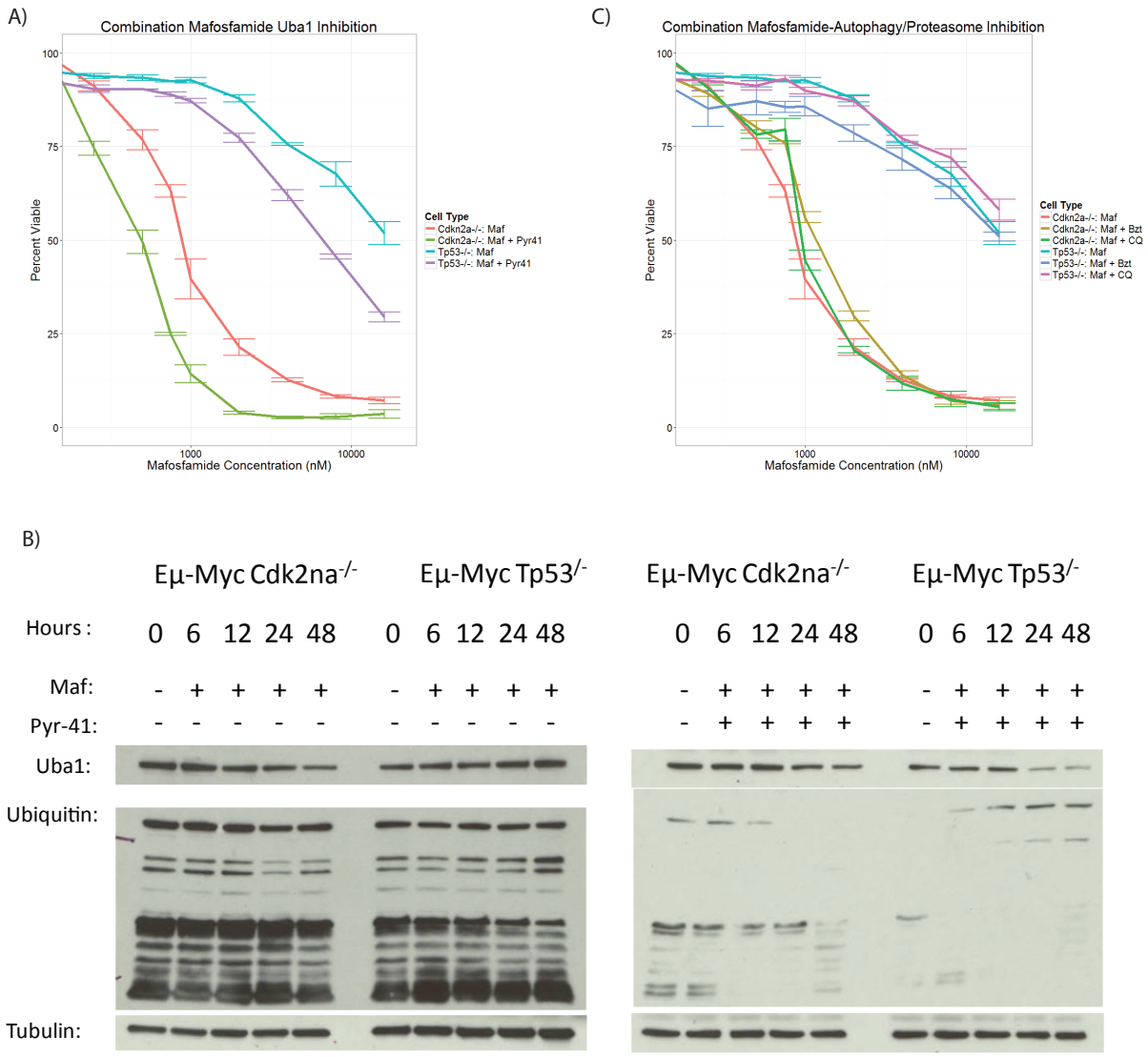
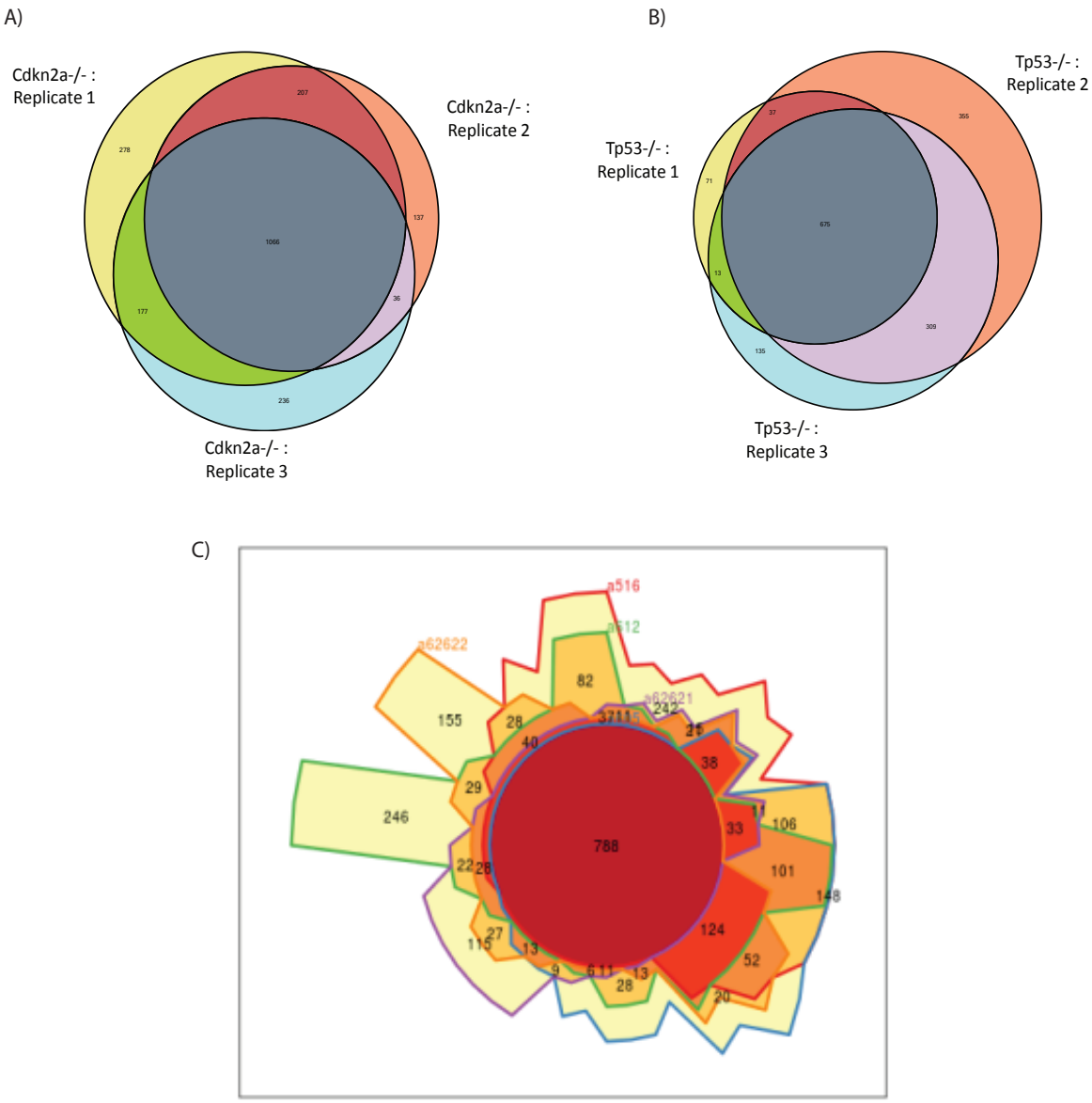


Figure 3: Combinitorial inhibition and mafosfamide treatment.
A) EC-50 of combinitorial treatment of mafosfamide and Uba1 inhibitor Pyr-41 (10μM) after 24 hours. B) Uba1 protein expression reproduces protein levels seen in LC-MS/MS experiments, with decrease expression over time in Cdk2na-/- cells and increase expression in Tp53-/- cells. After 48 hours of combinitorial mafosfamide Pyr-41 treatment, Uba1 protein levels are reduced in the Tp53-/- cells. Both lines show reduction in total protein ubiquitination after 48 hours of treatmet. C) EC-50 of Tp53-/- and Cdkn2a-/- lines 24 hours after either mafosfamide, mafosfamide and bortezomib (bzt) (1nM), or mafosfamide and chloroquine (CQ) (1μM).

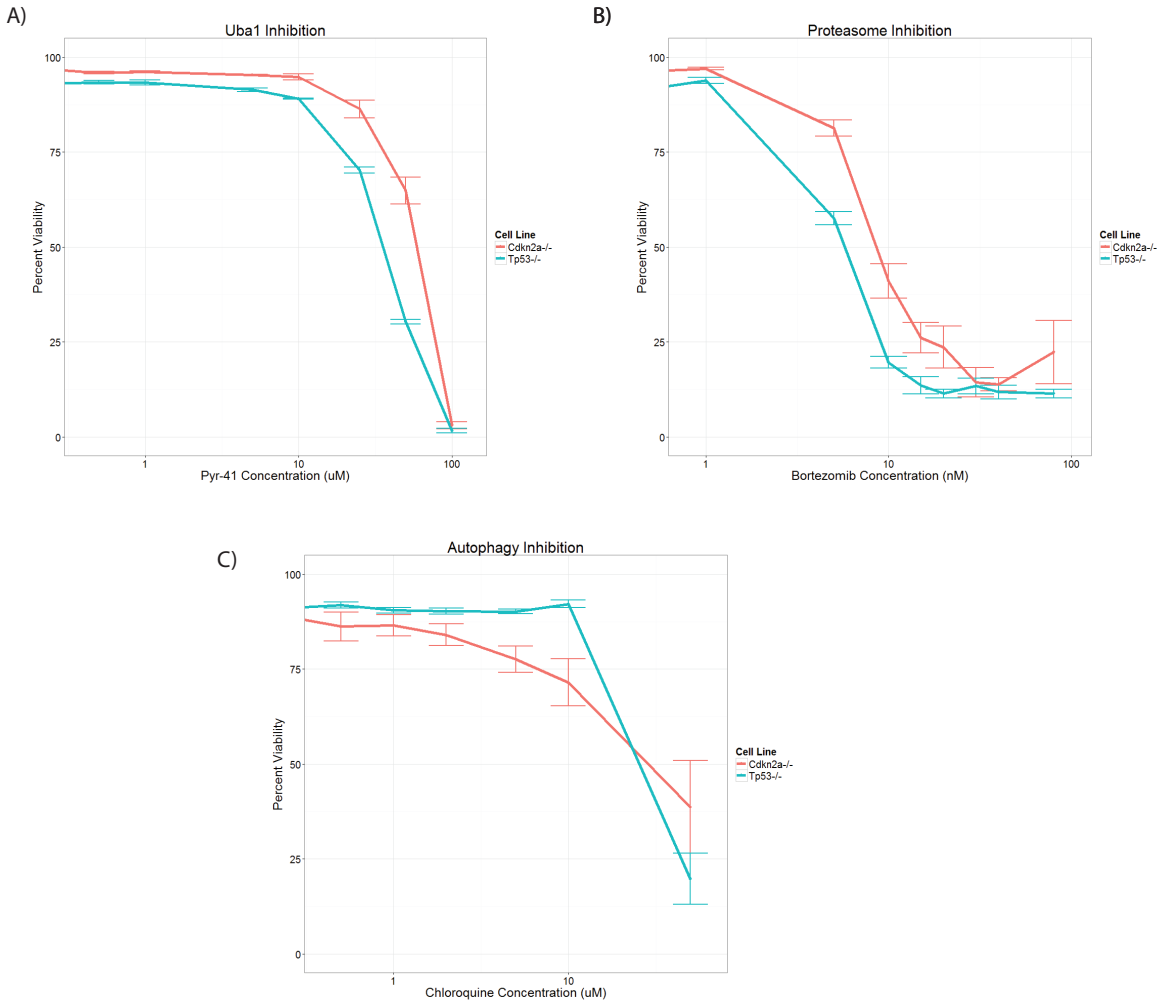
In this study we sought to identify pathways involved in chemotherapeutic resistance that are regulated at the post-transcriptional level. By comparing changes in mRNA and protein level in a temporal manner in response to the DNA alkylating agent mafosfamide we were able to identify Uba1 as a potential mechanism of resistance that is regulated at the post-translational level. Although individual analysis of changes of either the transcriptome or proteome highlighted the potential role of Myc regulation of cell cycle control and metabolism neither alone suggested a role for Uba1 mediated post-transcriptional regulation. Interestingly, our work suggests that in response to cytotoxic agents the control of cell cycle checkpoint and apoptosis pathways is regulated at the of transcription, while metabolic changes are regulated at the level of the proteome.

Combined transcriptomic-proteomic analysis yield enrichment of GO terms associated with the ubiquitin-proteasome pathway with overall greater significance and the individual analyses. In our Tp53-/- cell line, Uba1 showed increased protein expression compared to the drug sensitive Cdkn2a -/- cell line. Pharmacological inhibition of Uba1 resulted in increased cell death in the Tp53-/- cell line compared to the Cdkn2a-/- cell line. Furthermore, Uba1 inhibition led to lower levels of protein ubiquitination, suggesting that protein ubiquitination is in some part mediating drug resistance in these cell lines. Although both lines and in particular the Tp53-/- cell appear to be highly sensitive to proteasome inhibition, combinatorial inhibition of Uba1 and the proteasome were not synergistic. As such, Uba1 appears to be acting independently of downstream pathways since inhibition of either the proteasome or autophagy did not result in increased cell death when combined with mafosfamide.

The increase in protein expression and simultaneous decrease in mRNA expression in the Tp53-/- line suggests that either an increased translational rate or decreased degradation in the Tp53-/- line compared to the Cdkn2a-/- line may explain observed differences between the two lines in response to mafosfamide. Despite the high frequency of Tp53 mutations in various cancers its mutational status is of limited prognostic value[24]. This is partly due to mutations downstream of Tp53 which can inactivate aspects of its function. However, Tp53 inactivation of has also been shown to lead to an increase in ribosome biogenesis[32]. Given both the role that Myc and Tp53 play in the biogenesis, ribosomal activity and protein synthesis and that Myc over expressing cancers often contain Tp53 inactivating mutations it would be interesting to look further into the synergy that these two lesions have in protein synthesis and how it affects disease progress and drug resistance.



Supplemental Figure 1: Protein overlap between LC-MS/MS experiments.
A) Protein identification overlap between biological replicates of LC-MS/MS experiments from mafosfamide treated Cdkn2a-/- cells. B)Protein identification overlap between biological replicates of LC-MS/MS experiments from mafosfamide treated Tp53-/- cells. C) Protein identification overlap between all LC-MS/MS experiments.



Supplmental Figure 2: EC-50s of inhibitors

A) EC-50 of Uba1 inhibitor Pyr-41 after 24 hours of treatment in Cdkn2a-/- and Tp53-/- cells. B) EC-50 of proteasome inhibitor bortezomib (btz) after 24 hours of treatment in Cdkn2a-/- and Tp53-/- cells. C) EC-50 of autophagy inhibitor chloroquine (CQ) after 24 hours of treatment in Cdkn2a-/- and Tp53-/- cells.

Gene Name	p-value	Gene Name	p-value
H2-Eb1	1.45E-09	Pno1	1.25E-07
Wdr74	1.51E-09	Ddit4	1.38E-07
Lipt2	3.89E-09	Coro1c	1.40E-07
Ddx39	6.80E-09	B4galt5	1.49E-07
Cacna1e	7.38E-09	H2-DMa	1.52E-07
Mxd4	7.88E-09	Bcl11a	1.63E-07
Prmt5	9.10E-09	Prr5	1.66E-07
Smyd5	9.96E-09	1110001A16Rik	1.70E-07
Sap30bp	1.09E-08	Clpp	1.74E-07
Btg2	1.14E-08	Enoph1	2.05E-07
Tmem181a	2.07E-08	Thg1l	2.16E-07
Tmco4	2.26E-08	Ltv1	2.20E-07
1810013D10Rik	3.02E-08	Guk1	2.59E-07
Rrp12	3.81E-08	Aven	2.82E-07
Wars	4.52E-08	Zfp593	2.83E-07
Rnf138	7.02E-08	1110008F13Rik	2.86E-07
Mrpl40	7.17E-08	Prkrir	3.60E-07
Wdr34	7.38E-08	Tmem109	3.64E-07
Hddc2	7.58E-08	Timm8a1	3.70E-07
Phb	8.24E-08	Sfxn1	4.12E-07
Fam54b	9.67E-08	Rnaseh2b	4.24E-07
Rsl1d1	9.73E-08	Lyar	4.58E-07
Eif2s3y	1.02E-07	Gart	4.69E-07
Plk3	1.07E-07	Kcnab2	5.14E-07
Nme2	1.11E-07	Cisd1	5.38E-07

Supplemental Table 1: Top 50 mRNA genes by interaction significance.

Top 50 genes by interaction significance from the two-way ANOVA for mRNA expression in Tp53-/- and Cdkn2a-/- lines after 48 hours of mafosfamide treatment.

GO ID	GO Biological Process Term	Annotated	Significant	Expected	Rank in classic Fisher	Classic Fisher
GO:0034969	histone arginine methylation	8	2	0.03	1	0.00043
GO:0035247	peptidyl-arginine omega-N-methylation	8	2	0.03	2	0.00043
GO:0018216	peptidyl-arginine methylation	9	2	0.04	3	0.00056
GO:0035246	peptidyl-arginine N-methylation	9	2	0.04	4	0.00056
GO:0006974	cellular response to DNA damage stimulus	615	9	2.46	5	0.00066
GO:1901070	guanosine-containing compound biosynthet...	10	2	0.04	6	0.00069
GO:0018195	peptidyl-arginine modification	14	2	0.06	7	0.00139
GO:0015949	nucleobase-containing small molecule int...	16	2	0.06	8	0.00182
GO:0006479	protein methylation	84	3	0.34	9	0.00459
GO:0008213	protein alkylation	84	3	0.34	10	0.00459
GO:0010941	regulation of cell death	1179	11	4.72	11	0.00606
GO:0009411	response to UV	96	3	0.38	12	0.00666
GO:0010605	negative regulation of macromolecule met...	1372	12	5.5	13	0.0068
GO:0042771	intrinsic apoptotic signaling pathway in...	34	2	0.14	14	0.00813
GO:0016070	RNA metabolic process	3229	21	12.94	15	0.00934
GO:0007346	regulation of mitotic cell cycle	326	5	1.31	16	0.00955
GO:0090304	nucleic acid metabolic process	3686	23	14.77	17	0.00995
GO:0000075	cell cycle checkpoint	211	4	0.85	18	0.00997
GO:0006396	RNA processing	617	7	2.47	19	0.01096
GO:0044272	sulfur compound biosynthetic process	117	3	0.47	20	0.01143
GO:1901990	regulation of mitotic cell cycle phase t...	220	4	0.88	21	0.01149
GO:1901987	regulation of cell cycle phase transitio...	224	4	0.9	22	0.01221
GO:0016575	histone deacetylation	42	2	0.17	23	0.01223
GO:0060548	negative regulation of cell death	638	7	2.56	24	0.01301
GO:0009892	negative regulation of metabolic process	1494	12	5.99	25	0.01311
GO:0006283	transcription-coupled nucleotide-excisio...	46	2	0.18	26	0.01455
GO:0043414	macromolecule methylation	130	3	0.52	27	0.01516
GO:0030516	regulation of axon extension	48	2	0.19	28	0.01578
GO:0000077	DNA damage checkpoint	134	3	0.54	29	0.01643
GO:0009314	response to radiation	375	5	1.5	30	0.01669

Supplemental Table 2: Group set anlaysis of GO terms from mRNA two way ANOVA.
Top 30 GO terms for biological processes enriched for from the significant mRNA transcripts from the two way ANOVA for Cdkn2a-/- and Tp53-/- mafosfamide treated cells over 48 hours.

Gene Name	p-value	Gene Name	p-value
Lmnb1	4.71E-06	Rpl11	0.000885
Rpl34	1.80E-05	Pdia3	0.000896
Rps4x	2.23E-05	P4hb	0.000902
Tpi1	2.31E-05	Calr	0.000916
Ass1	2.57E-05	Hsp90b1	0.000973
Hdgf	4.32E-05	Ppib	0.000977
Tfrc	4.61E-05	Ywhae	0.001083
Uba1	4.77E-05	Ahcy	0.001123
Cycs	4.97E-05	Eef1g	0.001182
Sf3b1	5.09E-05	Rps29	0.001299
Alb	5.32E-05	Hmgb2	0.001551
U2surp	6.31E-05	Prdx1	0.001633
Tagln2	8.27E-05	Rps12	0.001734
Psmb1	0.000101	Slc25a3	0.001753
Hnrnpa0	0.000158	Anxa6	0.001781
Aldh2	0.00025	Prpf8	0.001788
Lcp1	0.000268	Taldo1	0.00234
Psme1	0.000325	Aldoa	0.002433
Fth1	0.000334	Bub3	0.002459
Hspa9	0.000379	Lrpprc	0.002529
Eef1b2	0.000506	Myl6	0.002566
Ptma	0.000637	Rpl13a	0.002642
Rpl39	0.000819	Adk	0.002677
Pgk1	0.00086	Sfpq	0.002717
Fubp1	0.000868	Phb2	0.002762

Supplemental Table 3: Top 50 proteins by interaction significance.
Top 50 proteins by interaction significance from the two-way ANOVA for protein expression in Tp53-/- and Cdkn2a-/- lines after 48 hours of mafosfamide treatment.

GO ID	GO Biological Process Term	Annotated	Significant	Expected	Rank in classic Fisher	Classic Fisher
GO:0051235	maintenance of location	14	6	1.21	1	0.00055
GO:0019725	cellular homeostasis	25	7	2.17	2	0.00344
GO:0006875	cellular metal ion homeostasis	9	4	0.78	3	0.00456
GO:0051238	sequestering of metal ion	5	3	0.43	4	0.00542
GO:0042592	homeostatic process	49	10	4.25	5	0.00582
GO:0051651	maintenance of location in cell	10	4	0.87	6	0.00712
GO:0006873	cellular ion homeostasis	10	4	0.87	7	0.00712
GO:0030003	cellular cation homeostasis	10	4	0.87	8	0.00712
GO:0055065	metal ion homeostasis	10	4	0.87	9	0.00712
GO:0002237	response to molecule of bacterial origin	6	3	0.52	10	0.01018
GO:0032496	response to lipopolysaccharide	6	3	0.52	11	0.01018
GO:0055080	cation homeostasis	11	4	0.95	12	0.01047
GO:0032075	positive regulation of nuclease activity	12	4	1.04	13	0.01471
GO:0050801	ion homeostasis	12	4	1.04	14	0.01471
GO:0006007	glucose catabolic process	12	4	1.04	15	0.01471
GO:0019320	hexose catabolic process	12	4	1.04	16	0.01471
GO:0009617	response to bacterium	7	3	0.61	17	0.01674
GO:0030001	metal ion transport	13	4	1.13	18	0.0199
GO:0032069	regulation of nuclease activity	13	4	1.13	19	0.0199
GO:0034976	response to endoplasmic reticulum stress	13	4	1.13	20	0.0199
GO:0046364	monosaccharide biosynthetic process	13	4	1.13	21	0.0199
GO:0016052	carbohydrate catabolic process	13	4	1.13	22	0.0199
GO:0044724	single-organism carbohydrate catabolic p...	13	4	1.13	23	0.0199
GO:0046365	monosaccharide catabolic process	13	4	1.13	24	0.0199
GO:0006096	glycolysis	8	3	0.69	25	0.02515
GO:0060249	anatomical structure homeostasis	14	4	1.21	26	0.0261
GO:0033993	response to lipid	14	4	1.21	27	0.0261
GO:0055082	cellular chemical homeostasis	14	4	1.21	28	0.0261
GO:0003012	muscle system process	14	4	1.21	29	0.0261
GO:0051239	regulation of multicellular organismal p...	52	9	4.51	30	0.02671

Supplemental Table 4: Group set anlaysis of GO terms from the proteintwo way ANOVA.
Top 30 GO terms for biological processes enriched for from the significant proteins from the two way ANOVA for Cdkn2a-/- and Tp53-/- mafosfamide treated cells over 48 hours.

References

1. Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med*. 2006;354(23):2419-2430.

2. Dave SS, Fu K, Wright GW, et al. Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med*. 2006;354(23):2431-2442.

3. O'Connell K, Prencipe M, O'Neill A, et al. The use of LC-MS to identify differentially expressed proteins in docetaxel-resistant prostate cancer cell lines. *Proteomics*. 2012;12(13):2115-2126.

4. Everton KL, Abbott DR, Crockett DK, Elenitoba-Johnson KS, Lim MS. Quantitative proteomic analysis of follicular lymphoma cells in response to rituximab. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2009;877(13):1335-1343.

5. de Sousa Abreu R, Penalva LO, Marcotte EM, Vogel C. Global signatures of protein and mRNA expression levels. *Mol Biosyst*. 2009;5(12):1512-1526.

6. Tchourine K, Poultney CS, Wang L, et al. One third of dynamic protein expression profiles can be predicted by a simple rate equation. *Mol Biosyst*. 2014;10(11):2850-2862.

7. Hecht JL, Aster JC. Molecular biology of Burkitt's lymphoma. *J Clin Oncol*. 2000;18(21):3707-3721.

8. van Riggelen J, Yetil A, Felsher DW. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer*. 2010;10(4):301-309.

9. Sewastianik T, Prochorec-Sobieszek M, Chapuy B, Juszczynski P. MYC deregulation in lymphoid tumors: molecular mechanisms, clinical consequences and therapeutic implications. *Biochim Biophys Acta*. 2014;1846(2):457-467.

10. Dang CV, O'Donnell KA, Zeller KI, Nguyen T, Osthus RC, Li F. The c-Myc target gene network. *Semin Cancer Biol*. 2006;16(4):253-264.

11. Schuhmacher M, Kohlhuber F, Holzel M, et al. The transcriptional program of a human B cell line in response to Myc. *Nucleic Acids Res*. 2001;29(2):397-406.

12. Soucek L, Whitfield J, Martins CP, et al. Modelling Myc inhibition as a cancer therapy. *Nature*. 2008;455(7213):679-683.

13. Choi PS, van Riggelen J, Gentles AJ, et al. Lymphomas that recur after MYC suppression continue to exhibit oncogene addiction. *Proc Natl Acad Sci U S A*. 2011;108(42):17432-17437.

14. Schmitt CA, McCurrach ME, de Stanchina E, Wallace-Brodeur RR, Lowe SW. INK4a/ARF mutations accelerate lymphomagenesis and promote chemoresistance by disabling p53. *Genes Dev*. 1999;13(20):2670-2677.

15. Wildes TM, Farrington L, Yeung C, et al. Rituximab is associated with improved survival in Burkitt lymphoma: a retrospective analysis from two US academic medical centers. *Ther Adv Hematol*. 2014;5(1):3-12.

16. Lee EJ, Petroni GR, Schiffer CA, et al. Brief-duration high-intensity chemotherapy for patients with small noncleaved-cell lymphoma or FAB L3 acute lymphocytic leukemia: results of cancer and leukemia group B study 9251. *J Clin Oncol*. 2001;19(20):4014-4022.

17. Blum KA, Lozanski G, Byrd JC. Adult Burkitt leukemia and lymphoma. *Blood*. 2004;104(10):3009-3020.

18. Schmitz R, Young RM, Ceribelli M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature*. 2012;490(7418):116-120.

19. Love C, Sun Z, Jima D, et al. The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet*. 2012;44(12):1321-1325.

20. Schmitt CA, Fridman JS, Yang M, Baranov E, Hoffman RM, Lowe SW. Dissecting p53 tumor suppressor functions in vivo. *Cancer Cell*. 2002;1(3):289-298.

21. Peller S, Rotter V. TP53 in hematological cancer: low incidence of mutations with significant clinical relevance. *Hum Mutat*. 2003;21(3):277-284.

22. Kruse JP, Gu W. Modes of p53 regulation. *Cell*. 2009;137(4):609-622.

23. Sturm I, Bosanquet AG, Hermann S, Guner D, Dorken B, Daniel PT. Mutation of p53 and consecutive selective drug resistance in B-CLL occurs as a consequence of prior DNA-damaging chemotherapy. *Cell Death Differ*. 2003;10(4):477-484.

24. Xu-Monette ZY, Medeiros LJ, Li Y, et al. Dysfunction of the TP53 tumor suppressor gene in lymphoid malignancies. *Blood*. 2012;119(16):3668-3683.

25. Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*. 2003;19(2):185-193.

26. Thompson A, Schafer J, Kuhn K, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem*. 2003;75(8):1895-1904.

27. Tu Y, Chen C, Pan J, Xu J, Zhou ZG, Wang CY. The Ubiquitin Proteasome Pathway (UPP) in the regulation of cell cycle control and DNA damage repair and its implication in tumorigenesis. *Int J Clin Exp Pathol*. 2012;5(8):726-738.

28. Suh KS, Tanaka T, Sarojini S, et al. The role of the ubiquitin proteasome system in lymphoma. *Crit Rev Oncol Hematol*. 2013;87(3):306-322.

29. Perez-Galan P, Roue G, Villamor N, Montserrat E, Campo E, Colomer D. The proteasome inhibitor bortezomib induces apoptosis in mantle-cell lymphoma through generation of ROS and Noxa activation independent of p53 status. *Blood*. 2006;107(1):257-264.

30. Yu D, Carroll M, Thomas-Tikhonenko A. p53 status dictates responses of B lymphomas to monotherapy with proteasome inhibitors. *Blood*. 2007;109(11):4936-4943.

31. Mezzaroba N, Zorzet S, Secco E, et al. New potential therapeutic approach for the treatment of B-Cell malignancies using chlorambucil/hydroxychloroquine-loaded anti-CD20 nanoparticles. *PLoS One*. 2013;8(9):e74216.

32. Marcel V, Ghayad SE, Belin S, et al. p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. *Cancer Cell*. 2013;24(3):318-330.

Conclusion

Over the past decade advancements in DNA sequencing and peptide mass spectrometry technology have resulted in the widespread use of these technologies by researchers. Where as previously these technologies were confined to highly specialized laboratories their decrease in expense and increase in robustness have allowed their use to become more widespread. As the use of these platforms increases and becomes more common place the number of studies leveraging multiple global measurement platforms as similarly increased. Global measurements spanning the genome, transcriptome, and proteome are collected on differing platforms all with differing levels of reliability, coverage depth, and quantitation. As a result, the integration of large biological datasets from multiple platforms in a relevant way that reveals new insights remains a challenge. However, integrating these data sets in such a fashion as to provide new insights into biology also represents an opportunity to understand not only fundamental aspects of molecular regulation but could also provide new insights into the development and treatment of disease.

Characterization of individual cancers based on gene expression patterns allowed for the segmentation of different cancer types based on molecular properties. This molecular definition of specific cancer types as enabled the stratification of patients and allows for a refinement of treatment tailored to an individual's specific cancer. Such an analysis allowed for a molecular definition of Burkitt's lymphoma based on gene expression patterns that distinguishes it from other high grade non-Hodgkin's lymphoma such as diffuse large B-cell lymphoma. However, individual cancers with similar gene expression profiles will still often have differential responses to treatment suggesting that gene expression patterns alone are incapable of fully segregating between different tumor types at a molecular level. The use of other global measurements of molecular status at levels other than the

transcriptome such as DNA methylation and quantitative proteomics could allow for new insights into cancer biology that could enable a more personalized treatment.

A major hurdle to the successful treatment of Burkitt's lymphoma is the acquisition of resistance. Up to approximately 80% of individuals who present with disease that initially responds to therapy later re-present with disease that is no longer amenable to treatment using the standard treatment regime R-CHOP. To gain insight into the mechanisms by which these individuals become resistant to treatment we investigated the acquisition of resistance as well as the response to treatment using a multi-omic approach.

In our initial studies we chose to investigate the regulation of mRNA-protein dynamics in two mouse models of Burkitt's lymphoma that display differing sensitivity to the DNA alkylating agent mafosfamide. By characterizing transcriptomic and proteomic changes in response to mafosfamide treatment we were able to show that not only are the transcriptome and proteome regulated in different manners, but the pathways which appear to be regulated at either level differ from one another. By performing a multi-omic analysis that took into account differences between the transcriptome and proteome over time between the two lines in response to treatment we were able to identify biological pathways that were unique to the combined analysis. Furthermore, the genes and pathways identified by the combined analysis showed greater significance than either analysis of the transcriptome or proteome alone. The combined analysis, suggested a role for the ubiquitin-proteasome pathway and in particular the ubiquitin activating enzyme Uba1. Inhibition of Uba1 resulted in decreased Uba1 protein expression and lead to decreased viability of the resistant Tp53-/- cell line. This work demonstrates that an analysis that takes into account both the transcriptome and proteome is capable of revealing novel drug resistance mechanisms and drug targets that is not revealed by analysis of either alone.

Acquired resistance arises through the acquisition of mutations or alteration of gene expression that leads to reduced treatment response. In our second study we sought to understand the rise of resistance using a model derived via dose escalation of an initially mafosfamide sensitive cell line. Analysis gene expression changes in response to treatment in the parental sensitive cell line over 48 hours suggested that an alteration of the histone 3 lysine 27 tri-methylation (H3K27me3) modifying mark was associated with drug response. Similar analysis of gene expression changes of the resistant lines derived from the parental cell line showed that genes whose expression decreased with increasing resistance were enriched for targets of the histone methyltransferase Polycomb Repressive Complex 2 (Prc2). Chromatin immuno-precipitation sequencing (ChIP-seq) of the parental and resistant lines confirmed that the resistant lines showed increased abundance of the Prc2 mark H3 lysine 27 tri-methylation. Furthermore, genes that showed increase H3K27me3 were enriched for gene ontology terms related to development. Those genes that showed increased H3K27me3 were also enriched for binding targets of the B-cell maturation transcription factor E2a. Prc2 and the addition of H3K27me3 has previously been shown to be associated with increased DNA methylation. By performing reduced representation bisulfite sequencing on our parental and resistant lines we were able to show that the DNA methylation data matched our initial gene expression analysis. Genes that showed an increase in DNA methylation were similarly enriched for targets of Prc2, thus confirming our initial hypothesis that alterations in H3K27me3 are associated with the development of resistance. Furthermore, principle component analysis of gene expression changes during normal B-cell maturation suggested that the development of resistance in our derived lines is accompanied by a regression in B-cell maturity.

Although multi-omic integration is capable of revealing new insights into biology and treatment response in cancer the analysis of multiple large datasets remains challenging. Much of the difficulty arises from differences in the measurement platforms themselves which gives rise to datasets of differing quantitation and coverage. While sequencing data can give absolute quantitative data in the

form of number of fragments sequenced, proteomic approaches are capable of only giving relative changes. This can present a problem in that recent advances in measuring the transcriptome such as RNA-sequencing allow for the measurement of precise number of mRNA transcripts for a particular gene. However, quantitative proteomics results only in the relative abundance between one sample and another. In order to better understand the dynamics of transcriptional and translational regulation it would be beneficial to be able to compare quantitative measurements across both platforms. In order to compare two datasets, one of which is absolutely quantified while the other is only relatively quantified requires either that each dataset is analyzed individually and the results compared or that the dataset that is absolutely quantified be converted to relative quantitation so that both dataset can be properly integrated. While there do exist methods to absolutely quantify protein abundance they are as of yet not robust enough to apply on a global level. As such for the time being until peptide mass spectrometry improves it will be necessary to use relative quantitation between experimental sample and control when comparing differing global datasets.

Besides quantitation level, coverage depth poses another difficultly in comparing multi-omic data. While DNA sequencing technology has progressed to the degree that whole genome sequences and whole RNA sequences can be achieved at a depth that coverage approaches 100% of sample coverage, proteomic measurements and in particular quantitative mass spectrometry of peptides is still limited by the depth of coverage. Reasonably achievable depth of coverage of the proteome is limited due to the wide biological dynamic range of protein expression compared to the dynamic range mass spectrometry instruments. The disparity in coverage between DNA sequencing and mass spectrometry technology results in data sets that have only partial overlap and thus reduce the amount of useful data collected. Furthermore, given that current mass spectrometry protocols are incapable of detecting low abundant proteins such as transcription factors and other DNA binding proteins, a great deal of

interesting biological is potentially being missed since measurements for given genes do not have data at both the transcriptomic and proteomic level.

In the studies described above the differences in coverage depth between platforms became acutely evident. The narrower dynamic range of peptide mass spectrometry compared to DNA microarray and sequencing greatly limited our ability to measure compare global transcriptomic-proteomic changes. In the initial study, we were able to identified over 2500 proteins quantitatively, however we were forced to reduce that number to those that were identified in all biological replicates yielding less than 800 proteins. Furthermore, this number was again reduced to approximately 600 because of disparities in coverage between the quantitative mass spectrometry experiments and microarray studies. Although peptide mass spectrometry technology continues to improve even the state of the art mass spectrometers are capable of reliably identifying less than 10,000 proteins quantitatively. Given that whole genome microarrays assay well over 30,000 transcripts a full two-thirds of genes will not have data at both levels.

Similarly, in our second study differences in coverage depth prevented a more in depth analysis. To sequence the methylome of our cell lines we used the reduced representation bisulfate sequencing (RRBS) method which preferentially sequences CpG islands. Since RRBS sequences only CpG islands methylated CpGs outside of CpG island are not measured. This prevented a greater in-depth analysis of overlap between DNA methylation sequencing and our ChIP-seq data which covered the entire genome. The limited overlap greatly inhibited our ability to assess the co-occupancy of these marks at particular genes of interest.