



# The Study of NSD2 Biochemical and Biological Activity Through Small-Molecule Profiling of Cell Lines

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:37944987>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

# **The study of NSD2 biochemical and biological activity through small-molecule profiling of cell lines**

A dissertation presented

by

Sixun Chen

to

The Committee on Higher Degrees in Chemical Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Chemical Biology

Harvard University

Cambridge, Massachusetts

July, 2016



## **The study of NSD2 biochemical and biological activity through small-molecule profiling of cell lines**

### **Abstract**

Histone methyltransferase NSD2 (WHSC1/MMSET) is correlated with several cancers, but promising NSD2 inhibitors are yet to be developed for the clinic, alternative strategies should be studied parallel to probe development in targeting NSD2-driven cancer. In this thesis, I employ traditional biochemical, cellular biology methods and small molecule profiling to study a recurrent activating mutation of NSD2, E1099K. This study sought identify orthogonal proteins that can be targeted to inhibit the growth of E1099K-NSD2 cancers. The body of work describes biochemical and cellular approaches in studying E1099K-NSD2 that revealed E1099K-NSD2 to mono- and dimethylate NSD2 native substrate of H3K36, leading to transcriptional signatures similar to cell lines with overexpression of NSD2. To identify NSD2 E1099K-induced dependencies of cell lines, a panel of hematopoietic cancer cell lines, were profiled against a set of small molecules to establish cell viability in response to treatment. HDAC 1/2 inhibitors, were identified to have greater potency for killing E1099K- compared to WT-NSD2 cell lines. CRISPR technology was applied to engineer isogenic cell lines with different NSD2 genomic status and profiled them against the set of small-molecules. RNA-seq identified oncogenic pathway and PRC2 targets to be upregulated in E1099K-NSD2 clones. Further analysis of clones revealed a small molecule study on mycoplasma colonized cells that resulted in a brief study described at the end of the thesis.

# Table of Contents

<b>Abstract</b> .....	<b>iii</b>
<b>Table of Contents</b> .....	<b>iv</b>
<b>Acknowledgements</b> .....	<b>vi</b>
<b>Chapter I: Histone methylation in cancer and NSD2</b> .....	<b>1</b>
1.1 Chromatin function and histone methylation .....	2
1.2 Histone methyltransferases and cancer .....	3
1.3 Role of histone methyltransferase NSD2 in physiology and cancer .....	5
1.3 Motivation and overview of the study .....	6
1.4 References.....	8
<b>Chapter II: Biochemical and cellular characterization of E1099K-NSD2</b> .....	<b>14</b>
<b>2.1 Introduction</b> .....	<b>15</b>
<b>2.2 Results and Discussion</b>	
2.2.1 Expression and purification of recombinant WT- and E1099K-NSD2.....	16
2.2.2 WT- and E1099K-NSD2 differ in activity depending on substrate .....	17
2.2.3 E1099K-NSD2 ALLs are heterozygous mutants .....	20
2.2.4 E1099K-NSD2 ALLs have distinct methylation states .....	21
2.2.5 E1099K-NSD2 ALLs have NSD2-activated gene expression signatures .....	23
2.2.6 Concluding remarks.....	26
<b>2.3 Experimental methods</b> .....	<b>26</b>
<b>2.4 References</b> .....	<b>31</b>
<b>Chapter III: Small-molecule profiling to identify dependencies of NSD2-driven CCLs</b> .....	<b>34</b>
<b>3.1 Introduction</b> .....	<b>35</b>
<b>3.2 Results and Discussion</b>	
3.2.1 HDAC1/2 selective inhibitors target E1099K-ALLs .....	36
3.2.2 IGFR inhibitors and FGFR3 inhibitors target t(4;14)+ MMs .....	43
3.2.3 Dependencies of NSD2-driven CCLs are more lineage dependent .....	47
3.2.4 Concluding remarks.....	51

2.3	Experimental methods .....	54
2.4	References.....	57
<b>Chapter IV: Engineering of CRISPR-NSD2 cell lines to identify novel NSD2 biology.....</b>		<b>62</b>
4.1	Introduction.....	63
4.2	<b>Results and Discussion</b>	
4.2.1	Engineering NSD2 cell lines using CRISPR.....	64
4.2.2	Characterization of clones obtained .....	66
4.2.3	Clones have different sensitivity towards NAMPT inhibitors.....	69
4.2.4	Acquisition of NAMPTi-resistant WT-NSD2 clones.....	72
4.2.5	Evidence of contamination .....	75
4.2.6	RNA-seq on clones.....	77
4.2.7	Concluding remarks.....	78
4.3	<b>Experimental methods .....</b>	<b>80</b>
4.4	<b>References.....</b>	<b>84</b>
<b>Chapter V: Using small-molecule profiling to study biology of <i>Mycoplasma hyorhinitis</i>-infected human cell lines .....</b>		<b>86</b>
5.1	Introduction.....	87
5.2	<b>Results and Discussion</b>	
5.2.1	Acquisition of infected clones and identification of agent of infection .....	88
5.2.2	RNA-seq results .....	89
5.2.3	Efforts in identification of factors involved in phenotype .....	91
5.2.4	Concluding remarks.....	92
5.3	<b>Experimental methods .....</b>	<b>93</b>
5.4	<b>References.....</b>	<b>95</b>
<b>Supplemental Information.....</b>		<b>98</b>
<b>Appendix.....</b>		<b>118</b>

## Acknowledgements

The most important person in my graduate school experience has been my dissertation advisor, Dr. Stuart Schreiber. Thank you Stuart, for giving me the freedom, courage to pursue scientific questions and support to do the best science that I might not have undertaken otherwise or even been possible in other circumstances. I would like to thank the members of my thesis committee for their guidance: David Liu, James Bradner, Alan Saghatelian and Andy Phillips, and my scholarship agency: A\*STAR Graduate Academy, for funding.

I have been immensely lucky to have worked alongside amazing people in the Schreiber lab. I started graduate school without any knowledge of chromatin and through the guidance and starting mentorship of Drew Adams and Joshiawa Paulk, I have come to wield techniques well and complete this dissertation. The entire Schreiber lab and JDRF team (past and present), especially Kai, Alicia, Tamara, Amedeo, Bennett, Elaine, Jake and Matt Ryan thank you for being such great labmates to work with, I have learnt so much of life and science from all of you.

Thank you Vasanthi, Michelle Stewart, Cherrie, and Liza, for being there when I am at my lowest, for your confidence in my ability and such great friendship. I will forever be grateful to you. I would like to thank the mentorship of the Broad community, Aly Shamji, Bridget Wagner, Paul Clemons, Joanne Kotz and Mathias Wawer for their time in my one-on-one meetings.

The Harvard community allowed me to interact with the kindest and most talented individuals. Lindsay, Yang Lin, Ghee Chuan, Margie, Stephanie, Ahmed, Kevin, Erin, Alexandra and Roody. You are a constant reminder of why the decision to undertake my PhD at Harvard has been one of the best in my life.

I am fortunate to have support from tens of thousands of miles away in Singapore, my brother and sister, Siheng and Sihui, my Mum and Dad, Lim Yen Kheng and Chen Guo Hua, and

my friends, Yixian, Shu Hui and Zhi Ping, for always being present and giving me your understanding for my absence back home.

Graduate school has been a bittersweet experience. It's accurate to say that I would not have been alive if not for these people. They have been an inspiration in my understanding of science and the world, immensely shaping my growth as an intellectual scholar and as a human being. Words are not enough to express my gratitude towards these exceptional individuals. I am extremely lucky to have crossed paths with them in my life.

# Chapter I

## Histone methylation in cancer and NSD2

## 1.1 Chromatin function and histone methylation

A human cell packs its gigabases of DNA and organizes it into units called chromatin, which comprise of DNA wrapped around a group of proteins named histones. The chromatin landscape in a cell is made up of post-translational modifications (PTMs) of histone proteins, such as acetylation, methylation, phosphorylation and ubiquitination. The variety of these modifications creates different permutations that would result in the manipulation of DNA such as its expression, replication and repair <sup>1</sup>. These cellular behavior can be controlled with specific modifications on chromatin and their subsequent interaction with chromatin-modifying enzymes such as methyltransferases and deacetylases, making these enzymes promising targets for the development of probes and therapeutics <sup>2-4</sup>.

The family of histone methyltransferases include proteins that methylate through a catalytic SET domain, such as EZH2 and the NSD family, and proteins that methylate through a SAM-binding domain, such as DOT1L and PRMT family <sup>5</sup>. These methyltransferases are able to methylate histones using lysine or arginine residue as substrates. Lysine residues can be mono-, di- or tri-methylated <sup>6</sup> while arginine residues can be mono- or di-methylated <sup>7</sup>. These different states of methylation create an extra layer of complexity within the chromatin landscape that allow the cellular function at the genomic level to be tightly regulated.

Cellular function that are known to be regulated by histone methylation include labeling regions of the genome into heterochromatin and euchromatin (parts of genome that are ‘silent’ and ‘active’ respectively), transcription of genes, repair of DNA lesions and replication of the genome <sup>8</sup>. These functions are associated with a set of chromatin modifications that directs relevant protein complexes to parts of the genome through reader domains that recognizes chromatin marks. Histone methyl marks are recognized by multiple domains, such as chromodomains <sup>9,10</sup> and PHD

fingers<sup>11,12</sup>, with high specificity on protein complexes that are recruited for downstream manipulation of the chromatin<sup>13</sup>.

Histone methylation plays a role in both activation and repression of transcription. H3K4<sup>14,15</sup>, H3K36<sup>16</sup> and H3K79<sup>17,18</sup> are known to be activating marks, with H3K4<sup>19</sup> and H3K36<sup>20</sup> methylation observed to associate with RNA polymerase II during transcriptional elongation. H3K9<sup>21</sup>, H3K27<sup>22,23</sup> and H4K20<sup>24</sup> methylation are known to be repressive marks, with K27me3 observed to be enriched in transcriptionally inactive genes such as the inactive X-chromosome<sup>25</sup>. However, a single methyl mark does not dictate the cell fate at that site, different permutations of methyl marks and other histone PTMs provides a secondary regulation in manipulation of DNA or recruitment of chromatin-modifying complexes.

The phenomenon describing the effects of other histone methyl marks and PTMs on histone modify enzymes is termed 'histone cross-talk'<sup>26</sup>. Several probable mechanisms to histone cross-talk have been mentioned in literature, including steric effects of adjacent histone marks, reader domains on chromatin-modifying complexes, and enhanced or lowered substrate binding affinity of histone-modifying enzymes depending on other modifications on the histone substrate<sup>1</sup>. Histone cross-talk prevents antagonistic marks, such as activating marks of H3K36 methylation and repressive marks of H3K27 methylation, to be located in the same genomic loci. There are exceptions to this, during the pluripotent state of embryonic stem-cells, bivalent domains with both H3K4 methylation and H3K9 methylation can be 'primed' to be activated or repressed depending on the cell fate upon differentiation<sup>27</sup>.

## **1.2 Histone methyltransferases and cancer**

The importance of regulating histone methylation can be reaffirmed by several diseases that are characterized by abnormal activities of histone methyltransferases, making them desirable targets

for therapeutics. H3K27 methyltransferase EZH2 is found to be overexpressed in many cancers<sup>28</sup> and activating mutations of EZH2 were also found in a subset of B-cell lymphomas<sup>29</sup>. This suggests that hypermethylating H3K27 promotes an oncogenic cell state. Probe development of an inhibitor against EZH2 has resulted in multiple compounds that were able to kill EZH2-driven cancer cell lines<sup>30 31</sup>. One of these compounds, EPZ-6438, is currently in clinical trials for patients with B-cell non-Hodgkin's lymphoma<sup>32</sup>.

In addition to overexpression and mutation, chromosomal translocation of histone methyltransferases can also lead to oncogenic behavior in the cell. H3K36 histone methyltransferase NSD1 gene, has been found to be fused with NUP98, nucleoporin 98 gene, in cases of childhood acute myeloid leukemia (AML)<sup>33</sup>. It has been observed that NUP98-NSD1 prevents the silencing of Hox-A gene that leads to the expression of proto-oncogenes, driving AML transformation<sup>34</sup>. Patient studies have identified NSD1-NUP98 fusion AML to be a subset of AML with a distinct gene expression signature and poor prognosis<sup>35</sup>.

There are several other histone methyltransferases identified to be somatically altered in cancer<sup>36</sup>. Developing inhibitors against these enzymes would be a good strategy in forming targeted therapy in cancer. Another approach could be identifying and targeting an orthogonal dependency of the altered histone methyltransferase, such as a subunit of the complex, instead. In mixed lineage leukemia, a H3K4 histone methyltransferase MLL1, was observed to be translocated. The aberrant activity of MLL1 led to the recruitment of another protein, DOT1L, and inducing a dependency of DOT1L catalytic activity in these cancers for survival<sup>37</sup>. Inhibition of DOT1L in these cancers was observed to be a sufficient strategy to ablate tumors in xenograph mice models<sup>38</sup>. A DOT1L inhibitor, EPZ-5676, is currently in clinical trials for leukemia bearing the MLL translocation<sup>39</sup>.

These studies recognize the validity of studying histone methyltransferase roles in the cell, expanding our understanding of their network and their contribution to cancer<sup>5,40</sup>.

### 1.3 Role of histone methyltransferase NSD2 in physiology and cancer

NSD2 (also known as WHSC1/MMSET), a member of the nuclear SET domain (NSD) family of histone methyltransferases, catalyzes the mono- and di- methylation of H3K36<sup>41,42</sup>, a marker in transcription elongation. NSD2 has also been observed to methylate H4K20 in response to DNA double strand breaks mediating 53BP1 recruitment to these sites and non-homologous end joining<sup>43,44</sup>. It has been suggested that NSD2 mediates Wnt<sup>45</sup>, androgen receptor<sup>46</sup>, and NF-kB signaling<sup>47</sup>. Haploinsufficiency of NSD2 leads to developmental defects that characterizes Wolf-hirshhorn syndrome (WHS)<sup>48</sup>.

NSD2's role in cancer has been frequently studied in the case of t(4;14)+ multiple myeloma (MM) which accounts for 15-20% of all cases of MM<sup>49</sup> and recently studied in the case of E1099K-NSD2 acute lymphoblastic leukemia (ALL) which accounts for 14% of t(12;21) ETV6-RUNX1-containing ALL<sup>50</sup>.

In t(4;14)+ MM, the cells undergoes a chromosomal translocation between chromosomes 4 and 14 placing NSD2 under the influence of a strong IgH enhancer that drives its overexpression in this cancers<sup>51</sup>. The high levels of NSD2 promotes several oncogenic characteristics in the cell - clonogenic growth, tumorigenicity and tumor aggressiveness<sup>52</sup>. Several studies with mice xenograph models have shown that the removal of translocated allele in t(4;14)+ MM cells leads to reduced NSD2 and H3K36me2 levels, curbing tumor growths in these mice<sup>41</sup>.

Apart from ALLs, E1099K-NSD2 has been observed in patients harboring distinct cancers<sup>53</sup>: acute lymphocytic leukemia<sup>50,54-56</sup>, chronic lymphocytic leukemia<sup>57</sup>, mantle cell lymphoma<sup>58</sup>, lung adenocarcinoma<sup>59</sup>, stomach adenocarcinoma<sup>60</sup>, multiple myeloma<sup>61</sup> and Wilms tumor<sup>62</sup>. Similar to t(4;14)+ MMs, studies with mice xenograph models have shown that knockdown of NSD2 in E1099K-ALLs successfully impaired growth of tumor<sup>50</sup>.

Beyond MM and ALL, NSD2 is broadly overexpressed in cancer, its expression is correlated with the severity of the disease and poor patient outcomes<sup>63,64</sup>. These provide strong evidence that NSD2 is involved in oncogenic growth in cancer and inhibition of NSD2 or its pathway can decrease this growth.

#### **1.4 Motivation and overview of the study**

As I have described, there are still unknowns in NSD2 biology, such as the complex it is part of and the redundancy of multiple H3K36 methylases (including NSD1). The understanding of NSD2 biology would be imperative in undertaking new strategies that could target NSD2 malignancies. As mice xenograph studies have suggested, an NSD2 inhibitor would be useful as a targeted therapy. However, development of NSD2 inhibitors are fraught with challenges, that I have personally encountered, such as lack of robust biochemical assays, use of complex substrates and abundance of false positives. I believe that the discovery of novel NSD2 complex and pathways would provide insights into developing an efficient and informative bioassay for assessing potential NSD2 inhibitor. In addition, it would illuminate orthogonal proteins that could be targeting downstream of NSD2.

In this thesis, I described my study into NSD2 and its mutant, E1099K-NSD2, through a series of biochemical, cellular and profiling experiments. The approach that I took to study NSD2 and E1099K-NSD2, is separated into 3 chapters.

In chapter 2, I described experiments that sought to explain that mechanism behind the enhanced catalytic activity of E1099K-NSD2 through a series of controlled biochemical experiments using recombinant NSD2-SET protein. Results of this experiment revealed that E1099K-SET NSD2 to have enhanced activity in complex substrates and not peptides, suggesting that the mutation might have additional interactions and affinity with higher-order structure substrates. In

addition, I studied the panel of ALL cell lines (6 WT-ALLs and 6 E1099K-ALLs) and show that E1099K-ALLs have increased H3K36me2 and decreased H3K27me3 levels, consistent with overexpression of WT-NSD2, suggesting that E1099K does not alter substrate specificity of NSD2. Further studies on gene expression using transcript levels obtained from CCLE database <sup>65</sup>, suggests that E1099K-ALLs drives oncogenic behavior that is similar to t(4;14)+ MMs.

In chapter 3, I described our efforts in a large-scale small-molecule profiling experiment <sup>59</sup> with several hematopoietic CCLs that has identified small-molecule inhibitors selective for HDAC1/2 that show greater potency for E1099K-ALLs relative to other hematopoietic CCLs. The experiment also identified FGFR3 inhibitors and IGF1R inhibitors to be selectively targeting t(4;14)+ MMs. Instead of targeting NSD2, our experiment was able to identify alternative pathways that can be targeted in these cancer which can be an alternative strategy in approaching NSD2 malignancies.

In chapter 4 and 5, I attempted to utilize small-molecule profiling in understanding NSD2 biology through a set of NSD2-CRISPR clones. I engineered a set of NSD2-CRISPR clones <sup>58,66</sup> to have an insertion of E1099K mutation or small deletions in NSD2-SET domain. RNA-seq was performed in this set of clones, and clones that remained to harbor WT-NSD2. Genesets related to oncogenesis and PRC2 complex were found to be upregulated in E1099K-NSD2 CRISPR clones. These clones were also profiled and found to have differential sensitivities towards treatment of NAMPT inhibitor (NAMPTi). Further examination of these clones revealed the phenotype to be mycoplasma dependent. I examined the chromatin landscape of the clones and observed that myco-(+) clones have higher levels of H3K27Ac compared to myco-(-) clones, suggesting an epigenetic nature for the mechanism of NAMPTi-resistance. The study has generated new tools for studying NSD2 biology and cancer biology, that can be further utilized by others in the cancer biology community.

## 1.5 References

1. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128**, 693–705 (2007).
2. Chi, P., Allis, C. D. & Wang, G. G. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. *Nat Rev Cancer* **10**, 457–469 (2010).
3. Popovic, R. & Licht, J. D. Emerging epigenetic targets and therapies in cancer medicine. *Cancer Discov* **2**, 405–413 (2012).
4. Zagni, C., Chiacchio, U. & Rescifina, A. Histone methyltransferase inhibitors: novel epigenetic agents for cancer treatment. *Curr Med Chem* **20**, 167–185 (2013).
5. Copeland, R. A., Solomon, M. E. & Richon, V. M. Protein methyltransferases as a target class for drug discovery. *Nature Reviews Drug Discovery* **8**, 724–732 (2009).
6. Martin, C. & Zhang, Y. The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* **6**, 838–849 (2005).
7. Morales, Y., Cáceres, T., May, K. & Hevel, J. M. Biochemistry and regulation of the protein arginine methyltransferases (PRMTs). *Arch. Biochem. Biophys.* **590**, 138–152 (2016).
8. Margueron, R. & Reinberg, D. Chromatin structure and the inheritance of epigenetic information. *Nat Rev Genet* **11**, 285–296 (2010).
9. Flanagan, J. F. *et al.* Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* **438**, 1181–1185 (2005).
10. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).
11. Wysocka, J. *et al.* A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86–90 (2006).
12. Shi, X. *et al.* ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* **442**, 96–99 (2006).

13. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res* **21**, 381–395 (2011).
14. Bernstein, B. E. *et al.* Methylation of histone H3 Lys 4 in coding regions of active genes. *Proceedings of the National Academy of Sciences* **99**, 8695–8700 (2002).
15. Santos-Rosa, H. *et al.* Active genes are tri-methylated at K4 of histone H3. *Nature* **419**, 407–411 (2002).
16. Lucio-Eterovic, A. K. *et al.* Role for the nuclear receptor-binding SET domain protein 1 (NSD1) methyltransferase in coordinating lysine 36 methylation at histone 3 with RNA polymerase II function. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16952–16957 (2010).
17. Steger, D. J. *et al.* DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol Cell Biol* **28**, 2825–2839 (2008).
18. Mohan, M. *et al.* Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). *Genes Dev.* **24**, 574–589 (2010).
19. Kim, T. & Buratowski, S. Dimethylation of H3K4 by Set1 Recruits the Set3 Histone Deacetylase Complex to 5' Transcribed Regions. *Cell* **137**, 259–272 (2009).
20. Kizer, K. O. *et al.* A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol Cell Biol* **25**, 3305–3316 (2005).
21. Volpe, T. A. *et al.* Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. *Science* **297**, 1833–1837 (2002).
22. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–1043 (2002).
23. Müller, J. *et al.* Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell* **111**, 197–208 (2002).

24. Vakoc, C. R., Sachdeva, M. M., Wang, H. & Blobel, G. A. Profile of histone lysine methylation across transcribed mammalian chromatin. *Mol Cell Biol* **26**, 9185–9195 (2006).
25. Plath, K. *et al.* Role of Histone H3 Lysine 27 Methylation in X Inactivation. *Science* **300**, 131–135 (2003).
26. Lee, J.-S., Smith, E. & Shilatifard, A. The Language of Histone Crosstalk. *Cell* **142**, 682–685 (2010).
27. Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125**, 315–326 (2006).
28. Simon, J. A. & Lange, C. A. Roles of the EZH2 histone methyltransferase in cancer epigenetics. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **647**, 21–29 (2008).
29. Bödör, C. *et al.* EZH2 mutations are frequent and represent an early event in follicular lymphoma. *Blood* **122**, 3165–3168 (2013).
30. McCabe, M. T. *et al.* EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* **492**, 108–112 (2012).
31. Campbell, J. E. *et al.* EPZ011989, A Potent, Orally-Available EZH2 Inhibitor with Robust in Vivo Activity. *ACS Med. Chem. Lett.* **6**, 491–495 (2015).
32. Knutson, S. K., Kawano, S. & Minoshima, Y. *Selective inhibition of EZH2 by EPZ-6438 leads to potent antitumor activity in EZH2-mutant non-Hodgkin lymphoma.* (Molecular cancer ..., 2014). doi:10.1016/j.febslet.2012.07.066/full
33. Jaju, R. J. *et al.* A novel gene, NSD1, is fused to NUP98 in the t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. *Blood* **98**, 1264–1267 (2001).
34. Wang, G. G., Cai, L., Pasillas, M. P. & Kamps, M. P. NUP98–NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nat Cell Biol* **9**, 804–812 (2007).

35. Hollink, I. H. I. M. *et al.* NUP98/NSD1 characterizes a novel poor prognostic group in acute myeloid leukemia with a distinct HOX gene expression pattern. *Blood* **118**, 3645–3656 (2011).
36. Albert, M. & Helin, K. Histone methyltransferases in cancer. *Seminars in Cell and Developmental Biology* **21**, 209–220 (2010).
37. Daigle, S. R. *et al.* Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell* **20**, 53–65 (2011).
38. Chen, L. *et al.* Abrogation of MLL-AF10 and CALM-AF10-mediated transformation through genetic inactivation or pharmacological inhibition of the H3K79 methyltransferase Dot1l. *Leukemia* **27**, 813–822 (2013).
39. Stein, E. M. *et al.* The DOT1L Inhibitor EPZ-5676: Safety and Activity in Relapsed/Refractory Patients with MLL-Rearranged Leukemia. *Blood* **124**, 387–387 (2014).
40. Arrowsmith, C. H., Bountra, C., Fish, P. V., Lee, K. & Schapira, M. Epigenetic protein families: a new frontier for drug discovery. *Nature Reviews Drug Discovery* **11**, 384–400 (2012).
41. Kuo, A. J. *et al.* NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Molecular Cell* **44**, 609–620 (2011).
42. Li, Y. *et al.* The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. *J Biol Chem* **284**, 34283–34295 (2009).
43. Hajdu, I., Ciccia, A., Lewis, S. M. & Elledge, S. J. Wolf-Hirschhorn syndrome candidate 1 is involved in the cellular response to DNA damage. *Proceedings of the National Academy of Sciences* **108**, 13130–13134 (2011).
44. Pei, H. *et al.* MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. *Nature* **470**, 124–128 (2011).
45. Toyokawa, G. *et al.* Histone lysine methyltransferase Wolf-Hirschhorn syndrome candidate 1 is involved in human carcinogenesis through regulation of the Wnt pathway. *Neoplasia* **13**,

- 887–898 (2011).
46. Kang, H.-B. *et al.* The histone methyltransferase, NSD2, enhances androgen receptor-mediated transcription. *FEBS Lett* **583**, 1880–1886 (2009).
  47. Yang, P. *et al.* Histone methyltransferase NSD2/MMSET mediates constitutive NF-kappaB signaling for cancer cell proliferation, survival, and tumor growth via a feed-forward loop. *Mol Cell Biol* **32**, 3121–3131 (2012).
  48. Stec, I. *et al.* WHSC1, a 90 kb SET Domain-Containing Gene, Expressed in Early Development and Homologous to a Drosophila Dysmorphia Gene Maps in the Wolf-Hirschhorn Syndrome Critical Region and is Fused to IgH in t(1;14) Multiple Myeloma. *Hum. Mol. Genet.* **7**, 1071–1082 (1998).
  49. Kalff, A. & Spencer, A. The t(4;14) translocation and FGFR3 overexpression in multiple myeloma: prognostic implications and current clinical strategies. *Blood Cancer Journal* **2**, e89 (2012).
  50. Jaffe, J. D. *et al.* Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nat Genet* **45**, 1386–1391 (2013).
  51. Chesi, M. *et al.* The t(4;14) Translocation in Myeloma Dysregulates Both FGFR3 and a Novel Gene, MMSET, Resulting in IgH/MMSET Hybrid Transcripts. *Blood* **92**, 3025–3034 (1998).
  52. Martinez-Garcia, E. *et al.* The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. *Blood* **117**, 211–220 (2011).
  53. Forbes, S. A. *et al.* COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucl. Acids Res.* **43**, D805–11 (2015).
  54. Holmfeldt, L. *et al.* The genomic landscape of hypodiploid acute lymphoblastic leukemia. *Nat Genet* **45**, 242–252 (2013).

55. Neumann, M. *et al.* Mutational spectrum of adult T-ALL. *Oncotarget* **6**, 2754–2766
56. Ma, X. *et al.* Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. *Nat Comms* **6**, 6604 (2015).
57. Houtkooper, R. H., Cantó, C., Wanders, R. J. & Auwerx, J. The Secret Life of NAD<sup>+</sup>: An Old Metabolite Controlling New Metabolic Signaling Pathways. *Endocrine Reviews* **31**, 194–223 (2010).
58. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**, 2281–2308 (2013).
59. Basu, A. *et al.* An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* **154**, 1151–1161 (2013).
60. Zeileis, A. & Grothendieck, G. zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software* **14**, 1–27 (2005).
61. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods* **10**, 1096–1098 (2013).
62. Gould, J., Getz, G., Monti, S., Reich, M. & Mesirov, J. P. Comparative gene marker selection suite. *Bioinformatics* **22**, 1924–1925 (2006).
63. Hudlebusch, H. R. *et al.* The histone methyltransferase and putative oncoprotein MMSET is overexpressed in a large variety of human tumors. *Clin Cancer Res* **17**, 2919–2933 (2011).
64. Kassambara, A., Klein, B. & Moreaux, J. MMSET is overexpressed in cancers: Link with tumor aggressiveness. *Biochemical and Biophysical Research Communications* **379**, 840–845 (2009).
65. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
66. Doudna, J. A. & Charpentier, E. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096–1258096 (2014).

# Chapter II

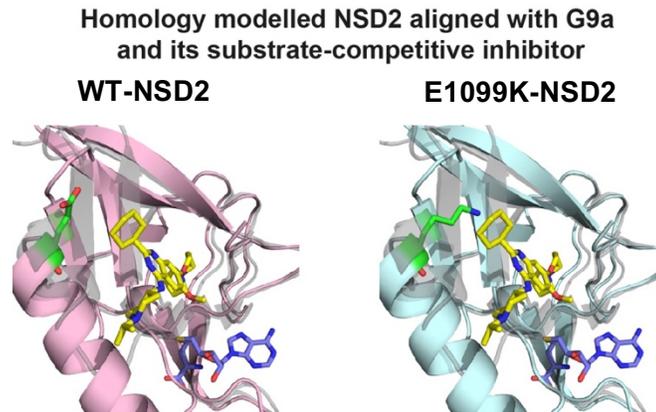
## Biochemical and cellular characterization of E1099K-NSD2

### Collaborators contributions:

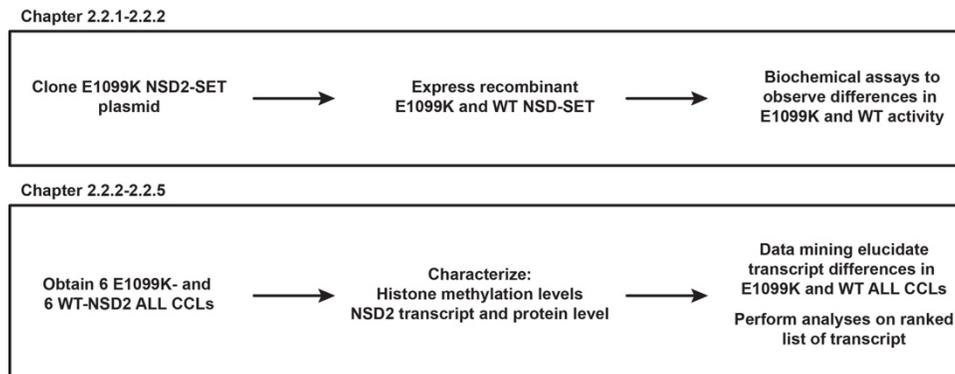
- **Dr. LaTese Briggs** optimized the expression of recombinant NSD2-SET.
- **Dr. Drew Adams** aided in the site directed mutagenesis and expression of recombinant NSD2.
- **Dr. Joshiawa Paulk** suggested radiometric methods for methylation assays.
- **Dr. Jaime Cheah** provided the cell lines used in the study.
- **Roodolph St. Pierre** shared buffer for methylation assays.

## 2.1 Introduction

In this chapter, I will be describing the biochemical and cellular characterization of E1099K-NSD2. The E1099K mutation lies within the catalytic SET domain of NSD2 and we hypothesize E1099K would activate NSD2 catalytic activity. Homology modelling of NSD2-SET domain using NSD1 crystal structure and G9a crystal structure revealed that E1099K residue lies within the substrate-binding region of the SET domain, predicting that E1099K might interact differently with NSD2 substrate (Figure 2.1a). This was also observed in literature <sup>1,2</sup>. We studied this hypothesis through biochemical and cellular experiments. Brief scheme of this chapter is shown below (Figure 2.1b).



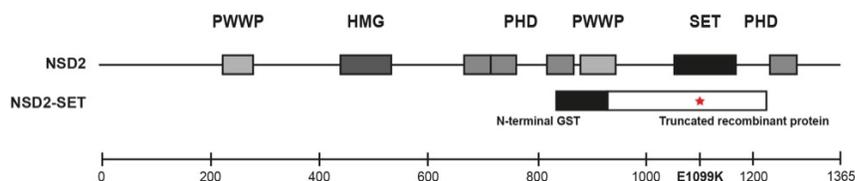
**Figure 2.1a.** NSD2 WT-SET and E1099K-SET generated by RaptorX <sup>3</sup> using crystal structure of NSD1 as template. Alignment of the homologous structures and G9a crystal structure including substrate-competitive inhibitor (in yellow) and co-factor (in blue) are shown. 1099 residue (green) of E1099K-NSD2 appears to interact with the substrate binding pocket more than WT-NSD2.



**Figure 2.1b** Scheme of experiments described in this chapter.

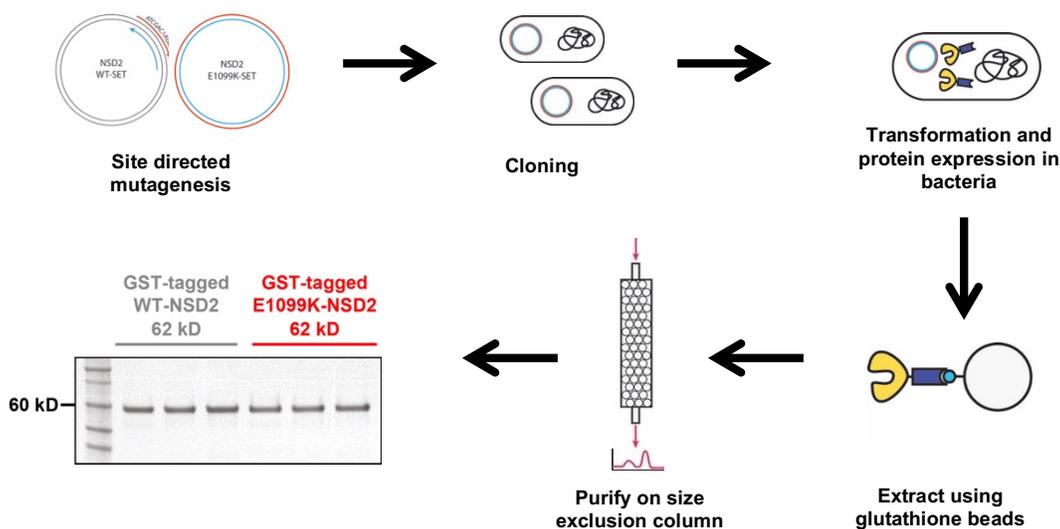
## 2.2 Results and Discussion

### 2.2.1 Expression and purification of recombinant WT- and E1099K-NSD2



**Figure 2.2.1a.** The truncated recombinant NSD2 SET expressed spans amino acids 941-1240 of the full-length protein with N-terminal Glutathione S-Transferase (GST) tag.

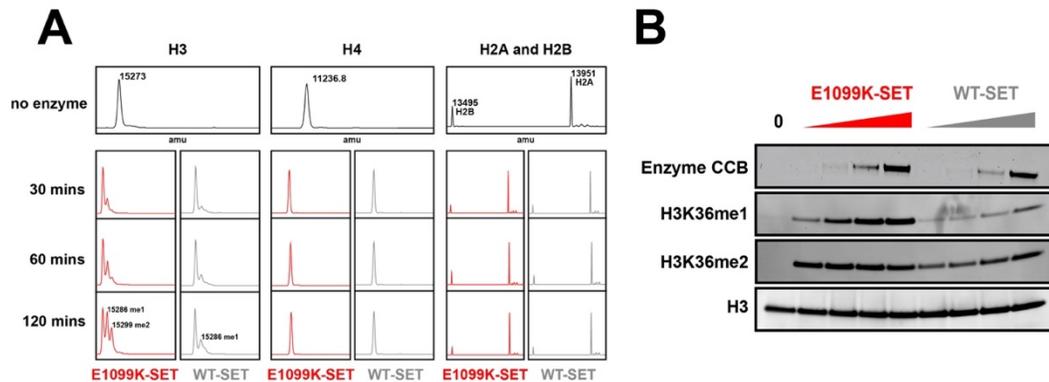
We expressed the truncated NSD2-SET protein used in the study from a plasmid that contains the catalytic SET domain of NSD2 with a N-terminal GST tag (Figure 2.2.1a). We cloned the NSD2 E1099K-SET plasmid using site directed mutagenesis from the WT plasmid. WT and E1099K was expressed in *e. coli*, extracted using glutathione beads and purified using size exclusion column to give a clean product (Figure 2.21b).



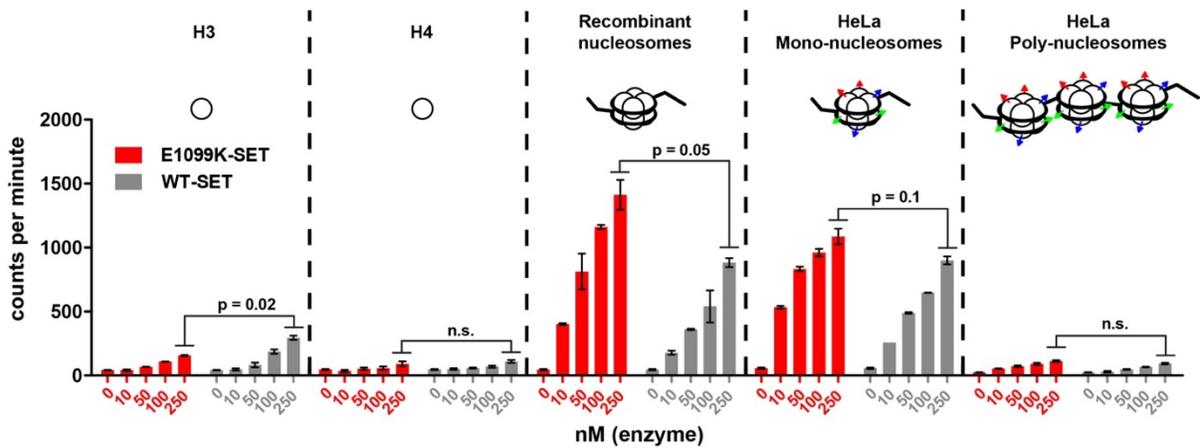
**Figure 2.2.1b** Expression scheme of recombinant NSD2 SET protein.

## 2.2.2 WT- and E1099K-NSD2 differ in activity depending on substrate

We tested the recombinant protein in various methylation assays to compare the activity of E1099K-SET against WT-SET. Methylation assays were performed in the presence of a substrate containing H3K36 residue (in this case, recombinant nucleosomes) and co-factor S-adenosyl methionine (SAM). The products of the assay were detected via electrospray ionization mass spectrometry (ESI-MS) over 30, 60 and 120 minutes. Both WT-SET and E1099K-SET methylated H3 with methylation detected in both reactions over the course of 120 minutes. E1099K-SET is more reactive than WT-SET, producing higher amounts of mono- and di-methylated H3 at each time point. Both E1099K-SET and WT-SET did not methylate other histone residues, H4, H2A and H2B, which suggests that the mutation did not change substrate specificity of NSD2 (Figure 2.2.2a A). We reconfirmed mono- and di-methylation of H3 with an immunoblot of the methylated-K36 residue. E1099K-SET was also observed to methylate H3K36 to a greater extent compared to WT-SET (Figure 2.2.2a B).



**Figure 2.2.2a A)** Spectra of nucleosomal subunits in methylation reaction detected via ESI-MS. **B)** Immunoblot of methylation reaction, blotting for H3K36 mono- and di-methylation.



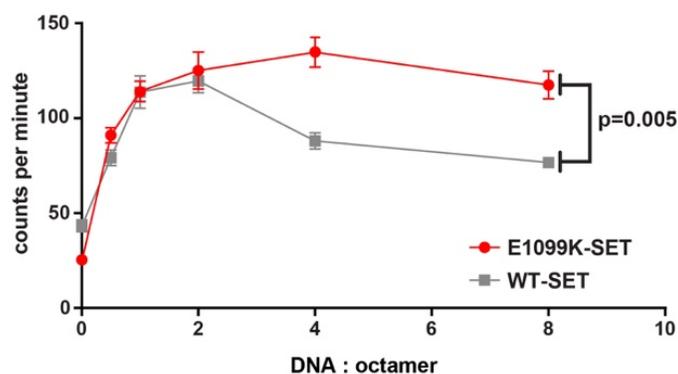
**Figure 2.2.2b** Scintillation counts of methylation assay of NSD2-SET with substrates of different complexity. Pictorial examples of substrate complexity are shown: circles represent histone proteins, lines wrapped around group of circles represent DNA and colored arrowheads represent modifications on histone proteins. Error bars represent standard error of mean of three replicates.

We further studied E1099K-SET enhanced activity in additional methylation assays.

Radiometric methylation assay was performed with a panel of different substrates. The panel of substrate was chosen with increase complexity: single histone proteins (H3 and H4), non-modified mono-nucleosome (recombinant nucleosomes), modified mono-nucleosomes (mono-nucleosomes extracted from HeLa cells) and modified poly-nucleosomes (poly-nucleosomes extracted from HeLa cells). The enzymes were not significantly active when using simpler substrates, histone H3 and H4. Corroborating with previous results, E1099K-SET is more active on recombinant nucleosomes compared to WT-SET. To test the possibility of other post-translational histone marks enhancing E1099K-SET activity, we repeated the assay with native HeLa mono-nucleosomes. HeLa mono-nucleosomes were observed to be less methylated compared to recombinant nucleosomes, possibly due to the lack of available methylation sites on the native substrate. The effects of a euchromatin environment in the cell was also explored using native HeLa poly-nucleosomes, a quaternary protein

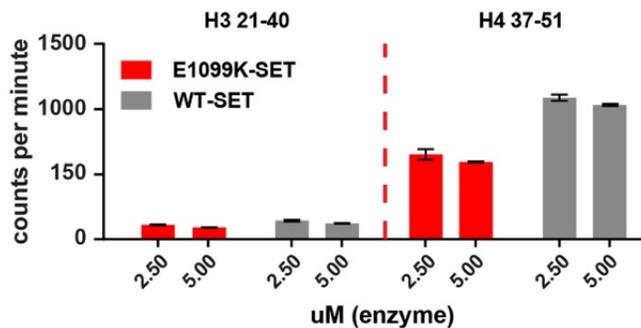
structure, but the methylation activity on this substrate was very low, possibly due to steric effects of the bulky substrate (Figure 2.2.2b).

We hypothesized that E1099K-SET's enhanced activity towards nucleosomes could be due to DNA of the nucleosome interacting E1099K-SET. It has been reported that DNA can act as an allosteric effector of NSD2, enhancing its activity and directing its specificity towards H3K36<sup>4</sup>. Another study on NSD1, suggests that close proximity of nucleosomal DNA to H3K36 might stabilize post-SET loop of NSD1, opening up access to H3K36<sup>5</sup>. We tested the hypothesis using a radiometric methylation assay with an octamer-DNA substrate. Different DNA concentrations were tested and the same enhancement of activity was recorded for both WT- and E1099K-SET with increasing concentrations of DNA. The difference in WT- and E1099K-SET was apparent at higher concentrations of DNA. At higher concentrations of DNA, we observed DNA to have an inhibitory effect on both enzyme's activity; however, WT-SET is more susceptible to this inhibition compared to E1099K-SET (Figure 2.2.2c). This suggests that E1099K-NSD2 is less inhibited by nucleosomal DNA in areas of the genome that were inaccessible to WT-NSD2, leading to higher levels of H3K36 methylation in cell lines.



**Figure 2.2.2c** Scintillation counts of methylation assay with differing DNA and octamer ratios. Error bars represent standard error of mean of three replicates.

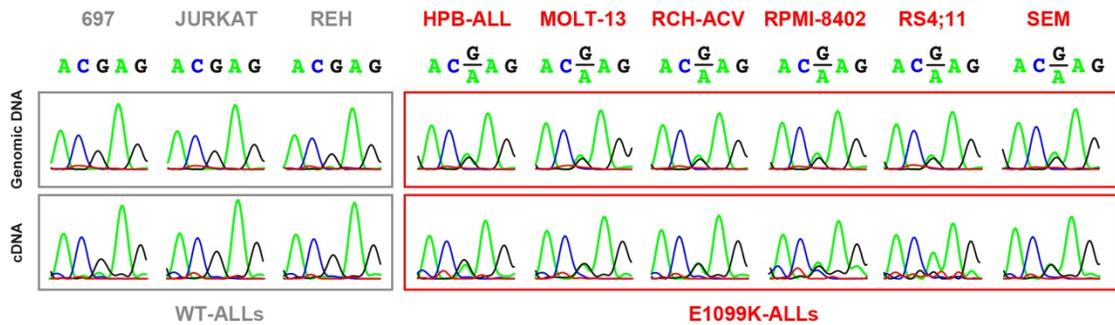
We also hypothesized that histone modifications could affect E1099K- and WT-SET differently. We tested this hypothesis using histone peptides in a radiometric methylation assay. NSD2-SET was more active when using a H4 peptide H4(37-51) compared to its native substrate H3 peptide H3(21-40) (Figure 2.2.2d). This suggests that peptide substrates might not be relevant for studying the catalytic activity using the truncated enzyme. The peptide assay does not reflect physiological activity of NSD2. It would be insufficient and inaccurate to study the interaction between different histone modification and NSD2-SET using this assay.



**Figure 2.2.2d** Scintillation counts of methylation assay with H3 and H4 peptide. Error bars represent standard error of mean of three replicates.

### 2.2.3 E1099K-NSD2 ALLs are heterozygous mutants

The E1099K mutation is enriched in ALLs, we studied 12 ALLs with different NSD2 mutation status (6 WT, 6 E1099K). The genomic DNA and RNA of the ALLs were genotyped to assess NSD2 status in their genome and transcriptome. E1099K mutation was observed to be heterozygous in the genome and in the transcript (Figure 2.2.3a). Homozygous mutation was not observed, which suggests that WT-NSD2 is not redundant in these cell lines and that E1099K-NSD2 could be functionally different from WT-NSD2.

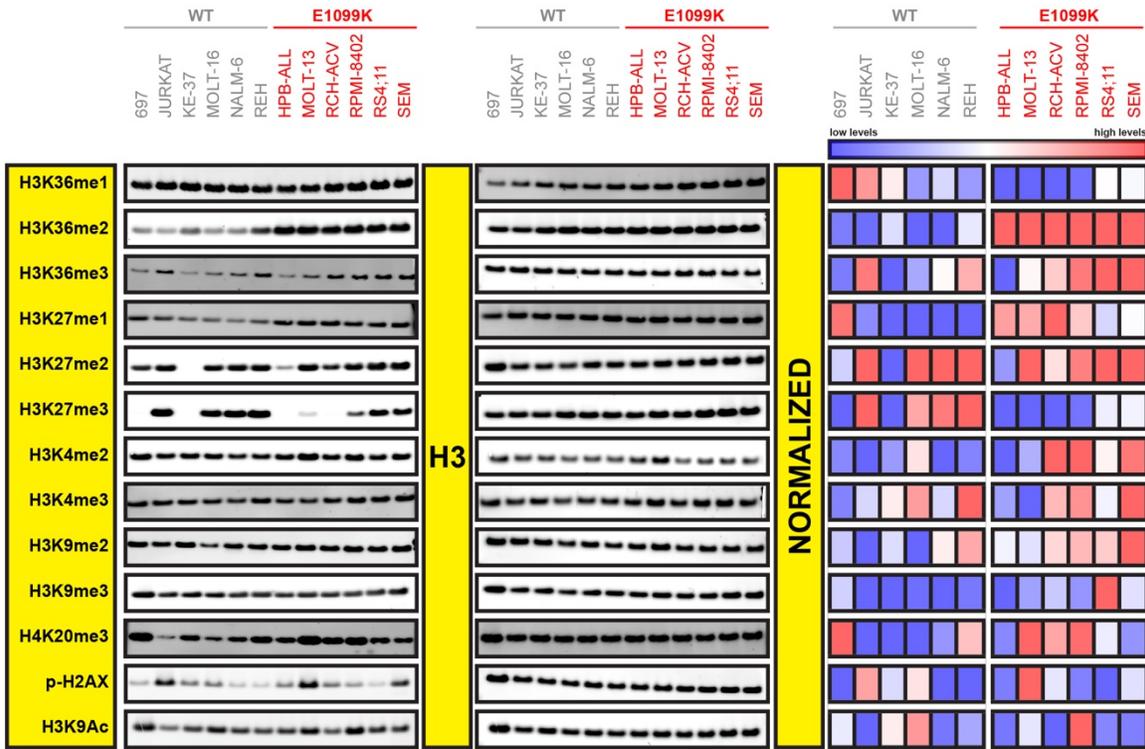


**Figure 2.2.3a** Chromatograph of sequencing results within the NSD2 sequence. 3 out of 6 WT-ALLs are shown as a representation. In one allele, guanine has been mutated to alanine, causing the E1099K mutation in NSD2.

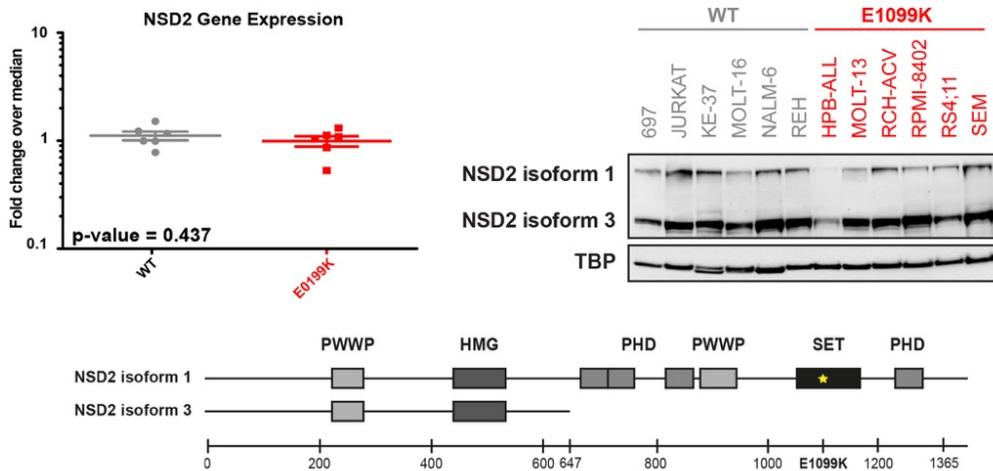
#### 2.2.4 E1099K-NSD2 ALLs have distinct methylation states

We evaluated the effect of the E1099K mutation on the chromatin landscape of E1099K-ALLs by profiling 13 histone modifications in 6 WT-ALLs and 6 E1099K-ALLs. E1099K-ALLs have elevated levels of H3K36me2 and lower levels of H3K27me3 relative to WT-ALLs. E1099K-ALLs also have elevated levels of H3K36me3 and H3K27me1 and lower levels of H3K36me1 relative to WT-ALLs. The difference in H3K36 and H3K27 methylation levels in these ALLs stemmed from the hypermethylation of H3K36me2 by E1099K-NSD2. There was no observable trend for the other histone modifications probed (Figure 2.2.4a).

We also assessed the level of NSD2 transcript and protein in these ALLs using qPCR and immunoblot respectively and observed that NSD2 levels do not correlate with the increased H3K36me2 levels in E1099K-ALLs (Figure 2.2.4b). This suggests E1099K to be an activating mutation of NSD2, enhancing the cell's ability to methylate H3K36 residue regardless of the level of NSD2 in the cell.



**Figure 2.2.4a** Left: Immunoblot of histone modifications probed. Center: Immunoblot of loading control H3. Right: Heatmap of quantitated blot normalized across all ALLs tested.

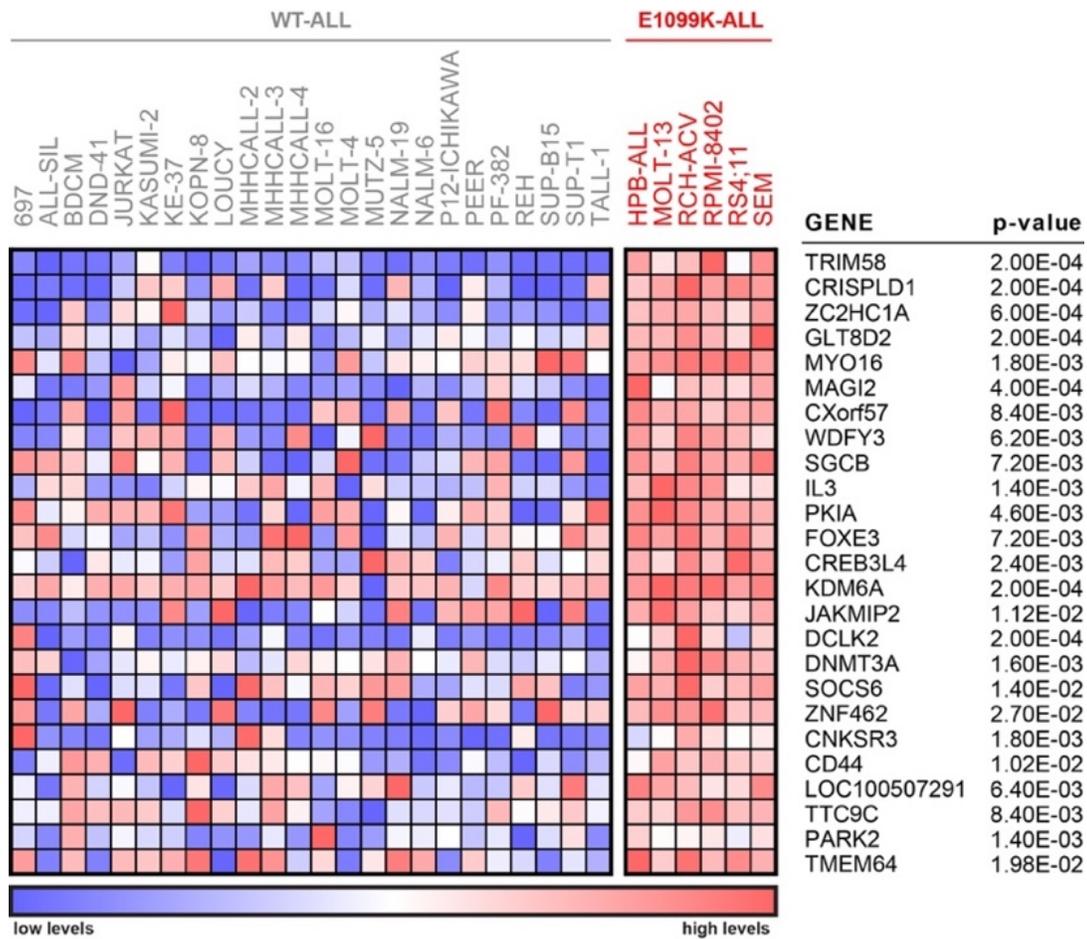


**Figure 2.2.4b** Top left: NSD2 gene expression measured through qPCR, lines represent mean and standard error of mean of two replicates. Top right: Immunoblot of NSD2, TBP was measured as a loading control. Bottom: A representation of NSD2 isoforms detected by the immunoblot, isoform

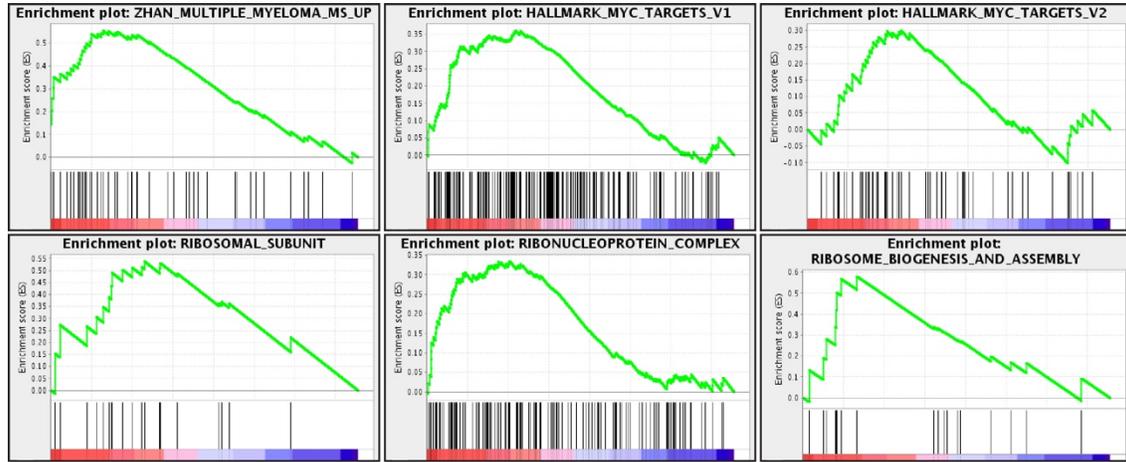
1 contains the catalytic SET domain of NSD2.

### 2.2.5 E1099K-NSD2 ALLs have NSD2-activated gene expression signatures

We obtained publicly available gene expression profiles <sup>6</sup> for available ALLs (24 WT-ALLs and 6 E1099K-ALLs) from the cancer cell line encyclopedia (CCLE). Differential expression analysis <sup>7</sup> was utilized to identify upregulated transcripts in E1099K-ALLs (Supplemental Table 2.1, Figure 2.2.5a). 18898 genes were included in this analysis and using bonferroni correction, genes that are significant have a p-value of  $< 2.65E-05$ . Based on this p-value, there were no genes that are statistically significantly upregulated in E1099K-ALLs. Instead of using p-value to rank transcripts, we used the score calculated from differential expression analysis as an input to perform secondary analyses and identify significant gene sets or protein networks within the E1099K ALLs.



**Figure 2.2.5a** 25 top scoring transcripts that are upregulated in E1099K-ALLs compared to WT-ALLs with their respective p-values.

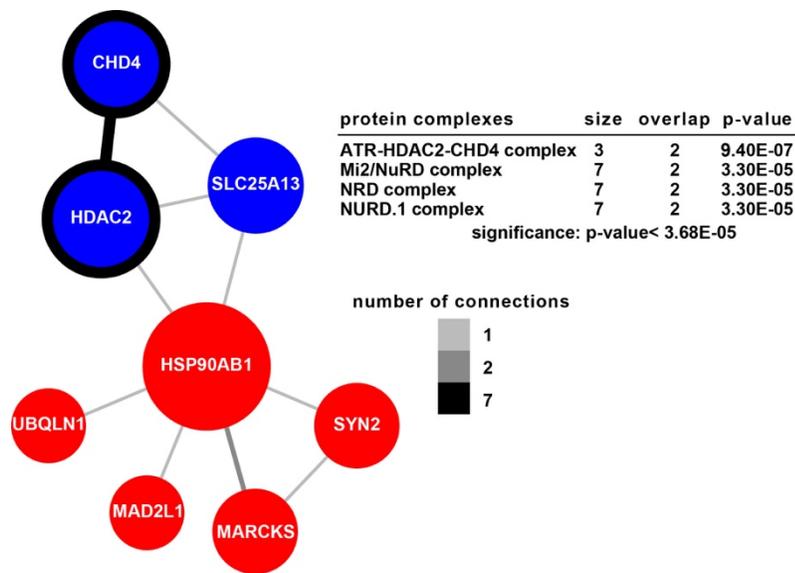


**Figure 2.2.5b.** Enrichment plots of E1099K-ALL transcript within gene sets of interest.

We performed secondary analysis using gene set enrichment analysis (GSEA)<sup>8</sup> to identify pre-annotated gene sets that are enriched in E1099K-ALLs transcript. A NSD2-centric gene set was included in the cluster of gene sets analyzed - ZHAN\_MULTIPLE\_MYELOMA\_MS\_UP, a gene set containing 50 top upregulated genes in t(4;14)+ MM<sup>9</sup>. GSEA found E1099K-ALL upregulated genes are significantly enriched in ZHAN\_MULTIPLE\_MYELOMA\_MS\_UP (Figure 2.2.5b, Supplemental Table 2.2). Other enriched gene sets include gene ontology (GO) sets that relates to ribosomal assembly and hallmark gene set that relates to c-Myc targets. Ribosomal assembly genes and c-MYC expression were previously linked to t(4;14)+<sup>10,11</sup>, suggesting E1099K-NSD2 could be behaving similar to overexpression of NSD2.

We utilized another secondary analysis tool, using visualization of genome networks (GeNets)<sup>12</sup>, to identify groups of classifier transcripts upregulated in E1099K-ALLs that share common annotations, such as being in the same biological pathway, the same protein complex, etc. Pathway analysis using InWeb protein-protein interaction<sup>13</sup> was applied to the top 100 scoring

upregulated transcripts in E1099K-ALLs. 2 communities of proteins with common annotations were identified, including HDAC2 and CHD4 (Figure 2.2.5c). Additional analysis within the GeNets platform was performed using the Comprehensive Resource of Mammalian protein complexes (CORUM) <sup>14</sup>. It revealed a number of HDAC2-CHD4 complexes to be significantly enriched in E1099K-ALLs. NSD2 have previously been shown to co-purify with HDAC2, supporting the idea that HDAC2 activity could be connected to E1099K-NSD2 <sup>15,16</sup>. The p-value for significance in CORUM analysis was calculated with bonferroni correction as 1359 protein complexes were tested.



**Figure 2.2.5c** GeNets analysis identified 2 communities (colored in red and in blue) upregulated in E1099K-ALLs. The color of the connectors between the genes represent the number of connections found in CORUM between the 2 genes. HDAC2 and CHD4 has the most connections as they are found together in multiple protein complexes.

Lastly, we used ChromNet<sup>17</sup> to identify chromatin signals that NSD2 might correlate with. ChromNet is an online analysis database that infer a network of chromatin interactions through ChIP-seq data from the encyclopedia of DNA elements (ENCODE) project <sup>18</sup>. The ENCODE database has one ChIP-seq dataset for NSD2 and it correlates significantly to the ChIP-seq dataset

for H3K9Ac. Both datasets were obtained from K562 cell line<sup>19</sup>. As H3K9 is a substrate for HDAC2<sup>20</sup>, this finding strengthens NSD2 possible relationship with HDAC2.

### **2.2.6 Concluding remarks**

In this chapter, I have detailed biochemical and cellular experiments that emphasized E1099K-NSD2 to be an activating mutation that causes the cell to have an expression profile similar to overexpression of WT-NSD2.

Biochemically, E1099K-SET has elevated catalytic activity against physiologically relevant nucleosomal substrates, but diminished activity against peptide and histone substrates, indicating that E1099K-SET benefits from interactions with nucleosomal DNA. We observed that increasing concentration of DNA in methylation assays did not inhibit activity of E1099K-SET, and did significantly impaired activity of WT-SET. It is possible that the E1099K mutation alters the enzyme structure in a way that changes its interaction with DNA and enhances its methylation activity. Further studies should be performed using full-length NSD2 enzyme and ‘designer’ nucleosomes<sup>21</sup> to fully elucidate the mechanism behind the activating mutation, E1099K-NSD2.

The biochemical results agreed with the increased levels of K36me2 observed in E1099K-ALLs, which was accompanied by the decreased levels of K27me3 – phenomenon that have been observed by others<sup>2,22-24</sup>. Despite the differences in histone methyl marks, the significant differences in the transcript levels of E1099K-ALLs and WT-ALLs seemed to be obscured by their common lineage. E1099K-NSD2’s effect on the cell could be loci specific that would be difficult to assessed globally through immunoblot or transcriptome profiling. Further studies should be performed using ChIP-seq of K36me2 and NSD2 to identify loci effects of E1099K-NSD2 in the cells.

## **2.3 Experimental methods**

### **Homology modelling of NSD2-SET**

RaptorX structure prediction tool<sup>3</sup> was used to build homology model of NSD2-SET using NSD1-SET crystal structure (PDB ID: 3OOI<sup>5</sup>) obtained from the protein data bank (PDB).

### **Expression and purification of NSD2-SET recombinant protein**

pGEX6p-1 NSD2-SET (amino acids 941-1240) plasmid was a generous gift from Dr. Danny Reinberg<sup>4</sup>. E1099K mutation was generated using QuikChange II Site-Directed Mutagenesis Kit (Agilent). GST-SET plasmids were transformed in BL21-CodonPlus(DE3)-RIPL *Escherichia coli* (Agilent). GST-SET recombinant proteins were expressed and purified as previously described in<sup>4</sup>.

### **Mass spectrometric analysis**

50 nM NSD2-SET enzymes were incubated with 100 nM recombinant mono-nucleosomes (Epiccypher) and quenched at appropriate time points with equal volume of 6M Urea. Methylation was determined through LC-MS (Agilent 6220 ESI-TOF) using a PLRP-S column (Agilent).

### **Methylation assays**

Histone H3 and H4 were purchased from New England Biolabs. Recombinant mono-nucleosomes, HeLa mono-nucleosomes, HeLa poly-nucleosomes and recombinant octamers were purchased from Epiccypher. DNA duplex sequence used in octamer methylation assay was: 5'-CTCTCTTTGAGGACACCAACCTGGCGGCCATCCACGCCAAG-3'<sup>4</sup>. NSD2-SET enzymes were incubated with 250 nM substrate, 1 uM 3H-SAM (Perkin-Elmer) and 6 uM SAM (New England Biolabs) in a buffer containing 50 mM 8.5 Tris, 1 M NaCl, 0.1% Tween20, 1 mM DTT (Sigma-Aldrich). Reaction was quenched at 120 minutes by addition to a well in a pre-washed 96-well HA-filter binding plate (Millipore) containing 100 uL of ice-cold 20% TCA (Sigma-Aldrich). Plate was incubated at 4C for 1 hour. Plate was washed 3 times with ice-cold 5% TCA using vacuum filtration and dried for 1 hour. 50 uL of MicroScint PS (Perkin-Elmer) was added to each well and

incubated for at least 30 minutes before reading on TopCount Scintillation Counter (Perkin-Elmer). Immunoblotting methylation assay was performed with 7  $\mu$ M of SAM (New England Biolabs) and analyzed using standard western blotting techniques as described below. Octamer-DNA assay was performed using the same protocol with 10 nM of NSD2-SET enzymes and 50 nM of octamer in a 2 hour reaction.

### **Radiometric peptide methylation assay**

H3 21-40 peptide (LATKAARKSAPATGGVKKPH) was purchased from Anaspec and H4 37-51 (RLARRGGVKRISGLI) was a generous gift from Roodolph St. Pierre and Dr. James Bradner. 25 nM of NSD2-SET enzymes were incubated with 25  $\mu$ M of N-terminal biotinylated peptides and 3H-SAM (1  $\mu$ M). Reaction was quenched with 1 mM SAM in a 96-well DELFIA streptavidin plate (Perkin-Elmer). Quenched reaction was incubated for 1 hour and washed 3 times with PBS. Peptides were eluted after 1 hour incubation with 100  $\mu$ L of buffer containing 70% acetonitrile, 5% formic acid and 1 mM biotin. 80  $\mu$ L of eluent was added to 5 mL of Ultima Gold Cocktail (Perkin Elmer) and read on Tri-Carb scintillation counter (Perkin-Elmer).

### **Antibodies**

Antibodies used in this study are listed in Appendix A.

### **Cell lines**

Cell lines used in expression analysis are listed in the Appendix B. 697, NALM-6, REH, JURKAT, KE-37, RCH-ACV, RPMI-8402 and RS4;11 were cultured in RPMI1640 (Gibco) supplemented with 10% fetal bovine serum (Gibco). HPB-ALL, MOLT-13 and MOLT-16 were cultured in RPMI1640 supplemented with 20% fetal bovine serum. SEM was cultured in IMDM (Gibco), supplemented

with 10% fetal bovine serum. All media was supplemented with 1% penicillin/streptomycin (Gibco).

### **Genotyping and quantitative RT-PCR analysis**

Genomic DNA and RNA of cell lines were obtained using AllPrep DNA/RNA Mini Kit (Qiagen).

cDNA was obtained via RT using SuperScript VILO MasterMix (Life Technologies). Allele

genotyping was studied by amplifying regions of NSD2 using primer pairs: 5'-

ACTGCTGACCCTGATGTATTT-3' and 5'-CAGATGAGGAAGTCTCAGCATC-3'. cDNA

genotyping was studied using primer pairs: 5'-TGAAGCACGAGATTGGAGAA-3' and 5'-

CATTGCCAGACAATCGAG-3'. Sanger sequencing were performed by Eton Bioscience Inc.

using the same primers and analyzed using plasmid editor, ApE. Quantitative RT-PCR was

performed on LightCycler 480 (Roche) using TaqMan Mastermix (Roche) following manufacturer's

protocol. TaqMan primers (Life Technologies) used in this study were: NSD2 (Hs00370212\_m1)

and GAPDH (Hs03929097\_g1).

### **Immunoblot and analysis of protein levels**

For histone extractions, cells were washed once with PBS and lysed with 0.1% Triton X-100 (Sigma-

Aldrich) in PBS supplemented with HALT protease inhibitor cocktail (Thermo Scientific) on ice for

10 minutes. Cells were spun down at 10,000 rcf for 5 minutes at 4C. The pellet was washed once

with 0.1% Triton X-100 in PBS and spun down again. Pellet was resuspended in 2N HCl and left

overnight in 4C. The next day, mixture was spun down at 10,000 rcf for 5 minutes at 4C. The

supernatant was transferred to a new tube and quenched with 40% volume of 1M 7.5 Tris (Life

Technologies). For nuclear extractions, lysate was obtained using NE-PER Nuclear and Cytoplasmic

Extraction Reagents (Thermo Scientific) supplemented with HALT protease inhibitor (Thermo

Scientific).

Protein levels in lysates were quantified using Bradford protein assay (Bio-Rad). Lysates were denatured and reduced, ran on acrylamide Bis-Tris gels (Life Technologies) and transferred using iBolt (Life Technologies) for analysis.

Blots were blocked with Odyssey TBS blocking buffer (LI-COR) and incubated overnight with primary antibody diluted in 50% blocking buffer in TBS-T. The next day, blots were washed 3 times with TBS-T and incubated with an appropriate IR-Dye conjugated secondary antibody (LI-COR). Blots were washed 3 times with TBS-T, 2 times with water and imaged with Odyssey CLx (LI-COR). NSD2 blots were blocked with 5% nonfat dry milk (Cell Signaling) in TBS-T and incubated overnight with primary antibody diluted in blocking buffer. The next day, NSD2 blots were washed and incubated with HRP-conjugated secondary antibody diluted in blocking buffer. NSD2 blots were developed using SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific) and imaged on Image Station 4000MM (Kodak).

Western blots of histone modifications and H3 in this chapter were quantified using ImageJ software. Values of each histone modification were normalized across cell lines to their respective H3 bands. In each set of histone modification, the values were further normalized to the sample with the lowest level of that histone modification. Values were saved as a gct file and heat map generated using GENE-E software <sup>25</sup>.

### **Gene expression analysis of E1099K and WT NSD2 cell lines**

CCLE (<http://www.broadinstitute.org/ccle/home>) gene expression data <sup>6</sup> for 30 ALL lines were downloaded. The expression data was analyzed using comparative marker selection module <sup>26</sup> in GenePattern <sup>7</sup> to rank differential transcripts in E1099K-ALLs compared to WT-ALLs. The ranked

list was then used as input for Gene Set Enrichment Analysis <sup>8</sup> using the C2 (curated gene sets), C5 (gene ontology) and H (hallmark gene sets) collections for the analysis. The top 100 enriched transcripts in E1099K-ALLs were used as input for GeNets <sup>12</sup> pathway analysis using InWeb protein-protein interaction network <sup>13</sup>.

## 2.4 References

1. Oyer, J. A. *et al.* Point mutation E1099K in MMSET/NSD2 enhances its methyltransferase activity and leads to altered global chromatin methylation in lymphoid malignancies. *Leukemia* **28**, 198–201 (2014).
2. Jaffe, J. D. *et al.* Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nat Genet* **45**, 1386–1391 (2013).
3. Källberg, M. *et al.* Template-based protein structure modeling using the RaptorX web server. *Nature Protocols* **7**, 1511–1522 (2012).
4. Li, Y. *et al.* The target of the NSD family of histone lysine methyltransferases depends on the nature of the substrate. *J Biol Chem* **284**, 34283–34295 (2009).
5. Qiao, Q. *et al.* The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. *J Biol Chem* **286**, 8361–8368 (2011).
6. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
7. Reich, M. *et al.* GenePattern 2.0. *Nat Genet* **38**, 500–501 (2006).
8. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545–15550 (2005).
9. Zhan, F. *et al.* The molecular classification of multiple myeloma. *Blood* **108**, 2020–2028 (2006).

10. Mirabella, F. *et al.* MMSET is the key molecular target in t(4;14) myeloma. *Blood Cancer Journal* **3**, e114 (2013).
11. Min, D.-J. *et al.* MMSET Stimulates Myeloma Cell Growth Through MicroRNA-Mediated Modulation of c-MYC. *Blood* **118**, 469–469 (2011).
12. Mercer, J., MacArthur, D., Mesirov, J. P., Regev, A. & Lage, K. GeNets: Broad Institute Web Platform for Genome Networks.
13. Lage, K. *et al.* A human phenome-interactome network of protein complexes implicated in genetic disorders. *Nat Biotechnol* **25**, 309–316 (2007).
14. Ruepp, A. *et al.* CORUM: the comprehensive resource of mammalian protein complexes--2009. *Nucl. Acids Res.* **38**, D497–501 (2010).
15. Marango, J. *et al.* The MMSET protein is a histone methyltransferase with characteristics of a transcriptional corepressor. *Blood* **111**, 3145–3154 (2008).
16. Kim, J.-Y. *et al.* Multiple-myeloma-related WHSC1/MMSET isoform RE-IIBP is a histone methyltransferase with transcriptional repression activity. *Mol Cell Biol* **28**, 2023–2034 (2008).
17. Lundberg, S. M. *et al.* ChromNet: Learning the human chromatin network from all ENCODE ChIP-seq data. *Genome Biology* **17**, 1 (2016).
18. Consortium, T. E. P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
19. Ram, O. *et al.* Combinatorial Patterning of Chromatin Regulators Uncovered by Genome-wide Location Analysis in Human Cells. *Cell* **147**, 1628–1639 (2011).
20. Dovey, O. M. *et al.* Histone deacetylase 1 and 2 are essential for normal T-cell development and genomic stability in mice. *Blood* **121**, 1335–1344 (2013).
21. Nguyen, U. T. T. *et al.* Accelerated chromatin biochemistry using DNA-barcoded nucleosome libraries. *Nature Methods* **11**, 834–840 (2014).

22. Kuo, A. J. *et al.* NSD2 Links Dimethylation of Histone H3 at Lysine 36 to Oncogenic Programming. *Molecular Cell* **44**, 609–620 (2011).
23. Martinez-Garcia, E. *et al.* The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. *Blood* **117**, 211–220 (2011).
24. Ezponda, T. *et al.* The histone methyltransferase MMSET/WHSC1 activates TWIST1 to promote an epithelial-mesenchymal transition and invasive properties of prostate cancer. *Oncogene* **32**, 2882–2890 (2013).
25. Gould, J. GENE-E. *GENE-E* Available at: (Accessed: 16 July 2015)
26. Gould, J., Getz, G., Monti, S., Reich, M. & Mesirov, J. P. Comparative gene marker selection suite. *Bioinformatics* **22**, 1924–1925 (2006).

# Chapter III

## Small-molecule profiling to identify dependencies of NSD2-driven CCLs

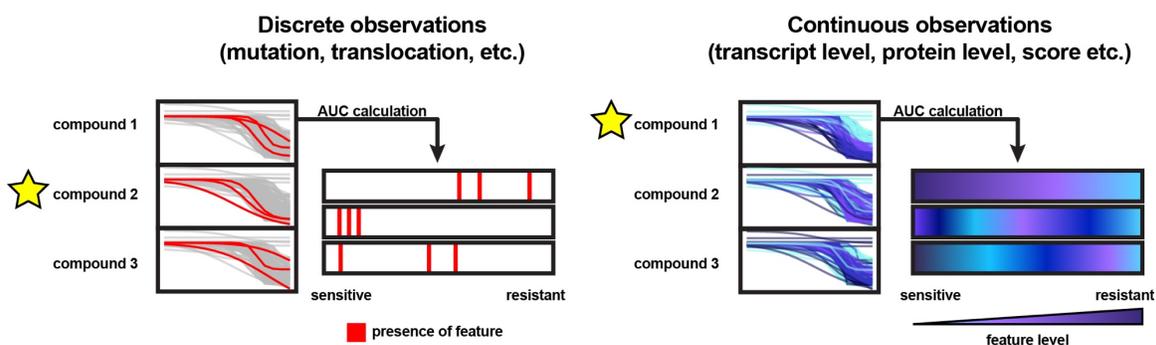
### Collaborators contributions:

- **Dr. Drew Adams** aided in setting up the study.
- **Dr. Jaime Cheah** for her contributions to the CTD<sup>2</sup> cell line small molecule profiling.
- **Dr. Paul Clemons** for the statistical analyses done on the CTD<sup>2</sup> dataset
- **Dr. Joshiawa Paulk** for his analysis on the CTD<sup>2</sup> dataset and E1099K-HDAC1/2 dependency.
- **Dr. Florence Wagner and Dr. Edward Holson** for providing the HDAC-selective inhibitors used in this study.

### 3.1 Introduction

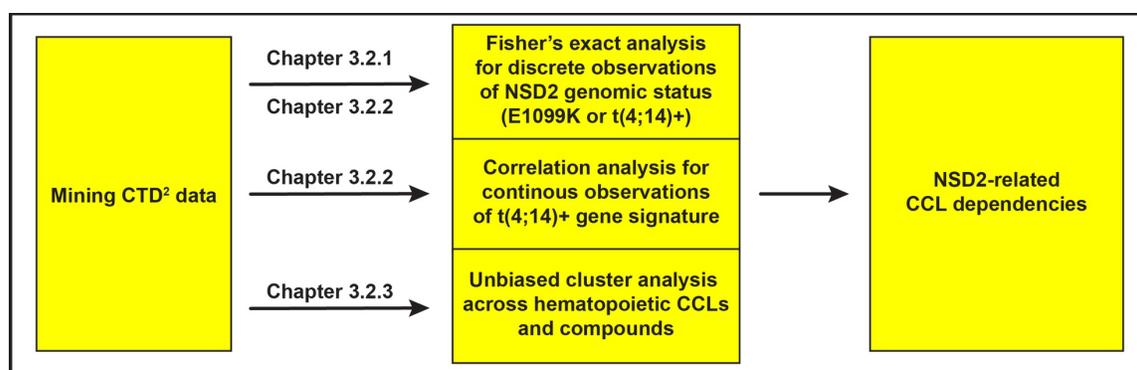
In this chapter, I describe the efforts of large-scale small-molecule profiling with hematopoietic CCLs to infer targets in NSD2-related CCLs. NSD2-related CCLs, E1099K-ALLs and t(4;14)+ MMs, have been observed to be unviable upon knockout of NSD2 or its activity<sup>1-3</sup>. Currently, there are no known inhibitors of NSD2, therefore present standard of care include non-targeted treatment such as general chemotherapy treatment<sup>4,5</sup> for MM and radiation and stem-cell transplant in ALL<sup>6</sup>. With current treatment options, t(4;14)+ MMs, remain to have low prognostic outlook<sup>7,8</sup>.

Dr Cheah and her team, using a large set of well-characterized CCLs from CCLE<sup>9</sup> and a panel of small molecules with known and unknown mechanism of action (MoA), performed small-molecule profiling with CCLs to generate a dataset of cell viability metrics in response to these compounds<sup>10</sup>. Using CCLE annotations, we were able to identify CCLs with E1099K mutations and t(4;14)+ translocation as well as their level of NSD2 transcripts, enabling us to perform statistical analyses on the dataset. Our analysis can be categorized into 2 types, analysis using discrete observations and analysis using continuous observations (Figure 3.1a).



**Figure 3.1a Left:** Analysis using discrete observations, compound 2 appears to target CCLs with our feature of interest. **Right:** Analysis using continuous observations, compound 1 appears to target CCLs with high levels of our feature of interest.

In discrete observations, CCLs are annotated with a binary feature (WT or mutant, WT or translocated). Our aim is to identify compounds that selectively target CCLs with a specific binary label. In continuous observations, CCLs are annotated with a range of a feature (transcript level, protein level, enrichment score, etc.). Our aim is to identify compounds that correlate viability with the feature <sup>11</sup>. Lastly, clustering of CCLs based on their viability metric was performed to analyze similarity between results from discrete analysis and continuous analysis. Brief scheme of this chapter is shown below (Figure 3.1b).



**Figure 3.1b** Scheme of experiments described in this chapter.

## 3.2 Results and Discussion

### 3.2.1 HDAC1/2 selective inhibitors target E1099K- ALLs

As part of ongoing efforts to identify novel predictors of cell-line sensitivity to small-molecule probes and drugs <sup>10</sup>, Dr Cheah and her team profiled 545 small molecules with 157 hematopoietic cancer cell lines, including 20 ALLs, in hopes of identifying oncogene-induced dependencies of E1099K-ALLs that can be targeted by small molecules. Each cell line was incubated with compound for 72 hours and ATP levels were measured as a surrogate of cell growth and viability. Dose-response curves (DRPs) were generated with ATP levels of compound-treated cells normalized to ATP levels of DMSO-only treated cells. Area under concentration-repose curves (AUC) was calculated as a metric of sensitivity. The results are available online in the database –

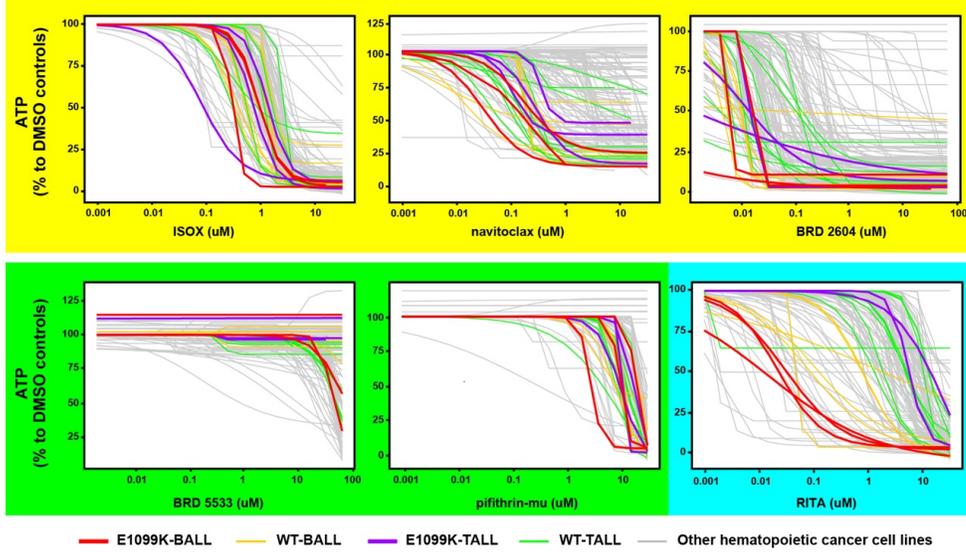
Cancer Therapeutics Response Portal v2 (CTRP2).

To determine enrichment for E1099K-ALLs among those most sensitive to each compound, Dr Clemons identified the AUC threshold corresponding to the best enrichment of mutant lines using Fisher's exact test implemented in MATLAB. Each threshold that changed the number of E1099K called sensitive (or non-responsive) was tested, we retained the most significant odds-ratio value for each compound and their corresponding p-values.

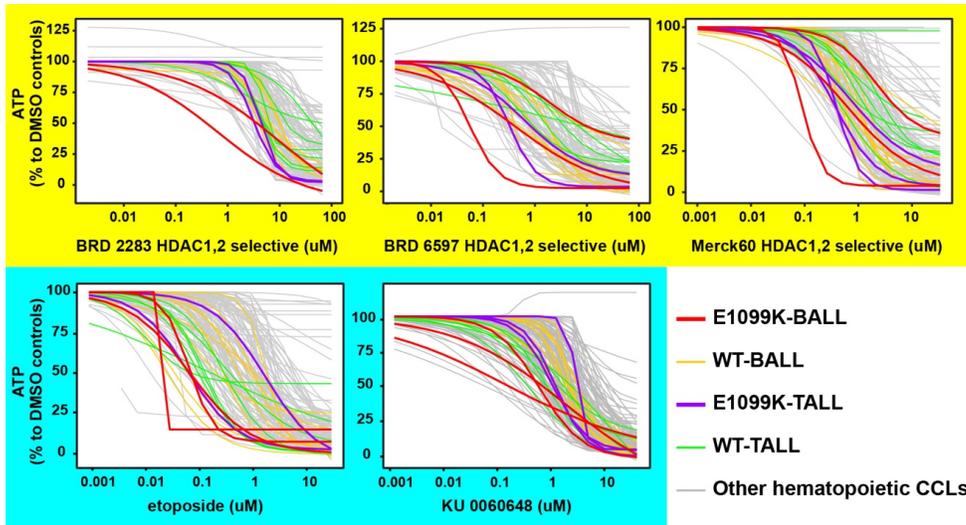
We identified 18 compounds that significantly target E1099K-ALLs (Figure 3.2.1a). DRPs of these compounds were reassessed to validate the results of the enrichment test. Compounds that target ALLs or have incomplete DRPs could appear to be enriching for targeting of E1099K-ALLs giving rise to false positives (Figure 3.2.1b).

compound	target	p-value
BRD-K24690302	HDAC1	1.23E-10
BRD-K66532283	HDAC1;HDAC2	5.94E-08
LY-2183240	FAAH	2.01E-07
navitoclax	BCL2;BCL2L1;BCL2L2	4.01E-07
BRD-K61166597	HDAC1;HDAC2	6.18E-07
ISOX	HDAC6	7.27E-07
PX-12	TXN	8.02E-07
etoposide	TOP2A	1.12E-06
BRD-K23984367	BRAF;RAF1;KDR;FLT3;FLT4;PDGFRB;KIT	7.21E-06
BRD-K55116708		1.10E-05
BRD-K45935533	RARA;RARB;RARG	1.91E-05
BRD-K85133207	HDAC1	1.96E-05
BRD-K34022604	POLA1;POLA2;POLE;RRM1	2.08E-05
BRD-K34222889		4.48E-05
KU 0060648	PRKDC	5.98E-05
RITA	TP53;MDM2	7.45E-05
pifithrin-mu	TP53;HSPA1A;HSPA1B;HSPA1KL	1.26E-04

**Figure 3.2.1a** List of compounds identified in the statistical test that significantly enriched for sensitivity of E1099K-ALLs. HDAC1/2 inhibitors are highlighted in red and other notable compounds are highlighted in yellow.



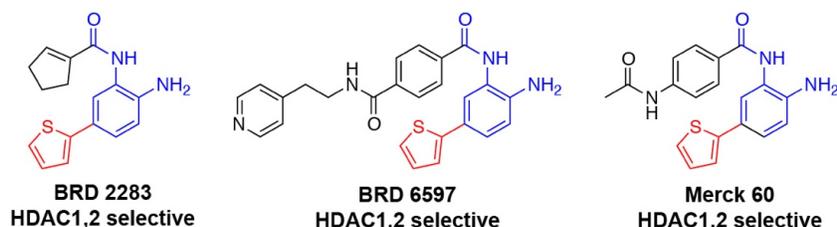
**Figure 3.2.1b** DRPs of false positives that were identified in the statistical test. DRPs in boxes shaded in yellow show E1099K-ALLs to have similar DRP to other ALLs. DRPs in boxes shaded in green are incomplete. DRPs in boxes shaded in blue show distinct separation between BALLs and TALLs leading to false enrichment of E1099K-ALLs.



**Figure 3.2.1c** DRPs in boxes shaded in yellow show HDAC1/2 inhibitors selectively target E1099K-ALLs, DRPs in boxes shaded in blue show etoposide and KU 0060648 also seemed to have an effect in targeting E1099K-ALLs.

We observed multiple HDAC1/2 inhibitors showed maximal potency for E1099K-ALLs relative to the rest of the hematopoietic CCLs tested. 5 E1099K-ALLs tested were among the most sensitive 21 CCLs treated with BRD 2283 (140 CCLs total), with a significant enrichment of  $p = 5.94E-08$ . Other notable compounds that were enriched include topoisomerase (TOP2A) inhibitor, etoposide ( $p = 1.12E-06$ , 4 E1099K-ALLs in 14 sensitive of 147 total CCLs), and PRKDC inhibitor KU 0060648 ( $p = 5.98E-05$ , 6 E1099K-ALLs in 33 sensitive of 142 total CCLs) (Figure 3.2.1c).

In Chapter 2, E1099K-ALLs had an enrichment of HDAC2 complexes, we explored the E1099K-ALLs' HDAC1/2 dependency with selective HDAC1/2 inhibitors that were developed by our colleagues at the Broad. BRD 2283 emerged from ongoing efforts to develop isoform-selective HDAC inhibitors. Wagner *et al.* previously reported that potent and selective inhibition of HDAC 6 does not require a surface binding motif often characteristic of known HDAC inhibitors<sup>12</sup>. A small linker motif, a cyclopentenyl, was coupled with a hydrozamic acid zinc binding group (ZBG) to achieve potent and selection inhibition of HDAC6 isoform<sup>13,14</sup>. Here, Dr Wagner and colleagues applied the same strategy by pairing the same small cyclopentenyl linker moiety with the isoform biasing *ortho*-aminoanilide ZBG<sup>15,16</sup> to achieve potent and selection inhibition of HDAC1 and HDAC2 by BRD 2283 (IC<sub>50</sub> = 0.003  $\mu$ M and 0.054  $\mu$ M respectively) in biochemical assays (Figure 3.2.1c and 3.2.1d)<sup>17</sup>.



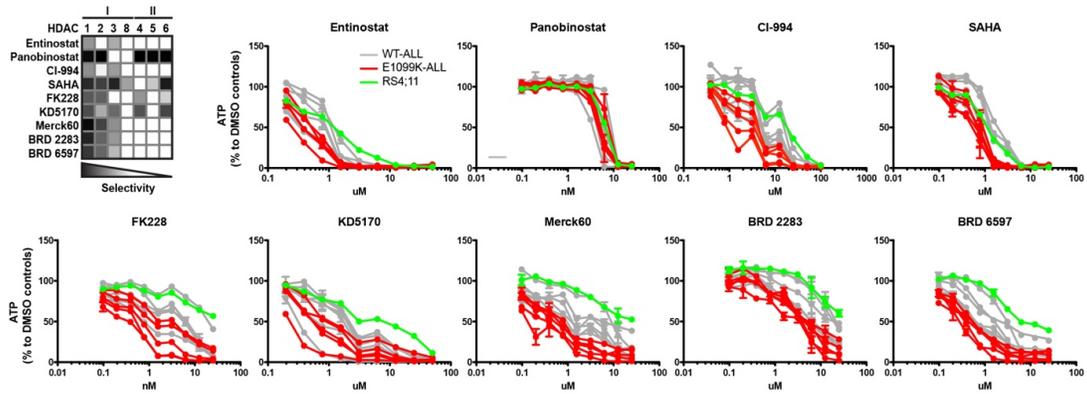
**Figure 3.2.1c** Structures of HDAC1/2 selective compounds, Merck 60, BRD 2283 and BRD 6597. Zinc-binding motif (ZBG) of structures are highlighted in blue. Cavity motif of structures is highlighted in red. Cavity motif occupies an internal cavity that is next to the catalytic zinc in the proteins, HDAC1 and HDAC2.

Name	IC50 ( $\mu$ M)								
	HDAC1 3h Preinc	HDAC2 3h Preinc	HDAC3 3h Preinc	HDAC4 0h Preinc	HDAC5 0h Preinc	HDAC6 0h Preinc	HDAC7 0h Preinc	HDAC8 0h Preinc	HDAC9 0h Preinc
SAHA	0.004 $\pm$ 0.0001	0.011 $\pm$ 0.0005	0.003 $\pm$ 0.0005	>33.33	8.75 $\pm$ 1.80	0.002 $\pm$ 0.0002	>33.33	1.02 $\pm$ 0.187	>33.33
Merck 60	0.001 $\pm$ 0.001	0.013 $\pm$ 0.009	0.398 $\pm$ 0.105	>33.33	>33.33	>33.33	>33.33	>33.33	>33.33
BRD6597	0.008 $\pm$	0.064 $\pm$	4.28 $\pm$	>33.33	>33.33	>33.33	>33.33	>33.33	>33.33
BRD2283	0.003 $\pm$ 0.002	0.054 $\pm$ 0.016	0.604 $\pm$ 0.039	>33.33	>33.33	>33.33	>33.33	>33.33	>33.33

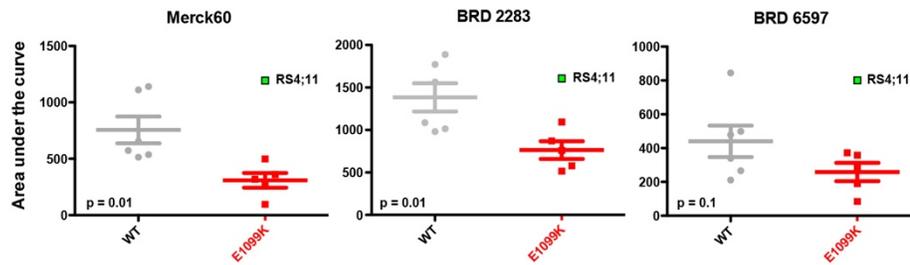
**Figure 3.2.1d** IC<sub>50</sub> values for multiple histone deacetylase inhibitors across 9 HDAC isoforms. Data are expressed as mean  $\pm$  standard deviation.

We further characterized a collection of 9 HDAC inhibitors of varying selectivity in a focused panel of 6 E1099K-ALLs and 6 WT-ALLs. Similarly to BRD 2283, BRD 6597 and Merck60<sup>18</sup>, both highly potent and selective HDAC1/2 inhibitors, showed selectivity for E1099K-ALLs (Figure 3.2.1e and 3.2.1f). A selective HDAC1/2/3 inhibitor, CI-994<sup>16</sup>, also appeared to selectively kill E1099K-ALLs over WT-ALLs (Figure 3.2.1e). Other non-selective HDAC inhibitors, such as SAHA<sup>19</sup> and Panobinostat<sup>20</sup> (broad class I and class II HDAC inhibitors), did not show a striking distinction in targeting E1099K-ALLs (Figure 3.2.1e). One E1099K-ALL, RS4;11, appeared to be resistant to HDAC1/2 inhibitors. By querying CCLE, we observed that RS4;11 has a frameshift deletion in HDAC2 (p.T553fs). HDAC2 frameshift mutations have previously been studied to cause resistance towards HDAC inhibition<sup>21</sup>, we hypothesize that the frameshift supersedes the effect of E1099K have on HDAC1/2 sensitivity in RS4;11. We have also observed E1099K-ALLs have increased levels of HDAC2 complexes (chapter 2), which might form the basis of the E1099K-specific dependency within ALLs.

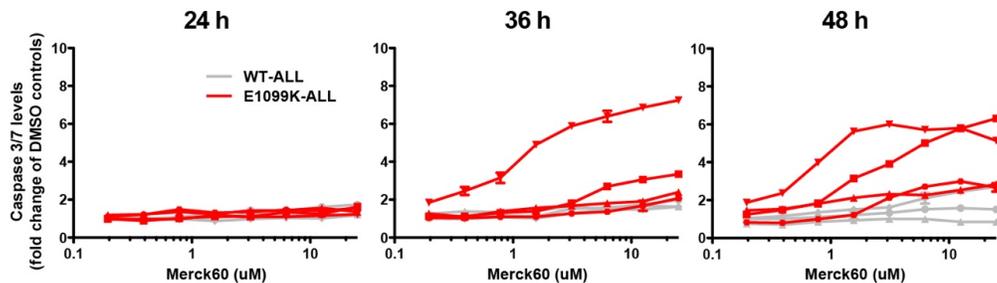
To confirm cell death in these assays, we measured caspase3/7 activity in these ALLs upon treatment with Merck60. Increasing levels of caspase3/7 activity observed in E1099K-ALLs over 48 hours confirmed the onset of apoptosis in these cell lines upon HDAC1/2 inhibition (Figure 3.2.1g).



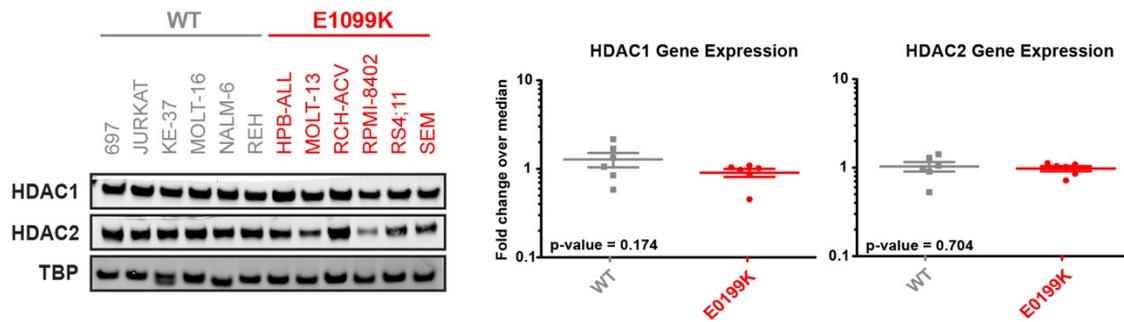
**Figure 3.2.1e** DRPs of selected ALL cell lines when treated with a panel of HDAC inhibitors. 9 HDAC inhibitors were profiled, their known selectivity towards each member of HDAC family are represented as a heat map with a color gradient from black (high inhibition) to white (no inhibition). RS4;11 is a HDAC1/2-resistant E1099K-ALL represented as a green line. Error bars represent standard error of mean over two replicates.



**Figure 3.2.1f** Scatter plot representation of Figure 3.2.1e. AUCs of HDAC1/2 selective compounds in Figure 3.2.1e were calculated and plotted as a scatter plot, bars represent mean and standard error of mean. p-values were calculated by performing student's t-test.



**Figure 3.2.1g** Caspase3/7 level of selected ALLs when treated with Merck60. 3 WT-ALLs (JURKAT, NALM-6, REH) are represented as gray lines and 4 E1099K-ALLs (RPMI-8402, SEM, HPB-ALL, RCH-ACV) are represented as red lines. E1099K-ALLs were undergoing more apoptosis at 48 h compared to WT-ALLs. Error bars represent standard error of mean over two replicates.

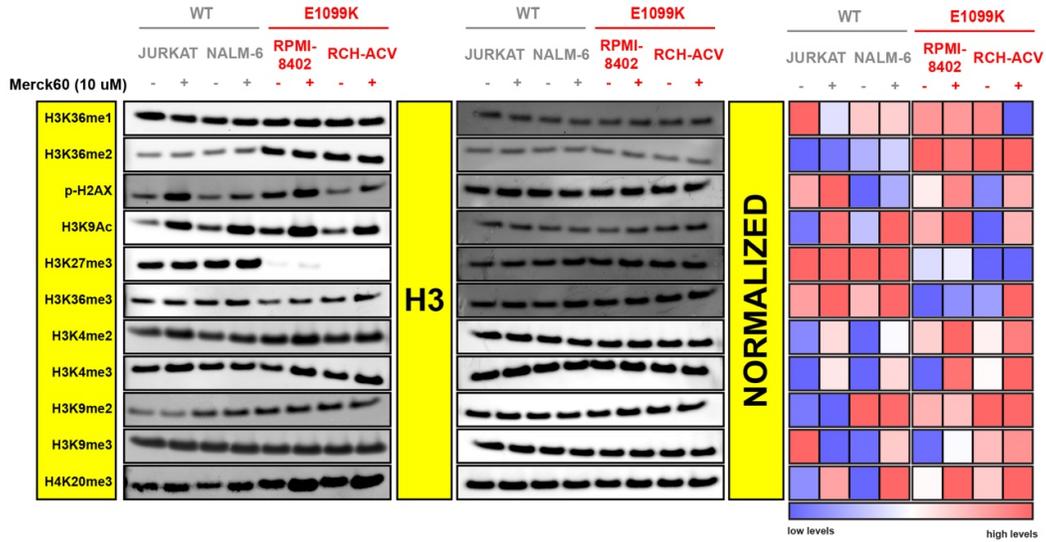


**Figure 3.2.1h Left:** Immunoblot of HDAC1 and HDAC2 in E1099K- and WT-ALLs. **Right:** Transcript levels of HDAC1 and HDAC2 in E1099K- and WT-ALLs. Bars represent mean and standard error of mean over three replicates. p-values were calculated by performing student's t-test.

To elucidate the mechanism of sensitivity of E1099K-ALLs to HDAC1/2 inhibition, we assessed levels of HDAC1 and HDAC2 in the same panel of ALLs. The protein and transcript levels of HDAC1 and HDAC2 do not differ between E1099K and WT-ALLs (Figure 3.2.1h), suggests that the mechanism of HDAC1/2 dependency did not depend on levels of HDAC1/2 proteins in ALLs.

To assess the response of the cell lines to HDAC1/2 inhibition, we treated a panel of E1099K-ALLs (RPMI-8402, RCH-ACV) and WT-ALLs (JURKAT, NALM-6) with 10  $\mu$ M of Merck60 for 24 hours. Histone modifications of treated and untreated cells were profiled and as expected, E1099K-ALLs have elevated H3K36me2 levels and reduced H3K27me3 levels, with and without Merck60 treatment (Figure 3.2.1i). Upon treatment, levels of H3K9Ac increased, which has been previously observed with loss of HDAC1/2 activity<sup>22</sup>. H3K4 methylation marks were also elevated upon treatment, which has been previously observed with an increase in H3 acetylation<sup>23</sup>. DNA damage markers, p-H2AX and H4K20me3, were elevated upon treatment, which was expected as it is known that HDAC inhibitors cause DNA damage<sup>24</sup>. Overall, both E1099K- and WT-ALLs did not display differential response when treated with Merck60. This suggests that HDAC1/2 dependency of E1099K-ALLs could be more complicated than that of affecting NSD2's

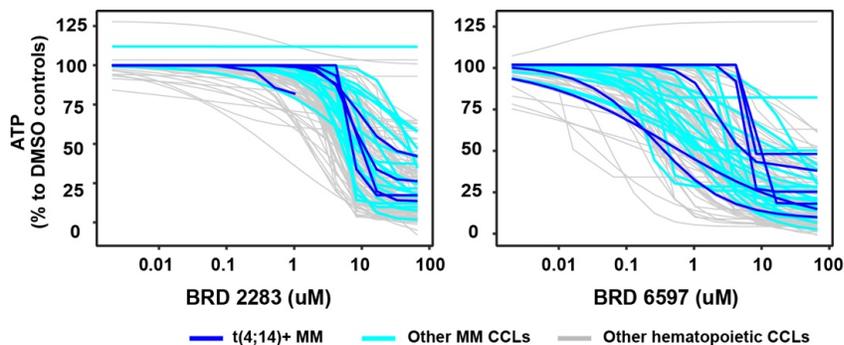
global methylation activity on the epigenetic landscape.



**Figure 3.2.1i Left:** Immunoblot of histone modifications in 2 WT- and E1099K-ALLs upon 24h of treatment with 10 uM of Merck60. **Center:** Immunoblot of loading control H3. **Right:** Heatmap of quantitated blot normalized across ALLs tested.

### 3.2.2 IGFR inhibitors and FGFR3 inhibitors target t(4;14)+ MMs

In Chapter 2, we observed E1099K-ALLs display similar phenotypic characteristic as NSD2-overexpressing t(4;14)+ MMs, which suggests that HDAC1/2 dependency could be a common characteristic among NSD2-dependent malignancies. Analysis of DRPs generated by t(4;14)+ MMs treatment with BRD 2283 and BRD 6597 do not show the same sensitivity compared to E1099K-ALLs (Figure 3.2.2a).



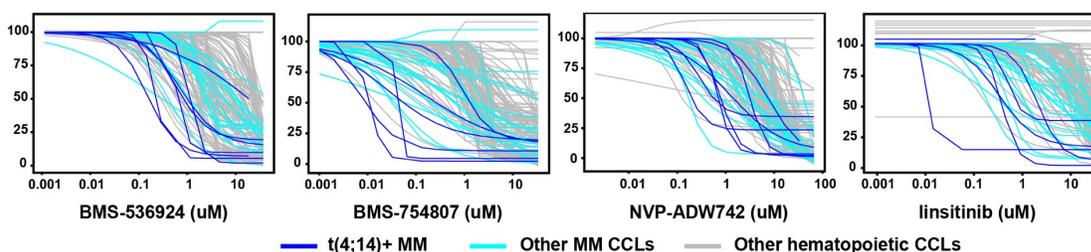
**Figure 3.2.2a** DRPs of hematopoietic CCLs with treatment of HDAC1/2 inhibitors did not show selective killing of t(4;14)+ MMs.

Dr Clemons performed Fisher's exact test as described in Chapter 3.2.1 to identify compounds that selectively target t(4;14)+ MMs. The test revealed that IGF1R inhibitors ranked high in selective killing of t(4;14)+ MMs compared to other hematopoietic CCLs. All 4 IGF1R inhibitors tested, BMS-536924, linsitinib, NVP-ADW742 and BMS-754807, significantly target t(4;14)+ MMs (Figure 3.2.2b).

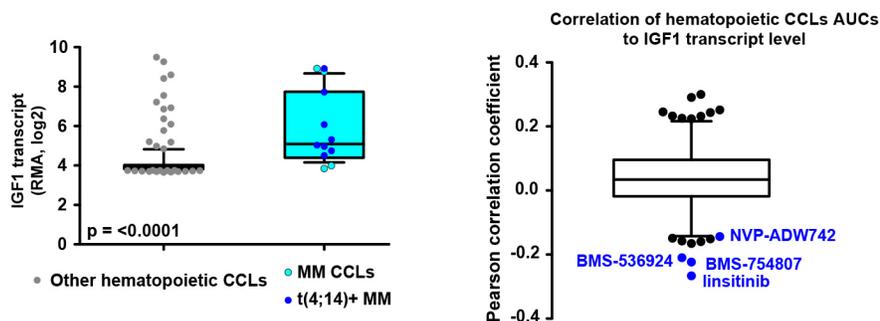
compound	target	p-value
<b>BMS-536924</b>	<b>IGF1R;INSR</b>	<b>1.54E-06</b>
<b>linsitinib</b>	<b>IGF1R;INSR</b>	<b>4.17E-06</b>
<b>NVP-ADW742</b>	<b>IGF1R;INSR</b>	<b>1.77E-05</b>
curcumin	PPID	4.96E-05
BRD 9647		4.99E-05
<b>BMS-754807</b>	<b>IGF1R;INSR</b>	<b>5.33E-05</b>
GDC-0941	PIK3CA;PIK3CB;PIK3CD;PIK3CG	1.02E-04
PI-103	MTOR;PIK3CA;PIK3CB;PIK3CD;PIK3CG;PRKDC	1.02E-04

**Figure 3.2.2b** List of compounds identified in the statistical test that significantly enriched for sensitivity of t(4;14)+ MMs. IGF1R inhibitors are highlighted in red.

To confirm t(4;14)+ sensitivity towards IGF1R inhibitors, we reassessed DRPs of CCLs upon IGF1R treatment. We observed that IGF1R inhibitors target MM, with t(4;14)+ MMs being more sensitive. This suggests that IGF1R inhibition could also be a lineage-driven dependency in addition to being a NSD2 dependency (Figure 3.2.2c). Publicly available IGF1 expression data of hematopoietic CCLs in CCLE revealed MMs to be highly enriched in IGF1 expression compared to all other hematopoietic cell lines. The expression data was also used in correlation analysis of profiling data to reveal IGF1 transcript level correlates with sensitivity towards IGF1R inhibitors (Figure 3.2.2d). It is known that MMs have aberrant expression of IGF1<sup>25</sup> and others have noted the correlation between high IGF1 levels and sensitivity towards IGF1R inhibitors<sup>9,26</sup>. Our results corroborate on that observation.



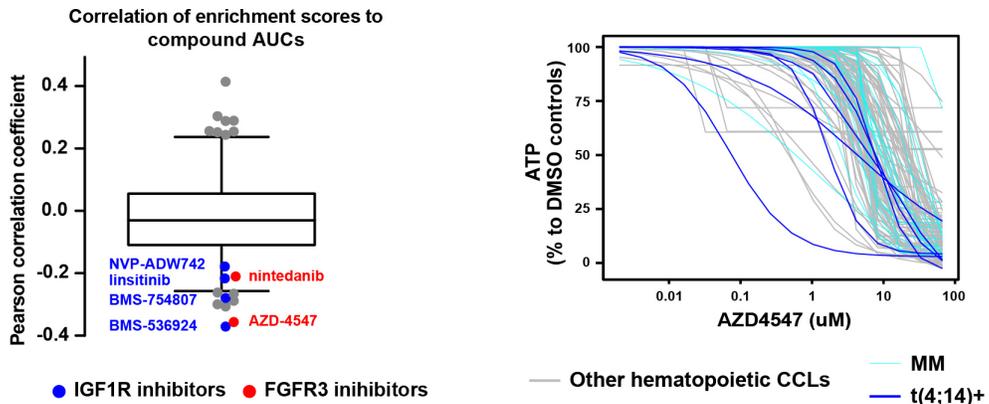
**Figure 3.2.2c** DRPs of hematopoietic CCLs with treatment of IGF1R inhibitors show selective killing of MM CCLs especially towards t(4;14)+ MMs.



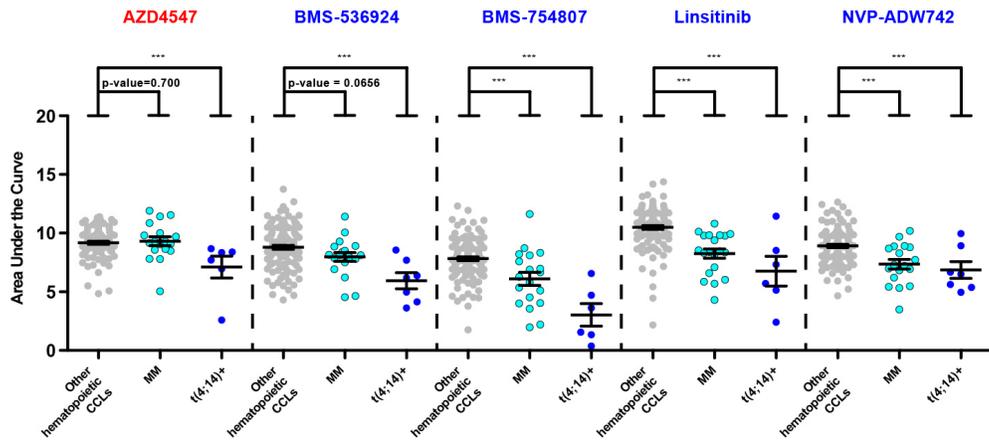
**Figure 3.2.2d Left:** Transcript levels of IGF1 in MMs are higher than in other hematopoietic CCLs. **Right:** Plot of Pearson correlation coefficient of AUC to IGF1 transcript levels. Sensitivity to IGF1R inhibitor treatment is ranked most negatively correlated to IGF1 transcript.

Another approach was used to identify compounds that target t(4;14)+ MMs. We employed single-sample GSEA (ssGSEA) on gene expression data of hematopoietic CCLs to assign enrichment scores for geneset ‘ZHAN\_MULTIPLE\_MYELOMA\_MS\_UP’ to each CCL (Supplemental Table 3.1). As expected, t(4;14)+ MMs had higher enrichment scores compared to other MMs and hematopoietic CCLs. Correlation analysis was performed on the enrichment scores and compound AUCs (Supplemental Table 3.2). BMS-536924, a IGF1R inhibitor, and AZD4547, a FGFR3 inhibitor, were most selective in killing CCLs with high t(4;14)+ enrichment score ( $r = -0.37$  and  $-0.36$  respectively). Other IGF1R inhibitors tested (BMS-754807, linsitinib and NVP-ADW742) and FGFR3 inhibitor tested (nintedanib) were also selective in killing CCLs with high t(4;14)+ enrichment score (Figure 3.2.2e). Compared to IGF1R inhibitors, AZD4547 only significantly

targets t(4;14)+ MMs and not the other MMs. This is to be expected as t(4;14)+ translocation upregulated the levels of FGFR3 in these cells as well as the levels of NSD2, suggesting FGFR3 to be another target in killing t(4;14)+ MMs<sup>27,28</sup>. IGF1R inhibitors significantly target MMs, with t(4;14)+ MMs being more sensitive to BMS-546924 and BMS-754807 (Figure 3.2.2f).



**Figure 3.2.2e Left:** Plot of Pearson correlation coefficient of AUC to enrichment scores of CCLs. Effect of FGFR3 and IGF1R inhibitors are most anti-correlated to enrichment scores. **Right:** DRPs of hematopoietic CCLs with treatment of IGF1R inhibitors show selective killing of t(4;14)+ MMs.



**Figure 3.2.2f** Dot plot of AUC of DRPs of CCLs with compound treatment. Error bars represent standard error of the mean.

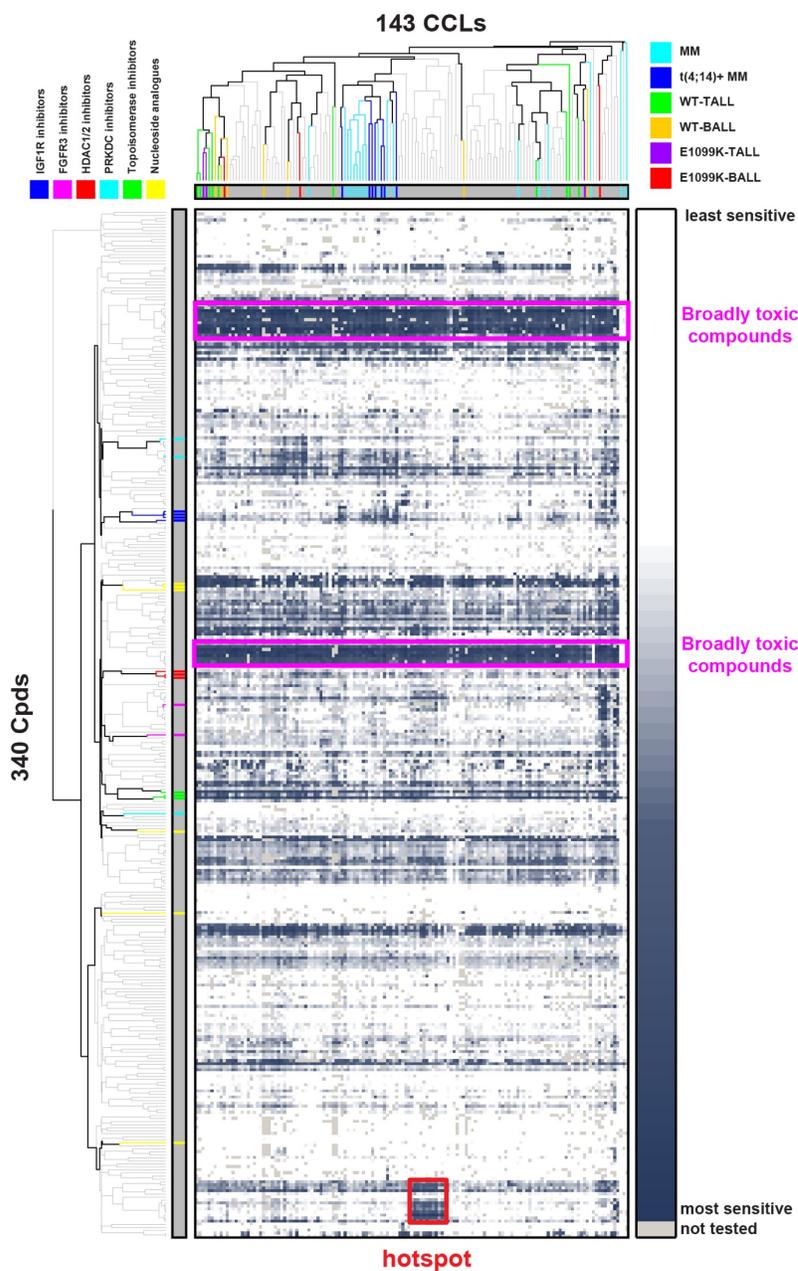
### **3.2.3 Dependencies of NSD2-driven CCLs are more lineage dependent**

We applied a previously published method, annotated cluster multidimensional enrichment (ACME) analysis<sup>26</sup> that could integrate sensitivity information from several compounds and CCLs. The unbiased, hierarchical clustering of the AUC data was performed to observe clustering of NSD2 affected CCLs (E1099K-ALLs and t(4;14)+ MMs) based on their similarity in responses across compound treatments. The resulting matrix revealed several horizontal stripes that correspond to compounds that were broadly toxic in general (Figure 3.2.3a). We annotated CCLs with their lineage, t(4;14)+ or E1099K status, and compounds with their targets or function, for annotation enrichment analysis, to determine if the clusters of compounds or CCLs were enriched for an annotation.

The top annotation that was found enriched within the same cluster is chronic myeloid leukemia (CML) (p-value = 1.05E-17) (Figure 3.2.3b). By observing the matrix of AUC clustering, there is a sensitivity ‘hotspot’ where CMLs are targeted by a small group of compounds (Figure 3.2.3a). AUC analyses of these compounds targeting CMLs confirmed that their selectivity towards CMLs is significant (Figure 3.2.3c). In addition, a number of these compounds (bosutinib, nilotinib, asatinib and imatinib) are approved drugs in the clinic for treatment of CML<sup>29</sup>. These observations for CML validated the use of unbiased clustering to identify cluster enrichment of annotations in compounds and CCLs.

NSD2-related annotations that are significantly enriched include: plasma cell myeloma (MM) (p-value = 2.56E-15) and acute T cell lymphoblastic leukemia (T-ALL) (p-value = 1.61E-07), t(4;14)+ MM (p-value = 3.71E-07) (Figure 3.2.3b). E1099K-ALLs and t(4;14)+ MMs were observed to not cluster together. t(4;14)+ MM is enriched in the same cluster as plasma cell myeloma (Figure 3.2.3d). Both sub-lineages of ALLs, T-ALL and B-ALL, are significantly enriched in different clusters. As annotations of E1099K-ALLs include both E1099K T-ALLs and B-ALLs, the

annotation of E1099K-ALLs was not significantly enriched in our analysis. However, they do cluster within their own respectively sub-lineages.



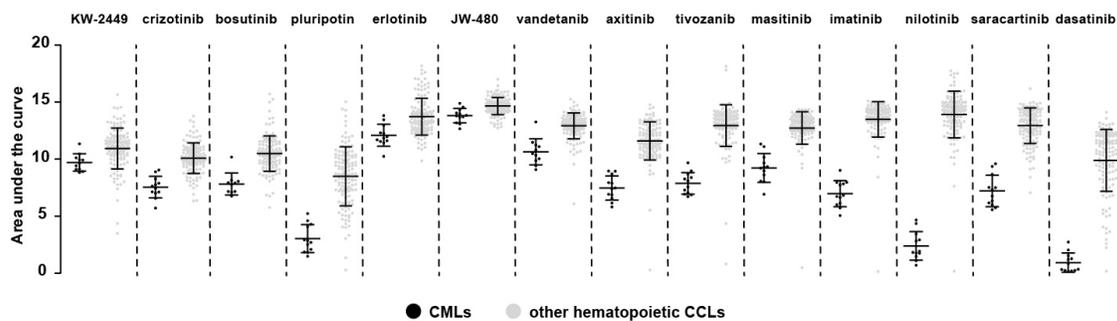
**Figure 3.2.3a** Matrix of AUC values after compound and CCL clustering. Dendrogram of CCLs are color-labeled to represent NSD2-related annotations. Dendrogram of compounds are color-labeled to represent compounds of interest mentioned in this chapter. Stripes of low AUC is labeled as broadly toxic compounds. CML hotspot is also marked out.

rank	Annotations	p-value	Cluster	CCLs in annotation		CCLs in cluster	
				Total	% in cluster	Total	% in annotation
1	chronic myeloid leukaemia	1.05E-17	108	12	100	12	100
2	plasma cell myeloma	2.56E-15	118	25	72	20	90
3	blast phase	1.00E-12	108	11	91	12	83
4	acute lymphoblastic T cell leukemia	1.61E-07	52	14	43	6	100
5	hematopoietic neoplasm	1.62E-07	108	43	28	12	100
6	Burkitt lymphoma	2.12E-07	107	9	67	8	75
7	t(4;14)+ MM	3.71E-07	118	7	100	20	35
8	acute lymphoblastic B cell leukemia	6.07E-05	81	10	40	5	80
9	acute myeloid leukaemia	6.89E-05	74	31	19	6	100
10	chronic lymphocytic leukaemia	2.95E-04	34	3	67	2	100
17	E1099K-ALL	1.48E-03	26	6	33	2	100

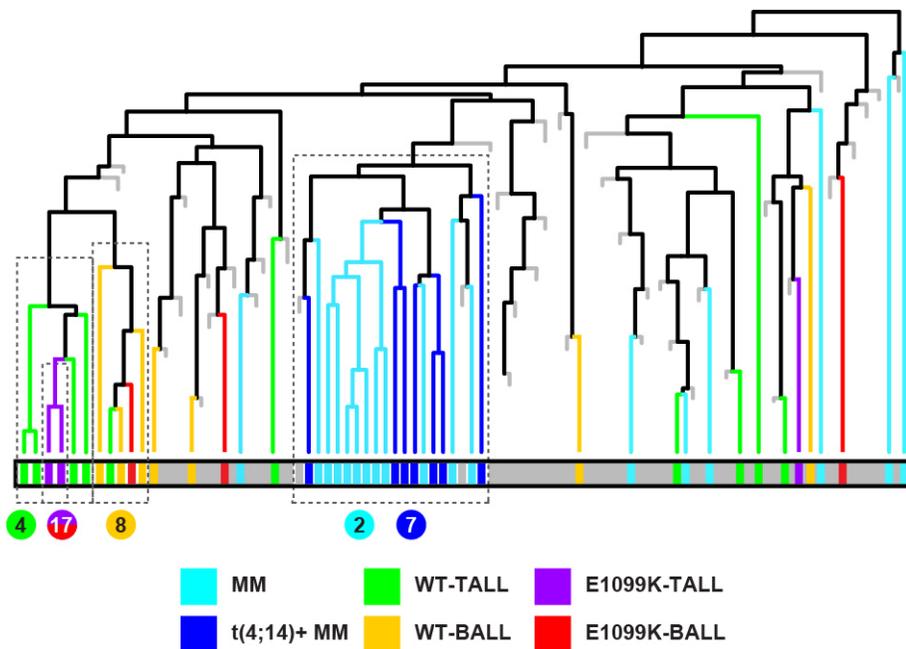
  

rank	Annotations	p-value	Cluster	Cpds in annotation		Cpds in cluster	
				Total	% in cluster	Total	% in annotation
1	KDR	3.22E-15	200	15	73	14	79
2	FLT3	5.56E-14	200	13	77	14	71
3	PIK3CD	2.81E-09	215	8	75	10	60
4	IGF1R	2.74E-08	209	4	100	6	67
5	MTOR	2.74E-08	81	6	67	4	100
18	HDAC1	1.61E-05	321	12	83	81	12
23	TOP2A	1.74E-05	31	2	100	2	100
34	HDAC2	1.85E-05	94	10	30	3	100
44	DNA	1.26E-04	198	9	33	5	60
60	PRKDC	1.84E-03	132	3	67	9	22
64	FGFR3	3.63E-03	271	2	100	21	10

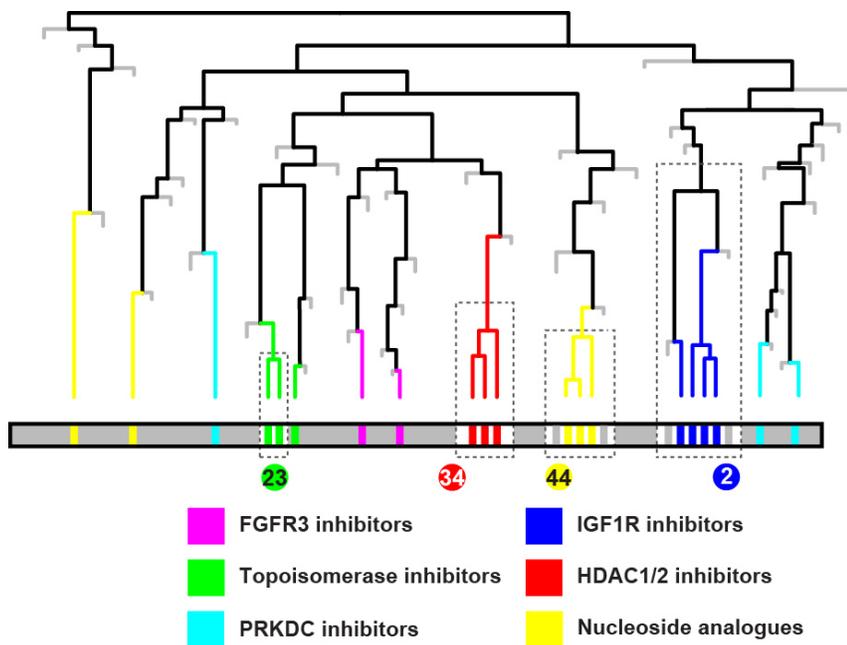
**Figure 3.2.3b** List of annotations (**Top:** CCL annotations **Bottom:** compound target annotation) that were significantly enriched within clusters and annotations that are of interest to the study.



**Figure 3.2.3c** Dot plot of AUC of DRPs of CMLs with compound treatment. Error bars represent standard error of the mean.



**Figure 3.2.3d** Simplified dendrogram of clustered CCLs. NSD2-related annotations are color-labeled. Numbers in circles color-labeled by annotation represents rank of enrichment.



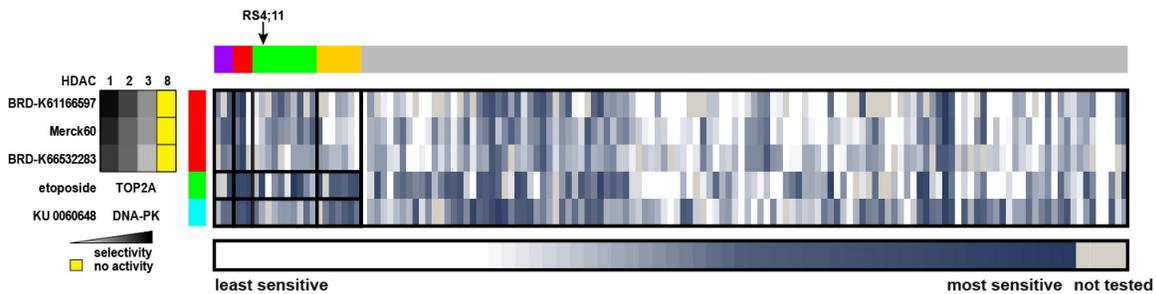
**Figure 3.2.3e** Simplified dendrogram of clustered compounds. Compounds of interest annotations are color-labeled. Number in circles color-labeled by annotation represents rank of enrichment.

Compound annotation enrichment analysis was more complicated due to the multiple annotations that a compound can be attached to (e.g. a pan-HDAC inhibitor can be annotated with 4 targets: HDAC1, HDAC2, HDAC3, HDAC8). With that in mind, we observed that our compounds of interest, IGF1R inhibitors, HDAC1/2 inhibitors and TOP2A inhibitors cluster within themselves while FGFR3 inhibitors do not (Figure 3.2.3e). This suggests that clustered well compounds (IGF1Ri, HDAC1/2i, TOP2Ai) behave similarly and identified closely in previous analysis done in chapter 3.2.1 and 3.2.2 (Figure 3.2.1a and 3.2.2b). FGFR3 inhibitors do not cluster together, which could explain for AZD4547 not identified closely with the other FGFR3 inhibitor, nintedanib in chapter 3.2.2 (Figure 3.2.2e).

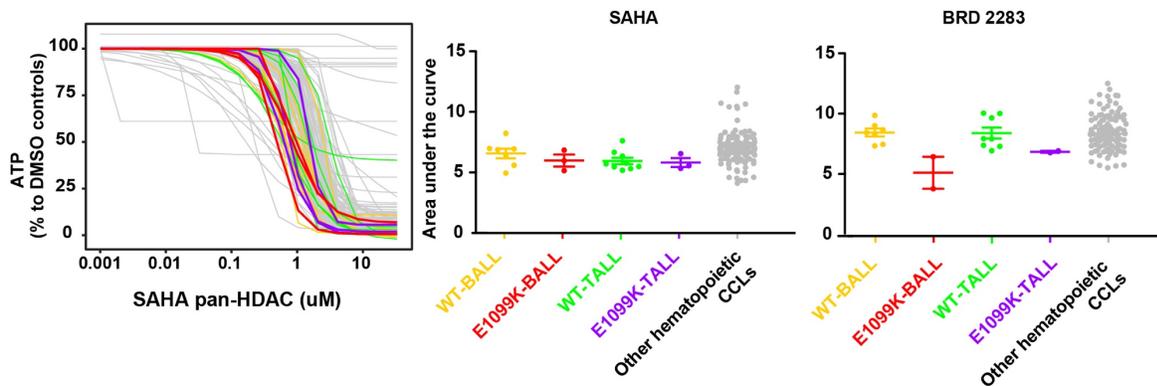
#### **3.2.4 Concluding remarks**

We applied cell-line profiling to identify NSD2-related CCLs' dependencies and observed that E1099K-ALLs showed greater sensitivity to multiple HDAC1/2-selective inhibitors, etoposide and KU0060648 (Figure 3.2.4a). There have been several studies published suggesting HDAC inhibitors as a therapeutic strategy for targeting ALLs or subsets of ALLs<sup>30-33</sup>. We do observe pan-HDAC inhibitor, SAHA, to kill ALLs in our cell-line profiling experiment; however, E1099K mutation further sensitizes the cells to HDAC1/2 inhibition, implying that E1099K induces a HDAC1/2 dependent cell state distinct from other ALLs (Figure 3.2.4b). I have mentioned in chapter 2 that E1099K-ALLs have an increased expression of HDAC2 complexes, supporting the idea of HDAC1/2 activity being an induced dependency of cells having E1099K mutation<sup>34,35</sup>. These results suggest an alternative way of targeting E1099K-ALLs using existing HDAC1/2 inhibitors in lieu of a NSD2 inhibitor. E1099K-ALLs sensitivity to HDAC1/2 inhibitors do not inhibit NSD2 catalytic activity of K36me2 in the cell, suggesting that the nature of dependency could be another mechanism that is unrelated to NSD2's role in transcription activation.

E1099K-ALLs sensitivity towards the topoisomerase-II inhibitor (etoposide) and a dual DNA-PK/P13K (PRKDC) inhibitor (KU0060648) suggests compromised genomic stability in these cell lines. The catalytic activity of WT-NSD2 has been observed to have a role promoting response to DNA-strand breaks<sup>36,37</sup>. We have also observed that both E1099K- and WT-ALLs exhibit DNA damage markers, p-H2AX and H3K20me3, when treated with Merck60 (HDAC1/2 inhibitor). It is possible that the mechanism that sensitizes E1099K-ALLs to HDAC1/2 inhibitor converges with one that sensitizes E1099K-ALLs to etoposide and KU0060648.

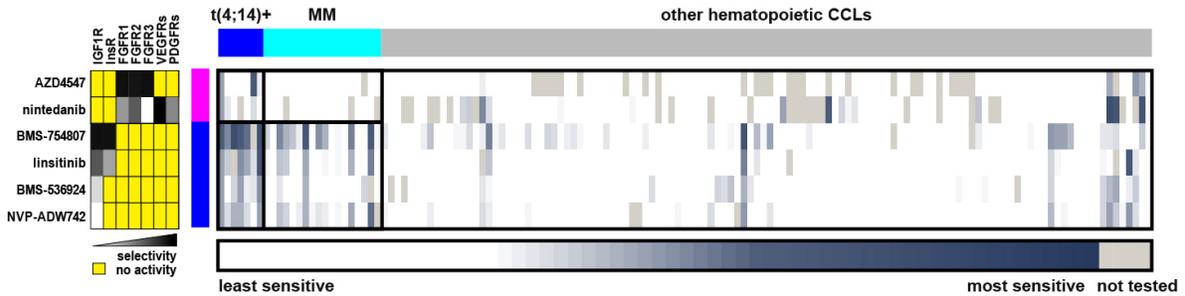


**Figure 3.2.4a** Matrix of AUC values of ALLs treatment with HDAC1/2 inhibitors, etoposide and KU 0060648 compared to other hematopoietic CCLs.



**Figure 3.2.4b** Left: DRPs of hematopoietic CCLs upon treatment of SAHA. Right: Scatter plot representations of AUCs of hematopoietic CCLs upon treatment of SAHA and BRD 2283. Error bars represent standard error of mean.

T(4;14)+ MM and MMs were found to be sensitive to IGF1R inhibitors, BMS-754807, NVP-ADW742, BMS-536924 and linsitinib, and FGFR3 inhibitor, AZD4547 (Figure 3.2.4c). The strategy to target IGF1R has been demonstrated in clinical trials against multiple malignancies including non-small cell lung cancer (NSCLC), and breast cancer<sup>38-40</sup>. Unfortunately, results from these trials have proved to be unfavorable due to the lack of a biomarker to predict efficacy and the side effect of IGF1R inhibition in disrupting endocrine signaling<sup>41</sup>. In MM, IGF1 is aberrantly expressed compared to primary hematopoietic cells<sup>25</sup>, allowing high potency of IGF1R inhibition at a lower dose. Mice xenograph studies on MM have shown that IGF1R inhibition with NVP-ADW742 to have a favorable in-vivo therapeutic window<sup>42</sup>. These studies, combined with our results, provided further evidence that overexpression of IGF1 could be a relevant biomarker for IGF1R inhibition treatment, elucidating the framework for future clinical trials using this strategy against MM.



**Figure 3.2.4c** Matrix of AUC values of MMs treatment with FGFR3 and IGF1R inhibitors compared to the other hematopoietic CCLs.

T(4;14)+ MM sensitivity towards FGFR3 inhibitor, AZD4547, suggests the dependency of t(4;14)+ MM on FGFR3 overexpression due to the translocation. There have been multiple studies on the efficacy of targeting FGFR3 in t(4;14)+ MM<sup>43-45</sup>. Small molecules that targets FGFR3 include several tyrosine kinase (RTK) inhibitors that are pan-inhibitors. However, we observed that these

compounds do not show the same efficacy as the FGFR3-selective inhibitor.

T(4;14)+ MMs and E1099K-ALLs do not display the same dependencies. The NSD2-affected CCLs behave more similarly to CCLs of the same lineage. Different strategies would have to be employed when targeting different NSD2-driven malignancies. Hopefully, the several selective HDAC and IGF1R inhibitors undergoing clinical trials<sup>46</sup> would lead to new strategies for patients with E1099K-ALLs and t(4;14)+ MM.

### **3.3 Experimental methods**

#### **Cell lines**

Cell lines used in profiling are listed in<sup>10</sup>. ALLs for retest were cultured as described in Chapter 2.3.

Cell lines used in correlation analysis are listed in the Appendix B.

#### **Profiling of hematopoietic cell lines for sensitivity to 545 small molecules**

Cells were plated at 500 per well in 1536-well plates (6  $\mu$ l recommended media), and 16 hours after plating compound (20 nl) was added using an Echo Liquid Handler system (Labcyte). After a 72-h incubation, CellTiter-Glo reagent (Promega) was added to each well (1.5  $\mu$ l) and luminescence was captured using a ViewLux  $\mu$ HTS Microplate Imager (Perkin-Elmer). Percent viability at a given compound concentration was calculated by comparing luminescence at that concentration to wells treated with only DMSO<sup>10</sup>. Concentration-response curves were fit to data using 3- or 4-parameter sigmoid functions in MATLAB. Sensitivities expressed as areas under concentration-response curves by numeric integration between a top concentration determined for each compound through 7 2-fold dilutions. To determine enrichment for E1099K and t(4;14) cell lines among those most (or least) sensitive to each compound, we identified the AUC threshold corresponding to the best enrichment of mutant lines using Fisher's exact test using a function from the MATLAB file

exchange (<http://www.mathworks.com/matlabcentral/fileexchange/15434-myfisher22>). Each threshold that changed the number of E1099K cell lines called sensitive (or non-responsive) was tested, and we retained the most significant odds-ratio value for each compound. ACME analysis were performed as described in <sup>26</sup>.

### **Retesting ALL panel of cell lines**

For DRPs, ALLs were plated at 2500 cells per well in 384-well plates (40 uL) and incubated overnight. The next day, compounds were added by pin transfer (100 nL in DMSO)(Cybio Vario). After a 72-hour incubation, CellTiter-Glo reagent (Promega) diluted 1:1 with PBS was added to each well (20 uL) and luminescence was measured using an Envision microplate reader (Perkin-Elmer). Percent viability was determined upon normalization with cells treated in DMSO. For measuring caspase activity, plates were incubated for an appropriate amount of time, Caspase-Glo reagent (Promega) was added to each well (20 uL) and luminescence was measured using a Spectramax microplate reader (Molecular Devices). For immunoblotting of proteins after treatment, ALLs were plated at 1.00E06 in 6-well plates (2 mL) and treated immediately with Merck60 or DMSO. After a 24-hour incubation, the cells were washed and lysed for protein collection.

### **Immunoblot and analysis of protein levels**

Experiments were performed as described in Chapter 2.3.

### **Antibodies**

Antibodies used in this study are listed in the Appendix A.

## Compounds

Full list of compounds used in profiling are listed in <sup>10</sup>. HDAC1/2 inhibitors are generously provided by Dr Florence Wagner and Dr Edward Holson. Other compounds used in retest are purchased from Cayman Chemical, Sigma-Aldrich and Tocris.

### Chemical characterization data of BRD2283 and BRD6597

#### BRD 2283

ESI+ MS:  $m/z$  285 ( $[M+H]^+$ ), <sup>1</sup>H NMR (500 MHz, d<sub>6</sub>-DMSO):  $\delta$  9.12 (s, 1H), 7.40 (d,  $J = 2$  Hz, 1H), 7.35 (dd,  $J = 5, 1$  Hz, 1H), 7.26 (dd,  $J = 8; 2$  Hz, 1H), 7.22 (dd,  $J = 4$  Hz; 1.5 Hz, 1H), 7.04 (dd,  $J = 5; 4.5$  Hz, 1H), 6.78 (d,  $J = 8.5$  Hz, 1H), 6.73-6.68 (m, 1H), 5.06 (s, 2H), 2.61-2.57 (m, 2H), 2.51-2.48 (m, 2H), 1.93-1.90 (m, 2H).

#### BRD 6597

ESI+ MS:  $m/z$  443 ( $[M+H]^+$ ), <sup>1</sup>H NMR (300 MHz, d<sub>6</sub>-DMSO):  $\delta$  9.85 (s, 1H), 8.78-8.76 (m, 1H), 8.48 (d,  $J = 6.0$  Hz, 2H), 8.09 (d,  $J = 9.0$  Hz, 2H), 7.94 (d,  $J = 9.0$  Hz, 2H), 7.48 (bs, 1H), 7.40-7.20 (m, 5H), 7.06 (t,  $J = 6.0$  Hz, 1H), 6.82 (d,  $J = 9.0$  Hz, 1H), 5.21 (bs, 2H), 3.64-3.52 (m, 2H), 2.91 (t,  $J = 9.0$  Hz, 2H).

### Genotyping and quantitative RT-PCR analysis

Experiments were performed as described in Chapter 2.3. Additional TaqMan primers (Life Technologies) used in this study were: HDAC1 (Hs02621185\_s1) and HDAC2 (Hs00231032\_m1).

### 3.4 References

1. Jaffe, J. D. *et al.* Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nat Genet* **45**, 1386–1391 (2013).
2. Kuo, A. J. *et al.* NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Molecular Cell* **44**, 609–620 (2011).
3. Lauring, J. *et al.* The multiple myeloma associated MMSET gene contributes to cellular adhesion, clonogenic growth, and tumorigenicity. *Blood* **111**, 856–864 (2008).
4. Rajkumar, S. V. Multiple myeloma: 2013 update on diagnosis, risk-stratification, and management. *American Journal of Hematology* **88**, 225–235 (2013).
5. Palumbo, A. *et al.* International Myeloma Working Group consensus statement for the management, treatment, and supportive care of patients with myeloma not eligible for standard autologous stem-cell transplantation. *J. Clin. Oncol.* **32**, 587–600 (2014).
6. PDQ Pediatric Treatment Editorial Board. Childhood Acute Lymphoblastic Leukemia Treatment (PDQ®): Patient Version. (2002).
7. Kassambara, A., Klein, B. & Moreaux, J. MMSET is overexpressed in cancers: Link with tumor aggressiveness. *Biochemical and Biophysical Research Communications* **379**, 840–845 (2009).
8. Hudlebusch, H. R. *et al.* The histone methyltransferase and putative oncoprotein MMSET is overexpressed in a large variety of human tumors. *Clin Cancer Res* **17**, 2919–2933 (2011).
9. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
10. Basu, A. *et al.* An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* **154**, 1151–1161 (2013).
11. Rees, M. G. *et al.* Correlating chemical sensitivity and basal gene expression reveals

- mechanism of action. *Nature Chemical Biology* **12**, 109–116 (2015).
12. Wagner, F. F., Wesmaly, M., Michel, U., Lewis, M. C. & Holson, E. B. Small molecule inhibitors of zinc-dependent histone deacetylases. *Neurotherapeutics* **10**, 589–604 (2013).
  13. Wagner, F. F. *et al.* Potent and selective inhibition of histone deacetylase 6 (HDAC6) does not require a surface-binding motif. *J. Med. Chem.* **56**, 1772–1776 (2013).
  14. Olson, D. E. *et al.* Discovery of the first histone deacetylase 6/8 dual inhibitors. *J. Med. Chem.* **56**, 4816–4820 (2013).
  15. Vannini, A. *et al.* Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor. *Proceedings of the National Academy of Sciences* **101**, 15064–15069 (2004).
  16. Moradei, O. M. *et al.* Novel aminophenyl benzamide-type histone deacetylase inhibitors with enhanced potency and selectivity. *J. Med. Chem.* **50**, 5543–5546 (2007).
  17. Holson, E., WAGNER, F. F. & Weiwer, M. Inhibitors of histone deacetylase. (2012).
  18. Methot, J. L. *et al.* Exploration of the internal cavity of histone deacetylase (HDAC) with selective HDAC1/HDAC2 inhibitors (SHI-1:2). *Bioorg Med Chem Lett* **18**, 973–978 (2008).
  19. Marks, P. A. & Breslow, R. Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nat Biotechnol* **25**, 84–90 (2007).
  20. Atadja, P. Development of the pan-DAC inhibitor panobinostat (LBH589): Successes and challenges. *Cancer Letters* **280**, 233–241 (2009).
  21. Ropero, S. *et al.* A truncating mutation of HDAC2 in human cancers confers resistance to histone deacetylase inhibition. *Nat Genet* **38**, 566–569 (2006).
  22. Dovey, O. M. *et al.* Histone deacetylase 1 and 2 are essential for normal T-cell development and genomic stability in mice. *Blood* **121**, 1335–1344 (2013).
  23. Nightingale, K. P. *et al.* Cross-talk between histone modifications in response to histone

- deacetylase inhibitors: MLL4 links histone H3 acetylation and histone H3K4 methylation. *J Biol Chem* **282**, 4408–4416 (2007).
24. Lee, J. H., Choy, M. L., Ngo, L., Foster, S. S. & Marks, P. A. Histone deacetylase inhibitor induces DNA damage, which normal but not transformed cells can repair. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 14639–14644 (2010).
  25. Sprynski, A. C. *et al.* The role of IGF-1 as a major growth factor for myeloma cell lines and the prognostic relevance of the expression of its receptor. *Blood* **113**, 4614–4626 (2009).
  26. Seashore-Ludlow, B. *et al.* Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. *Cancer Discov* **5**, 1210–1223 (2015).
  27. Chesi, M. *et al.* The t(4;14) Translocation in Myeloma Dysregulates Both FGFR3 and a Novel Gene, MMSET, Resulting in IgH/MMSET Hybrid Transcripts. *Blood* **92**, 3025–3034 (1998).
  28. Kalff, A. & Spencer, A. The t(4;14) translocation and FGFR3 overexpression in multiple myeloma: prognostic implications and current clinical strategies. *Blood Cancer Journal* **2**, e89 (2012).
  29. PDQ Adult Treatment Editorial Board. Chronic Myelogenous Leukemia Treatment (PDQ®): Patient Version. (2002).
  30. Jasmine Zain, O. A. O. Targeting histone deacetylases in the treatment of B- and T-cell malignancies. *Invest New Drugs* **28**, 58–78 (2010).
  31. Romanski, A. *et al.* Use of a novel histone deacetylase inhibitor to induce apoptosis in cell lines of acute lymphoblastic leukemia. *Haematologica* **89**, 419–426 (2004).
  32. Scuto, A. *et al.* The novel histone deacetylase inhibitor, LBH589, induces expression of DNA damage response genes and apoptosis in Ph- acute lymphoblastic leukemia cells. *Blood* **111**, 5093–5100 (2008).
  33. Stumpel, D. J. P. M. *et al.* Connectivity mapping identifies HDAC inhibitors for the treatment

- of t(4;11)-positive infant acute lymphoblastic leukemia. *Leukemia* **26**, 682–692 (2012).
34. Kim, J.-Y. *et al.* Multiple-myeloma-related WHSC1/MMSET isoform RE-IIBP is a histone methyltransferase with transcriptional repression activity. *Mol Cell Biol* **28**, 2023–2034 (2008).
  35. Marango, J. *et al.* The MMSET protein is a histone methyltransferase with characteristics of a transcriptional corepressor. *Blood* **111**, 3145–3154 (2008).
  36. Hajdu, I., Ciccia, A., Lewis, S. M. & Elledge, S. J. Wolf-Hirschhorn syndrome candidate 1 is involved in the cellular response to DNA damage. *Proceedings of the National Academy of Sciences* **108**, 13130–13134 (2011).
  37. Pei, H. *et al.* MMSET regulates histone H4K20 methylation and 53BP1 accumulation at DNA damage sites. *Nature* **470**, 124–128 (2011).
  38. Gombos, A., Metzger-Filho, O., Dal Lago, L. & Awada-Hussein, A. Clinical development of insulin-like growth factor receptor—1 (IGF-1R) inhibitors: At the crossroad? *Invest New Drugs* **30**, 2433–2442 (2012).
  39. Klinakis, A. *et al.* Igf1r as a therapeutic target in a mouse model of basal-like breast cancer. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2359–2364 (2009).
  40. S John Weroha, P. H. **IGF-1 Receptor Inhibitors in Clinical Trials—Early Lessons.** *J Mammary Gland Biol Neoplasia* **13**, 471–483 (2008).
  41. Gualberto, A. & Pollak, M. Emerging role of insulin-like growth factor receptor inhibitors in oncology: early clinical trial results and future directions. *Oncogene* **28**, 3009–3021 (2009).
  42. Mitsiades, C. S. *et al.* Inhibition of the insulin-like growth factor receptor-1 tyrosine kinase activity as a therapeutic strategy for multiple myeloma, other hematologic malignancies, and solid tumors. *Cancer Cell* **5**, 221–230 (2004).
  43. Martinelli, G., Tosi, P., Ottaviani, E., Soverini, S. & Tura, S. Molecular therapy for multiple myeloma. *Haematologica* **86**, 908–917 (2001).

44. de Brito, L. R. *et al.* Comparative pre-clinical evaluation of receptor tyrosine kinase inhibitors for the treatment of multiple myeloma. *Leukemia Research* **35**, 1233–1240 (2011).
45. Paterson, J. L. *et al.* Preclinical studies of fibroblast growth factor receptor 3 as a therapeutic target in multiple myeloma. *British Journal of Haematology* **124**, 595–603 (2004).
46. Fassnacht, M. *et al.* Linsitinib (OSI-906) versus placebo for patients with locally advanced or metastatic adrenocortical carcinoma: a double-blind, randomised, phase 3 study. *The Lancet Oncology* **16**, 426–435 (2015).

# Chapter IV

## Engineering of CRISPR-NSD2 cell lines to identify novel NSD2 biology

### Collaborators contributions:

- **Dr. Michelle Stewart** for her comments and insights on the experiments performed in this chapter.
- **Broad Technology Labs** for their assistance in performing RNA-seq in this chapter.

## 4.1 Introduction

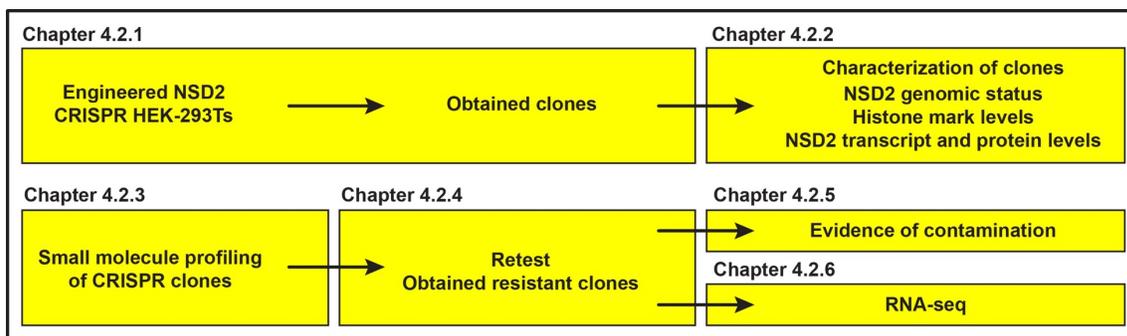
Studies on NSD2 in the field has focused on NSD2 in cancer, and we still have limited knowledge on the cellular biology of NSD2. Current methods to study the biology of NSD2 are limited by the quality of antibodies targeting NSD2 and its substrate/product <sup>1</sup>, the lack of selective-NSD2 inhibitors and low abundance of NSD2 in non-transformed cell lines <sup>2</sup>. There have been efforts, including in this lab, to develop a selective-NSD2 inhibitor for the clinic <sup>3</sup>, however, I believe that in understanding more about NSD2 pathway, we would be able to identify other alternative strategies in targeting these NSD2-dependent malignancies. I have described possible NSD2-dependencies in chapter 3 and in this chapter, I describe my efforts in exploring more of the NSD2 pathway in the cell using isogenic cell lines and small molecule profiling.

Isogenic cell lines have previously been engineered for NSD2, with the knockout of overexpression allele in t(4;14)+ MM CCL, KMS11 <sup>4</sup>. The pair of cell lines, KMS11 and its knockout (TKO), have been used in several studies on NSD2, including the study I performed in Chapter 3 <sup>5-8</sup>. I generated a set of NSD2 isogenic cell lines that would have a neutral lineage background without known malignancies of NSD2 (MM and ALL CCLs were thus excluded). I picked HEK293T due to its use versatile role in studying protein functions <sup>1</sup>. As a continuation of my study on NSD2-E1099K in Chapter 2 and 3, I inserted the E1099K mutation into the WT-NSD2 loci of HEK293T cells via CRISPR technology to obtain several clones that contain different NSD2 mutants.

CRISPR (Clustered regularly interspaced short palindromic repeats) technology consists of delivering or expressing a RNA-guided endonuclease, Cas9, that cuts specific parts of the genome <sup>9</sup>. The cell has two methods of repairing these double-stranded breaks. It could repair using non-homologous end joining (NHEJ), a less accurate repair pathway that aims to ligate the breaksite with higher probability of causing insertions, deletions and mutations, or homology directed repair (HR), a more accurate repair pathway that uses a template (usually the cell's other copy of the

chromosome) to repair the breaksite<sup>10</sup>. Engineering the E1099K mutation into NSD2 exon requires a template for HR, and HR happens at a lower rate compared to NHEJ, resulting in a collection of clones that has different forms of NSD2. This set of NSD2 clones are presumed to be isogenic, hence ideal for characterization of NSD2 functions and use in small molecule profiling in assessing differences in sensitivity depending on their NSD2 status.

In this chapter I'll describe my strategy in engineering E1099K mutation into HEK293T cells and characterizing clones in the study. In addition, RNA-seq was used to explore the expression differences in clones in the study. Brief scheme of this chapter is shown below (Figure 4.1).



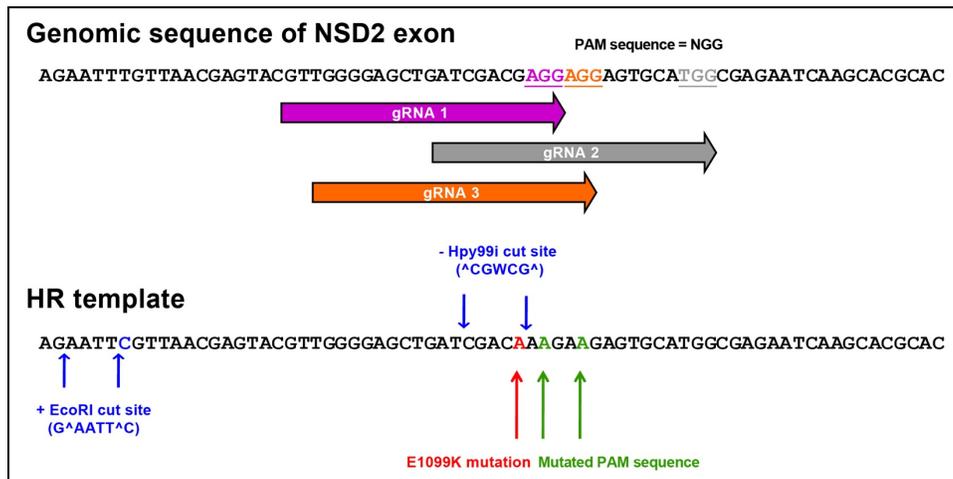
**Figure 4.1** Scheme of experiments described in this chapter.

## 4.2 Results and Discussion

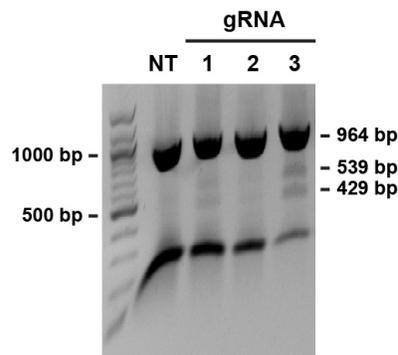
### 4.2.1 Engineering NSD2 cell lines using CRISPR

We performed CRISPR with homologous recombination (HR) on HEK293T cells. In this CRISPR experiment, we designed 3 gRNAs with sequences in close proximity to NSD2-1099 site in the genome with an oligo HR template that include several mutations: E1099K mutation, mutation to insert a EcoRI restriction site and mutations that remove the PAM sequence required for Cas9 targeting (Figure 4.2.1a). The cells were transfected and selected through puromycin. A sample of

each population of transfected cells was then assessed in their HR via an amplification of the genome at NSD2 exon loci followed by an EcoRI restriction site digest. Amplification of the genomic NSD2 exon yields a product of 964 bp and EcoRI digest of successful HR clones show up as bands of 429 and 539 bp. gRNA3 had the most prominent digest band (Figure 4.2.1b). We decided that the population cells engineered with gRNA3 would be used to seed clones to obtain isogenic NSD2 cell lines.

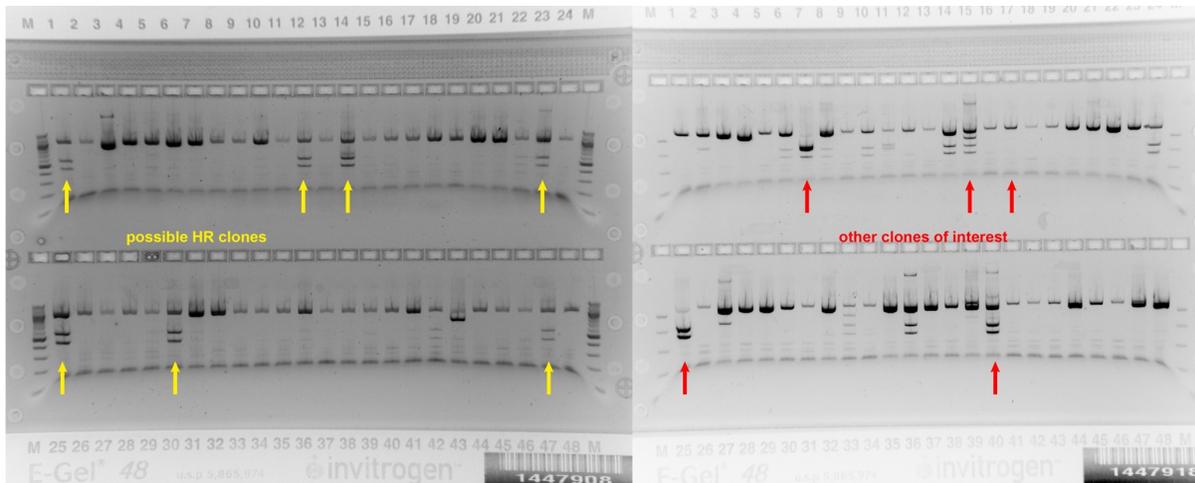


**Figure 4.2.1a** gRNA sequence targeting NSD2 exon in HEK cells and corresponding homologous recombination (HR) oligo template that introduces the E1099K mutation, a EcoRI cut site and mutated PAM sequences into the genome.



**Figure 4.2.1b** EcoRI restriction digest pattern of amplified gDNAs from population of HEK293T cells that were transfected with different gRNAs. Cells that underwent HR have a digest band at 429 and 539 bp.

We seeded single cell colonies in 96-well plates and as they grew, we transferred to 48-well and 24-well plates. Genomic status of each clone was evaluated when clones reached confluency in a 24-well plate. The clones were assessed using the same method as described previously. Selection of clones include clones with/without EcoRI digest pattern, clones with abnormal amplification and clones with abnormal digest pattern (Figure 4.2.1c).

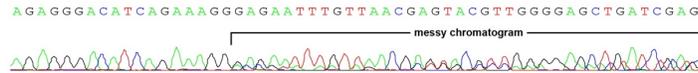


**Figure 4.2.1c** EcoRI restriction digest pattern of gDNA from singular clones. A mixture of clones with different digest patterns was expanded. Possible HR clones (yellow arrows) were identified to have an engineered EcoRI cut site. Other clones of interest (red arrows) were identified to have abnormal digest patterns, suggesting a non-WT NSD2 status.

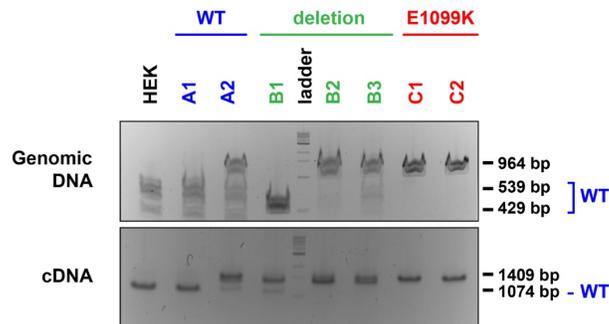
#### **4.2.2 Characterization of clones obtained**

We picked several clones and evaluated their NSD2 genomic and transcript status through gDNA and cDNA sequencing. 2 clones expressed WT-NSD2 (A1 and A2), 3 clones expressed deletions within NSD2 (B1, B2 and B3) and 2 clones expressed E1099K-NSD2 (C1 and C2) (Figure 4.2.2a). A2 projected a messy sequencing chromatogram that is typical of an NHEJ repaired allele, suggesting that the alleles in this clone were repaired non-uniformly. It is difficult to discern sequences at each allele of A2 through sequencing alone. Hence, the genomic and transcript status

of NSD2 in A2 was determined via Hpy99i restriction digest that would yield bands at 539 bp and 429 bp for allele containing WT-NSD2 and 1074 bp for a transcript containing WT-NSD2. The digests revealed that A2 is expressing WT-NSD2 to a small extent (Figure 4.2.2b).

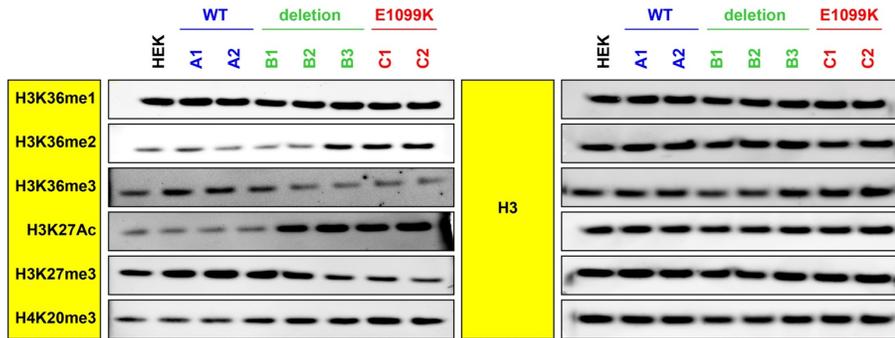
clone	RNA sequence of NSD2 transcript	status
HEK	caagagggacatcagaaaggagaatttgttaacgagtagcttggggagctgatcgacgaggaggagtgcattggcgagaatc aagcagcgacacgagaacgacatcaccacttctacatgctcactatagacaaggaccgtataatagacgctggcc	Homozygous WT
A1	caagagggacatcagaaaggagaatttgttaacgagtagcttggggagctgatcgacgaggaggagtgcattggcgagaatc aagcagcgacacgagaacgacatcaccacttctacatgctcactatagacaaggaccgtataatagacgctggcc	Homozygous WT
A2	<p>AGAGGGACATCAGAAAGGGAGAATTTGTTAACGAGTAGCTTGGGGAGCTGATCGAG</p> 	Heterozygous WT (NHEJ)
B1	caagagggacatcagaaagg-----caagaccgtataatagacgctggcc	Deletion
B2	caagagggacatcagaaa-----ggaccgtataatagacgctggcc	Deletion
B3	caagagggacatcagaaagg-----ac-----c-----ga-----at-----a-----t-----a-----a----- -----acgctggcc	Deletion
C1	caagagggacatcagaaaggagaattcgttaacgagtagcttggggagctgatcgacaagaagagtagcattggcgagaatc aagcagcgacacgagaacgacatcaccacttctacatgctcactatagacaaggaccgtataatagacgctggccc	Homozygous E1099K
C2	caagagggacatcagaaaggagaattcgttaacgagtagcttggggagctgatcgacaagaagagtagcattggcgagaatc aagcagcgacacgagaacgacatcaccacttctacatgctcactatagacaaggaccgtataatagacgctggccc	Homozygous E1099K

**Figure 4.2.2a** Mutations and deletions in the sequences are highlighted in red. HEK represents the parental line that underwent CRISPR. Clones A have expression of WT-NSD2 transcript. Clones B have deletions in the transcript. Clones C underwent homologous recombination with donor E1099K oligo to have expression of E1099K-NSD2 transcript. Sequence chromatography of clone A2 is shown to demonstrate sequence chromatography of a clone that has undergone non-homologous end joining (NHEJ) to have a messy chromatogram.

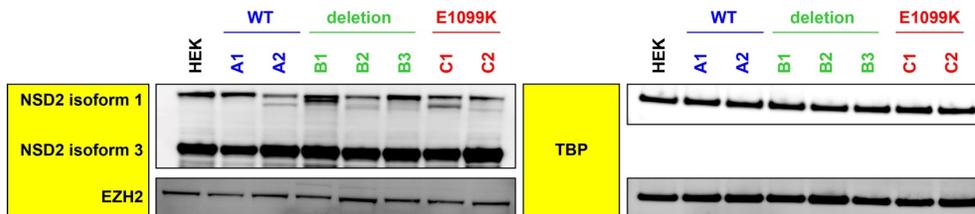


**Figure 4.2.2b** Genomic DNA and cDNA were collected from selected clones and PCR the region of CRISPR target. NSD2 WT status was assessed via the presence of the Hpy99i cut site. gRNA of WT-NSD2 is expected to have digestion bands of 539 and 429 bp. cDNA of WT-NSD2 is expected to have digestion band of 1074 bp.

We assessed levels of chromatin modifications in the clones via immunoblot and observed that E1099K clones have higher H3K36me2 levels and lower H3K27me3 levels, supporting our findings in chapter 2 (Figure 4.2.2c). NSD2 protein from different clones were observed to be of different molecular weight, supporting the differences in transcript sequence in each clone, especially in the deletion clones (Figure 4.2.2d). For example, B1 has a deletion in its NSD2 transcript that translated to a protein of lower molecular weight. None of the deletion clones resulted in the ablation of NSD2 protein or H3K36 methylation marks. The levels of NSD2 protein does not correlate with the levels of H3K36 methylation, suggesting that E1099K mutation indeed enhanced catalytic activity of these clones. There is no difference in levels of EZH2 protein in the clones, the difference in H3K27me3 levels in these clones cannot be attributed to any difference in the levels of EZH2 protein (Figure 4.2.2d).



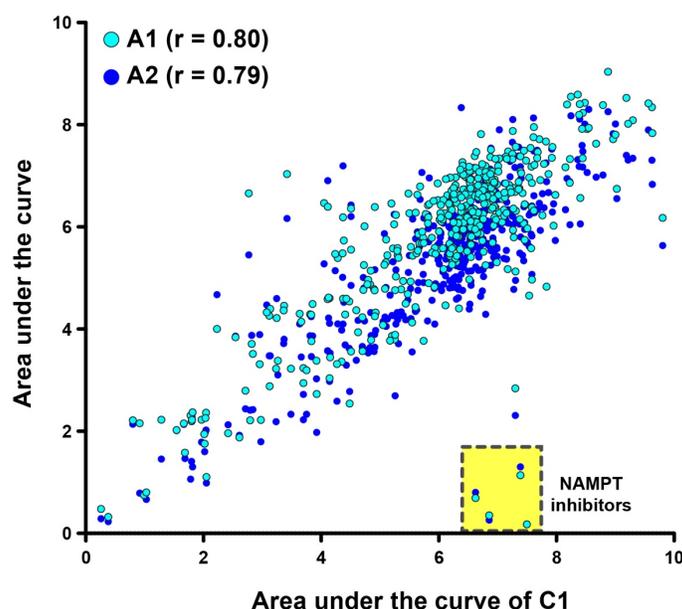
**Figure 4.2.2c** Chromatin profile of clones. Clones with E1099K mutation inserted have higher levels of H3K36me2 and lower levels of H3K27me3. H3 was blotted as a loading control.



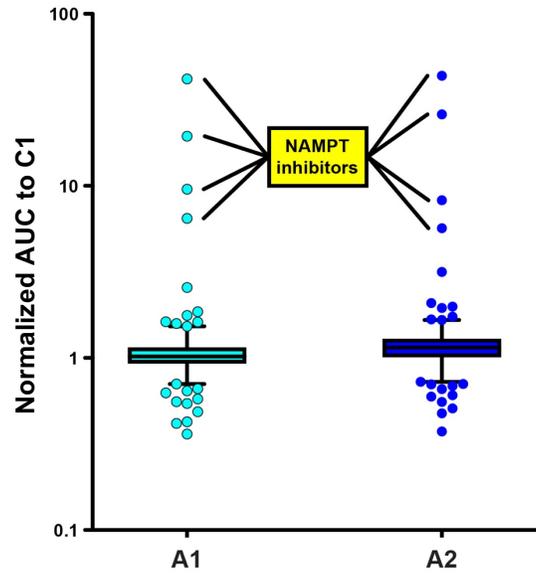
**Figure 4.2.2d** NSD2 levels are different among different clones due to different CRISPR cuts in the NSD2 genome, EZH2 levels are similar among clones. TBP was blotted as a loading control

### 4.2.3 Clones have different sensitivity towards NAMPT inhibitors

Using methods described in Chapter 3, we picked three clones (A1, A2 and C1) and treated them with 440 small molecules. AUC was calculated and normalized across compound concentrations as a surrogate to cell viability. We used calculated AUC values to analyze differences in viability that could be attributed to change in NSD2 genomic/transcript status. AUCs of C1 v A1 and C1 v A2 were plotted to identify outliers that affected C1 differently compared to A1 or A2 (Figure 4.2.3a). AUCs of C1, A1 and A2 correlated highly to each other (C1 v A1  $r = 0.80$ , C1 v A2  $r = 0.79$ , A1 v A2  $r = 0.92$ ), demonstrating that most compounds have the same effect on the clones and suggests that the clones have a highly similar cellular environment. To clearly identify outliers that show the biggest difference in affecting viability of C1 v A1, A2, we calculated the differences between the AUCs of C1 v A1 and C1 v A2 and plotted them (Supplementary Table 4.1, Figure 4.2.3b). Nicotinamide phosphoribosyltransferase inhibitors (NAMPTi) were identified as the top outliers, killing A1 and A2 more than C1.

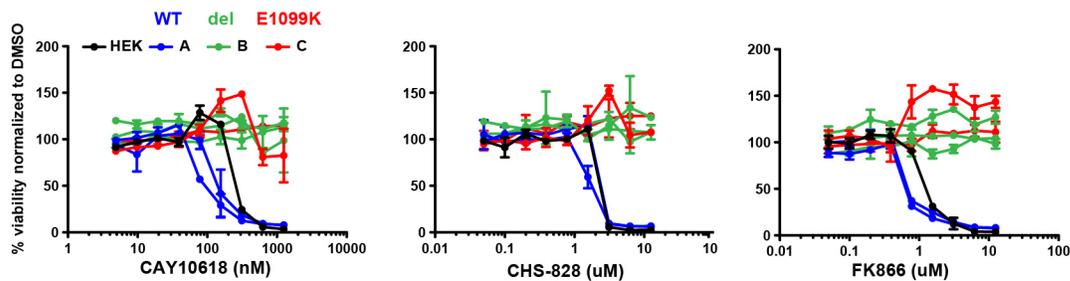


**Figure 4.2.3a** Scatter plot of C1 AUC values plotted against A1 and A2 AUC values. A1 and A2 AUC values overlap with each other well. NAMPTi are singled out as outliers that confers sensitivity to A1 and A2 and resistance to C1.



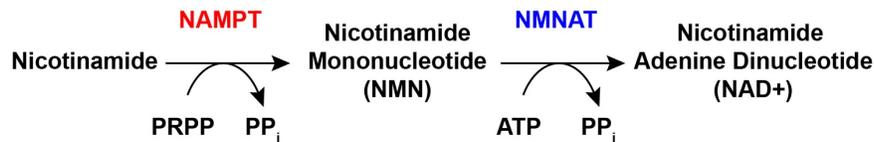
**Figure 4.2.3b** AUC values were normalized as ratios of C1:A1 and C1:A2 with compounds that selectively kill A1 and A2 ranked as high outliers. The top 4 compounds are NAMPTi.

We reconfirmed the observation by treating the all the clones (A1, A2, B1, B2, B3, C1 and C2) with a panel of NAMPTi. Clones with WT-NSD2 expressed are sensitive to NAMPT inhibition, while clones without WT-NSD2 expressed (deletions and E1099K) had the opposite sensitivity towards these compounds.

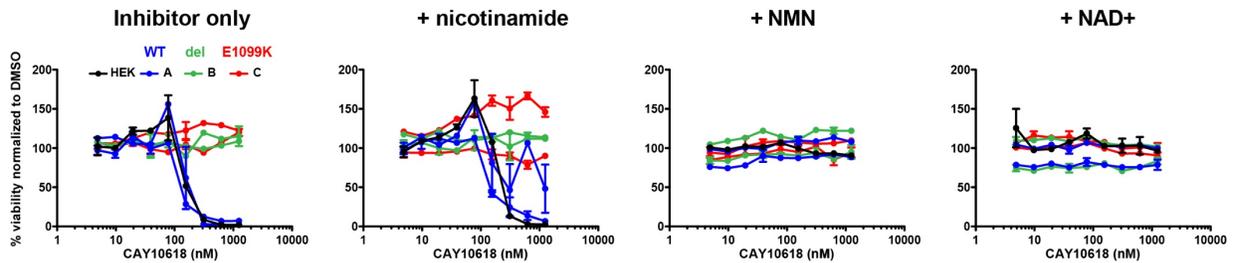


**Figure 4.2.3c** Treatment of NAMPT inhibitors against the set of CRISPR clones. Clones with WT-NSD2 (HEK and A) are sensitive towards NAMPT inhibition while clones with deletions in NSD2 (B) and E1099K-NSD2 (C) are resistant towards NAMPT inhibition.

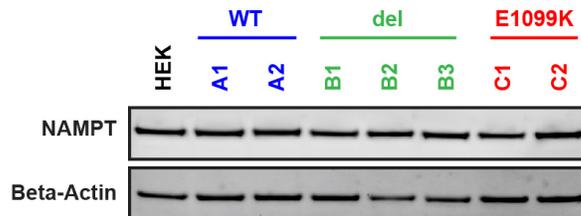
NAMPT is an essential enzyme in the biosynthesis of nicotinamide adenine dinucleotide (NAD<sup>+</sup>) (Figure 4.2.3d).<sup>2,3</sup> We performed rescue experiments by co-treating clones with NAMPT inhibitor (CAY10618) and downstream metabolites in the NAMPT pathway, nicotinamide mononucleotide (NMN) and NAD<sup>+</sup>. Cell death in WT-NSD2 clones were rescued when co-treated with NMN and NAD<sup>+</sup> but not when co-treated with nicotinamide (substrate of NAMPT) (Figure 4.2.3e), validating the dependency of NAMPT activity in these sensitive clones. CRISPR engineering might have affected NAMPT protein levels in the clones. We assessed NAMPT protein levels using immunoblot and observed there to be no different between clones (Figure 4.2.3f).



**Figure 4.2.3d** Pathway of NAD<sup>+</sup> biosynthesis from nicotinamide. NAMPT catalyzes the rate-limiting step of converting nicotinamide to NMN<sup>3</sup>.

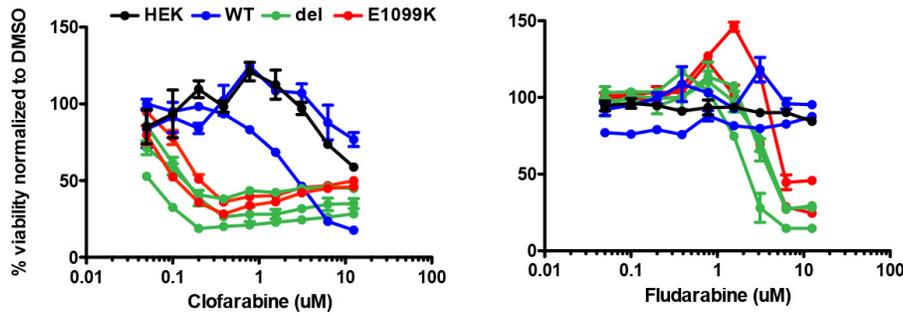


**Figure 4.2.3e** Co-treatment of NAMPT inhibitor with downstream NAMPT metabolites are able to rescue cell death from NSD-WT clones but not with NAMPT substrate.



**Figure 4.2.3f** Immunoblotting of clones reveal that clones have similar levels of NAMPT protein. Beta-actin was blotted as a loading control.

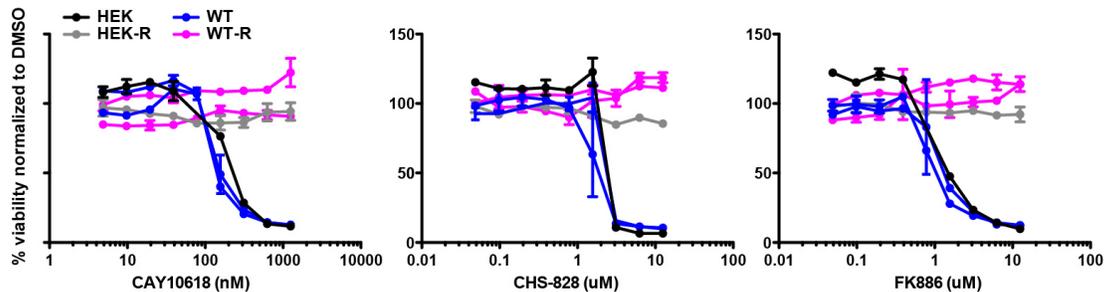
We also observed other compounds that had different activity towards these clones. WT-NSD2 clones are resistant to clofarabine and fludarabine while clones without WT-NSD2 expressed (deletions and mutants) had the opposite sensitivity towards these compounds (Figure 4.2.3g).



**Figure 4.2.3g** Treatment of clofarabine and fludarabine against the set of CRISPR clones. Clones with WT-NSD2 (HEK and A) are resistant while clones with deletions in NSD2 (B) and E1099K-NSD2 (C) are sensitive.

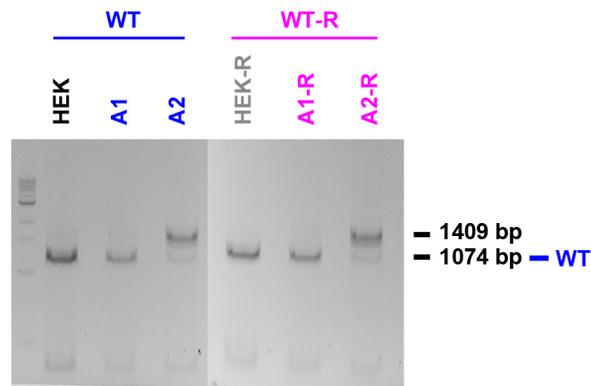
#### 4.2.4 Acquisition of NAMPTi-resistant WT-NSD2 clones

NAMPTi-resistant WT-NSD2 (WT-R) clones were obtained after several passages of parental HEK293T and WT-NSD2 clones (A1 and A2) (Figure 4.2.4a). We did not purposely plan for the creation of these clones but they provided a set of controls to study a possible NAMPT-NSD2 relationship.

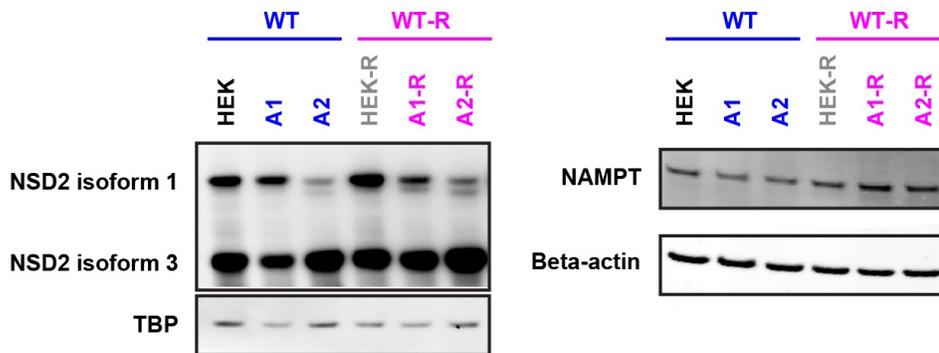


**Figure 2.2.4a** Treatment of NAMPTi with HEK, HEK-R, WT and WT-R clones.

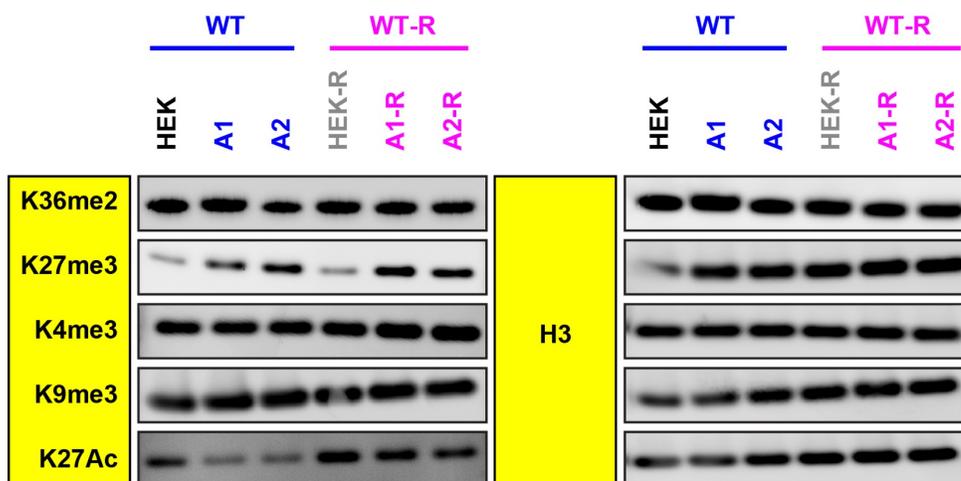
We assessed NSD2 transcript status of WT and WT-R clones as described in 4.2.3. The clones were observed to have no changes to NSD2 sequence (Figure 4.2.4b), suggesting that it was not an introduction of mutations or deletions that conferred the NAMPTi-resistance. NSD2 and NAMPT protein levels were also observed to be the same (Figure 4.2.4c). K4me3, K9me3, K27me3 and K36me2 levels of these isogenic clones were assessed via immunoblot and there was no observable difference. K27Ac levels in WT-R clones are higher than the non-resistant clones.



**Figure 4.2.4b** cDNA was obtained from RNA of WT and WT-R clones. Hpy99i digest patterns of amplified NSD2 transcript from WT and WT-R clones show that they have the same NSD2 status in transcript.

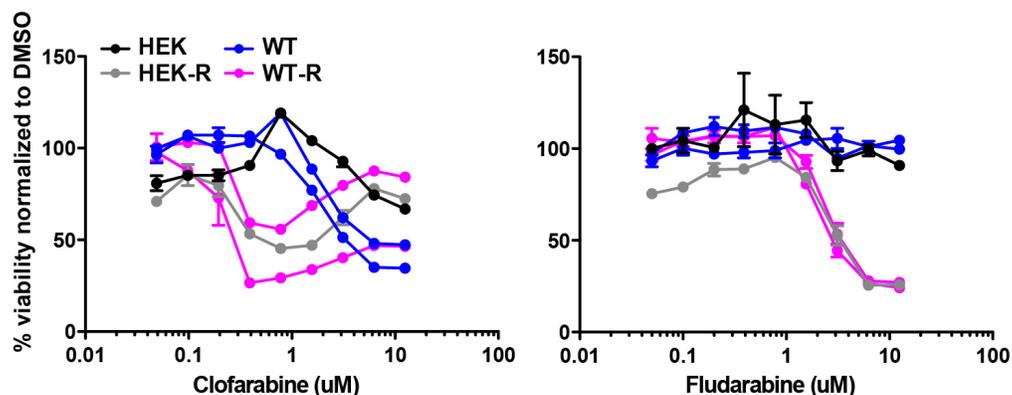


**Figure 4.2.4c Left:** Immunoblot of NSD2 protein levels in WT and WT-R clones. TBP was blotted as a loading control. **Right:** Immunoblot of NAMPT protein levels in WT and WT-R clones. Beta-actin was blotted as a loading control.



**Figure 4.2.4d** Immunoblot of chromatin marks in WT and WT-R clones. K27Ac levels are higher in WT-R compared to WT clones. H3 was blotted as a loading control.

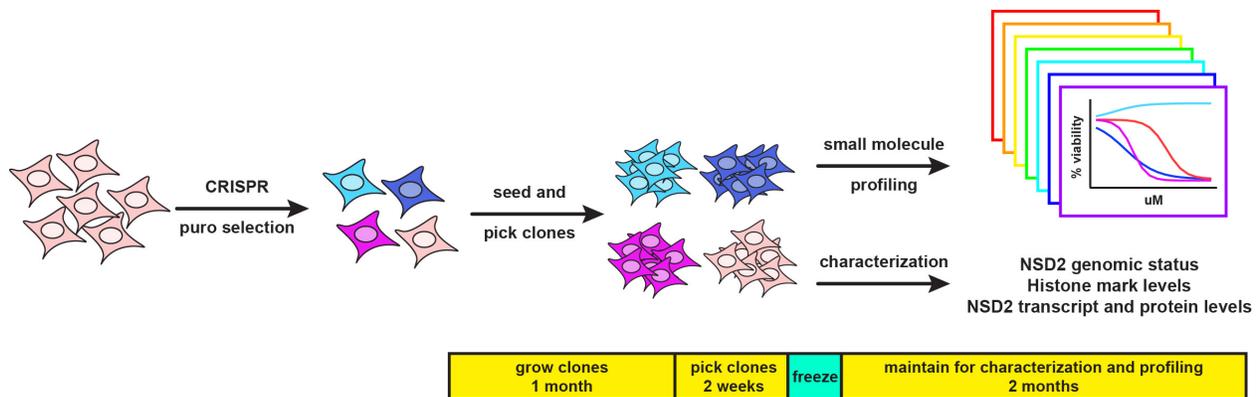
These new clones also behaved differently when treated with clofarabine and fludarabine. The WT-R clones are now sensitive to these compounds while the WT clones remains resistant (Figure 4.2.4e). This suggests a possible relationship between NAMPT and targets of clofarabine/fludarabine.



**Figure 4.2.4e** Treatment of clofarabine and fludarabine against the set of isogenic clones. NAMPTi-sensitive (HEK/A1/A2) clones are resistant while NAMPTi-resistant clones (HEK-R/A1-R/A2-R) are resistant.

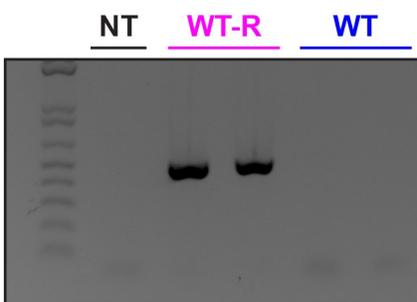
#### 4.2.5 Evidence of contamination

At this point, we have cultured WT and HEK clones for awhile and it is known that there is an increase in probability of spontaneous changes in cellular state or contamination of the culture when cells have been kept for too long<sup>4,5</sup>. The average duration of study for each batch of freshly thawed culture was 2 months. In addition to that, the clones were grown from single cell colony which spans a duration of 2 months. It can be safely assumed clones are maintained for 2-3 months after splitting from their parental CRISPR population (Figure 4.2.5a).

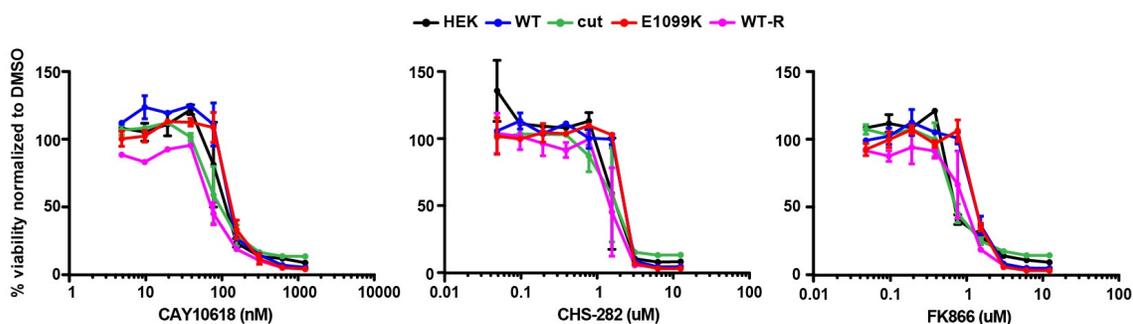


**Figure 4.2.5a** Scheme of clone maintenance after CRISPR engineering.

The NAMPTi resistance in WT-NSD2 clones arose spontaneously without changes in genomic or transcript status of NSD2, led us to hypothesize that NSD2 might not be the main driver in this phenotype. We tested the different clones for mycoplasma contamination using PCR with primers spanning a wide range of mycoplasma species<sup>6</sup> and results were that all clones that were NAMPTi resistant were mycoplasma positive (Figure 4.2.5b). Cells were treated two weeks with anti-mycoplasma reagent and retreated with NAMPTi. Mycoplasma-free cells regardless of their NSD2 status all tested sensitive towards NAMPTi (Figure 4.2.5c).



**Figure 4.2.5b** Gel of PCR product obtained after PCR with DNA extracted from media of clones. Presence of mycoplasma can be discerned by the positive PCR band. NT was ran with primers only.

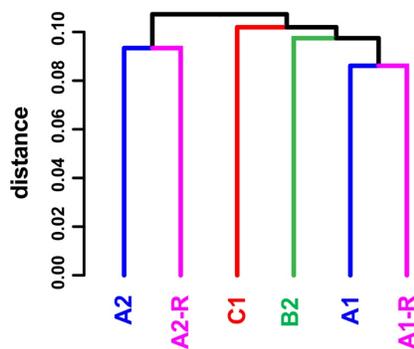


**Figure 4.2.5c** Treatment of NAMPTi against a variety of clones. Previously resistant clones (cut, E1099K and WT-R) display sensitive phenotype after treatment with anti-mycoplasma reagent.

Keeping in mind the contamination of clones and its relation to the phenotype, sequencing experiments that have been performed would be analyzed with augmented methods to best separate effects of contamination from effects of NSD2 status. RNA-seq explained later in the same chapter (4.2.6) would focus on differential expression analysis on WT-NSD2 vs non-WT-NSD2. Further exploration of the contamination would be described in later chapter (Chapter 5). Difference in K27Ac levels was observed to arise upon contaminated cell lines, leading us to believe that the change in K27Ac levels in clones was a consequence of mycoplasma contamination.

#### 4.2.6 RNA-seq on clones

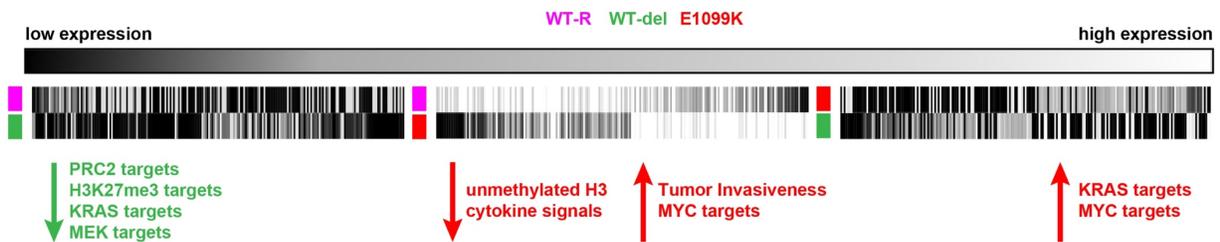
We extracted RNA in duplicates from clones A1, A2, A1-R, A2-R, B2 and C1 and sent them in for sequencing by the Broad Technologies Lab (BTL) using their protocol for SMART-Seq<sup>7</sup> (12 samples in total). Sequences were aligned using TopHat and differential expression were analyzed using the Cufflinks suite<sup>8</sup>. Aligned reads were also checked for quality using FastQC, a quality control tool for high-throughput sequencing data, and our sequences behaved as per expected for a RNA-seq experiment<sup>9</sup>. Dendrograms of the expression profiles were used to assess similarity between the clones and as expected, pairs of isogenic clones (A1/A1-R and A2/A2-R) clustered closely, suggesting their transcript profile to be most similar (Figure 4.2.6a).



**Figure 4.2.6a** Dendrogram of RNA-seq samples calculated from replicates of each sample. Jensen-Shannon distances were calculated using the cummeRbund package<sup>8</sup>.

For assessing NSD2's impact on transcript expression, we focused on comparing transcript levels between NAMPTi-resistant cell lines: WT-NSD2 NAMPTi-resistant (A1-R) vs del-NSD2 (B2), WT-NSD2 NAMPTi-resistant (A1-R) vs E1099K-NSD2 (C1) and del-NSD2 (B2) vs E1099K-NSD2 (C1) using Cuffdiff package (Supplementary Table 4.2)<sup>10</sup>. To calculate differential expression between pairs, I used ranked list of significant genes (removing novel isoforms identified by cuffdiff) as calculated by cuffdiff, ranking them by (q-value x log<sub>2</sub>(fold\_change)) to best represent the significance and difference in fold\_change, however, that resulted in a limited list of genes with

extreme values that did not process well with preranked-GSEA. Instead, I used read counts generated by the analysis as input for GSEA using curated gene sets (C2) module, oncogenic signatures (C6) module and hallmark gene sets (H) module (Supplementary Table 4.3)<sup>11</sup>. Del-NSD2 samples have lower levels of transcripts related to PRC2 and H3K27me3 targets compared to WT-NSD2 NAMPTi-resistant samples. Transcript levels of oncogenic markers, KRAS and MEK targets were also observed to be lower in del-NSD2 samples. E1099K-NSD2 samples have lower levels of transcripts related to unmethylated H3 and cytokine signals and higher levels of transcripts related to tumor invasiveness and c-Myc. When comparing del-NSD2 and E1099K-NSD2, E1099K-NSD2 samples had higher level of transcripts relating to oncogenic behavior, KRAS and c-Myc targets (Figure 4.2.6b).



**Figure 4.2.6b** Heatmaps of differentially expressed transcripts in each comparison set. From left to right: WT-NSD2 vs WT-del, WT-NSD2 vs E1099K-NSD2, del-NSD2 vs E1099K-NSD2. Enriched gene sets were identified using GSEA<sup>11</sup>.

#### 4.2.7 Concluding remarks

Our efforts in this chapter have generated multiple hypotheses that one could explore in expanding NSD2 role in the cell. We applied the use of CRISPR engineering to understand NSD2 biology by creating clones with different versions of the NSD2 transcript. Engineering NSD2 E1099K mutation into a WT-NSD2 cell line increased NSD2's catalytic activity in the cell, as observed with an increase of K36me2 levels. This suggests that CRISPR technology can be used in rapidly characterize gain-of-function or loss-of-function mutations in diseases. However, this

method would only be limited to known activities of the enzyme, it would not be helpful in the efficient discovery of mutations that have new roles in the cell unrelated to the WT.

Combining CRISPR with small-molecule profiling is a useful hypothesis generating tool that can be applied to other isogenic cell line systems. We observed differential response of the NSD2 clones to NAMPTi and clofarabine. Sensitivity to NAMPTi was initially hypothesized to be related to the presence of WT-NSD2. However, the rise of NAMPTi-resistant WT-NSD2 clones proved that hypothesis to be false. These isogenic NAMPTi-sensitive and NAMPTi-resistant clones provided useful pairs of cell lines to study this phenomenon. Closer observation of epigenetic marks of NAMPTi-resistant vs NAMPTi-sensitive clones revealed that K27Ac levels are elevated in NAMPTi-resistant cell lines. Upon treatment of NAMPTi, NAMPTi-sensitive cell lines have a decrease in K27Ac but not NAMPTi-resistant cell lines. We hypothesize that the resistance is epigenetic in nature. Further testing of WT-NSD2 NAMPTi-resistant cell lines revealed that the NAMPTi-resistant phenotype is driven by the presence of mycoplasma in culture.

Mycoplasma contamination in cell culture is a wide-spread phenomenon first documented in 1956<sup>12</sup>. Since then, there have been efforts to detect and eradicate mycoplasma contamination<sup>13</sup>, but the percentages of contaminated cultures appears to remain significant{Armstrong:2010cu}. Mycoplasma contamination has been observed to transform cell states including inducing oncogenic transformation in benign cells<sup>14,15,16</sup>. Due to the prevalence of mycoplasma pathogenicity in humans, there have been several documentations of mycoplasma associated with tumor tissues from cancer patients<sup>17</sup>. We are the first to observe an epigenetic change with mycoplasma infection as well as a drastic difference in NAMPTi sensitivity. Further research into this phenotype could yield insights into host-pathogen characteristics that could be targeted as potential leads for anti-cancer therapeutics.

RNA-seq performed on the clones served to concur with NSD2 function published in literature and as studied in previous chapters of this thesis. del-NSD2 clones have a downregulation of transcript related to NSD2 function (H3K27me3 levels and PRC2 targets). Previous literature have connected NSD2 function with EZH2 (a component of PRC2 complex)<sup>18,19</sup>, an unfunctional NSD2 in the clone exhibited expected decrease in transcripts of PRC2 target genes. E1099K-NSD2 has an upregulation of transcripts related to oncogenic behavior, that we have observed in our cellular studies of E1099K-ALL cell lines in Chapter 2.

Moving forward, it would be helpful to reconfirm these findings using mycoplasma-free clones. Experiments, such as small molecule profiling, will have to be repeated on mycoplasma-free clones to obtain informed hypotheses on novel NSD2 biological role in the cell.

### **4.3 Experimental methods**

#### **Cell lines**

HEK-293Ts were purchased from Takara Bio USA and cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco) supplemented with 10% fetal bovine serum (Gibco) and 1% penicillin/streptomycin (Gibco).

#### **Generating CRISPR clones**

Cloning was performed as described in<sup>20</sup>. Plasmid PX459 was a generous gift from Dr Feng Zhang. gRNA was designed using CRISPR design tool at [crispr.mit.edu](http://crispr.mit.edu). gRNA sequence used in this study were (PAM sequence NGG are in bold):

gRNA1: CGTTGGGGAGCTGATCGACG**AGG**

gRNA2: GATCGACGAGGAGGAGTGCAT**TGG**

gRNA3: TGGGGAGCTGATCGACGAGG**AGG**

HR sequence:

```
CCCGACACTGAGGATTGGTCAGCACGCTTTTGGTCATGGCCACATGCTTGTGATTTCC  
AGGGAGAATTTGTTAACGAGTACGTTGGGGAGCTGATCGACAAAGAAGAGTGCATG  
GCGAGAATCAAGCACGCACACGAGAACGACATCACCCACTTCTACATGCTCACTATAG  
ACAAGGTAATGCGGAACTCCACTGTGAG
```

HEK293T cells were transfected with 5 ug of px459 vector and 1 uL of 100 uM HR template using Neon transfection system (Thermo Fisher Scientific). Cells were electroporated at 1100V for 2 ms using pulse 2. Puromycin (Gibco) was added to the cells at concentration of 1 ug/mL 24 hours after transfection.

Genomic DNA was harvested from global CRISPR population of each gRNA using AllPrep DNA/RNA Mini Kit (Qiagen). Allele genotyping was studied by amplifying regions of NSD2 using primer pairs: 5'-ACTGCTGACCCTGATGTATTT-3' and 5'-CAGATGAGGAAGTCTCAGCATC-3' using PrimeSTAR® GXL DNA Polymerase (Takara Bio USA). Amplified product was digested overnight with EcoRI restriction enzyme (New England Biolabs). Digested product was analyzed on 10-well 2% agarose e-gel (Thermo Fisher Scientific).

Global population of gRNA3 transfected HEKs were diluted and plated into 96-well plates at approximately 2 cells/well. Single colonies were allowed to grow for a week before observing under the microscope for the presence of colonies in each well. Wells with one growing colony were transferred to a 24-well plate to expand the culture. Upon confluency, clones were expanded to a 12-well plate and were split into 2 duplicate sets of 12-well plates. gDNA of 96 clones were extracted using Wizard® SV 96 Genomic DNA Purification System (Promega). The clones were amplified

and digested with EcoRI as described above. Digested product was analyzed on 48-well 1% agarose e-gel (Thermo Fisher Scientific).

Selected clones were further expanded onto flasks with larger volume. gDNA and RNA of clones were extracted and analyzed as described in Chapter 2.3. Hpy99I (New England Biolabs) digestion of amplified products was performed at 37C for 1 hour.

### **Antibodies**

Antibodies used in this study are listed in the Appendix A.

### **Immunoblot and analysis of protein levels**

Protocol for immunoblot and analysis of protein levels was done as described in Chapter 2.3.

### **Profiling of CRISPR clones for sensitivity to 440 small molecules**

Cells were plated at 2500 per well in 384-well plates (40 uL) in duplicates. The next day, compounds were added by pin transfer (100 nL in DMSO) (Cybio Vario). After a 72-hour incubation, CellTiter-Glo reagent (Promega) diluted 1:1 with PBS was added to each well (20 uL) and luminescence was measured using an Envision microplate reader (Perkin-Elmer). Percent viability at a given compound concentration was calculated by comparing luminescence at that concentration to wells treated with only DMSO <sup>21</sup>. Area under the curve was calculated using zoo package in R <sup>22</sup>.

Compounds for the retest were purchased from Cayman Chemical.

## **Mycoplasma testing and treatment**

Media from culture was collected and spun down at 16,000 x g for 5 minutes. Pellet was washed with PBS once. Genomic DNA was extracted using Wizard® Genomic DNA Purification Kit (Promega). PCR was performed using primers described in <sup>6</sup> and PrimeSTAR® GXL DNA Polymerase (Takara Bio USA). PCR product was purified using MinElute Reaction Cleanup Kit (Qiagen) and analyzed on 10-well 2% agarose e-gel (Thermo Fisher Scientific). Contaminated cells were treated with 25 ug/mL Plasmocin™ (Invivogen) for 2 weeks, passaging every 48 hours, and retested for mycoplasma contamination.

## **RNA-seq and analysis**

RNA from clones were extracted using RNeasy mini kit (Qiagen) with DNase digestion. RNA samples were quantified using Qubit assay (Thermo Fisher Scientific) and submitted to Broad Technology Platform for their SMART-seq2 protocol <sup>7</sup>. Results from the RNA-seq was provided as raw data in FASTA files, TopHat-aligned reads in bam files and cufflinks analyzed normalized fragments per kilobase of transcript per million mapped reads (fpkm) counts. FPKM counts were analyzed using GSEA module <sup>10</sup> in GenePattern <sup>23</sup> to rank differential transcripts clones using the C2 (curated gene sets), C6 (oncogenic signature), and H (hallmark gene sets) collection for the analysis.

## **4.4 References**

1. Lin, Y.-C. *et al.* Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nat Comms* **5**, 4767 (2014).
2. Peiró, C., Romacho, T., Carraro, R. & Sánchez-Ferrer, C. F. Visfatin/PBEF/Nampt: A New Cardiovascular Target? *Front. Pharmacol.* **1**, (2010).
3. Garten, A., Petzold, S., Körner, A., Imai, S.-I. & Kiess, W. Nampt: linking NAD biology,

- metabolism and cancer. *Trends in Endocrinology & Metabolism* **20**, 130–138 (2009).
4. Masters, J. R. & Stacey, G. N. Changing medium and passaging cell lines. *Nature Protocols* **2**, 2276–2284 (2007).
  5. Stepanenko, A. A. & Dmitrenko, V. V. HEK293 in cell biology and cancer research: phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. *Gene* **569**, 182–190 (2015).
  6. Uphoff, C. C. & Drexler, H. G. Detecting mycoplasma contamination in cell cultures by polymerase chain reaction. *Methods Mol. Biol.* **731**, 93–103 (2011).
  7. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods* **10**, 1096–1098 (2013).
  8. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562–578 (2012).
  9. Andrews, S. *FastQC: a quality control tool for high throughput sequence data.* Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. (Accessed: 6 July 2016)
  10. Gould, J., Getz, G., Monti, S., Reich, M. & Mesirov, J. P. Comparative gene marker selection suite. *Bioinformatics* **22**, 1924–1925 (2006).
  11. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545–15550 (2005).
  12. ROBINSON, L. B., WICHELHAUSEN, R. H. & ROIZMAN, B. Contamination of Human Cell Cultures by Pleuropneumonia-like Organisms. *Science* **124**, 1147–1148 (1956).
  13. Stanbridge, E. Mycoplasmas and cell cultures. *Bacteriological Reviews* **35**, 206–227 (1971).
  14. Namiki, K. *et al.* Persistent Exposure to Mycoplasma Induces Malignant Transformation of Human Prostate Cells. *PLoS One* **4**, e6872 (2009).

15. Feng, S. H., Tsai, S., Rodriguez, J. & Lo, S. C. Mycoplasmal infections prevent apoptosis and induce malignant transformation of interleukin-3-dependent 32D hematopoietic cells. *Mol Cell Biol* **19**, 7995–8002 (1999).
16. Tsai, S., Wear, D. J., Shih, J. W. & Lo, S. C. Mycoplasmas and oncogenesis: persistent infection and multistage malignant transformation. *Proceedings of the National Academy of Sciences* **92**, 10197–10201 (1995).
17. Vande Voorde, J., Balzarini, J. & Liekens, S. Mycoplasmas and cancer: focus on nucleoside metabolism. *EXCLI Journal* **13**, 300–322 (2014).
18. Asangani, I. A. *et al.* Characterization of the EZH2-MMSET histone methyltransferase regulatory axis in cancer. *Molecular Cell* **49**, 80–93 (2013).
19. Popovic, R. *et al.* Histone Methyltransferase MMSET/NSD2 Alters EZH2 Binding and Reprograms the Myeloma Epigenome through Global and Focal Changes in H3K36 and H3K27 Methylation. **10**, e1004566 (2014).
20. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**, 2281–2308 (2013).
21. Basu, A. *et al.* An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. *Cell* **154**, 1151–1161 (2013).
22. Zeileis, A. & Grothendieck, G. zoo: S3 Infrastructure for Regular and Irregular Time Series. *Journal of Statistical Software* **14**, 1–27 (2005).
23. Reich, M. *et al.* GenePattern 2.0. *Nat Genet* **38**, 500–501 (2006).

## Chapter V

### Using small-molecule profiling to study biology of *Mycoplasma hyorhinitis*-infected human cell lines

#### Collaborators contributions:

- **Dr. Joshiawa Paulk** for his comments and insights on the experiments performed in this chapter.

## 5.1 Introduction

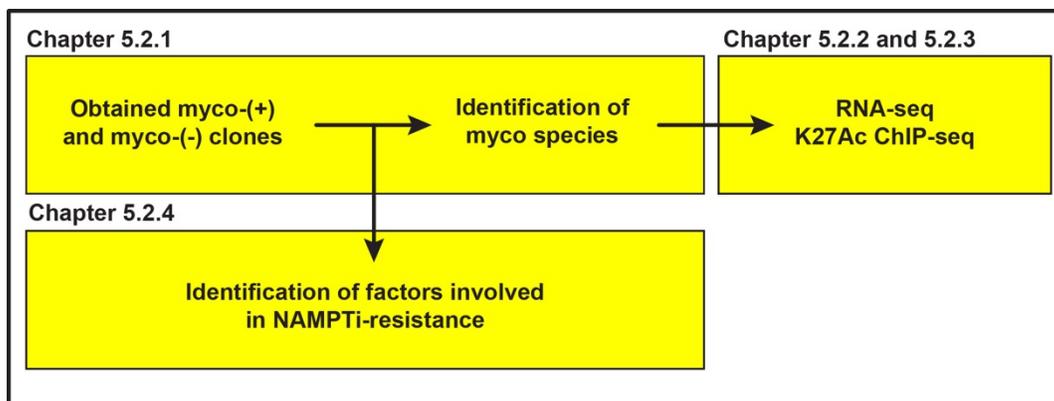
Mycoplasma is a genus of bacteria from the family *Mycoplasmataceae* characterized by the lack of a cell wall. There are about 150 species of Mycoplasma with a total of around 500 strains documented in NCIB's taxonomy database <sup>1</sup>. Mycoplasma have minimal genome and lack genes involving in amino acid and cofactor synthesis, hence they are often observed to have parasitic (and sometimes pathogenic) relationships with other complex organisms such as humans and related cell lines <sup>2</sup>. Several pathogenic mycoplasma have been identified to have cause diseases in human such as urethritis and other pelvic inflammatory conditions due to colonization of *M. genitalium* and *M. homominis* in the urogenital tract <sup>3</sup>. Due to its pathogenicity in humans, mycoplasma have also been isolated from tumor tissues obtained from cancer patients <sup>4</sup>. Studies on cancer *in vitro* have also observed oncogenic transformation of benign human cells upon mycoplasma infection connecting mycoplasma infection and cancer <sup>5,6,7</sup>.

The ability of mycoplasma to colonize and transform human cells have been a troubling issue confounding conclusions on observations involving the culture of these cells. In fact, mycoplasma contamination is a persisting problem in cell culture and it has been estimated that up to 35% of cultures in biopharmaceutical research are contaminated <sup>8</sup>. As mentioned, mycoplasma co-culture with human cell lines causes a significant transformation in the host, including inducing cytokine expression that could affects cellular proliferation and signal transduction <sup>3</sup>. It would be difficult to separate effects observed in these studies, intended perturbation of the cell, from those that were induced by mycoplasma infection of host cells.

I have unknowingly contaminated HEK293T cells with mycoplasma when scaling up clones from a CRISPR experiment. The contaminated cells were part of a small molecule profiling experiment described in Chapter 4, which exhibited drastic differences in their small molecule sensitivity. In addition, we have observed a difference in H3K27Ac levels in these myco-(+) and

myco(-) cell lines. We are the first to observe an epigenetic change with mycoplasma infection as well as a drastic difference in NAMPTi sensitivity. Further research into this phenotype could yield a cautionary tale of mycoplasma contamination in studies involving epigenetics or small molecule profiling. We can be hopeful that this study might also reveal insights into host-pathogen characteristics between mycoplasma and human cells that could be targeted as potential leads for anti-cancer therapeutics.

In this chapter I'll describe my strategy in studying this phenotype, including identifying the species of mycoplasma, RNA-seq and ChIP-seq analysis. Small molecule profiling and characterization of these cell lines were described in Chapter 4 and the results would not be repeated in this Chapter. Brief scheme of this chapter is shown below (Figure 5.1).



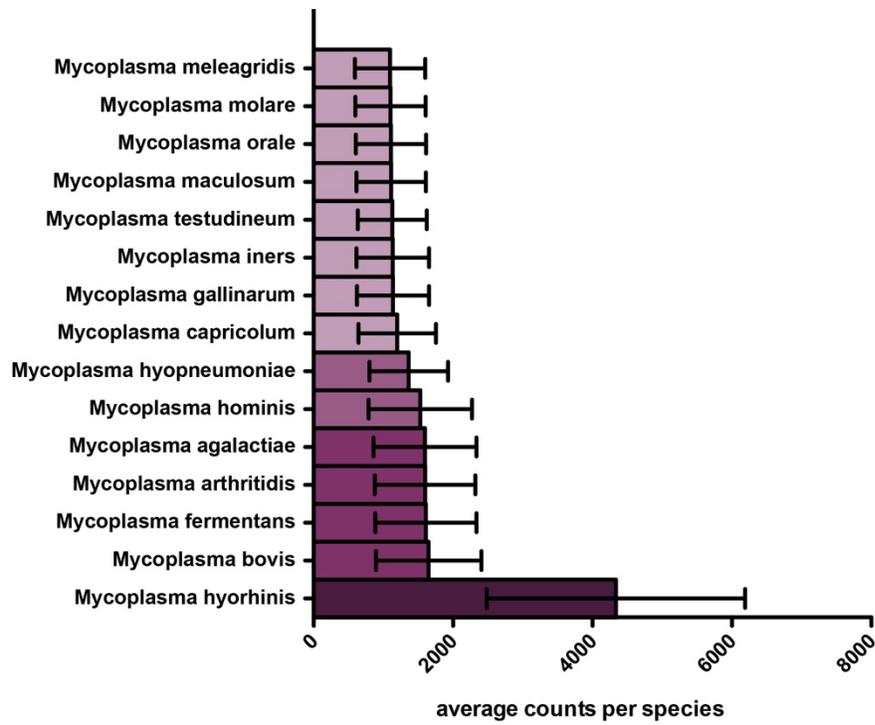
**Figure 5.1** Scheme of experiments described in this chapter.

## 5.2 Results and Discussion

### 5.2.1 Acquisition of infected clones and identification of mycoplasma species

As described in Chapter 4.2.4 and 4.2.5, we acquired a pair of myco-(+) and myco(-) HEK293Ts after culturing the clones over a period of time. We have also observed that the pair have vastly different sensitivities towards NAMPTi treatment and clofarabine treatment. We identified the species of mycoplasma responsible for the contamination using methods described in <sup>9</sup>, isolated single colonies of PCR

product and sequenced them. Sequencing results were processed through BLAST<sup>10</sup>. Top ranked match in BLAST were species *Mycoplasma Hyorihinis* (Figure 5.2.1).

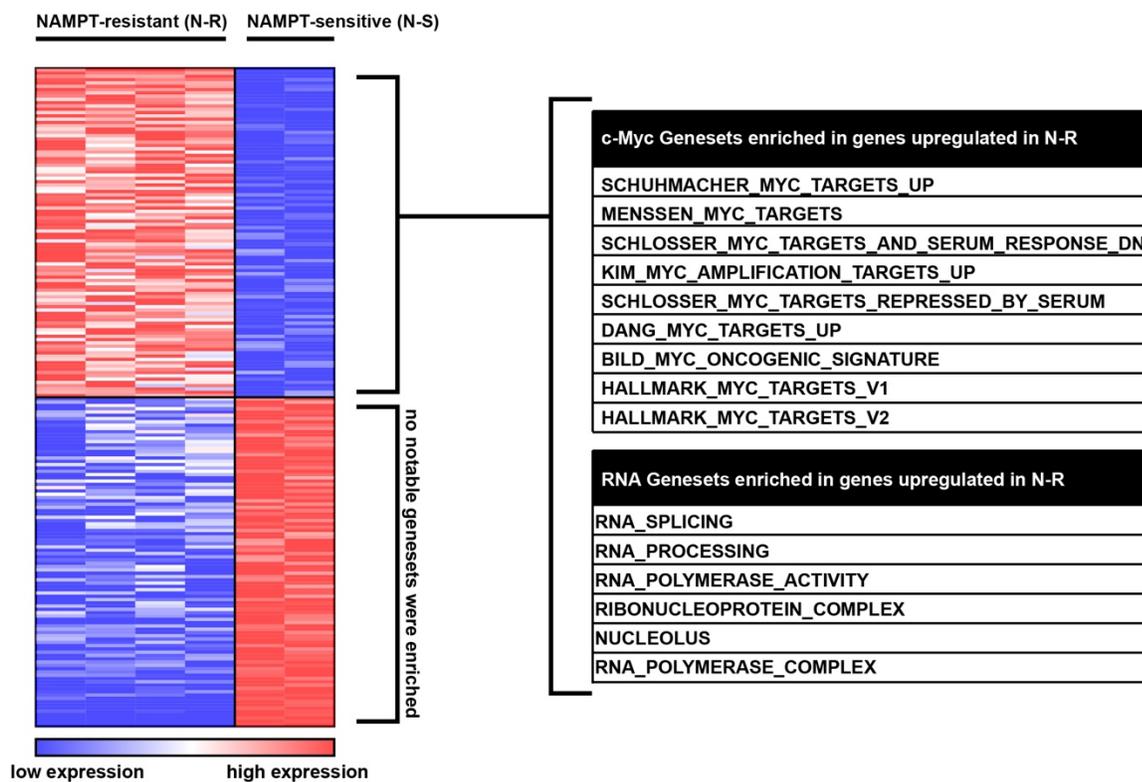


**Figure 5.2.1** Counts calculated using BLAST scores and averaged over 20 colonies that were collected and sequenced over 2 samples.

### 5.2.2 RNA-seq results

We used NAMPTi-resistance as a phenotype for myco-(+) cells and analyzed RNA-seq results of cell lines sequenced as described in Chapter 4. We performed comparative differential analysis on 2 groups of expression data: NAMPTi-resistant (B2,C1,A1-R,A2-R) and NAMPTi-sensitive (A1,A2) (Figure 5.2.2a) (Supplementary Table 5.1)<sup>11</sup>. The score generated by the analysis was ranked and used as input for GSEA<sup>12</sup>. Gene sets that relate to c-Myc function were significantly enriched in GSEA using the curated geneset (C2) module and gene sets that relate to RNA

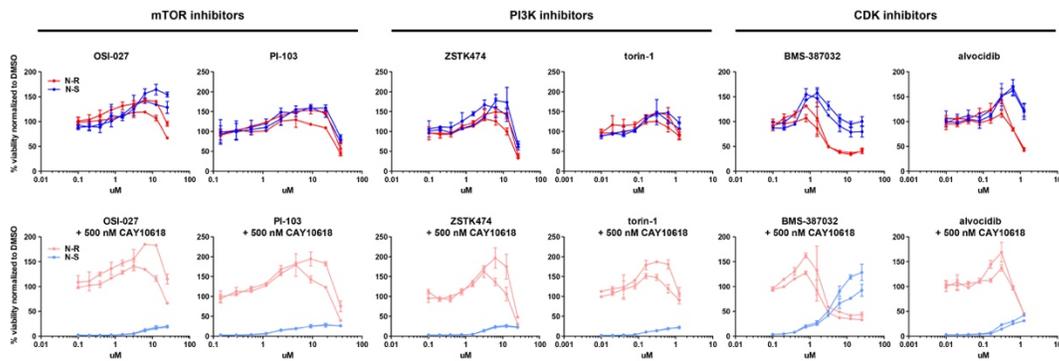
processing were significantly enriched when using the gene ontology (GO C5) module (Supplementary Table 5.2).



**Figure 5.2.2a.** Heatmap showing expression profiles of top 100 differential gene expression (upregulated and downregulated) in N-R and N-S samples. Gene sets that were significantly enriched by upregulated genes include gene sets regulated by Myc, as well as gene sets related to RNA processes in the cell.

We took the top 100 ranked upregulated genes and top 100 ranked downregulated genes for connectivity map (CMAP) analysis using L1000 characteristic direction signature search engine on the LINCS platform. The LINCS database includes gene-expression profiles obtained from small molecule cell line profiling. Analysis was done to identify compound-treated gene expression that most resembles my input. Compound treatments identified were highly enriched in compounds targeting the mTOR pathway and CDK (Supplementary Table 5.3).

NAMPTi-sensitive and NAMPTi-resistant clones (HEK/HEK-R and A2/A2-R) were pre-treated with a set of mTOR, PI3K and CDK inhibitors for 24 hours and treated with NAMPTi for an additional of 72 hours. m-TOR and PI3K inhibitors have small effect on NAMPTi-sensitive cell lines. CDK inhibitors, alvocidib and BMS-387032, had a dose-dependent response in activating NAMPTi-resistance in sensitive cell lines (Figure 5.2.2b). Our RNA-seq results concurred with previous studies made in the c-Myc and NAMPT field: linking c-Myc to NAMPT<sup>13</sup> and cdk being a target of c-Myc<sup>14,15</sup>.

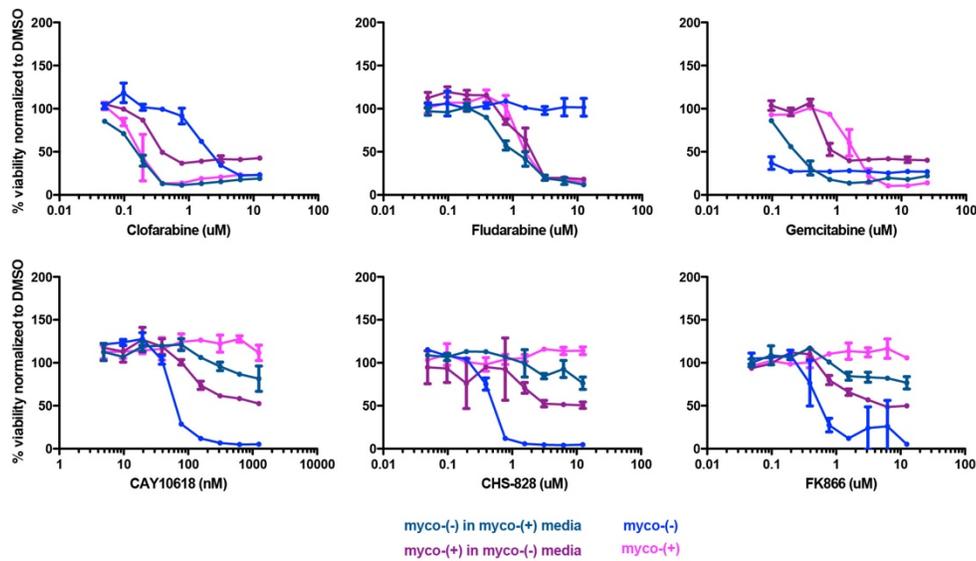


**Figure 5.2.1b Top:** DRPs of NAMPTi-resistant clones (N-R: HEK-R, A2-R) and NAMPTi-sensitive clones (N-S: HEK, A2) cell lines treated with mTOR, P1K3 and CDK inhibitors **Bottom:** Same set of cell lines pre-treated with m-TOR, P1K2 and CDK inhibitors for 24 hours and treated with 500 nM of NAMPTi (CAY10618) for another 72 hours. Markedly rescue of cell death can be observed in N-S cell lines pre-treated with CDK inhibitors.

### 5.2.3 Efforts in identification of factors involved in phenotype

We hypothesized that mycoplasma in myco-(+) cultures could be secreting factors or proteins that were interacting with the drugs to cause the NAMPTi-resistance and clofarabine sensitivity that we have observed. Myco-(+) and myco-(-) cells were cultured and their media collected. Mycoplasma was removed from myco-(+) media by high-speed centrifugation. Myco-(-) cells were then incubated with myco-(+) media (and vice versa), treat with NAMPTi, clofarabine, fludarabine and gemcitabine and their viability assessed after 72 hours. We observed that the phenotype of myco-(+) and myco-(-) cell lines switched upon incubation with the other's media,

with the exception of fludarabine treatment in myco(-) cells. This suggests that these cells that NAMPTi resistance and clofarabine sensitivity is reversible and dependent on extracellular environment despite the transformation within the cell with evidence of differing K27Ac levels and transcript levels between the myco-(+) and myco(-) cell lines. The phenotype was not observed to be as extreme as myco(-) cells or myco(+) cells in their own. This could be due to degradation of factors or protein in the media upon storage in the fridge.



**Figure 5.2.3** DRPs of myco-(+) and myco(-) cells incubated with the other’s media and treated with nucleoside analogues (top) and NAMPTi (bottom).

### 5.2.4 Concluding remarks

Study of mycoplasma colonization in human cell lines are often shunned upon due to the reputation of mycoplasma as a cell culture contaminant. The opportunity of exploring mycoplasma-infected cells and their effect on the cells’ sensitivity towards certain treatment was serendipitously offered when our long term culture became contaminated, leading to extreme phenotypes. As we have already made great strides in characterizing the phenotype and ran a RNA-seq on these cell

lines, we decided to embrace the contamination and took the opportunity to study a topic that to us, might have immense impact on cancer biology and treatment. Despite being a taboo, mycoplasma are often found to be colonizing tumors in patient samples, their numbers correlating with the severity of the disease <sup>4</sup>.

We found that mycoplasma infected cell lines are highly sensitive to clofarabine and its analogue, fludarabine, but not towards other nucleoside analogues such as gemcitabine. This observation concurred with literature precedence where other groups have observed a change in cellular nucleoside metabolism upon mycoplasma infection <sup>16-18</sup> The difference in sensitivity between different nucleoside analogues could be due to mycoplasma effecting only specific pathways of nucleoside metabolism in the cell. In addition, mycoplasma's effect in the cell's redox pathway involving NAD<sup>+</sup> have not been published, this finding opens up a new field of understanding the relationship between mycoplasma and its host.

Moving forward, tools described in this chapter could be used to further explore the effects of mycoplasma infectivity in human cells, especially the change in K27Ac. As the difference in small molecule sensitivity is striking and reversible, these findings introduce a novel way of targeting cancer that not currently being pursued and could disrupt the field of cancer therapy.

### **5.3 Experimental methods**

#### **Identification of mycoplasma species**

Media from culture was collected and spun at 16,000 x g for 5 minutes. Pellet was washed with PBS once. Genomic DNA was extracted using Wizard® Genomic DNA Purification Kit (Promega). PCR was performed using primers described in <sup>6</sup> and PrimeSTAR® GXL DNA Polymerase (Takara Bio USA). PCR product was cloned into pMiniT 2.0 vector using NEB® PCR cloning kit and plated on Ampicillin LB plate. 10 colonies were picked per PCR reaction for sequencing.

Sequence results from PCR cloning were analyzed using BLAST<sup>10</sup> with 16S microbial sequence library obtained from NCIB BLAST FTP site<sup>19</sup>.

### **Cell lines**

Cell lines were cultured as described in Chapter 4.

### **RNA-seq analysis**

Analysis was performed in Cufflinks suite as described in Chapter 4. Counts obtained from Cufflinks were input into comparative marker selection module<sup>11</sup> in GenePattern<sup>20</sup> to rank differential transcripts in N-R clones compared to N-S clones. The ranked list was then used as input for GSEA<sup>12</sup> using the C2 (curated gene sets) and C5 (gene ontology) collection for the analysis. The top 100 upregulated and downregulated transcripts were used as input in LINC analysis (<http://amp.pharm.mssm.edu/L1000CDS2/#/index>).

### **Small molecule pre-treatment**

For retest of compounds with pre-treatment, cells were plated at 2500 per well in 384-well plates (40 uL) in duplicates. The next day, compounds for pre-treatment were added by pin transfer (100 nL in DMSO) (Cybio Vario). After a 24-hour incubation, media was aspirated from the wells and the wells was washed once with 40 uL of PBS. After aspirating the PBS wash, 40 uL of media with DMSO or CAY10618 was added into the wells. Plates treated with CAY10618 have wells column 1,2,23 and 24 treated with DMSO instead to provide a value for normalization in future analysis. The cells were then incubated for another 72 hours and viability was assessed as described in Chapter 4.

Compounds were purchased from Cayman Chemical.

## Media exchange of cell lines

Media was collected from myco-(+) and myco-(-) cell lines over multiple passages, spun at 16,000 x g for 5 minutes to remove mycoplasma and stored at 4C. Cells to be treated were washed once with PBS and counted to be seeded at 2500 cells per 40 uL well. Cells were pelleted, resuspended in appropriate media and seeded into 384 well plates. Treatment of cells were as described in Chapter 4.

## 5.4 References

1. NCIB Taxonomy Browser Mycoplasma. Available at: <https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=2093>. (Accessed: 6 January 2017)
2. Waites, K. B. & Xiao, L. in *Molecular Medical Microbiology* 1587–1609 (Elsevier, 2015). doi:10.1016/B978-0-12-397169-2.00089-5
3. Razin, S., Yogev, D. & Naot, Y. Molecular Biology and Pathogenicity of Mycoplasmas. *Microbiol. Mol. Biol. Rev.* **62**, 1094–1156 (1998).
4. Vande Voorde, J., Balzarini, J. & Liekens, S. Mycoplasmas and cancer: focus on nucleoside metabolism. *EXCLI Journal* **13**, 300–322 (2014).
5. Namiki, K. *et al.* Persistent Exposure to Mycoplasma Induces Malignant Transformation of Human Prostate Cells. *PLoS One* **4**, e6872 (2009).
6. Feng, S. H., Tsai, S., Rodriguez, J. & Lo, S. C. Mycoplasmal infections prevent apoptosis and induce malignant transformation of interleukin-3-dependent 32D hematopoietic cells. *Mol Cell Biol* **19**, 7995–8002 (1999).
7. Tsai, S., Wear, D. J., Shih, J. W. & Lo, S. C. Mycoplasmas and oncogenesis: persistent infection and multistage malignant transformation. *Proceedings of the National Academy of Sciences* **92**, 10197–10201 (1995).

8. Armstrong, S. E., Mariano, J. A. & Lundin, D. J. The scope of mycoplasma contamination within the biopharmaceutical industry. *Biologicals* **38**, 211–213 (2010).
9. Uphoff, C. C. & Drexler, H. G. Detecting mycoplasma contamination in cell cultures by polymerase chain reaction. *Methods Mol. Biol.* **731**, 93–103 (2011).
10. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403–410 (1990).
11. Gould, J., Getz, G., Monti, S., Reich, M. & Mesirov, J. P. Comparative gene marker selection suite. *Bioinformatics* **22**, 1924–1925 (2006).
12. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545–15550 (2005).
13. Menssen, A. *et al.* The c-MYC oncoprotein, the NAMPT enzyme, the SIRT1-inhibitor DBC1, and the SIRT1 deacetylase form a positive feedback loop. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E187–96 (2012).
14. Obaya, A. J., Mateyak, M. K. & Sedivy, J. M. Mysterious liaisons: the relationship between c-Myc and the cell cycle. , *Published online: 13 May 1999; | doi:10.1038/sj.onc.1202749* **18**, 2934–2941 (1999).
15. Obaya, A. J., Kotenko, I., Cole, M. D. & Sedivy, J. M. The proto-oncogene c-myc acts through the cyclin-dependent kinase (Cdk) inhibitor p27(Kip1) to facilitate the activation of Cdk4/6 and early G(1) phase progression. *J Biol Chem* **277**, 31263–31269 (2002).
16. Bronckaers, A., Balzarini, J. & Liekens, S. The cytostatic activity of pyrimidine nucleosides is strongly modulated by Mycoplasma hyorhinitis infection: Implications for cancer therapy. *Biochemical Pharmacology* **76**, 188–197 (2008).
17. Voorde, J. V., Gago, F., Vrancken, K., Liekens, S. & Balzarini, J. Characterization of

- pyrimidine nucleoside phosphorylase of *Mycoplasma hyorhinis*: implications for the clinical efficacy of nucleoside analogues. *Biochemical Journal* **445**, 113–123 (2012).
18. JETTÉ, L., BISSOON-HAQQANI, S., Le François, B., MAROUN, J. A. & BIRNBOIM, H. C. Resistance of colorectal cancer cells to 5-FUdR and 5-FU caused by *Mycoplasma* infection. *Anticancer Res* **28**, 2175–2180 (2008).
  19. Tao, T., Madden, T., Christiam, C. & Szilagyi, L. BLAST FTP Site. (2016).
  20. Reich, M. *et al.* GenePattern 2.0. *Nat Genet* **38**, 500–501 (2006).

## Supplemental Information

Supplemental Table 2.1

Top 100 genes upregulated in E1099K-ALLs			
Feature	Score	Feature P	Fold Change
TRIM58	7.07	2.00E-04	1.53
CRISPLD1	6.40	2.00E-04	1.67
ZC2HC1A	5.87	1.40E-03	1.35
GLT8D2	5.42	4.00E-04	1.06
MAGI2	5.33	2.80E-03	1.26
CREB3L4	5.28	6.00E-04	1.11
LOC100506652	5.21	4.00E-04	1.05
ACOXL	5.10	1.60E-03	1.12
LOC100505490	4.78	3.60E-03	1.08
VPS72	4.67	2.40E-03	1.03
IL3	4.61	5.20E-03	1.04
MYO16	4.60	2.40E-03	1.05
SGCB	4.52	8.80E-03	1.27
WDFY3	4.49	1.82E-02	1.25
TBX3	4.44	4.60E-03	1.03
CNTD1	4.39	3.40E-03	1.06
TSC22D3	4.33	1.80E-03	1.33
POLR3C	4.21	1.00E-03	1.05
KDM6A	4.19	2.20E-03	1.12
TMEM64	4.18	3.80E-03	1.25
CD44	4.14	1.22E-02	1.33
CXorf57	4.14	3.00E-02	1.47
SOCS6	4.11	1.22E-02	1.27
METTL20	4.09	1.18E-02	1.05
TTC9C	3.98	1.18E-02	1.04
UBQLN1	3.96	3.20E-03	1.04
TXLNG	3.88	4.20E-03	1.05
DERL1	3.78	4.00E-03	1.06
ATP5G2	3.77	1.14E-02	1.03
MORC4	3.75	7.60E-03	1.38
ADCK3	3.71	5.40E-03	1.07
RAB10	3.68	1.10E-02	1.03
MMP12	3.66	8.00E-03	1.06
KDELC2	3.64	2.70E-02	1.35
DNMT3A	3.64	2.40E-03	1.10
ASPH	3.62	1.30E-02	1.21
FOXE3	3.61	1.34E-02	1.04
HOXA5	3.59	9.80E-03	1.43
CD93	3.58	3.40E-03	1.35
ZBTB3	3.55	5.80E-03	1.08
ISCA1	3.49	1.06E-02	1.05
LOC100507291	3.49	1.20E-02	1.06
FAM209B	3.46	1.48E-02	1.08
C1RL	3.44	1.28E-02	1.07
MRPS30	3.43	1.20E-02	1.04
CNKSR3	3.39	9.40E-03	1.26
MAD2L1BP	3.38	1.00E-03	1.04
MYH7	3.38	1.28E-02	1.07
OAZ1	3.36	1.80E-03	1.01
MBTPS2	3.35	1.80E-03	1.07
CDH3	3.35	2.96E-02	1.06

Top 100 genes downregulated in E1099K-ALLs			
Feature	Score	Feature P	Fold Change
CLEC16A	-7.64	8.00E-04	1.09
GJA3	-5.83	1.60E-03	1.13
DKFZP586I1420	-5.42	6.00E-04	1.12
KIAA0226L	-5.35	5.20E-03	1.83
PDXK	-5.15	2.60E-03	1.08
TEX29	-5.14	1.80E-03	1.11
GDPD5	-4.80	2.20E-03	1.10
FAM193B	-4.66	1.00E-03	1.13
C15orf5	-4.49	3.20E-03	1.19
STK36	-4.40	5.00E-03	1.11
ASIC1	-4.38	1.10E-02	1.18
ANKS3	-4.34	1.80E-03	1.09
UGP2	-4.33	7.00E-03	1.05
INTS1	-4.32	3.20E-03	1.08
AGAP3	-4.20	1.50E-02	1.14
TTC13	-4.19	5.00E-03	1.09
HRK	-4.16	3.20E-03	1.35
AMELY	-4.06	8.80E-03	1.06
ADM	-4.05	2.32E-02	1.40
ZNF556	-4.05	3.20E-03	1.07
ATG2B	-3.94	2.60E-03	1.05
PPP1R37	-3.94	8.20E-03	1.07
CYP2S1	-3.92	8.40E-03	1.08
ZDHHC6	-3.92	8.80E-03	1.03
ACOT12	-3.91	5.80E-03	1.06
LYSMD4	-3.90	5.20E-03	1.08
PROX2	-3.90	7.20E-03	1.06
SYNJ1	-3.90	7.00E-03	1.16
LOC100507506	-3.88	1.02E-02	1.06
OSBPL7	-3.88	4.80E-03	1.07
RGCC	-3.86	3.22E-02	1.65
USP15	-3.80	4.60E-03	1.07
RAD21L1	-3.77	4.00E-03	1.05
PRPF38B	-3.76	1.70E-02	1.05
GABRA6	-3.74	1.10E-02	1.04
IGSF9B	-3.74	1.78E-02	1.08
FLJ40288	-3.71	7.00E-03	1.07
SREK1	-3.68	1.38E-02	1.06
GNG2	-3.67	9.20E-03	1.21
AAGAB	-3.67	9.00E-03	1.06
ANO8	-3.65	3.00E-03	1.08
SERPINB7	-3.58	7.60E-03	1.05
WDFY2	-3.56	1.12E-02	1.10
SARDH	-3.55	5.40E-03	1.08
COG3	-3.54	2.60E-03	1.10
TNFRSF10A	-3.54	9.40E-03	1.12
ABCG2	-3.54	1.00E-02	1.29
TNNC1	-3.52	1.16E-02	1.09
LOC145474	-3.51	6.40E-03	1.19
CHRM2	-3.51	2.14E-02	1.04
WFDC8	-3.49	9.20E-03	1.06

Feature	Score	Feature P	Fold Change
SLC25A13	3.34	2.84E-02	1.29
MOS	3.33	1.04E-02	1.04
RIPK2	3.32	4.44E-02	1.23
ZNF792	3.30	8.60E-03	1.16
JAKMIP2	3.29	4.32E-02	1.27
SNHG6	3.28	3.52E-02	1.02
LOC494558	3.28	7.20E-03	1.05
RNF139	3.28	7.80E-03	1.03
C17orf79	3.27	1.92E-02	1.06
MYF5	3.27	5.20E-03	1.03
KIAA0368	3.27	3.74E-02	1.23
LOC100506342	3.24	1.44E-02	1.10
C9orf23	3.24	3.00E-03	1.06
APH1A	3.23	3.52E-02	1.04
IFNGR2	3.18	1.62E-02	1.15
LOC201477	3.17	1.44E-02	1.04
QDPR	3.17	1.90E-02	1.10
PRDM11	3.16	6.40E-03	1.05
C8orf82	3.15	1.06E-02	1.21
RGS12	3.15	3.44E-02	1.08
HAND1	3.15	1.82E-02	1.05
KRT84	3.14	9.00E-03	1.07
B4GALT4	3.12	1.04E-02	1.07
CD55	3.12	5.40E-02	1.23
LOC100506190	3.11	1.86E-02	1.08
UBE2B	3.10	5.00E-03	1.04
SEC14L3	3.09	1.10E-02	1.03
RAB3C	3.08	5.00E-03	1.09
MLK7-AS1	3.07	1.32E-02	1.08
LOC100505828	3.07	1.80E-02	1.06
KCTD12	3.07	1.44E-02	1.38
SLC25A6	3.07	2.42E-02	1.04
CLEC3A	3.06	8.60E-03	1.08
CLMN	3.05	1.30E-02	1.17
SAP30	3.05	1.36E-02	1.11
EEF1D	3.05	1.04E-02	1.02
ASXL3	3.05	2.00E-03	1.13
DCLK2	3.04	5.80E-03	1.16
CLDN6	3.04	4.80E-03	1.05
DAZAP1	3.03	3.98E-02	1.02
TAGLN2	3.03	1.92E-02	1.10
FLJ41350	3.02	2.52E-02	1.05
ADNP	3.02	2.90E-02	1.03
RNF135	3.02	8.82E-02	1.24
RAP2B	3.02	4.40E-03	1.14
HAUS4	3.02	2.58E-02	1.05
MFSD6L	3.01	1.06E-02	1.03
CRNN	3.00	2.16E-02	1.05
NKX2-1-AS1	3.00	7.00E-03	1.06

Feature	Score	Feature P	Fold Change
NOD1	-3.48	1.12E-02	1.12
PCNXL3	-3.44	5.00E-03	1.07
CSF2	-3.43	1.16E-02	1.05
ADAT2	-3.41	2.80E-03	1.06
BAZ2B	-3.41	1.94E-02	1.20
USF1	-3.40	2.08E-02	1.08
SYT5	-3.39	1.20E-02	1.04
ZNF346	-3.38	3.60E-03	1.06
KIRREL2	-3.35	3.56E-02	1.23
DAOA	-3.35	2.02E-02	1.03
PIP5K1C	-3.35	1.20E-02	1.07
ZNF540	-3.34	1.50E-02	1.12
UAP1L1	-3.34	3.00E-02	1.17
IZUMO1	-3.34	3.08E-02	1.05
IQCF2	-3.33	1.52E-02	1.05
LOC283177	-3.31	3.08E-02	1.13
GABRR3	-3.31	1.16E-02	1.05
FAM214A	-3.31	3.70E-02	1.09
PCDHA9	-3.31	4.42E-02	1.04
COX4I2	-3.29	1.02E-02	1.04
LECT2	-3.27	1.10E-02	1.06
BRD9	-3.26	1.26E-02	1.05
NPDC1	-3.26	2.46E-02	1.21
FGF11	-3.25	4.00E-03	1.06
LOC100507520	-3.24	3.78E-02	1.11
STK17A	-3.24	1.80E-02	1.17
ODZ1	-3.24	5.24E-02	1.15
CD69	-3.23	7.40E-03	1.56
BECN1	-3.21	1.96E-02	1.03
AHCYL2	-3.20	4.20E-03	1.10
LOC100133039	-3.20	2.10E-02	1.14
KIAA0415	-3.19	5.50E-02	1.08
IGIP	-3.18	2.68E-02	1.16
LOC100506758	-3.18	1.66E-02	1.08
LOC100189589	-3.17	1.90E-02	1.04
NOXA1	-3.16	1.48E-02	1.08
KCNAB3	-3.16	3.80E-02	1.08
CDK12	-3.16	1.58E-02	1.04
LDHAL6B	-3.15	3.26E-02	1.04
DDX26B	-3.15	1.80E-02	1.10
AGAP11	-3.15	4.10E-02	1.04
METRNL	-3.14	2.30E-02	1.16
PLCH1	-3.13	2.14E-02	1.24
FAM40B	-3.13	8.00E-03	1.16
S1PR3	-3.12	7.82E-02	1.22
FAM161B	-3.12	2.38E-02	1.12
EPS8L2	-3.11	1.62E-02	1.10
PEAK1	-3.11	1.52E-02	1.09
TNIK	-3.10	8.42E-02	1.22

**Supplemental Table 2.1** Differential expression between E1099K-ALLs and WT-ALLs using CCLE transcripts as input in the comparative marker selection suite in genepattern. Upregulated genes in E1099K-ALLs are ranked positively while downregulated genes in E1099K-ALLs are ranked negatively.

## Supplemental Table 2.2

### C2 gene sets

NAME	ES	NOM p-val	FDR q-val
JAATINEN_HEMATOPOIETIC_STEM_CELL_UP	0.42	0	2.36E-02
REACTOME_3_UTR_MEDIATED_TRANSLATIONAL_REGULATION	0.55	0	1.56E-02
ZHAN_MULTIPLE_MYELOMA_MS_UP	0.55	0	2.30E-02
REACTOME_PEPTIDE_CHAIN_ELONGATION	0.57	0	2.85E-02
BILANGES_RAPAMYCIN_SENSITIVE_VIA_TSC1_AND_TSC2	0.49	0	3.06E-02
FARMER_BREAST_CANCER_CLUSTER_2	0.59	0	2.88E-02
SCHLOSSER_MYC_TARGETS_REPRESSED_BY_SERUM	0.42	0	3.28E-02
BIDUS_METASTASIS_UP	0.41	0	3.06E-02
BHATTACHARYA_EMBRYONIC_STEM_CELL	0.47	0	3.40E-02
GRAHAM_CML_QUIESCENT_VS_NORMAL_QUIESCENT_UP	0.48	0	3.10E-02
BOYALT_LIVER_CANCER_SUBCLASS_G3_UP	0.40	0	4.41E-02
KEGG_RIBOSOME	0.53	0	4.30E-02
ALCALAY_AML_BY_NPM1_LOCALIZATION_UP	0.42	0	6.00E-02
AKL_HTLV1_INFECTION_UP	0.59	0	7.37E-02
NIKOLSKY_BREAST_CANCER_8Q12_Q22_AMPLICON	0.41	0	8.64E-02
ZHAN_VARIABLE_EARLY_DIFFERENTIATION_GENES_DN	0.54	2.55E-03	8.63E-02
SMITH_LIVER_CANCER	0.50	0	8.74E-02
GENTILE_UV_RESPONSE_CLUSTER_D6	0.51	0	9.59E-02
SEIDEN_MET_SIGNALING	0.60	0	1.23E-01
FOURNIER_ACINAR_DEVELOPMENT_LATE_DN	0.58	4.47E-03	1.41E-01
ZHENG_RESPONSE_TO_ARSENITE_DN	0.60	0	1.51E-01
CERVERA_SDHB_TARGETS_1_DN	0.50	0	1.44E-01
REACTOME_NONSENSE_MEDIATED_DECAY_ENHANCED_BY_THE_EXON_JUNCTION_COMPLEX	0.44	2.79E-03	1.70E-01
LE_SKI_TARGETS_UP	0.60	0	1.85E-01
REACTOME_CHOLESTEROL_BIOSYNTHESIS	0.57	0	1.80E-01

## C5 gene sets

NAME	ES	NOM p-val	FDR q-val
ACTIN_FILAMENT_BINDING	0.55	2.00E-03	7.35E-01
STRUCTURAL_CONSTITUENT_OF_RIBOSOME	0.48	0	5.18E-01
RRNA_METABOLIC_PROCESS	0.60	7.00E-03	5.59E-01
RIBOSOME_BIOGENESIS_AND_ASSEMBLY	0.58	1.20E-02	4.93E-01
RIBOSOME	0.48	0	4.00E-01
NITROGEN_COMPOUND_BIOSYNTHETIC_PROCESS	0.52	0	3.51E-01
AMINE_BIOSYNTHETIC_PROCESS	0.59	1.60E-02	3.24E-01
RRNA_PROCESSING	0.59	5.00E-03	3.40E-01
HYDROLASE_ACTIVITY_ACTING_ON_CARBON_NITROGEN NOT_PEPTIDEBONDSIN_CYCLIC_AMIDINES	0.58	1.20E-02	3.04E-01
RIBOSOMAL_SUBUNIT	0.54	1.40E-02	3.12E-01
DRUG_BINDING	0.57	1.70E-02	3.14E-01
MITOCHONDRIAL_RIBOSOME	0.52	5.00E-03	3.11E-01
REGULATION_OF_MITOTIC_CELL_CYCLE	0.50	1.20E-02	3.31E-01
TRANSFERASE_ACTIVITY_TRANSFERRING_PENTOSYL_GROUPS	0.52	2.90E-02	3.41E-01
ORGANELLAR_RIBOSOME	0.52	1.60E-02	3.23E-01
NUCLEAR_PORE	0.47	1.40E-02	3.34E-01
EMBRYONIC_DEVELOPMENT	0.40	5.00E-03	4.03E-01
RIBONUCLEOPROTEIN_COMPLEX	0.33	3.00E-03	4.17E-01
METALLOENDOPEPTIDASE_ACTIVITY	0.46	3.30E-02	3.98E-01
PROTEASOME_COMPLEX	0.50	3.70E-02	3.86E-01
NUCLEAR_BODY	0.44	1.70E-02	3.87E-01
PORE_COMPLEX	0.43	2.10E-02	3.78E-01
MITOCHONDRIAL_MEMBRANE_PART	0.39	1.50E-02	3.92E-01
NUCLEAR_MEMBRANE_PART	0.41	2.50E-02	4.67E-01
GAMETE_GENERATION	0.34	6.00E-03	5.17E-01

## H gene sets

NAME	ES	NOM p-val	FDR q-val
HALLMARK_MYC_TARGETS_V1	0.36	0	2.11E-02
HALLMARK_OXIDATIVE_PHOSPHORYLATION	0.35	0	1.06E-02
HALLMARK_UV_RESPONSE_DN	0.30	6.73E-03	1.01E-01
HALLMARK_G2M_CHECKPOINT	0.28	3.39E-03	9.93E-02
HALLMARK_E2F_TARGETS	0.28	6.62E-03	9.34E-02
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	0.25	4.61E-02	3.27E-01
HALLMARK_MYC_TARGETS_V2	0.30	1.90E-01	4.15E-01
HALLMARK_WNT_BETA_CATENIN_SIGNALING	0.32	2.31E-01	3.85E-01
HALLMARK_CHOLESTEROL_HOMEOSTASIS	0.27	1.86E-01	3.95E-01
HALLMARK_SPERMATOGENESIS	0.24	1.68E-01	4.32E-01
HALLMARK_ANDROGEN_RESPONSE	0.25	2.39E-01	4.57E-01
NAME	ES	NOM p-val	FDR q-val
HALLMARK_ADIPOGENESIS	0.22	2.73E-01	5.50E-01
HALLMARK_DNA_REPAIR	0.22	2.46E-01	5.13E-01
HALLMARK_UV_RESPONSE_UP	0.20	4.87E-01	8.13E-01
HALLMARK_COMPLEMENT	0.19	5.80E-01	8.68E-01
HALLMARK_HEME_METABOLISM	0.18	7.50E-01	1.00E+00
HALLMARK_PROTEIN_SECRETION	0.20	6.85E-01	9.86E-01
HALLMARK_HEDGEHOG_SIGNALING	0.25	6.50E-01	9.50E-01
HALLMARK_ANGIOGENESIS	0.24	7.06E-01	9.59E-01
HALLMARK_REACTIVE_OXIGEN_SPECIES_PATHWAY	0.22	7.52E-01	9.83E-01
HALLMARK_MTORC1_SIGNALING	0.16	9.58E-01	1.00E+00
HALLMARK_NOTCH_SIGNALING	0.21	8.69E-01	1.00E+00
HALLMARK_MITOTIC_SPINDLE	0.14	1.00E+00	9.82E-01

**Supplemental Table 2.2** Top GSEA gene sets enriched in upregulated E1099K-ALLs.

Supplemental Table 3.1

Cell line	Score	Cell line	Score	Cell line	Score
KMS11	2611.98	MOLT16	-2110.34	SUPHD1	-3803.87
HDMYZ	719.31	ME1	-2125.15	EM2	-3832.43
SKM1	666.82	DEL	-2130.35	JEKO1	-3837.19
KMS34	605.65	697	-2161.50	AMO1	-3932.19
JURLMK1	462.41	KYO1	-2221.14	MHHCALL4	-3957.40
OPM2	393.20	KMS20	-2262.64	BV173	-3977.79
HEL	321.50	OCIAML3	-2391.06	OCILY19	-3994.01
RPMI8402	269.74	KO52	-2396.55	JJN3	-4079.46
JURKAT	222.42	SUDHL1	-2438.98	L363	-4122.40
THP1	99.02	REC1	-2478.05	PL21	-4145.61
KU812	95.97	KARPAS620	-2492.28	JM1	-4159.01
LP1	-58.17	OCIAML2	-2540.04	P31FUJ	-4249.66
RCHACV	-148.25	SIGM5	-2560.20	KIJK	-4294.84
SET2	-155.24	KMS21BM	-2598.29	TALL1	-4462.79
HEL9217	-167.97	RPMI8226	-2600.80	KMS12BM	-4512.43
KMS26	-315.25	L428	-2681.19	KARPAS299	-4533.73
LAMA84	-362.72	KE97	-2731.61	SUDHL6	-4561.02
MOLM16	-476.13	U937	-2740.09	NUDHL1	-4640.90
KE37	-494.47	MHHCALL3	-2781.17	RI1	-4710.87
F36P	-636.69	SKMM2	-2808.41	MOLP2	-4714.41
NCIH929	-805.30	EOL1	-2814.40	SUDHL8	-4772.00
PF382	-838.38	P12ICHIKAWA	-2851.98	ST486	-4807.38
NCO2	-957.41	RS411	-2871.48	L540	-4833.58
OCIM1	-985.81	KMH2	-2953.69	NUDUL1	-4835.54
HPBALL	-1037.26	SUPT1	-3008.75	BL70	-4915.66
MM1S	-1072.98	SUPT11	-3074.72	KMS27	-4938.95
KASUMI2	-1097.65	NALM6	-3079.56	HS611T	-4952.63
CMK	-1211.56	EJM	-3102.28	CA46	-4968.37
KMS18	-1313.66	AML193	-3110.43	HDM2	-4981.53
JK1	-1337.12	EHEB	-3164.36	GA10	-5016.99
MOLP8	-1440.74	GRANTA519	-3180.44	MOTN1	-5023.88
M07E	-1454.59	NB4	-3187.16	KARPAS422	-5123.26
SEM	-1513.48	U266B1	-3246.74	MINO	-5151.89
KMM1	-1609.87	MOLM13	-3268.66	DAUDI	-5236.31
TF1	-1805.02	MV411	-3318.81	RAJI	-5241.15
PEER	-1872.50	DND41	-3372.72	DB	-5243.44
MOLM6	-1918.20	SUPB15	-3435.99	SUDHL10	-5325.44
KASUMI1	-1955.84	JVM3	-3459.85	CI1	-5393.92
OCIAML5	-1962.66	SUPM2	-3474.38	EB1	-5400.52
GDM1	-2019.84	SR786	-3477.62	HT	-5431.40
A3KAW	-2021.87	HL60	-3583.17	MC116	-5457.53
KCL22	-2031.52	OCILY3	-3599.68	OCILY10	-5516.75
MONOMAC1	-2039.84	A4FUK	-3643.02	WSUDLCL2	-5521.61
REH	-2080.11	HUT78	-3682.10	SUDHL4	-5543.35
CMLT1	-2099.49	MEC1	-3740.94	RL	-5591.84
K562	-2107.01	KHM1B	-3785.09	NAMALWA	-5606.85
		PFEIFFER	-3802.97	P3HR1	-5733.44

Supplemental Table 3.1 ssGSEA performed on expression data of hematopoietic CCLs, scoring their expression using the geneset signature ZHAN\_MULTIPLE\_MYELOMA\_MS\_UP.

Supplemental Table 3.2

Top 50 anti-correlated		Top 50 correlated	
Compound	Pearson's r	Compound	Pearson's r
BMS-536924	-0.371	YM-155	0.414
AZD4547	-0.356	BYERS	0.322
BRD-K01121114	-0.307	ciclopirox olamine	0.303
KHS101	-0.299	TAUBE	0.292
BRD-M46138460	-0.287	AZD6482	0.289
SP600125	-0.279	BRD-K51490254	0.288
BMS-754807	-0.265	obatoclax	0.256
BRD-K50799972	-0.262	bendamustine	0.254
ruxolitinib	-0.255	SNS-032	0.252
JQ-1	-0.241	bexarotene	0.244
NVP-BSK805	-0.240	GROGER	0.244
NVP-231	-0.234	parthenolide	0.235
BRD-M98616540	-0.234	cisplatin	0.227
nilotinib	-0.234	sirolimus	0.216
BRD-K22828899	-0.232	BRD-M02488208	0.211
BRD-M09269305	-0.228	TG100-115	0.209
itraconazole	-0.225	TGX-221	0.205
triptolide	-0.225	BRD-M37545453	0.195
TG-101348	-0.222	AA-COCF3	0.191
BRD-M39644375	-0.221	NVP-LDE225	0.191
axitinib	-0.219	necrostatin-7	0.19
trifluoperazine	-0.217	SRT-1720	0.186
linsitinib	-0.216	BIX-01294	0.181
nintedanib	-0.210	MK-2206	0.17
RITA	-0.210	CR-1-31B	0.164
EU-5346	-0.204	CAL-101	0.162
selumetinib	-0.203	YL54	0.159
VER-155008	-0.203	PIK-93	0.158
AZD7545	-0.196	NSC632839	0.155
BRD4132	-0.196	ASN-05257430	0.154
AZ-3146	-0.194	temozolomide	0.153
compound 1B	-0.191	topotecan	0.152
crizotinib	-0.191	vincristine	0.151
BRD-M94343540	-0.190	GANT-61	0.15
bosutinib	-0.189	PYR-41	0.141
tacrolimus	-0.185	BRD-M68535767	0.137
SCH-79797	-0.183	NSC23766	0.136
peranib-2	-0.180	necrostatin-1	0.134
PD318088	-0.179	PAC-1	0.133
BRD-M45185124	-0.179	PI-103	0.13
BRD-M65857522	-0.178	GDC-0941	0.13
NVP-ADW742	-0.178	BRD-K64610608	0.129
MLN-4924	-0.177	canertinib	0.129
pifithrin-mu	-0.177	EGCG	0.126
MGCD-265	-0.177	BRD-M48845972	0.125
BRD-K80183349	-0.175	curcubitacin I	0.124
BRD-K66532283	-0.174	MST-312	0.122
foretinib	-0.173	EX-527	0.12
JW-480	-0.173	BRD-M60574774	0.12
BRD-M86771886	-0.173	OSI-027	0.119

**Supplemental Table 3.2** List of top 50 compounds that are anti-correlated and correlated to  $t(4;14)+$  enrichment score of hematopoietic CCLs.

Supplemental Table 4.1

Top 25 compounds that target WT clones (AUC of DRPs)						
	A1	A2	C1		C1/A2	C1/A1
Daporinad	0.042	0.041	1.546		37.757	36.463
CAY10618	0.407	0.395	11.251		28.517	27.621
GMX-1778	0.008	0.008	0.161		19.314	19.120
STF-31	1.398	1.404	26.493		18.876	18.945
Neratinib	7.878	4.255	26.188		6.155	3.324
RG-108	114.442	97.000	289.441		2.984	2.529
bafilomycin A1	0.044	0.053	0.114		2.144	2.589
MK-2206	7.743	5.425	10.811		1.993	1.396
Ciclopirox	13.604	9.566	18.723		1.957	1.376
ML258	31.860	18.760	35.195		1.876	1.105
SID 26681509	104.730	69.887	128.390		1.837	1.226
marinopyrrole A	10.359	8.878	16.242		1.829	1.568
BRD-K63431240	3.541	2.555	4.519		1.769	1.276
L-685458	22.599	20.379	35.228		1.729	1.559
BRD-K30748066	15.885	10.829	18.378		1.697	1.157
Afatinib	12.354	11.441	19.328		1.689	1.564
oligomycin A	9.742	7.765	12.978		1.671	1.332
MI-2	26.410	23.360	38.435		1.645	1.455
importazole	11.694	11.743	19.316		1.645	1.652
LBH-589	3.547	1.570	2.523		1.607	0.711
niclosamide	6.649	5.143	8.228		1.600	1.237
GANT-61	27.726	21.422	34.263		1.599	1.236
Masitinib	21.044	22.378	35.571		1.590	1.690
Ibrutinib	28.312	22.963	36.494		1.589	1.289
Lovastatin	90.314	50.536	79.084		1.565	0.876

Top 25 compounds that target non-WT clones (AUC of DRPs)						
	A1	A2	C1		C1/A2	C1/A1
SJ-172550	26.642	29.003	26.819		0.925	1.007
C6-ceramide	49.523	55.167	50.918		0.923	1.028
SKI-II	82.666	92.914	85.647		0.922	1.036
semagacestat	55.280	55.914	51.403		0.919	0.930
PHA-793887	17.894	16.415	15.009		0.914	0.839
olaparib	101.182	100.558	91.166		0.907	0.901
AGK-2	17.081	16.969	15.305		0.902	0.896
BMS-195614	44.520	52.578	46.849		0.891	1.052
vincristine	15.494	13.209	11.746		0.889	0.758
AM-580	196.452	227.205	201.335		0.886	1.025
LE-135	36.040	35.073	31.060		0.886	0.862
KX2-391	7.275	5.892	5.161		0.876	0.709
BRD-K70511574	8.300	8.612	7.515		0.873	0.905
parbendazole	15.157	12.822	11.174		0.872	0.737
serdemetan	9.390	9.042	7.776		0.860	0.828
methotrexate	13.745	14.873	12.626		0.849	0.919
AZ-3146	17.094	17.530	14.783		0.843	0.865
silmitasertib	14.234	14.158	11.890		0.840	0.835
BIBR-1532	119.839	109.935	91.941		0.836	0.767
pevonedistat	7.240	6.541	5.369		0.821	0.741
CHM-1	7.349	7.459	6.102		0.818	0.830
teniposide	1.728	1.785	1.456		0.816	0.843
fumonisin B1	272.352	334.766	271.149		0.810	0.996
SU11274	9.649	9.704	7.828		0.807	0.811
NSC95397	15.158	12.363	9.940		0.804	0.656

**Supplemental Table 4.1:** List of AUC of top 50 compounds that differentially target NSD2-WT clones and top 50 compounds that differentially target NSD2 non-WT clones as analyzed by the ratios of AUC calculated.

## Supplemental Table 4.2

### WT-R vs del-NSD2

	score	log2(fold_change)	p_value	q_value		score	log2(fold_change)	p_value	q_value
CLDN6	4.76	-5.37	8.11E-03	4.63E-02	ITM2A	-4.53	4.33	7.45E-05	1.11E-03
TNFRSF10D	4.75	-5.27	4.44E-16	5.67E-14	FAM84B	-4.46	4.13	0.00E+00	0.00E+00
CXCL12	4.66	-4.84	1.85E-05	3.29E-04	SNX19	-4.42	4.03	1.41E-09	6.68E-08
GALC	4.60	-4.59	8.12E-09	3.28E-07	HAAO	-4.40	3.98	9.68E-05	1.39E-03
TRPA1	4.51	-4.27	1.16E-12	9.55E-11	ZNF35	-4.19	3.50	1.01E-07	3.21E-06
PDGFD	4.46	-4.12	1.58E-04	2.10E-03	RASSF9	-4.14	3.41	1.26E-03	1.15E-02
CALCA	4.30	-3.74	5.04E-14	5.09E-12	CFI	-4.02	3.20	2.92E-04	3.48E-03
BARX2	4.26	-3.65	1.15E-08	4.47E-07	JAKMIP1	-4.02	3.20	2.57E-08	9.33E-07
MXRA8	4.22	-3.57	1.62E-03	1.41E-02	PPP1R13L	-3.96	3.11	4.97E-07	1.34E-05
MEGF10	4.21	-3.55	6.04E-07	1.61E-05	BCHE	-3.91	3.03	3.55E-15	4.15E-13
HSPB8	4.17	-3.46	0.00E+00	0.00E+00	TMEM47	-3.80	2.88	2.10E-08	7.78E-07
CALCB	4.04	-3.24	0.00E+00	0.00E+00	NELL1	-3.80	2.88	2.71E-05	4.59E-04
IGFBPL1	4.04	-3.24	0.00E+00	0.00E+00	CPS1	-3.69	2.72	1.55E-15	1.93E-13
LINC00842	4.03	-3.23	6.69E-05	1.01E-03	C1orf61	-3.65	2.68	2.61E-03	2.04E-02
PLP2	3.83	-2.92	1.07E-06	2.68E-05	SLC30A3	-3.52	2.52	3.62E-04	4.16E-03
LMCD1	3.78	-2.85	1.31E-10	7.74E-09	TMEM59L	-3.49	2.49	3.99E-09	1.72E-07
P2RX1	3.77	-2.84	3.82E-03	2.69E-02	CCDC89	-3.29	2.27	7.42E-04	7.50E-03
FAM84A	3.76	-2.82	2.01E-07	5.96E-06	BAI2	-3.21	2.20	4.35E-07	1.20E-05
C16orf71	3.75	-2.81	1.26E-03	1.15E-02	CDR1	-3.21	2.19	0.00E+00	0.00E+00
TMEM163	3.73	-2.78	1.93E-05	3.42E-04	LAMB2P1	-3.18	2.17	4.81E-03	3.22E-02
SPARC	3.66	-2.69	0.00E+00	0.00E+00	ARHGFB6	-3.15	2.14	2.41E-12	1.89E-10
PTGER4	3.66	-2.69	1.99E-04	2.54E-03	ZIC3	-3.14	2.13	4.96E-12	3.74E-10
OTX2	3.63	-2.66	2.01E-03	1.67E-02	DGCR5	-3.13	2.12	1.16E-04	1.62E-03
SNAR-C2	3.63	-2.65	2.86E-05	4.82E-04	SAMD13	-3.12	2.11	3.16E-05	5.25E-04
TRPV2	3.59	-2.60	5.03E-06	1.05E-04	PNPLA4	-3.02	2.02	2.32E-10	1.31E-08
SNAR-C1	3.58	-2.60	7.20E-09	2.95E-07	GPC4	-3.00	2.00	1.30E-10	7.72E-09
ZNF883	3.57	-2.58	1.03E-07	3.25E-06	MAGEB2	-2.98	1.98	7.52E-10	3.79E-08
BATF2	3.55	-2.56	1.12E-03	1.04E-02	DIAPH2	-2.97	1.97	2.85E-13	2.59E-11
SPINK5	3.51	-2.51	2.23E-05	3.86E-04	PPP2R2C	-2.88	1.89	1.18E-04	1.64E-03
USP51	3.51	-2.51	2.28E-06	5.28E-05	COL4A2	-2.76	1.80	6.79E-07	1.78E-05
DPYSL3	3.49	-2.50	3.30E-06	7.28E-05	RPL10L	-2.74	1.78	1.23E-08	4.75E-07
PPAPDC1A	3.47	-2.47	3.87E-03	2.72E-02	MAPK15	-2.68	1.73	5.14E-03	3.38E-02
CCKBR	3.46	-2.46	3.73E-06	8.11E-05	FAM228B	-2.67	1.72	1.54E-03	1.36E-02
SNAR-C4	3.45	-2.45	1.67E-06	3.99E-05	FGF13	-2.67	1.72	1.34E-03	1.21E-02
DOK5	3.45	-2.45	5.14E-06	1.07E-04	RPL22L1	-2.65	1.70	0.00E+00	0.00E+00
SNAR-B2	3.45	-2.44	0.00E+00	0.00E+00	ZNF132	-2.61	1.67	5.57E-03	3.56E-02
MAGEL2	3.43	-2.43	4.74E-05	7.49E-04	TMEM200B	-2.60	1.66	3.67E-03	2.61E-02
SNAR-C3	3.43	-2.42	5.63E-07	1.50E-05	PPFIA4	-2.60	1.66	8.39E-05	1.23E-03
IFI35	3.42	-2.42	7.51E-10	3.79E-08	ZNF75D	-2.56	1.63	4.11E-07	1.13E-05
SFTA3	3.42	-2.41	1.47E-04	1.97E-03	NPAS2	-2.55	1.62	2.38E-03	1.90E-02
NEDD9	3.41	-2.40	3.48E-04	4.01E-03	HOXA4	-2.53	1.61	3.19E-03	2.38E-02
IRF7	3.39	-2.38	3.89E-04	4.42E-03	ASAP2	-2.53	1.61	1.85E-06	4.37E-05
SNAR-E	3.38	-2.37	3.31E-06	7.30E-05	TMEM187	-2.53	1.61	1.32E-07	4.06E-06
PDE3A	3.35	-2.34	9.17E-06	1.80E-04	KRTAP19-1	-2.52	1.60	1.78E-03	1.52E-02
L3MBTL4	3.35	-2.34	3.62E-05	5.90E-04	RPL10	-2.50	1.59	0.00E+00	0.00E+00
CITED1	3.33	-2.32	6.88E-15	7.66E-13	ZNF25	-2.50	1.58	2.47E-04	3.03E-03
APOBEC3G	3.33	-2.32	4.45E-06	9.43E-05	GNAZ	-2.48	1.57	3.15E-05	5.23E-04
CXorf31	3.32	-2.31	6.58E-03	4.01E-02	COL4A1	-2.46	1.56	1.15E-08	4.47E-07
DOCK8	3.32	-2.31	1.45E-06	3.53E-05	MPP1	-2.46	1.55	4.44E-16	5.67E-14
ACTRT3	3.31	-2.29	1.35E-08	5.14E-07	SMIM10	-2.45	1.55	1.78E-05	3.19E-04
EPST11	3.30	-2.29	3.04E-03	2.29E-02	CARS2	-2.45	1.54	5.16E-04	5.57E-03
ABCB1	3.29	-2.28	7.03E-12	5.27E-10	TP53I13	-2.42	1.53	9.45E-05	1.37E-03
MYD88	3.29	-2.28	4.44E-14	4.50E-12	PCDH10	-2.42	1.52	7.21E-10	3.68E-08
APBA1	3.27	-2.26	1.39E-05	2.58E-04	CAMK2B	-2.41	1.52	3.57E-06	7.80E-05
KCNV1	3.27	-2.25	8.72E-05	1.27E-03	PREP	-2.38	1.49	1.03E-09	5.00E-08
PTRPM	3.27	-2.25	4.69E-06	9.88E-05	B4GALNT4	-2.38	1.49	4.32E-03	2.97E-02
C21orf88	3.26	-2.25	3.65E-03	2.60E-02	SLC16A2	-2.37	1.49	2.05E-06	4.77E-05
RORB	3.26	-2.25	8.82E-07	2.26E-05	ZNF385B	-2.35	1.47	1.58E-04	2.10E-03
UNC13D	3.24	-2.23	7.83E-04	7.83E-03	SDC1	-2.32	1.45	8.43E-05	1.24E-03
ZNF285	3.23	-2.22	1.24E-03	1.14E-02	HTATSF1	-2.31	1.44	0.00E+00	0.00E+00
LTK	3.23	-2.22	2.97E-03	2.25E-02	HSPG2	-2.29	1.43	1.33E-05	2.48E-04
WFDC2	3.21	Inf	1.17E-03	1.09E-02	PLAC1	-2.28	1.42	4.98E-05	7.83E-04
FZD10	3.20	-2.19	8.47E-09	3.41E-07	IER2	-2.27	1.41	5.73E-05	8.83E-04
MYH3	3.20	-2.19	1.28E-03	1.17E-02	JUN	-2.26	1.41	7.51E-10	3.79E-08
DES	3.17	-2.15	1.83E-03	1.55E-02	CXorf40B	-2.26	1.41	9.46E-10	4.64E-08
PLK2	3.13	-2.12	0.00E+00	0.00E+00	CETN2	-2.24	1.39	7.37E-10	3.75E-08
PALM3	3.12	-2.11	2.74E-04	3.30E-03	FAM127A	-2.23	1.39	2.27E-10	1.29E-08
SHISA2	3.12	-2.11	0.00E+00	0.00E+00	SCCPDH	-2.21	1.37	2.23E-08	8.21E-07
FAM214B	3.10	-2.09	1.15E-03	1.07E-02	MGAT3	-2.20	1.36	1.42E-03	1.27E-02
NME5	3.09	-2.08	4.12E-03	2.86E-02	PLXNA3	-2.20	1.36	4.24E-04	4.74E-03
WNT11	3.08	-2.07	5.33E-05	8.31E-04	TSPAN7	-2.19	1.36	8.54E-10	4.23E-08
ESAM	3.07	-2.07	8.09E-04	8.04E-03	MOXD1	-2.19	1.36	1.03E-06	2.59E-05
SLC27A6	3.06	-2.06	1.16E-07	3.60E-06	TMBIM1	-2.19	1.36	1.28E-04	1.75E-03
FZD9	3.04	-2.03	4.21E-03	2.90E-02	SRC	-2.18	1.35	2.43E-05	4.17E-04
TP53I3	3.00	-2.00	7.75E-14	7.64E-12	ST3GAL1	-2.16	1.33	1.13E-10	6.80E-09
SEPP1	2.97	-1.97	0.00E+00	0.00E+00	OLFML2A	-2.14	1.32	1.23E-04	1.69E-03
SATB1	2.95	-1.96	8.88E-16	1.12E-13	HACE1	-2.13	1.31	2.58E-05	4.39E-04
GRAPL	2.95	-1.95	5.33E-03	3.46E-02	GD11	-2.07	1.27	2.31E-10	1.31E-08
ALDH3A1	2.94	-1.95	1.62E-04	2.14E-03	TCEANC2	-2.05	1.25	1.01E-08	3.99E-07

	score	log2(fold change)	p value	q value		score	log2(fold change)	p value	q value
KCNT2	2.93	-1.94	1.64E-05	2.96E-04	CKMT1A	-2.03	1.24	2.98E-10	1.64E-08
FIBIN	2.93	-1.94	3.38E-04	3.92E-03	CA12	-2.01	1.23	4.00E-04	4.51E-03
C10orf10	2.93	-1.93	1.23E-12	1.01E-10	ZNF275	-2.01	1.23	3.36E-06	7.40E-05
DHRS2	2.92	-1.93	1.04E-07	3.27E-06	CKMT1B	-2.00	1.22	7.27E-10	3.70E-08
ADRB2	2.86	-1.88	1.58E-05	2.87E-04	NAP1L3	-1.99	1.21	5.53E-07	1.48E-05
SH2D3C	2.85	-1.87	4.53E-03	3.07E-02	ANXA4	-1.98	1.21	5.70E-08	1.91E-06
CTRL	2.85	-1.87	7.35E-03	4.33E-02	FUK	-1.97	1.20	1.53E-03	1.35E-02
MLC1	2.84	-1.86	1.76E-03	1.51E-02	UBL4A	-1.97	1.20	1.80E-09	8.29E-08
DSCR8	2.84	-1.86	1.28E-06	3.15E-05	ATP11C	-1.96	1.20	2.97E-13	2.69E-11
ADAMTS15	2.82	-1.84	1.29E-03	1.17E-02	ATP1A3	-1.96	1.19	3.47E-06	7.59E-05
BASP1	2.81	-1.83	2.53E-04	3.08E-03	QDPR	-1.96	1.19	4.02E-06	8.63E-05
BMPER	2.79	-1.82	4.69E-05	7.41E-04	BCRP2	-1.96	1.19	7.48E-04	7.54E-03
MAP1A	2.79	-1.82	3.75E-13	3.31E-11	SOSTDC1	-1.96	1.19	6.57E-03	4.01E-02
RAB37	2.79	-1.82	1.36E-03	1.22E-02	MRPS26	-1.96	1.19	5.39E-10	2.83E-08
LINC00152	2.76	-1.79	1.10E-03	1.03E-02	LGALS1	-1.95	1.19	3.72E-07	1.04E-05
LOC645638	2.76	-1.79	5.73E-03	3.64E-02	ITGA4	-1.94	1.18	2.13E-04	2.67E-03
DCHS1	2.73	-1.77	2.58E-04	3.14E-03	HFM1	-1.94	1.18	5.84E-03	3.69E-02
DOHH	2.72	-1.76	1.00E-03	9.57E-03	SF3B14	-1.91	1.16	1.59E-11	1.10E-09
KCNJ2	2.70	-1.74	1.89E-03	1.60E-02	MATK	-1.91	1.16	1.14E-03	1.06E-02
TMEFF2	2.69	-1.74	1.35E-03	1.22E-02	SPEF2	-1.91	1.16	3.62E-03	2.59E-02

### WT-R vs E1099K-NSD2

	score	log2(fold change)	p value	q value
SFRP1	4.92	-6.89	8.83E-06	8.83E-06
DSCR8	4.58	-4.52	3.58E-08	3.58E-08
SNAI2	4.38	-3.91	0.00E+00	0.00E+00
LOC100289187	4.19	-3.51	2.80E-04	2.80E-04
AMBN	4.04	-3.23	1.56E-03	1.56E-03
CLDN6	3.96	-3.11	1.36E-04	1.36E-04
PRKCB	3.90	-3.02	1.44E-09	1.44E-09
C10orf82	3.88	-2.98	9.05E-04	9.05E-04
KCNH5	3.85	-2.95	9.12E-04	9.12E-04
L3MBTL4	3.83	-2.92	9.13E-06	9.13E-06
IFI35	3.82	-2.91	4.68E-10	4.68E-10
CXCL6	3.81	-2.88	5.69E-03	5.69E-03
FZD9	3.81	-2.88	1.98E-03	1.98E-03
TLE6	3.80	-2.87	4.07E-03	4.07E-03
AMHR2	3.78	-2.85	8.59E-10	8.59E-10
KLK8	3.78	-2.84	4.33E-04	4.33E-04
GRAPL	3.75	-2.80	4.29E-04	4.29E-04
GIP	3.67	-2.71	4.54E-03	4.54E-03
COL5A2	3.67	-2.71	5.79E-09	5.79E-09
TEX40	3.60	-2.62	5.56E-03	5.56E-03
TMEM163	3.59	-2.60	1.12E-04	1.12E-04
FAAH2	3.57	-2.59	3.09E-03	3.09E-03
CALCA	3.54	-2.55	4.27E-09	4.27E-09
ZNF487P	3.51	-2.51	8.24E-04	8.24E-04
FZD10	3.50	-2.50	1.25E-08	1.25E-08
EFCAB4A	3.49	-2.49	7.24E-05	7.24E-05
LINC00842	3.48	-2.48	5.61E-04	5.61E-04
ARSJ	3.46	-2.46	6.42E-04	6.42E-04
C16orf45	3.46	-2.45	1.85E-05	1.85E-05
FANK1	3.43	-2.42	3.45E-03	3.45E-03
FAM84B	3.41	-2.41	7.96E-11	7.96E-11
CILP2	3.41	-2.40	2.76E-08	2.76E-08
ZNF467	3.40	-2.39	5.93E-03	5.93E-03
BARX2	3.39	-2.39	3.55E-06	3.55E-06
P2RY2	3.39	-2.38	1.11E-03	1.11E-03
IGFBPL1	3.38	-2.38	2.42E-12	2.42E-12
C10orf10	3.36	-2.35	1.81E-13	1.81E-13
CALCB	3.36	-2.35	0.00E+00	0.00E+00
C12orf68	3.35	-2.34	2.38E-03	2.38E-03
BHLHE40	3.34	-2.33	0.00E+00	0.00E+00
TRPV2	3.34	-2.32	3.89E-05	3.89E-05
ASTN1	3.32	-2.31	3.07E-04	3.07E-04
CYP4F22	3.31	-2.30	4.25E-03	4.25E-03
NOS2	3.28	-2.27	4.14E-06	4.14E-06
LMCD1	3.28	-2.27	1.13E-07	1.13E-07
RHEBL1	3.24	-2.23	2.14E-04	2.14E-04
TMEM100	3.23	-2.22	4.56E-04	4.56E-04
MGC12916	3.23	-2.21	3.74E-03	3.74E-03
EFCAB10	3.22	-2.21	4.44E-04	4.44E-04
UNC13D	3.20	-2.18	2.05E-03	2.05E-03
FOXO4	3.17	-2.16	2.66E-06	2.66E-06
FUT8-AS1	3.15	-2.14	1.00E-03	1.00E-03
ANKRD34B	3.14	-2.13	1.30E-05	1.30E-05
RASA4CP	3.06	-2.06	1.28E-03	1.28E-03
FGD3	3.05	-2.05	2.11E-03	2.11E-03
MMP2	3.04	-2.03	1.18E-05	1.18E-05
BATF2	3.03	-2.03	4.78E-03	4.78E-03
ZNF883	3.01	-2.01	2.18E-05	2.18E-05

	score	log2(fold change)	p value	q value
FGF13	-4.79	5.53	0.00E+00	0.00E+00
ITM2A	-4.64	4.72	1.87E-05	1.87E-05
SNX19	-4.45	4.11	1.31E-09	1.31E-09
LOC643648	-4.40	3.97	1.20E-03	1.20E-03
NELL1	-4.27	3.67	3.94E-08	3.94E-08
CPS1	-4.01	3.19	0.00E+00	0.00E+00
HAAO	-3.95	3.09	4.02E-03	4.02E-03
FAM133A	-3.94	3.08	9.95E-06	9.95E-06
RIPPLY2	-3.90	3.01	1.40E-03	1.40E-03
MAP1LC3C	-3.88	2.99	0.00E+00	0.00E+00
PTN	-3.78	2.84	3.11E-05	3.11E-05
TMEM47	-3.75	2.81	1.19E-07	1.19E-07
BEX1	-3.68	2.71	0.00E+00	0.00E+00
GAP43	-3.64	2.66	1.66E-03	1.66E-03
ZIK1	-3.64	2.66	6.00E-05	6.00E-05
CCDC89	-3.63	2.66	9.46E-05	9.46E-05
KLHDC9	-3.62	Inf	3.74E-05	3.74E-05
CDR1	-3.57	2.58	0.00E+00	0.00E+00
PTGFR	-3.47	2.47	2.90E-05	2.90E-05
DPYD	-3.45	2.45	1.12E-05	1.12E-05
SCG3	-3.42	2.41	1.55E-04	1.55E-04
SAMD13	-3.41	2.41	4.60E-06	4.60E-06
DIAPH3-AS1	-3.38	2.37	1.70E-04	1.70E-04
ZIC3	-3.18	2.17	1.27E-11	1.27E-11
GABBR2	-3.14	2.13	8.78E-05	8.78E-05
FGF12	-3.10	2.09	1.83E-07	1.83E-07
MYRIP	-3.08	2.08	1.07E-04	1.07E-04
SMIM10	-3.08	2.07	9.63E-09	9.63E-09
PCBP3	-3.08	2.07	1.28E-03	1.28E-03
BST2	-3.07	2.06	1.41E-03	1.41E-03
NAP1L3	-3.06	2.06	0.00E+00	0.00E+00
CALB1	-3.06	2.05	2.83E-03	2.83E-03
HTATSF1	-3.02	2.02	0.00E+00	0.00E+00
MSX1	-2.98	1.98	1.03E-12	1.03E-12
ZNF25	-2.97	1.97	9.47E-06	9.47E-06
RPL22L1	-2.91	1.92	0.00E+00	0.00E+00
TMCO3	-2.85	1.87	6.57E-07	6.57E-07
KRTAP19-1	-2.78	1.81	6.64E-04	6.64E-04
SLC25A29	-2.78	1.81	5.98E-03	5.98E-03
PPP2R2C	-2.71	1.75	7.38E-04	7.38E-04
NHSL2	-2.62	1.68	3.68E-03	3.68E-03
WDR78	-2.61	1.67	1.27E-03	1.27E-03
NGEF	-2.50	1.59	3.55E-03	3.55E-03
FAM127B	-2.50	1.59	0.00E+00	0.00E+00
FLNA	-2.49	1.58	0.00E+00	0.00E+00
GPC4	-2.45	1.55	3.94E-06	3.94E-06
RAB38	-2.40	1.51	3.93E-03	3.93E-03
MAGEH1	-2.39	Inf	1.16E-03	1.16E-03
HS6ST2	-2.35	1.47	0.00E+00	0.00E+00
COL4A6	-2.33	1.46	4.15E-05	4.15E-05
COL4A1	-2.32	1.45	5.20E-07	5.20E-07
KIAA0232	-2.30	1.44	5.97E-14	5.97E-14
COL4A2	-2.30	1.43	2.37E-04	2.37E-04
CXorf40B	-2.26	1.41	6.37E-09	6.37E-09
SCN9A	-2.26	1.41	1.51E-04	1.51E-04
LOC389906	-2.26	1.40	1.91E-03	1.91E-03
FAM127A	-2.22	1.37	3.68E-09	3.68E-09
ZNF75D	-2.20	1.36	7.62E-05	7.62E-05

	score	log2(fold_change)	p_value	q_value
ANGPT1	2.96	-1.96	2.08E-03	2.08E-03
ARRDC2	2.96	-1.96	2.83E-03	2.83E-03
PLAU	2.95	-1.96	2.56E-04	2.56E-04
HMGCL1	2.95	-1.96	2.24E-04	2.24E-04
C3	2.95	-1.95	2.69E-03	2.69E-03
NEDD9	2.94	-1.95	1.04E-03	1.04E-03
C21orf88	2.94	-1.95	5.45E-03	5.45E-03
TENC1	2.93	-1.94	1.58E-03	1.58E-03
CYP26A1	2.93	-1.94	0.00E+00	0.00E+00
ITGB4	2.93	-1.94	8.32E-04	8.32E-04
RSG1	2.92	-1.93	5.02E-03	5.02E-03
TMEM37	2.92	-1.93	1.19E-05	1.19E-05
DES	2.91	-1.92	4.76E-03	4.76E-03
SCIN	2.90	-1.92	5.56E-06	5.56E-06
TRPA1	2.90	-1.91	6.45E-06	6.45E-06
TMEM171	2.87	-1.89	4.70E-04	4.70E-04
FAM84A	2.85	-1.87	1.04E-04	1.04E-04
APOBEC3G	2.85	-1.87	1.13E-04	1.13E-04
MFAP3L	2.83	-1.85	8.70E-07	8.70E-07
DKF5	2.81	-1.84	1.88E-04	1.88E-04
CXCR4	2.81	-1.84	1.91E-11	1.91E-11
MEGF10	2.81	-1.83	8.67E-04	8.67E-04
SHISA3	2.80	-1.83	4.33E-03	4.33E-03
COMP	2.80	-1.83	1.16E-03	1.16E-03
ATHL1	2.80	-1.82	9.58E-07	9.58E-07
SYNGR3	2.79	-1.82	1.76E-04	1.76E-04
PRKCH	2.78	-1.81	7.55E-04	7.55E-04
FAT4	2.72	-1.76	6.50E-06	6.50E-06
ADAMTS15	2.72	-1.76	3.24E-03	3.24E-03
SYNM	2.72	-1.76	4.15E-04	4.15E-04
OXGR1	2.69	-1.74	5.88E-03	5.88E-03
SEMA3A	2.69	-1.73	1.50E-07	1.50E-07
TRIM7	2.67	-1.72	2.62E-04	2.62E-04
DLL4	2.67	-1.72	4.73E-03	4.73E-03
HAPLN3	2.66	-1.71	7.74E-08	7.74E-08
IRF7	2.65	-1.70	2.66E-04	2.66E-04
AMPD3	2.64	-1.70	2.02E-03	2.02E-03
HOXA1	2.62	-1.68	6.66E-16	6.66E-16
CASC1	2.62	-1.68	1.59E-03	1.59E-03
DOCK8	2.58	-1.65	3.27E-04	3.27E-04
ZFR2	2.58	-1.65	7.07E-04	7.07E-04

	score	log2(fold_change)	p_value	q_value
TAPT1	-2.19	1.35	2.30E-03	2.30E-03
HMGNS5	-2.16	1.34	5.57E-11	5.57E-11
C4orf48	-2.14	1.32	5.75E-05	5.75E-05
ADPRHL1	-2.11	1.30	2.23E-03	2.23E-03
CNRIP1	-2.11	1.30	3.74E-05	3.74E-05
MMGT1	-2.11	1.30	2.47E-10	2.47E-10
LONRF2	-2.10	1.29	7.48E-06	7.48E-06
DIAPH2	-2.10	1.29	7.21E-06	7.21E-06
ZNF71	-2.10	1.29	1.99E-03	1.99E-03
ZBTB49	-2.06	1.27	1.01E-03	1.01E-03
QDPR	-2.04	1.25	3.81E-06	3.81E-06
SLC16A2	-2.03	1.25	1.97E-04	1.97E-04
SLC16A14	-2.03	1.24	1.49E-05	1.49E-05
HPRT1	-2.03	1.24	1.81E-09	1.81E-09
LCORL	-2.01	1.23	1.16E-12	1.16E-12
ASAP2	-1.99	1.22	1.73E-04	1.73E-04
RB1	-1.97	1.20	1.60E-07	1.60E-07
KLHL13	-1.96	1.20	0.00E+00	0.00E+00
BLOC1S4	-1.96	1.19	2.58E-04	2.58E-04
INA	-1.95	1.19	8.67E-04	8.67E-04
TBC1D14	-1.94	1.18	5.02E-07	5.02E-07
CERKL	-1.94	1.18	2.24E-03	2.24E-03
INPP1	-1.94	1.18	4.93E-04	4.93E-04
NAPRT1	-1.93	1.18	1.00E-06	1.00E-06
ARHGEF6	-1.91	1.16	6.19E-04	6.19E-04
BEX2	-1.85	1.12	0.00E+00	0.00E+00
EVC	-1.84	1.12	1.98E-03	1.98E-03
ME1	-1.81	1.10	1.93E-06	1.93E-06
CELSR3	-1.81	1.09	2.33E-04	2.33E-04
PRKAA2	-1.79	1.08	2.18E-07	2.18E-07
TCEANC2	-1.77	1.07	4.83E-06	4.83E-06
LAP3	-1.77	1.07	3.31E-07	3.31E-07
C3orf14	-1.77	1.07	1.58E-06	1.58E-06
RNF4	-1.76	1.06	2.48E-09	2.48E-09
FAM127C	-1.75	1.06	2.06E-05	2.06E-05
MAGI3	-1.75	1.06	1.12E-07	1.12E-07
WDR1	-1.75	1.05	4.92E-08	4.92E-08
AMMECR1	-1.74	1.05	4.89E-06	4.89E-06
NAP1L2	-1.74	1.05	5.73E-03	5.73E-03
SNX9	-1.73	1.04	3.82E-04	3.82E-04
ARL4C	-1.73	1.04	7.85E-04	7.85E-04

### del-NSD2 vs E1099K-NSD2

	score	log2(fold_change)	p_value	q_value
FAM84B	4.89	-6.55	0.00E+00	0.00E+00
SFRP1	4.88	-6.37	4.43E-05	1.68E-03
SLC6A19	4.84	-5.96	2.00E-01	1.00E+00
SPINT3	4.75	-5.27	8.31E-01	1.00E+00
CGB1	4.74	-5.22	5.21E-01	1.00E+00
C3orf55	4.67	-4.88	9.04E-01	1.00E+00
TUSC5	4.62	-4.67	8.04E-01	1.00E+00
LINC00570	4.55	-4.40	4.19E-01	1.00E+00
SNAI2	4.54	-4.38	0.00E+00	0.00E+00
LINC00595	4.54	-4.36	8.87E-01	1.00E+00
NKAIN4	4.52	-4.32	8.16E-01	1.00E+00
PTPRN2	4.52	-4.30	8.17E-01	1.00E+00
CYP24A1	4.50	-4.26	7.60E-01	1.00E+00
FBN3	4.49	-4.21	4.31E-02	1.00E+00
PCDHB12	4.48	-4.18	1.55E-01	1.00E+00
PCDHB8	4.47	-4.15	2.43E-01	1.00E+00
EPHA1	4.45	-4.11	6.30E-01	1.00E+00
APLN	4.44	-4.07	3.51E-01	1.00E+00
LOC283867	4.41	-4.00	5.22E-01	1.00E+00
CRB2	4.39	-3.95	3.67E-01	1.00E+00
IL27	4.38	-3.93	8.99E-01	1.00E+00
CHDC2	4.37	-3.90	2.15E-01	1.00E+00
ARHGEF33	4.35	-3.84	6.86E-01	1.00E+00
WFIKK1	4.34	-3.81	4.43E-02	1.00E+00
ANKRD37	4.33	-3.79	7.05E-02	2.59E-01
NOD2	4.32	-3.79	7.01E-02	1.00E+00
FAAH2	4.32	-3.78	5.43E-06	2.87E-04
LOC100505817	4.31	-3.76	7.31E-01	1.00E+00
PLA2G4D	4.30	-3.74	4.25E-01	1.00E+00
ST8SIA2	4.30	-3.73	7.45E-02	1.00E+00
PADI3	4.29	-3.71	7.66E-02	1.00E+00
EPS8L2	4.27	-3.67	4.99E-02	1.00E+00
GRAP	4.26	-3.65	2.36E-01	1.00E+00
CCIN	4.26	-3.64	5.01E-02	1.00E+00
TMOD4	4.25	-3.63	1.88E-01	1.00E+00
TLR4	4.25	-3.62	3.68E-01	1.00E+00
TAGLN3	4.23	-3.58	7.39E-01	1.00E+00
TIPARP-AS1	4.21	-3.55	8.94E-02	1.00E+00
ATP2B3	4.21	-3.55	1.27E-03	1.00E+00

	score	log2(fold_change)	p_value	q_value
MTRNR2L2	-5.00	0.00	1.00E+00	1.00E+00
DHFR	-4.93	0.00	1.00E+00	1.00E+00
TOX3	-4.86	6.11	4.63E-01	1.00E+00
LOC643669	-4.72	5.11	6.56E-01	1.00E+00
CHL1	-4.68	4.93	6.74E-01	1.00E+00
PCSK2	-4.66	4.82	4.18E-01	1.00E+00
GALC	-4.65	4.79	2.57E-09	3.35E-07
COL4A2-AS1	-4.60	4.58	4.34E-01	1.00E+00
TNFRSF10D	-4.53	4.34	9.95E-11	1.76E-08
THBS1	-4.52	4.30	8.98E-02	1.00E+00
LRLT3	-4.50	4.25	1.45E-01	1.00E+00
PPP2R2B	-4.48	4.19	2.52E-01	1.00E+00
ZNF705B	-4.48	4.18	5.61E-01	1.00E+00
LOC100129620	-4.48	4.18	1.21E-01	1.00E+00
PDGFD	-4.46	4.12	1.92E-04	5.62E-03
CALB1	-4.45	4.10	1.28E-03	2.48E-02
GLTPD2	-4.43	4.05	5.53E-02	2.25E-01
USP17L20	-4.43	4.04	7.31E-01	1.00E+00
SNAR-D	-4.40	3.96	2.33E-01	4.64E-01
MGP	-4.39	3.93	7.73E-01	1.00E+00
SPANXC	-4.38	3.93	5.05E-01	1.00E+00
CXCL12	-4.37	3.89	7.33E-04	1.63E-02
ZNF837	-4.36	3.88	4.38E-01	1.00E+00
NAT2	-4.36	3.88	9.11E-02	1.00E+00
NLRP9	-4.34	3.82	6.38E-02	1.00E+00
SH3TC1	-4.33	3.79	2.05E-01	1.00E+00
MCSR	-4.32	3.77	6.61E-01	1.00E+00
USP17L9P	-4.31	3.76	8.98E-01	1.00E+00
FGF13	-4.31	3.76	0.00E+00	0.00E+00
ACTBL2	-4.29	3.72	5.21E-01	1.00E+00
GABRB2	-4.29	3.72	2.63E-04	1.00E+00
DLGAP1	-4.29	3.71	1.51E-02	1.00E+00
USP17L15	-4.29	3.70	7.71E-01	1.00E+00
C6orf7	-4.28	3.68	2.20E-01	1.00E+00
HYDIN	-4.26	3.65	8.61E-03	1.00E+00
LANL3	-4.24	3.61	2.34E-01	1.00E+00
RNU6-76	-4.22	inf	9.45E-02	3.03E-01
TMPRSS9	-4.22	3.57	8.11E-02	1.00E+00
DACT1	-4.22	3.56	2.64E-02	1.00E+00

	score	log2(fold_change)	p_value	q_value		score	log2(fold_change)	p_value	q_value
PDE2A	4.19	-3.50	2.42E-02	1.00E+00	FOXL1	-4.21	3.55	4.19E-01	1.00E+00
PP14571	4.16	-3.44	2.14E-01	1.00E+00	HMGCS2	-4.21	3.55	4.01E-01	1.00E+00
SOWAHD	4.16	-3.44	6.94E-02	1.00E+00	DUX4L4	-4.19	3.49	7.75E-01	1.00E+00
STAB1	4.15	-3.44	2.74E-01	1.00E+00	SLC51B	-4.16	3.45	7.38E-01	1.00E+00
CTSZ	4.15	-3.43	3.74E-02	1.00E+00	SSC5D	-4.16	3.44	1.97E-01	1.00E+00
CFI	4.14	-3.40	3.59E-03	4.74E-02	SERPIN8	-4.12	3.37	1.49E-01	1.00E+00
TTY8B	4.13	-3.40	7.74E-01	1.00E+00	MTVR2	-4.12	3.36	3.41E-02	1.66E-01
SLCO3A1	4.13	-3.39	9.99E-02	1.00E+00	NR2E1	-4.11	3.35	1.67E-01	1.00E+00
CYP4F22	4.13	-3.39	2.97E-06	1.72E-04	NEFH	-4.11	3.35	2.56E-01	1.00E+00
ZNF280A	4.12	-3.37	3.45E-02	1.00E+00	CDYL2	-4.10	3.34	1.10E-01	1.00E+00
NTNG2	4.11	-3.36	1.59E-01	1.00E+00	FAM57B	-4.10	3.33	3.43E-02	1.00E+00
KRT4	4.11	-3.35	8.44E-01	1.00E+00	MAP1LC3C	-4.09	3.32	0.00E+00	0.00E+00
CASQ2	4.08	-3.31	5.07E-01	1.00E+00	HUS1B	-4.07	3.29	2.73E-01	1.00E+00
LRRC17	4.08	-3.31	2.55E-03	3.85E-02	MYD88	-4.07	3.29	0.00E+00	0.00E+00
LOC729177	4.06	-3.27	5.21E-02	1.00E+00	ACSM2A	-4.06	3.28	4.05E-01	1.00E+00
FUT5	4.05	-3.26	5.55E-01	1.00E+00	DUSP6	-4.05	3.26	2.39E-02	1.00E+00
GAB3	4.05	-3.26	7.45E-02	1.00E+00	ZNF648	-4.03	3.23	2.83E-01	1.00E+00
ERP27	4.04	-3.24	4.75E-01	1.00E+00	OR51B5	-4.03	3.22	1.95E-02	1.19E-01
HNF4A	4.04	-3.23	6.39E-01	1.00E+00	TCEB3B	-4.02	3.20	7.75E-01	1.00E+00
GCNT7	4.03	-3.22	2.45E-01	1.00E+00	MAP2	-4.01	3.19	6.56E-04	1.00E+00
RAG2	4.03	-3.22	3.09E-01	1.00E+00	SLC15A3	-4.01	3.19	2.77E-01	1.00E+00
MAFA	4.03	-3.22	1.32E-01	1.00E+00	KCNA1	-3.99	3.16	6.06E-01	1.00E+00
AQP4	4.03	-3.21	8.66E-03	1.00E+00	ACSM2B	-3.99	3.15	4.97E-01	1.00E+00
C10orf107	4.00	-3.16	9.20E-02	1.00E+00	ZSCAN10	-3.98	3.14	4.87E-01	1.00E+00
NXP3	3.99	-3.16	9.30E-02	1.00E+00	SLIT2-IT1	-3.97	3.13	6.70E-03	6.95E-02
CRYAB	3.98	-3.14	5.73E-02	2.30E-01	CACNA2D3-AS1	-3.97	3.13	5.94E-01	1.00E+00
AKR1E2	3.98	-3.14	4.70E-01	1.00E+00	LINC00602	-3.97	3.12	8.43E-01	1.00E+00
WFDCC11	3.97	-3.12	1.45E-01	1.00E+00	PTN	-3.96	3.11	1.48E-05	6.63E-04
SLC45A1	3.94	-3.08	2.74E-01	1.00E+00	PCDHB14	-3.95	3.10	7.17E-01	1.00E+00
C10orf114	3.94	-3.08	1.54E-01	1.00E+00	GBP2	-3.95	3.10	5.40E-01	1.00E+00
IL17D	3.94	-3.07	1.47E-01	1.00E+00	GPX7	-3.95	3.09	1.38E-01	1.00E+00
NGFR	3.93	-3.07	1.47E-01	1.00E+00	RFPL3	-3.94	3.08	8.49E-01	1.00E+00
KCNH3	3.93	-3.07	1.06E-03	1.00E+00	NTRK3	-3.94	3.07	8.35E-01	1.00E+00
TMPPRS13	3.92	-3.05	1.96E-01	1.00E+00	AQP7	-3.93	3.06	3.06E-01	1.00E+00
MACC1-AS1	3.92	-3.04	7.56E-01	1.00E+00	SIRP3	-3.93	3.06	7.52E-01	1.00E+00
FOXO4L1	3.91	-3.03	4.18E-02	1.00E+00	C6orf163	-3.93	3.06	1.44E-01	1.00E+00
UBA7	3.90	-3.02	7.10E-02	1.00E+00	LINGO2	-3.92	3.05	5.62E-02	1.00E+00
SCARNA16	3.90	-3.02	3.51E-01	5.74E-01	COL6A5	-3.91	3.03	1.47E-02	1.00E+00
LOC728084	3.88	-2.99	6.84E-02	1.00E+00	PCDHB7	-3.91	3.03	6.36E-01	1.00E+00
TMEM179	3.88	-2.99	1.13E-01	1.00E+00	MXRA8	-3.91	3.03	9.37E-03	8.35E-02
ALS2CR11	3.86	-2.96	1.15E-01	1.00E+00	LHFPL3	-3.89	3.01	8.97E-01	1.00E+00
ITIH2	3.85	-2.95	3.09E-01	1.00E+00	MYRIP	-3.89	3.00	3.96E-06	2.17E-04
HAR1A	3.84	-2.94	8.11E-01	1.00E+00	CAMK2N2	-3.89	3.00	1.56E-01	1.00E+00
LOC100128568	3.84	-2.93	2.96E-01	1.00E+00	NAGPA-AS1	-3.87	2.97	1.62E-01	1.00E+00
CSMD3	3.82	-2.91	1.82E-02	1.00E+00	RNY4	-3.86	Inf	3.20E-01	5.45E-01
PLEKHG7	3.82	-2.90	1.54E-01	1.00E+00	BMX	-3.86	2.96	1.68E-01	1.00E+00
NKAIN2	3.81	-2.89	1.99E-01	1.00E+00	TRIM53AP	-3.86	2.96	5.78E-01	1.00E+00
FAM83C	3.81	-2.89	6.87E-01	1.00E+00	UPP2	-3.86	2.96	7.03E-02	1.00E+00
INHA	3.80	-2.88	1.30E-01	1.00E+00	TMEM132E	-3.86	2.96	4.53E-02	1.00E+00
VCX2	3.79	-2.87	8.16E-01	1.00E+00	OTX2	-3.86	2.95	6.99E-04	1.58E-02
TRPV6	3.79	-2.86	8.46E-02	1.00E+00	PTPN7	-3.85	2.94	2.52E-01	1.00E+00
TLE6	3.79	-2.86	4.53E-03	5.49E-02	CACNA1E	-3.84	2.93	2.54E-01	1.00E+00
DLGAP3	3.78	-2.85	1.47E-01	1.00E+00	FOLR3	-3.84	2.93	5.06E-01	1.00E+00
AMBN	3.78	-2.85	6.47E-03	1.00E+00	PIGZ	-3.83	2.92	4.15E-02	1.00E+00
TPRG1-AS2	3.77	-2.84	8.12E-02	2.80E-01	MAB21L2	-3.83	2.92	9.60E-01	1.00E+00
SYNPO	3.77	-2.84	4.68E-04	1.00E+00	SMYD1	-3.81	2.88	8.71E-01	1.00E+00
ADCYAP1	3.77	-2.84	9.19E-01	1.00E+00	PLCB2	-3.80	2.88	6.65E-02	1.00E+00
LOC100129055	3.76	-2.82	4.15E-02	1.00E+00	TBX5-AS1	-3.80	2.88	8.61E-01	1.00E+00
C19orf77	3.75	-2.81	9.12E-02	1.00E+00	TEKT4	-3.79	2.87	9.39E-02	1.00E+00
TTY8	3.74	-2.80	7.38E-01	1.00E+00	PAPPA2	-3.79	2.86	4.56E-01	1.00E+00

Supplemental Table 4.2: Differential expression between E1099K-ALLs and WT-ALLs using Cuffdiff and GSEA.

Supplementary Table 4.3

C2 gene sets (WT-R vs del-NSD2)			
NAME	ES	NOM p-value	FDR q-value
REACTOME_NONSENSE_MEDIATED_DECAY_ENHANCED_BY_THE_EXON_JUNCTION_COMPLEX	-0.580	0.0000	0.02491
BENPORATH_ES_WITH_H3K27ME3	0.482	0.0000	0.00000
BENPORATH_PRC2_TARGETS	0.517	0.0000	0.00000
BENPORATH_EED_TARGETS	0.454	0.0000	0.00000
BENPORATH_SUZ12_TARGETS	0.443	0.0000	0.00000
REACTOME_GPCR_LIGAND_BINDING	0.681	0.0000	0.00000
MEISSNER_BRAIN_HCP_WITH_H3K4ME3_AND_H3K27ME3	0.423	0.0000	0.00000
MIKKELSEN_NPC_HCP_WITH_H3K4ME3_AND_H3K27ME3	0.657	0.0000	0.00012
CHIANG_LIVER_CANCER_SUBCLASS_CTNNB1_DN	0.599	0.0000	0.00039
SATO_SILENCED_BY_METHYLATION_IN_PANCREATIC_CANCER_1	0.468	0.0000	0.00132
VECCHI_GASTRIC_CANCER_EARLY_DN	0.492	0.0000	0.00119
KINSEY_TARGETS_OF_EWSR1_FLII_FUSION_DN	0.462	0.0000	0.00122
NABA_SECRETED_FACTORS	0.643	0.0000	0.00119
GROSS_HYPOXIA_VIA_ELK3_AND_HIF1A_UP	0.575	0.0000	0.00183
REACTOME_SIGNALING_BY_GPCR	0.471	0.0000	0.00220
MIKKELSEN_MEF_HCP_WITH_H3K27ME3	0.536	0.0000	0.00548
RIGGI_EWING_SARCOMA_PROGENITOR_UP	0.446	0.0000	0.00849
REACTOME_GPCR_DOWNSTREAM_SIGNALING	0.485	0.0000	0.00818
NABA_MATRISOME	0.395	0.0000	0.00808
HOOI_ST7_TARGETS_DN	0.540	0.0000	0.01206
NABA_MATRISOME_ASSOCIATED	0.426	0.0000	0.01249
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_3D_UP	0.607	0.0000	0.01223
MARTORIATI_MDM4_TARGETS_NEUROEPITHELIUM_DN	0.597	0.0000	0.01175
HELLER_SILENCED_BY_METHYLATION_UP	0.481	0.0000	0.01214
GROSS_HYPOXIA_VIA_ELK3_DN	0.517	0.0000	0.01209
SCHUETZ_BREAST_CANCER_DUCTAL_INVASIVE_UP	0.407	0.0000	0.01354
VART_KSHV_INFECTION_ANGIOGENIC_MARKERS_UP	0.573	0.0028	0.01344
BOQUEST_STEM_CELL_UP	0.471	0.0000	0.01415
DELYS_THYROID_CANCER_UP	0.395	0.0000	0.01423
TAKEDA_TARGETS_OF_NUP98_HOXA9_FUSION_8D_UP	0.539	0.0014	0.01464
HAMAI_APOPTOSIS_VIA_TRAIL_DN	0.514	0.0013	0.01463
MEISSNER_NPC_HCP_WITH_H3_UNMETHYLATED	0.438	0.0012	0.01719
SMID_BREAST_CANCER_NORMAL_LIKE_UP	0.404	0.0000	0.01875
SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_DN	0.477	0.0013	0.02024
KIM_RESPONSE_TO_TSA_AND_DECITABINE_UP	0.537	0.0027	0.02907
YANG_BCL3_TARGETS_UP	0.370	0.0024	0.03600
ACEVEDO_FGFR1_TARGETS_IN_PROSTATE_CANCER_MODEL_DN	0.435	0.0038	0.04121
SHEDDEN_LUNG_CANCER_GOOD_SURVIVAL_A4	0.506	0.0013	0.04871

C6 gene sets (WT-R vs del-NSD2)			
NAME	ES	NOM p-value	FDR q-value
KRAS.600_UP.V1_UP	0.601	0.0000	0.00531
BMI1_DN_MEL18_DN.V1_DN	0.548	0.0000	0.00428
ATF2_UP.V1_DN	0.530	0.0000	0.01180
MEK_UP.V1_UP	0.513	0.0000	0.01585
MEL18_DN.V1_DN	0.493	0.0000	0.02163
KRAS.DF.V1_DN	0.547	0.0000	0.03216
SNF5_DN.V1_DN	0.491	0.0025	0.03510

**H gene set (WT-R vs del-NSD2)**

NAME	ES	NOM p-value	FDR q-value
HALLMARK_KRAS_SIGNALING_UP	0.531	0.0000	0.00710
HALLMARK_INTERFERON_ALPHA_RESPONSE	0.541	0.0014	0.01544
HALLMARK_INTERFERON_GAMMA_RESPONSE	0.432	0.0025	0.04476

**C2 gene sets (WT-R vs E1099K-NSD2)**

NAME	ES	NOM p-value	FDR q-value
MEISSNER_NPC_HCP_WITH_H3_UNMETHYLATED	0.546	0.0000	0.00093
MEISSNER_BRAIN_HCP_WITH_H3K4ME3_AND_H3K27ME3	0.440	0.0000	0.00046
MARTENS_TRETINOIN_RESPONSE_UP	0.550	0.0000	0.00278
WINTER_HYPOXIA_METAGENE	0.544	0.0000	0.00488
REACTOME_GPCR_LIGAND_BINDING	0.603	0.0000	0.00409
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	0.656	0.0000	0.00743
REACTOME_SIGNALING_BY_GPCR	0.512	0.0000	0.00863
ONDER_CDH1_TARGETS_2_DN	0.458	0.0000	0.01229
SMID_BREAST_CANCER_NORMAL_LIKE_UP	0.461	0.0000	0.01496
YOSHIMURA_MAPK8_TARGETS_UP	0.416	0.0000	0.01754
KEGG_WNT_SIGNALING_PATHWAY	0.621	0.0000	0.02594
PLASARI_TGFB1_TARGETS_10HR_UP	0.555	0.0000	0.03067
REACTOME_GPCR_DOWNSTREAM_SIGNALING	0.499	0.0014	0.02831
SHEPARD_BMYB_MORPHOLINO_DN	0.501	0.0028	0.03216
HAMAI_APOPTOSIS_VIA_TRAIL_DN	0.581	0.0014	0.03119
NABA_MATRISOME	0.402	0.0000	0.03069
MEISSNER_NPC_HCP_WITH_H3K4ME2_AND_H3K27ME3	0.553	0.0044	0.03389
BENPORATH_ES_WITH_H3K27ME3	0.388	0.0000	0.03366
FEVR_CTNNB1_TARGETS_UP	0.399	0.0000	0.03278
MCLACHLAN_DENTAL_CARIES_UP	0.496	0.0014	0.03867

**C2 gene sets (del-NSD2 vs E1099K-NSD2)**

NAME	ES	NOM p-value	FDR q-value
REACTOME_PEPTIDE_CHAIN_ELONGATION	0.633	0.0000	0.01026
REACTOME_BETA_DEFENSINS	-0.717	0.0000	0.03179
REACTOME_LIGAND_GATED_ION_CHANNEL_TRANSPORT	-0.790	0.0000	0.02724

**Supplementary Table 4.3** Significant gene sets (by FDR q-values < 0.05) enriched in pairs analyzed.

Supplementary Table 5.1

Top 100 genes upregulated in N-R cells			
Description	Score	Feature P	Fold Change
HSP90AB1	19.19	1.47E-01	1.35
PTTG3P	17.89	1.47E-01	55.02
TOMM5	14.16	1.47E-01	1.18
SLC1A3	13.37	1.47E-01	1.81
PA2G4	10.78	1.47E-01	1.41
POLR3G	10.63	1.47E-01	1.76
ENOPH1	10.32	1.47E-01	1.34
CCT8	10.30	1.47E-01	1.26
NKAPL	9.65	1.47E-01	18.31
NKD2	9.50	1.47E-01	2.44
MYOG	9.31	1.47E-01	3.88
POLR3K	9.16	1.47E-01	1.35
PLS1	9.02	1.47E-01	1.32
FAM118A	8.80	1.47E-01	1.28
HNRNPF	8.51	1.47E-01	1.12
NOP16	8.43	1.47E-01	1.44
PPA1	8.41	1.47E-01	1.17
C11orf24	8.40	1.47E-01	1.20
TMEM133	7.99	1.47E-01	3.62
NHP2L1	7.97	1.47E-01	1.30
GCN1L1	7.60	1.47E-01	1.19
NAE1	7.58	1.47E-01	1.29
IGHMBP2	7.58	1.47E-01	1.13
SUPV3L1	7.56	1.47E-01	1.18
MARS2	7.51	1.47E-01	1.31
LOC100652999	7.46	1.47E-01	1.13
DDX24	7.41	1.47E-01	1.10
NIP7	7.33	1.47E-01	1.29
GPR85	7.26	1.47E-01	3.44
C16orf55	7.13	1.47E-01	1.19
SLITRK3	7.07	1.47E-01	3.23
KIAA0020	7.03	1.47E-01	1.25
SULT1A2	6.90	1.47E-01	12.13
C10orf2	6.89	1.47E-01	1.65
LYSMD2	6.82	1.47E-01	3.07
HLA-DRA	6.81	1.47E-01	12.01
CD79B	6.81	1.47E-01	45.46
METTL2B	6.79	1.47E-01	1.22
MALSU1	6.71	1.47E-01	1.30
CXCL2	6.70	1.47E-01	6.94
BCCIP	6.67	1.47E-01	1.21
NCBP2	6.56	1.47E-01	1.24
MTMR11	6.54	1.47E-01	1.76
NUP155	6.47	1.47E-01	1.31
CCT6A	6.46	1.47E-01	1.33
PAK1IP1	6.45	1.47E-01	1.29
MCAT	6.38	1.47E-01	1.48
SERPINE3	6.37	1.47E-01	7.47
HSPA4	6.33	1.47E-01	1.34
CDC123	6.29	1.47E-01	1.24
PFDN6	6.20	1.47E-01	1.41
PGM5	6.19	1.47E-01	9.92
PPAT	6.16	1.47E-01	1.28

Top 100 genes downregulated in N-R cells			
Description	Score	Feature P	Fold Change
ARL5C	-85.14	2.00E-04	6.78
FLJ41649	-71.01	2.00E-04	1.02
TRIM43	-44.93	2.00E-04	5.63
C1QC	-42.16	2.00E-04	8.36
GRIA3	-39.85	2.00E-04	34.69
MSLN	-34.85	2.00E-04	9.39
GRIN3B	-33.31	2.00E-04	4.83
MYT1L	-28.54	2.00E-04	0.51
ESRP1	-23.38	2.00E-04	29.97
SLC6A9	-19.82	2.00E-04	2.32
XKR5	-18.93	2.00E-04	4.63
GNRH2	-17.59	2.00E-04	53.92
FAM24B-CUZD1	-17.18	2.00E-04	131.96
GAREM	-15.94	2.00E-04	1.54
MAGEB18	-15.89	2.00E-04	2.17
APAF1	-14.84	2.00E-04	1.64
IFI27L2	-14.55	2.00E-04	1.39
ATXN1	-13.08	2.00E-04	1.62
SLC7A11	-12.79	2.00E-04	2.37
ATXN2L	-12.77	2.00E-04	1.37
USP30-AS1	-12.71	2.00E-04	11.06
ATG10	-12.53	2.00E-04	1.34
FZD2	-11.74	2.00E-04	1.63
HNRNPCL1	-10.67	2.00E-04	28.75
CEBPG	-10.57	2.00E-04	1.77
TK2	-10.15	2.00E-04	1.30
LOC100130275	-10.11	2.00E-04	1.91
SLC9A8	-9.75	2.00E-04	1.25
HHAT	-9.60	2.00E-04	1.70
SLC7A3	-9.36	2.00E-04	1.52
BEST1	-9.05	2.00E-04	4.83
CYP21A1P	-8.67	2.00E-04	7.28
SLC35D2	-8.55	2.00E-04	12.57
CARS	-8.45	2.00E-04	1.39
NHLH1	-8.40	2.00E-04	5.07
PTPDC1	-8.32	2.00E-04	1.73
LOC440173	-8.31	2.00E-04	1.63
TAS2R30	-8.26	2.00E-04	9.34
FDFT1	-8.23	2.00E-04	1.27
SPINK4	-8.15	2.00E-04	38.68
SKOR1	-8.10	2.00E-04	9.25
LINC00534	-8.10	2.00E-04	2.41
AXDND1	-7.88	2.00E-04	3.26
CCDC67	-7.88	2.00E-04	1.23
AMER3	-7.69	2.00E-04	2.41
DYNAP	-7.65	2.00E-04	2.51
ARL9	-7.64	2.00E-04	7.68
ZNF439	-7.62	2.00E-04	3.26
ZBED3	-7.62	2.00E-04	1.65
CLEC12A	-7.61	2.00E-04	15.77
CCL27	-7.51	2.00E-04	12.35
RPH3A	-7.50	2.00E-04	15.57
CRNDE	-7.43	2.00E-04	1.46

Description	Score	Feature P	Fold Change
PITRM1-AS1	6.16	1.47E-01	1.36
GOLGA6L6	6.14	1.47E-01	2.65
GLB1L2	6.14	1.47E-01	1.28
CDC20B	6.03	1.47E-01	8.15
DYRK2	6.03	1.47E-01	1.47
RSL1D1	6.03	1.47E-01	1.22
CC2D1A	6.01	1.47E-01	1.20
SLC4A11	6.00	1.47E-01	1.45
HERC2P3	5.98	1.47E-01	2.35
TMEM109	5.97	1.47E-01	1.19
MRM1	5.95	1.47E-01	1.71
IRS2	5.92	1.47E-01	1.24
SET	5.89	1.47E-01	1.18
SLC6A15	5.87	1.47E-01	1.53
FKBP4	5.85	1.47E-01	1.33
ELL2	5.85	1.47E-01	1.40
CAMKV	5.80	1.47E-01	1.50
ARHGAP22	5.78	1.47E-01	1.89
TAGLN2	5.77	1.47E-01	1.22
EIF3M	5.75	1.47E-01	1.21
PAICS	5.70	1.47E-01	1.23
ANKRD13B	5.67	1.47E-01	1.85
CPEB3	5.67	1.47E-01	1.27
CD40LG	5.67	1.47E-01	1.55
VDAC1	5.66	1.47E-01	1.24
FDXACB1	5.64	1.47E-01	1.24
CEBPA	5.62	1.47E-01	2.05
ZNF330	5.60	1.47E-01	1.37
NANP	5.59	1.47E-01	1.27
TMEM138	5.56	1.47E-01	1.33
SPAG17	5.54	1.47E-01	1.67
MRPL50	5.54	1.47E-01	1.37
ANKRD40	5.53	1.47E-01	1.18
WDR36	5.51	1.47E-01	1.15
GNL3L	5.47	1.47E-01	1.30
SUZ12P1	5.46	1.47E-01	1.75
CHORDC1	5.46	1.47E-01	1.44
ZNF582-AS1	5.38	1.47E-01	2.58
TLR6	5.37	1.47E-01	2.39
EN2	5.34	1.47E-01	1.21
PKP2	5.28	1.47E-01	1.56
GMPS	5.25	1.47E-01	1.20
LDHA	5.24	1.47E-01	1.30
SETD6	5.23	1.47E-01	1.28
GPN3	5.21	1.47E-01	1.23
NOLC1	5.20	1.47E-01	1.31
CYCS	5.18	1.47E-01	1.38

Description	Score	Feature P	Fold Change
TRIM66	-7.4	2.00E-04	1.47
ATG4D	-7.32	2.00E-04	1.40
CCDC17	-7.10	2.00E-04	3.31
APLF	-7.02	2.00E-04	2.78
KHDC1	-6.92	2.00E-04	2.11
LOC100505964	-6.87	2.00E-04	1.48
H3F3A	-6.84	2.00E-04	1.31
SLC17A7	-6.82	2.00E-04	2.39
VN1R10P	-6.68	2.00E-04	3.38
FOXC1	-6.67	2.00E-04	1.25
HIST2H2BF	-6.65	2.00E-04	1.78
LINC00324	-6.55	2.00E-04	2.09
SLC1A4	-6.50	2.00E-04	2.66
PTGER2	-6.50	2.00E-04	3.43
LRRFIP2	-6.46	2.00E-04	1.61
FBXW4	-6.42	2.00E-04	1.32
TSIX	-6.4	2.00E-04	1.24
TRAM1	-6.35	2.00E-04	1.44
LOC100499489	-6.31	2.00E-04	2.32
ZNF777	-6.31	2.00E-04	1.42
CDKN1A	-6.29	2.00E-04	2.05
RAMP2-AS1	-6.27	2.00E-04	1.35
ALLC	-6.20	2.00E-04	10.42
KIAA1683	-6.16	2.00E-04	4.87
SPNS2	-6.09	2.00E-04	1.29
KLHL24	-6.06	2.00E-04	1.39
JAG1	-6.06	2.00E-04	2.00
TOP2A	-6.04	2.00E-04	1.08
RFPL3S	-6.04	2.00E-04	3.09
PCDHGC5	-6.03	2.00E-04	5.75
HYI	-6.02	2.00E-04	1.52
PBXIP1	-6.00	2.00E-04	1.38
CLCN4	-5.99	2.00E-04	4.52
GMCL1P1	-5.93	2.00E-04	2.25
FAM189B	-5.87	2.00E-04	1.17
CREB3L4	-5.87	2.00E-04	1.48
CPED1	-5.85	2.00E-04	4.91
IL17RC	-5.82	2.00E-04	1.38
SLC35E1	-5.79	2.00E-04	1.34
GLG1	-5.79	2.00E-04	1.14
NFE2L2	-5.78	2.00E-04	1.60
HMGCS1	-5.72	2.00E-04	1.40
ATF3	-5.69	2.00E-04	1.64
DGCR6	-5.69	2.00E-04	1.34
TNFRSF10B	-5.67	2.00E-04	1.46
TPRN	-5.64	2.00E-04	1.61
VTI1A	-5.64	2.00E-04	1.29

**Supplemental Table 4.2:** Differential expression between E1099K-ALLs and WT-ALLs using RNA-seq as input in the comparative marker selection suite in genepattern. Upregulated genes in NAMPT-resistant (N-R) clones are ranked positively while downregulated genes are ranked negatively.

## Supplementary Table 5.2

### C5 gene sets

NAME	ES	NOM p-val	FDR q-val
SPLICEOSOME	0.64	0	0
RNA_SPLICING	0.57	0	1.59E-03
UNFOLDED_PROTEIN_BINDING	0.65	0	1.42E-03
PROTEIN_FOLDING	0.61	0	1.34E-03
RNA_PROCESSING	0.49	0	2.12E-03
RNA_POLYMERASE_ACTIVITY	0.79	0	1.95E-03
RIBONUCLEOPROTEIN_COMPLEX	0.51	0	2.41E-03
NUCLEOLUS	0.51	0	3.97E-03
RNA_POLYMERASE_COMPLEX	0.73	1.81E-03	5.29E-03
MITOCHONDRIAL_PART	0.49	0	6.34E-03
NITROGEN_COMPOUND_BIOSYNTHETIC_PROCESS	0.67	0	5.86E-03
CELLULAR_COMPONENT_DISASSEMBLY	0.64	0	7.31E-03
DNA_DIRECTED_RNA_POLYMERASE_COMPLEX	0.73	0	7.48E-03
TRANSLATION_REGULATOR_ACTIVITY	0.59	0	1.20E-02
HORMONE_METABOLIC_PROCESS	0.63	0	1.12E-02
NUCLEAR_DNA_DIRECTED_RNA_POLYMERASE_COMPLEX	0.73	0	1.34E-02
SMALL_NUCLEAR_RIBONUCLEOPROTEIN_COMPLEX	0.65	1.77E-03	1.81E-02
MITOCHONDRIAL_LUMEN	0.55	1.76E-03	2.27E-02
RIBONUCLEOPROTEIN_COMPLEX_BIOGENESIS_AND_ASSEMBLY	0.50	0	2.20E-02
TRANSLATION_FACTOR_ACTIVIT_NUCLEIC_ACID_BINDING	0.58	1.72E-03	2.18E-02
MITOCHONDRIAL_MATRIX	0.55	0	2.32E-02
TRANSLATION_INITIATION_FACTOR_ACTIVITY	0.63	6.78E-03	3.25E-02
MRNA_PROCESSING_GO_0006397	0.49	0	3.13E-02
TRANSCRIPTION_FROM_RNA_POLYMERASE_III_PROMOTER	0.65	0	3.53E-02

### H gene sets

NAME	ES	NOM p-val	FDR q-val
HALLMARK_MYC_TARGETS_V1	0.61	0	0
HALLMARK_MYC_TARGETS_V2	0.71	0	0
HALLMARK_OXIDATIVE_PHOSPHORYLATION	0.40	0	4.52E-03
HALLMARK_GLYCOLYSIS	0.33	4.89E-03	1.42E-01
HALLMARK_E2F_TARGETS	0.33	1.16E-02	1.28E-01
HALLMARK_ANDROGEN_RESPONSE	0.36	4.38E-02	1.20E-01
HALLMARK_ALLOGRAFT_REJECTION	0.31	4.07E-02	2.18E-01

Supplementary Table 5.2 Top 25 gene sets enriched in NAMPT-resistant vs NAMPT-sensitive clones.

Supplementary Table 5.3

Rank	score	Perturbation	Cell-line	Dose	Time
1	0.048	Emetine Dihydrochloride Hydrate (74)	MCF7	0.63um	24.0h
2	0.048	BRD-U82589721	ASC	10.0um	24.0h
3	0.048	GDC-0980	A375	0.37um	24.0h
4	0.048	torin-2	A549	10um	24.0h
5	0.048	WYE-125132	A549	1.11um	24.0h
6	0.048	WYE-125132	HME1	10um	24.0h
7	0.04	VALYLTRYPTOPHAN	HCC515	10.0um	24.0h
8	0.04	Kenpallone	VCAP	10.0um	24.0h
9	0.04	FLURANDRENOLIDE	HCC515	10.0um	24.0h
10	0.04	Methyl 2,5-dihydroxycinnamate	HA1E	10.0um	24.0h
11	0.04	trichostatin A	ASC	10.0um	24.0h
12	0.04	NVP-BEZ235	A375	0.04um	24.0h
13	0.04	buparlisib	A549	10um	24.0h
14	0.04	CGP-60474	A549	0.12um	24.0h
15	0.04	ON-01910	A375	1.11um	24.0h
16	0.04	PIK-93	HEPG2	10um	24.0h
17	0.04	torin-1	A375	1.11um	24.0h
18	0.04	GSK-2126458	A375	0.37um	24.0h
19	0.04	ZSTK-474	A375	3.33um	24.0h
20	0.04	AZD-8055	A375	10um	24.0h
21	0.04	WYE-125132	A549	3.33um	24.0h
22	0.04	AZD-5438	HA1E	3.33um	24.0h
23	0.04	PHA-793887	HA1E	10um	24.0h
24	0.04	QL-X-138	HA1E	3.33um	24.0h
25	0.04	GSK-2126458	HEPG2	0.12um	24.0h
26	0.04	AZD-8055	HS578T	1.11um	24.0h
27	0.04	KIN001-043	HS578T	3.33um	24.0h
28	0.04	GSK-2126458	HT29	3.33um	24.0h
29	0.04	torin-1	LNCAP	0.12um	24.0h
30	0.04	GDC-0941	LNCAP	3.33um	24.0h
31	0.032	Clobetasol propionate	HCC515	10.0um	24.0h
32	0.032	ICI-162,846	HCC515	10.0um	6.0h
33	0.032	Aloisine	VCAP	10.0um	6.0h
34	0.032	Tranilast	HA1E	10.0um	24.0h
35	0.032	wortmannin	HA1E	10.0um	6.0h
36	0.032	Flunisolide	HCC515	10.0um	24.0h

Rank	score	Perturbation	Cell-line	Dose	Time
37	0.032	Beclomethasone dipropionate	HCC515	10.0um	24.0h
38	0.032	U-0126	VCAP	10.0um	24.0h
39	0.032	VINBLASTINE SULFATE	HA1E	10.0um	24.0h
40	0.032	DEXAMETHASONE	HCC515	10.0um	24.0h
41	0.032	KETOROLAC TROMETHAMINE	HCC515	10.0um	24.0h
42	0.032	IOPANIC ACID	HCC515	10.0um	24.0h
43	0.032	SKF 96365 hydrochloride	HCC515	10.0um	24.0h
44	0.032	PENICILLIN G POTASSIUM	HCC515	10.0um	24.0h
45	0.032	RIZATRIPTAN BENZOATE	HCC515	10.0um	24.0h
46	0.032	NOCODAZOLE	PC3	10.0um	24.0h
47	0.032	FENOTEROL HYDROBROMIDE	VCAP	10.0um	6.0h
48	0.032	HYDROCORTISONE HEMISUCCINATE	A375	10.0um	6.0h
49	0.032	calyculin A	PC3	10.0um	24.0h
50	0.032	PI 103 hydrochloride	A375	11.1um	24.0h

**Supplementary Table 4.4** Small molecule treatments that mimic gene expression in NAMPT-resistant clones as analyzed in LINCS.

## Appendix A

### List of antibodies used in study

<b>Anti-</b>	<b>vendor</b>	<b>product number</b>
H3	cell signaling	4499
Btn-H3	cell signaling	5748
H3K36me1	abcam	ab9048
H3K36me2	cell signaling	2901
H3K36me3	cell signaling	4909
H3K27me1	cell signaling	7693
H3K27me2	cell signaling	9728
H3K27me3	cell signaling	9733
H3K27Ac	abcam	ab177178
H3K4me2	cell signaling	9725
H3K4me3	cell signaling	9727
H3K9me2	cell signaling	9753
H3K9me3	abcam	ab8898
H3K9Ac	millipore	ABE18
H4K20me3	cell signaling	5737
p-H2AX	cell signaling	9718
TBP	abcam	ab51841
NSD2	abcam	ab75359
HDAC1	abcam	ab31263
HDAC2	abcam	ab124974
NAMPT	cayman chem	10813
beta-Actin	abcam	ab75186
mouse (IR-dye)	LI-COR	926-32212
rabbit (IR-dye)	LI-COR	926-32213
Streptavidin (IR-dye)	LI-COR	926-68031
mouse (HRP)	cell signaling	7076

## Appendix B

List of cell lines from CCLE used in analysis

Cell Line	Lineage	E1099K
BDCM	B-ALL	-
KASUMI2	B-ALL	-
KOPN8	B-ALL	-
MHHCALL2	B-ALL	-
MHHCALL3	B-ALL	-
MHHCALL4	B-ALL	-
MUTZ5	B-ALL	-
NALM19	B-ALL	-
SUPB15	B-ALL	-
697	B-ALL	-
NALM6	B-ALL	-
REH	B-ALL	-
ALLSIL	T-ALL	-
DND41	T-ALL	-
LOUCY	T-ALL	-
MOLT4	T-ALL	-
P12ICHIKAWA	T-ALL	-
PEER	T-ALL	-
PF382	T-ALL	-
SUPT1	T-ALL	-
TALL1	T-ALL	-
JURKAT	T-ALL	-
KE37	T-ALL	-
MOLT16	T-ALL	-
RCHACV	B-ALL	+
RS411	B-ALL	+
SEM	B-ALL	+
HPBALL	T-ALL	+
MOLT13	T-ALL	+
RPMI8402	T-ALL	+

Cell Line	Lineage	t(4:14)+
AMO1	MM	-
COLO677	MM	-
EJM	MM	-
HUNS1	MM	-
JJN3	MM	-
KARPAS620	MM	-
KE97	MM	-
KHM1B	MM	-
KMM1	MM	-
KMS12BM	MM	-
KMS20	MM	-
KMS21BM	MM	-
KMS27	MM	-
L363	MM	-
MOLP2	MM	-
MOLP8	MM	-
PCM6	MM	-
RPMI8226	MM	-
SKMM2	MM	-
U266B1	MM	-
MM1S	MM	-
KMS11	MM	+
KMS18	MM	+
KMS26	MM	+
KMS28BM	MM	+
KMS34	MM	+
LP1	MM	+
NCIH929	MM	+
OPM2	MM	+