Vitamin C induced epigenomic remodelling in IDH1 mutant acute myeloid leukemia

by

Matthew Mingay

BSc., The University of Western Ontario, 2013

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE and POSTDOCTORAL STUDIES

(MICROBIOLOGY & IMMUNOLOGY)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2017

Abstract

The genomes of myeloid malignancies are characterized by epigenomic abnormalities. Heterozygous, inactivating TET2 mutations and neomorphic IDH mutations are recurrent and mutually exclusive in acute myeloid leukemia genomes. Ascorbic Acid (vitamin C) has been shown to stimulate the catalytic activity of TET2 *in vitro* and thus we sought to explore its effect in a leukemic model expressing $IDH1^{R132H}$. Vitamin C treatment induced a reduction in cell proliferation and an increase in the expression of genes involved in leukocyte differentiation in $IDH1^{R132H}$ expressing cells. Genome-wide assessment of 5mC and 5hmC revealed a set of vitamin C induced differentially methylated regions (DMRs), many of which displayed demethylation at myeloid enhancers and regions containing binding elements for the hematopoietic transcription factors RUNX1 and PU.1. We observed a significant loss of PU.1 DNA binding and an increase of RUNX1 binding that was accompanied by demethylation at RUNX1 binding sites within enhancers. Genome-wide profiling of histone modifications revealed a negative correlation between H3K4 methylation and DNA methylation at vitamin C induced DMRs. In addition, vitamin C induced an increase in H3K27ac flanking sites bound by RUNX1. Taken together our results suggest that vitamin C is able to stimulate epigenetic remodelling of transcription factor binding sites and regulatory elements to drive differentiation in a leukemic model.

Preface

I was the lead investigator for this project where I was responsible for all major areas of concept formation, technical development, data collection and analysis, as well as the majority of manuscript composition. Michael Heuser and Anuhar Chaturvedi obtained, prepared and characterized the cell lines used in this study. Michelle Moksa and Martin Hirst oversaw the generation of DNA libraries and sequencing data and Qi Cao performed immunofluorescence experiments. Tony Hui and Linda Jackson helped construct sequencing libraries and Misha Bilenky and Alireza Moussavi contributed to the processing of this data. Martin Hirst was the supervisory author on this project and was involved throughout the project in concept formation and manuscript edits.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2-OGDD – 2-OG dependent dioxygenase
5caC – 5-carboxylcytosine
5fC – 5-formylcytosine
5hmC – 5-hydroxymethylcytosine
5mC – 5-methylcytosine
AML – Acute Myeloid Leukemia
CGI – CpG island
ChIP-seq – chromatin-immunoprecipitation sequencing
CLP – common lymphoid progenitor
deMR- demethylated region
DHS – DNAse hypersensitivity site
DMR – differentially methylated region
DNA – deoxyribonucleic acid
DNMT – DNA methyltransferase
DxMR – intersection of deMR and iHR
FACS – fluorescence activated cell sorting
GMP – general myeloid progenitor
hmeDIP- hydroxyl-methylated DNA immunoprecipitation
HSC – Hematopoietic stem cell
iHR – increased hydroxymethylated region
iMR – increased methylated region
K – lysine
K-S test – Kolmogorov-Smirnov test
meDIP - methylated DNA immunoprecipitation
MEP – megakaryocyte-erythroid progenitor
mESC – mouse embryonic stem cell
R-2HG – *R*-2-hydroxyglutarate
ROS – reactive oxygen species
S-2HG – *S*-2-hydroxyglutarate
ST-HSC – short-term hematopoietic stem cell
TET – Ten-eleven translocation enzymes
TSS – transcriptional start site
Vitamin C – Ascorbic Acid
VitC - 2-phosphate ascorbic acid
WGBS – whole genome bisulfite sequencing

Acknowledgements

I offer my enduring gratitude to the faculty, staff and my fellow students at the University of British Columbia who have inspired my work in this field.

I owe special thanks to Dr. Martin Hirst and Ms. Michelle Moksa for enhancing my knowledge and appreciation of science and providing guidance and answers to my endless questions. I also thank the various members of my lab and thesis committee for providing guidance and constructive discussions throughout my studies.

This work was supported by grants from the Terry Fox Research Institute Program Project (Grant #1039) to Martin Hirst and R. Keith Humphries and the Canadian Cancer Society Research Institute (Grant # 703489) to Martin Hirst.

Special thanks are owed to my parents, who have supported me throughout my years of education, both morally and financially.

Chapter 1: Introduction

1.1 A brief history of epigenetics

Nature's engineering elegancy is perhaps best exemplified by the evolution of its information storage and transfer system encoded within polymeric deoxy-ribonucleic acid (DNA) molecules. In its simplest form, four organic compounds, known as bases or nucleotides, comprise DNA. The sequence of these bases form a cypher in which the majority of the information required for life is encoded. The process of DNA replication produces an identical copy of the original template DNA that is passed on to a cell's progeny during propagation. The definition of the term "epigenetics" reflects an implication of this fundamental biological process: the understanding that cellular identity, driven by gene expression patterns, differs drastically between genetically identical cell types within multicellular organisms. The word "epigenetics", in which the prefix epi- originates from the latin word for "above", was coined by Waddington (1) in 1942 to describe the poorly understood, highly regulated processes guiding the development of a mature organism from a fertilized, diploid zygote. By 1996, Riggs and Porter described epigenetics as "the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence" (2).

Within eukaryotic nuclei, linear DNA is wrapped around core histone octamers to form nucleosomes, chains of which are further packaged into 30 nm fibres that are arranged into a macromolecule referred to as chromatin. This three dimensional packaging of DNA offers an additional layer of complexity and opportunity for information storage beyond what is encoded in the underlying sequence of bases. The inclusion of distinct chemical groups at specific amino acids within histone subunits or directly to the DNA are known as epigenetic modifications and

can alter chromatin conformation and thus accessibility to the DNA-binding proteins that drive gene expression programs.

1.2 Epigenetic modifications

1.2.1 DNA methylation

5-methylcytosine (5mC) is one of the most widely studied epigenetic modifications. In vertebrates, 5mC is generated by cytosine methylation, catalyzed by the highly conserved DNA methyltransferase enzymes (DNMTs) and provides a heritable, reversible mechanism for cells to functionally mark genomic DNA. The vast majority of cytosine methylation occurs symmetrically at cytosines followed by a (3') guanine (G), a sequence pattern known as a CpG dinucleotide. DNA methylation is made heritable in the CpG context by the enzyme DNMT1, also known as the "maintenance" methyltransferase, which leverages the inherent sequence and methylation symmetry of CpGs within complementary, double stranded DNA to copy methylation patterns. During DNA replication, UHRF1 recruits DNMT1 to regions where the only template strand is methylated (3) to methylate the newly synthesized strand, conserve methylation symmetry and thus allow for the conservation of methylation patterns in the daughter cell and subsequent progeny. The discovery of the ten-eleven translocation (TET) family of proteins, which subsequently oxidize 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) intermediates (4, 5), provides a mechanism for active, replication-independent DNA demethylation, a process previously thought to be restricted to passive loss during replication. 5-caC, the terminal product of 5mC oxidation, is recognized and replaced with un-methylated cytosine by thymine DNA glycosylase-mediated base excision repair during active DNA demethylation (6). The oxidized derivatives of 5mC can also facilitate passive methylation by inhibiting DNMT1 binding during DNA replication (6).

Recent advances in sequencing technology have produced maps of genome-wide methylation patterns at single nucleotide resolution and have shown that active regulatory sequences, such as enhancers and promoters, are generally un-methylated (7-9). The presence of methylation within the binding sites of many transcription factors reduces their binding affinity and restricts transcriptional activation (10). Interestingly, while spontaneous deamination of cytosine produces uracil, which is efficiently removed by uracil-DNA-glycosylase, spontaneous deamination of 5mC results in thymine, a proper genomic base (11). This reaction leads to increased frequency of C to T transitions and, over time, genome-wide CpG depletion. CpG islands (CGI) are mammalian genomic elements that maintain ancestral CpG density due to positive selection against transition events and thus are thought to have a functional role in the genome. A subset of CGIs exist within gene promoter regions and, in this context, are generally maintained in an un-methylated state. The presence of 5mC within CGIs, however, is associated with gene repression that can be stable or dynamic at different genes during cellular division and development. Failure to accurately control the distribution of 5mC has been repeatedly shown to disturb development and contribute to disease in model organisms and humans, highlighting the importance of this fundamental regulatory mechanism.

1.2.2 Histone modifications

Epigenetic information is also encoded within covalent chemical modifications to the histones that package DNA. These chemical modifications primarily affect histone H3, one of the four core histone subunits, and are often covalently attached to lysine (K) residues within the N-terminal tail of this histone (H3K-). Genome-wide mapping of modified histones by chromatin-immunoprecipitation sequencing (ChIP-seq) (12), has revealed correlations between specific epigenetic marks and regulatory features of the genome. For example, H3K4me1 is

associated with enhancer elements while H3K4me3 marks gene promoters, where its density correlates with promoter activity (13). The addition of negatively charged acetyl groups to lysine residues at position 27 of the H3 histone tail (H3K27ac) by acetyl-transferases produces a local negative charge that repels the negatively charged DNA backbone leading to a more relaxed chromatin structure and the permission of DNA binding proteins. Thus, H3K27ac represents an activating epigenetic mark that is found co-existing with H3K4me1 and H3K4me3 to promote the activity of enhancers and promoters, respectively (13).

Lysine tri-methylation also occurs at position 9 (H3K9me3), 27 (H3K27me3) and 36 (H3K36me3) of the H3 histone tail. H3K9me3 is a silencing mark that marks both constitutive and facultative heterochromatin (14). H3K27me3 is also a silencing mark that localizes to inactive gene promoters, especially those silenced during development, and plays an important role in X chromosome inactivation (15). H3K36me3 is found primarily within gene bodies where it is correlated with gene expression (16) and is thought to promote methylation at CpGs within actively transcribed genes by recruiting DNMT3b, a *de novo* methyltransferase (17).

1.3 Epigenetics and hematopoiesis

Hematopoiesis is a highly conserved process during which quiescent hematopoietic stem cells (HSCs), through a hierarchical differentiation cascade, reconstitute all cell types of the blood while maintaining multi-potent and self-renewal capacities. The equilibrium between multi-potency and self-renewal in HSCs is maintained by intrinsic and extrinsic mechanisms including niche-associated factors, transcription factors, signal transduction pathways and epigenetic modifiers (18). CpG methylation is a critical regulatory mechanism during these processes and aberrant DNA methylation patterns are frequently observed in malignancies, especially those of hematopoietic origin. Indeed deletion of DNMT1 (19) and DNMT3a (20)

leads to persistent self-renewal and a differentiation block in hematopoietic stem and progenitor cells, emphasizing the importance of this epigenetic mark in cell fate decisions.

1.4 Epigenetics and leukemogenesis

Acute myeloid leukemia (AML) is the most common acute leukemia in adults, accounting for approximately 80% of cases in this group, with 3-5 cases per 100,000 people in the United States (21). AML arises as a result of the abnormal proliferation and differentiation of a clonal population of myeloid stem cells, resulting in the accumulation of poorly differentiated myeloid cells within the bone marrow, peripheral blood and occasionally peripheral organs (22). In the case of AML, induction therapy, defined as the first in a series of therapeutic measures taken to treat a disease, consists of a chemotherapeutic regimen that is comprised of 7 days of continuous infusion cytarabine with 3 days of anthracycline. Although hypomethylating agents such as azacitidine and decitabine, which irreversibly inhibit DNMT1 and promote DNA hypomethylation, have led to marginal increases in overall patient survival within elderly patients (22) their clinical efficacies have been generally disappointing. Even though about 60-80% of patients with *de novo* AML will achieve complete remission with induction therapy, residual disease often persists and relapse will inevitably occur if treatment is discontinued. It is also important to consider the severe side effects of the aforementioned chemotherapeutic agents, which include: bone marrow/immune suppression, cardiotoxicity, neurotoxicity, vomiting, diarrhea, liver problems and general malaise. These side effects contribute to substantially reduced quality of life in patients and, in combination with the high instance of treatment failure and disease relapse, warrant investigation into alternative, targeted therapies in genetically stratified cohorts.

Recent large-scale efforts to characterize recurrent mutations in AML revealed a significant proportion of patients harboring inactivating mutations to Ten-Eleven Translocation 2 (TET2; 8.5%), DNA methyltransferase 3a (DNMT3a; 22%) and neomorphic mutations to Isocitrate Dehydrogenase 1 and 2 (IDH1/2; 19.5%) (23, 24). Both TET2 and IDH1/2 normally contribute to DNA methylation homeostasis and *de novo* neomorphic IDH1/2 mutations promote aberrant DNA methylation through increased production of 2-hydroxyglutarate (25), which inhibits the hydroxylation of 5mC, the reaction catalyzed by the TET enzymes during active demethylation (26). In support of this notion, multiple reports have shown that the *R* enantiomer of 2-hydroxyglutarate (*R*-2HG), specifically, initiates leukemogenesis in a reversible manner and contributes to DNA hypermethylation (27, 28). AML patient genomes also demonstrate mutual exclusivity in mutations to IDH and TET2; and TET2-mutant AMLs show an overlapping promoter hypermethylation signature with IDH1/2-mutant AMLs (26). Survival data associated with TET2 and IDH1/2 mutations in AML patients are inconsistent and no prognostic significance for these mutations has been conclusively established (29, 30). Epigenetic alterations are theoretically reversible and inhibitors targeting some of these epigenetic pathways are currently being applied in the clinic. However, the existence of therapeutic compounds able to activate epigenetic pathways has yet to be explored in the context of acute myeloid leukemia.

1.5 The dichotomy between vitamin c and IDH1/TET2 mutations

Ascorbic acid (vitamin C) is a water-soluble essential nutrient and common cell culture medium supplement. In the human body, vitamin C is highly concentrated in immune and brain cells where it acts as an electron donor and maintains iron in the Fe(II) state, a requirement for the catalytic activity of 2-oxoglutarate-dependent dioxygenases (2-OGDD) (31). The importance of vitamin C in human health is exemplified in the disease scurvy, where its absence impairs

collagen formation by rendering collagen prolyl hydroxylase, a 2-OGDD, catalytic inactive (32). Most mammals are able to synthesize ascorbic acid endogenously. Primates, guinea pigs and bats, however, harbour evolutionarily conserved mutations in the *GULO* gene, which encodes the last enzyme in the ascorbic acid biosynthetic pathway, rendering ascorbic acid an essential nutrient in these species. In addition to its requirement as a cofactor, *in vitro* studies have provided evidence of a physical interaction between vitamin C and the catalytic domain of TET2 that enhances the enzymatic oxidation of 5mC to 5hmC (33). The biological significance of the epigenetic modulation induced by vitamin C has been illustrated through its ability to improve induced pluripotent stem cell generation (34) and induce a blastocyst like state in mouse embryonic stem cells (mESC) (35) by promoting demethylation of H3K9me3 and 5mC, respectively. In addition to TET DNA demethylases, many histone demethylases are also 2-OGDD enzymes. These include JmjC histone demethylases, many of which are putative tumor suppressors that are responsible for the demethyation of H3K36, H3K27, H3K9 and H3K4. *R*-2-HG, the oncogenic metabolite produced by $IDH1^{R132H}$ expression, has been shown to inhibit JmjC histone demethylases including KDM2A, KDM4A, KDM4C, and KDM7A (25). As 2-OGDDs, these enzymes require vitamin C for optimal activity (36, 37) but, unlike the TET enzymes, catalytic enhancement by vitamin C has yet to be explored.

Multiple lines of evidence have suggested that vitamin C, in its native form, is toxic in culture through the formation of extracellular H_2O_2 (38, 39). This excess extracellular H_2O_2 participates in the Fenton Reaction and reacts with Fe(II) to generate Fe(III) and hydroxyl radicals and reactive oxygen species (ROS). The ability of vitamin C to recycle Fe(III) to Fe(II) promotes this reaction and leads to the extracellular accumulation of toxic ROS and cell death (40). 2-phosphate ascorbic acid was developed as an oxidatively stable form

of vitamin C that is dephosphorylated and converted to ascorbic acid during transport across the cell membrane(41) and is used frequently as an alternative to ascorbic acid in cell culture experiments.

Considering the importance of DNA methylation homeostasis in normal hematopoiesis and the functional dichotomy between vitamin C and $IDH1^{R132H}$ on the pathways which regulate it, we hypothesize that vitamin C will activate the same components of the DNA demethylation pathway that are recurrently disturbed in AML to reinstate DNA methylation homeostasis and facilitate proper cellular differentiation.

Chapter 2: Results

2.1 Vitamin c modulates phenotype in a leukemic model expressing IDH1^{R132H}

2.1.1 Induction of 5hmC by vitamin c

To explore the effect of vitamin C in a leukemic context we utilized a previously established model in which murine bone marrow cells were immortalized by the overexpression of *HOXA9* (42) and subsequently transfected with a retroviral vector expressing either *IDH1*^{WT} (IDH1^{WT} cells) or *IDH1*^{R132H} (IDH1^{R132H} cells) (43). In this model, *IDH1*^{R132H} expression increases the ratio of *R*-2-HG:*S*-2-HG (43) and provides an environment to explore the effects of vitamin C in the context of this so-called "oncometabolite". We utilized 2-phosphate L-ascorbic acid (vitC), which, unlike native ascorbic acid, remains oxidatively stable under standard cell culture conditions (44). Compared to IDH1^{WT} cells, of which 74% stained positive for 5hmC (mean intensity = 5), IDH1^{R132H} mutant cells had low 5hmC levels with only 4% of cells positive for 5hmC (mean intensity = 0.151) (Figure 1 and Table 1). Interestingly, vitC treatment (0.345mM; the concentration used throughout, unless otherwise indicated) in IDH1^{R132H} cells induced a marked increase in 5hmC signal after 12 hours (mean intensity = 3.84) and resulted in 84% of cells staining positive for 5hmC (Figure 2 and Table 1). Since the TETs are the only enzymes known to catalyze 5hmC formation, these results are consistent with the notion that increased *R*-2HG levels, resulting from IDH1^{R132H} expression, lead to reduced TET activity and that vitamin C stimulates TET activity.

2.1.2 Vitamin c induces a reduction in cell proliferation

Daily vitC treatment (0.345mM) over the course of 10 days induced a significant reduction in the proliferation rate of IDH1^{R132H} cells and to a lesser extent IDH1^{WT} cells (Figure

3). Titration experiments at lower vitC concentrations revealed a striking dose dependency and robust effectiveness (Figure 4). This reduction in cell proliferation likely results from attenuation of the effects of $IDH1^{R132H}$ expression rather than vitC induced cytotoxicity.

2.1.3 Vitamin c enhances the expression of cell surface antigens associated with mature myeloid cell types

Decreased cell proliferation is a hallmark of terminal differentiation and thus we sought to investigate changes in cell surface antigens, a measure of cellular identity, between cell types and treatments. The Sca1 surface antigen is expressed at high levels in early hematopoietic stem and progenitor cells (45) whereas Mac-1 is expressed in most myeloid cells (46) and marks mature granulocyte and macrophage cells when co-expressed with Gr-1 or F4/80, respectively (47). We used fluorescence-activated cell sorting (FACS) to identify untreated cells expressing common hematopoietic cell surface antigens and observed that, compared to $IDH1^{WT}$ cells, fewer $IDH1^{R132H}$ cells expressed $Gr\text{-}1^{+}/Mac\text{-}1^{+}$ (23% reduction) and $Mac1^{+}/F4.80^{+}$ (25% reduction), suggesting that $IDH1^{R132H}$ expression promotes less mature myeloid populations (Figure 5). Interestingly, upon treatment with vitC we observed an increase in the proportion of these populations ($Mac1^{+}/F4.80^{+}$ (25% gain); $Gr\text{-}1^{+}/Mac\text{-}1^{+}$ (11% gain)) and the myeloid committed $Gr\text{-}1^{-}/Mac\text{-}1^{+}$ (12% gain) cells in $IDH1^{R132H}$ cells. VitC also decreased the percentage of $Sca1^{+}$ early progenitor populations to a similar degree in both cell types (8% and 9% reduction in $IDH1^{R132H}$ and $IDH1^{WT}$, respectively) (Figure 5). Improper maturation of myeloid stem and progenitor cells is a characteristic feature of AML and can be driven by disturbances in gene expression patterns and DNA methylation homeostasis through reduced TET activity (48). The increased proportion of cells expressing mature myeloid surface antigens

after vitC treatment in IDH1^{R132H} cells suggests that vitC is able to promote the differentiation of myeloid stem and progenitor cells.

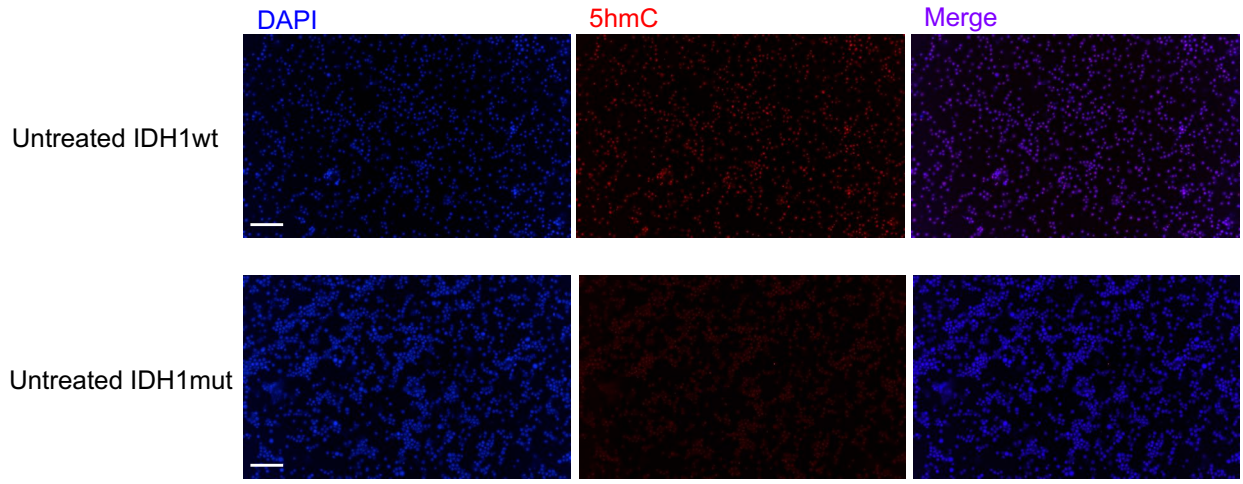

Figure 1. 5hmC immunofluorescence images in untreated IDH1^{WT} (top) and untreated IDH1^{R132H} cells (bottom). Scale bar, 200 µm.

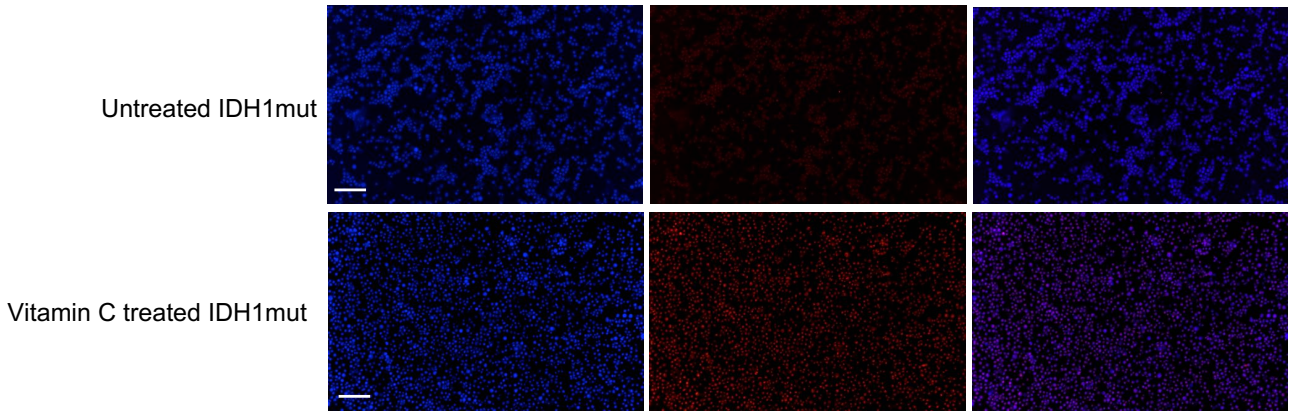

Figure 2. Immunofluorescence for 5hmC in IDH1^{R132H} cells 12h after growth in the absence (top) or presence of vitamin C (0.345mM). Scale bar, 200 µm.

Cell Type	% 5hmC positive	Mean Fluorescence Intensity	VitC Dosage
IDH1 WT	74%	5	0mM
IDH1 WT	86%	4.5	0.345mM
IDH1 R132H	4%	0.15	0mM
IDH1 R132H	84%	3.8	0.345mM

Table 1. The percentage of cells staining positive for 5hmC and the intensity of 5hmC signal derived from the quantitative analysis of immunofluorescence experiments in untreated and vitamin C treated (12h) $IDH1^{WT}$ and $IDH1^{R132H}$ cells.

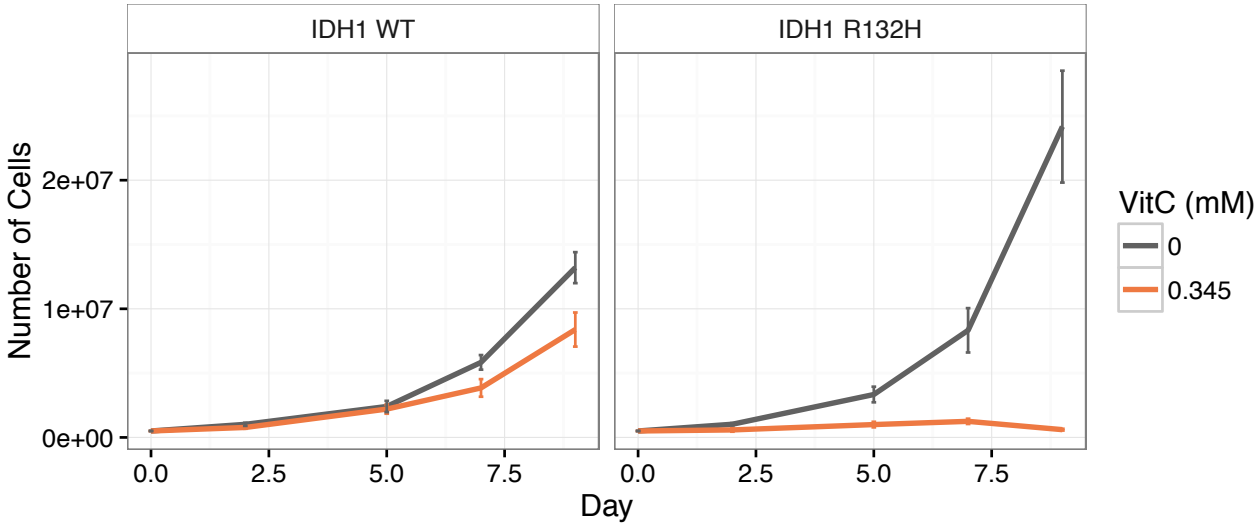

Figure 3. A growth curve for cell counts, represented as mean ± SD, for untreated (gray) or vitamin C treated (0.345mM) $IDH1^{WT}$ and $IDH1^{R132H}$ cells over 9 days.

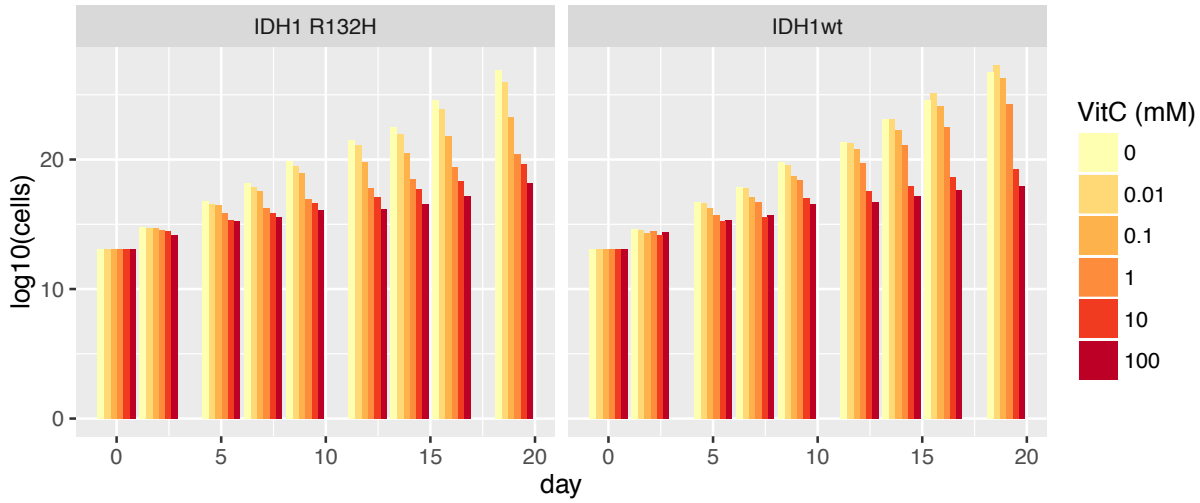

Figure 4. Cell counts represented by log10(cell number), for IDH1^{R132H} (left) and IDH1^{WT} (right) cells over 20 days at decreasing vitamin C concentrations.

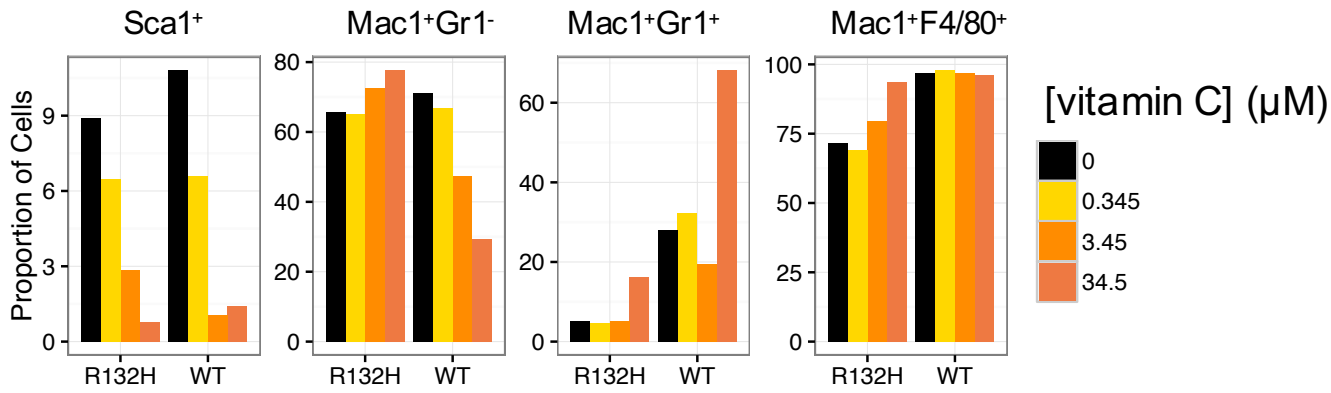

Figure 5. The proportion of cells positive for different marks or combinations of marks after FACS sorting untreated (black) or vitamin C treated cells after 19 days of growth at different vitamin C concentrations.

2.2 Vitamin c induces the expression of a myeloid-specific gene signature

To begin to investigate molecular changes induced by vitC we performed mRNA-seq on IDH1^{R132H} cells and observed 450 up-regulated and 221 down-regulated genes upon vitC treatment (RPKM > 0.01; FDR <0.05) (Figure 6) using a custom java program developed in-house. Up-regulated genes were significantly enriched in Gene Ontology (GO) terms (47 terms,

Bonferroni p.value < 0.01). The top enriched GO terms for up-regulated genes included positive regulation of leukocyte differentiation, chemotaxis and migration and negative regulation of cell proliferation (Figure 7). We also observed a significant association with KEGG terms for hematopoietic cell lineage, cytokine-cytokine receptor interaction and Toll-like receptor signalling (Figure 8). We observed the upregulation of multiple cytokines and cytokine receptors, the expression of which are known to regulate self-renewal and multi-potency in HSCs (49). Multiple genes implicated in myeloid differentiation were upregulated including *CSF2RA* (50), *CSF1R*(51), *TLR2*, *CCL3* and *MYADM* (52) and the genes *MEIS1*(53) and *CDK1*(54), which are overexpressed in HSCs and AML, were among those downregulated. These observations were validated across 3 additional mRNA-seq libraries generated from biological replicates (Replicates 2-4; Figure 9). Consistent with the FACS data, vitC also increased the expression of genes encoding cell surface markers that define mature myeloid cells and reduced the expression of c*KIT,* a gene expressed highly in hematopoietic stem and progenitor cells, specifically (55). These observations were also confirmed in the biological replicate mRNA-seq libraries (Figure 10). Surprisingly the *Sca1* gene, which is usually expressed in hematopoietic stem cells, showed a slight increase in expression upon vitC treatment. In contrast to upregulated genes, genes downregulated upon vitC treatment revealed no significantly enriched GO terms. Collectively, these results suggest that vitC activates the expression of myeloid lineage specific genes that are usually repressed in hematopoietic stem and progenitor cells.

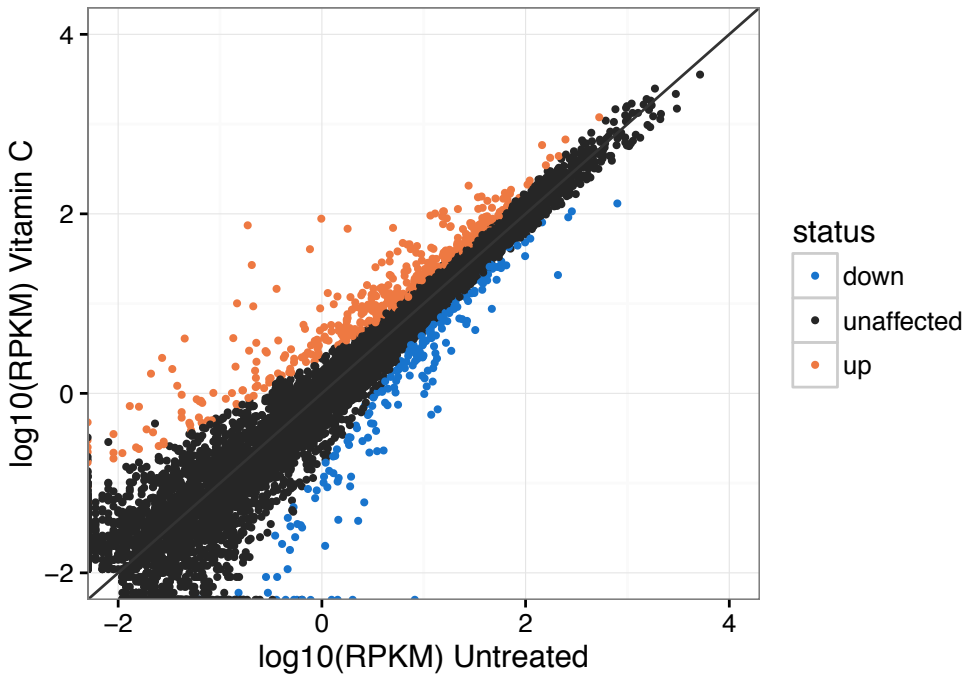

Figure 6. Expression values, represent as log10 (rpkm), of each gene in mm10 Ensembl v71 (RPKM > 0.01) for untreated (x-axis) and vitamin C treated (y-axis) cells. Upregulated (up), downregulated (down) and unaffected genes (RPKM > 0.01; FDR <0.05) are represented by orange, blue and black dots, respectively.

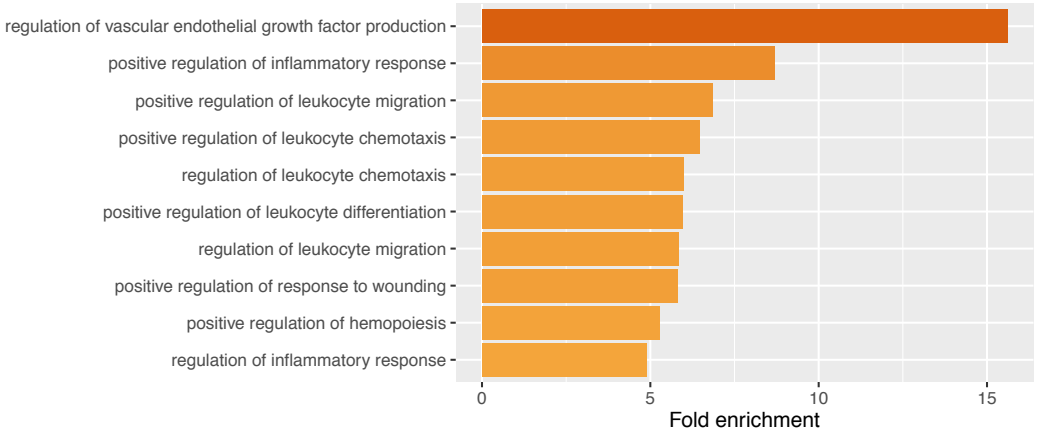

Figure 7. The top 10 gene ontology results represented by hyper-geometric fold enrichment for upregulated genes upon vitamin C treatment.

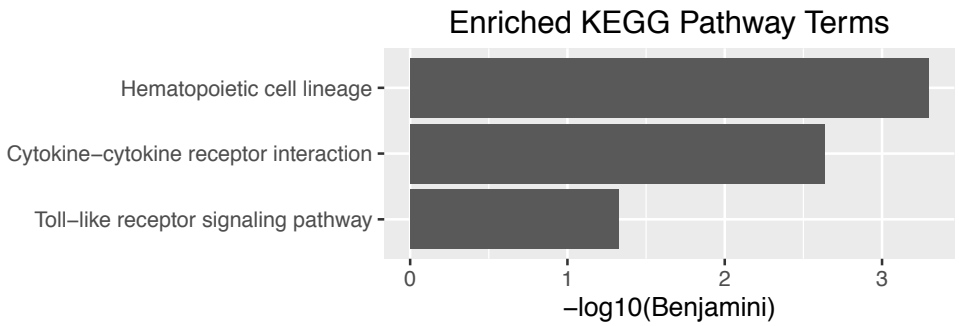

Figure 8. The 6 most enriched KEGG pathways (-log10(Benjamini q-value)) for genes up-regulated by vitC in IDH1^{R132H} expressing cells.

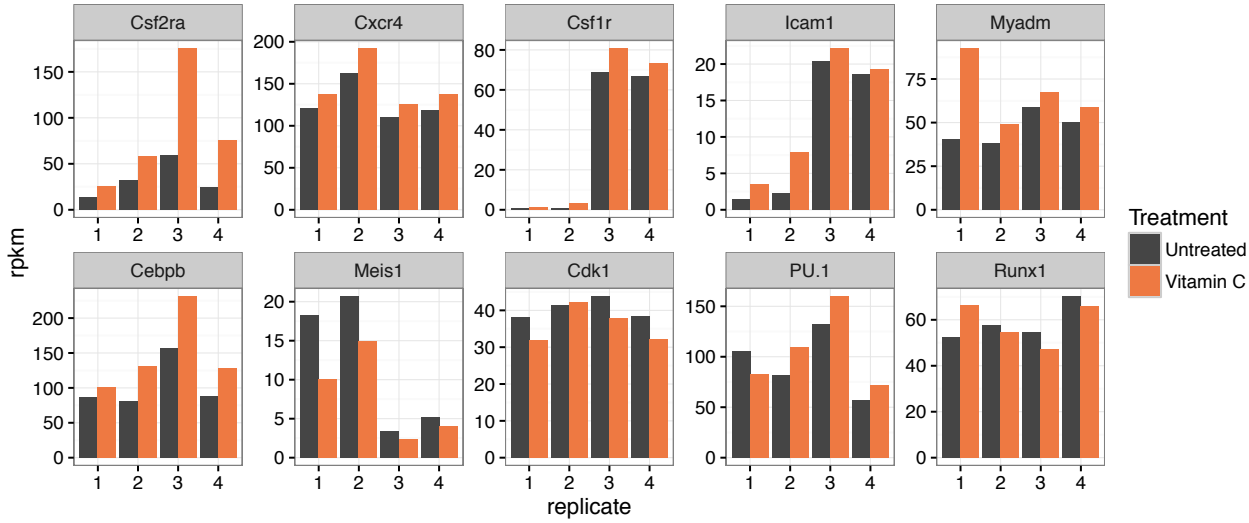

Figure 9. Expression values (RPKM) from 4 biological replicate RNA-seq libraries for several genes of interest in untreated (black) and vitC (orange) treated IDH1^{R132H} cells.

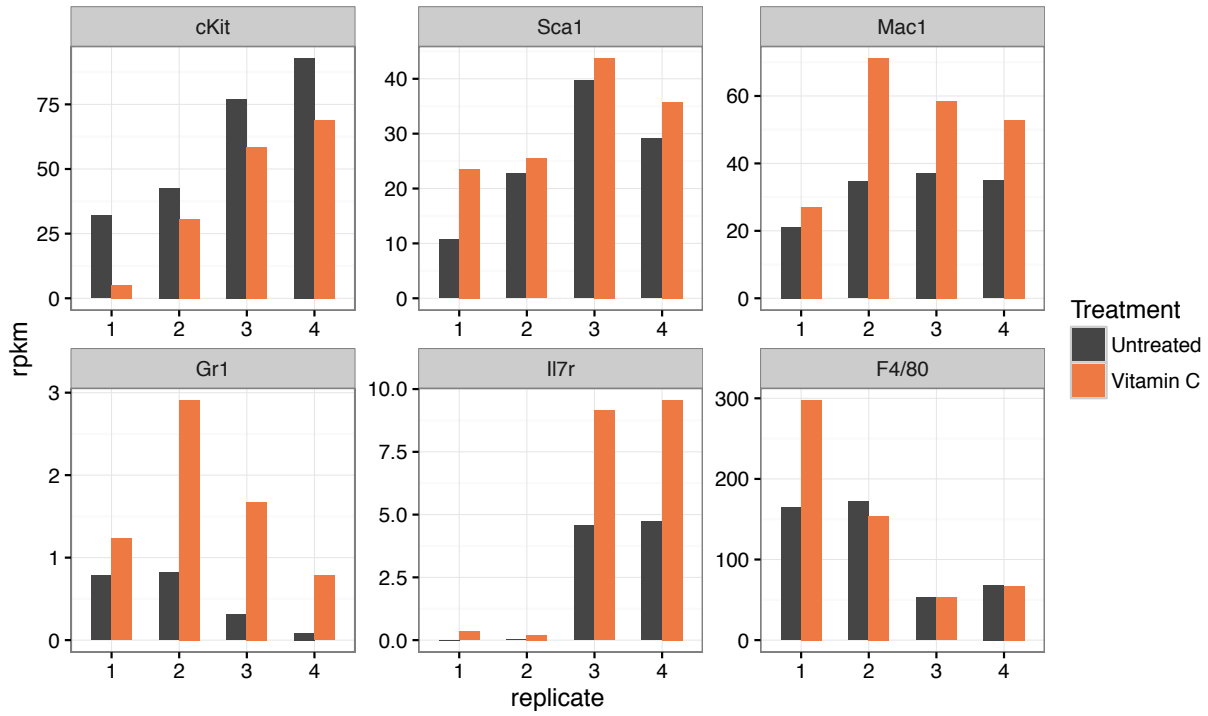

Figure 10. RNA-seq expression values (RPKM) for the genes encoding cell surface markers commonly used to separate hematopoietic cell types in untreated (black) and vitC treated IDH1^{R132H} cells across 4 biological replicates.

2.3 Vitamin c stimulates methlyome remodeling in the context of IDH1^{R132H} expression

2.3.1 Qualitative vitamin c induced changes in 5mC and 5hmC are associated with hematopoiesis and AML

Since IDH1^{R132H} expression leads to increased production of *R*-2-HG (43), a TET2 inhibitor (56), and vitC stimulates TET activity (33), we hypothesized that vitC may induce changes in the genome-wide distribution of DNA methylation in our model. To explore qualitative changes in DNA methylation upon vitC treatment, genomic DNA was extracted from untreated and vitC treated IDH1^{R132H} and IDH1^{WT} cells and subjected to methylated DNA or hydroxyl-methylated DNA immunoprecipitation sequencing (meDIP-seq or hmeDIP-seq, respectively) (Table 2). MACS (57) was used to identify differentially methylated regions

(DMRs; $p<1e^{-4}$) and the R package called MEDIPS (58) was used to calculate significant changes ($p<0.01$) in meDIP-seq and hmeDIP-seq signal between treatments in 500bp windows genome wide (Table 3). Genomic regions that lost 5mC in IDH1^{R132H} cells upon vitC treatment (meDIP deMRs; 3789 regions) showed increased 5hmC signal flanking their centre ($p \sim 0$) (Figure 11, right), a pattern consistent with the notion that TET2-mediated hydroxylation of 5mC to 5hmC drives site-specific demethylation. Interestingly, we observed a similar 5hmC pattern flanking the centre of regions that gained methylation in untreated IDH1^{R132H} cells compared to untreated IDH1^{WT} cells (IDH1^{R132H} iMRs; 4182 regions; Figure 11, left). This suggests that some of the hypermethylation incurred as a result of *IDH1^{R132H}* expression and increased *R*-2-HG production can be attenuated by vitamin C, potentially through TET catalytic enhancement. VitC induced deMRs (n=3928) and iMRs (n=3964) identified by meDIP-seq in IDH1^{WT} cells were unique (<1% overlap with IDH1^{R132H} equivalents) and devoid of statistically significant gene ontology associations related to hematopoiesis or cancer suggesting that the effect of vitC is specific to IDH1^{R132H} cells.

2.3.2 Quantitative analysis of vitamin c induced changes in 5mC

Having established specific qualitative differences in 5mC and 5hmC upon vitamin C treatment we sought to quantitatively assess 5mC changes in response to vitC by whole genome shotgun bisulfite sequencing (WGBS) of genomic DNA extracted from untreated and vitC treated IDH1^{R132H} cells. We identified 6496 deMRs and 8374 regions with increased methylation (iMRs) induced by vitC in IDH1^{R132H} cells by WGBS analysis (Table 3). In accordance with the (h)meDIP-seq experiments we observed increased 5hmC (hmeDIP) signal and decreased 5mC (meDIP) levels within WGBS deMRs (Figure 12). In agreement with our meDIP based analysis, WGBS deMRs were associated with genes implicated in myeloid

differentiation and hematopoiesis, a subset of which showed promoter demethylation and were up-regulated upon vitC treatment including *CSF1R*, *CSF2RA*, *TLR2*. Curiously, we also observed deMRs near the transcriptional start site (TSS) of the hematopoietic transcription factors *PU.1* and *RUNX1*, although neither displayed significant or consistent differences in expression upon treatment across replicates (Figure 9).

Functional enrichment of genes associated with (<20kb from) vitC deMRs(59) and regions that gained 5hmC upon vitC treatment in IDH1^{R132H} cells (vitC iHRs; n=7829) revealed statistically significant ($q<0.01$) enrichment for gene sets related to hematopoiesis including the MAPK and CXCR4 signalling pathways and gene ontology terms related to myeloid cell differentiation (Figure 13). Similarly, regions that gained 5hmC upon vitamin C treatment in IDH1^{R132H} cells (VitC iHRs; n=7829) were found to be associated with genomic regions involved in chemokine ($1.289e^{-9}$), CXCR4 ($q=3.36e^{-9}$) and PDGFRB signalling ($q=1.5e^{-5}$) (Figure 13).

We hypothesized that there may exist genomic regions where vitC, by enhancing TET activity, stimulates both a gain of 5hmC (iHR) and loss of 5mC (deMR) to reverse hypermethylation induced by *IDH1^{R132H}* expression (IDH1^{R132H} iMRs). To this end we identified a set of vitC deMRs that were within 1kb of both an IDH1^{R132H} iMR and a vitC iHRs (n=23). Interestingly, this relatively small set of regions associated significantly with genes that are repressed in primitive hematopoietic cells overexpressing either NUP98-HOXA9 ($q=0.003$) or RUNX1-RUNX1T1 (AML1-ETO9a) ($q=0.009$) (Figure 13). Extending this approach to further examine regions potentially undergoing active, TET mediated demethylation (through a 5hmC intermediate) upon vitamin C treatment we defined a set of regions comprised of the union of deMRs identified by both meDIP-seq and WGBS that were within 1.5kb of an iHR

(DxMRs; n=712). Functional enrichment analysis of the genes associated with (<20kb from) DxMRs reinforced the pathway associations observed for deMRs and increased the gene set fold enrichments for genes down-regulated upon expression of NUP98-HOXA9 (+3.7 hyperFE vs. deMR alone) and AML-ETO9a (RUNX1-RUNX1T1) (+1.64 hyperFE vs. deMR alone). The increased association between DxMRs and gene sets related to leukemogenesis (compared to deMRs alone) suggests that the biologically relevant changes in methylation occur at regions that gain 5hmC and lose 5mC which represent potential TET target regions. This pattern of biologically relevant, co-occurring gain of 5hmC and loss of 5mC is exemplified at the promoter of *CSF2RA,* a gene that becomes up-regulated upon vitC treatment and encodes a subunit of the GM-CSF receptor and is required for GM-CSF induced myeloid differentiation (Figure 14). Reduced fractional methylation upon vitamin C treatment within the deMR at the *CSF2RA* promoter was confirmed by targeted bisulfite sequencing in biological triplicates, along with 5 other deMRs near genes of interest and positive (H19 ICR) and negative (GAPDH promoter) controls (Figure 15).

The majority of vitC iMRs (6684; 79.8%) were distal (>5kb) from gene TSSs and a fraction overlapped CpG islands (938; 11.2%) where CpG fractional methylation is typically low (11) (Figure 16). Functional annotation of genes associated with vitC iMRs revealed terms unrelated to hematopoiesis but implicated nervous system function and development (Figure 17). Furthermore, GREAT (59) functional enrichment analysis of genes with promoter CpG islands containing iMRs revealed an enrichment for regions marked by H3K27me3 and H3K4me3 in embryonic stem cells and the “adult tissue stem module”; a set of genes that are coordinately up-regulated in somatic stem and leukemic cells that includes *MEIS1* (53) and *CDK1* (54), both of which are down-regulated by vitC. Since TET functions as a DNA demethylase, gains in 5mC

would not be expected to be a direct consequence of vitC induced TET activation but rather indirect as the consequence of changes in cellular identity or differentiation, for example. Together these observations suggest that, in *IDH1*^{R132H} expressing cells, vitC treatment leads to a specific loss of DNA methylation at regulatory regions implicated in the control of hematopoiesis. Additionally, gene ontology analysis supports the notion that vitamin C induced demethylation, especially in the context of increased 5hmC, may activate regulatory regions associated with genes that are usually repressed in AML but activated in more mature myeloid populations.

Library Strategy	Cell Type	Treatment	Target Epigenetic Mark	Number of Reads
meDIP-seq	IDH1 WT	Untreated	5-methylcytosine	102927422
meDIP-seq	IDH1 WT	Vitamin C	5-methylcytosine	108200352
meDIP-seq	IDH1 R132H	Untreated	5-methylcytosine	96378804
meDIP-seq	IDH1 R132H	Vitamin C	5-methylcytosine	96282208
hmeDIP-seq	IDH1 R132H	Untreated	5-hydroxymethylcytosine	99191630
hmeDIP-seq	IDH1 R132H	Vitamin C	5-hydroxymethylcytosine	103338942
Whole Genome Bisulfite	IDH1 R132H	Untreated	5-methylcytosine	660404508
Whole Genome Bisulfite	IDH1 R132H	Vitamin C	5-methylcytosine	737187172
ChIP-seq (H3K27ac)	IDH1 R132H	Untreated	H3K27ac	36754742
ChIP-seq (H3K27ac)	IDH1 R132H	Vitamin C	H3K27ac	34838944
ChIP-seq (H3K4me1)	IDH1 R132H	Untreated	H3K4me1	115262818
ChIP-seq (H3K4me1)	IDH1 R132H	Vitamin C	H3K4me1	130725370
ChIP-seq (H3K4me3)	IDH1 R132H	Untreated	H3K4me3	43882302
ChIP-seq (H3K4me3)	IDH1 R132H	Vitamin C	H3K4me3	47123064
ChIP-seq (PU.1)	IDH1 R132H	Untreated	PU.1	34814708
ChIP-seq (PU.1)	IDH1 R132H	Vitamin C	PU.1	22954835
ChIP-seq (Runx1)	IDH1 R132H	Untreated	RUNX1	27397561
ChIP-seq (Runx1)	IDH1 R132H	Vitamin C	RUNX1	35128160
polyA+ RNA-seq	IDH1 R132H	Untreated	mRNA	105066515
polyA+ RNA-seq	IDH1 R132H	Vitamin C	mRNA	130658459

Table 2. A summary of sequencing libraries generated for this study.

Name	Number of Regions	Total Genomic Occupancy (kilobases)	Method of Detection	Description
deMR	6496	1283	WGBS	regions losing methylation upon vitamin C treatment in IDH1 R132H cells
iMR	8374	2159	WGBS	regions gaining methylation upon vitamin C treatment in IDH1 R132H cells
iHR	7829	4531	hmeDIP	regions gaining hydroxy-methylation upon vitamin C treatment in IDH1 R132H cells
deHR	6179	3548	hmeDIP	regions losing hydroxy-methylation upon vitamin C treatment in IDH1 R132H cells
DxMR	712	233	medIP/hmedIP	deMRs within 1.5kb of an iHR
IDH1 R132H iMR	4182	2580	meDIP	regions hypermethylated in IDH1 R132H cells compared to IDH1WT cells
IDH1 R132H deMR	4519	2685	meDIP	regions hypomethylated in IDH1 R132H cells compared to IDH1WT cells
meDIP deMR	3789	2242	meDIP	regions losing methylation upon vitamin C treatment in IDH1 R132H cells
meDIP iMR	4564	2755	meDIP	regions gaining methylation upon vitamin C treatment in IDH1 R132H cells

Table 3. An enumeration of the regions that become differentially marked (DMRs) upon vitamin C treatment or upon IDH1^{R132H} expression and their total genomic occupancy in kilobases (kb).

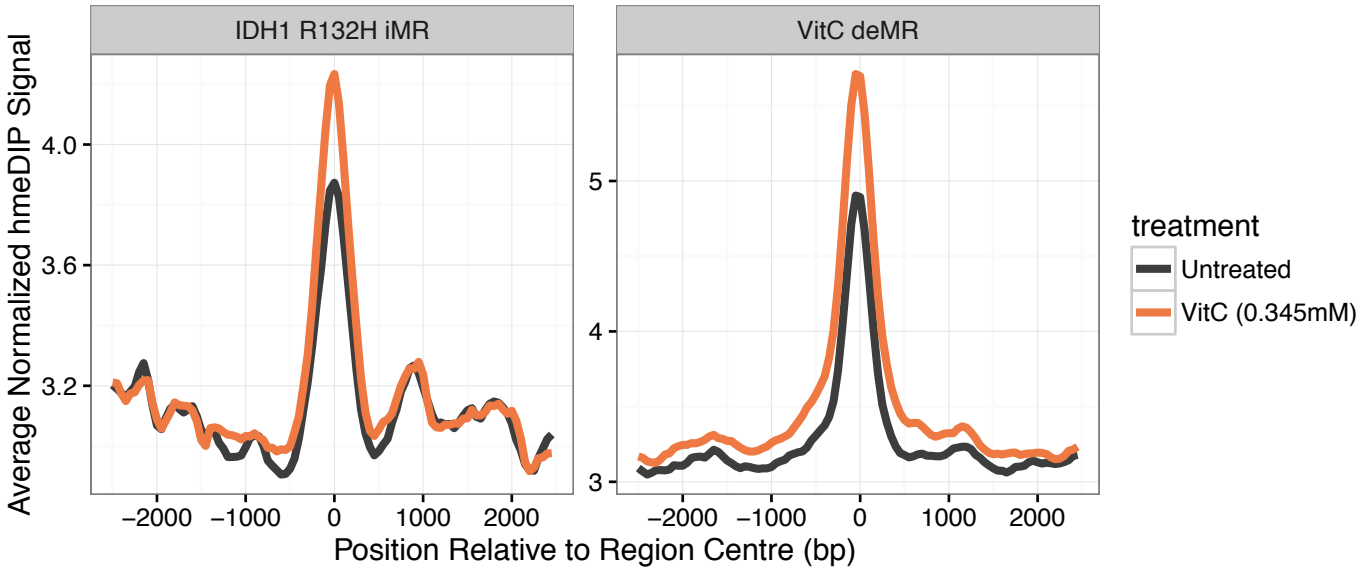

Figure 11. Average normalized 5hmC signal for untreated (black) and vitamin C treated (orange) IDH1^{R132H} cells ± 1.5kb from the centre of regions hypermethylated in IDH1^{R132H} cells vs. IDH1^{WT} cells (IDH1^{R132H} iMRs; left) and regions losing demethylated upon vitamin C treatment (deMRs; right).

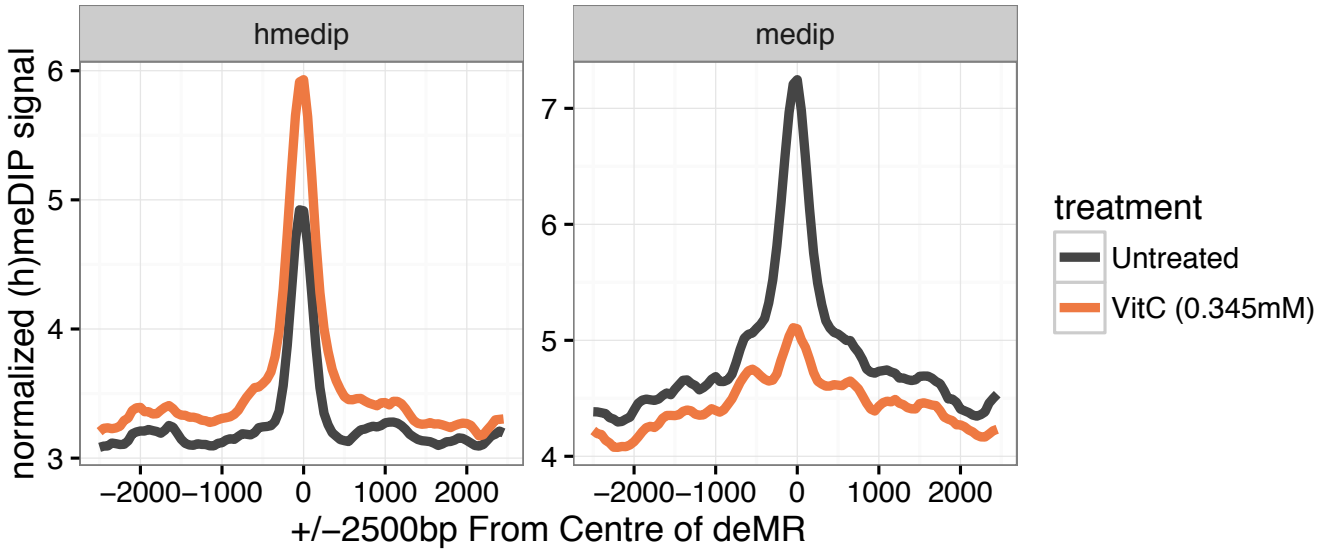

Figure 12. Average normalized meDIP (right) and hmeDIP (left) signal ± 2.5kb from the centre of vitamin C deMRs identified by WGBS analysis.

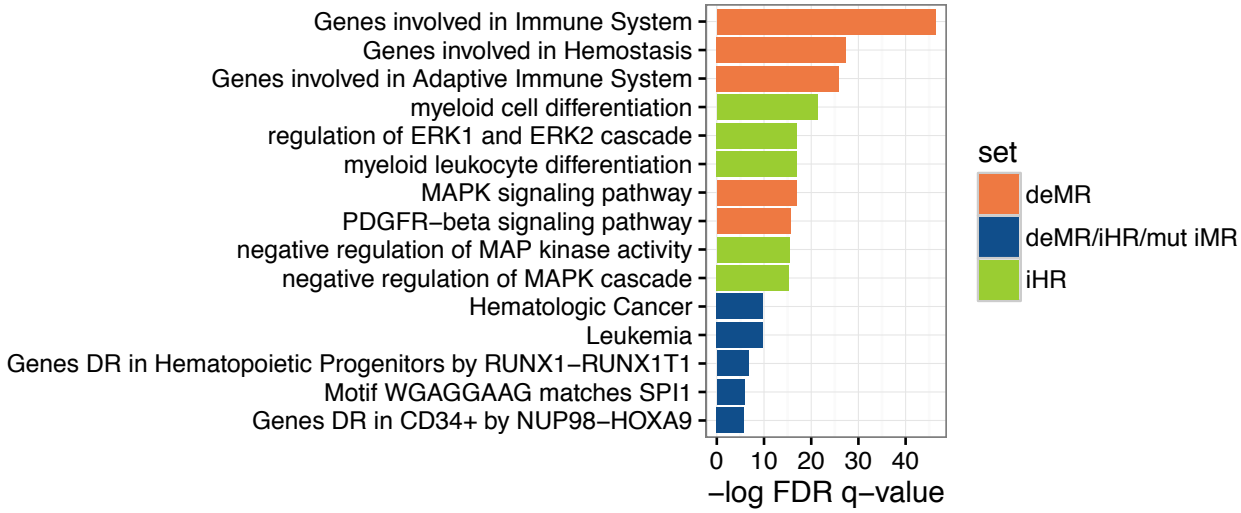

Figure 13. Top 5 GREAT gene ontology results for vitamin C deMRs (orange), iHRs (green) and the intersection between deMRs, iHRs and IDH1^{R132H} iMRs (blue; n=23).

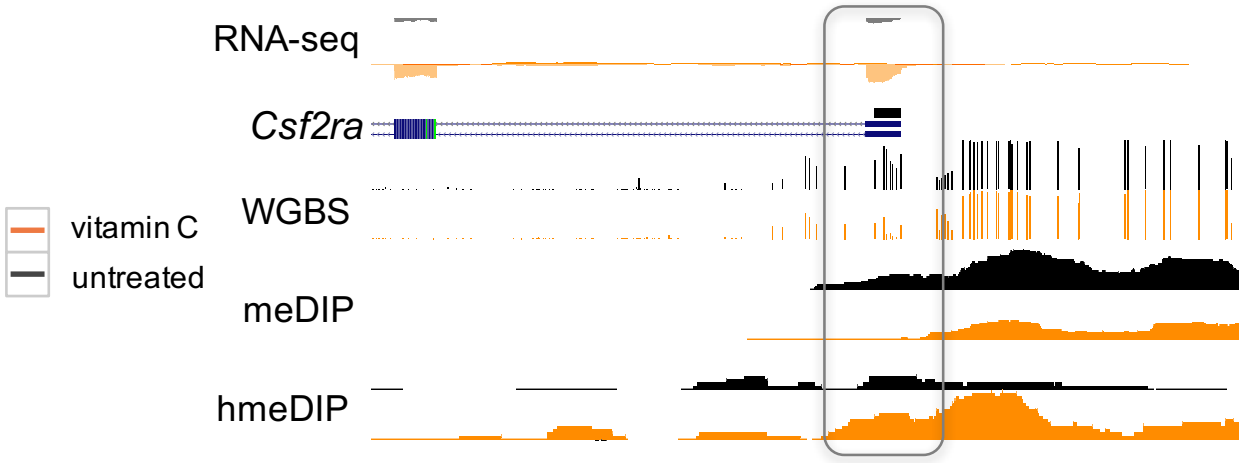

Figure 14. A screen shot from the UCSC genome browser depicting signal for RNA-seq (Rep 1), WGBS, meDIP-seq and hmeDIP-seq at the *Csf2ra* promoter.

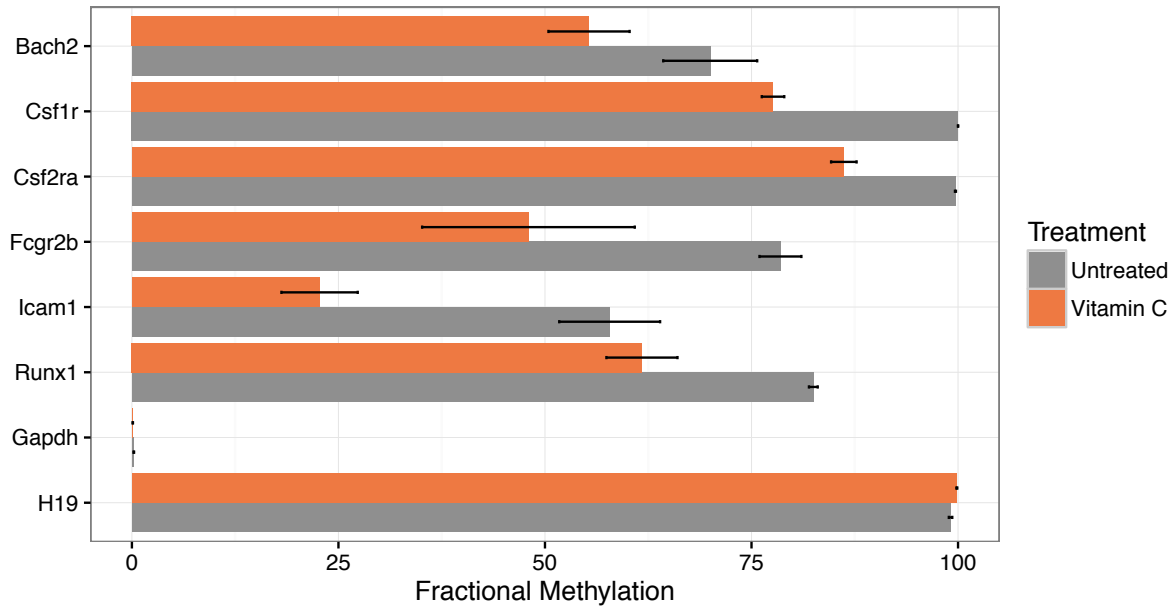

Figure 15. Average fractional methylation from targeted bisulfite sequencing in vitC deMRs associated with genes of interest in untreated and vitamin C treated IDH1^{R132H} cells. H19 ICR and GAPDH promoter represent positive and negative controls, respectively

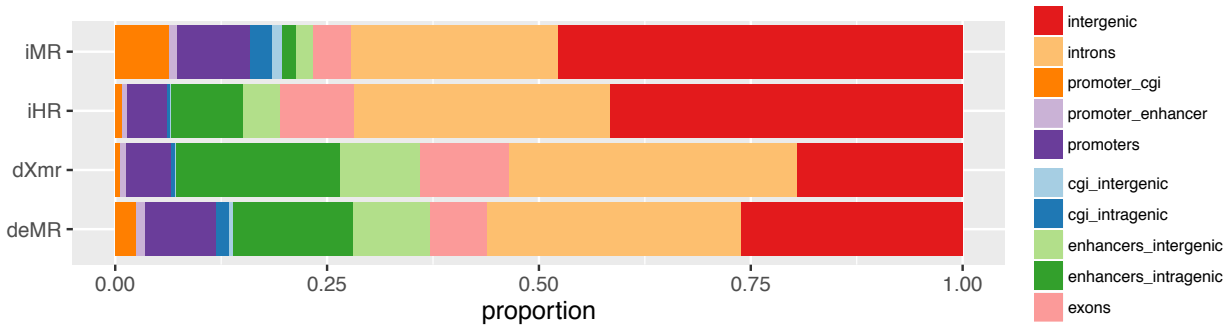

Figure 16. The proportion of various sets of DMRs that overlap previously annotated genomic features (mm10).

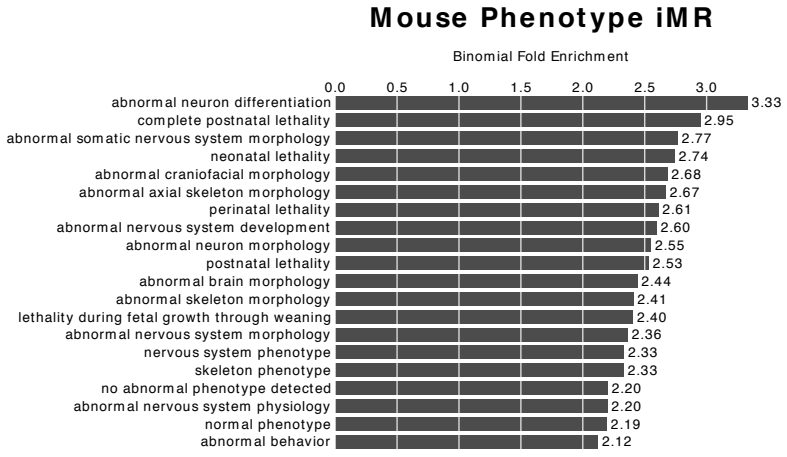

Figure 17. Statistically significant GREAT gene ontology results (mouse phenotype) for regions gaining methylation (iMRs) upon vitamin C treatment in IDH1^{R132H} cells.

2.4 Vitamin c induces demethylation at hematopoietic transcription factor binding sites

Since DNA methylation has previously been shown to influence the binding affinities of transcription factors(60), which play an essential role in normal hematopoiesis, we sought to examine the link between vitC induced changes in methylation and transcription factor binding sites by identifying enriched motifs within vitC deMRs. Consistent with the observed gene

ontology associations, the top *de novo* motifs identified by HOMER(61) motif enrichment analysis within deMRs were for PU.1 (34%), CEBPβ (15%), HIF1α and RUNX1 (30%) (Figure 18). Restricting our analysis to DxMRs increased the proportion of regions with *de novo* motifs for PU.1 (41%) and CEBPβ (41%) motifs while the proportion containing a RUNX1 motif decreased to 5%. In contrast, no *de novo* motifs were present in >1% of our iMR regions, suggesting vitC induced increases in methylation do not occur at transcription factor binding motifs.

To further explore the relationship between vitC DMRs and transcription factor binding we sought to calculate the proportion of different DMRs that overlapped previously identified regions of interest. Based on the motif enrichment results and their relevancy in myelopoiesis we included previously published PU.1 and CEBPβ binding sites, as identified by ChIP-seq, as well as DNase hypersensitvity sites (DHSs) identified in hematopoietic progenitors and macrophages(62). Both active regulatory regions and regions bound by transcription factors are sensitive to DNase cleavage are annotated DHSs, offering a general, non-specific estimation of the active regulatory landscape in different cell types. We included binding sites for the embryonic transcription factors Nanog (63) and Rad21 (64) in mESC as negative controls. VitC deMRs significantly overlapped mature myeloid (macrophage) DHSs (1019; 16%), PU.1 (1846; 28%) and CEBPβ (1428; 22%) binding sites individually (Figure 19), with 768 deMRs (12%) overlapping regions bound by both PU.1 and CEBPβ binding sites in macrophages. Furthermore, the average 5mC levels decreased in response to vitC treatment at macrophage (MAC) PU.1 binding sites (t-test; p~0), CEBPβ binding sites (t-test; p~0) and DHS (t-test; p~0), but not progenitor (HP) DHS (t-test; p=0.447) (Figure 20). In contrast, vitC iMRs shared minimal overlap with transcription factor binding sites in myeloid cells (Figure 19). However, a

significant proportion (226; 31%) of iMR containing CpG islands overlap hematopoietic progenitor DHSs supporting the notion that iMRs represent progenitor-specific regulatory regions that become deactivated during vitC treatment. Taken together, these results support a model of vitC dependent demethylation at regulatory regions and within the binding sites of TFs required for myeloid development.

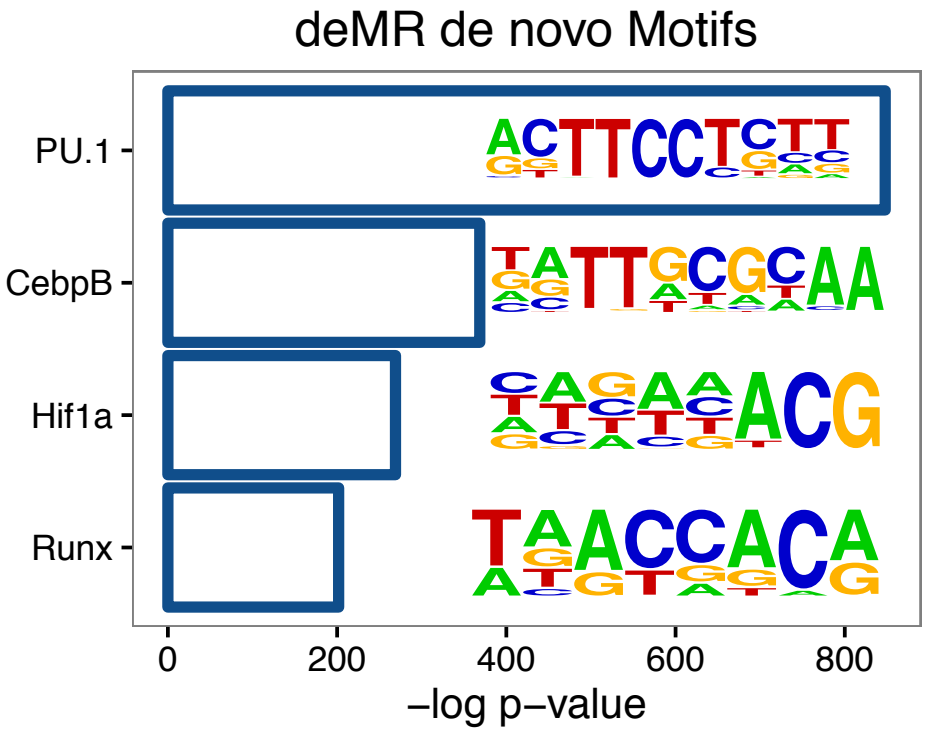

Figure 18. Top hits for HOMER de novo motif analysis performed on vitamin C deMRs identified by WGBS.

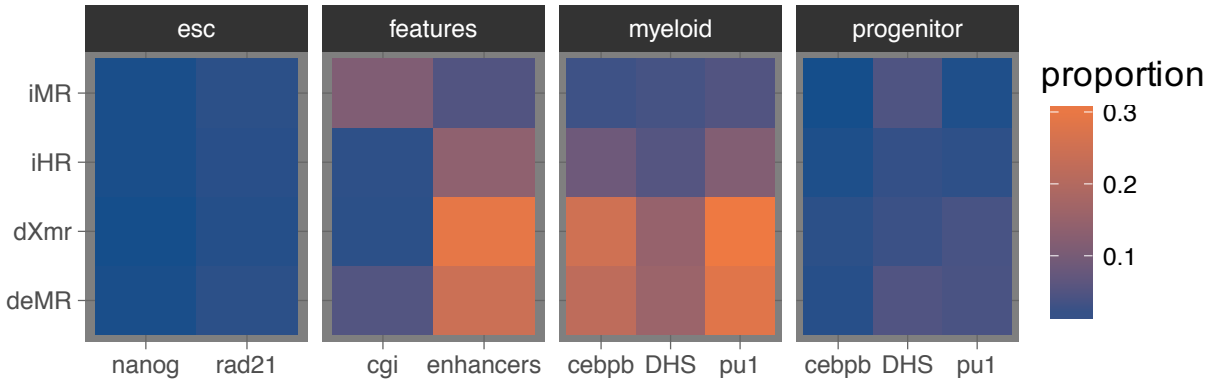

Figure 19. A heatmap displaying the proportion of regions in different DMR sets (y-axis) that overlap transcription factor binding sites and genomic features (x-axis).

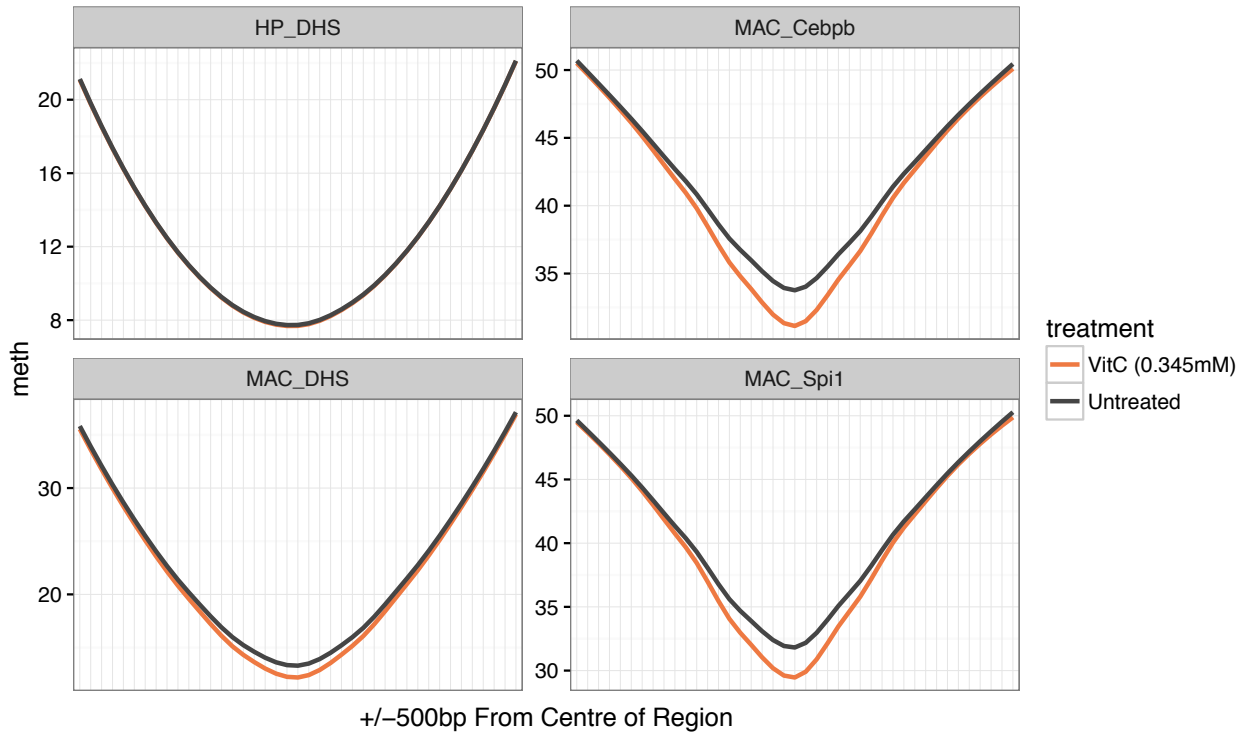

Figure 20. Average (smoothed) fractional methylation for untreated (black) and vitamin C treated (orange) cells within hematopoietic progenitor (HP) and mature myeloid (MAC) DNAse hypersensitivity sites (DHS) and MAC Cebpß and Spi1 (PU.1) binding sites.

2.5 Vitamin c induced demethylation is enriched at myeloid specific enhancers

Dynamic epigenetic remodelling at enhancer regions is a crucial step in the cell fate determination of hematopoietic progenitor cells, a process that is often disturbed in hematological malignancies. Hypermethylation at enhancer elements coincides with leukemogenesis in hematopoietic progenitor cells of *TET2*$^{-/-}$ mice expressing the *AML-ETO9a* fusion(65). Additionally, the presence of DNA methylation has been shown to negatively correlate with enhancer activity(66) and alter the ability of tissue specific DNA binding proteins to bind their target sites(60). Enhancer activity, represented by H3K4me1 signal, typically becomes established at the root of lineage commitment and increases as lineage specific progenitor cells mature. Previously published H3K4me1 processed signal at 48,415 hematopoietic enhancers in 16 unique cell types (67) was leveraged to explore the relationship between vitC induced demethylation and enhancer dynamics during hematopoiesis. Intersection of vitC deMRs with these hematopoietic enhancers revealed a significant overlap (1584; 24.4%; t-test p~0), 43% of which contained previously published PU.1 binding sites. Plotting the mean H3K4me1 signal at DMR containing enhancers during lineage commitment from hematopoietic stem cells (ST-HSC) to the general myeloid progenitors (GMP), common lymphoid progenitors (CLP) and megakaryocyte erythroid progenitors (MEP) revealed a progressive increase in H3K4me1 signal (enhancer activity) at deMR associated enhancers during the progression from ST-HSC towards the committed myeloid precursor GMP, but not CLP or MEP cells (Figure 21a). Conversely, enhancers overlapping iMRs became less active, losing H3K4me1 signal in committed progenitors of all three lineages (Figure 21a) suggesting that these enhancers may be active only in primitive hematopoietic progenitors and may undergo inactivation via DNA methylation during hematopoiesis. In mature hematopoietic cell types H3K4me1 signal

increased at vitC deMR enhancers (Kolmogorov-Smirnov (K-S) test; p ~ 0) and decreased at vitC iMR containing (n=391) enhancers (K-S test; p = $1.25e^{-12}$) compared to random enhancers (n=1000) in the myeloid lineage (Figure 21b and Figure 22) but showed no statistically significant difference in the lymphoid and erythroid lineages. These results suggest that vitC-induced demethylation occurs at enhancers that become active as hematopoietic progenitor cells differentiate into mature cells of the myeloid lineage, specifically.

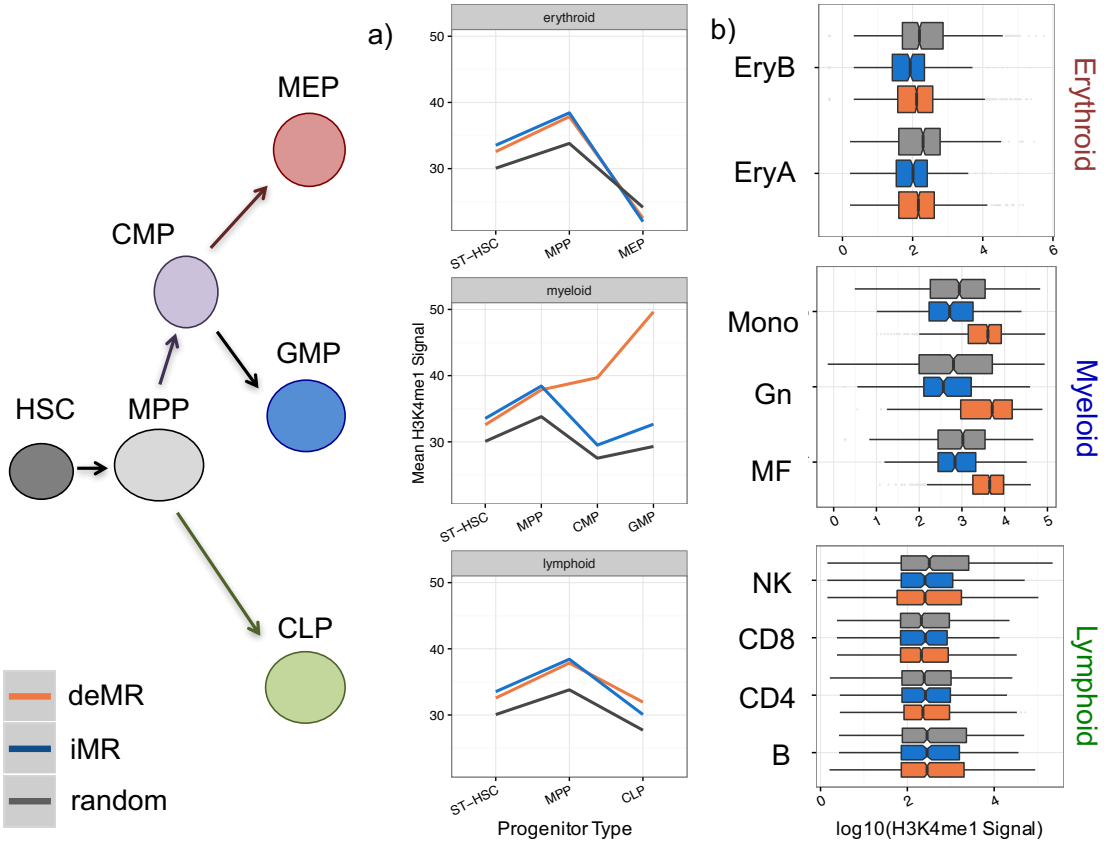

Figure 21. VitC treatment leads to demethylation at enhancers that become active in the myeloid lineage specifically. A depiction of the hierarchal hematopoietic differentiation pathway from ST-HSC into mature cells of the erythroid (top), myeloid (middle) and lymphoid (bottom) lineages; a) A linear representation of mean H3K4me1 signal during lineage commitment and; b) A boxplot of H3K4me1 signal at enhancers in mature cell types of a given lineage at previously identified enhancers overlapping deMRs (orange), iMRs (blue) or a random set (n=1000; gray).

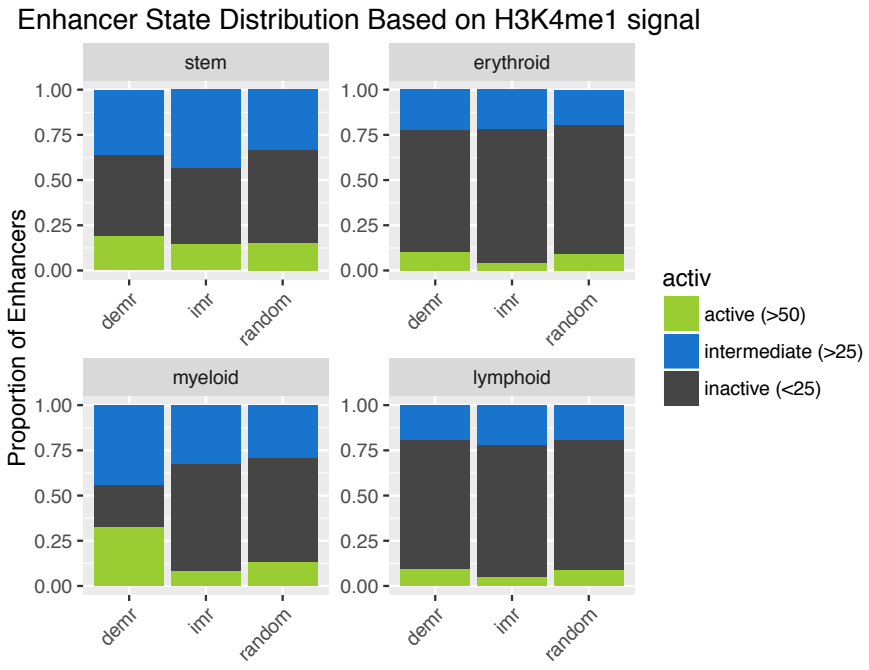

Figure 22. The proportion of previously identified enhancers overlapping deMRs, iMRs or a random set (n=1000) that are active (> 50 H3K4me1 signal; green), intermediate (25-50 H3K4me1 signal; blue) or inactive (<25 H3K4me1 signal; black) as previously described in the original publication (67).

2.6 Vitamin c treatment results in the re-distribution of histone modifications

2.6.1 H3K4 methylation and H3K27 acetylation dynamics at vitamin c DMRs

In addition to the absence of 5mC, the presence of H3K4me1 and H3K27ac are hallmarks of active enhancers. To explore changes in the distribution of histone modifications and enhancer activity upon vitC treatment we performed H3K27ac, H3K4me1 and H3K4me3 ChIP-seq on IDH1^{R132H} cells grown in the presence and absence of vitC. By calculating the change in normalized ChIP-seq signal between treatments in 50bp bins across a 500bp window centered on the middle of vitC induced DMRs we identified an increase in H3K27ac at deMRs, iHRs and DxMRs (t-test; $p<1.34e^{-9}$), but not iMRs (t-.test; p=0.472; Figure 23). We also observed increased H3K4me1 signal at all DMRs (t-test; $p<3.37e^{-6}$) and increased H3K4me3

signal at deMRs (t-test; $p=1.34e^{-4}$) and DxMRs (t-test; p=0.03) but not iHRs (t-test; p=0.42) or iMRs (t-test; p=0.81). This pattern of change in H3K4me1/me3 and H3K27ac was validated in a separate ChIP-seq experiment in a biological replicate (Figure 24). These results support the previously developed notion that deMRs, DxMRs and iHRs, unlike iMRs, overlap with regulatory elements that undergo multifaceted epigenetic remodelling during vitC treatment.

2.6.2 Enhancer activity correlates with changes in transcription and DNA methylation

We next examined the link between changes in histone modifications, DNA methylation and gene expression upon vitC treatment. The majority of up (373/452; 83%) and down-regulated (165/221; 75%) gene TSSs are associated with (<20kb from) one or more previously identified hematopoietic enhancers (67). Enhancers associated with genes that are transcriptionally induced upon vitC treatment showed a significant increase in H3K27ac signal (K-S test; p=0.0086) (Figure 25a), suggesting that increased enhancer activity, represented by the acquisition of H3K27ac, coincides with increased expression of nearby genes. We next identified treatment specific enhancers within $IDH1^{R132H}$ cells from our ChIP-seq data as those regions marked by H3K4me1 and H3K27ac (MACS2; $p<1e^{-5}$) in either untreated or vitC treated cells, removing those present in both cell types. This produced a set of 4389 and 3650 enhancers specifically activated in untreated (Untreated only) and vitC (vitC only) treated cells, respectively. The integration of fractional methylation values from WGBS data revealed significant loss of 5mC (t-test; $p=7.94e^{-5}$) at enhancers active in vitC treated cells specifically, but not in those specific to untreated $IDH1^{R132H}$ cells (t-test**;** p=0.8357; Figure 25b). To explore the activity of these $IDH1^{R132H}$ cell enhancers during hematopoiesis we intersected them with the catalogue of hematopoietic enhancers (67) as previously described. A significant fraction of $IDH1^{R132H}$ cell enhancers overlapped the catalogue of hematopoietic enhancers (1358/4389

(31%) and 1276/3650 (35%) for untreated and vitamin C, respectively). Interestingly, although the majority of vitamin C specific enhancers were devoid of deMRs (1138/1276; 89.2%) they displayed a pattern of activation similar to deMR containing enhancers (Figure 21) during myeloid lineage commitment (Figure 26**;** centre panel) and had higher H3K4me1 signal in mature myeloid cells compared to random regions (K-S test; p ~ 0; Figure 27). These results suggest that changes in histone modifications upon vitC treatment, like changes in DNA methylation, occur at regulatory elements that become activated during the myelopoiesis.

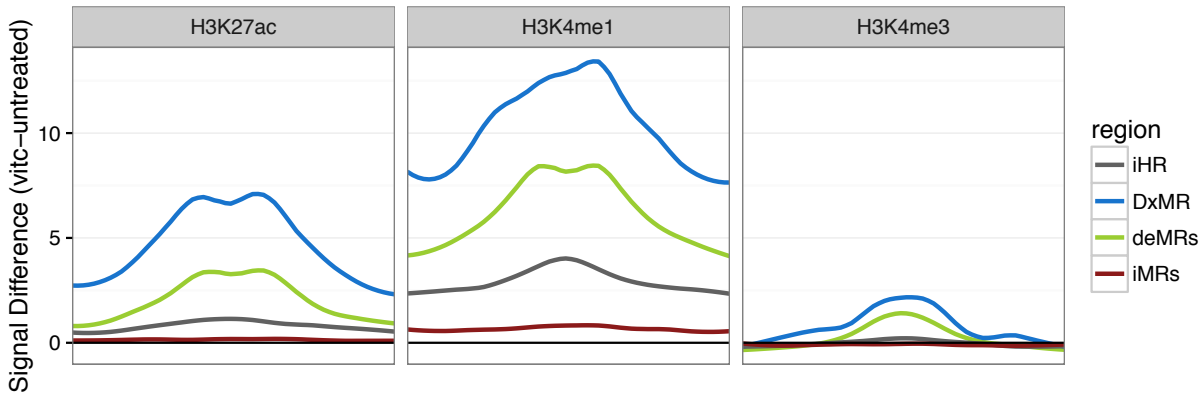

Figure 23. Average normalized signal difference between treatments (vitC – untreated) for H3K27ac (left) H3K4me1 (middle), and H3K4me3 (right) ± 1.5kb from the centre of vitamin C DMRs.

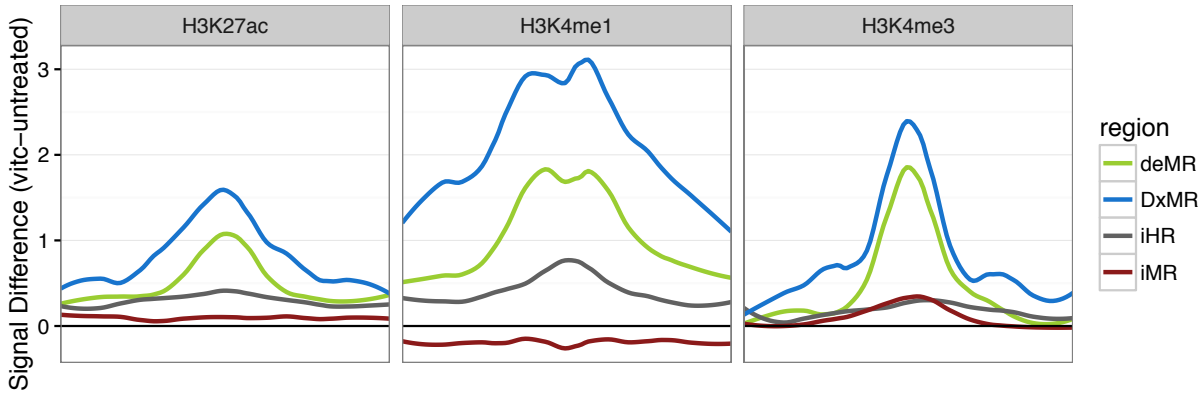

Figure 24. Validation experiment showing normalized signal difference in a biological replicate between treatments (vitC – untreated) for H3K27ac (left) H3K4me1 (middle), and H3K4me3 (right) ± 1.5kb from the centre of vitamin C DMRs.

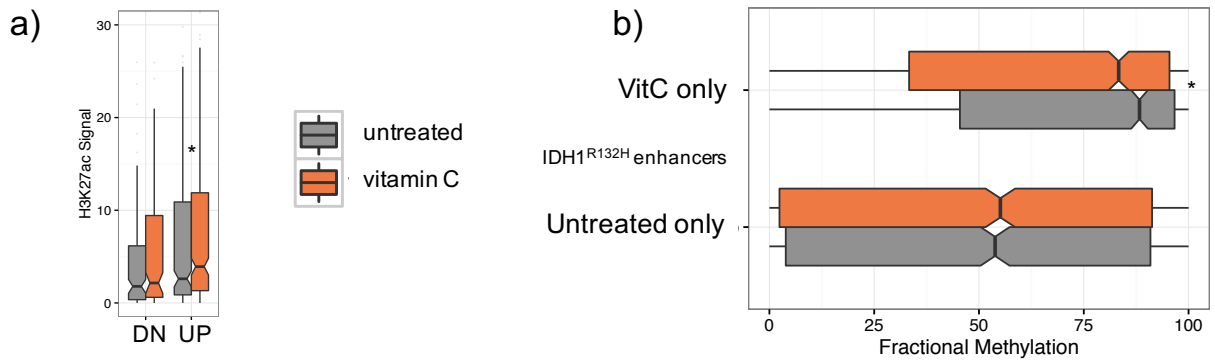

Figure 25. a) Average normalized H3K27ac signal in the nearest enhancers (within 20kb) to differentially expressed genes with * indicating statistically significant ($p<0.01$) differences between untreated (gray) and vitamin C treated (orange). b) Average fractional methylation for vitamin C treated (orange) or untreated (gray) cells at CpG sites overlapping $IDH1^{R132H}$ enhancers identified only in vitamin C treated (top) or untreated (bottom) cells.

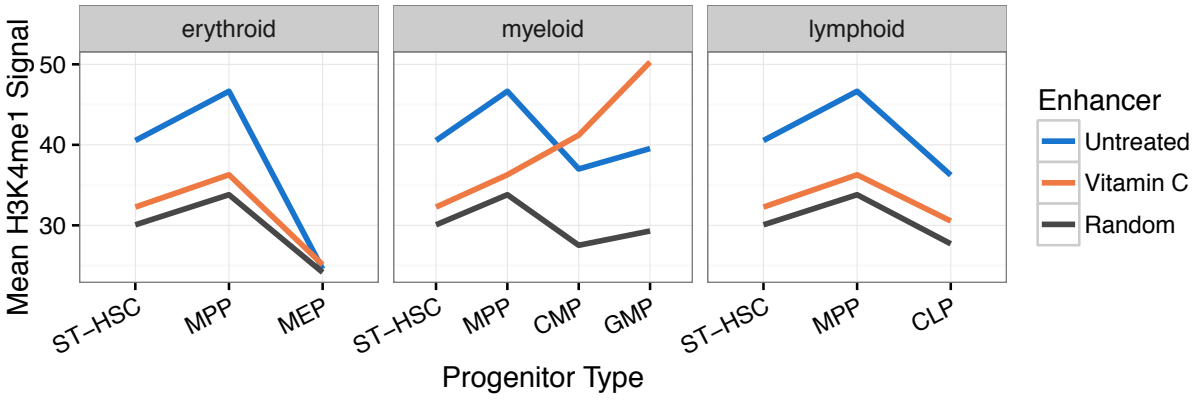

Figure 26. A linear representation of mean H3K4me1 signal during lineage commitment at previously published enhancers that overlap $IDH1^{R132H}$ enhancers identified in our study in specific to Untreated (blue) or Vitamin C treated (orange) cells.

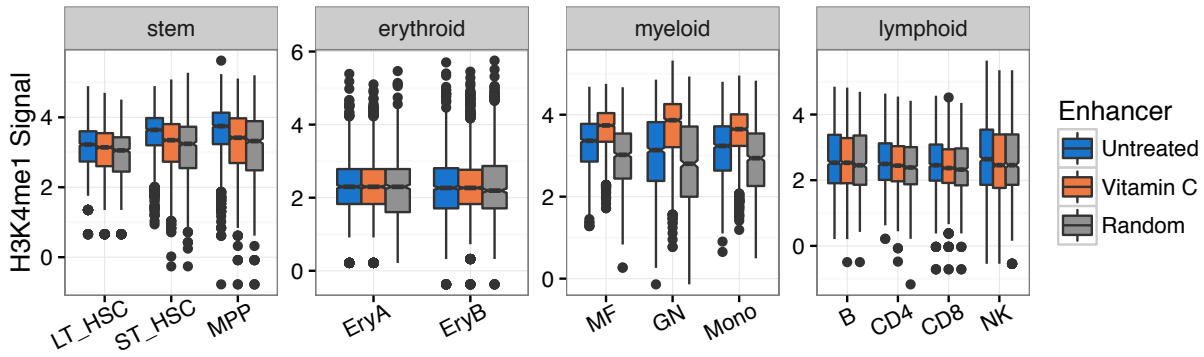

Figure 27. A boxplot of log10 H3K4me1 signal at enhancers in mature cell types of a given lineage at previously identified enhancers overlapping IDH1^{R132H} enhancers identified during our study in untreated cells only (Blue), vitamin C treated cells only (orange) or a random set (n=1000; gray).

2.7 Vitamin c alters the binding patterns of PU.1 and RUNX1

PU.1 is a pioneering transcription factor that is essential for hematopoiesis and, unlike most transcription factors, can bind chromatin in the presence of restrictive epigenetic modifications (68). Reduced PU.1 expression in hematopoietic cells leads to reduced growth and transformative capability in several leukemic mouse models including AML (69). Additionally, PU.1 has been implicated in the control of the expression of genes that were differentially expressed upon vitC treatment (e.g. *cKIT*, *CSF1R*, *CSF2RA*). RUNX1 is a critical regulator of hematopoiesis that frequently participates in leukemia-associated translocations including the *RUNX1-RUNX1T1* (*AML1-ETO9a*) fusion, which is recurrent in myeloid malignancies(70).

To explore the effect of vitC treatment on the binding patterns of these key hematopoietic transcription factors we performed ChIP-seq using antibodies against *RUNX1* and *PU.1* (TF-ChIP) on IDH1^{R132H} cells grown in the presence or absence of vitC. We identified enriched regions using MACS2 (57) (narrow peaks; $p<1e^{-5}$) and observed 62,770 and 23,781 PU.1 enriched regions in untreated and vitC treated cells, respectively. Interestingly, the majority of

PU.1 binding sites in vitC treated cells were also found in untreated cells (22,646/23,781; 95%) suggesting a significant loss of PU.1 binding upon vitC treatment. Consistent with the previously described physical association between TET2 and PU.1 (71), we found that a significant proportion of vitC DMRs (30%; Figure 28) and IDH1^{R132H} enhancers (1252/3650; 34% vitC only and 1651/4389; 38% untreated only) were located within 1kb of PU.1 binding sites. These results suggest that PU.1 may guide the epigenomic remodelling events that occur upon vitC treatment.

We identified 4047 and 6587 RUNX1 enriched regions in untreated and vitC treated cells, respectively. In contrast to PU.1, a large fraction of RUNX1 binding sites were specific to vitC treated cells (3134/6587; 48%). A majority of vitamin C specific RUNX1 binding events were within CpG islands (2128/3134; 68%) with a subset localizing to IDH1^{R132H} cell enhancers (214/3134; 7%). Examining DNA methylation values within vitC specific RUNX1 enriched regions revealed significant demethylation flanking RUNX1 enriched summits (+/-100bp) within enhancers (K-S test; p=0.0024), but not at CpG islands (K-S test; p=0.2372; Figure 29).

RUNX1 acts in concert with epigenetic modifiers, most notably histone acetyl-transferases, and perturbation of this ability can facilitate transformation (72). We examined the relationship of H3K27ac and H3K4me1 in vitC induced RUNX1 binding by plotting the difference in signal in the presence and absence of vitC (vitC – untreated). VitC specific RUNX1 enriched regions showed a significant gain of H3K4me1 and H3K27ac signal (K-S test; p~0) in the 500bp flanking the centre of RUNX1 binding sites (Figure 30 and Figure 31). In contrast, we observed a significant loss of H3K4me1/me3 (K-S test; p ~ 0), and to a lesser extent H3K27ac signal (K-S test; p=1.626e-6), upon vitC treatment in the region +/-500bp from the centre of RUNX1 enriched regions lost in response to vitC treatment (Figure 30). This pattern of change in H3K4me1/me3 and H3K27ac relative to the centre of treatment specific RUNX1 binding sites

was also observed in a biological replicate (Figure 32). Taken together these results suggest that vitC treatment increases RUNX1 binding and increases H3K27ac in the genomic regions surrounding its binding sites.

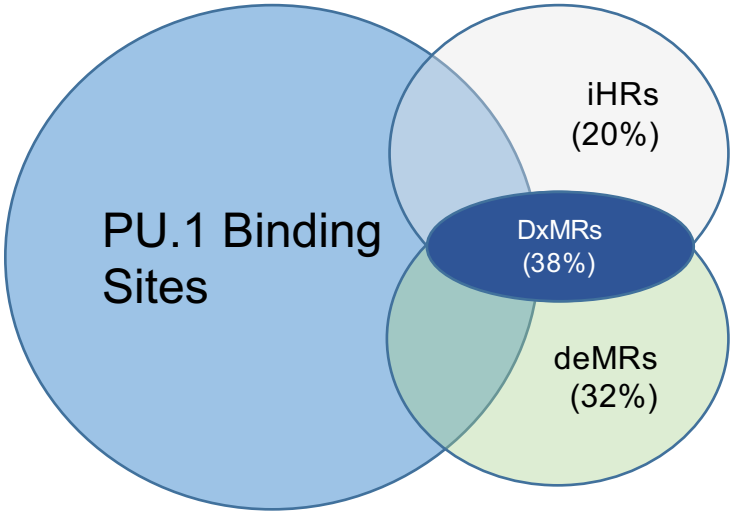

Figure 28. The percentage of different DMR types that exist within 1kb of at least one PU.1 binding site identified by ChIP-seq.

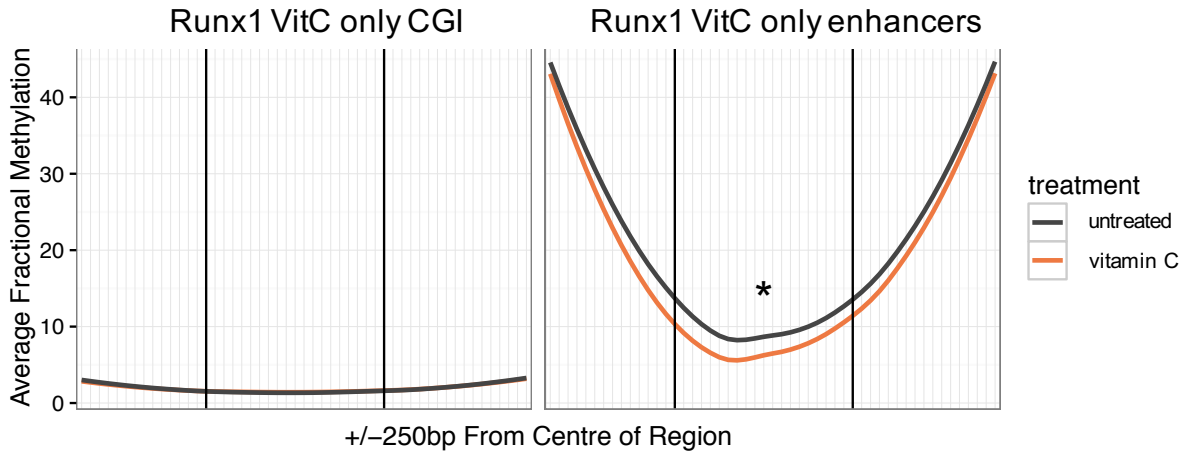

Figure 29. Average (smoothed) fractional methylation for untreated (black) and vitamin C treated (orange) cells at CpG sites ± 1.5kb from the centre of vitamin C specific Runx1 binding sites that overlapped either CpG islands (left) or enhancers (right). *Indicates a statistically significant difference (K-S test; $p < 0.05$) between treatments.

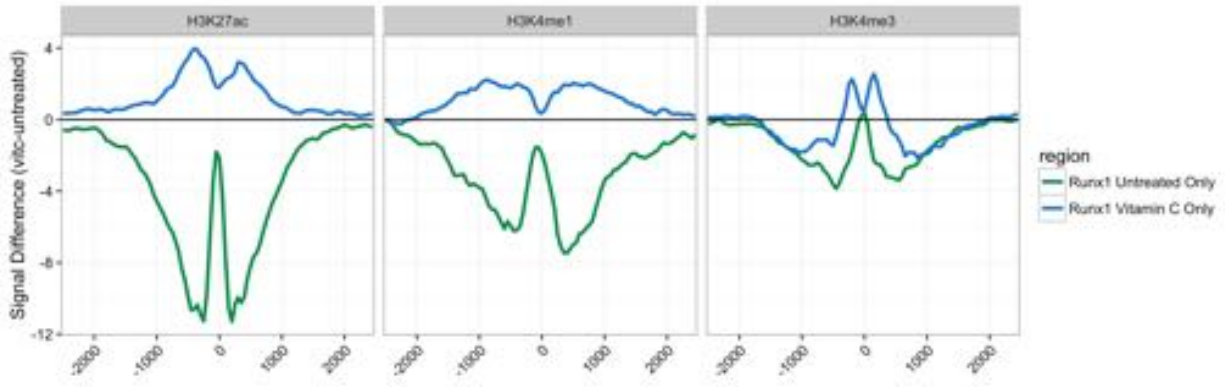

Figure 30. Mean normalized H3K27ac (left), H3K4me1 (middle) and H3K4me3 (right) signal difference (vitC – untreated) ± 2kb from the centre of untreated only (green) or vitamin C only (blue) RUNX1 binding sites.

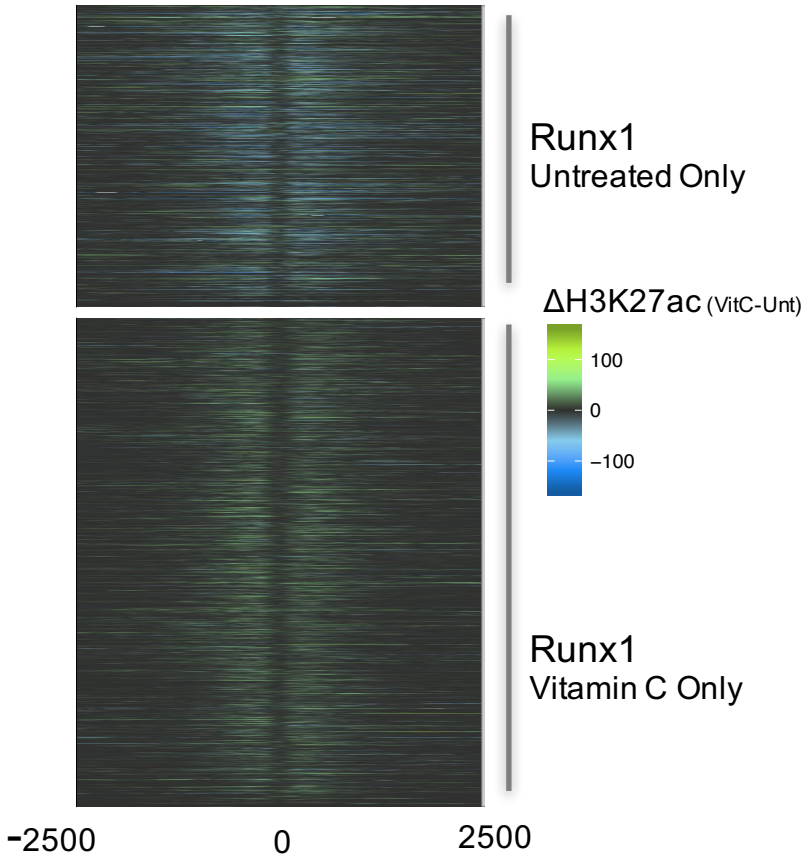

Figure 31. A heatmap representing normalized H3K27ac signal difference (vitamin C – untreated) +/- 2.5 kb from the summit of untreated only (top) or vitamin C only (bottom) Runx1 binding sites.

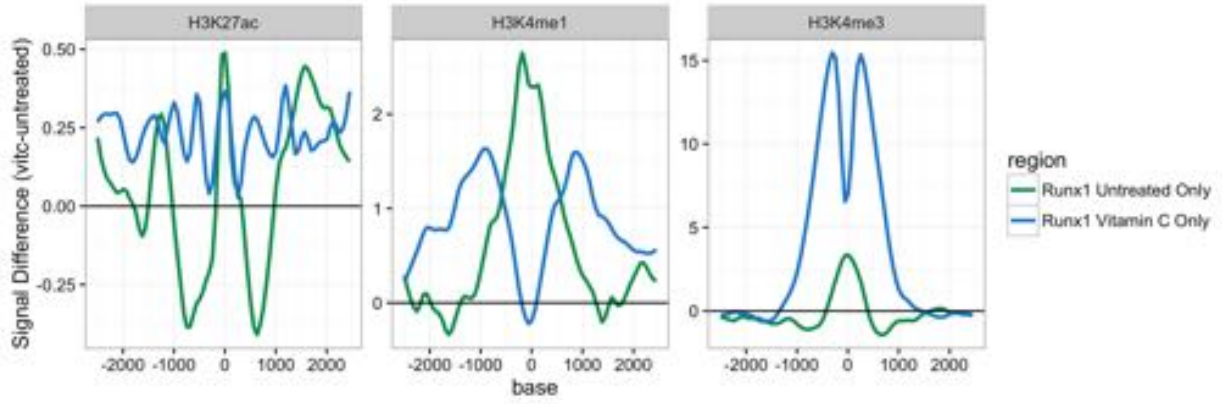

Figure 32. Mean normalized signal difference (vitC – untreated) from biological replicate H3K27ac (left), H3K4me1 (middle) and H3K4me3 (right) chIP-seq libraries in the regions ± 2kb from the center of regions flanking the summit of previously identified untreated only (green) or vitamin C only (blue) RUNX1 binding sites.

2.8 Materials and methods

2.8.1 Laboratory methods

2.8.1.1 Retroviral vectors and preparation of mouse bone marrow cells

Retroviral vectors MSCV-HoxA9-PGKneo, pSF91-IRES-eGFP, pSF91-IDH1^{R132H}-IRES-eGFP, pSF91-IDH1^{WT}-IRES-eGFP have been described previously (43). C57BL/6J mice were injected intraperitoneally with 5-Fluorouracil (Medac, Hamburg, Germany) at a dose of 150 mg/kg. Five days later, bone marrow cells were harvested and transduced first by co-cultivation with a HoxA9 viral producer cell line followed by either control (CTL) vector or IDH1^{WT} or IDH1^{R132H} viral producer GP+E86 cells (43). Cells were then sorted for GFP expression and maintained in DMEM medium (Stem Cell Technologies, Vancouver, Canada) supplemented with 10% heat inactivated fetal bovine serum (Life Technologies) and 6 ng/ml murine IL-3, 10 ng/ml human IL-6, and 20 ng/ml murine SCF (Peprotech, Hamburg, Germany).

2.8.1.2 In vitro proliferation assay

Cells were plated at a density of $5x10^5$ cells/ml in flat bottom 24 well plates and incubated under light protective conditions. Vitamin C in the form of L-Ascorbic acid 2-phosphate sesqui-magnesium salt hydrate (vitC; A8960; Sigma-Aldrich, Steinheim, Germany) was added to the culture medium daily at a concentration of 100 µg/mL (0.345 mM). Live cells were counted at the specified days with trypan blue.

2.8.1.3 Lineage staining and FACS antibodies

Cells were plated at a density of $5x10^5$ cells/ml in flat bottom 6 well plates (5ml/well) under light protective conditions and vitamin C was added at the specified concentrations. Monoclonal antibodies for lineage staining used were Gr1-PE (clone-RB6-8CS), Sca-1-PE (D7), ckit-APC (2B8; BD Biosciences, Heidelberg, Germany), CD11b-APC (M1/70) and F4/80-PE (BM8; eBioscience, Frankfurt, Germany). Lineage distribution was determined by FACS analysis (FACS Calibur, Becton Dickinson, Heidelberg, Germany) as previously described (43).

2.8.1.4 Immunofluorescence staining

Untreated and vitamin C treated (0.345mM; 12h) $IDH1^{R132H}$ cells were fixed in 4% paraformaldehyde for 15 min. After washing three times with PBS, cells were blocked with 5% FBS in PBST (PBS + 0.5% Tween 20) for 2 h at room temperature (20–25 °C). Primary antibody for 5-hydroxymethylcytosine (1:100, Active Motif, Cat #39769) was diluted in blocking solution and incubated with cells overnight at 4 °C. Cells were then washed three times in PBS and incubated for 2 h at room temperature with secondary antibodies diluted in blocking solution. Cells were washed three times in PBS and mounted (ProLong Diamond Antifade Mountant with DAPI, Molecular Probes) before imaging with Alexa Fluor 568-conjugated goat anti-mouse (1:1,000, Life Technologies) secondary antibody.

2.8.1.5 RNA-sequencing

Briefly, RNA was extracted from untreated and vitamin C treated (0.345mM added daily) *IDH1*^{R132H} cells 72h after initial treatment and subjected to polyA enrichment RNA-seq library construction according to standard operating procedures (http://www.epigenomes.ca/protocols-and-standards). RNA-seq library construction involves the following standard operating procedures (SOPs) in order: (1) Purification of polyA+ mRNA and mRNA(−) flowthrough total RNA using MultiMACS 96 separation unit; (2) strand-specific 96-well complementary DNA (cDNA) synthesis; (3) strand-specific 96-well library construction for Illumina sequencing. In brief, polyA+ RNA was purified using the MACS mRNA isolation kit (Miltenyi Biotec, Bergisch Gladbach, Germany), from 2–10 µg of total RNA with a RIN >=7 (Agilent Bioanalyzer) as per the manufacturer's instructions. The process included on-column DNaseI treatment (Invitrogen, Carlsbad, CA, USA). Double-stranded cDNA was synthesized from the purified polyA+ RNA using the Superscript II Double-Stranded cDNA Synthesis kit (Invitrogen) and 200 ng random hexamers (Invitrogen). After first strand synthesis, dNTPs were removed using 2 volumes of AMPure XP beads (Beckman Genomics, Danvers, MA, USA). GeneAmp (Invitrogen) 12.5 mM dNTPs blend (2.5 mM dCTP, 2.5 mM dGTP, 2.5 mM dATP, 5.0 mM dUTP) was used in the second-strand synthesis mixture in the presence of 2 µg of ActinomycinD. Double-stranded cDNA was purified using 2 volumes of Ampure XP beads, fragmented using Covaris E series shearing (20% duty cycle, Intensity 5, 55 s), and used for paired-end sequencing library preparation (Illumina). Before library amplification uridine digestion was performed at 37 °C for 30 min following with 10 min at 95 °C in Qiagen Elution buffer (10 mM Tris–Cl, pH 8.5) with 5 units of Uracil-N-Glycosylase (UNG: AmpErase). The resulting single-stranded sequencing library was amplified by PCR (10–13 cycles) to add

Illumina P5 and P7 sequences for cluster generation. PCR products were purified on Qiaquick MinElute columns (Qiagen, Mississauga, ON) and assessed and quantified using an Agilent DNA 1000 series II assay and Qubit fluorometer (Invitrogen, Carlsbad, CA), respectively. Libraries were pooled and sequenced on the Illumina NextSeq 500 platform.

2.8.1.6 Methylated (meDIP) and hydroxyl-methylated (hmeDIP) DNA immunoprecipitation sequencing

Genomic DNA was harvested from untreated and vitamin C treated (0.345mM added daily) IDH1^{R132H} cells 72h after initial treatment and subjected to low-input meDIP-seq and hmeDIP-seq as previously described (73). Briefly, genomic DNA was isolated from untreated and vitamin C treated HOXA9-IDH1^{WT} and HOXA9-IDH1^{R132H} after 72h and sonicated to approximately 200–400 bp with a Covaris E-220 Sonicator (Covaris) before end-repair and adapter ligation. 40ng of adapter ligated DNA was immuno-precipitated with anti-5-methylcytidine antibody (1µg/mL; Eurogentec; BI-MECY-01000) or an anti-hydroxymethylcytosine (1µg/mL; Active Motif; #39769), respectively. To complete library preparation, two 10-cycle PCR reactions using 16 µl of DNA were performed for each sample with paired-end and indexed Illumina PCR primers. A final size selection was performed to select all fragments <700 bp in size by electrophoresis in 8% polyacrylamide gel. The indexed libraries were pooled and sequenced as paired-end 125nt reads on the Illumina MiSeq and Illumina Hiseq 2500 platforms (v3 chemistry).

2.8.1.7 Whole genome bisulfite sequencing

Genomic DNA was harvested from untreated and vitamin C treated (0.345mM added daily) *IDH1^{R132H}* cells 72h after initial treatment. Bisulfite conversion was performed using the MethylEdge Kit (Promega) according to the manufacturer recommendations. To monitor under

and over conversion, 1ng of fully methylated T7 phage and unmethylated lambda DNA were spiked into 100ng of sample genomic DNA prior to bisulfite conversion. The resulting pool was suspended in 5ul of EB buffer, added to 32.5 µl of MethylEdge Bisulfite Conversion Reagent and incubated at 98°C for 6 min followed by 54°C for 60 min. Bisulfite converted DNA was isolated with MagSi-DNA all-around beads (MagNa Medics) and desulfonation was performed on the beads for 15 min before 2 washes 80% ethanol and elution for 15 min @ 56C. cDNA was generated from the resulting single stranded template by adding 2.5 µl of 10 mM dNTP mix (NEB) and 2 µl of 500 µM random primer to the DNA and denaturing at 98°C for 1 min before quickly moving to ice and adding 5 µl of 10X NEB2 buffer (NEB) and 1 µl of 50 U/µl Klenow fragment (3′→5′ exo−) (NEB). The mixture was then incubated as follows: 4°C for 30 min; 4°C-37°C (1°C/min); 37°C for 30 min; 70°C for 10 min. After cDNA synthesis the DNA was purified using SeraMag magnetic beads (20% PEG; 1M NaOH) a 75ul**:**50ul bead:DNA ratio, and in 35 µl before standard Illumina library preparation beginning at the adapter ligation stage. Completed libraries were pooled and sequencing on the HiSeq 2500 sequencing platform (v3 chemistry) following the manufacturer's protocols (Illumina, Hayward CA.). Targeted bisulfite amplicons were created to validate several DMRs. Genomic DNA was extracted from three biological replicate vitamin C treated and untreated samples as described above. This DNA was bisulfite converted as per the process described above and subjected to PCR using bisulfite converted primers targeting deMRs identified by WGBS. These amplicons were tested for proper size using an Agilent Bioanalyzer, purified using SeraMag beads and pooled for sequencing on the Illumina Miseq sequencing platform (v3 chemistry) following the manufacturers protocols.

2.8.1.8 Histone mark chromatin immunoprecipitation sequencing (ChIP-seq)

ChIP-seq was performed on chromatin harvested from untreated and vitamin C treated (1mM added at day 0) IDH1^{R132H} cells. Standard operating procedures for native ChIP-seq library construction are also available (http://www.epigenomes.ca/protocols-and-standards) or by request. Cells were lysed in mild non-ionic detergents (0.1% Triton X-100 and Deoxycholate) and protease inhibitor cocktail (Calbiochem) in order to preserve the integrity of histones harbouring epitope of interest during cell lysis. Cells were digested by Microccocal nuclease (MNase) at room temperature for 5 minutes and 0.25mM EDTA was used to stop the reaction. Antibodies, H3K4me1 (Diagenode: Catalogue #C15410037, lot A1657D), H3K4me3 (Cell Signaling Technology: Catalogue# 9751S, lot #8), and H3K27ac (Hiroshi Kimura, Cell Biology Unit, Tokyo Institute of Technology) were incubated with anti-IgA magnetic bead (Dynabeads from Invitrogen) for 2 hours. Digested chromatin was incubated with magnetic beads alone for 1.5 hours. Digested chromatin was separated from the beads and incubated with antibody-bead complex over night in IP buffer (20mM Tris-HCl pH 7.5, 2mM EDTA, 150mM NaCl, 0.1% Triton X-100, 0.1% Deoxycholate). The resulting IPs were washed 2 times by low salt (20mM Tris-HCl pH 8.0, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.1% SDS) and then high salt (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.1% SDS) wash buffers. IPs were eluted in elution buffer (1% SDS, 100mM Sodium Bicarbonate) for 1.5 hours at 65°C. Remaining histones were digested by Protease (Invitrogen) for 30min at 50°C and DNA fragments were purified using SeraMag magnetic beads in 30% PEG. Illumina sequencing libraries were generated by end repair, 3' A-addition, and Illumina sequencing adaptor ligation (New England BioLabs, E6000B-10). Libraries were then indexed, PCR amplified (10 cycles)

and sequenced on Illumina HiSeq 2500 platform following the manufacturer's protocols (Illumina, Hayward CA.).

2.8.1.9 Transcription factor chromatin immunoprecipitation sequencing (TF-ChIP-seq)

Formaldehyde cross-linked TF-ChIP-seq library construction was performed on chromatin harvested from untreated and 2-phosphate ascorbic acid treated (0.345mM added daily) $IDH1^{R132H}$ cells 72h after initial treatment using RUNX1 (Abcam; ab23980) and PU.1 (SCBT; sc-22805) antibodies as previously described (74) with minor modifications. Snap frozen cell pellets ($10e^7$ cells) were cross-linked with formaldehyde (1% final concentration) and washed by PBS. Cells were lysed and chromatin was sheared using the Covaris E-220 Sonicator (Covaris). Chromatin fragments were incubated with protein A/G Sepharose beads (GE Healthcare) for two hours at 4^oC to eliminate non-specific binding. A bead-antibody mixture was generated by combining protein A/G beads with primary antibody with mixing for 2h at 4°C. Unbound chromatin fragments were separated from the beads and incubated with a bead-Runx1 (Abcam; ab23980) or PU.1 (SCBT; sc-22805) at 4°C overnight in IP buffer (10mM Tris-HCl pH 7.5, 2mM EDTA, 90mM NaCl, 0.1% Triton X-100, 0.1% Deoxycholate, 0.1% SDS). After incubation, IPs were washed twice with ChIP wash buffer (20mM Tris-HCl pH 8.0, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.1% SDS) and once with final ChIP wash buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.1% SDS). IPs were eluted in elution buffer (1% SDS, 100mM Sodium Bicarbonate) and incubated at 68°C for 2 hours to reverse crosslinks. DNA fragments were stripped from histones and purified using QIAquick PCR Purification Kit (Qiagen) and subjected to 8 cycles of paired-end indexing PCR. Illumina sequencing libraries were generated as previously described. Libraries were PCR

amplified and sequenced on the Illumina NextSeq 500 platform following the manufacturer's protocols (Illumina, Hayward CA.)

2.8.2 Bioinformatics methods

2.8.2.1 RNA-sequencing

Sequencing reads were processed as previously described (75) (http://www.epigenomes.ca/data/CEMT/methods/RNA-Seq.html). Sequence reads were aligned to a transcriptome reference generated by JAGuaR (version 2.0.2) (76) using the mm10 reference genome supplemented by read-length-specific exon–exon junction sequences (Ensemble v71 gene annotations) and bam files were generated using SAMtools (77) (version 0.1.16). To quantify exon and gene expression reads per kilobase per million mapped reads (RPKM) metrics were calculated. Differentially expressed genes were identified using a custom DEfine matlab tool (FDR cutoff=0.05) considering only those genes with an RPKM > 0.01. Gene Ontology (GO) enrichment terms for DE genes were obtained using Immuno-Navigator (78) (Bonferonni p.value < 0.01) and KEGG (79) pathway analysis (Bonferonni p.value < 0.1).

2.8.2.2 hmeDIP and meDIP sequencing

Sequencing reads were aligned to mm10 using Burrows-Wheeler Aligner (80) (BWA; version 0.7.5a) and converted to the bam format by SAMtools(77) (version 0.1.16). PCR duplicates were marked using Picard (version 1.52). Pairwise DMRs were identified from aligned reads using the MEDIPS (58) R package (v1.16; 250bp window; p<0.01) and MACS (v1.4.2; p<1e-5). All gene ontology results for meDIP and WGBS DMRs were generated using GREAT(59) which associates DMRs with the 2 nearest genes within 20kb. A custom java program (bamToWig) was used to generate wig files for downstream analysis and visualization. Reads with BWA

mapping quality scores <5 were discarded and reads that aligned to the same genomic coordinate were counted only once in the profile generation.

2.8.2.3 Whole genome bisulfite sequencing

Bisulfite converted reads were aligned to a bisulfite converted version of mm10 (NovoIndex –b) using NovoAlign version 3.02.08 (Novocraft Technologies). Bisulfite conversion rates were > 99.5% for both libraries. Fractional methylation calls were generated using a custom shell script which leverages SAMtools(77) (version 0.1.16) mpileup command. DMRs were determined from WGBS data by scanning the genome for adjacent, differentially methylated CpGs (Δ0.25; $p < 0.05$) within 300bp windows. Genome profiles were generated calculating mean CpG methylation scores in specified bin sizes for CpGs with >8x coverage in both libraries. *De novo* enriched motifs were called using HOMER(61) version 4.7.2. The genomic distribution of different deMRs was generated using an in-house shell script, which utilized BEDTools version 2.24.0(81) to intersect deMRs with mm10 features from the UCSC genome browser(82). The DMRs were intersected with different feature sets in order of increasing genomic occupancy (ie. smallest feature to largest) until all DMRs had been assigned to a region. For the comparative analysis with previously published enhancers(67), H3K4me1 values at each enhancer in each cell type were downloaded from GEO (GSE60103) and utilized for subsequent analyses. The genomic coordinates for these enhancers were converted from mm9 to mm10 using the UCSC liftover tool(83) and intersected with deMRs and iMRs, with a random selection of 1000 enhancers used as a control.

2.8.2.4 Transcription factor ChIP-seq

Sequencing reads were aligned to mm10 using Burrows-Wheeler Aligner(80) (BWA; version 0.7.5a) and converted to the bam format by SAMtools (77) (version 0.1.16). Duplicate reads were marked using Picard (version 1.52). 20 million reads were randomly subsampled from each library using SAMtools (v0.1.2) before calling peaks MACS2 version 2.1.0.2 (narrow peak; p<1e-5) (57) and generating wig files (custom java script). Reads with BWA mapping quality scores <5 were discarded and reads that aligned to the same genomic coordinate were counted only once in the profile generation.

2.8.2.5 Histone modification ChIP-seq

Sequencing reads were aligned to mm10 using Burrows-Wheeler Aligner (80) (BWA; version 0.7.5a) and converted to the bam format by SAMtools (77) (version 0.1.16). Duplicate reads were marked using Picard (version 1.52). A custom java program (bamToWig) was used to generate wig files for downstream analysis and visualization. Reads with BWA mapping quality scores <5 were discarded and reads that aligned to the same genomic coordinate were counted only once in the profile generation. For ChiP-seq data, peaks were called using MACS2 version 2.1.0.2 ($p<1e^{-5}$) in broad mode for H3K4me1 and narrow mode for H3K4me3 and H3K27ac. Profiles were generated from wig files by calculating average coverage in 50bp bins using custom java scripts. $IDH1^{R132H}$ enhancers were called by intersecting H3K4me1 H3K27ac enriched regions with common enhancers removed to identify condition specific enhancers. For the comparative analysis with previously published enhancers(67), H3K4me1 values at each enhancer in each cell type were downloaded from GEO (GSE60103) and utilized for subsequent analyses. The genomic coordinates for these enhancers were converted from mm9 to mm10

using the UCSC liftover tool (83) and intersected with $IDH1^{R132H}$ enhancers with a random selection of 1000 enhancers used as a control.

2.8.2.6 Plots and statistical methods

All plots were generated in RStudio (84) version 3.2.3. Bedtools (81) was used for comparing and overlapping the genomic coordinates of peaks, DMRs and existing genomic features described in the manuscript. Boxplots represent the mean (centre line), the first and third quartiles (top and bottom or box, respectively) and confidence intervals (95%; black lines). P-values were generated in R using two-sided students t-tests (t-test) or two-sided Kolmogorov-Smirnov (K-S) tests. Browser screen shots were downloaded from the UCSC genome browser (83).

Chapter 3: Conclusion

In this thesis I present evidence supporting a model of vitC induced epigenomic remodelling that coincides with the differentiation and maturation of myeloid progenitor cells, a process frequently disrupted in myeloid malignancies (22). Using a murine leukemic model expressing $IDH1^{R132H}$ we show that vitC treatment promotes DNA demethylation and H3K4me1 deposition at enhancers implicated in myeloid differentiation. We show that regions demethylated upon vitC treatment are enriched in putative regulatory elements associated with genes implicated in hematopoiesis and leukemic transformation and describe several key hematopoietic genes that show increased expression and promoter demethylation. We observe significant overlap between regions that change methylation states upon vitC treatment (DMRs) and PU.1 binding sites, consistent with the previously described physical association between PU.1 and TET2 (71). These observations support a model in which PU.1 guides TET2 to a specific subset of genomic regions to facilitate the oxidation of 5mC to 5hmC, a reaction that is inhibited in the $IDH1^{R132H}$ context by *R*-2HG (56) and enhanced by vitamin C (33) (Figure 33). Although demethylation events were observed in association with PU.1 sites, PU.1 binding did not increase upon vitC treatment, consistent with the ability of PU.1 to bind regions marked by 5mC, unlike other ETS family members (85).

The oncogenic t(8;21) *AML-ETO9a* translocation is recurrent in AML and fuses *RUNX1* to *RUNX1T1* leading to aberrant HDAC2 recruitment and histone deacetylation(86). We observe a relationship between vitC induced DMRs and genes known to be dysregulated by this fusion protein and show that H3K27ac and H3K4me1 signal increases in the regions flanking vitC induced RUNX1 binding sites. Interestingly, many regions that acquire H3K4me1 and H3K27ac upon vitamin C treatment overlap with previously described enhancers that become activated

during myeloid development, but are devoid of deMRs. This suggests that vitamin C elicits multifaceted, myeloid specific, epigenomic reprogramming that is not restricted to changes in 5mC and 5hmC.

Historically, observational studies using oral and intravenous vitamin C originally suggested therapeutic benefit for a non-specific subset of cancer patients (87-89) but subsequent randomized, placebo-controlled trials revealed insignificant efficacy (90, 91). However, these latter studies utilized oral vitamin C, which induces transient micromolar increases in plasma levels(92) significantly below the concentration used in our study (0.345mM). Interestingly, a recent report in a small cohort of patients (n=24) with haematological malignancies showed that >90% had decreased plasma vitamin C concentrations and 58% were classified as vitamin C deficient (93). In contrast, only 7.1% of the general United States population are vitamin C deficient (94). Additionally, a recent case study in which a child with a chemotherapy resistant neurofibromatosis type 1 and optic nerve glioma was treated with intravenous vitamin C for 30 months (95). By the end of the treatment tumours in the optic chiasm, hypothalamus and left optic nerve had stabilized and the right optic nerve mass disappeared completely. Interestingly, gliomas and AMLs represent two cancer types with the highest frequency of recurrent, neomorphic IDH mutations and although the tumors were not genotyped the fact that $IDH1^{R132H}$ mutations are found in approximately 80% of grade 2 and grade 3 gliomas(96) makes it likely that a subpopulation within this patient could have harboured this mutation.

Our study provides mechanistic insight into a “first in class” epigenetic modulator capable of re-activating an epigenetic pathway recurrently disrupted in cancer. This mechanistic insight, combined with our refined understanding of the role of epigenetic disruption in leukemia, suggests that further investigation of the effects of millimolar concentrations of

vitamin C in AML patients harbouring IDH and TET2 mutations, achievable in plasma through intravenous administration (97), is warranted. Indeed, the lack of an appropriate, genotypically stratified cohort of patients may have contributed to the previously observed patient dependent response to vitamin C treatment.

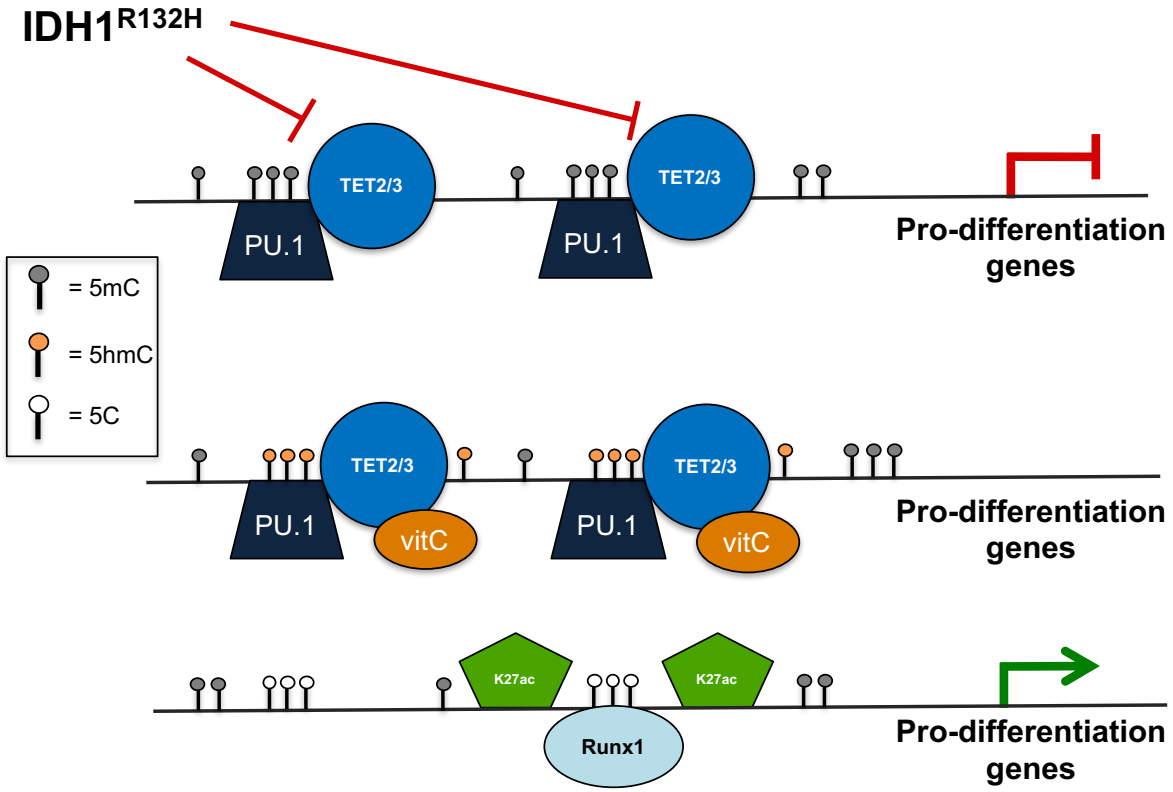

Figure 33. A proposed model of vitamin C induced epigenetic remodeling.

Bibliography

1. Waddington CH. The Epigenotype. *Endeavour* 1942;: 18-20.

2. Riggs AD, Porter TN. Overview of epigenetic mechanisms. *Cold Spring Harbor Monograph Archive* 1996; **32**: 29-45.

3. Liu X, Gao Q, Li P, Zhao Q, Zhang J, Li J, *et al*. UHRF1 targets DNMT1 for DNA methylation through cooperative binding of hemi-methylated DNA and methylated H3K9. *Nature communications* 2013; **4**: 1563.

4. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, *et al*. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 2009; **324**: 930-5.

5. Iyer LM, Tahiliani M, Rao A, Aravind L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell cycle* 2009; **8**: 1698-710.

6. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, *et al*. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 2011; **333**: 1303-7.

7. Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, *et al*. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat Genet* 2013; **45**: 1198-206.

8. Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Schöler A, *et al*. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* 2011;.

9. Hodges E, Molaro A, Dos Santos CO, Thekkat P, Song Q, Uren PJ, *et al*. Directional DNA methylation changes and complex intermediate states accompany lineage specificity in the adult hematopoietic compartment. *Mol Cell* 2011; **44**: 17-28.

10. Schbeler D. Function and information content of DNA methylation. *Nature* 2015; **517**: 321-6.

11. Schübeler D. Function and information content of DNA methylation. *Nature* 2015; **517**: 321-6.

12. Johnson DS, Mortazavi A, Myers RM, Wold B. Genome-wide mapping of in vivo protein-DNA interactions. *Science* 2007; **316**: 1497-502.

13. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? *Mol Cell* 2013; **49**: 825-37.

14. Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, *et al*. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Current Biology* 2003; **13**: 1192-200.

15. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, *et al*. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 2006; **441**: 349-53.

16. Barski A, Cuddapah S, Cui K, Roh T, Schones DE, Wang Z, *et al*. High-resolution profiling of histone methylations in the human genome. *Cell* 2007; **129**: 823-37.

17. Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, *et al*. Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* 2015; **520**: 243-7.

18. Hu D, Shilatifard A. Epigenetics of hematopoiesis and hematological malignancies. *Genes Dev* 2016; **30**: 2021-41.

19. Trowbridge JJ, Snow JW, Kim J, Orkin SH. DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell stem cell* 2009; **5**: 442-9.

20. Challen GA, Sun D, Jeong M, Luo M, Jelinek J, Berg JS, *et al*. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet* 2012; **44**: 23-31.

21. Yamamoto JF, Goodman MT. Patterns of leukemia incidence in the United States by subtype and demographic characteristics, 1997–2002. *Cancer Causes & Control* 2008; **19**: 379-90.

22. De Kouchkovsky I, Abdul-Hay M. Acute myeloid leukemia: a comprehensive review and 2016 update. *Blood Cancer Journal* 2016; **6**: e441.

23. Chan SM, Majeti R. Role of DNMT3A, TET2, and IDH1/2 mutations in pre-leukemic stem cells in acute myeloid leukemia. *Int J Hematol* 2013; **98**: 648-57.

24. Network CGAR. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 2013; **368**: 2059.

25. Losman J, Kaelin WG. What a difference a hydroxyl makes: mutant IDH,(R)-2-hydroxyglutarate, and cancer. *Genes Dev* 2013; **27**: 836-52.

26. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, *et al*. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer cell* 2010; **18**: 553-67.

27. Chaturvedi A, Cruz MA, Jyotsana N, Sharma A, Goparaju R, Schwarzer A, *et al*. Enantiomer-specific and paracrine leukemogenicity of mutant IDH metabolite 2-hydroxyglutarate. *Leukemia* 2016;.

28. Losman JA, Looper RE, Koivunen P, Lee S, Schneider RK, McMahon C, *et al*. (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. *Science* 2013; **339**: 1621-5.

29. Gaidzik VI, Paschka P, Spath D, Habdank M, Kohne CH, Germing U, *et al*. TET2 mutations in acute myeloid leukemia (AML): results from a comprehensive genetic and clinical analysis of the AML study group. *J Clin Oncol* 2012; **30**: 1350-7.

30. Abbas S, Lugthart S, Kavelaars FG, Schelen A, Koenders JE, Zeilemaker A, *et al*. Acquired mutations in the genes encoding IDH1 and IDH2 both are recurrent aberrations in acute myeloid leukemia: prevalence and prognostic value. *Blood* 2010; **116**: 2122-6.

31. Monfort A, Wutz A. Breathingâ€ in epigenetic change with vitamin C. *EMBO Rep* 2013; **14**: 337-46.

32. Hutton JJ, Tappel A, Udenfriend S. Requirements for α-ketoglutarate, ferrous ion and ascorbate by collagen proline hydroxylase. *Biochem Biophys Res Commun* 1966; **24**: 179-84.

33. Yin R, Mao S, Zhao B, Chong Z, Yang Y, Zhao C, *et al*. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc* 2013; **135**: 10396-403.

34. Chen J, Liu H, Liu J, Qi J, Wei B, Yang J, *et al*. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nat Genet* 2013; **45**: 34-42.

35. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martnez JA, Goyal P, Mahapatra S, *et al*. Vitamin [thinsp] C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature* 2013; **500**: 222-6.

36. Klose RJ, Yamane K, Bae Y, Zhang D, Erdjument-Bromage H, Tempst P, *et al*. The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* 2006; **442**: 312-6.

37. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, *et al*. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* 2006; **439**: 811-6.

38. Chen Q, Espey MG, Krishna MC, Mitchell JB, Corpe CP, Buettner GR, *et al*. Pharmacologic ascorbic acid concentrations selectively kill cancer cells: action as a pro-drug to deliver hydrogen peroxide to tissues. *Proc Natl Acad Sci U S A* 2005; **102**: 13604-9.

39. Kawada H, Kaneko M, Sawanobori M, Uno T, Matsuzawa H, Nakamura Y, *et al*. High concentrations of L-ascorbic acid specifically inhibit the growth of human leukemic cells via downregulation of HIF-1α transcription. *PloS one* 2013; **8**: e62717.

40. Du J, Cullen JJ, Buettner GR. Ascorbic acid: chemistry, biology and the treatment of cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* 2012; **1826**: 443-57.

41. Takamizawa S, Maehata Y, Imai K, Senoo H, Sato S, Hata R. Effects of ascorbic acid and ascorbic acid 2-phosphate, a long-acting vitamin C derivative, on the proliferation and differentiation of human osteoblast-like cells. *Cell Biol Int* 2004; **28**: 255-65.

42. Calvo KR, Sykes DB, Pasillas M, Kamps MP. Hoxa9 immortalizes a granulocyte-macrophage colony-stimulating factor-dependent promyelocyte capable of biphenotypic differentiation to neutrophils or macrophages, independent of enforced meis expression. *Mol Cell Biol* 2000; **20**: 3274-85.

43. Chaturvedi A, Cruz MMA, Jyotsana N, Sharma A, Yun H, Grlich K, *et al*. Mutant IDH1 promotes leukemogenesis in vivo and can be specifically targeted in human AML. *Blood* 2013; **122**: 2877-87.

44. Austria R, Semenzato A, Bettero A. Stability of vitamin C derivatives in solution and topical formulations. *J Pharm Biomed Anal* 1997; **15**: 795-801.

45. Holmes C, Stanford WL. Concise review: stem cell antigen-1: expression, function, and enigma. *Stem Cells* 2007; **25**: 1339-47.

46. Solovjov DA, Pluskota E, Plow EF. Distinct roles for the alpha and beta subunits in the functions of integrin alphaMbeta2. *J Biol Chem* 2005; **280**: 1336-45.

47. Austyn JM, Gordon S. F4/80, a monoclonal antibody directed specifically against the mouse macrophage. *Eur J Immunol* 1981; **11**: 805-15.

48. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C*, et al*. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer cell* 2011; **20**: 11-24.

49. Zhang CC, Lodish HF. Cytokines regulating hematopoietic stem cell function. *Curr Opin Hematol* 2008; **15**: 307-11.

50. de Groot RP, Coffer PJ, Koenderman L. Regulation of proliferation, differentiation and survival by the IL-3/IL-5/GM-CSF receptor family. *Cell Signal* 1998; **10**: 619-28.

51. Stanley ER, Chitu V. CSF-1 receptor signaling in myeloid cells. *Cold Spring Harb Perspect Biol* 2014; **6**: 10.1101/cshperspect.a021857.

52. Wang Q, Li N, Wang X, Shen J, Hong X, Yu H*, et al*. Membrane protein hMYADM preferentially expressed in myeloid cells is up-regulated during differentiation of stem cells and myeloid leukemia cells. *Life Sci* 2007; **80**: 420-9.

53. Argiropoulos B, Yung E, Humphries RK. Unraveling the crucial roles of Meis1 in leukemogenesis and normal hematopoiesis. *Genes Dev* 2007; **21**: 2845-9.

54. Radomska HS, Alberich-Jorda M, Will B, Gonzalez D, Delwel R, Tenen DG. Targeting CDK1 promotes FLT3-activated acute myeloid leukemia differentiation through C/EBPalpha. *J Clin Invest* 2012; **122**: 2955-66.

55. Edling CE, Hallberg B. c-Kit—a hematopoietic cell essential receptor tyrosine kinase. *Int J Biochem Cell Biol* 2007; **39**: 1995-8.

56. Xu W, Yang H, Liu Y, Yang Y, Wang P, Kim S*, et al*. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases. *Cancer cell* 2011; **19**: 17-30.

57. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE*, et al*. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008; **9**: 1.

58. Lienhard M, Grimm C, Morkel M, Herwig R, Chavez L. MEDIPS: genome-wide differential coverage analysis of sequencing data derived from DNA enrichment experiments. *Bioinformatics* 2014; **30**: 284-6.

59. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB*, et al*. GREAT improves functional interpretation of cis-regulatory regions. *Nat Biotechnol* 2010; **28**: 495-501.

60. Domcke S, Bardet AF, Ginno PA, Hartl D, Burger L, Schbeler D. Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* 2015;.

61. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P*, et al*. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010; **38**: 576-89.

62. Goode DK, Obier N, Vijayabaskar M, Lie-A-Ling M, Lilly AJ, Hannah R, *et al*. Dynamic Gene Regulatory Networks Drive Hematopoietic Specification and Differentiation. *Developmental cell* 2016; **36**: 572-87.

63. Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, *et al*. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 2008; **133**: 1106-17.

64. Nitzsche A, Paszkowski-Rogacz M, Matarese F, Janssen-Megens EM, Hubner NC, Schulz H, *et al*. RAD21 cooperates with pluripotency transcription factors in the maintenance of embryonic stem cell identity. *PloS one* 2011; **6**: e19470.

65. Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, *et al*. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev* 2015; **29**: 910-22.

66. Wiench M, John S, Baek S, Johnson TA, Sung MH, Escobar T, *et al*. DNA methylation status predicts cell type-specific enhancer activity. *EMBO J* 2011; **30**: 3028-39.

67. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, Zaretsky I, Jaitin DA, David E, *et al*. Chromatin state dynamics during blood formation. *Science* 2014; **345**: 943-9.

68. Feng R, Desbordes SC, Xie H, Tillo ES, Pixley F, Stanley ER, *et al*. PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells. *Proc Natl Acad Sci U S A* 2008; **105**: 6057-62.

69. Zhou J, Wu J, Li B, Liu D, Yu J, Yan X, *et al*. PU. 1 is essential for MLL leukemia partially via crosstalk with the MEIS/HOX pathway. *Leukemia* 2014; **28**: 1436-48.

70. Okuda T, Nishimura M, Nakao M, Fujitaa Y. RUNX1/AML1: a central player in hematopoiesis. *Int J Hematol* 2001; **74**: 252-7.

71. de la Rica L, Rodríguez-Ubreva J, García M, Islam AB, Urquiza JM, Hernando H, *et al*. PU. 1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol* 2013; **14**: 1.

72. Wang L, Gural A, Sun XJ, Zhao X, Perna F, Huang G, *et al*. The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science* 2011; **333**: 765-9.

73. Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D, *et al*. Methylome analysis using MeDIP-seq with low DNA concentrations. *Nature protocols* 2012; **7**: 617-36.

74. Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, Odom DT. ChIP-seq: using high-throughput sequencing to discover protein–DNA interactions. *Methods* 2009; **48**: 240-8.

75. Gascard P, Bilenky M, Sigaroudinia M, Zhao J, Li L, Carles A, *et al*. Epigenetic and transcriptional determinants of the human breast. *Nature communications* 2015; **6**.

76. Butterfield YS, Kreitzman M, Thiessen N, Corbett RD, Li Y, Pang J, *et al*. JAGuaR: junction alignments to genome for RNA-seq reads. *PloS one* 2014; **9**: e102398.

77. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al*. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078-9.

78. Vandenbon A, Dinh VH, Mikami N, Kitagawa Y, Teraguchi S, Ohkura N, *et al*. Immuno-Navigator, a batch-corrected coexpression database, reveals cell type-specific gene networks in the immune system. *Proc Natl Acad Sci U S A* 2016; **113**: E2393-402.

79. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000; **28**: 27-30.

80. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009; **25**: 1754-60.

81. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010; **26**: 841-2.

82. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, *et al*. The UCSC Genome Browser Database. *Nucleic Acids Res* 2003; **31**: 51-4.

83. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, *et al*. The human genome browser at UCSC. *Genome Res* 2002; **12**: 996-1006.

84. Team R. RStudio: integrated development for R. *RStudio, Inc., Boston, MA.URL* http://www.RStudio.com/ide 2014;.

85. Stephens DC, Poon GM. Differential sensitivity to methylated DNA by ETS-family transcription factors is intrinsically encoded in their DNA-binding domains. *Nucleic Acids Res* 2016;.

86. Hu Z, Gu X, Baraoidan K, Ibanez V, Sharma A, Kadkol S, *et al*. RUNX1 regulates corepressor interactions of PU.1. *Blood* 2011; **117**: 6498-508.

87. Cameron E, Campbell A. The orthomolecular treatment of cancer II. Clinical trial of high-dose ascorbic acid supplements in advanced human cancer. *Chem Biol Interact* 1974; **9**: 285-315.

88. Cameron E, Pauling L. Supplemental ascorbate in the supportive treatment of cancer: Prolongation of survival times in terminal human cancer. *Proc Natl Acad Sci U S A* 1976; **73**: 3685-9.

89. Cameron E, Pauling L. Supplemental ascorbate in the supportive treatment of cancer: reevaluation of prolongation of survival times in terminal human cancer. *Proc Natl Acad Sci U S A* 1978; **75**: 4538-42.

90. Creagan ET, Moertel CG, O'Fallon JR, Schutt AJ, O'Connell MJ, Rubin J, *et al*. Failure of high-dose vitamin C (ascorbic acid) therapy to benefit patients with advanced cancer: a controlled trial. *N Engl J Med* 1979; **301**: 687-90.

91. Moertel CG, Fleming TR, Creagan ET, Rubin J, O'Connell MJ, Ames MM. High-dose vitamin C versus placebo in the treatment of patients with advanced cancer who have had no prior chemotherapy: a randomized double-blind comparison. *N Engl J Med* 1985; **312**: 137-41.

92. Levine M, Padayatty SJ, Espey MG. Vitamin C: a concentration-function approach yields pharmacology and therapeutic discoveries. *Adv Nutr* 2011; **2**: 78-88.

93. Liu M, Ohtani H, Zhou W, Orskov AD, Charlet J, Zhang YW, *et al*. Vitamin C increases viral mimicry induced by 5-aza-2'-deoxycytidine. *Proc Natl Acad Sci U S A* 2016;.

94. Schleicher RL, Carroll MD, Ford ES, Lacher DA. Serum vitamin C and the prevalence of vitamin C deficiency in the United States: 2003-2004 National Health and Nutrition Examination Survey (NHANES). *Am J Clin Nutr* 2009; **90**: 1252-63.

95. Mikirova N, Hunnunghake R, Scimeca RC, Chinshaw C, Ali F, Brannon C, *et al*. High-Dose Intravenous Vitamin C Treatment of a Child with Neurofibromatosis Type 1 and Optic Pathway Glioma: A Case Report. *Am J Case Rep* 2016; **17**: 774-81.

96. Cohen AL, Holmen SL, Colman H. IDH1 and IDH2 mutations in gliomas. *Current neurology and neuroscience reports* 2013; **13**: 1-7.

97. Padayatty SJ, Sun H, Wang Y, Riordan HD, Hewitt SM, Katz A, *et al*. Vitamin C pharmacokinetics: implications for oral and intravenous use. *Ann Intern Med* 2004; **140**: 533-7.