

The biological impact of novel dual histone methyltransferase inhibitors

Ian Lewis Green

MSci (Hons) University of Aberdeen

CID: 00718112

Division of Cancer

Department of Surgery & Cancer

Imperial College London

Thesis submitted for the degree of Doctor of Philosophy

2015

Declaration of Originality

I, Ian Green, hereby declare that this PhD thesis is my own work. In the preparation of this manuscript all references have been consulted by me. Except where specifically stated, the work presented in this thesis was performed by me.

Copyright Declaration

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives license. Researchers are free to copy distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the license terms of this work.

Abstract

Background: EZH2 is a histone methyltransferase (HKMT) responsible for the maintenance of epigenetic silencing of genes through maintenance of the repressive H3K27me3 mark and it is aberrantly regulated in numerous cancers, including breast cancer where it is linked to aggressive phenotypes and poor clinical outcomes. EHMT2 is a related HKMT responsible for gene silencing by mediating H3K9me3 levels. EHMT2 is also responsible for H3K27me1 and has been shown to physically interact with EZH2. Specific inhibitors of EZH2 are available and have been shown to be effective in cancers with EZH2 mutation driven phenotypes (e.g. follicular lymphoma) but have shown limited efficacy in epithelial cancers. Here we present the characterisation of novel dual HKMT inhibitors targeting both EZH2 and EHMT2, which we believe will have a greater impact than individual inhibitors in reversing EZH2 mediated silencing.

Results: Utilising publicly available data, we show expression of EZH2 and related subunits of the PRC2 complex and related EHMT2/EHMT1 complex range greatly in normal tissue, but EZH2 and EHMT2 expression are consistently up-regulated in numerous cancers. We show that CNV and mutation of EZH2 and EHMT2 infrequently occur in breast cancer- however, in breast cancer high expression of EZH2 is linked to reduced RFS and OS of patients. In breast cancer cell lines, dual HKMT inhibitors up-regulate EZH2 target genes, in gene specific and genome wide manner, to a greater degree than EZH2 or EHMT2 inhibition alone and induce expression of genes associated with apoptotic pathways. This up-regulation of silenced genes occurs concurrently with a decrease in H3K27me3 and H3K9me3 levels on target genes. In breast cancer cells and ovarian cancer cells, dual HKMT inhibitors reduce cell clonogenicity,

cancer stem cell activity, cancer stem cell self-renewal capacity, and sensitise cancer stem cells to Paclitaxel and Cisplatin treatment.

Conclusions: Novel dual inhibitors of EZH2 and EHMT2 alter gene expression and inhibit cell growth and cancer stem cell activity in wild-type EZH2 tumour cells. These data support the further preclinical and clinical evaluation of such inhibitors in triple negative breast cancer and epithelial ovarian cancer.

Acknowledgements

I would like to acknowledge foremost Professor Bob Brown and Dr Ed Curry, who have supervised me throughout this project. Their unceasing support, guidance, and belief have allowed this project to move forward- I cannot express my gratitude enough.

Nadine Chapman-Rothe acted as my secondary supervisor during the initial phases of this project and provided help and collaboration with ChIP-PCR experiments, and was succeeded by Constanze Zeller whose enthusiasm and support was a great resource. Elham Shamsaei and Sarah Kandil both worked a great deal on this project, and helped drive it forward to where it is now. MRes students Emma Bell and Luke Payne both worked on this project as part of their studies, and their input is something for which I am very grateful.

Collaborators Anthony Uren from the MRC Clinical Sciences Centre, Gillian Farnie and Amrita Shergill from University of Manchester all provided wonderful expertise in their fields and their collaboration allowed this project to move in interesting and exciting directions. Any acknowledgements to specific experimental work are highlighted within this thesis.

Fanny Cherblanc, Thota Ganesh, Nitipol Srimongkolpithak, Joachim Caron, Fengling Li, James P Snyder, Masoud Vedadi, and Pete Dimaggio have all worked around the chemistry of these novel inhibitors, and without them this work would not have been possible- their efforts were orchestrated by Matt Fuchter, whose enthusiasm for the project has helped unearth many avenues of subsequent research.

The wonderful collection of postdocs in the epigenetics group (or nearby...) were an invaluable source of knowledge, ideas, and coffee- Erick Loomis, Kirsty Flower, Charlotte Wilhelm-Benartzi, Paula Cunnea, Elaina Maginn, Fieke Froeling, Nair Bonito- thank you all.

Fellow students Jane Borley, Natalie Shenker, Angela Wilson, Kevin Brennan, Alun Passy, Kayleigh Davis and David Phelps have all be lovely with their time and feedback and friendship.

Nahal Masrour has been a constantly helpful presence, and James Flanagan has been more than helpful with his input and critical eye.

CRUK provided me with my studentship, administered by Jennifer Podesta, without which this work would have been impossible, and OCA and Imperial College provided me with the space, environment, and colleagues which allowed this work to be completed. Copenhagen Biosciences subsidised my attendance to the Copenhagen Biosciences Stem Cell Niche conference 2014 in Copenhagen, which was a wonderful opportunity to see some first class research.

My parents and brothers have shown unfailing support and encouragement and patience, and their belief has been a continuing source of comfort and resilience.

Finally Abi, without who I would have probably died of scurvy at about the 18 month mark, and for all the other obvious reasons.

Contents

Declaration of Originality.....	2
Copyright Declaration	2
Abstract.....	3
Acknowledgements	5
Contents	7
List of figures.....	12
List of tables	13
Abbreviations.....	14
Peer reviewed publications and presentations	17
Chapter 1: Introduction	18
1.1 Overview of epigenetics and cancer	18
1.1.1- Overview	18
1.1.2- Epigenetic therapies and pathways in cancer	19
1.2 The HKMT EZH2	20
1.2.1- H3K27me3 and HKMTs	20
1.3 EZH2 and cancer	22
1.3.1 EZH2 and cancer	22
1.3.2 EZH2 and EHMT2	25

1.4 Cancer stem cells and EZH2	27
1.5 Identification of novel dual HKMT	30
Hypothesis	34
Aims	34
Chapter 2: Materials and methods	35
Cell culture	35
RNA preparation	35
QRT-PCR	37
Compound batch data	39
Calculation of differential expression (Harvard Centre for Computational & Integrative biology)	39
Correlation analysis	41
CancerMA Forest Plots	41
Mutation rate, CNV, and expression of target genes in TCGA data	42
Comparison of gene expression, clinical data, and CNV in TCGA data	42
Cox proportional hazard modelling	43
Survival analysis utilising combined data sources	43
Gene expression microarray	44
Enrichment analysis	44
Correlation of gene expression after compound treatment	46
ConsensusPathDB pathway enrichment analysis	46
SiRNA knockdown experiments	47

Chromatin immunoprecipitation.....	47
Cell proliferation assay	51
Clonogenic assay	52
CSC activity and self-renewal capacity	52
Xenograft culture	55
Secondary xenograft culture	55
Extreme limiting dilution analysis.....	55
Chapter 3: Evaluation of EZH2 and EHMT2 as therapeutic targets in cancer utilising publicly available data	56
3.1 Introduction.....	56
3.2 Expression in normal tissues of EZH2, EHMT2, and related genes	59
3.3 Expression of EZH2 and EHMT2 in cancerous tissues.....	67
3.4 Mutations in EZH2 and EHMT in cancerous tissues.....	71
3.5 EZH2 and EHMT2 CNV in cancerous tissues	77
3.6 Relationship between target gene CNV, target gene expression, and clinical characteristics in cancerous tissues.....	79
3.7 Target gene expression and survival.....	86
3.8 Summary.....	93
Chapter 4: Impact of novel dual HKMT inhibitors on the epigenetic state of cancer cells	95
4.1 Introduction and Aims	95
4.2 Impact of dual HKMT inhibitors on EZH2 target gene expression.....	97

4.3	Comparison of inhibitors' impact on gene expression	110
4.4	Functional signatures of dual HKMT inhibition.....	114
4.5	Identification of putative pharmacodynamic biomarkers & examination of chromatin state of target genes after dual HKMT inhibition.....	116
4.6	Summary.....	126
Chapter 5: Effect of dual HKMT inhibition on cancer cell phenotype and cancer stem cells ..		129
5.1	Introduction.....	129
5.2	Effect of dual HKMT inhibition on cancer cell proliferation	131
5.3	Effect of dual HKMT inhibition on cancer stem cell activity, self-renewal, and chemosensitivity in <i>in vitro</i> models.....	134
5.4	Effect of dual HKMT inhibition on cancer stem cell activity, self-renewal, and chemosensitivity in <i>in vivo</i> models.....	150
5.5	Summary.....	155
Chapter 6: General discussion		159
Chapter 6: Discussion.....		159
6.1-	Introduction	159
6.2-	Evaluation of EZH2/EHMT2 as targets utilising publicly available data.....	161
6.2.1-	Discussion	161
6.2.2-	Future work	163
6.3-	Impact of novel dual HKMT inhibitors on the epigenetic state of cancer cells.....	165
6.3.1-	Discussion	165
6.3.2-	Future work	166

6.4- Effect of dual HKMT inhibition on cancer cell phenotype and cancer stem cells.....	168
6.4.1- Discussion	168
6.4.2- Future work	169
6.5- General discussion and Conclusions	171
6.5.1- General discussion.....	171
6.5.2- Conclusions	173
Chapter 7: List of references	174
Chapter 8: Supplementary data.....	185
APPENDIX	211

List of figures

Figure	Page number	Figure	Page number
Figure 1.1	21	Figure 4.8	112
Figure 1.2	26	Figure 4.9	117
Figure 1.3	28	Figure 4.10	119
Figure 1.4	33	Figure 4.11	120
Figure 2.1	48	Figure 4.12	121
Figure 2.2	59	Figure 4.13	122
Figure 3.1	60	Figure 4.14	124
Figure 3.2	62	Figure 4.15	132
Figure 3.3	64	Figure 5.1	135
Figure 3.4	67	Figure 5.2	137
Figure 3.5	68	Figure 5.3	139
Figure 3.6	69	Figure 5.4	141
Figure 3.7	75	Figure 5.5	143
Figure 3.8	79	Figure 5.6	145
Figure 3.9	90	Figure 5.7	147
Figure 3.10	99	Figure 5.8	150
Figure 4.1	102	Figure 5.9	151
Figure 4.2	104	Figure 5.10	152
Figure 4.3	105	Figure 5.11	153
Figure 4.4	107	Figure 5.12	185
Figure 4.5	108	Figure 8.2	186
Figure 4.6	110	Figure 8.3	206
Figure 4.7	111	Figure 8.4	208

List of tables

Table	Page Number		
Table 1.1	24	Table 5.2	132
Table 1.2	25	Table 5.3	136
Table 1.3	27	Table 5.4	137
Table 2.1	31	Table 5.5	138
Table 2.2	41	Table 5.6	142
Table 3.1	72	Table 5.7	143
Table 3.2	74	Table 5.8	146
Table 3.3	76	Table 5.9	153
Table 3.4	77	Table 5.10	154
Table 3.5	80	Table 8.1	184
Table 3.6	81	Table 8.2	187
Table 3.7	83	Table 8.3	188
Table 3.8	85	Table 8.4	189
Table 3.9	86	Table 8.5	190
Table 3.10	88	Table 8.6	191
Table 3.11	88	Table 8.7	192
Table 3.12	89	Table 8.8	194
Table 3.13	91	Table 8.9	195
Table 3.14	91	Table 8.10	205
Table 4.1	97	Table 8.11	206
Table 4.2	98	Table 8.12	207
Table 4.3	114	Table 8.13	208
Table 4.4	115	Table 8.14	209
Table 5.1	130	Table 8.15	209

Abbreviations

ABC	ATP Binding Cassette Family
ARID1A	AT-rich interactive domain-containing protein 1A
BCL2	B-cell lymphoma 2
BMI1	B cell-specific Moloney murine leukaemia virus integration site 1
CBX	Chromobox Protein Homolog 1
CD133	Prominin-1
ChIP	Chromatin immunoprecipitation
ChIP-PCR	Chromatin immunoprecipitation polymerase chain reaction
ChiSq	Chi-square
CNS	Central nervous system
CNV	Copy number variation
CpG	Cytosine-phosphate-Guanine
CRUK	Cancer Research UK
CSC	Cancer stem cell
CYP1B1	Cytochrome P450 1B1
CYP3A43	Cytochrome P450 3A43
DB	Diffuse B lymphoblast large cell lymphoma
DF	Density function
DMEM	Dulbecco's Modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DNMT1	DNA (cytosine-5)-methyltransferase 1
DNMT3A	DNA (cytosine-5)-methyltransferase 3a
DNMT3B	DNA (cytosine-5)-methyltransferase 3b
DOHH2	Non-Hodgkin's lymphoma cell line
DOT1	Disruptor of telomeric silencing
DZNep	3-Deazaneplanocin A hydrochloride
EED	Embryonic ectoderm development
EHMT1	Euchromatic histone-lysine N-methyltransferase 1
EHMT2	Euchromatic histone-lysine N-methyltransferase 2
ELDA	Extreme limiting dilution analysis
ER	Oestrogen receptor
ES	Embryonic stem cell

EZH2	Histone-lysine N-methyltransferase
FBS	Fetal bovine serum
FBXO32	F-box protein 32
G9a	Histone-lysine N-methyltransferase, H3 lysine-9 specific 3
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GATA4	GATA binding protein 4
GISTIC	Genomic Identification of Significant Targets in Cancer
H2AK119	Histone 2A lysine 119
H2AK119ub1	Histone 2A lysine 119 monoubiquitination
H3	Histone 3
H3K27	Histone 3 lysine 27
H3K27me1	Histone 3 lysine 27 monomethylation
H3K27me2	Histone 3 lysine 27 dimethylation
H3K27me3	Histone 3 lysine 27 trimethylation
H3K9	Histone 3 lysine 9
H3K9me1	Histone 3 lysine 9 monomethylation
H3k9me2	Histone 3 lysine 9 dimethylation
H3K9me3	Histone 3 lysine 9 trimethylation
HDAC9	Histone deacetylase 9
HGNC	HUGO Gene Nomenclature Committee
HKMT	Histone methyltransferase
HKMT-I-005	Histone methyltransferase inhibitor 5
HKMT-I-011	Histone methyltransferase inhibitor 11
HKMT-I-022	Histone methyltransferase inhibitor 22
IC50	Inhibitory concentration 50%
IGROV1	Ovarian carcinoma cell line
IL24	Interleukin 24
IP	Immunoprecipitation
Jmjd3	Histone 3 lysine 27 demethylase
KRT17	Keratin 17
lg2FC	Log base 2 fold change
LIMMA	Linear Models for Microarray Data
MCF10a	Mammary epithelial cells
MCF-7	Breast cancer cell line
MDA-MB-231	Breast cancer cell line
MFE	Mammosphere formation efficiency
MLL2	Histone-lysine N-methyltransferase 2D
MRC	Medical Research Council
mRNA	Messenger ribonucleic acid
ncRNA	Non coding ribonucleic acid
OCA	Ovarian Cancer Action
OS	Overall survival
PBS	Phosphate-buffered saline

PFS	Progression free survival
PH1	Pairing homologous 1
PR(>ChiSq)	Pairwise tests for differences in Chi-Square distribution
PRC1	Polycomb Repressive Complex 1
PRC2	Polycomb Repressive Complex 2
PSTI	Type II restriction endonuclease
qRT-PCR	Quantitative real time polymerase chain reaction
RbAp48	Histone-binding protein RBBP4
RFS	Relapse free survival
RHOQ	Ras homolog family member Q
RING1	ring finger protein 1
RNA Pol II	RNA polymerase II
RNAi	RNA interference
RNase	Ribonuclease
RPMI	Roswell Park Memorial Institute medium
SAM	S-Adenosyl methionine
SET	protein domain present in drosophila su(var)3-9 and Enhancer of zeste proteins
SFE	Spheroid formation efficiency
siRNA	small interfering ribonucleic acid
SPINK1	Pancreatic secretory trypsin inhibitor
SUDHL8	Lymphoblast-like B lymphocyte cell line
SUV39H1	Suppressor of variegation 3-9 homolog 1 (Drosophila)
SUV39H2	Suppressor of variegation 3-9 homolog 2 (Drosophila)
SUZ12	SUZ12 polycomb repressive complex 2 subunit
SYBR	Asymmetrical cyanine dye used as a nucleic acid stain in molecular biology
TCGA	The Cancer Genome Atlas
Th1	Type 1 helper T cells
Th2	Type 2 helper T cells
TSS	Transcription start site
TUBB	Tubulin beta chain
UTX	Ubiquitously transcribed tetratricopeptide repeat, X chromosome
WILL1	CD5 and CD10 double-positive mature B-cell line
WSU-FSCLL	Low-grade follicular small cleaved cell lymphoma cell line

Peer reviewed publications and presentations

National Cancer Research Institute Cancer Conference, Liverpool, UK, November 2–5 , 2014-

Dual inhibition of EZH2 and EHMT2 as a targeted therapy in breast cancer- Ian Green, Ed Curry, Elham Shamsaei, Matt Fuchter, Robert Brown

The Stem Cell Niche (Copenhagen Biosciences), Copenhagen, Denmark, May 18-22, 2014-

Dual histone methyltransferase inhibitors activate apoptosis pathways , inhibit cell growth, and reduce cancer stem cell activity in breast and ovarian cancer cells- Ian Green, Ed Curry, Elham Shamsaei, Luke Payne, Gillian Farnie, Matt Fuchter, Robert Brown

Precision Medicines in Breast Cancer, London, United Kingdom, May 09-10, 2013 -

Phenocopying EZH2 knockdown with novel histone methyltransferase inhibitors- Ian Green, Ed Curry, Elham Shamsaei, Robert Brown

Chapter 1: Introduction

1.1 Overview of epigenetics and cancer

1.1.1- Overview

Cancer is a disease of uncontrolled growth, spurred on by genomic instability, epigenomic alterations, and the microenvironment in which the cancerous cells exist (e.g. inflammation). Together these factors conspire to produce a situation where continued growth can occur. The following hallmarks have been identified as key steps in the commencement and continuation of neoplastic disease ¹:

- Sustaining proliferative signalling,
- Evading growth suppressors
- Resisting cell death
- Enabling replicative immortality
- Inducing angiogenesis
- Activating invasion and metastasis
- Reprogramming of energy metabolism
- Evading immune destruction

As well as occurring through genomic instability, these key capabilities of the nascent neoplasm can be conferred or accompanied by alterations to the epigenetic landscape, and from this knowledge an emerging therapeutic field is forming. The link between epigenetics and cancer has long been established (e.g. unusual patterns of DNA methylation were observed in cancer cells relative to non-cancerous tissue ²). Since then, the link between epigenetics and cancer has been extensively explored from a variety of directions.

1.1.2- Epigenetic therapies and pathways in cancer

Epigenetics is normally defined as the study of reversible, heritable changes to gene expression which occur without alteration of the genetic code ³. There are a number of ways in which these alterations can occur, ranging from the methylation of DNA itself to the modification of the lysine tails of the nucleosome forming histone proteins.

These histone proteins are the structures that DNA is wrapped around in the nucleus, a combination known as chromatin, and by modifying the lysine tails of the histone proteins (the lysine tails of histone proteins can be ubiquitinated, acetylated, sumoylated, phosphorylated, or methylated) gene expression can be altered ⁴.

As with genomic instability and alterations, these epigenetic modifications (as well as epigenetic modifiers and related pathways) are commonly shown to be altered within cancer ⁵, and these changes vary in form and magnitude between cancer types. Further exploration of specific cancer epigenomes offers the possibility to stratify cancer types as potentially susceptible to tailored epigenetic intervention- for example, follicular lymphoma has been found to contain recurrent mutations of the histone methyltransferase MLL2 in roughly 90% of cases ⁶, and in as many as 12 distinct cancers the histone demethylase UTX is mutated ⁷.

Already there are several approved drugs used routinely in cancer treatment based upon the premise of targeting the cancer epigenome. These include 5-azacytidine ⁸ and 5-aza-2'-deoxycytidine ⁹, which hypomethylate DNA by chemically inhibiting DNA methyltransferase activity and are used in treatment of myelodysplastic syndromes, and histone deacetylase inhibitors such as Vorinostat ¹⁰ and Romidespin ¹¹ which can be utilised in the treatment of cutaneous T-cell lymphoma.

1.2 The HKMT EZH2

1.2.1- H3K27me3 and HKMTs

A mark which is truly epigenetic in that it is both reversible and heritable is H3K27me3 (it can be inherited somatically during cell division, where EZH2 stably associates with DNA during replication to re-establish the H3K27me3 levels post-replication ¹²). In terms of identifying potential epigenetics targets, targeting this mark could be key in reversing aberrant epigenetic silencing in many cancers ¹³. In cancer, abnormal epigenetic silencing can occur on multiple tumour suppressor genes via mechanisms associated with H3K27me3 and this can occur independently of DNA methylation ¹⁴. The degree of H3K27me3 is largely mediated by the methylation of H3K27 by the PRC2 complex, containing the HKMT EZH2.

H3K27me1 (monomethylation) is associated with active transcription (and targeted according to Setd2-dependent H3K36me3 deposition); H3K27me2 (di-methylation) is associated with inactive transcription, and the protection of enhancer regions from acetylation (and occurs concurrently with a reduction in H3K36me levels); Finally, H3K27me3 is associated with repression of promoter regions, resulting in a reduction in gene expression ¹⁵.

HKMTs catalyse the methylation of lysines at the carboxy-terminus of histones such as H3 and H4 (histone lysine tails), and almost all of the HKMT proteins that have been identified thus far (with the exception of DOT1 HKMT proteins) belong to the SET-domain superfamily ¹⁶. The SET-domain is the catalytic domain within the HKMT that recognises the S-adenosylhomocysteine hydrolase (SAH) methyl donor and the histone substrate, and global inhibitors of HKMT such as DZNep ¹⁷ have been identified- DZNep inhibits the activity of S-adenosylhomocysteine hydrolase, which indirectly inhibits numerous S-adenosylmethionine (SAM) dependent methylation reactions including the methylation leading to the H3K27me3 state ¹⁸.

As an HKMT, the PRC2 complex member EZH2 (which contains SET-domain) catalyses the dimethylation and trimethylation (utilising SAM as a resource) of H3K27¹⁹ and the resulting H3K27me3 leads to chromatin condensation and a reduction in gene expression. This methylation of H3K27 can be reversed by histone lysine demethylases such as JmjD3²⁰.

The H3K27me3 that is caused by EZH2 (as part of the PRC2 complex) is recognised and bound by the PRC1 complex subunit CBX- upon this binding the catalytic RING1 subunit of PRC1 monoubiquitylates H2AK119- this represses gene transcription (Fig.1.1) by preventing RNA Pol II dependent transcriptional elongation²¹.

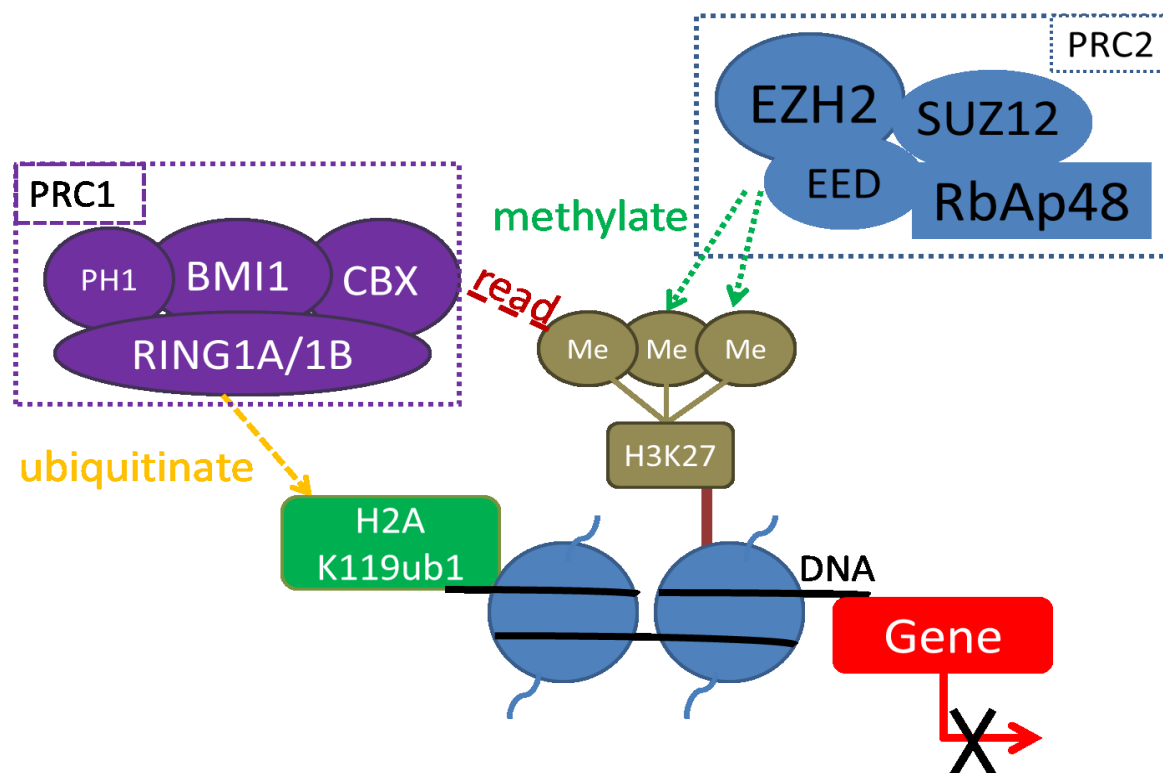


Figure 1.1- Summary diagram of PRC2 mediated PRC1 recruitment leading to gene silencing

The PRC2 complex plays a key role in development, catalysing H3K27me3 and also physically interacting with and recruiting DNA methyltransferases, DNMT1, DNMT3A, and DNMT3B, which methylate CpG points on EZH2 target genes and establish stable repressive chromatin structures²². Functional mutations in the PRC2 complex can lead to a loss of pluripotency in

embryonic stem cells ²³, and PRC2 is required for Hox gene silencing ²⁴- proper regulation of Hox genes is developmentally vital in order to properly allocate segmental identity along different body axes in mammals, and PRC2 in combination with PRC1 regulates Hox gene targets during development ²⁵. EZH2 can also methylate non-histone targets such as transcription factors (e.g. GATA4)²⁶.

The action of EZH2 as a mediator of epigenetic silencing is complex. When this silencing is aberrantly regulated, rather than maintain the delicate balance of gene expression required in development, EZH2 can help drive undesirable phenotypes.

1.3 EZH2 and cancer

1.3.1 EZH2 and cancer

High expression of EZH2 (including some cases of gene amplification) was initially reported in prostate cancer ²⁷ and breast cancer ²⁸. Since then, high levels of EZH2 have been shown to be a marker of aggressive breast cancer ²⁹⁻³¹, and associated with difficult to treat basal or triple negative breast cancer ³². High levels of EZH2 expression are associated with high proliferation rate and aggressive tumour subgroups in cutaneous melanoma and cancers of the endometrium and prostate ³⁰.

High EZH2 expression has now also been linked to bladder cancer ³³, poor prognosis and metastasis ³⁴ as well as cisplatin resistance ³⁵ in ovarian cancer, progression of lung cancer ³⁶ and liver cancer ³⁷, higher stage of brain tumours ³⁸, poor prognosis in renal cancer ³⁹, poor prognosis in gastric cancer ⁴⁰, poor prognosis in oesophageal cancer ⁴¹, proliferation and chemoresistance in pancreatic cancer ⁴², and is linked to invasion in nasopharyngeal carcinoma ⁴³. Linking back to the initial findings mentioned where EZH2 showed high expression in prostate cancer, overexpression of EZH2 has been shown to be a driver for metastasis in animal

models of prostate cancer ⁴⁴. Gene knockdown studies of EZH2 have shown that EZH2 knockdown reduces growth in a variety of these tumour cell types ^{27,45}.

One important role of EZH2 is its involvement in the maintenance of the CSC population, the population of cells that are theorised to drive cancer initiation, progression, metastasis, recurrence and drug resistance ⁴⁶. Reduction of EZH2 levels by siRNA treatment has been shown to lead to the loss of a side-population of CSC like cells that overexpress ABC drug transporters and sustain the growth of drug-resistant cells during chemotherapy in ovarian cancer models ⁴⁷, and EZH2 is essential to maintain CSC populations in glioblastoma ⁴⁸ as well as pancreatic cancer and breast cancer ⁴⁹

EZH2 mediated gene silencing plays a role in numerous cancers- as summarised in Table 1.1, the expression of EZH2 is known to be regulated by various tumour suppressor miRNAs and oncogenic transcription factors and the access by EZH2 of specific DNA sites has been shown to be regulated by numerous DNA binding proteins, transcription factors, and ncRNAs ⁵⁰.

Table 1.1- Summary of known regulators of EZH2 expression and DNA targeting in cancer

Transcription regulators	miRNAs	DNA binding proteins/transcription factors	ncRNAs
Myc	miR-25	YY1	HOTAIR
E2F	miR-26a	Snail	HEIH
EWS-DLI1	miR-30d	Myc	PCAT-1
SOX4	miR-98	SAFB1	H19
NF-Y	miR-101	HIC1	linc-UBC1
ANCCA	miR-124	PER2	
STAT3	miR-137		
ETS	miR-138		
EIK-1	miR-144		
HIF-1	miR-214		
	miR-let7		

As well as general overexpression driven by regulators of EZH2, mutations of EZH2 have been identified in lymphomas. In lymphoma⁵¹ within the catalytic SET-domain of EZH2 a heterozygous missense mutation has been identified by high throughput transcriptome sequencing at amino acid Y641- a variety of heterozygous mutations at Y641 were found in 7% of lymphomas and 22% of diffuse large cell B-cell lymphomas with germinal centre origin. Mutations were not observed elsewhere in EZH2. These Y641 mutations confer an enhanced catalytic efficiency for H3K27me2 and H3K27me3, and as such increase the degree of H3K27me3 mediated silencing⁵².

In an effort to target the PRC2 complex chemically and thus reverse the H3K27me3 mediated silencing related to so many negative clinical outcomes, many groups have developed small molecule inhibitors that target EZH2 (Table 1.2)

These inhibitors all focus on the inhibition of the PRC2 complex, the majority of them sharing the target of the EZH2 SET-domain SAM cofactor binding pocket, mostly following an established chemotype (Fig.1.4). They show a reduction in H3K27me3 levels and a reduction in

growth of Y641 EZH2 mutant cells, but on the whole have been relatively ineffective in EZH2 wild type cells.

Table 1.2: Examples of EZH2 specific inhibitors

	Mode of action	EZH2 K _i (nM)	Reference
DZNep	S- adenosylhomocysteine hydrolase inhibitor	Not applicable	Tan et al. 2007
GSK126	SAM competitive EZH2 inhibitor	0.5–3	McCabe et al. 2012
EPZ005687	SAM competitive EZH2 inhibitor	24	Knutson et al. 2012
EPZ-6438	SAM competitive EZH2 inhibitor	2.5	Knutson et al. 2013
EI1	SAM competitive EZH2 inhibitor	13	Qi et al. 2012
GSK926	SAM competitive EZH2 inhibitor	7.9	Verma et al. 2012
GSK343	SAM competitive EZH2 inhibitor	1.2	Verma et al. 2012

1.3.2 EZH2 and EHMT2

EHMT2 (also known as G9a) and the highly homologous EHMT1 (also known as GLP) are HKMTs responsible for H3K9me1, H3k9me2, and H3K9me3 in heterochromatin- H3k9me1/2/3 are transcriptionally repressive chromatin marks that are typically found on the promoter regions of silenced genes, and this silencing occurs frequently in cancer ⁵⁹. EHMT2 is amplified and highly expressed in a number of cancers including prostate carcinoma, lung cancer, and leukaemia, and the growth of these tumours can be reduced by gene knockdown of EHMT2 ^{60,61}.

EHMT2 is, like EZH2, a member of the SET-domain superfamily and the two both have a catalytic SET-domain responsible for the methylation of their respectively targeted lysine residues ¹⁶.

As well as methylating H3K9, EHMT2 has also shown the capacity to methylate H3K27^{62,63}. It has been theorised that this may provide cells a method to compensate for loss of EZH2⁶⁴. Recently it has become clear that EHMT2 actually physically interacts with the PRC2 complex, and shares targets with EZH2 for epigenetic silencing⁶⁵, which leads to a proposed model of function (**Fig.1.2**) where inhibition of EZH2 alone may not be sufficient to wholly reverse aberrant H3K27me3 mediated epigenetic silencing.

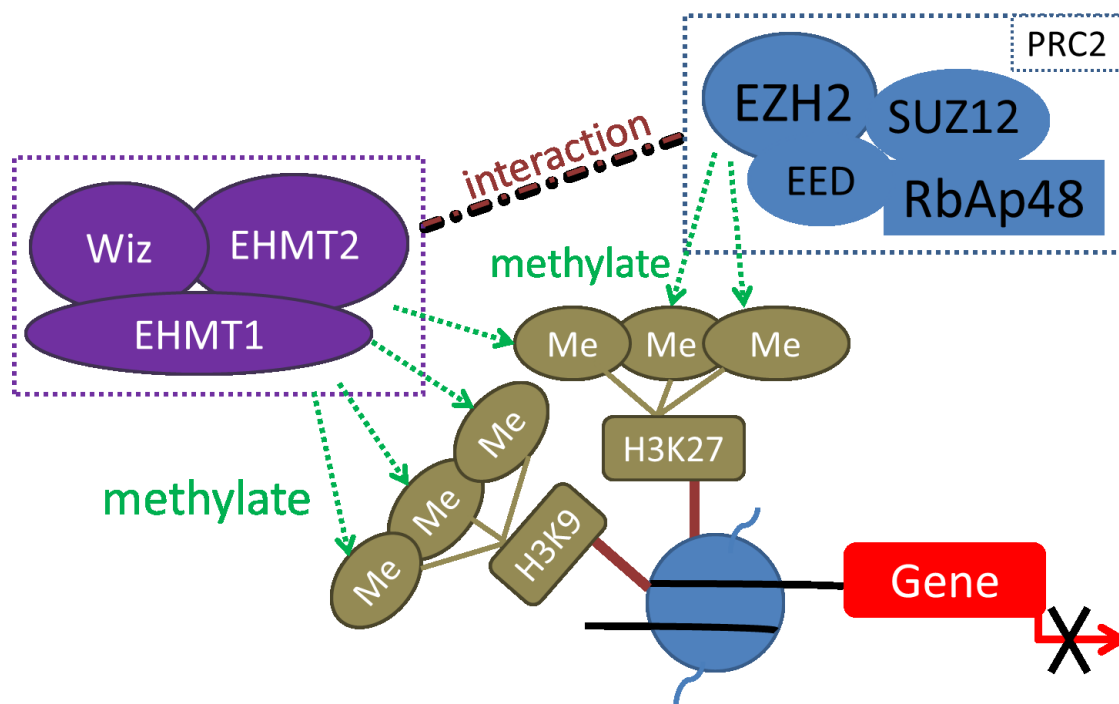


Figure 1.2- Summary diagram of theorised interaction between PRC2 and the EHMT2/EHMT1 complex

1.4 Cancer stem cells and EZH2

Another area under investigation will be the impact of HKMT inhibition on CSC activity. CSCs (known as cancer stem cells, cancer stem-like cells, tumour initiating cells, or tumour propagating cells) are a sub-population of cells exhibiting stem-like characteristics- that is to say within a cancer type, the CSC should be capable of either maintaining its undifferentiated state or giving rise to any cell type found within that cancer. They are normally characterised as cancer cells capable of long-term clonal repopulation with long-term self-renewal capacity⁶⁶. First identified in leukaemia⁶⁷, CSCs have subsequently been identified in numerous cancers (examples shown in Table 1.3).

Table 1.3: Example cancer types with identified CSC populations

Cancer type	Reference
Acute Myeloid Leukaemia	Bonnet & Dick 1997
Brain	Singh et al. 2003; Ignatova et al. 2002
Breast	Al-Hajj et al. 2003
Ovarian	Zhang et al. 2008
Colorectal	Ricci-Vitiani et al. 2007
Skin squamous cell	Malanchi et al. 2008
Head & neck	Prince et al. 2007
Lung	Eramo et al. 2008
Pancreatic	Li et al. 2007
Melanoma	Schatton et al. 2008
Prostate	Collins et al. 2005

In the CSC model of tumour growth⁷⁹, the tumour is a hierarchically organised structure within which the CSC population sustains tumour growth (Fig 1.3). This CSC model of growth does not maintain the tumour as homogenous entity- somatic mutations can occur within the CSC population, which will lead to clonal diversity and increased tumour heterogeneity. However, even within genetically identical cell populations this epigenetically different CSC population can be observed. There is as yet no universal identifying cell surface marker for CSCs- CD133 has been associated with CSCs in many different types of tumours⁸⁰, but is not always

applicable, and currently the only way of consistently identifying a CSC is through the capacity for long-term clonal repopulation and long-term self-renewal capacity.

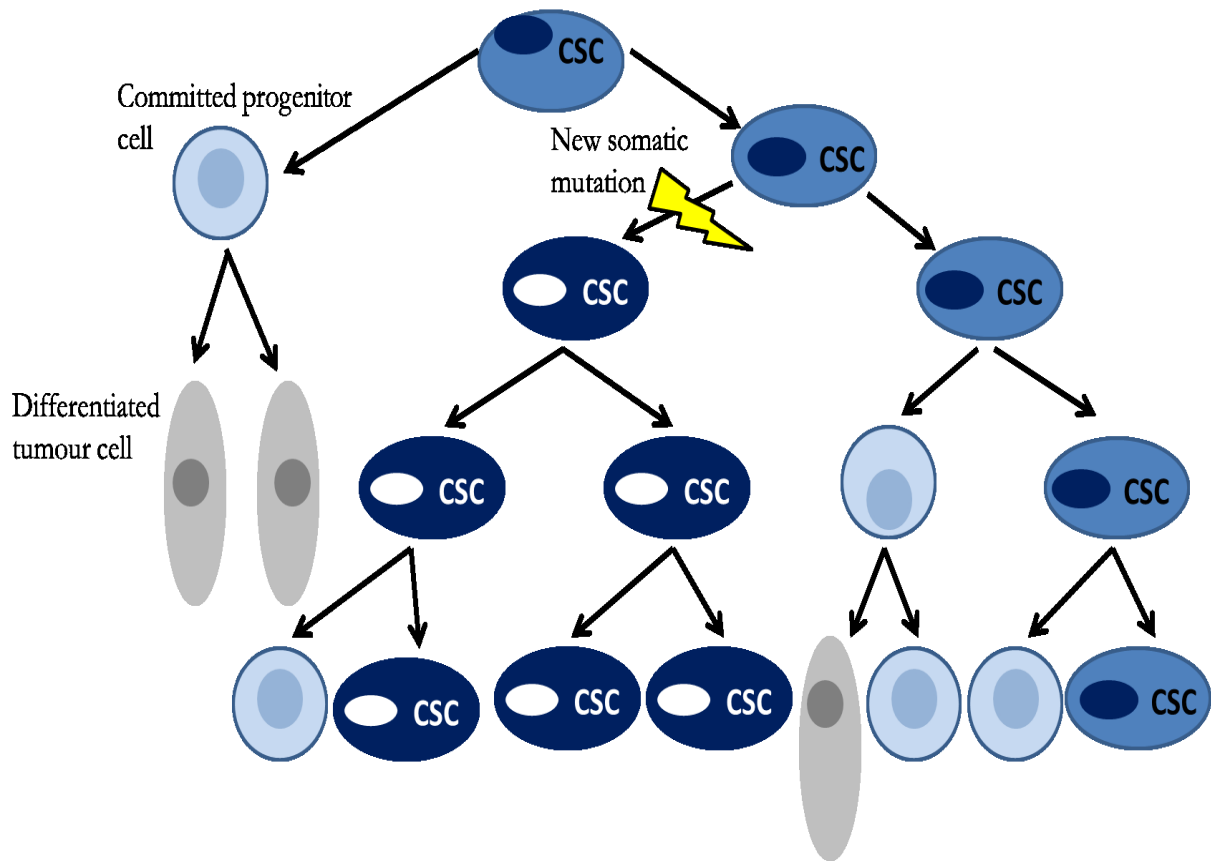


Figure 1.3: In the proposed CSC model of tumour growth, only a subset of tumour cells have the ability for long-term self-renewal and these cells give rise to progenitors with limited proliferative potential that will eventually terminally differentiate (modified from ⁷⁹)- it is noteworthy that in this model somatic mutations can occur, which will lead leading to clonal diversity which may increases tumour heterogeneity

This CSC population is of therapeutic note for two main reasons- firstly, it is theorised to sustain the growth of the tumour ⁷⁹. Secondly, the CSC population is characteristically resistant to chemotherapy (relative to the bulk of the cancer) ^{81,82} – as many chemotherapies traditionally target rapidly dividing cell populations, relatively quiescent CSC populations may not be killed by the applied doses. In addition, some CSC populations have been shown to highly express ABC transporters that can cause drug efflux. This resistance is theorised to lead to a sustained

survival of CSCs throughout therapy, which then potentially lead to tumour recurrence post chemotherapy.

Examples where CSC levels are enriched following chemotherapy or radiotherapy include colorectal cancer⁸³, brain cancer⁸⁴, breast cancer⁸⁵, and ovarian cancer^{86,87} - this indicates the CSC population is not being targeted effectively by conventional therapy, which may be leading to relapse. Indeed, by combining traditional therapy with CSC targeting therapy, it was recently shown that this combination could lead to drastically lowered growth of glioblastoma in *in vivo* models⁸⁸.

From the perspective of putative dual HKMT inhibition, CSCs may be susceptible to the targeting of EZH2 mediated silencing. As part of the PRC2 complex, EZH2 is known to be required for the maintenance of embryonic stem cells⁸⁹. Induction of EZH2 expression in haematopoietic stem cells can lead to the accumulation and induction of the epigenetic changes required for these stem cells to progress the development of leukaemia⁹⁰. In CSCs, increasing levels of EZH2 expression can lead to the expansion of CSC population in breast cancer⁹¹, and increased expression of EZH2 also leads to the maintenance of a stem cell like phenotype in some cancers⁹².

This indicates that EZH2 inhibition may target the CSC population, and indeed this has been shown in several cases- in ovarian cancer models, siRNA mediated reduction in EZH2 expression leads to a loss of a CSC population that has been characterised as overexpressing ABC drug transporters and sustaining chemotherapy resistant growth⁴⁷, and reduction in EZH2 levels also leads to a reduction of CSC population in glioblastoma⁴⁸ and a reduction in CSC population in prostate cancer⁹³.

EZH2 clearly plays a role in the maintenance of the CSC populations in numerous cancer types- as such dual inhibitors of EZH2 and EHMT2 which should strongly reverse EZH2 mediated epigenetic silencing may have a powerful impact on CSC activity.

1.5 Identification of novel dual HKMT

Based upon the premise of EHMT2 acting to support EZH2 activity, and as EHMT2 and EZH2 both contain catalytic SET-domains, it is theorised that by chemically targeting the SET-domain it may be possible to inhibit both EZH2 and EHMT2 with one compound.

The SET-domain is comprised of two binding pockets- one for the protein substrate, the other for the cofactor SAM. Occupancy of either of these binding pockets with a small molecule inhibitor is an established effective strategy to block HKMT mediated methylation⁹⁴. High throughput screening identified the first substrate competitive inhibitor of EHMT2, BIX-01294⁹⁵, and since then a number of derivatives and analogues have been developed including UNC0638⁹⁶.

As mentioned, EHMT2 has also shown the capacity to methylate H3K27^{62,63} -as BIX-01294⁹⁵ was shown to bind to the substrate binding pocket within the SET domain and it is known that protein recognition motifs for histone binding at repressive sites are similar⁹⁷, it was deemed possible that there are common aspects to the histone binding pockets of the repressive HKMTs EZH2 and EHMT2.

In collaboration with Jim Snyder (Department of Chemistry, Emory University, Atlanta) and Masoud Vedadi (Structural Genomics Consortium, University of Toronto), **Matt Fuchter**, **Fanny Cherblanc**, and **Nitipol Srimongkolpithak** (Department of Chemistry, Imperial College London) derived a compound library from the quinazoline template of BIX-01294 in an attempt to discover of dual (substrate competitive) inhibitors⁹⁸.

This library was screened by **Elham Shamsaei** utilising a QRT-PCR screen (**Materials and methods: QRT-PCR**), with EZH2 inhibitory capacity measured by re-expression of the KRT17 and FBXO32 genes, which are known to be silenced in an EZH2 dependent manner⁵³. Within the library three compounds were identified (Fig.1.4) as up-regulating the expression levels of KRT17 and FBXO32: HKMT-I-005, HKMT-I-011, and HKMT-I-022 (**APPENDIX 1**- Manuscript of Curry et al '*Dual EZH2 and EHMT2 histone methyltransferase inhibition increases biological efficacy in breast cancer cells*').

HKMT-I-005, HKMT-I-011, and HKMT-I-022 up-regulated expression of KRT17 and FBXO32 (Supplementary Table 8.1) Known EHMT2 inhibitors BIX-01294 and UNC0638 did not up-regulate KRT17, but did up-regulate FBXO32, though FBXO32 has previously been shown to be regulated via multiple mechanisms⁹⁹.

The specific EZH2 inhibitor GSK343 had no effect on all the target genes studied when examined up to 72 hours following treatment and at concentrations up to 10 μ M. Representative compounds that failed this qRT-PCR screen are included were included for reference (Supplementary Table 8.1).

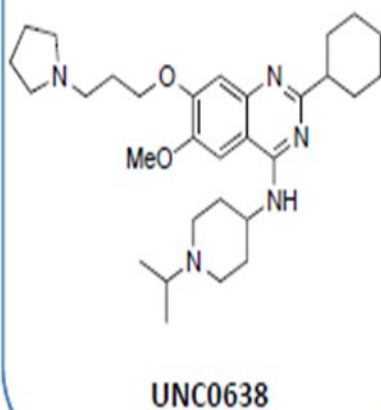
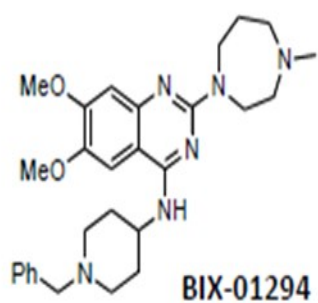
Using a scintillation proximity assay (SPA) which monitors the transfer of a tritium-labelled methyl group from SAM to a biotinylated-H3 (1-25) peptide substrate, mediated by EHMT2, the EHMT2 IC₅₀ of HKMTI-1-005, HKMTI-1-011 and HKMTI-1-022 was found to be 0.10 μ M, 3.19 μ M, and 0.47 μ M respectively (Srimongkolpithak et al. 2014).

Matt Fuchter, Fanny Cherblanc, and Nitipol Srimongkolpithak carried out a PRC2 enzymatic assay monitoring transfer of biotinylated-H3 (21-44) peptide substrate groups from the cofactor SAM to assess biochemical inhibitory activity of the hits against EZH2 (comparable to the assay performed for EHMT2⁹⁸), and found compounds HKMTI-1-005, HKMTI-1-011 and HKMTI-1-022 to have PRC2 IC₅₀ values of 24 μ M, 12 μ M and 16 μ M

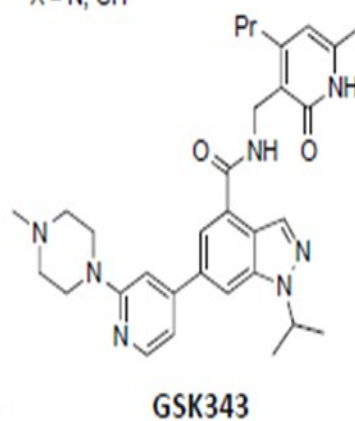
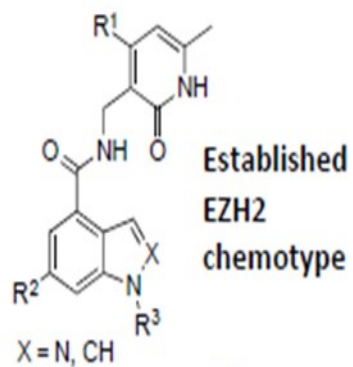
respectively (Supplementary Figure 8.1). A methyltransferase selectivity assay was also performed comparing the binding capacity of the inhibitors to different HKMT (Supplementary Figure 8.2) - this data indicates EHMT2/1 and EZH2 were inhibited significantly up to a dose of 100 μ M by HKMT-I-005.

Compound batch data is shown in **Materials & Methods: Compound Batch data**. Each batch was tested using the aforementioned MDA-MB-231 proliferation assay and QRT-PCR screen, and only used if results were comparable between batches.

Known EHMT2 inhibitors



Known EZH2 inhibitors



Hits

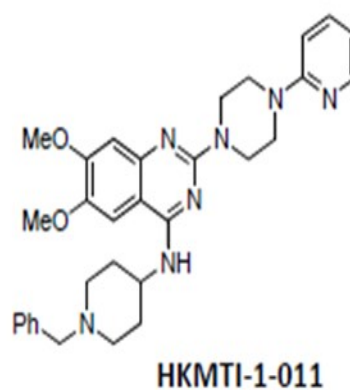
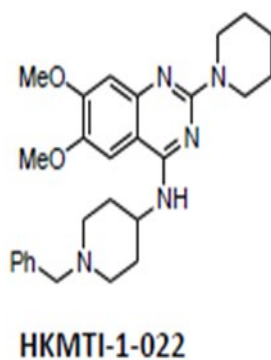
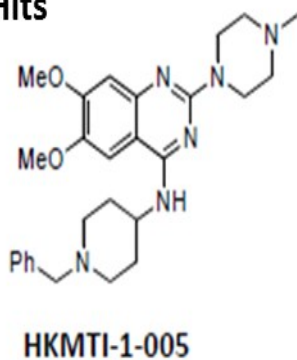


Figure 1.4- Chemical structure of Histone Lysine Methyltransferase inhibitors

Hypothesis

Aberrant EZH2 mediated epigenetic silencing has been observed in multiple cancer types and is linked to negative clinical outcomes and aggressive phenotypes. It appears that this silencing is supported by the HKMT EHMT2. We hypothesise that by targeting both EZH2 and EHMT2 a greater reversal of EZH2 mediated epigenetic silencing will occur relative to targeting EZH2 or EHMT2 individually, and that this dual HKMT inhibition will have a stronger impact on HKMT mediated cancer cell phenotypes than individual HKMT inhibition.

Aims

- Utilise publicly available data to examine the degree to which EZH2/EHMT2 expression, CNV, and mutation status vary between cancer types and within cancer subtypes and patients to establish if stratification by EZH2/EHMT2 expression, CNV or mutations at a patient and disease level is viable
- Characterise the impact of novel dual HKMT inhibitors on gene expression levels in cancer cell models, and examine how this relates to the chromatin state of target genes with regards to silencing marks H3K27me3 and H3k9me3
- Examine the effect of dual HKMT inhibition on cancer cell phenotypes linked to HKMT expression (e.g. cancer stem cell activity, cancer cell proliferation, sensitivity to chemotherapeutic treatment)

Chapter 2: Materials and methods

Cell culture

The breast cancer cell line MDA-MB231¹⁰⁰ and ovarian cancer cell line IGROV1¹⁰¹ were maintained in DMEM (Sigma) or RPMI (Sigma) respectively, containing 10% FBS (First Link, UK), 2mM L-Glutamine (Gibco), 100U/ml Penicillin and 100µg/ml Streptomycin (Gibco) at 37°C in a humidified incubator, under 5% CO₂. Cells were tested for Mycoplasma contamination regularly by using sensitive bioluminescence based MycoAlert Mycoplasma detection kit (Lonza).

RNA preparation

RNA extraction comprises of five main stages: cell lysis and dissolution, removal of proteins, denaturation and inactivation of RNases, removal of cellular components, and precipitation of RNA. TRIzol (Invitrogen) was used to extract the RNA (based on acid guanidinium thiocyanate-phenol-chloroform extraction¹⁰²).

MDA-MB-231 cells were plated on 6 well plates and at 90% confluence were treated with hit compounds. Each well contains ~200,000 cells at this percentage of confluence. Culture media was removed, the cells were washed in PBS, and TRIzol added (1ml per well) to lyse the cells. Cells were manually pipetted up and down to ensure homogenisation of the sample, and then left for 5 minutes at room temperature to permit dissociation of the nucleoprotein complexes.

TRIzol inhibits RNase activity whilst simultaneously dissolving cell components and disrupting cells- having incubated in TRIzol for 5 minutes, 200µl of chloroform was added to each 1ml reaction, capped, and vigorously mixed. After a 3 minute room temperature incubation, this mixture was centrifuged at 12,000 x g at 4°C for 15 minutes. At this point the mixture has separated into three phases- a lower phenol chloroform phase that is red and contains the proteins and cellular components, an interphase containing DNA, and an upper aqueous phase that is colourless and contains the RNA. This upper aqueous phase is carefully removed, and from this RNA can be precipitated, washed, and eluted.

The RNA is precipitated from this aqueous solution with 100% isopropanol- 0.5ml 100% isopropanol was added to each aqueous phase isolated in the previous step, incubated for 10 minutes at room temperature, and then centrifuged at 12,000 x g at 4°C for 10 minutes. The resulting RNA pellet is then washed.

1ml of 75% ethanol was added to each pellet, and vortexed briefly to resuspend the pellet. This sample is then centrifuged at 7500 × g for 5 minutes at 4°C and the supernatant discarded. The RNA pellet was then air dried for 10 minutes at room temperature prior to resuspension in RNase free water.

Having isolated and prepared these RNA samples, quality control was performed to ensure the isolation was successful and there was no carry-over of phenols or contaminants.

A Nanodrop-2000 spectrophotometer (Thermoscientific) was used to produce absorption spectra for each sample, and to calculate the ratio of absorbance at 260/280nm and 260/230nm. RNA absorbs at 260nm, whilst many contaminants like phenol and proteins absorb strongly near 280nm. A 260/280nm absorption ratio of ≥ 1.8 was deemed satisfactory to indicate that these samples contained a high purity of RNA. Phenol can also absorb at 230nm, and so the 260/230nm absorbance ratio was also calculated and deemed to contain negligible

contamination at values ≥ 1.9 . The absorbance spectra trace of each sample was compared to reference traces from ThermoScientific and if the pattern was not as expected for pure RNA the sample was deemed unfit for use.

The samples which underwent microarray analysis went through a second quality control check at Oxford Gene Technologies prior to use, consisting of analysis using an Agilent Bioanalyzer (Agilent Technologies). This system uses electrophoretic separation on micro-fabricated chips, RNA samples are separated and then detected via laser induced fluorescence detection to compile an electrophelogram of the RNA, and calculate an RNA Integrity Number (RIN). This system assigns a value of 1-10 to the sample, with 1 being wholly degraded RNA and 10 being completely intact RNA. For microarray analysis a RIN of ≥ 7 is recommended, and any samples falling significantly below this were dropped from the microarray study (as per recommendations of Oxford Gene Technology).

QRT-PCR

Reverse transcription of RNA (isolated as described in Materials and Methods: RNA preparation) was completed using the SuperScript III First-Strand Synthesis System (Invitrogen) according to the manufactures instructions, using 7 μ l of purified RNA as starting material. For qRT-PCR measurements the 2x iQ SYBR Green Supermix (Bio-Rad), 200nM Primers and 0.4 μ l of cDNA /per 20 μ l reaction was used. The measurement was done in low-white 96-well plates (Bio-Rad) on a CFX96 Real-time System/C1000 Thermal Cycler (Bio-Rad) with the following protocol: 95°C for 3'; 95°C for 10'', 56°C for 10'', 72°C for 30'' 42 cycles followed by a melting curve from 72°C to 95°C in order to control for primer dimer or unwanted products. Each measurement was done in triplicate, and the list of primers can be found in Table 2.1. For normalisation we have used GAPDH and RNA pol II. In order to

account for preparation/handling differences during drug treatment and qRT-PCR measurement, we are using a second GAPDH (GAPDH_2) primer pair, and would count an experiment as valid if the difference between these primer pairs is not greater than ± 0.15 fold for each run. Experiments were also done with the 'Fast Sybr Green Cell-to-CTTM-Kit' according to the manufacturer's instructions (Applied Biosystem). 15,000 cells per 96 well were plated and after 24h treated with compounds at various concentrations. Conditions were used as described above for qRT-PCR measurement.

Table 2.1: QRT-PCR primers used

Name of gene	Forward primer	Reverse primer	Product	Pubmed REF
GAPDH_1	CCTGTTCGACAGTCAG CCG	CGACCAAATCCGTT GACTCC	101bp	12615716
GAPDH_2	CCCCTTCATTGACCTC AACTACAT	CGCTCCTGGAAGAT GGTGA	135bp	PMC2517635
KRT17	CAACACTGAGCTGGA GGTGA	GGTGGCTGTGAGG ATCTTGT	124bp	
FBXO32	TGTTGCAGCCAAGAAG AGAA	CAATATCCATGGCG CTCTTT	120bp	Primer 3
JMJD3	CCTCGAAATCCCATCA CAGT	GTGCCTGTCAGATC CCAGTT		
EZH2	AGTGTGACCCTGACCT CTGT	AGATGGTGCCAGC AATAGAT	122bp	RTPrimerDB probe ID: 4521
RUNX3	CAGAAGCTGGAGGAC CAGAC	TCGGAGAATGGGT TCAGTTC		
RUNX3	TTCCTAACTGTTGGCT TTCC	TAGGTGCTTTCCTG GGTTTA	95bp	RTPrimerDB probe ID: 4757
SPINK1	GGTAAGTGCGGTGCA GTTTT	TAGACTCAACAGG GCCAAGG	101bp	

Compound batch data

The following batches of the hit compounds HKMT-I-005, HKMT-I-011, and HKMT-I-022 were used in this study. Each batch was tested as per the compound screening procedure (described in Chapter 1) and any batch that did not replicate the previous findings was disregarded:

HKMTI-1-005: TG3-178-2 (synthesized 30/09/2008); NS-011 (synthesized 7/5/2011); NS-080 (synthesized 23/4/12); JC-087 (HCl salt formulation, synthesized 1/10/2012); NS-382 (synthesized 22/08/14)

HKMTI-1-011: TG3-214-1 (synthesized 13/11/2008); NS-014 (synthesized 20/6/11); NS-081 (synthesized 26/4/12)

HKMTI-1-022: TG3-179-1 (synthesized 20/10/2008); NS-015 (synthesized 28/6/11); NS-082 (synthesized 26/4/12)

Calculation of differential expression (Harvard Centre for Computational & Integrative biology)

Gene expression of target genes in normal human tissues was assessed utilising the online platform supplied by the Harvard Centre for Computational & Integrative biology¹⁰³. This platform contains 126 normal primary human tissues (represented by 557 different microarrays) obtained from Affymetrix U133A chips. Raw CEL files were obtained and normalized as a single experiment (microarray normalization was performed using the GCRMA module and present/absent calls were calculated using Affymetrix MAS5 package in Bioconductor. For the purpose of computing the enrichment scores, only probes with at least 1 present call across the entire dataset for which the expression value was above $\log_2(100)$ were retained). This

provided information on the differential expression of target genes. Differential expression having been calculated using the LIMMA module of Bioconductor ¹⁰⁴, The heatmap depicts linear coefficients derived from a pairwise comparison of expression values - Red denotes relatively high expression and Green denotes relatively low expression compared to all of the other tissues in the heatmap figures 3.1/2. In cases where multiple probes for the same genes exist, only the higher scoring probe is utilised.

Gene expression of target genes in cancerous human tissues was assessed utilising the online platform supplied by the Harvard Centre for Computational & Integrative biology ¹⁰³. This platform contains 16 cancerous human tissues (represented by 92 different microarrays) using Affymetrix U133A chips. Raw CEL files were obtained and normalized as a single experiment, providing information on the differential expression of target genes. In cases where multiple probes for the same genes exist, only the higher scoring probe is utilised.

Differential expression was calculated using the LIMMA module of Bioconductor ¹⁰⁴, with Red denoting high expression and Green denoting low expression compared to all of the other tissues.

Table 2.2- Probe IDs and gene names of target analysed in normal human tissues

Gene Name	Probe ID
RHOQ	212119_at
SPINK1	206239_s_at
KRT17	205157_s_at
JMJD3	41387_r_at
EHMT2	202326_at
EZH2	203358_s_at
SUZ12	212287_at
EED	209572_s_at
RBBP4	210371_s_at

Correlation analysis

Normalised expression data was retrieved for target probes (Table 2.2) from the platform described in **Material and Methods: Calculation of differential expression (Harvard Centre for Computational & Integrative biology)**. Pearson correlation coefficients and their statistical significance estimates were calculated using GraphPad Prism (Version 5.00 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com).

CancerMA Forest Plots

The CancerMA integrated bioinformatic analytical pipeline ¹⁰⁵ was utilised to investigate the relative expression of target genes in different cancers. The plots shown comprise of log 2-fold change values for individual studies as well as the total values for all cancer types in the study combined. Each study is illustrated by a diamond and the position on the x-axis represents the measure estimate (lg2FC ratio) - the size of the diamond is proportional to the weight of the study, and the horizontal line through the diamond is the confidence interval of the estimated expression within each study.

Mutation rate, CNV, and expression of target genes in TCGA data

Cbio portal for cancer genomics allows the interrogation of over 69 cancer genomics studies including 17584 samples using whole genome or whole exome sequencing^{106,107}. The number of reported non-synonymous mutations was derived. This portal allows the visualisation of mutational information as assessed by high throughput next generation sequencing.

CNV profiles estimated by the GISTIC algorithm¹⁰⁸ were available through the CBio portal for cancer genomics, along with mRNA expression calculated with reference to a normal adjacent tissue.

Copy number and mRNA expression in 570 ovarian serous cystadenocarcinoma cases was visualised using the CBio portal for cancer genomics, showing mRNA z-Scores (Agilent microarray) compared to the expression distribution of each gene in tumours that are diploid for this gene. Putative copy-number calls on 570 cases determined using GISTIC 2.0.

The results shown here are in whole or part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>.

Comparison of gene expression, clinical data, and CNV in TCGA data

Raw expression data, clinical data, and copy number data were accessed from TCGA data portal for ovarian, breast, colon, glioblastoma multiforme, kidney renal clear cell, kidney renal papillary cell, low grade glioma, lung, rectal, and uterine corpus endometrioid cancers. The results shown are in whole or part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>. Pearson correlation coefficients and estimates of their statistical significance were calculated using GraphPad Prism (Version 5.00 for Windows, GraphPad Software, San Diego California USA, www.graphpad.com).

Cox proportional hazard modelling

Cox proportional hazard modelling¹⁰⁹ was performed on TCGA data to evaluate associations between patient survival times and target gene expression. Model fits were obtained and evaluated using the coxph function provided in the ‘survival’ package of the statistical programming environment R¹¹⁰. The results shown are in whole or part based upon data generated by the TCGA Research Network: <http://cancergenome.nih.gov/>. Probe IDs summarised in Supplementary table 8.4.

Survival analysis utilising combined data sources

The relationship between target gene expression and overall survival/relapse free survival in breast and ovarian cancer were estimated utilising the online portal KMplotter¹¹¹. The expression of target genes was related to overall survival (OS) and relapse free survival (RFS) in breast cancer patients and overall survival and progression free survival (PFS) in ovarian cancer patients.

This resource sources Affymetrix expression data from The Cancer Genome Atlas, the Genome Expression Omnibus, and The European Genome-phenome Archive. Patient samples are split into two groups according to quartile expressions of the proposed biomarker. For each array all percentiles between lower and upper quartiles are computed and the best performing expression threshold is used as a cut-off. The two patient cohorts are compared by a Kaplan-Meier survival plot with a hazard ratio with 95% confidence intervals and logrank P value are calculated. This system allowed the interrogation of gene expression compared to survival data of 4142 breast cancer patients and 1464 ovarian cancer patients, and allowed for sub-division by clinical data such as oestrogen or progesterone receptor status or grade.

Gene expression microarray

Agilent 80k two-colour microarrays were used to profile gene expression changes induced by treatment with drug compounds in MDA-MB-231 cells, both at 24 hours and 48 hours. In the initial microarray experiment 3 replicates were used for each drug/time combination and in the validation study 4 replicates were used. A separate untreated control sample was used for comparison with each replicate. Sample labelling, array hybridization and scanning were performed by Oxford Gene Technologies, according to manufacturer's instructions. Feature Extracted files were imported into GeneSpring (Agilent) and data was normalised to produce log₂ ratios of treated/untreated for each replicate of each drug, time combination. For both arrays RNA was extracted after treatment using TRIzol® (Life Technologies) and quantified using NanoDrop 3300 (ThermoScientific).

Differential expression caused by drug treatment was statistically ascertained using normalised log₂ pre- vs post-treatment gene expression ratios, analysed using LIMMA¹¹² to obtain empirical Bayes moderated t-statistics reflecting statistical significance of differential expression across the replicates for each drug, time combination. Multiple testing adjustment was made using the Benjamini-Hochberg method, following which a threshold of $p < 0.1$ was used to denote significant differential expression in the initial microarray experiment and a threshold of $p < 0.05$ was used in the validation experiment.

Enrichment analysis

Enrichment analysis- a list of EZH2 silenced and activated targets in the MDA-MB-231 cell line was obtained from a previous study¹¹³ and a list of EZH2 silenced targets in the MCF-7 cell line was also obtained¹¹⁴.

Statistical significance of the observed shift towards induced transcriptional upregulation or downregulation of these EZH2 was evaluated using the Wilcoxon rank-sum test implemented in the ‘GeneSetTest’ method from the Bioconductor package LIMMA. What the enrichment analysis of the microarray data shows is if a randomly-chosen Target- gene is more likely to be differentially expressed more than any randomly-chosen non-target gene following treatment.

A meta-analysis was performed by MRes student **Emma Bell** to identify consensus target genes based on 18 independent EZH2 siRNA studies. Raw data for 18 microarray experiments profiling RNA from EZH2 RNAi treated cells were downloaded from Gene Expression Omnibus ¹¹⁵. These datasets were processed individually to minimise cross-array platform bias. A list of the study accession numbers is provided in **Supplementary table 8.7**. For each study, a linear regression model relating probe intensity values to the presence or absence of EZH2-targeting RNAi was fit using the R package LIMMA ¹⁰⁴. This generated empirical-Bayes moderated t-statistics for the EZH2 RNAi induced differential expression. To reconcile cross-platform probe IDs, HGNC gene symbols were used to identify genes. For genes with multiple probes on an array platform, the most statistically significant differentially expressed probe was used and all others discounted. Three meta-analysis approaches were taken to find genes with consistent upregulation following knock-down of EZH2: Fisher’s method of combining P-values, the Rank Product method ¹¹⁶, and a Random Effects Model ¹¹⁷. The top 300 consistently EZH2 silenced and EZH2 activated genes were used for meta-analysis (gene list in Supplementary Table 8.9).

Correlation of gene expression after compound treatment

The similarities between the inhibitors were explored using the 'heatmap' function in R to visualise the pair-wise Pearson correlation coefficients relating the genome-wide transcriptional effects of each treatment.

Having established what genes were differentially expressed (**Materials and Methods: Gene expression microarray**) the column-wise dendrogram shown is the result of complete-linkage hierarchical clustering based on the pairwise Euclidean distances of the treatment-wise vectors of correlation coefficients. For this unsupervised hierarchical clustering, a correlation-based distance metric was calculated for each pair of samples, defined as 1 minus the Pearson correlation coefficient between the vectors of expression values from each sample. Hierarchical clustering was performed using the 'hclust' function provided in R, using complete linkage.

ConsensusPathDB pathway enrichment analysis

Having established what genes were differentially expressed (**Materials and Methods: Gene expression microarray**), enrichment analysis- pathway analysis data was explored utilising the ConsensusPathDB database¹¹⁸. By mapping each probe to a pathway, statistical significance of the observed shift towards induced transcriptional upregulation or downregulation of these pathways could be evaluated using the Wilcoxon rank-sum test implemented in the 'GeneSetTest' method from the Bioconductor package LIMMA.

SiRNA knockdown experiments

SiRNA experiments were carried out on the MDA-MB-231 cell line using Qiagen reagents, according to the manufacturer's instructions. In brief, cells were seeded at a density of 1×10^5 cells/6 cm well and siRNA treated for 48h. HiPerfect, Optimem and 50nM of G9a (SI00091189 HS_BAT8 1, SI03083241 HS_EHMT2), SUV39H1 (SI02665019 HS_SUV39H1 6, SI00048685 HS_SUV39H1 4) and EZH2 (SI00063973 HS_EZH2 4, SI02665166 HS_EZH2 7) siRNA were used for transfections according to the manufacturer's instructions (siRNA sequences given in Supplementary table 8.10). The transfection mixture was added drop-wise onto 30% confluent cells and incubated for 48h after which RNA was extracted as described above.

Chromatin immunoprecipitation

In summary, the ChIP-PCR assay was based upon the 'fast-CHIP' protocol¹¹⁹ - in addition, additional purification with QIAquick PCR Purification Kit (Qiagen) was performed after the Chelex-100 stage of this protocol.

The cells used in this experiment were MDA-MB-231 cells grown to 90% confluency ($\sim 1 \times 10^6$ cells were used per immunoprecipitation). These cells were treated with the HKMT inhibitors (as described below), and chromatin immunoprecipitation was performed. The overview of this technique is as follows:

- protein-DNA complexes are fixed by cross-linking with formaldehyde
- chromatin is sheared, by sonication to into DNA fragment sizes of ~ 200 – $1,000$ base pairs
- Complexes containing the factor of interest are immunoprecipitated using an antibody specific to that protein

- DNA is purified from the isolated chromatin, and specific genomic regions are detected using PCR

To cross-link the protein-DNA complexes, 40 μ l of 37% (wt/vol) formaldehyde was added per 1 ml of cell culture medium to obtain a final concentration of 1.42%. The cells were incubated for 15 min at room temperature. The formaldehyde reaction was quenched after this point by the addition of 125mM glycine at room temperature for 5 minutes (for every 1ml of culture medium, 141 μ l of 1 M glycine was added). The cells were then scraped and collected by centrifugation (2000 x g for 5 min at 4 °C) and then washed with cold PBS twice.

Having cross-linked the protein-DNA complexes and harvested the cells, lysis was performed using IP buffer (150 mM NaCl, 50 mM Tris-HCl (pH 7.5), 5 mM EDTA, NP-40 (0.5% vol/vol), Triton X-100 (1.0% vol/vol). For 500 ml, add 4.383 g NaCl, 25 ml of 100 mM EDTA (pH 8.0), 25 ml of 1 M Tris-HCl (pH 7.5), 25 ml of 10% (vol/vol) NP-40 and 50 ml of 10% (vol/vol) Triton X-100).

This buffer was added (1ml per 10cm dish, containing protease inhibitor cocktail mix) to lyse the cells- the cell pellet was agitated by pipetting up and down repeatedly. This mixture was then centrifuged (12000 x g for 1 min at 4 °C) and the supernatant discarded. The nuclear pellet was washed in IP buffer with protease inhibitor cocktail, and then sonicated to shear the chromatin into DNA fragments of 200-1000bp in size using a Bioruptor Standard (Diagenode). Sonication conditions were previously tested using a variety of cell numbers and timings (Fig.2.1) and based upon this data chromatin was sonicated 3x 5 minutes at 4 °C (20 second pulses at high power sonication with 20 seconds rest for 5 minutes, ice water refreshed, repeated three times).

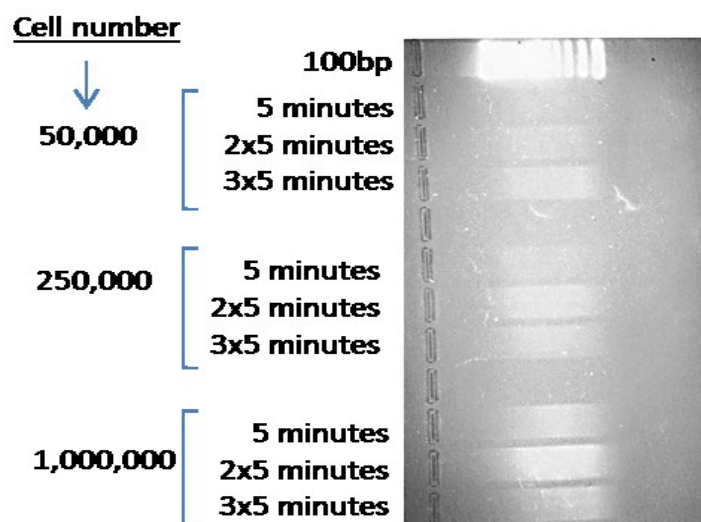


Figure 2.1- Sonication test using chromatin from IGROV1 cells (performed by Daniel Lieber) run on 1% (wt/vol) agarose gel (stained with EtBr) compared to a 100bp DNA ladder. Chromatin from 50000, 250000 or 1 million cells was sonicated for 5 minutes, 2x 5 minutes, or 3x 5 minutes

Lysate was cleared by centrifugation (12000 x g for 10 minutes at 4 °C). This sheared chromatin was now immunoprecipitated (one equivalent cell number was treated identically but with no antibody added to act as a control). Antibody was added to samples and incubated in overnight at 4 °C on a shaker (300rpm)- mock IP did not have antibody added.

Protein A Dynabeads were blocked overnight with sheared salmon sperm DNA and BSA. (30µl/ sample). To the 3% BSA solution in IP buffer (0.1 % NaAz) a corresponding volume of Salmon sperm DNA per (2µl (4µg) per 100ul of 3%BSA) was added.

The next day the samples were cleared by centrifugation (12000 x g for 10 min at 4°C). The Dynabeads were washed 3 times with IP buffer- a wash consists of resuspending the beads in 1ml IP buffer, placing the tubes into the magnetic rack, and removing the supernatant after the beads have attached to the magnet. The top 90% of the cleared chromatin was added to new tubes, and the Dynabeads added to these tubes. These tubes were rotated for 1 hour (20-30 rotations per minute) at 4 °C, and then washed 3 times with cold IP buffer.

100 µl 10% (wt/vol) Chelex 100 slurry was added directly to the washed beads, which were then vortexed briefly and boiled for 10 minutes. Chelex stops the action of DNases and after this boiling step the DNA is stable and can be stored long term.

At this point additional purification was performed with the QIAquick PCR Purification Kit (Qiagen). This kit contains a silica membrane assembly that binds DNA in high-salt buffer and elutes with low-salt buffer or water- after a series of washes using the provided buffer through this silica membrane, the DNA can be eluted with no carry-over of the Chelex resin.

In collaboration with **Nadine Chapman-Rothe**, Sybr green real-time PCR measurement of the FBXO32 transcription start site and KRT17 promoter region following Chromatin Immunoprecipitation was performed, using antibodies to the histone marks shown, of MDA-MB-231 cells treated with 3 selected compounds at 5µM for 72h. Shown are representative examples of a series of ChIP experiments which consistently showed similar changes. The abundance relative to the untreated sample is shown. Each IP-value has been determined as the relative increase to the no-antibody control and then normalised to GAPDH levels.

Sybr green real time PCR measurement of the SPINK1 transcription start site was performed following Chromatin Immunoprecipitation, using antibodies to the histone marks shown, of MDA-MB-231 cells treated with HKMT-I-005 at 2.5µM or 7.5µM, HKMT-I-011 at 2.5µM for 24h. Each IP-value has been determined as the relative increase to the no-antibody control and is shown as abundance relative to the untreated control. Supplementary Table 8.11 contains ChIP QRT primer details.

Analysis of the ChIP QRT-PCR results in this case was performed relative to a no-antibody mock IP (and in the case of FBXO32 and KRT17 also normalised to GAPDH levels observed). Another layer of analysis would be the inclusion of an input control- in this case, post-sonication 1/10th of the sheared DNA from each sample is removed, and can then be quantified

for concentration and analysed for shearing efficiency. This would allow normalisation of each sample to its own internal control, accounting for differences observed that are due to variations in DNA extracted or shearing efficiency.

Cell proliferation assay

Lymphoma cells from established lymphoma cell lines (Anthony Uren) were plated at 20,000 cells in 200µl per well in Ubottom 96 well plates in RPMI medium + 20% FCS. 48 hours later cells were resuspended, diluted 10 fold in PBS + propidium iodide (PI), and the concentration of PI negative cells was counted using an Attune flow cytometer with autosampler. Breast cancer cells from established breast cancer cell lines (Elham Shamsaei) and ovarian cancer cells from established ovarian cancer cell lines (Sarah Kandil) were seeded at a density of 10000 cells/well in a sterile 96 clear-well plate with 150 µl of DMEM (+10% FCS and 2mM L-Glutamine). Each compound treatment was performed in triplicate for 72h at concentrations of 100nM, 1µM, 5µM, 10µM and 50µM in 100µl of full-medium. After 72h, 20µl of MTT solution (3mg of MTT Formazan, Sigma/1ml PBS) was added to the medium, and incubated for 4h at 37°C in a CO₂-incubator. The MTT-product was solubilised with 100µl DMSO and for 1h incubated in the dark at room-temperature. The optical density was read at 570nm with PHERAstar.

Lymphoma study was performed by **Anthony Uren**, breast cancer study by **Elham Shamsaei**, ovarian cancer study by **Sarah Kandil**. MDA-MB-231 study combining GSK343 and UNC0638 was performed by **Luke Payne** under above conditions, but after 48 hours treatment with drugs. Statistical significance of difference between treatments was calculated by unpaired 2-tailed Student's T-test.

Mutation data for EZH1, EZH2, EHMT1, and EHMT2 for these cell lines was accessed using COSMIC ¹²⁰.

Clonogenic assay

The breast cancer cell line MDA-MB231 and ovarian cancer cell line IGROV1 are maintained in DMEM (Sigma) or RPMI (Sigma) containing 10% FBS (First Link, UK), 2mM L-Glutamine (Gibco), 100U/ml Penicillin and 100µg/ml Streptomycin (Gibco) at 37°C in a humidified incubator, under 5-10% CO₂. Cells are tested for Mycoplasma contamination regularly by using sensitive bioluminescence based MycoAlert Mycoplasma detection kit (Lonza).

Cells were treated with drugs/control conditions under investigation for 24 hours, and then re-plated in 10cm³ culture dishes at a density of 1000 cells per plate. Colonies were left for 10-12 days with media refreshed every 4 days. Colony formation efficiency following each treatment was calculated relative to DMSO control.

CSC activity and self-renewal capacity

Ovarian cancer spheroid formation and breast cancer mammosphere formation were used as a proxy measure of CSC activity ¹²¹:

- Culture and detach cells at 70–80 % confluency according to standard protocols
- Centrifuge at 580 g for 2 min, remove supernatant and resuspend in 1–5 ml of ice-cold PBS
- Use a 25 G needle to syringe the cell suspension three times, to ensure a single cell suspension has formed

- Use a haemocytometer to confirm a single cell suspension is present (if it is not a single cell suspension, syringe a further three times) and calculate the number of viable cells per ml using trypan blue. Add 2 ml of spheroid media (detailed below) to each well in a low attachment 6-well plate
- Plate out cell suspension at 5000 cells per well
- Incubate in a humidified atmosphere at 37°C and 5 % CO₂ for 5 days without moving or disturbing the plates and without replenishing the media
- After 5 days, count the number of spheroids/mammospheres (at x40 magnification) which are greater than 50 µm diameter using a microscope fitted with a graticule
- Mammosphere/spheroid forming efficiency (%) is calculated as follows (mammosphere used as example):

$$(\text{number of mammospheres per well} / \text{number of cells seeded per well}) \times 100$$

Media- phenol red-free DMEM/F12 (Gibco, Paisley, UK; 21041) containing B27 supplement (no vitamin A; Invitrogen, Paisley, UK; 12587) and rEGF (20 ng/ml; Sigma Aldrich, Poole, UK; E-9644)

Low-attachment plates- Corning® Costar® Ultra-Low attachment multiwell plates coated in hydrogel (CLS3471-24EA, Sigma-Aldrich)

Second generation spheroid/mammosphere generation was used as a measure of CSC self-renewal capacity¹²¹:

- Pipette the media containing the spheroids/mammospheres from each well into a centrifuge tube
- Wash the wells with PBS, adding each wash to the collected media
- Centrifuge at 115 g for 5 min

- Discard supernatant and resuspend pellet in 300 µl of 0.5 % trypsin/0.2 % EDTA. A pellet may not be visible at this point and care must be taken when removing the supernatant so as not to dislodge the pellet. Incubate at 37°C for 2–3 min
- Disaggregate the mammospheres/spheroids using a 25 G needle and syringe until a single cell suspension is produced
- Neutralize trypsin with double the volume of serum-containing media
- Centrifuge at 580 g for 5 min
- Discard supernatant and resuspend pellet in a small volume (100–200 µl) of ice-cold PBS
- Check cells with haemocytometer. If a single cell suspension has not been achieved, syringe three more times using a 25 G needle
- Plate out the entire single cell suspension into low attachment plates (2 ml of spheroid media per well) at the same seeding density that was used in the primary generation
- Incubate in a humidified atmosphere at 37°C and 5 % CO₂ for 5 days without replenishing the media.
- Following the culture period, count the number of mammospheres/spheroids (at x40 magnification) which are greater than 50 µm diameter.
- Calculate CSC self-renewal capacity (example of mammospheres used for demonstration:

CSC self-renewal capacity= (number of second generation mammospheres/number of first generation mammospheres) e.g. if 50 mammospheres are counted in generation 1, they are dissociated, re-plated, and if 50 mammospheres generate then CSC self-renewal capacity=1

Xenograft culture

Initial xenograft study – 50,000 MDA MB 231 cells are injected sub-cutaneously into MSG mice (suspended in 1:1 Mammosphere media and Matrigel)

Tumours grow to approximately 200mm³ over 3 weeks before the start of treatment (day 1 on graph) - HKMT 40mg/kg give i.p once daily and Paclitaxel give once weekly (24hours after the first treatment of HKMT)

Graphs represent the fold change in tumour size ($\pm$ SEM) from the size of the tumour at day 1 of treatment (each point represents 10 tumours (apart from control where 8 tumours were used))

Secondary xenograft culture

Secondary implantation- MDA-MB-231 cells were extracted from primary treated tumour and 10 or 5 cells were re-injected sub-cutaneously into the flank of MSG mice. Each point represents mean $\pm$ SEM of tumour size mm³ - tumour size calculation = $L \times (W \times W)/2$

Extreme limiting dilution analysis

Using an online extreme limiting dilution analysis (ELDA) calculator

(<http://bioinf.wehi.edu.au/software/elda/>) the tumour take rates across the dilutions (10 and 5 cells) are input to calculate the approximate number of CSCs and the changes after treatment¹²². This estimates confidence intervals for 1/(stem cell frequency). The likelihood ratio test is designed to test whether the single-hit model is correct. The score test is designed to test whether the different cultures (assays) have the same active cell proportion.

Chapter 3: Evaluation of EZH2 and EHMT2 as therapeutic targets in cancer utilising publicly available data

3.1 Introduction

Expression, CNV, and mutational status of EZH2 and EHMT2 were investigated to explore what tumour types/subtypes may theoretically respond well to dual EZH2/EHMT2 inhibition- such information can inform on what tumour types may be most suitable for future research. At a patient level, understanding how EZH2 and EHMT2 expression, CNV, and mutational status links with clinical outcomes will allow stratification between patients, reducing the risk of unnecessary or ineffectual treatment.

To identify potential clinical settings in which dual HKMT inhibition may prove most beneficial, a variety of publicly available data were interrogated utilising both direct manipulation of data sets and analysis through data portals. These portals and datasets are detailed in the materials and methods section (as referenced throughout this chapter). The strength of utilising these resources lie in the large quantity of data available (e.g. Hazard modelling utilising data from over 3000 breast cancer patients) which lends power to the analysis and generality of results. Limitations of utilising such resources are the lack of control in the original experimental design and the lack of oversight of the initial data processing. As such, where possible, multiple sources have been interrogated in an attempt to assess reproducibility of findings. This work is based on the assumption that the expression level or mutational state of EZH2 or EHMT2 may affect the sensitivity of a cancer to treatment with the HKMT inhibitors (which may not necessarily be the case).

The expression of EHMT2, EZH2, and EZH2's fellow members of the PRC2 complex EED, SUZ12, and RbAp48¹²³ were investigated, as well as H3K27 demethylase JMJD3¹²⁴. In addition the expression of SPINK1 and RHOQ were investigated- these genes are putatively silenced by EZH2 mediated H3K27me3 and were identified as potential EZH2 targets in microarray studies (see Chapter 4).

The gene expression levels of EZH2, EHMT2, and the related genes detailed above were assessed in normal tissues to ascertain how their expression correlates with each other in different tissue types and to determine how consistent any correlation observed is. This also serves to highlight any potential normal tissues that show relatively high expression of EZH2 and EHMT2 and may be sensitive to dual inhibition of EZH2/EHMT2 (and thus potentially be sources for negative clinical side effects of the dual inhibitors).

Mutation and CNV of EZH2 and EHMT2 were queried as factors that may be driving EZH2/EHMT2 gene expression. Y641n somatic point mutations of EZH2 have been shown to drive high expression of EZH2 and high levels of H3K27me3 and occur frequently in follicular lymphoma and aggressive diffuse large B-cell lymphoma, contributing to the pathogenesis of the lymphomas¹²⁵. These EZH2 mutant lymphomas have been shown to be vulnerable to EZH2 specific inhibition⁵⁴. A pan-cancer review of EZH2 and EHMT2 mutation data was undertaken to establish if mutation may be driving high EZH2 expression in other cancer models, and if this mutation may be suitable to use as a tool for patient stratification for dual HKMT inhibition.

EZH2 and EHMT2 CNV were examined to determine this CNV may provide a potential stratification approach to selecting tumour types or patients for intervention by dual HKMT inhibition. The degree and variance of CNV of EZH2 and EHMT2 in different cancer types and

tumour types was assessed. This was related to gene expression, and the relationship of gene expression and CNV were examined with relation to clinical characteristics and outcomes.

Finally, the relationship between the gene expression of EZH2, EHMT2, and related molecular subunits and OS/RFS was assessed using Cox proportional hazard modelling of TCGA data as well as Kaplan-Meier analysis of a mixture of TCGA, the Genome Expression Omnibus, and The European Genome-phenome Archive data. These databases contain sequencing, gene expression, and mutation data as well as clinical data for large patient cohorts across multiple cancer types. These large cohorts allow the interrogation of relatively less common cancer subtypes, as well as giving power to statistical analyses. It is intended to attempt to identify cancer types or subtypes that may benefit most from dual HKMT inhibition and potential patient stratification, as well as reinforcing the case for utilising dual HKMT inhibitors in cancers where EZH2 expression has been linked to aggressive phenotypes (such as breast cancer^{126,127}).

3.2 Expression in normal tissues of EZH2, EHMT2, and related genes

To assess the relationship between EZH2 and EHMT2 expression across as many normal tissue types as possible, differential expression of target genes was investigated using a platform made available by the Harvard Centre for Computational & Integrative biology¹⁰³. This platform contains 126 normal primary human tissues (Detailed in **Materials and Methods: Calculation of differential expression (Harvard Centre for Computational & Integrative biology)**).

The expression of genes of interest (EZH2, EHMT2, SPINK1, RHOQ, KRT17, JMJD3, EED, RbAp48, and SUZ12) was profiled. KRT17 is repressed by EZH2 inhibition, RHOQ and SPINK1 are putative targets of EZH2 inhibition, and EED, RbAp48, and SUZ12 are subunits of the PRC2 complex along with EZH2. JMJD3 is a histone demethylase targeting H3K27me3.

In ES cells, haematopoietic stem cells, B cells, T cells, and most myeloid tissues, EZH2 shows a high level of expression (Fig.3.1). PRC2 subunits EED, RbAp48, and SUZ12 show a similar pattern of expression to EZH2. EHMT2 also displays a similar pattern of expression as EZH2 but notably shows generally lower expression in stem cells and myeloid cells.

EHMT2 shows high expression across some tissues in the central nervous system (Fig.3.2) but along with EZH2 and PRC2 sub-units EED, RbAp48, and SUZ12, shows low expression across most other tissues surveyed.

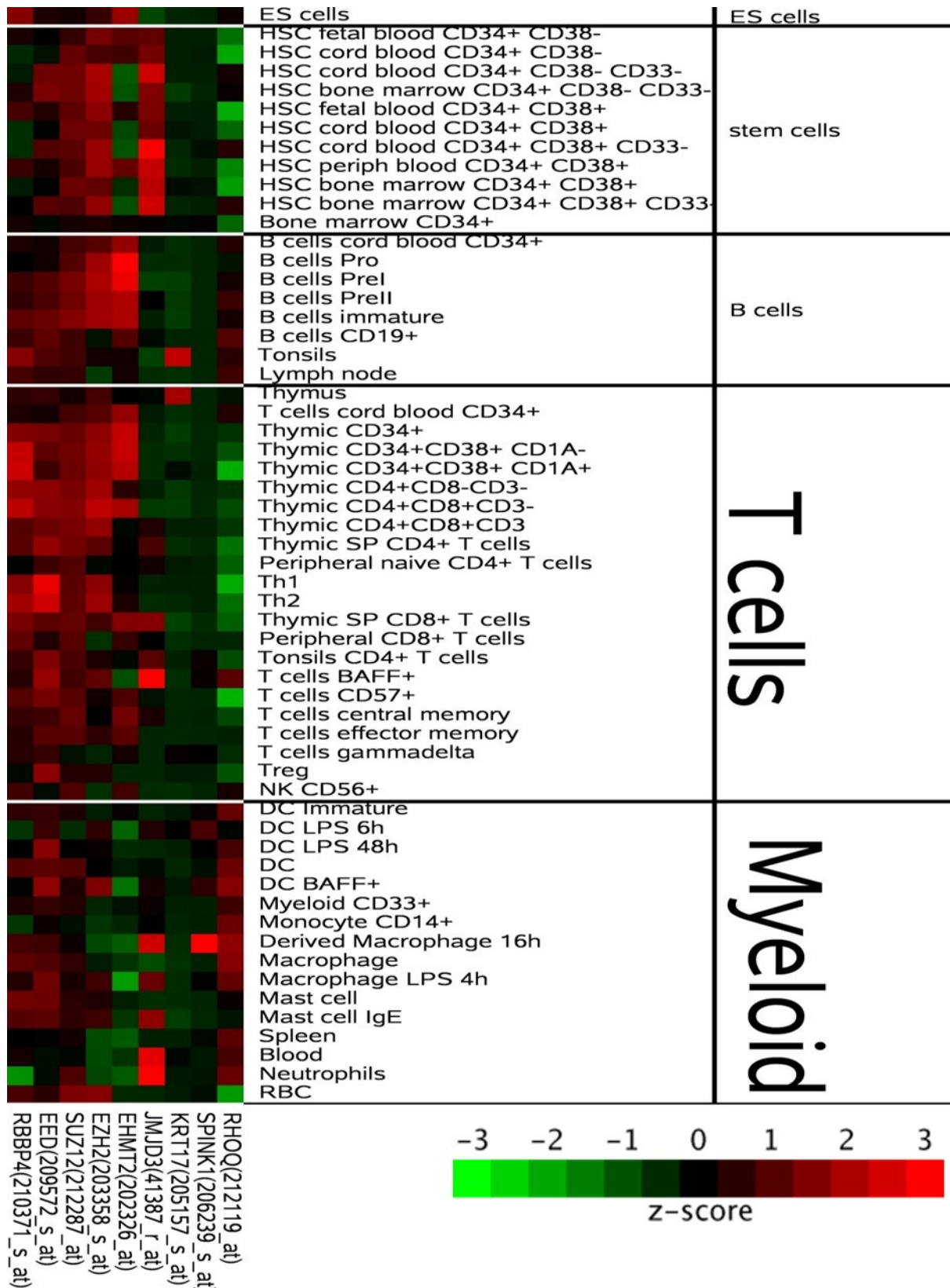


Figure 3.1- Differential expression of target genes across human normal primary tissues (ES cells, stem cells, B cells, T cells, and myeloid tissues) with Green representing low expression and Red representing high expression

This study of differential expression in normal primary human tissues highlights several points: EZH2 expression appears to largely correlate with the expression of other PRC2 components, low expression of SPINK1 and RHOQ appears to occur when EZH2 and EHMT2 are both highly expressed, and EZH2 and EHMT2 are both highly expressed in a number of tissues related to the immune system and haematopoietic system.

In order to quantify these relationships, the correlation between EZH2 expression, EHMT2 expression, and the expression levels of the other target genes was calculated (**Materials and methods: Correlation analysis (Harvard Centre for Computational & Integrative biology)**)).

EZH2 showed consistent negative expression correlation with RHOQ (Fig.3.3A) with the exception of B cells. This correlation was only statistically significant in Stem cells and Muscle cells (Supplementary Table 8.2). EZH2 expression correlated with SPINK1 expression (Fig.3.3A) positively in some cases (significantly (Supplementary Table 8.2) in Muscle, Airways, and Testis) and negatively in others (significantly (Supplementary Table 8.2) in Stem cells and B cells, also a trend shown in T cells/CNS).

For target genes KRT17 and JMJD3, EZH2 correlation varied in strength and direction across tissues (Fig.3.3B, significance in Supplementary Table 8.2) and similarly EZH2 expression correlation varied in strength and direction across tissues with SUZ12, EED, and RbAp48 expression (Fig.3.3C, significance in Supplementary Table 8.2). This highlights the variety in relationships between these subunits at a tissue level and indicates that the relationship between EZH2 and its related PRC2 subunits may be tissue specific.

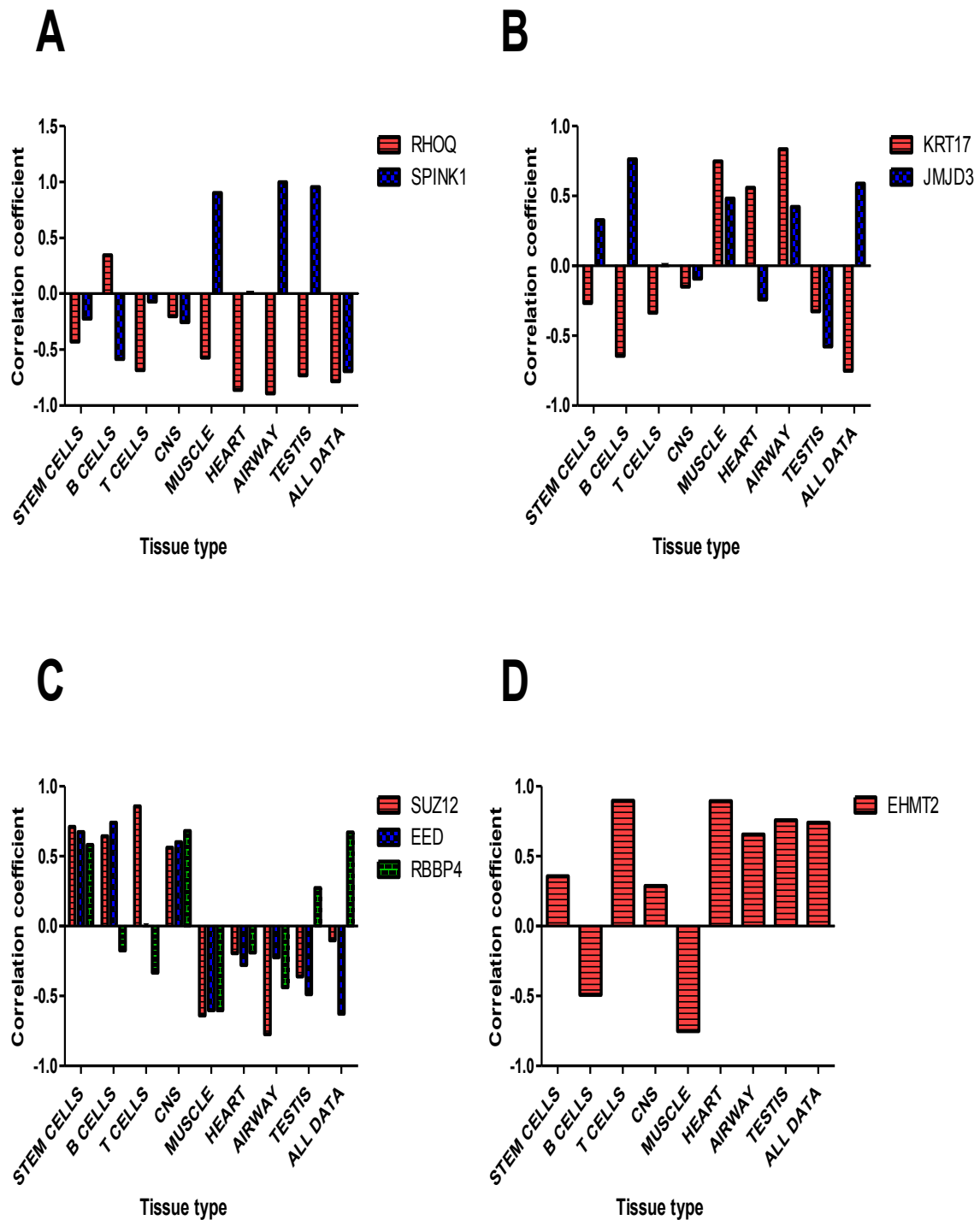


Figure 3.3- EZH2 correlation of expression in normal human tissue with expression of target genes A) putative dual HKMTi targets RHOQ and SPINK1 B) canonical targets KRT17 and JMJD3 C) PRC2 subunit components SUZ12, EED, and RbAp48 (RBBP4) D) EHMT2

Interestingly, with the exception of B cells and Muscle tissue, EZH2 expression positively correlates with EHMT2 expression (Fig.3.3D). This correlation is statistically significant in stem cells, T cells, and muscle cells (p-values in Supplementary Table 8.2). This result reinforces the hypothesis that the function of EZH2 and EHMT2 are intricately linked across numerous tissue types.

Interestingly one of the most significant expression correlations is the negative correlation seen in muscle cells between EZH2 and EHMT2. This data shows the heterogeneity in the relationship of EZH2 and these subunits in different tissues, highlighting the potential for different tissues to react in different manner to inhibition of HKMTs.

EHMT2 showed consistent significant negative expression correlation with RHOQ (Fig.3.4A) with the exception of CNS cells (Supplementary Table 8.3). EHMT2 expression correlated with SPINK1 expression (Fig.3.4A) significantly with most tissues (Fig.3.4.2A, Supplementary Table 8.3). Most of these significant correlations were negative, with the only positive correlations being in Heart and Airway tissues. These positive correlations were not significant.

In a manner similar to that shown with EZH2 (Fig.3.3B), EHMT2 showed a varied relationship with target genes KRT17 and JMJD3 both in terms of direction (Fig.3.4B) and significance, though overall the data pointed to a negative correlation in most tissue types. Again in a similar manner to the relationship between EZH2 expression and PRC2 subunit expression, EHMT2 expression correlated strongly with the expression of PRC2 subunits SUZ12, EED, and RbAp48, but the direction of this correlation was tissue type dependent.

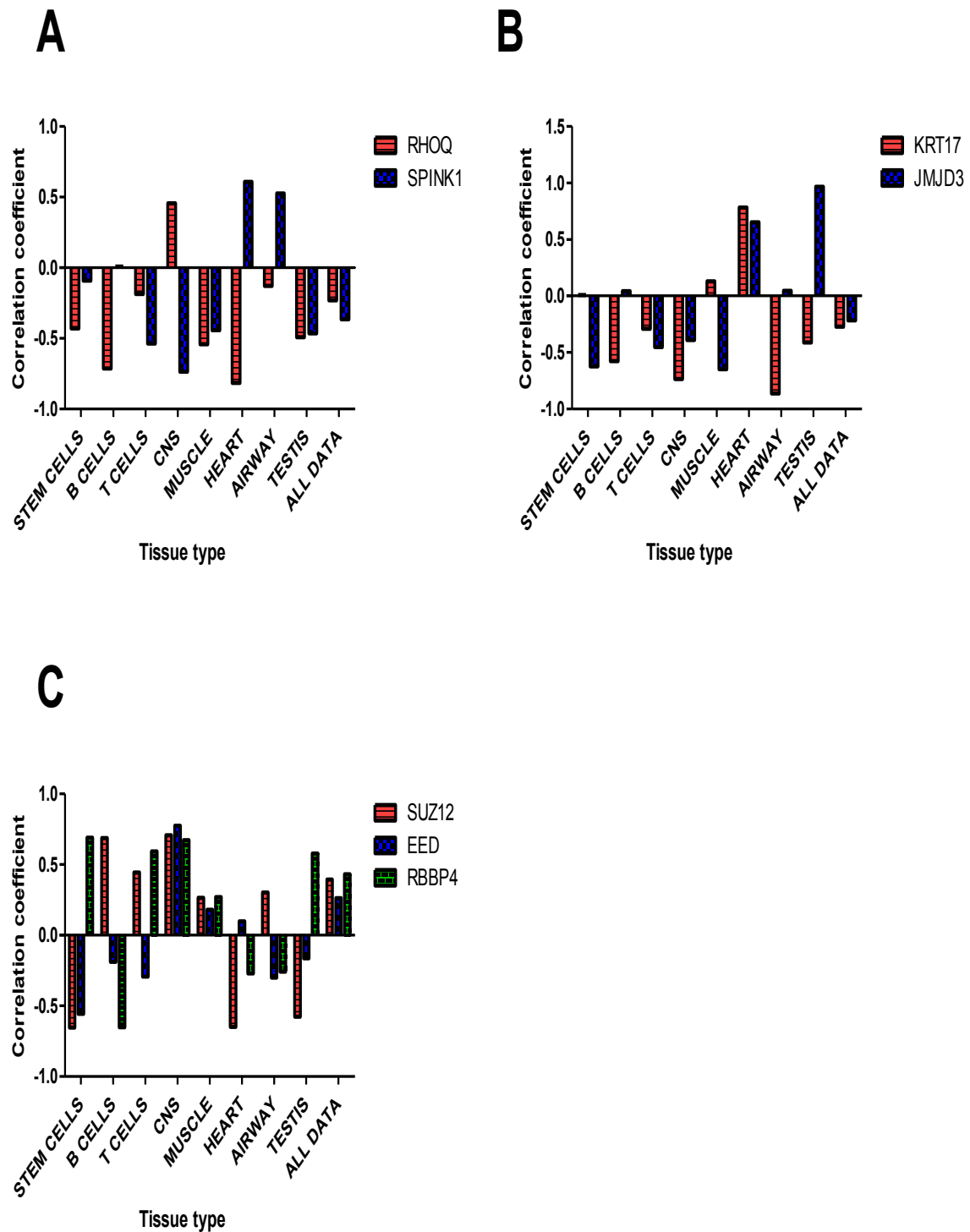


Figure 3.4- EMT2 correlation of expression in normal human tissue with expression of target genes A) putative dual HKMTi targets RHOQ and SPINK1 B) canonical targets KRT17 and JMJD3 C) PRC2 subunit components SUZ12, EED, and RbAp48 (RBBP4)

The levels of expression of EZH2, EHMT2, and related genes in normal human tissues appear to vary greatly dependent on the tissue in question (Figures 3.1/3.2). This indicates that these genes are by no means homogenous in terms of expression, and as such different tissues may respond to HKMT inhibition (either singular or dual) in differing manners dependent on the expression pattern.

The correlation of gene expression of these genes shows a similar range between tissues. Clear significant positive correlations can be seen between EZH2 and EHMT2 in most tissues studied (Fig.3.3D). However, it is clear that the correlation between EZH2 or EHMT2 and the chosen target genes/related subunits is heterogeneous in nature, varying in intensity and direction depending on the tissue type (Fig.3.3A, B, C/Fig.3.4A, B, C).

These results indicate a large degree of heterogeneity of EZH2/EHMT2 expression across tissue types in normal human tissues. Identifying if this tissue based heterogeneity persists in cancer phenotypes may help identify cancer types/subtypes that would most benefit from dual HKMT inhibition. Whilst the expression patterns of these genes are heterogeneous, EZH2 and EMT2 appear to be consistently linked in expression, reinforcing the potential impact of dual inhibition of EZH2 and EHMT2.

3.3 Expression of EZH2 and EHMT2 in cancerous tissues

Relative to normal tissue, expression of EZH2 has been observed as high and is linked to aggressive phenotypes in a number of cancers^{30,128–131}. Similarly, high expression of EHMT2 has been linked to aggressive phenotypes and poor clinical outcomes^{60,61,132,133}.

Utilising the CancerMA analysis tool (**Materials and methods: CancerMA Forest plots**) the expression of EZH2 and EHMT2 was analysed in 80 cancer microarray data sets covering 13 cancer types sourced from ArrayExpress and the Gene Expression Omnibus.

Analysis of EHMT2 expression in these microarrays shows that in 4 of the 13 cancer types studied (Lung, Adrenal, Brain, and prostate) EHMT2 shows an increase in expression in comparison to normal tissue (Fig.3.5A-D) with a log2 Fold Change increase in expression of ~0.5-1.5 in these four cancers.

Analysis of EZH2 expression in these microarrays shows that in 9 of the 13 cancer types studied (Renal, Ovarian, Brain, Thyroid, Adrenal, Colorectal, Lung, Breast, and Prostate); EZH2 shows an increase in expression in comparison to normal tissue (Fig.3.6) with a log2 Fold Change increase in expression of ~1.0-3.0 in these nine cancers. It should be noted that this platform shows that EZH2 or EHMT2 are strongly up-regulated across a number of cancers, but does not provide robust statistical analyses of these changes in expression.

In addition it shows us where EZH2 and EHMT2 both show up-regulation of expression in cancer in comparison to normal tissue: Adrenal, Brain, Lung, and Prostate cancers.

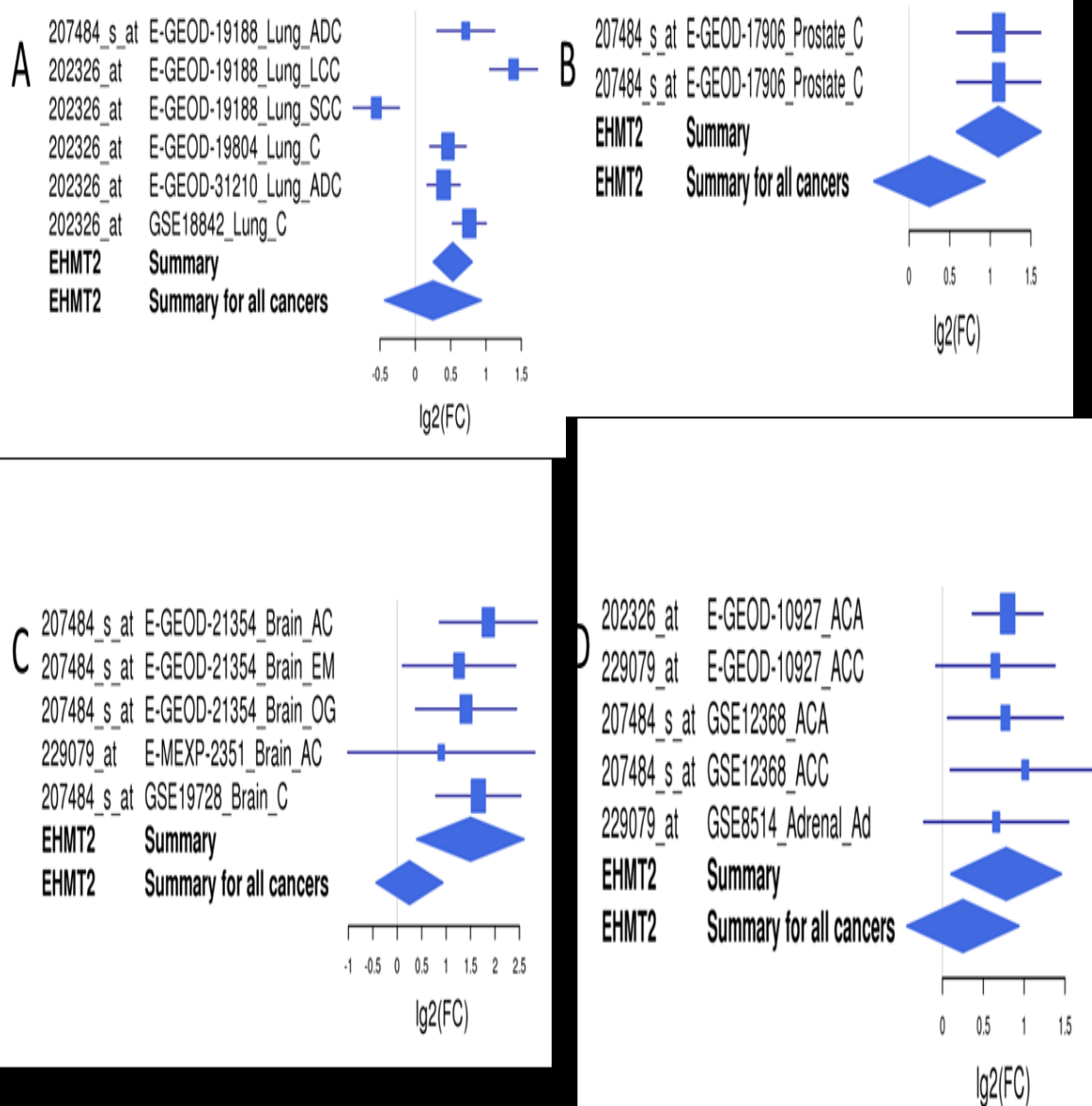


Figure 3.5- Differential expression of EHMT2 calculated as the meta-log 2-fold change in cancerous tissue relative to matched normal tissue - EHMT2 shows increased expression in the following cancers: A) Lung B) Prostate C) Brain D) Adrenal

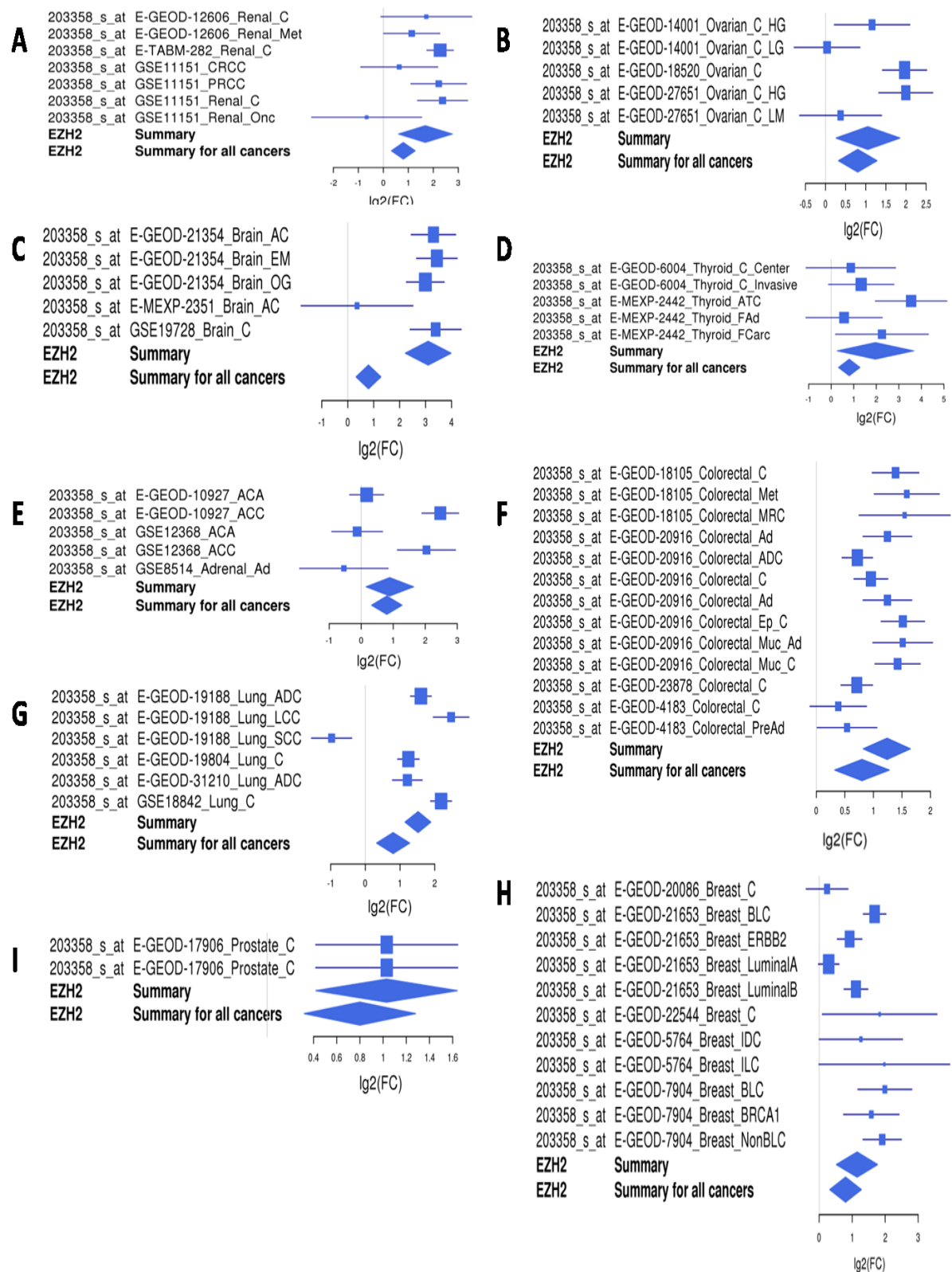


Figure 3.6- Differential expression of EZH2 calculated as the meta-log 2-fold change in cancerous tissue relative to matched normal tissue - EZH2 shows increased expression in the following cancers: A) Renal B) Ovarian C) Brain D) Thyroid E) Adrenal F) Colorectal G) Lung H) Breast I) Prostate

To confirm these results a second data-set was utilised (**Materials and methods: Calculation of differential expression (Harvard Centre for Computational & Integrative biology)**) examining 16 cancerous human tissues (represented by 92 different microarrays), expression levels in cancer were visualised for the genes EZH2, RbAp48, SUZ12, EED, EHMT2, SPINK1, and KRT17 (Fig. 3.7).

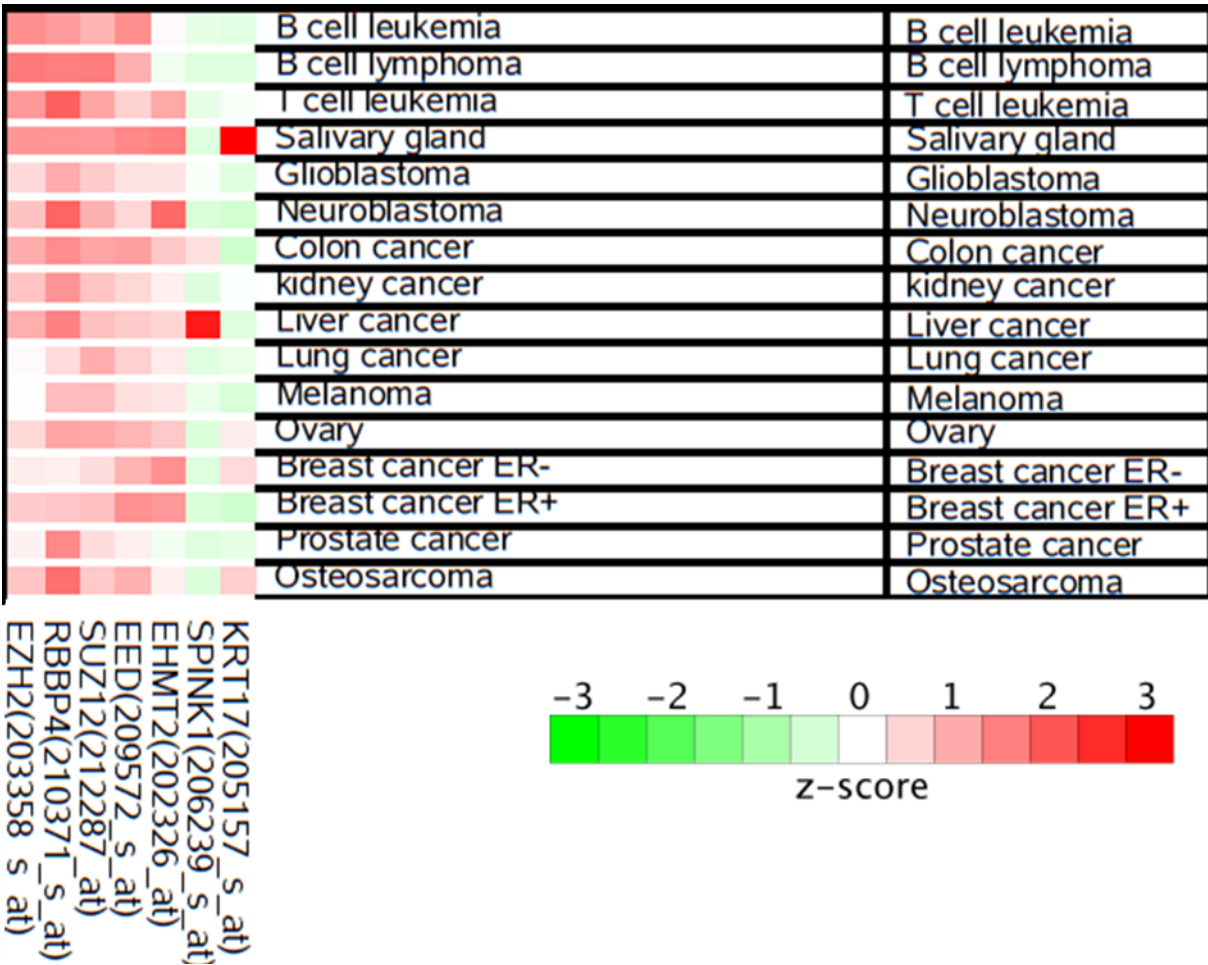


Figure 3.7- Differential expression of target genes across human cancer tissues with Green representing low expression and Red representing high expression

The HKMTs EZH2 and EHMT2 (and PRC2 subunits RbAp48, SUZ12, and EED) show either average or high expression in nearly every cancer present in the dataset (Figure 3.7).

It is clear that whilst EZH2, EHMT2, and related subunits show a large degree of variety in expression profiles across normal human tissues (Figures 3.1-4) but in the setting of cancer a slightly more homogenous expression profile can be seen (Figures 3.5-7). Adrenal, Brain, Lung, and Prostate cancers all show high expression of both EZH2 and EHMT2 (Figures 3.5/6), though it is worth noting that in order for the theorised HKMT driven disease phenotype to be present, high expression may only be required by one of these HKMT, with the other being expressed at a normal physiological level.

The general high expression of EZH2 and EHMT2 across a number of cancer types and datasets indicates that targeted inhibition of EZH2 and EHMT2 may have potential to impact on numerous cancer types. However, it is worth noting that where subtype information is available (such as ER-/ER+ breast cancer in Fig.3.7) the expression of these targets is not always consistent between subtypes. This highlights the need to stratify patient data using available clinical criteria in order to ascertain the best application of potential inhibitors of EZH2 and EHMT2.

3.4 Mutations in EZH2 and EHMT in cancerous tissues

Having shown the general up-regulation of EZH2 and EHMT2 across different cancers, the driving force behind this expression is unclear. One postulated factor that could impact EZH2 and EHMT2 expression and potentially help stratify patients for treatment is the mutational status of these genes. As previously mentioned some mutations (e.g. Y641n mutation in EZH2 in follicular lymphoma ¹²⁵) lead to high levels of EZH2 expression and increased levels of H3K27me3. With the advent of large consortia such as the International Cancer Genome Consortium ¹³⁴ and The Cancer Genome Atlas (TCGA) Research Network, large cancer datasets are available to be probed for information on target genes.

Somatic mutations were investigated across multiple cancer types (**Materials and methods: Mutation rate, CNV, and expression of target genes in TCGA data**) utilising the Cbio portal to TCGA datasets, allowing the degree of somatic mutational alterations to be quantified (Tables 3.1/2) and the visualisation of the location of these mutations (Fig.3.8A/B). The mutational status of EZH2 observed in TCGA data is summarised in Table 3.1.

Overall, mutations (sequence variants) of EZH2 appear to be infrequent, never encompassing more than 5% of the cases within a given cohort, and those cancer types shown to have high levels of EZH2 expression (such as Renal, Ovarian, Brain, Thyroid, Adrenal, Colorectal, Lung, Breast, and Prostate (Fig.3.8) show few or no reported mutations.

Table 3.1- Reported EZH2 mutations in TCGA data (only cancers with observed mutations included)

Study abbreviation	Study name	Number of cases altered	Percentage of cases altered
Uterine (TCGA pub)	Uterine Corpus Endometrioid Carcinoma (TCGA, Nature 2013)	12	4.80%
Uterine (TCGA)	Uterine Corpus Endometrial Carcinoma (TCGA, Provisional)	12	4.80%
Head & neck (Broad)	Head and Neck Squamous Cell Carcinoma (Broad, Science 2011)	3	4.10%
Melanoma (TCGA)	Skin Cutaneous Melanoma (TCGA, Provisional)	11	4%
Melanoma (Broad)	Skin Cutaneous Melanoma (Broad, Cell 2012)	4	3.30%
Melanoma (Yale)	Skin Cutaneous Melanoma (Yale, Nature Genetics 2012)	3	3.30%
Colorectal (Genentech)	Colorectal Adenocarcinoma (Genentech, Nature 2012)	2	2.80%
Lung adeno (TCGA)	Lung Adenocarcinoma (TCGA, Provisional)	6	2.60%
Cervical (TCGA)	Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (TCGA, Provisional)	1	2.60%
Lung SC (JHU)	Small Cell Lung Cancer (Johns Hopkins, Nature Genetics 2012)	1	2.40%
Bladder (TCGA pub)	Bladder Urothelial Carcinoma (TCGA, Nature 2014)	3	2.30%
Stomach (TCGA)	Stomach Adenocarcinoma (TCGA, Provisional)	5	2.30%
Lung squ (TCGA)	Lung Squamous Cell Carcinoma (TCGA, Provisional)	4	2.30%
Lung squ (TCGA pub)	Lung Squamous Cell Carcinoma (TCGA, Nature 2012)	4	2.20%
Esophagus (Broad)	Esophageal Adenocarcinoma (Broad, Nature Genetics 2013)	3	2.10%
Colorectal (TCGA)	Colorectal Adenocarcinoma (TCGA, Provisional)	4	1.80%
Colorectal (TCGA pub)	Colorectal Adenocarcinoma (TCGA, Nature 2012)	4	1.80%
Uterine CS (TCGA)	Uterine Carcinosarcoma (TCGA, Provisional)	1	1.80%
Lung adeno (TCGA pub)	Lung Adenocarcinoma (TCGA, Nature, in press)	4	1.70%
NCI-60	NCI-60 Cell Lines (NCI, Cancer Res. 2012)	1	1.70%
AML (TCGA)	Acute Myeloid Leukemia (TCGA, Provisional)	3	1.50%
AML (TCGA pub)	Acute Myeloid Leukemia (TCGA, NEJM 2013)	3	1.50%
Pancreas (TCGA)	Pancreatic Adenocarcinoma (TCGA, Provisional)	1	1.10%
GBM (TCGA)	Glioblastoma Multiforme (TCGA, Provisional)	3	1.10%
Liver (AMC)	Liver Hepatocellular Carcinoma (AMC, Hepatology in press)	2	0.90%
ccRCC (TCGA)	Kidney Renal Clear Cell Carcinoma (TCGA, Provisional)	3	0.70%
ccRCC (TCGA pub)	Kidney Renal Clear Cell Carcinoma (TCGA, Nature 2013)	3	0.70%
GBM (TCGA 2013)	Glioblastoma (TCGA, Cell 2013)	2	0.70%
Lung adeno (Broad)	Lung Adenocarcinoma (Broad, Cell 2012)	1	0.50%
Prostate (TCGA)	Prostate Adenocarcinoma (TCGA, Provisional)	1	0.40%
Head & neck (TCGA pub)	Head and Neck Squamous Cell Carcinoma (TCGA, in revision)	1	0.40%
Head & neck (TCGA)	Head and Neck Squamous Cell Carcinoma (TCGA, Provisional)	1	0.30%
Breast (TCGA)	Breast Invasive Carcinoma (TCGA, Provisional)	3	0.30%
Breast (TCGA pub)	Breast Invasive Carcinoma (TCGA, Nature 2012)	1	0.20%
CCLE	Cancer Cell Line Encyclopedia (Novartis/Broad, Nature 2012)	1	0.10%

EHMT2 also shows largely low levels of somatic mutations in cancer cases (Table 3.2) with the exception of Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (though this study contains very few cases and as such can only be regarded provisionally). As with EZH2, cancers that have shown upregulation of EHMT2 such as Adrenal, Brain, Lung, and Prostate, show little in the way of mutations, and never above 5% of the cases within each given cohort.

The location of the mutations that were observed (Fig.3.8) illustrates the wide range of locations of reported missense and nonsense mutations within EZH2 and EHMT2. Notably, missense mutation at Y641n (Fig.3.8A) has previously been shown in follicular lymphoma as driving increased expression and activity of EZH2¹³⁵, and it is this mutation that shows the highest number of reported cases.

Table 3.2- Reported EHMT2 mutations in TCGA data (only cancers with observed mutations included)

Study abbreviation	Study name	Number of cases altered	Percentage of cases altered
Cervical (TCGA)	Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (TCGA, Provisional)	3	7.70%
Melanoma (Broad)	Skin Cutaneous Melanoma (Broad, Cell 2012)	5	4.10%
Melanoma (TCGA)	Skin Cutaneous Melanoma (TCGA, Provisional)	11	4%
Pancreas (TCGA)	Pancreatic Adenocarcinoma (TCGA, Provisional)	3	3.30%
Bladder (TCGA pub)	Bladder Urothelial Carcinoma (TCGA, Nature 2014)	4	3.10%
chRCC (TCGA)	Kidney Chromophobe (TCGA, Provisional)	2	3%
pRCC (TCGA)	Kidney Renal Papillary Cell Carcinoma (TCGA, Provisional)	5	3%
Stomach (TCGA)	Stomach Adenocarcinoma (TCGA, Provisional)	6	2.70%
ACC (TCGA)	Adrenocortical Carcinoma (TCGA, Provisional)	2	2.20%
Head & neck (TCGA)	Head and Neck Squamous Cell Carcinoma (TCGA, Provisional)	6	2%
Head & neck (TCGA pub)	Head and Neck Squamous Cell Carcinoma (TCGA, in revision)	5	1.80%
Prostate (MICH)	Prostate Adenocarcinoma, Metastatic (Michigan, Nature 2012)	1	1.60%
Esophagus (Broad)	Esophageal Adenocarcinoma (Broad, Nature Genetics 2013)	2	1.40%
Head & neck (Broad)	Head and Neck Squamous Cell Carcinoma (Broad, Science 2011)	1	1.40%
Colorectal (TCGA)	Colorectal Adenocarcinoma (TCGA, Provisional)	3	1.30%
Colorectal (TCGA pub)	Colorectal Adenocarcinoma (TCGA, Nature 2012)	3	1.30%
Lung squ (TCGA pub)	Lung Squamous Cell Carcinoma (TCGA, Nature 2012)	2	1.10%
Lung adeno (Broad)	Lung Adenocarcinoma (Broad, Cell 2012)	2	1.10%
Lung adeno (TCGA pub)	Lung Adenocarcinoma (TCGA, Nature, in press)	2	0.90%
Uterine (TCGA pub)	Uterine Corpus Endometrioid Carcinoma (TCGA, Nature 2013)	2	0.80%
Uterine (TCGA)	Uterine Corpus Endometrial Carcinoma (TCGA, Provisional)	2	0.80%
ccRCC (TCGA)	Kidney Renal Clear Cell Carcinoma (TCGA, Provisional)	3	0.70%
ccRCC (TCGA pub)	Kidney Renal Clear Cell Carcinoma (TCGA, Nature 2013)	3	0.70%
GBM (TCGA)	Glioblastoma Multiforme (TCGA, Provisional)	2	0.70%
GBM (TCGA 2013)	Glioblastoma (TCGA, Cell 2013)	2	0.70%
Lung squ (TCGA)	Lung Squamous Cell Carcinoma (TCGA, Provisional)	1	0.60%
MM (Broad)	Multiple Myeloma (Broad, Cancer Cell 2014)	1	0.50%
Lung adeno (TCGA)	Lung Adenocarcinoma (TCGA, Provisional)	1	0.40%
Breast (TCGA)	Breast Invasive Carcinoma (TCGA, Provisional)	4	0.40%
Prostate (TCGA)	Prostate Adenocarcinoma (TCGA, Provisional)	1	0.40%
Glioma (TCGA)	Brain Lower Grade Glioma (TCGA, Provisional)	1	0.30%
Ovarian (TCGA)	Ovarian Serous Cystadenocarcinoma (TCGA, Provisional)	1	0.30%
Ovarian (TCGA pub)	Ovarian Serous Cystadenocarcinoma (TCGA, Nature 2011)	1	0.30%
Thyroid (TCGA)	Thyroid Carcinoma (TCGA, Provisional)	1	0.20%



Figure 3.8- Visualisation of mutations observed across cancer types in A) EZH2 B) EHMT2 with catalytic SET domain highlighted- the number of mutations recorded across all TCGA data is shown on the y axis, the location of mutation on the target on the x axis, and the type of mutation indicated by colour (green= predicted missense mutation, red= predicted nonsense mutation)

Clearly mutations of EZH2 and EHMT2 are not common at a pan-cancer level. Whilst cancers with certain EZH2 mutations have been shown to be susceptible to treatment with EZH2 inhibitors (such as follicular lymphoma with the aforementioned Y641n point mutation), the relative scarcity of these mutations and the lack of overlap with indicates that whilst mutational status of EZH2 or EHMT2 may indicate susceptibility to HKMT inhibition, it is likely not the driving force behind the increased EZH2/EHMT2 expression in most cancer tissues and in most cases will be unsuitable as a tool for stratification. In specific cancers such as the reported follicular lymphoma, intervention against EZH2 mutation driven epigenetic silencing with dual HKMT inhibitors may prove beneficial.

3.5 EZH2 and EHMT2 CNV in cancerous tissues

One potential factor that could explain the widespread up-regulation of EZH2 and EHMT2 expression is CNV, and potentially the degree of CNV could act as a stratification tool for identification of cancerous tissues that may benefit from dual HKMT inhibition.

Utilising CNV data from TCGA datasets (**Materials and Method: Mutation rate, CNV, and expression of target genes in TCGA data**), the number of cases showing CNV of EZH2 was estimated (Table 3.3).

Table 3.3- Reported EZH2 copy number variation in TCGA data (only cancers showing <2% cases altered included) n.b. Provisional denotes published data with additional cases added post publication

Study name	Number of cases altered	Percentage of cases altered
Ovarian Serous Cystadenocarcinoma (TCGA, Provisional)	67	11.80%
Ovarian Serous Cystadenocarcinoma (TCGA, Nature 2011)	29	5.90%
Skin Cutaneous Melanoma (TCGA, Provisional)	18	5.40%
Cancer Cell Line Encyclopaedia (Novartis/Broad, Nature 2012)	50	5%
Prostate Adenocarcinoma, Metastatic (Michigan, Nature 2012)	3	4.90%
Glioblastoma Multiforme (TCGA, Provisional)	22	4.40%
Brain Lower Grade Glioma (TCGA, Provisional)	8	3%
Acute Myeloid Leukemia (TCGA, NEJM 2013)	5	2.60%
Acute Myeloid Leukemia (TCGA, Provisional)	5	2.60%
Lung Adenocarcinoma (TCGA, in revision)	6	2.60%
Lung Adenocarcinoma (TCGA, Provisional)	6	2.60%
Sarcoma (TCGA, Provisional)	2	2.40%
Head and Neck Squamous Cell Carcinoma (TCGA, Provisional)	7	2.30%

In the cancers that previously (Section 3.3) showed high levels of expression of EZH2 and EHMT2 (Adrenal, Brain, Lung, and Prostate), Glioblastoma Multiforme (TCGA, Provisional) shows 4.4% CNV, Prostate Adenocarcinoma, Metastatic (Michigan, Nature 2012) shows 9%, and Lung Adenocarcinoma (TCGA, in revision) shows 2.6%.

The cancer that shows the highest degree of CNV is ovarian serous cystadenocarcinoma (TCGA, Provisional), with an estimated 11.8% of cases with some form of copy number variation.

EHMT2 also shows a degree of CNV, as summarised in Table 3.4.

Table 3.4- Reported EHMT2 copy number variation in TCGA data (only cancers showing <2% cases altered included) n.b. Provisional denotes published data with additional cases added post publication

Study name	Number of cases altered	Percentage of cases altered
Ovarian Serous Cystadenocarcinoma (TCGA, Provisional)	34	6%
Cancer Cell Line Encyclopaedia (Novartis/Broad, Nature 2012)	61	6.10%
Skin Cutaneous Melanoma (TCGA, Provisional)	14	4.20%
Prostate Adenocarcinoma (Broad/Cornell, Cell 2013)	2	3.60%
Lung Adenocarcinoma (TCGA, in revision)	8	3.50%
Lung Adenocarcinoma (TCGA, Provisional)	8	3.50%
Prostate Adenocarcinoma, Metastatic (Michigan, Nature 2012)	2	3.30%
Stomach Adenocarcinoma (TCGA, Provisional)	9	3.10%
Liver Hepatocellular Carcinoma (TCGA, Provisional)	4	2.90%
Ovarian Serous Cystadenocarcinoma (TCGA, Nature 2011)	13	2.70%
Pancreatic Adenocarcinoma (TCGA, Provisional)	1	2%

In the cancers that previously (Section 3.3) showed high levels of expression of EZH2 and EHMT2 (Adrenal, Brain, Lung, and Prostate), Lung Adenocarcinoma (TCGA, in revision) shows 3.5% CNV, and Prostate Adenocarcinoma (Broad/Cornell, Cell 2013) showed 3.6% alteration.

Interestingly, whilst Ovarian Serous Cystadenocarcinoma (TCGA, Provisional) shows the highest degree of EHMT2 CNV according to TCGA GISTIC analysis with 6% of cases altered,

previous analysis (section 3.3) did not highlight EHMT2 as being upregulated in ovarian cancer tissue.

From a therapeutic standpoint, utilising CNV to highlight tissues that may be susceptible to dual HKMT inhibition depends on said CNV conferring some alteration to the expression level of EZH2 or EHMT2. As such the relationship between observed CNV and gene expression will be investigated.

3.6 Relationship between target gene CNV, target gene expression, and clinical characteristics in cancerous tissues

As CNV of EZH2 and EHMT2 appear relatively common, the question as to if this CNV is driving expression of these target genes must be addressed in order to evaluate CNV as a potential tool for stratifying patients potentially susceptible to dual HKMT inhibition. In addition, how these factors relate to clinical characteristics may highlight particular clinical phenotypes linked to either EZH2/EHMT2 CNV or expression.

Ovarian serous cystadenocarcinoma showed the largest degree of CNV for EZH2 (11.8% of cases, Table 3.3) and EHMT2 (6% of cases, Table 3.4). In ovarian serous cystadenocarcinoma when expression data is correlated with estimated CNV status (**Materials and Methods: Mutation rate, CNV, and expression of target genes in TCGA data**) a trend toward higher expression with CNV gain and amplification can be observed in EZH2 (Fig.3.9A) and EHMT2 (3.9B).

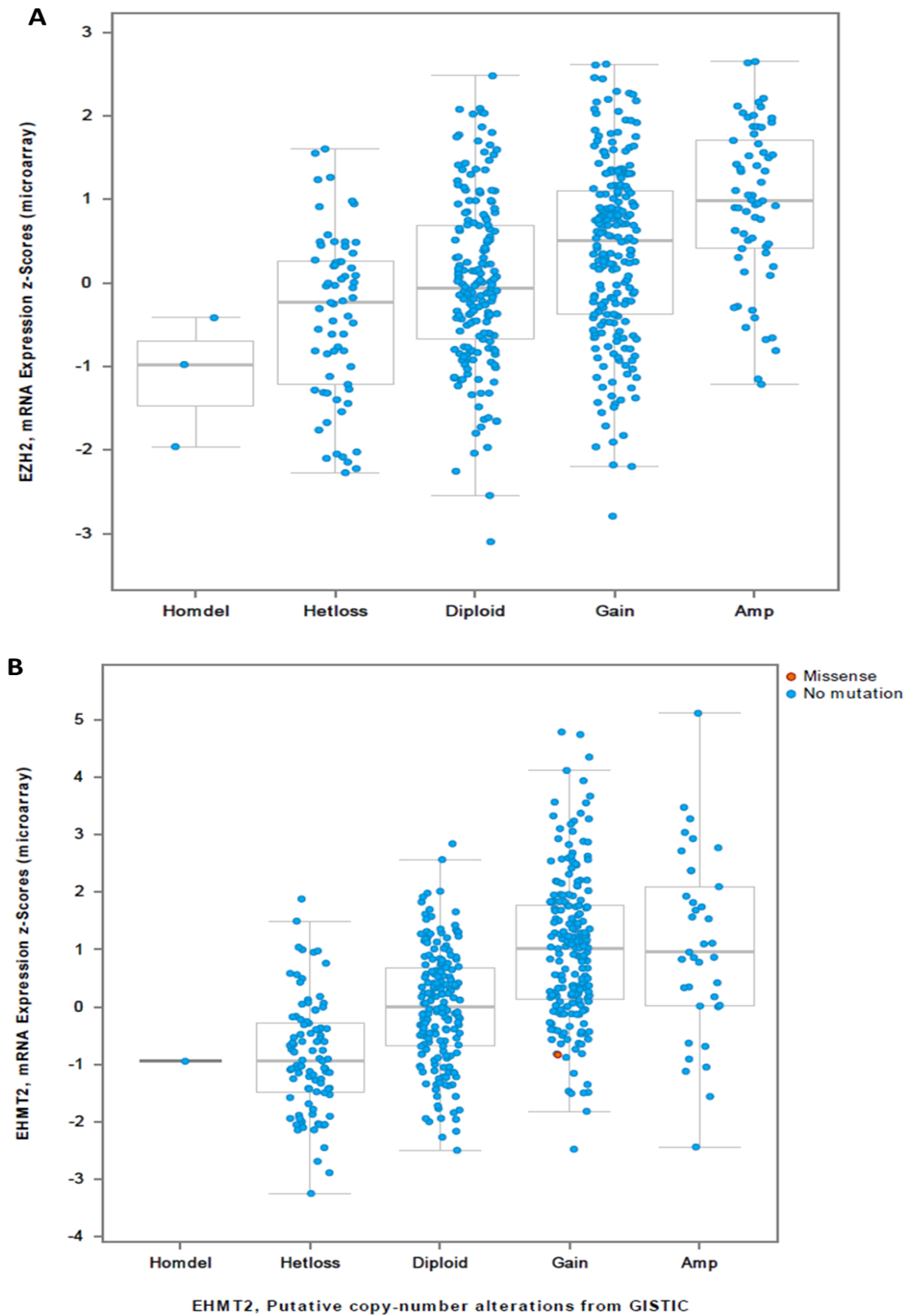


Figure 3.9- copy number and mRNA expression in 570 ovarian serous cystadenocarcinoma cases for A) EZH2 B) EHMT2 (mRNA z-Scores (Agilent microarray) compared to the expression distribution of each gene in tumours that are diploid for this gene, putative copy-number calls on 570 cases determined using GISTIC 2.0). Error bars are SEM.

The relationship between EZH2/EHMT2 CNV, expression, and clinical characteristics was quantified (**Materials and Methods: Comparison of gene expression, clinical data, and CNV in TCGA data**) using raw TCGA data (number of cases for each cancer examined shown in Table 3.5).

Table 3.5- Summary of data analysed for CNV/expression correlation obtained from TCGA

Cancer type	Number of cases
Ovarian	513
Breast	484
Colon	166
Glioblastoma multiforme	163
Kidney renal clear cell	70
Kidney renal papillary cell	12
Low grade glioma	27
Lung	32
Rectal	72
Uterine corpus endometrioid	54

In ovarian cancer data obtained from TCGA (Table 3.6), an examination of the correlation between expression and CNV was carried out (**Materials and Methods: Comparison of gene expression, clinical data, and CNV in TCGA data**). EZH2 CNV correlated significantly and positively with EZH2 expression in Ovarian, Breast, Colon, Glioblastoma multiforme, and rectal cancers.

EHMT2 CNV correlated significantly and positively with EHMT2 expression in all cancers studied with the exception of Kidney renal papillary cell, which showed no significance (though still showed positive correlation), reinforcing the results seen in normal tissues (Section 3.3). EZH2 CNV showed no significant correlation with EHMT2 expression or CNV. However, EZH2 expression correlated positively with EZH2 CNV in all cancer types, significantly so in all cancers studied except Low Grade Glioma and Lung (both of which have very low case numbers in the TCGA cohort which may explain the lack of significance shown).

Table 3.6- Pearson correlation of EZH2 and EHMT2 expression and copy number in TCGA cancer data (number of cases in Table 3.5), correlations with significance $p < 0.05$ highlighted yellow

	EZH2 CNV/EZH2 Expression	EHMT2 CNV/EHMT2 EXPRESSION	EZH2 CNV/EHMT2 EXPRESSION	EZH2 CNV/EHMT 2 CNV	EZH2 EXPRESSION/EHMT 2 EXPRESSION
Ovarian	0.5018275	0.6417792	0.0867746	0.0231867	0.1984618
Breast	0.3917432	0.4708633	0.07528444	-0.0271219	0.4281819
Colon	0.3678733	0.4537508	0.1462429	0.1228475	0.4984117
Glioblastoma multiforme	0.2853246	0.2529441	-0.05254869	0.0031743	0.2044092
Kidney renal clear cell	0.1571775	0.5022047	-0.1382232	-0.1764195	0.4706167
Kidney renal papillary cell	0.2210136	0.3838006	-0.037976	-0.2732419	0.6609735
Low grade glioma	0.3574323	0.4099649	-0.3785759	-0.0820694	0.2818888
Lung	0.2921452	0.5534537	-0.06399533	-0.0409352	0.2473632
Rectal	0.5105617	0.4714538	0.1230174	-0.0525343	0.2699952
Uterine corpus endometrioid	0.1088912	0.7435015	-0.09312409	-0.0572996	0.3813081

It appears that CNV of either EZH2 or EHMT2 tends to correlate with expression of said gene, but the significance and amplitude of this effect is dependent on cancer type. As the degree of EZH2 CNV is not always significantly or strongly associated with expression of EZH2, and as the degree of CNV of these genes within each cancer cohort is consistently low, utilising CNV as a stratification tool may not be a sound clinical strategy. It is clear however that expression of the targets of the dual HKMT inhibition EZH2 and EHMT2 are consistently correlated across numerous cancer types, which reinforces the potential of dual inhibition.

EZH2 expression has been linked to aggressive phenotypes in breast cancer¹²⁶ and ovarian cancer³⁵ - in order to support these findings the degree to which EZH2 CNV/expression and EHMT2 CNV/expression correlate with clinical outcomes (**Materials and Methods:**

Comparison of gene expression, clinical data, and CNV in TCGA data) was studied across

TCGA cancer data to elucidate if there are any clinical characteristics associated with these CNV or expression states (Table 3.7).

In Breast cancer, expression levels of EZH2 and EHMT2 significantly correlate with negative progesterone receptor status and negative oestrogen receptor status. This data support the case for intervention by dual HKMT inhibition in breast cancer by further illustrating links between EZH2, EHMT2, and negative clinical phenotypes (such as PR-/ER- breast cancer tumours).

EHMT2 CNV negatively correlates with progression free status in ovarian cancer (to a significant degree), but otherwise no relation between targets and clinical outcomes are seen in ovarian cancer.

High expression of EZH2 and EHMT2 significantly correlate with higher tumour stage in Kidney renal clear cell carcinoma. In rectal cancer, higher levels of EZH2 copy number significantly correlates with a higher age at initial pathologic diagnosis, however in uterine corpus endometrioid cancer EZH2 CNV negatively correlates with age at initial pathologic diagnosis. EHMT2 expression positively significantly correlates with age at initial pathologic diagnosis, and EHMT2 CNV positively correlates with the number of days to death.

Table 3.7- Pearson correlation of EZH2 and EHMT2 expression and copy number with clinical outcomes in TCGA cancer data (number of cases in Table 3.5), correlations with $p < 0.05$ highlighted yellow

Cancer type	Clinical characteristic	EZH2 expression	EZH2 copy number variation	EHMT2 expression	EHMT2 copy number variation
Ovarian	Age at diagnosis	0.0017776	-0.0351069	0.033942	-0.0343262
	Tumour grade	0.0741607	-0.0076188	-0.003931	-0.0257424
	Progression free status	-0.0151903	0.0509963	-0.0503167	-0.1077526
Breast	Age at initial pathologic diagnosis	-0.0778291	-0.105157	-0.0560611	-0.054501
	Days to death	-0.1366333	0.087835	-0.049679	-0.2177092
	Progesterone receptor status	-0.3059418	-0.0622482	-0.1747921	-0.1907064
	Oestrogen receptor status	-0.3130029	-0.0330698	-0.1957213	-0.2184403
Colon cancer	Age at initial pathologic diagnosis	-0.1489656	-0.1165508	-0.0520317	-0.0673751
	Days to death	-0.6524487	-0.4292221	-0.1357656	-0.201848
	Gender	-0.0565765	0.0280704	0.0183175	0.1254216
	Lymphatic invasion	0.0905995	0.2324266	0.1593187	0.1918415
Glioblastoma multiforme	Age at initial pathologic diagnosis	-0.0479766	0.1975299	-0.0687595	-0.0847677
	Days to death	-0.0221752	-0.0392693	-0.1181533	0.1625619
	Gender	0.0392475	-0.1515636	0.0316237	0.0809222
	Karnofsky score	-0.0240932	-0.1641532	-0.1609492	0.0845675
	Age at initial pathologic diagnosis	-0.0969832	-0.0075666	-0.2122888	-0.1564509
Kidney renal clear cell	Days to death	0.285603	-0.4353912	0.258888	0.1792746
	Gender	0.0669302	0.1071293	-0.1944817	-0.1526822
	Neoplasm histologic grade	0.0923653	-0.0490063	0.1733236	-0.1568577
	Tumour stage	0.2502406	-0.0438139	0.3775744	0.0744828
Kidney renal papillary cell	Age at initial pathologic diagnosis	0.2702908	0.0347612	0.2407317	0.2844299
	Gender	-0.0245357	0.3244777	-0.0046924	0.3275513
	Tumour stage	-0.1502799	-0.651768	-0.1491883	0.2629639
Low grade glioma	Age at initial pathologic diagnosis	0.1539625	0.3539578	-0.0663081	-0.1331918
	Days to death	-0.5352765	-0.5120105	0.4440297	0.2700757
	Gender	-0.1633157	-0.0822697	-0.0725039	0.0972517
Lung cancer	Age at initial	-0.1981139	-0.1729417	-0.1079096	0.3181156

	pathologic diagnosis				
	Days to death	0.8806854	0.4177005	0.8291734	0.8302324
	Gender	0.1953928	0.2961802	0.147857	-0.1378627
	Tumour stage	0.1929077	-0.0317282	-0.041544	0.0736851
Rectal cancer	Age at initial pathologic diagnosis	0.2024235	0.253082	0.0737415	-0.002
	Days to death	-0.5068186	-0.3988767	-0.5725386	-0.5331392
	Gender	0.0145101	-0.0816457	0.1637384	0.05204
	Lymphatic invasion	0.0452352	-0.0061905	-0.0081451	0.2315601
	Number of lymph nodes positive	0.0160364	-0.0874937	-0.0103846	0.0430844
	Venous invasion	0.062359	0.0264162	-0.0426753	-0.1056704
	Tumour stage	0.0903432	0.1866453	0.0796167	0.1110372
Uterine corpus endometrioid	Age at initial pathologic diagnosis	0.210155	-0.2779008	0.271219	0.2156729
	Days to death	0.3901652	0.606402	0.2232847	0.781775

These results indicate the complex relationship between EZH2 and EHMT2 expression and CNV and varying clinical characteristics across cancer types. However, high EZH2 expression only appears to significantly correlate with negative clinical characteristics. Similarly EHMT2 correlates with several negative outcomes but is related to a higher age at initial pathologic diagnosis in uterine corpus endometrioid cancer.

This study reinforces the case for dual HKMT inhibition in breast cancer and highlights the potential impact of dual HKMT inhibition in settings such as Colon cancer (where increased lymphatic invasion is linked to EHMT2 expression and CNV) and Kidney renal clear cell cancer, where EZH2 and EHMT2 expression both correlate with advanced tumour stages.

3.7 Target gene expression and survival

Cox proportional hazard modelling was performed on TCGA data to establish if expression of EHMT2, EZH2, and EZH2 related subunits correlates with survival (**Material and methods: Survival analysis utilising combined data sources**).

Expression of histone methyltransferases related to H3K9 methylation EHMT2, SUV39H1, and SUV39H2 were studied, as well as PRC2 subunits EED, EZH2, and SUZ12 (Probe IDs in Supplementary table 3.8).

Ovarian cancer, breast cancer, colon adenocarcinoma, glioblastoma multiforme, kidney renal clear cell, and rectal cancer were the cancer types with enough data to process for Cox proportional hazard modelling (number of cases in each cancer summarised in Table 3.8).

Table 3.8- Summary of data analysed for Cox proportional hazard modelling obtained from TCGA

Cancer type	Number of cases
Ovarian	487
Breast	484
Colon	164
Glioblastoma multiforme	162
Kidney renal clear cell	70
Rectal	69

For each probe, a hazard ratio and p value was calculated and these hazard ratios are tabulated in Table 3.9.

Table 3.9- Cox proportional hazard modelling of target probes in cancer data sets (Table 3.8)

- Hazard ratios with a significance of $p < 0.05$ are highlighted

Gene Symbol	Probe ID	Cancer Type					
		Ovarian	Breast	Colon	Glioblastoma multiforme	Kidney renal clear cell	Rectal
EED	AK026908_1_3458	0.799	0.00152	1.22	0.874	1.14	NA
EED	AK026908_1_3596	0.841	0.0341	1.71	1.1	0.932	NA
EHMT2	A_32_P122580	1.02	1.41	0.645	1.16	1.07	0.221
EHMT2	NKI_NM_004456	0.883	1.01	0.89	1.23	1.31	0.194
EHMT2	A_23_P422193	1.06	0.568	1.71	1.01	2.28	NA
EHMT2	A_23_P422195	1.09	0.626	1.76	1.07	2.08	NA
EZH2	NM_004456_3_2455	0.874	0.555	1.12	0.976	2.19	2.69
EZH2	NM_004456_3_2590	0.934	0.704	1.15	0.99	2.28	1.28
EZH2	A_23_P202392	0.962	2.26	0.936	1.04	0.855	1.14
EZH2	A_23_P202394	0.964	2.33	1.21	1.08	0.63	0.802
EZH2	A_32_P24223	1.33	0.832	1.05	0.975	2.3	2.55
EZH2	A_32_P4321	1.18	0.688	1.07	1.02	3.3	0.591
EZH2	A_32_P4324	1.26	0.556	1.14	0.972	2.12	1.9
JMJD4	A_23_P53216	1.19	1.65	0.751	1.09	1.88	0.625
JMJD4	A_23_P53217	1.17	1.39	0.808	1.11	1.86	0.143
JMJD4	A_24_P303389	0.853	2.99	3.18	1.25	2.21	11
JMJD4	A_24_P303390	0.838	3.34	1.35	1.31	1.95	14.2
SUV39H1	A_23_P115522	1.02	0.322	0.765	1.18	1.28	2.71
SUV39H1	A_23_P115523	0.988	0.316	0.573	1.06	1.96	5.83
SUV39H2	A_23_P259641	0.931	1.19	0.721	0.947	0.938	24.1
SUV39H2	A_23_P259643	0.874	0.35	1.56	0.692	0.369	51.9
SUV39H2	A_32_P122579	1.01	0.492	1.68	0.698	0.47	94.1
SUZ12	A_23_P214638	0.889	0.531	1.14	0.863	0.265	0.419
SUZ12	A_23_P214639	0.827	1.76	1.65	0.866	0.306	0.513
SUZ12	A_23_P202390	0.955	0.0495	1.23	0.969	2.76	6.53
SUZ12	A_23_P100883	1.34	0.63	1.26	1.08	3.38	3.37
SUZ12	A_23_P100885	1.26	1.99	1.29	1.21	3.51	2.09
SUZ12	A_24_P873263	0.987	2.41	2.03	0.836	0.511	0.132
SUZ12	A_32_P24215	1.16	0.855	1.41	0.956	0.533	3.64

No significant relationship between expression of these genes and survival was observed in Breast, Colon, or Rectal cancers. Glioblastoma multiforme showed a significant increase in

survival probability (Table 3.9) with the higher expression of two of three probes to the SUV39H2 gene.

In ovarian cancer, 2 of 7 SUZ12 probes significantly indicated a decrease in survival probability with higher expression. 1 of 7 EZH2 probes significantly indicated a decrease in survival probability with higher expression, and 1 of 2 EED probes significantly indicated an increase in survival probability with higher expression.

In Kidney renal clear cell cancer, 4 of 7 SUZ12 probes significantly related high expression with decreased survival probability and 5 of 7 EZH2 probes significantly related high expression with decreased survival probability, with hazard ratios consistently greater than 2.1 in significant probes.

The potential issues of low sample numbers and the relatively short follow up period in the currently available TCGA data indicates that these findings may not be indicative of all of the relationships present. In order to access greater patient numbers, the KMplot platform was accessed (**Materials and methods: Survival analysis utilising combined data sources**), allowing survival in ovarian and breast cancer to be assessed on a larger scale. Due to the low numbers of patients presenting certain clinical characteristics the data generated was of variable reliability (Table 3.10 shows the recommended reliability of different results based on the number of samples available for each clinical sub-grouping). As can be seen, RFS and PFS studies allow a greater reliability than most OS studies in this system.

RFS data in breast cancer will be highly reliable due to large patient numbers, but in terms of overall survival reliable analysis is only possible at a total cancer level for breast cancer. In ovarian cancer the only reliable analysis possible is for total ovarian serous and ovarian serous grade 3. The genes focused on in this study included HKMTs EZH2 and EHMT2 and PRC2 complex member EED.

Table 3.10- Suggested reliability of each study group in the KMplot platform based on the number of patients fitting clinical criteria for each sub-grouping

Study	OS	Reliability	RFS/PFS	Reliability
	# of patients		# of patients	
breast cancer (all)	1115	highly reliable analysis	3455	highly reliable analysis
breast cancer (ER negative)	140	preliminary analysis	668	highly reliable analysis
breast cancer (ER positive)	377	neutral	1767	highly reliable analysis
breast cancer (PR positive)	0	N/A	525	highly reliable analysis
breast cancer (PR negative)	0	N/A	481	reliable analysis
ovarian endometrioid	28	explorative analysis	28	explorative analysis
ovarian serous (all)	1058	highly reliable analysis	939	highly reliable analysis
ovarian serous (grade 3)	799	highly reliable analysis	696	highly reliable analysis
ovarian serous (grade 1)	27	explorative analysis	25	explorative analysis
ovarian serous (grade 2)	215	preliminary analysis	203	preliminary analysis

Table 3.11- Relationship between target gene expression and RFS in breast cancer patients, p values <0.05 highlighted yellow

	Gene	EZH2	EHMT2	EED
	Probe ID	203358_s_at	207484_s_at	209572_s_at
Breast cancer(all)	p-value	3.30E-16	1.30E-08	1.50E-08
	hazard ratio	1.83	0.69	1.43
Breast cancer (ER -)	p-value	0.2049	0.1186	0.2536
	hazard ratio	0.82	0.8	0.86
Breast cancer (ER +)	p-value	1.40E-07	0.0042	0.1111
	hazard ratio	1.73	0.76	0.86
Breast cancer (PR -)	p-value	0.0279	0.2173	0.0094
	hazard ratio	1.51	1.24	0.65
Breast cancer (PR +)	p-value	3.20E-08	0.0192	0.2429
	hazard ratio	2.69	0.65	0.8

EZH2 expression significantly relates with relapse free survival in total breast cancer, ER+ breast cancer, PR+ breast cancer, and PR- breast cancer, with high expression of EZH2 linked to earlier relapse (Table 3.11).

EHMT2 expression significantly relates with relapse free survival in total breast cancer, ER+ breast cancer, and PR+ breast cancer, but interestingly higher expression is linked to lengthier time to relapse.

High EED expression in total breast cancer and PR- breast cancer significantly relates to a reduced time to relapse and increased time to relapse relatively.

When OS is studied rather than RFS in breast cancer, there is not sufficient data to compute reliable analyses for PR+ and PR- cases- however, high expression of EZH2 is linked to a decreased probability of survival in total breast cancer and ER+ breast cancer (Table 3.12).

High expression of EED however showed a significant relationship with greater chance of overall survival in ER- and ER+ breast cancer.

Table 3.12- Relationship between target gene expression and OS in breast cancer patients, *p* values <0.05 highlighted yellow

	Gene	EZH2	EHMT2	EED
	Probe ID	203358_s_at	207484_s_at	209572_s_at
Breast cancer(all)	significance	1.90E-06	2.92E-01	2.11E-01
	hazard ratio	2.1	1.16	0.85
Breast cancer (ER -)	significance	0.0604	0.2311	0.0145
	hazard ratio	0.47	0.7	0.47
Breast cancer (ER +)	significance	2.40E-05	0.14	0.0057
	hazard ratio	2.49	1.4	0.54

At a total breast cancer level, it is clear that EZH2 expression and OS/RFS are related (Fig.3.10).

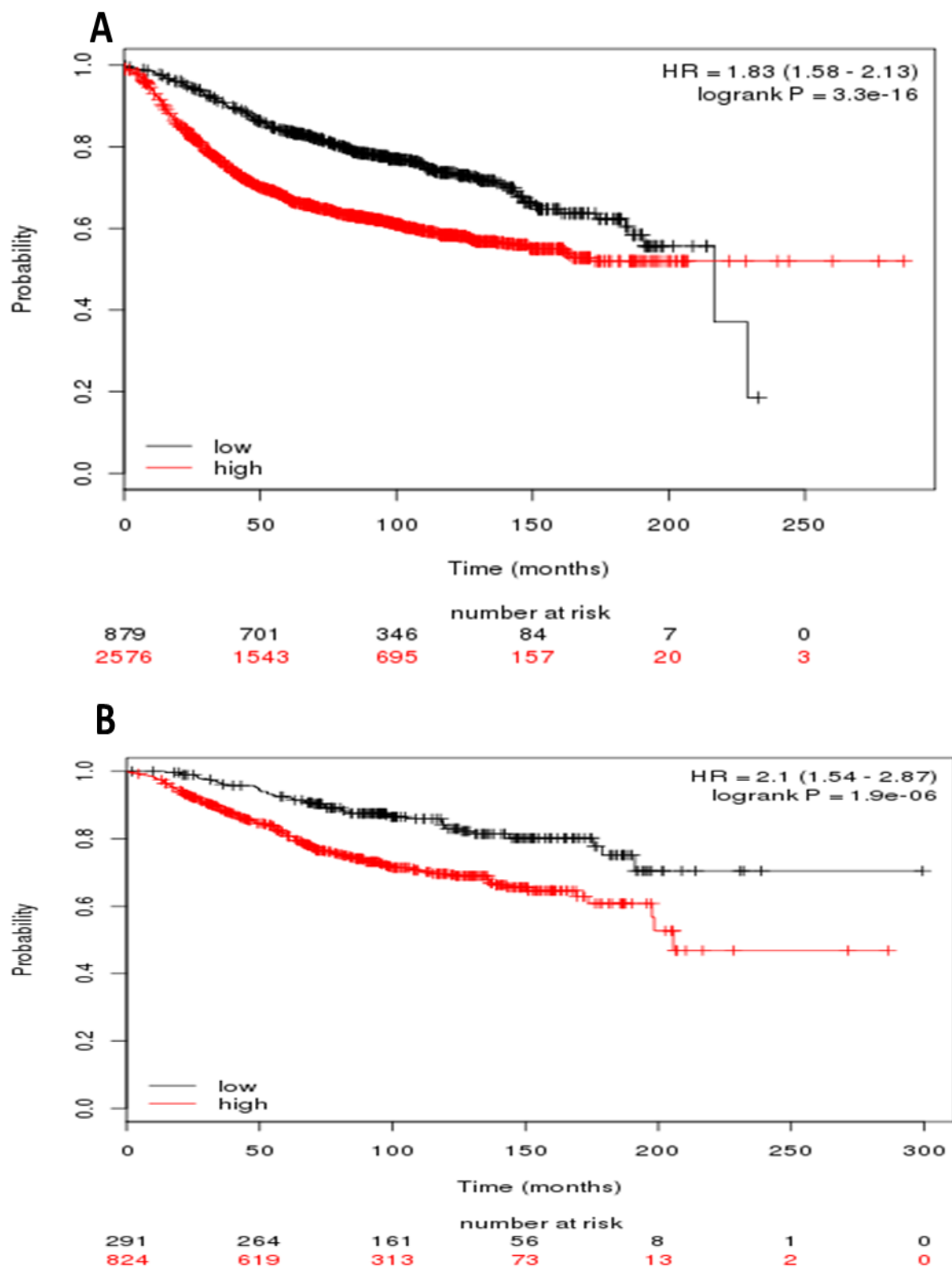


Figure 3.10- Kaplan-Meier plot of EZH2 expression split on the expression median (high expression in red) compared to A) Relapse free survival of 3455 breast cancer patients B) Overall survival of 1115 breast cancer patients

In ovarian cancer, low patient numbers mean PFS can only be calculated reliably in total ovarian serous adenocarcinoma and grade 3 ovarian serous adenocarcinoma (Table 3.13).

Table 3.13- Relationship between target gene expression and PFS in ovarian cancer patients, *p* values <0.05 highlighted yellow

	Gene	EZH2	EHMT2	EED
	Probe ID	203358_s_at	207484_s_at	209572_s_at
Ovarian serous (all)	Significance	0.0054	0.4113	0.0504
	hazard ratio	0.77	0.93	0.84
Ovarian serous (grade 3)	Significance	0.0005	0.1551	0.092
	hazard ratio	0.67	0.87	0.84

Interestingly, high expression of EZH2 appears to relate significantly to increased progression free survival in total and grade 3 ovarian serous adenocarcinoma. EHMT2 and EED expression did not significantly relate to PFS.

In terms of OS (Table 3.14), higher expression of EHMT2 was slightly significantly linked to higher survival probability in grade 3 patients, but was not significant when all grades are included. High expression of EED related to a lower survival probability (HR 1.22) at a total cancer level.

Table 3.14- Relationship between target gene expression and OS in ovarian cancer patients, *p* values <0.05 highlighted yellow

	Gene	EZH2	EHMT2	EED
	Probe ID	203358_s_at	207484_s_at	209572_s_at
Ovarian serous (all)	Significance	0.0864	0.0592	0.0195
	hazard ratio	1.16	0.85	1.22
Ovarian serous (grade 3)	Significance	0.111	0.0123	0.0648
	hazard ratio	0.84	0.78	1.2

3.8 Summary

EZH2, EHMT2, and related subunits show a diverse range of expressions in normal human tissue (3.1) and the correlation of expression of these genes varies in direction and strength depending on the tissue studied, indicating that tissue type could play a factor in determining response to dual HKMT inhibition if expression levels are important for biological effect of the inhibitors. High expression of both EZH2 and EHMT2 in tissue types such as the immune system, haematopoietic system, and CNS, indicate these tissues may be particularly dependent on these HKMTs and may react strongly to intervention with dual HKMT inhibitors. This can highlight potential tissues where dual HKMT inhibition may have an impact outside of the intended therapeutic target.

EZH2 and EHMT2 are highly upregulated in a number of cancer tissues (3.3) compared to normal tissues, and specifically they both show up-regulation of expression in Adrenal, Brain, Lung, and Prostate cancers. This data also indicates the need to examine the expression profiles of these targets in different clinical sub-types, as expression seems to vary between different cancer sub-types where data is available.

The frequency and location of recorded mutation of EZH2 and EHMT2 (3.4) indicate that whilst in some cases (like follicular lymphoma) well characterised mutations may help stratify patients for HKMT inhibition, in the majority of solid cancers EZH2 and EHMT2 mutation is not overly common and known pathogenic drivers are very uncommon.

To investigate if EZH2/EHMT2 CNV was a better indicator of potential receptivity to dual HKMT inhibition, the frequency of these CNVs was studied (3.5) and their relation to gene expression and clinical outcomes was characterised (3.6). CNV of EZH2 does not always appear to correlate with expression of EZH2, meaning that at a cancer level it is unsuitable to

stratify patients, but may still have potential as a tool in cancers where a strong relationship between CNV and expression was observed (such as ovarian cancer and breast cancer).

Investigating the correlation with expression, CNV, and clinical characteristics (3.6) has highlighted potential novel cancer types that may benefit from dual HKMT inhibition such as Colon cancer and Kidney renal clear cell cancer.

Utilising public data (3.7) again emphasised Kidney renal clear cell cancer as a potential future target for dual HKMT inhibition. Finally, large scale analysis of combined datasets showed that EZH2 is strongly linked to RFS and OS in breast cancer, reinforcing previous findings that EZH2 is linked to aggressive phenotypes in this disease setting.

In summation, multiple cancer types show negative clinical characteristics and outcomes to be linked to expression of EZH2/EHMT2, and reversal of epigenetically mediated gene silencing may prove therapeutically beneficial. This mechanism appears to be aberrantly regulated in multiple cancer types, and whilst differences between cancer types and sub-types may alter efficacy of treatment, targeted intervention with dual HKMT inhibitors has the potential to bring about significant clinical impact if this inhibition impacts on the cancer phenotypes observed. It is clear that expression of EZH2 and EHMT2 strongly positively correlate in numerous settings, further reinforcing the concept of their shared roles and potential redundancy.

Chapter 4: Impact of novel dual HKMT inhibitors on the epigenetic state of cancer cells

4.1 Introduction and Aims

As part of the PRC2 complex, EZH2 catalyses the addition of methyl groups to H3K27¹⁹ and the resulting H3K27me3 leads to chromatin condensation and a reduction in gene expression by recruitment of PRC1²¹ as detailed in Chapter 1 (Section 1.2, summarised in Fig. 1.1). Chapter 3 indicated that whilst mutations in EZH2 play a role in the pathology of specific diseases (e.g. follicular lymphoma¹³⁵).

Large scale analysis of combined datasets (3.7) showed that EZH2 is strongly linked to RFS and OS in breast cancer, and previously reported findings show high levels of EZH2 expression²⁸ in breast cancer, with high levels of EZH2 acting as markers of aggressive breast cancer^{29–31}, and expression of EZH2 associated with the often difficult to treat triple negative/basal phenotypes³². High EZH2 expression is linked to poor RFS and OS in breast cancer (Chapter 3), and combined with the published literature highlight breast cancer as a potential solid tumour target for dual HKMT inhibition of EZH2 mediated silencing, and as such the impact of novel dual HKMT inhibitors (HKMT-I-005, HKMT-I-011, and HKMT-I-022) on gene expression was studied using MDA-MB-231 triple negative breast cancer cells, which have been characterised as having relatively high EZH2 expression¹²⁷ and siRNA knockdown of EZH2 expression has been shown to reduce motility and block invasion in breast cancer cell lines¹³⁶.

The impact of novel dual HKMT inhibitors on gene expression was studied using gene expression microarray platforms with several goals in mind:

To establish if dual HKMT inhibitors induce up-regulation of expression of genes known to be silenced by EZH2 in MDA-MB-231 breast cancer cells, a list of genes showing significantly altered expression levels following siRNA mediated reduction in EZH2 levels was obtained from ¹¹³. Enrichment analysis of genes known to be EZH2 targets in MDA-MB-231 cells was performed after 24 hours or 48 hours of treatment with dual HKMT inhibitors as well as known specific inhibitors of EZH2 or EHMT2- these time points were chosen to allow any impact on chromatin state to have led to alterations in gene expression. In addition, this enrichment analysis was performed using EZH2 targets derived from another breast cancer line, MCF-7 (targets known to have significantly altered gene expression following siRNA mediated reduction in EZH2 levels ¹¹⁴), to establish if the compounds impact on expression of target genes differs greatly between target genes generated in different cell lines. Enrichment analysis was also performed on a list of EZH2 target genes generated by a meta-analysis of multiple studies in which EZH2 expression was artificially lowered- here, genes that consistently showed altered expression after siRNA/shRNA mediated EZH2 reduction in multiple cell lines were found (meta-analysis target list generated by MRes student **Emma Bell (Materials & Methods: Enrichment analysis)**)).

Having established the impact of novel HKMT inhibitors on EZH2 target gene expression, the differences and similarities in the pattern of genes whose expression was affected by treatment with these novel inhibitors and known EZH2 and EHMT2 inhibitors was examined, as well as the degree of similarity between novel dual HKMT inhibitors that passed the selection screen (Chapter 1, 1.5) and examples from the chemical library that did not pass.

The changes in gene expression (outside of the EZH2 targets described above) caused by treatment with HKMT inhibitors was studied through analysis of functional annotation enrichments, to highlight pathways showing altered expression after treatment.

In an effort to refine the initial compound selection process (Chapter 1, 1.5) analysis was performed to select potential pharmacodynamic markers of response to dual HKMT inhibitors that may be used either in the compound selection process, or in future downstream studies such as response of tumour cell *in vivo* to compound treatment. The levels of repressive chromatin marks H3K27me3 and H3K9me3 at potential biomarker SPINK1 were examined in parallel with known EZH2 silenced genes FBXO32 and KRT17.

4.2 Impact of dual HKMT inhibitors on EZH2 target gene expression

Two gene expression microarrays were performed after treatment with HKMT inhibitors and putative dual HKMT inhibitors (Table 4.1). MDA-MB-231 cells were treated with inhibitors and RNA was harvested- an initial array was performed, followed by a second validation array to replicate the key findings (**Materials & Methods: Gene expression microarray**). Both of these arrays included 24 hour and 48 hour time points for sample collection post-treatment.

These time points were chosen as the initial compound screen showed impact of these inhibitors on cell proliferation of MDA-MB-231 cells at 48 hours, as well as up-regulation of target genes FBXO32 and KRT17. This indicated that by 48 hours these drugs were impacting gene expression. The 24 hour time point was included to investigate the progression of gene expression alteration, to investigate if these changes at expression occurred at this earlier time and if they were different than the changes in expression seen at 48 hours. Doses were based upon published data for compounds UNC0638 and GSK343 (dose equivalent or higher than published IC₅₀ of EHMT2 and EZH2 respectively), and the doses of the hit compounds were based on their IC₅₀ in the initial compound screen (detailed in Table 4.2).

Table 4.1- Summary of inhibitors utilised in microarray analysis

Inhibitor	Description	Reference
HKMT-I-005	Dual EZH2/EHMT2 inhibitor	Developed in-house
HKMT-I-011	Dual EZH2/EHMT2 inhibitor	Developed in-house
HKMT-I-022	Dual EZH2/EHMT2 inhibitor	Developed in-house
TG3-259-1	Potential dual inhibitor (failed compound selection due to lack of up-regulation of EZH2 target genes FBXO32/KRT17)	Developed in-house
TG3-184-1	Potential dual inhibitor (failed compound selection due to lack of up-regulation of EZH2 target genes FBXO32/KRT17)	Developed in-house
UNC0638	EHMT2 specific inhibitor	Vedadi et al. 2011
GSK343	EZH2 specific inhibitor	Verma et al. 2012

The array included treatments by hit dual HKMT inhibitors, specific EZH2/EHMT2 inhibitors, and a compound from the chemical library that failed the selection screen (Chapter 1, 1.5.1).

The validation array included the two dual HKMT inhibitors which showed greatest up-regulation of EZH2 targets in the initial array, and another compound that failed the selection screen- summations of treatments, doses, and time points of each array detailed in Table 4.2.

Doses were selected based on published data for GSK343 and UNC0638, and based upon performance in the initial compound selection screen for the dual HKMT inhibitors.

Table 4.2- Microarray study design

Array	Replicates	Drug	Dose(s) (μM)	Time points (hours after treatment)
Initial Array	n=3	HKMT-I-005	2.5, 7.5	24, 48
		HKMT-I-011	2.5	24, 48
		HKMT-I-022	2.5	24, 48
		TG3-259-1	2.5	24, 48
		UNC0638	2.5, 7.5	24, 48
		GSK343	2.5	24, 48
Validation Array	n=4	HKMT-I-005	7.5	24, 48
		HKMT-I-011	2.5	24, 48
		TG3-184-1	5	24, 48

Statistical significance of differential expression induced by drug treatments was estimated (**Materials & Methods: Gene expression microarray**) for each gene expression probe, at each treatment and time point. The statistical significance of the systematic shift towards induced transcriptional upregulation or downregulation of the list of known EZH2 targets was established using enrichment analysis (**Materials & Methods: Enrichment analysis**). The calculated significance of enrichment of EZH2 target genes (from MDA-MB-231 cells target gene list- Supplementary table 8.9) after inhibitor treatment in MDA-MB-231 cells shows significant up-regulation of EZH2 silenced genes (significance- Supplementary table 8.5). Statistical significance of specific, systematic up-regulation of the EZH2 silenced genes after 24 hours treatment with dual HKMT inhibitors is shown (Fig.4.1 A (Enrichment p-values are plotted as inverse $\log_{10}$ values, where a p-value of 0.001 would be equal to 3 on the y-axis- any inverse $\log_{10}$ p-value >3 is very statistically significant))- this result was validated on both arrays for HKMT-I-005 and HKMT-I-011.

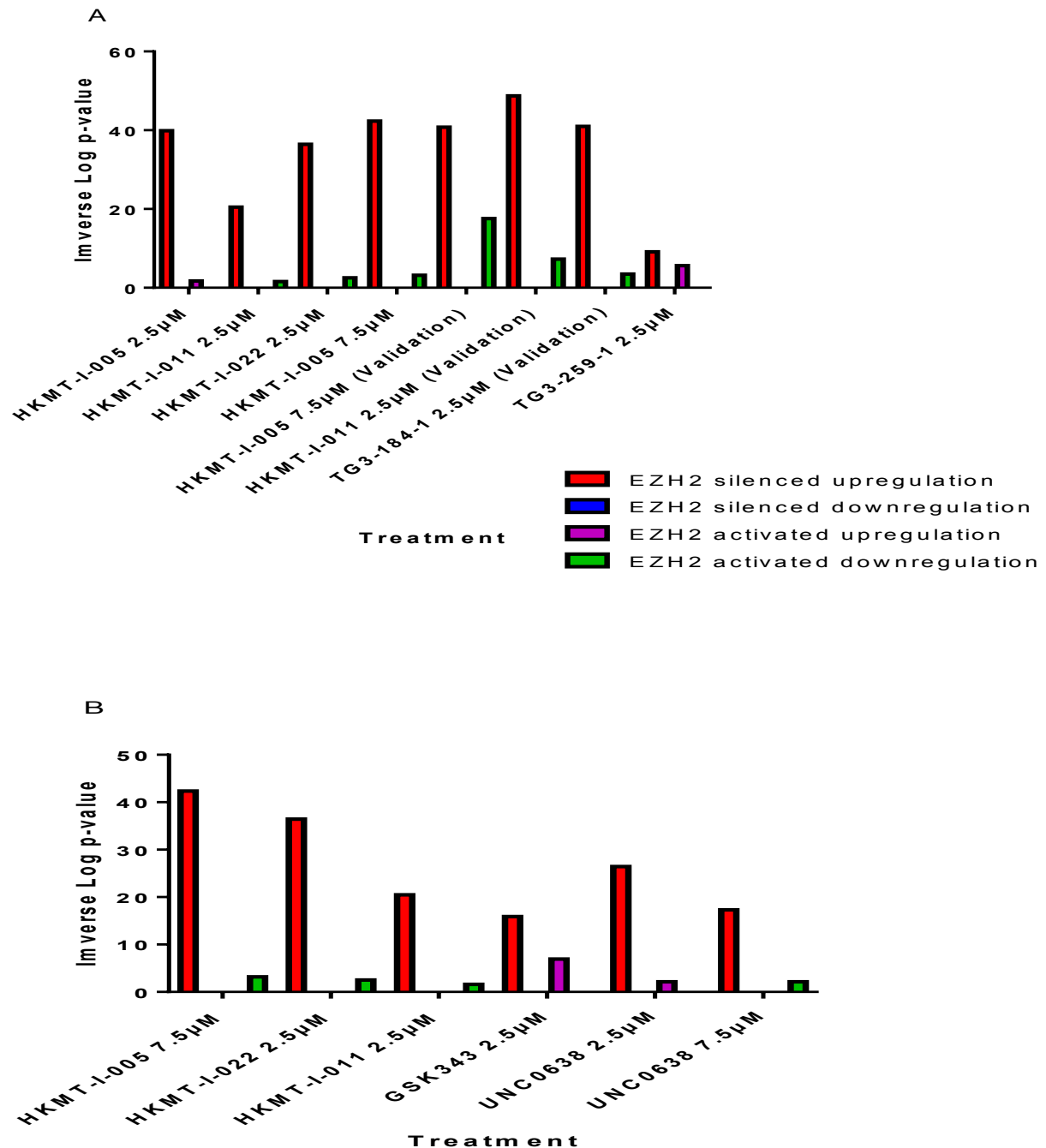


Figure 4.1- Enrichment of MDA-MB-231 EZH2 targets after 24 hour treatment with A) dual HKMT (including validation array results) B) dual HKMT and specific EZH2 inhibitor GSK343 and specific EHMT2 inhibitor UNC0638- enrichment p-values are plotted as inverse log₁₀ values, where a p-value of 0.001 would be equal to 3 on the y-axis

HKMT-I-005 showed a P-value of $p=4.53\text{E-}43$ for up-regulation of EZH2 silenced genes with a dose of $7.5\mu\text{M}$ for 24 hours, and HKMT-I-011 had a P-Value of $p=3.27\text{E-}21$. HKMT-I-005 showed similar impact at $2.5\mu\text{M}$ or $7.5\mu\text{M}$ treatment, but HKMT-I-011 showed a lesser (though still significant) upregulation of EZH2 silenced genes at the lower dose of $2.5\mu\text{M}$.

HKMT-I-011 and HKMT-I-005 also showed some capacity to down regulate expression of EZH2 activated targets (genes that showed significant decrease in expression after reduction of EZH2 levels by siRNA knockdown in MDA-MB-231 cells ¹¹³), though this was not replicated as strongly in both microarrays.

Interestingly, the compounds that failed the compound selection screen also showed up-regulation of EZH2 silenced targets. TG3-259-1 showed a significant up-regulation of EZH2 silenced genes ($p=7.20\text{E-}10$), which is highly significant (though substantially lower than hit compounds HKMT-I-005, HKMT-I-011, and HKMT-I-022). TG3-184-1 also showed a significant up-regulation of EZH2 silenced genes ($p=1.12\text{E-}41$) to a comparable level as hit compounds HKMT-I-005, HKMT-I-011, and HKMT-I-022. This highlights that the compound selection screen as stands (Chapter 1, 1.5.1) may be missing compounds that *in vitro* could have a significant impact- this is potentially due to the fact only two EZH2 target genes are being used in this screen, and at this preliminary stage understanding of the pharmacodynamics is not clear enough to know if these two genes will both be consistently, stably upregulated expression at the time point used in the compound screen (24 hour treatment). Development of further biomarkers and further time courses may allow adaptation of the existing screen to a more suitable form.

Treatment with the specific inhibitors of EZH2 and EHMT2 GSK343 and UNC0638 (respectively) resulted in (Fig. 4.2) significant specific, systematic up-regulation of MDA-MB-231 EZH2 silenced genes (GSK3434 $p= 1.28\text{E-}16$, UNC0638 $p= 3.68\text{E-}27$) – this up-regulation

of EZH2 target genes observed after treatment using the specific EHMT2 inhibitor supports the theory that EHMT2 plays a supporting role (via H3K9me1 and direct physical interaction with EZH2) in EZH2 repression, and that targeting EHMT2 will affect EZH2 mediated repression.

This firstly shows that dual HKMT HKMT-I-005, HKMT-I-011, and HKMT-I-022 all show a more significant up-regulation of MDA-MB-231 EZH2 silenced genes than specific EZH2 inhibitor GSK343 or the EHMT2 inhibitor UNC0638 at this time point and doses. Indeed, further analysis of the difference in systematic upregulation at 24 hours (based on the difference between the Wilcoxon Rank-Sum statistics across the target genes, for each treatment, **performed by Ed Curry**) showed that HKMTI-1-005 upregulated EZH2 silenced genes significantly more than either GSK343 ($p=5.8E-5$) or UNC0638, ($p=1.7E-4$).

UNC0638 is reported as having an inhibition IC50 for EZH2as $>10\mu\text{M}$ ⁹⁶, indicating that this up-regulation of EZH2 silenced genes may be through to the action of EHMT2 inhibition, further supporting the theorised overlap in targets of these HKMT- though UNC0638 could also be inhibiting EZH2, though to a much smaller degree than that of the IC50.

The clear portrait of EZH2 targets being impacted by HKMT inhibition is complicated slightly at the 48 hour time point (p-values shown in Supplementary table 8.5). At 48 hours, in the initial array there was no significant up-regulation of EZH2 silenced genes (Fig. 4.2 A). In the validation array however, a similar pattern is seen as in the 24 hour time point.

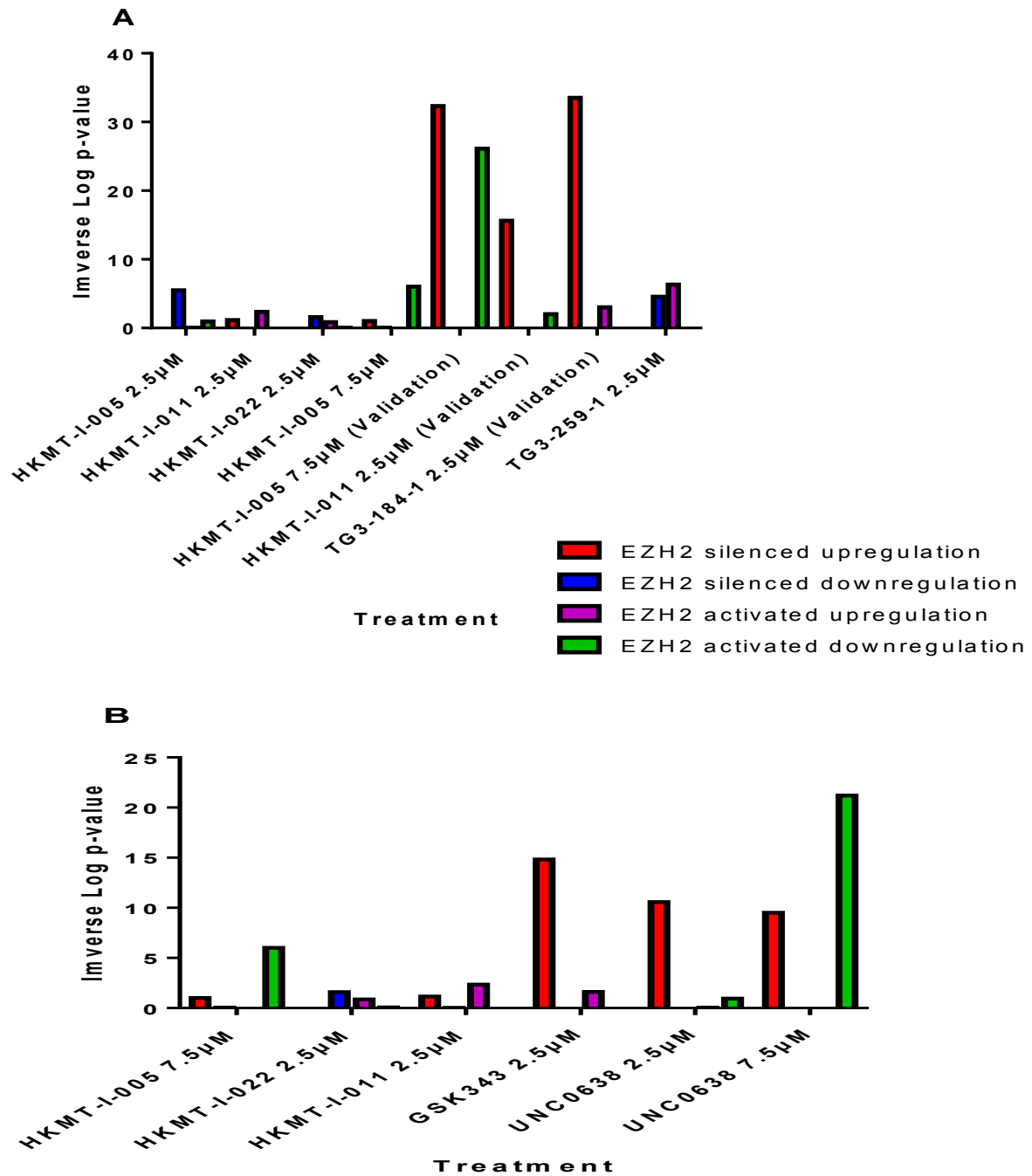


Figure 4.2- Enrichment of MDA-MB-231 EZH2 targets after 48 hour treatment with A) dual HKMT (including validation array results) B) dual HKMT and specific EZH2 inhibitor GSK343 and specific EHMT2 inhibitor UNC0638- enrichment p-values are plotted as inverse log10 values, where a p-value of 0.001 would be equal to 3 on the y-axis

In comparison, GSK343 and UNC0638 both showed (Fig. 4.2 B) significant up-regulation of MDA-MB-231 EZH2 target genes after 48 hour treatment (GSK343 $p=1.48E-15$, UNC0638 $p=2.65E-11$ at $2.5\mu\text{M}$ and $3.07E-10$ at $7.5\mu\text{M}$).

It is unclear as to why this is this case, though it may be to the lower toxicity of these treatments relative to the dual HKMT inhibitors. In an effort to investigate the similarity between EZH2 targets between cell types, the above analysis was repeated using an alternate gene list of MCF-7 EZH2 targets (**Materials & Methods: Enrichment analysis**), though notably only EZH2 silenced targets were available for use from this cell line.

At 24 hours, when examining up-regulation of MCF-7 derived EZH2 targets in MDA-MB-231 cells that have been treated with the varying HKMT inhibitors (Fig.4.3), HKMT-I-005 in the first microarray shows a marginally significant up-regulation ($p=0.025$) but this was not seen in the validation arrays. UNC0638 also induced a significant up-regulation of the MCF-7 EZH2 targets at this time point at the dose of $2.5\mu\text{M}$ ($p=0.038$), though considerably less so than that seen using the MDA-MB-231 derived EZH2 target list.

After 48 hours treatment (Fig.4.4), the only significant up-regulation of MCF7 EZH2 silenced genes in the MDA-MB-231 cells was by GSK343, the EZH2 specific inhibitor ($p=0.008$).

These results indicate that the targets of EZH2 differ between cell types, raising the issue of developing cancer or cancer sub-type specific biomarkers in order- biomarkers derived from studies in different cell types may not always be applicable.

Based upon these results, a meta-analysis was performed by MRes student **Emma Bell** to identify consensus target genes based on 18 independent EZH2 siRNA studies (details of meta-analysis: **Material & Methods: Enrichment analysis**). This meta-analysis provided a list of consistently EZH2 silenced and EZH2 activated genes across multiple tissue types.

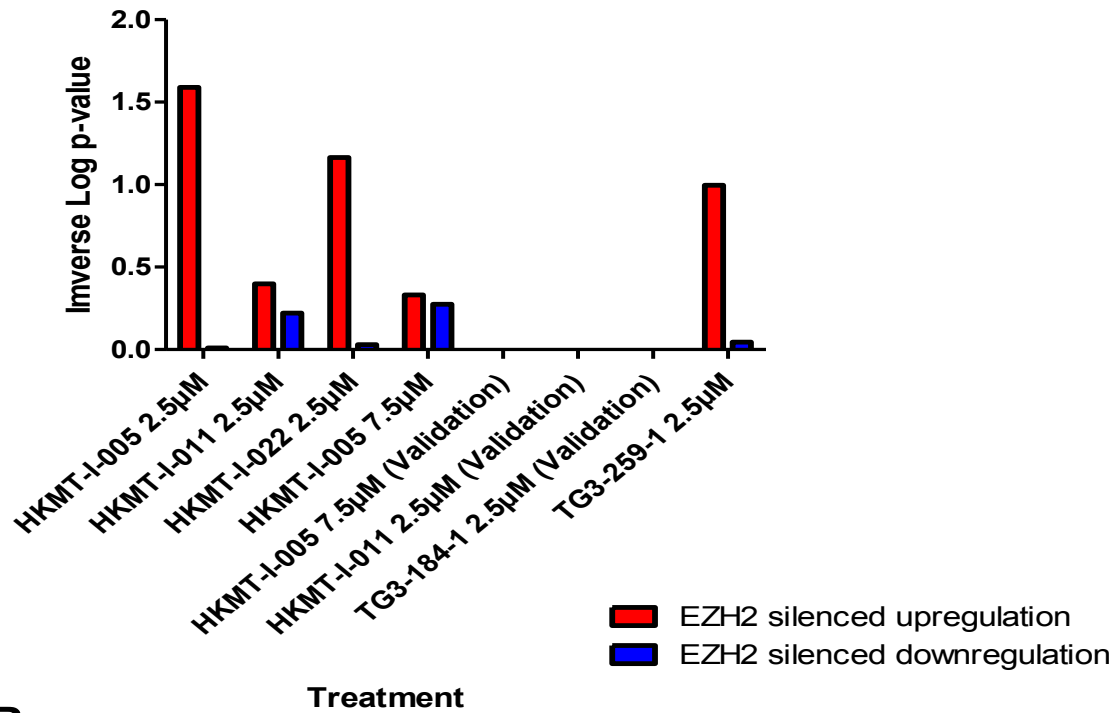
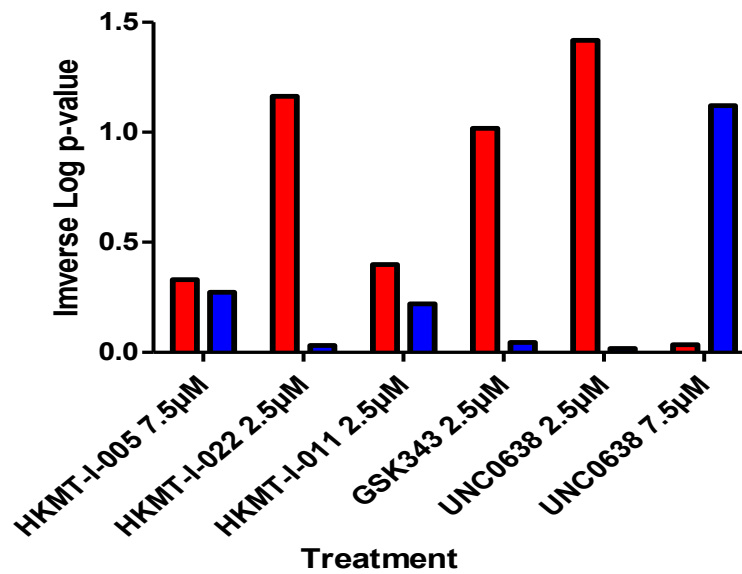
A**B**

Figure 4.3- Enrichment of MCF-7 EZH2targets in MDA-MB-231 cells after 24 hour treatment with A) dual HKMT inhibitors B) EZH2/EHMT2 specific inhibitors- enrichment p-values are plotted as inverse log₁₀ values, where a p-value of 0.001 would be equal to 3 on the y-axis

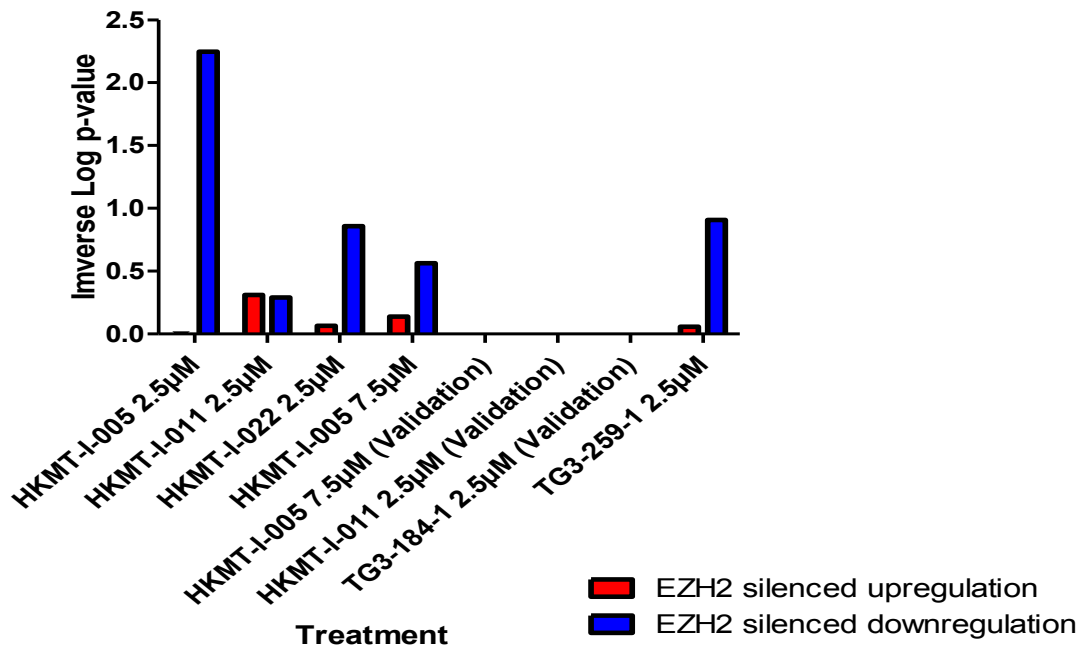
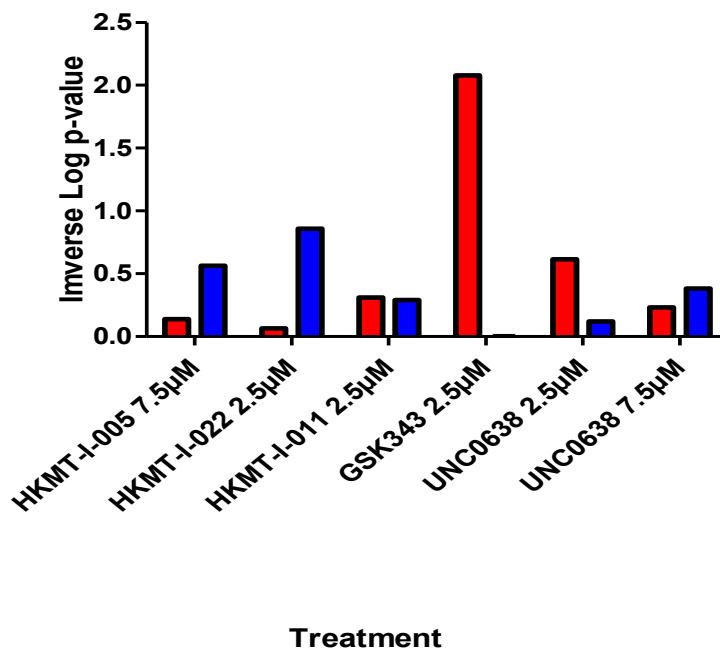
A**B**

Figure 4.4- Enrichment of MCF-7 EZH2 targets in MDA-MB-231 cells after 48 hour treatment with A) dual HKMT inhibitors B) EZH2/EHMT2 specific inhibitors- enrichment p-values are plotted as inverse log10 values, where a p-value of 0.001 would be equal to 3 on the y-axis

The impact of the dual HKMT inhibitors, compounds that failed the chemical screen, and specific EZH2 and EHMT2 inhibitors were analysed for the enrichment of upregulation of EZH2 silenced genes based upon the meta-analysis target list.

Encouragingly, at 24 hours the data appears to follow the pattern observed when this analysis was performed using the MDA-MB-231 data, with a great degree of significant upregulation of EZH2 silenced genes (Supplementary table 8.8).

At 24 hours (Fig.4.5), HKMT-I-005, HKMT-I-011, and HKMT-I-022 all showed very significant upregulation (Fig. 4.5 A) of the meta-analysis defined EZH2 silenced genes ($p=3.65E-26$, $p=1.18E-28$, $p=3.87E-23$ respectively).

EZH2 specific inhibitor GSK343 also showed (Fig.4.5 B) very significant upregulation ($p=1.79E-16$) of EZH2 meta-analysis defined repressed genes, as did EHMT2 inhibitor UNC0638 ($p=2.37E-29$).

TG3-184-1, which failed the original compound screening due to insufficient activation of specific EZH2 target genes FBXO32 and KRT17, showed significant impact upon of EZH2 target genes as defined by the meta analysis, indicating that some potent inhibitors may be falling through the screen due to lack of appropriate biomarkers.

This strong-upregulation of EZH2 target genes was also seen at 48 hours- interestingly, whilst the MDA-MB-231 derived EZH2 target genes showed little up-regulation of silenced genes at 48 hours with treatment from the dual HKMT, the meta analysis derived list of EZH2 silenced genes were significantly up-regulated at the 48 hour time point after some treatments- notably, HKMT-I-005 at a dose of $7.5\mu\text{M}$ ($p=3.85E-28$) and HKMT-I-011 at $2.5\mu\text{M}$ ($p=2.43E-15$).

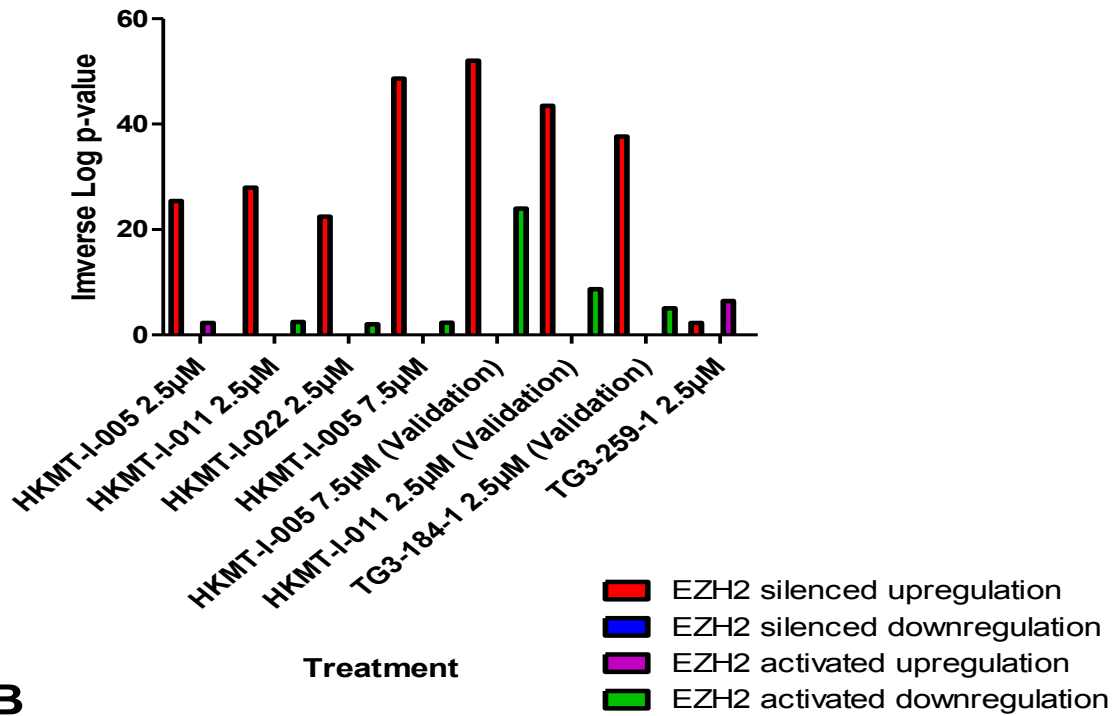
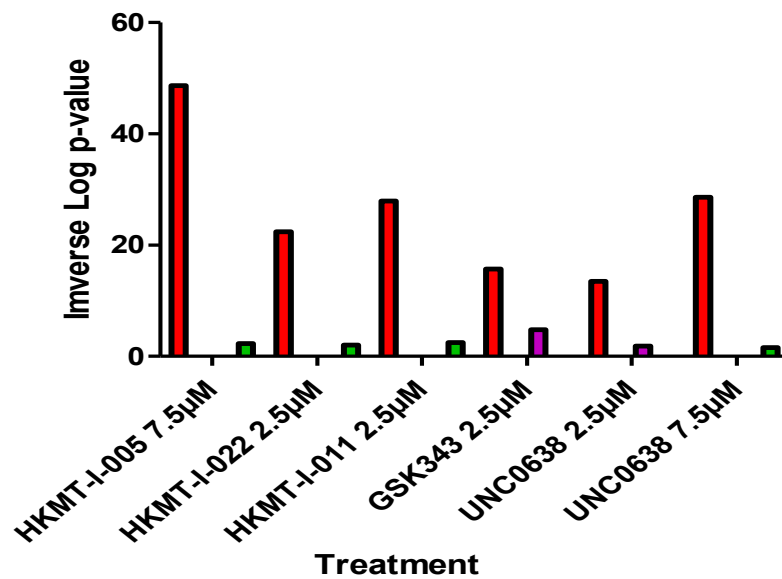
A**B**

Figure 4.5- Enrichment of meta-analysis EZH2 targets in MDA-MB-231 cells after 24 hour treatment with A) dual HKMT inhibitors B) EZH2/EHMT2 specific inhibitors- enrichment p-values are plotted as inverse log₁₀ values, where a p-value of 0.001 would be equal to 3 on the y-axis

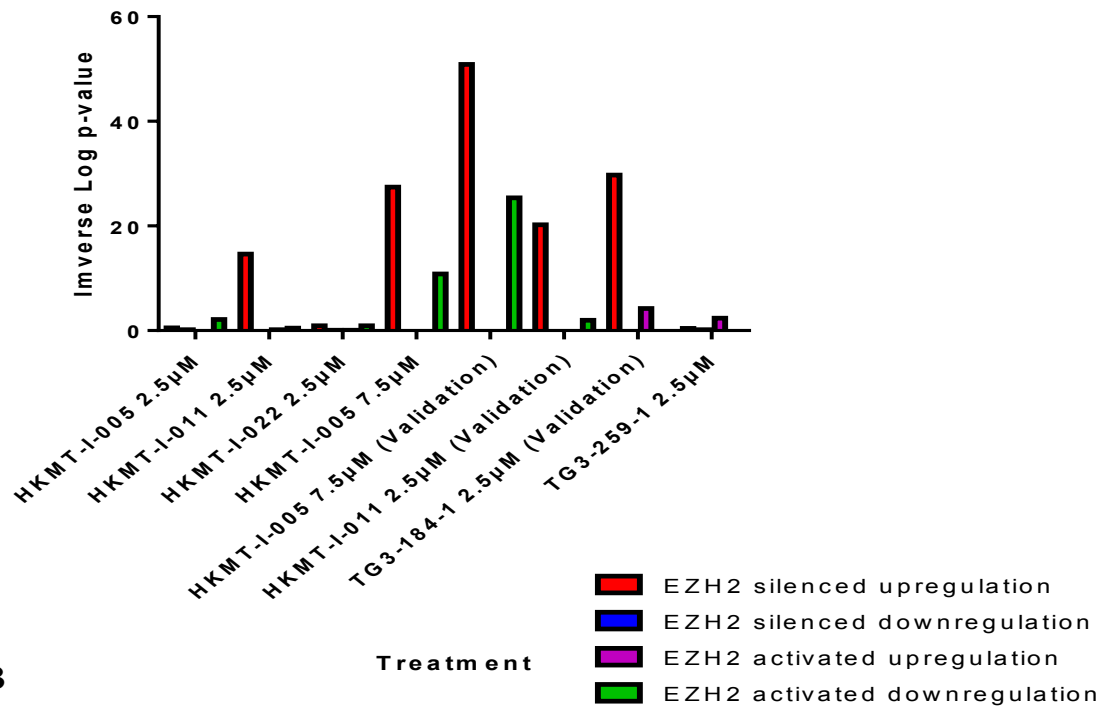
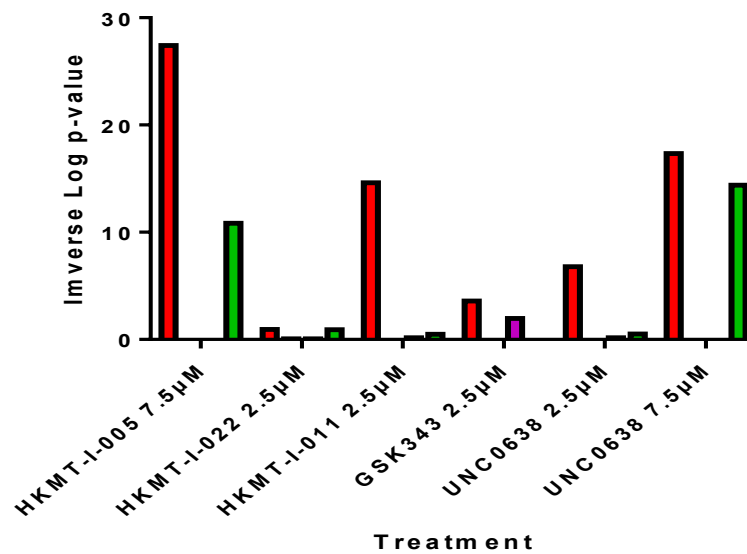
A**B**

Figure 4.6- Enrichment of meta-analysis EZH2targets in MDA-MB-231 cells after 48 hour treatment with A) dual HKMT inhibitors B) EZH2/EHMT2 specific inhibitors- enrichment p-values are plotted as inverse log10 values, where a p-value of 0.001 would be equal to 3 on the y-axis

This result is encouraging; as the meta-analysis represents a panel of genes that are consistently affected by EZH2 across numerous cell types, it should be more reliable than the MDA-MB-231 EZH2 targets which are derived from a single study.

The dual HKMT inhibitors HKMT-I-005 and HKMT-I-011 are capable of strongly upregulating the expression of EZH2 silenced genes in the MDA-MB-231 cells, to a significantly greater degree than EZH2 specific inhibitor GSK343 or EHMT2 inhibitor UNC0638- HKMT-I-022 is also a capable inhibitor of EZH2, though appear to be less potent in its action.

TG3-184-1 also seems to be capable of inducing expression of EZH2 target genes, highlighting the need for refinement within the chemical screen so potentially potent compounds are not bypassed.

4.3 Comparison of inhibitors' impact on gene expression

Having established that the dual HKMT inhibitors are capable of reversing EZH2 mediated gene silencing, the relative similarity of these inhibitors will be studied from the perspective of the gene expression changes observed following treatment. HKMT-I-005, HKMT-I-011, and HKMT-I-022 are to be compared to a known EZH2 inhibitor (GSK343) and a known EHMT2 inhibitor (UNC0638) as well as compounds that failed the chemical screen (TG3-259-1 and TG3-184-1) at an array wide and EZH2 target specific level in the hope of establishing potential commonalities.

Utilising the array data generated (**Materials & Methods: Gene expression microarray**, detailed in Chapter 4, 4.2) Correlation heatmaps were generated (**Materials & Methods: Correlation of gene expression after compound treatment**) comparing the genome-wide transcriptional effects of each treatment, at a whole-array level and utilising the target gene lists that were used previously (Supplementary Table 8.9)- these heatmaps are based on pair-wise

Pearson correlation coefficients, where 1= perfect correlation (shown here as red) and 0= no correlation (shown here in blue)- colour keys are shown for each heatmaps provided, as are column-wise dendrograms based upon complete unsupervised hierarchical clustering.

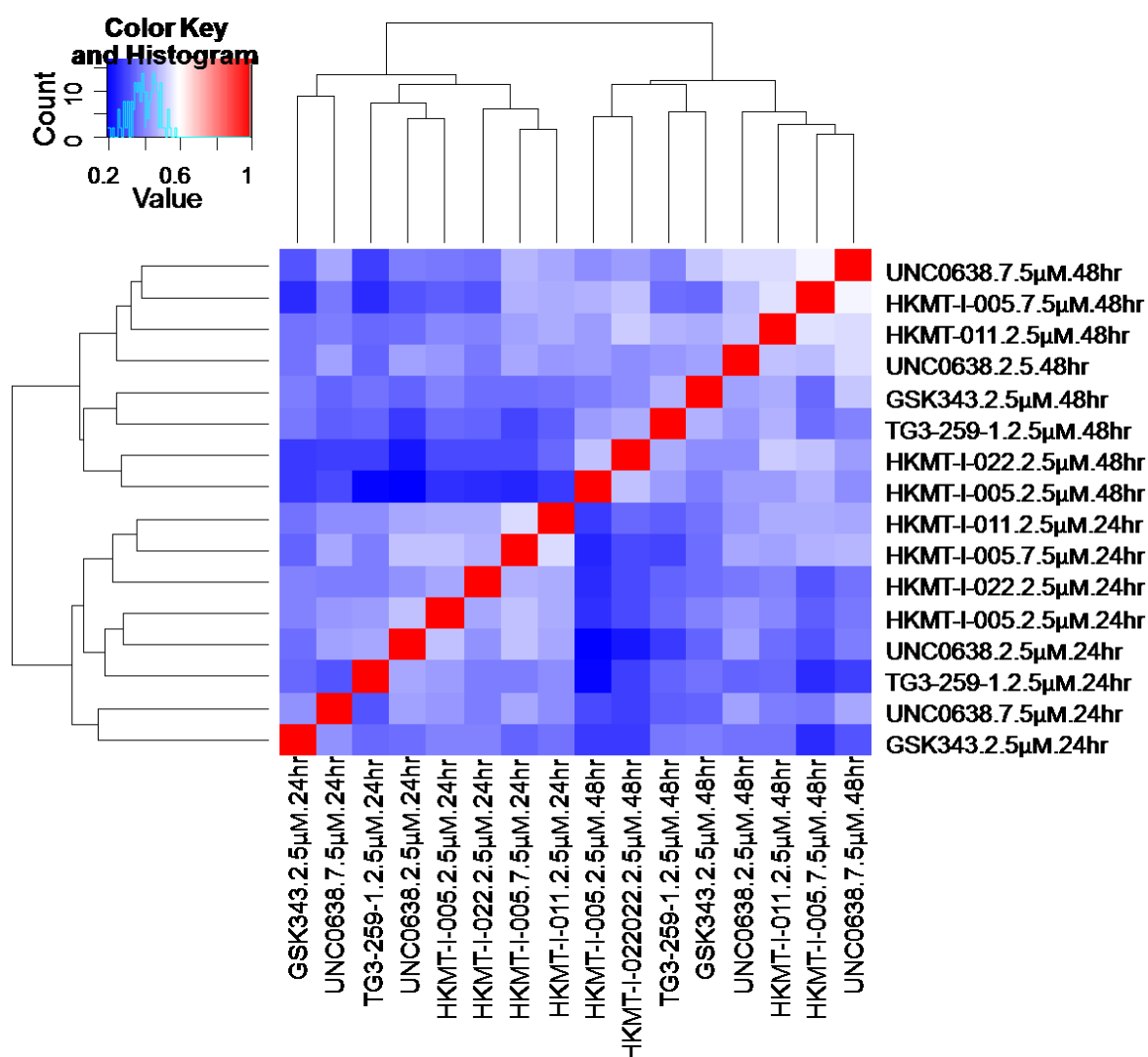


Figure 4.7- Correlation heatmap of gene expression in MDA-MB-231 after treatment with HKMT inhibitors at an array wide level

At an array wide level (Fig.4.7), there appear to be no strong correlations between treatments, and though there is a degree of clustering between the 24 hour and 48 hour samples, it does not indicate a strong separation- this could possibly be a batch effect, or a perhaps consistent later-onset effects of all the compounds.

When this analysis is performed using only MDA-MB-231 EZH2 silenced genes (as described in **Materials & Methods: Enrichment analysis**) a different pattern emerges (Fig. 4.8). Here, two primary clusters are seen- the first contains TG3-259-1 and GSK343, and the second contains UNC0638, HKMT-I-005, HKMT-I-022, and HKMT-I-011. When this is related back to the enrichment analysis performed on these target genes (Fig 4.1-4), it is clear that on the whole the inhibitors classed in this second cluster were those that induced the greatest reversal of EZH2 mediated silencing.

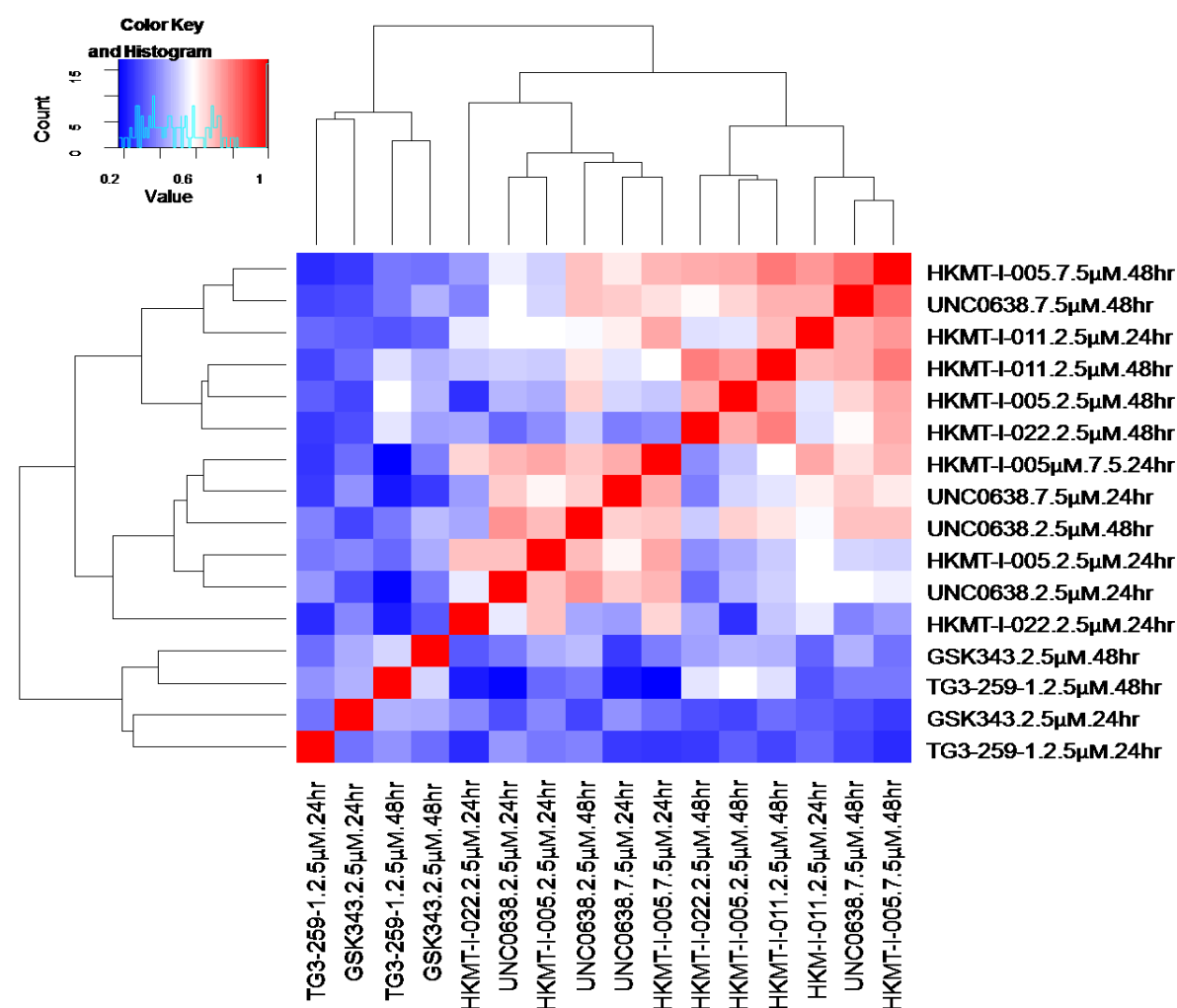


Figure 4.8- Correlation heatmap of gene expression in MDA-MB-231 after treatment with HKMT inhibitors for MDA-MB-231 EZH2 silenced genes

This analysis was repeated using the validation array (Fig.4.9 A), showing no strong correlations between treatments at an array wide level (though HKMT-I-005 and HKMT-I-011 cluster together).

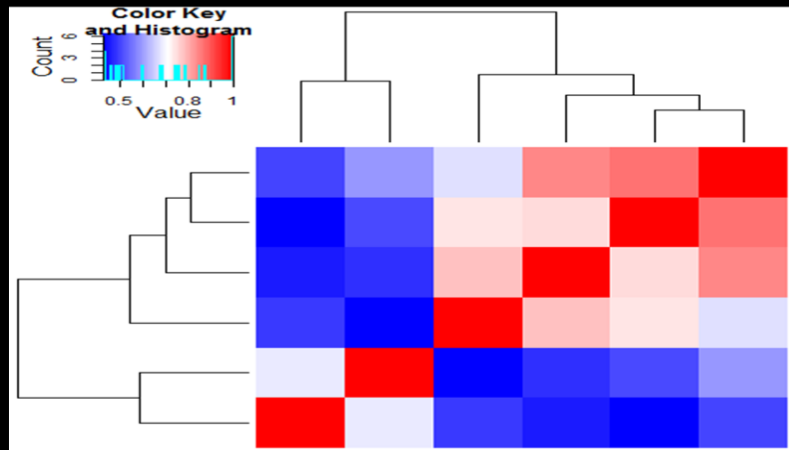
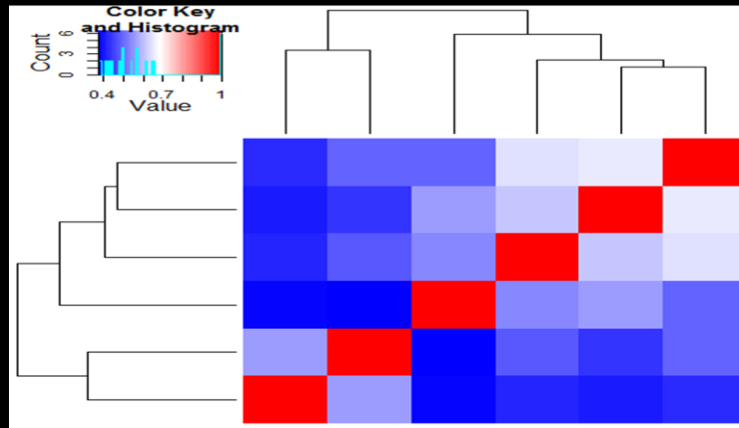


Figure 4.9- Correlation heatmap of gene expression in MDA-MB-231 after treatment with HKMT inhibitors for A) all genes on array B) MDA-MB-231 EZH2 silenced genes

When only MDA-MB-231 EZH2 silenced genes are investigated (Fig. 4.9 B), HKMT-I-005 and HKMT-I-011 correlate very strongly with each other, whilst compound TG3-184-1 clusters separately. This is surprising based upon the aforementioned capacity of TG3-184-1 to significantly up-regulate this set of EZH2 silenced genes in the MDA-MB-231 cells, and indicates that TG3-184-1 is having a different overall pattern of effect on these genes compared to HKMT-I-011 and HKMT-I-005.

The clear up-regulation of EZH2 silenced genes has been demonstrated, and the dual HKMT inhibitors show a similar pattern of induced expression change in EZH2 target genes as that shown by UNC0638, which also induces a strong reversal of silencing on these EZH2 targets in these MDA-MB-231 cells. What other genes are impacted by treatment with the HKMT inhibitors will be examined.

4.4 Functional signatures of dual HKMT inhibition

Differential expression caused by drug treatments were statistically ascertained to establish what genes showed a change in expression (**Materials & Method: Gene expression microarray**) at each treatment and time point. In order to assess the up or down regulation of cellular pathways, enrichment analysis for pathways annotated in ConsensusPathDB database (**Materials & Methods: ConsensusPathDB pathway enrichment analysis**) was performed.

Utilising the initial array data, in MDA-MB-231 cells 24 hours following treatment with HKMT-005, HKMT-011, and HKMT-022 apoptosis related pathways were the most significantly enriched, whilst protein processing in the endoplasmic reticulum was the most enriched pathway after treatment with both GSK343 or UNC0638 (Table 4.3).

Table 4.3: Top pathways enriched in the ConsensusPathDB database activated after 24 hours treatment

		Top pathways activated	
Treatment	Apoptosis Modulation and Signalling	Apoptosis	Protein processing in endoplasmic reticulum
HKMT-005	p<0.01	p<0.01	p<0.01
HKMT-011	p<0.01	p<0.01	p<0.01
HKMT-022	p<0.01	p<0.01	p<0.01
UNC0638	p=0.048	p=0.344	p<0.01
GSK343	p=0.106	p=0.422	p<0.01
TG3-259-1	p<0.01	p=0.032	p<0.01

As apoptosis related pathways were the most significantly enriched after treatment with HKMT-I-005, HKMT-I-022, and HKMT-I-011 after 24 hours, further analysis was performed on every pathway including the term apoptosis in the ConsensusPathDB database. When those pathways in which at least one treatment induced a significant enrichment (Table 4.4) are investigated, it is clear that HKMT-I-005, HKMT-I-011, HKMT-22, and UNC0638 all strongly impact the expression of genes related to multiple apoptosis pathways.

The specific EZH2 inhibitor GSK343 shows no significant activation of any apoptosis pathways studied. The impact of the dual HKMT inhibitors on cell clonogenicity, growth, and apoptosis will be further examined in Chapter 5.

Table 4.4: Top pathways enriched in the ConsensusPathDB database activated after 24 hours treatment (N.S=not statistically significant)

	Treatment					
Pathway	HKMT-005	HKMT-011	HKMT-022	TG3-259-1	UNC0638	GSK343
Intrinsic Pathway for Apoptosis	P<0.05	N.S	N.S	N.S	P<0.05	N.S
Apoptosis Modulation and Signalling	P<0.05	P<0.05	P<0.05	P<0.05	P<0.05	N.S
Apoptosis	P<0.05	P<0.05	P<0.05	P<0.05	P<0.05	N.S
Apoptosis	P<0.05	N.S	P<0.05	P<0.05	P<0.05	N.S
Apoptosis - Homo sapiens	P<0.05	P<0.05	N.S	P<0.05	P<0.05	N.S
Caspase Cascade in Apoptosis	N.S	N.S	N.S	N.S	P<0.05	N.S

4.5 Identification of putative pharmacodynamic biomarkers & examination of chromatin state of target genes after dual HKMT inhibition

In an effort to refine the initial compound selection process (Chapter 1, 1.5), analysis was performed to select potential pharmacodynamic markers of response to dual HKMT inhibitors that may be used either in the compound selection process, or in future downstream studies.

Utilising the lists of significantly differentially expressed genes generated during statistical analysis of the microarray data, the identification of potential biomarkers was undertaken.

Differential expression caused by drug treatments were statistically ascertained (**Materials & Methods: Gene expression microarray**) for each treatment and time point.

Genes significantly upregulated after treatment with HKMT-I-005, HKMT-I-022, and HKMT-I-011 were overlapped with the meta-analysis derived list of EZH2 silenced genes (**Material & Methods: Enrichment analysis**) to produce a shortlist of potential pharmacodynamic biomarkers that showed consistent expression upregulation following a diminishment of EZH2 levels- four genes were initially identified: RHOQ, IL24, HDAC9, and SPINK1.

One of these genes, SPINK1, was selected to be taken forward as an initial candidate pharmacodynamic biomarker. It was upregulated after treatment with HKMT-I-005, HKMT-I-011, and HKMT-I-022 at multiple doses and time points. SPINK1 is also known as pancreatic secretory trypsin inhibitor (PSTI) and is a potent protease inhibitor¹³⁷. In collaboration with **Luke Payne** (MRes student) QRT-PCR Primers were designed (primer details in Table 2.1) around the transcription start site of SPINK1 and QRT-PCR performed using treated MDA-MB-231 cells (**Materials & Methods: QRT-PCR**).

In this treatment, dose ranges of HKMT-I-005, GSK343, and UNC0638 were applied to MDA-MB-231 breast cancer cells. In addition, a dose range of GSK343 was applied in addition to a dose of 7.5 μ M of UNC0638 in the hopes of simulating dual knockdown of EZH2 and EHMT2.

HKMT-I-005 dose dependently increase expression of EZH2 target genes KRT17 and FBXO32 (Fig.4.10 A). UNC0638 increases expression levels of only FBXO32, and GSK343 has no discernible impact on the expression of these target genes. SPINK1 shows upregulation after treatment with HKMT-I-005 (Fig.4.10 A), no upregulation from treatment with GSK343 (Fig.4.10 C) or UNC0638 (Fig.4.10 B), but when UNC0638 and GSK343 are given in combination upregulation of SPINK1 expression occurs (Fig.4.10 D).

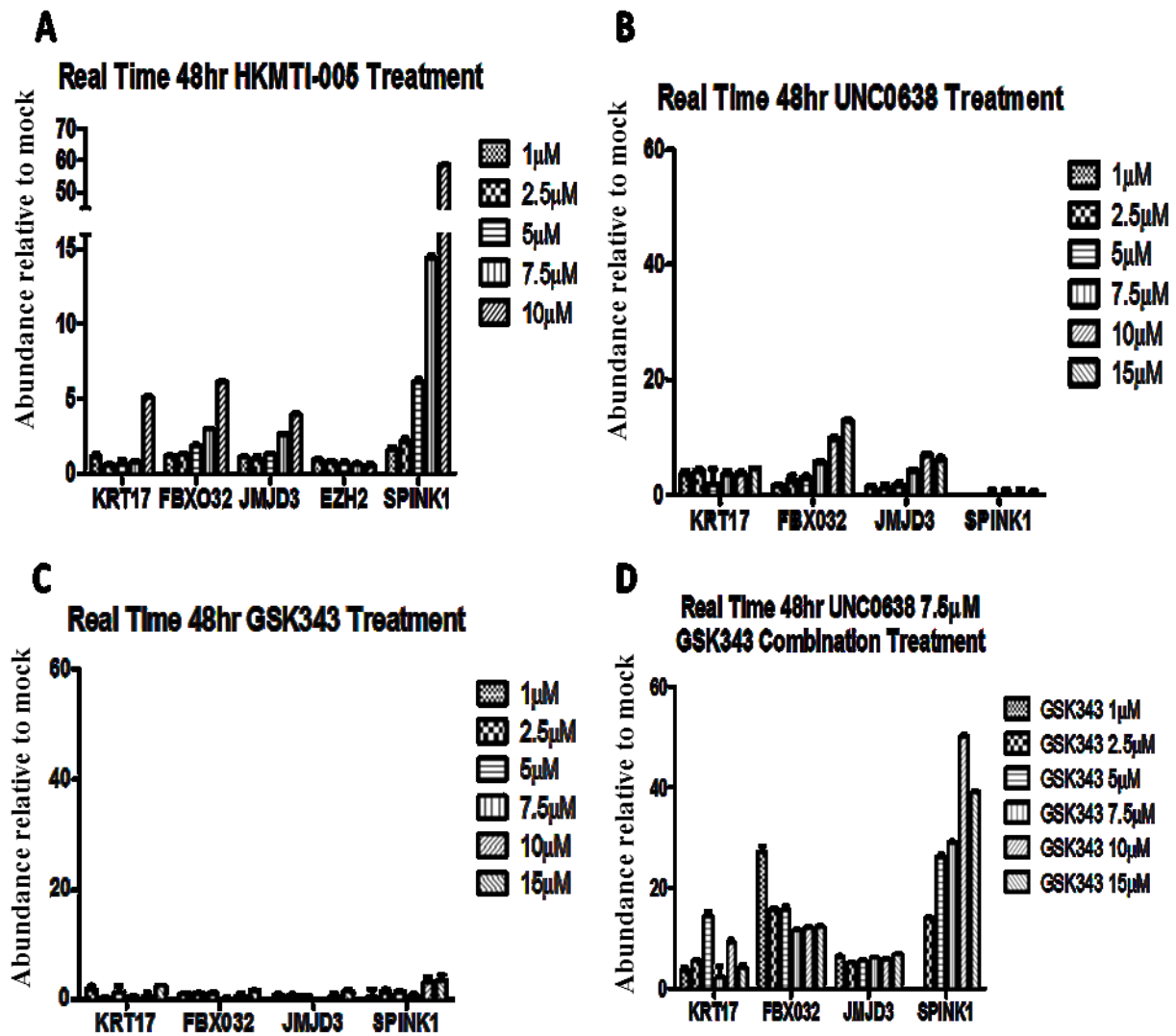


Figure 4.10 -QRT-PCR performed by Luke Payne using RNA from MDA-MB-231 cells after 48 hours treatment with A) HKMT-005 B) UNC0638 C) GSK343 and D) UNC0638 + GSK343. Error bars SEM of technical replicates (n≥3).

This preliminary data highlights the possibility of SPINK1 as a biomarker- it not only shows strong upregulation after treatment with HKMT-I-005, but the fact it is only upregulated by GSK343 and UNC0638 when they are given in combination indicates this may be a gene only upregulated by dual inhibition of EZH2 and EHMT2.

In collaboration with **Elham Shamsaei** work was performed to induce siRNA knockdowns of EZH2 and EHMT2 expression and measured SPINK1 mRNA levels using QRT-PCR (**Materials & Methods : siRNA knockdown experiments**).

In the MDA-MB-231 breast cancer cell line, SiRNA knockdown was used to examine the effect of combined inhibition of EZH2 and EHMT2 expression on SPINK1 levels (Fig. 4.11).

Individual siRNA knockdown of EZH2 (Fig.4.11 A) had no impact on SPINK1 expression.

Individual siRNA knockdown of EHMT2 (G9a) (Fig.4.11 B) had no impact on SPINK1 expression. Dual siRNA knockdowns of both EZH2 and EHMT2 (Fig.4.11 C) led to strong upregulation of SPINK1, reinforcing the findings shown by chemical dual inhibition (Fig.4.10 A).

So it is established that SPINK1 showed upregulation of expression following treatment with HKMT-I-005 (Fig.4.10 A), dual inhibition with GSK343 and UNC0638 (Fig.4.10 D) and treatment with HKMT-I-005, HKMT-I-011, and HKMT-I-022 all induced upregulation of expression of previously identified target genes KRT17 and FBXO32 (supplementary table 8.1).

In order to verify that the upregulation in target gene expression is due to chromatin remodelling as theorised, ChIP qRT-PCR experiments (**Materials & Methods: Chromatin immunoprecipitation**) were performed. Initially, in collaboration with **Nadine Chapman-Rothe**, the chromatin state of the promoter region of KRT17 and the TSS of FBXO32 were investigated for the levels of known repressive chromatin marks H3K9me3 and H3K27me3 as well as the known ‘activating’ chromatin marks H3K4me3, H3K4me2, H3K27ac and H3K9ac. Also investigated was the level of the H3K27 demethylase JMJD3.

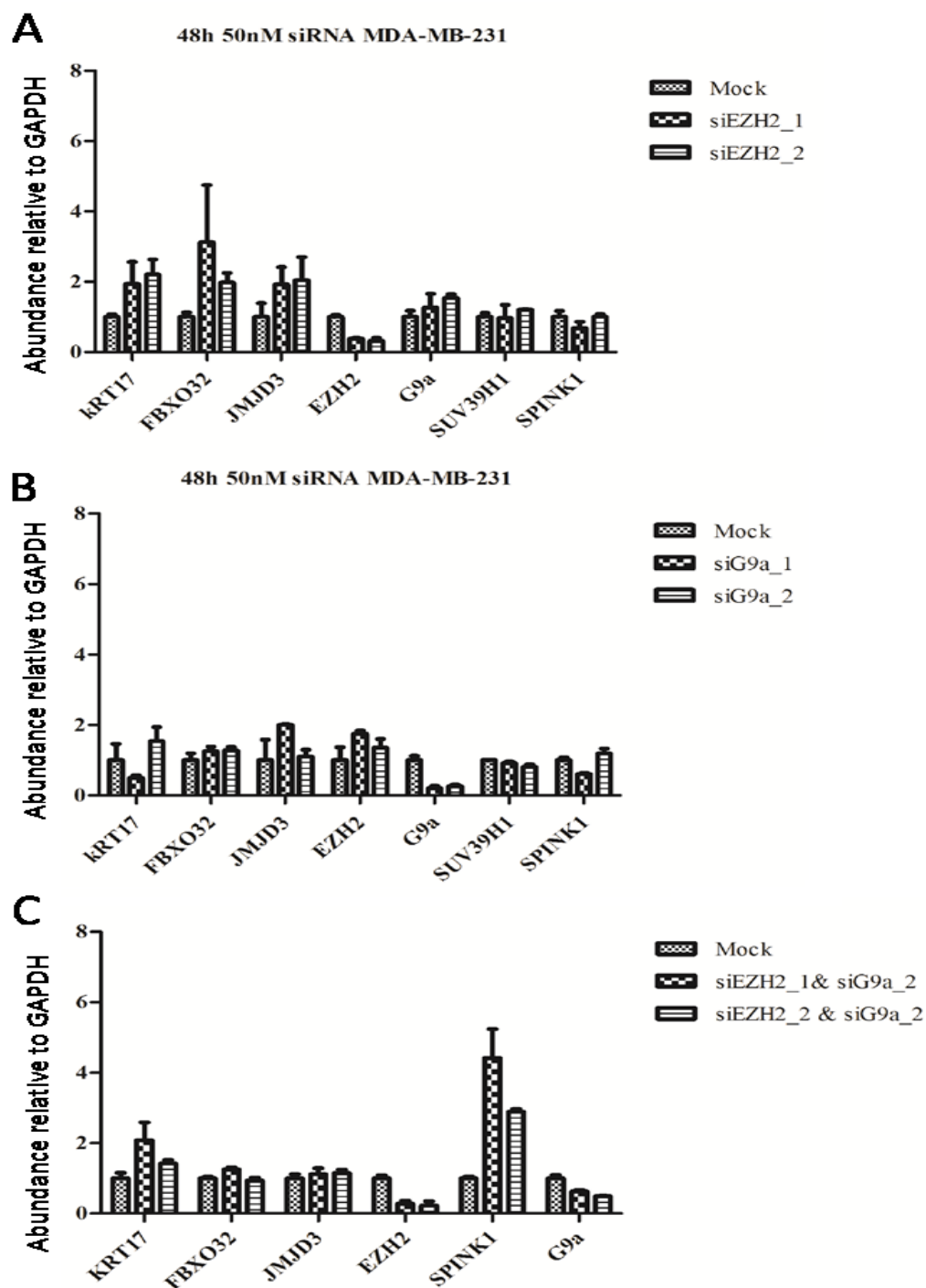


Figure 4.11- QRT-PCR of target genes performed by Elham Shamsaei using RNA from MDA-MB-231 cells after siRNA knockdown of A) EZH2 B) EHMT2 (G9a) C) EZH2 and EHMT2 (G9a). Error bars SEM of technical replicates (n≥3).

At a dose of 5 μ M, HKMT-I-005, HKMT-I-022, and HKMT-I-011 all showed a decrease in levels of H3K27me3 and H3K9me3 repressive chromatin marks at the KRT17 promoter region

(Fig.4.12A) and the TSS of the FBXO32 gene (Fig.4.12B). This is consistent with the capacity of the inhibitors to target both EZH2 and EHMT2, the primary enzymes responsible for H3K27me3 and H3K9me3 deposition (respectively).

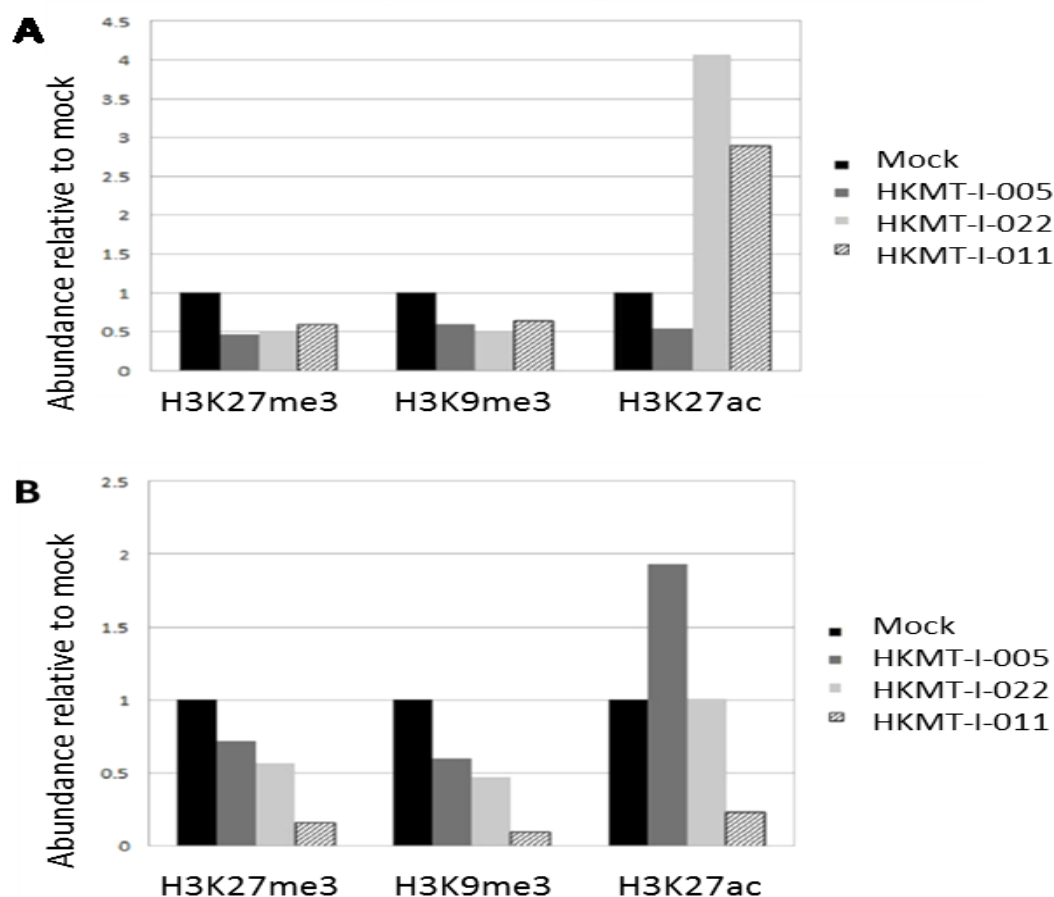


Figure 4.12- representative examples of a series of ChIP experiments which consistently showed similar changes in collaboration with *Nadine Chapman-Rothe*- ChIP qRT-PCR for H3K27me3, H3K9me3, and H3K27me3 after 72 hour treatment with 5 μ M of HKMT-I-005, HKMT-I-011, HKMT-I-022 or Mock (DMSO) at A) KRT17 promoter region B) FBXO32 TSS

Also shown was an increase in the level of some activating marks- HKMT-I-005 showed an increase in H3K4me3 at the KRT17 promoter (Fig.4.13A) and an increase in H3K4me2, H4K4me3, and H3K9ac at the FBXO32 TSS (Fig.4.13B); HKMT-I-022 and HKMT-I-011 treatment led to increased H3K27ac (Fig.4.12) and H3K4me, H3K4me3, and H3K9ac levels at the KRT17 promoter and FBXO32 TSS (Fig.4.13).

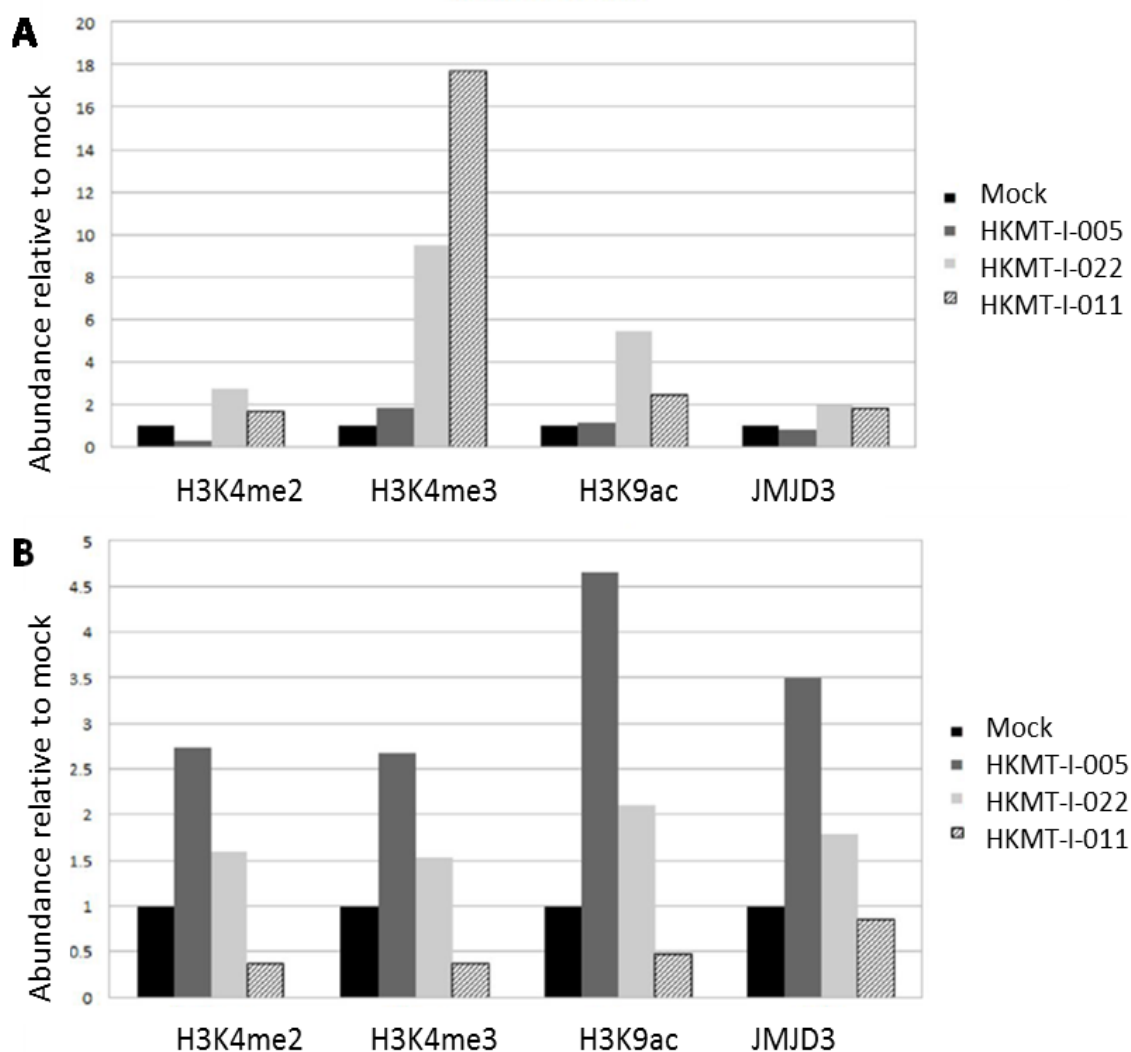
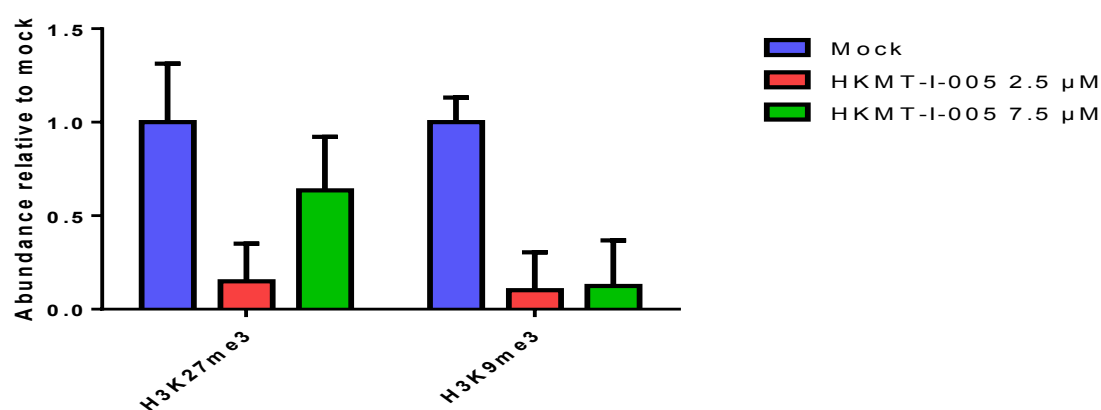


Figure 4.13- representative examples of a series of ChIP experiments which consistently showed similar changes in collaboration with *Nadine Chapman-Rothe*- ChIP qRT-PCR for H3K24me2, H3K4me3, H3K9ac, and JMJD3 after 72 hour treatment with 5 μ M of HKMT-I-005, HKMT-I-011, HKMT-I-022 or Mock (DMSO) at A) KRT17 promoter region B) FBXO32 TSS

Some increase in the levels of H3K27 demethylase JMJD3 was also observed after some treatments at the KRT17 promoter region and FBXO32 TSS (Fig.4.13) but this change was not consistent.

Overall this shows a decrease in levels of repressive chromatin marks at these two target gene loci and an increase in the levels of transcriptionally permissive chromatin marks. Further investigation of the repressive chromatin marks H3K27me3 and H3K9me3 was performed on the TSS of the putative pharmacodynamic biomarker SPINK1 after treatment with HKMT-I-011 and HKMT-I-005 (the two inhibitors that showed the strongest up-regulation of EZH2 silenced genes in the expression array).

A



B

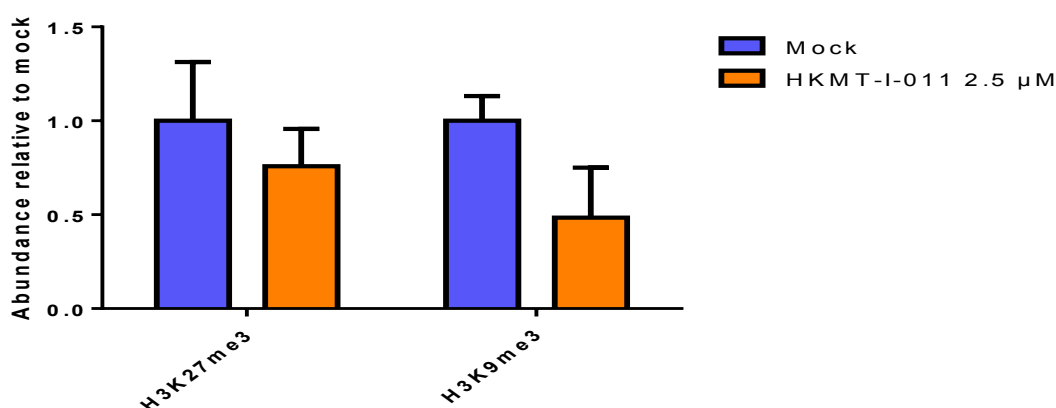


Figure 4.14- ChIP qRT-PCR for H3K27me3 and H3K9me3 at SPINK1 TSS after 24 hour treatment with A) HKMT-I-005 B) HKMT-I-011. Error bars SEM of biological replicates (n $\geq$ 2), Student's t-test not significant between conditions.

In MDA-MB-231 cells, ChIP qRT-PCR (**Materials & Methods: Chromatin immunoprecipitation**) was performed- 24 hours treatment with HKMT-I-005 at 2.5 μ M led to a decrease in H3K27me3 and H3K9me3 at the SPINK1 TSS (Fig.4.14A). Treatment with HKMT-I-005 for 24 hours at a dose of 7.5 μ M also showed a decrease in H3K9me3 and H3K27me3 at the SPINK1 TSS (Fig.4.14B), but notably the decrease in H3K27me3 was of a smaller size than the decrease observed after treatment with 2.5 μ M- the reason for this is presently unclear, but further pharmacodynamic studies may provide insight into the potential longevity of effect of these inhibitors. 24 hours of treatment with 2.5 μ M of HKMT-I-011 led to a slight decrease in H3K27me3 and a larger decrease in H3K9me3 at the SPINK1 TSS (Fig.4.14B). Together, these results support that a decrease in H3K27me3 and H3K9me3 leads to the upregulation of SPINK1 expression, and taken with the FBXO32 and KRT17 results, that these drugs are inducing upregulation of expression by means of chromatin remodelling.

MDA-MB-231 cells were treated with EZH2 specific inhibitor GSK343 at a dose of 2.5 μ M (Fig.4.15A) and no impact was observed on the levels of H3K27me3 or H3K9me3, despite this dose of GSK343 being capable of inducing significant upregulation of known EZH2 silenced target genes (Fig.4.2). UNC0638 reduced H3K9me3 levels at the SPINK1 TSS dramatically (Fig.4.15B) and also showed a strong reduction in the levels of H3K27me3- as an established EHMT2 specific inhibitor, this induced reduction in H3K27me3 supports the theorised supporting role of EHMT2 in establishing and maintaining EZH2 mediated H3K27me3 levels. This does not however lead to an increase in SPINK1 expression after UNC0638 treatment- this may be a pharmacodynamic effect. Further examination of the chromatin state across the SPINK1 gene would indicate if this alteration in H3K27me3 and H3K9me3 levels is consistent across the promoter regions and TSS.

This result is further supported by Western blot analysis performed by **Sarah Kandil** showing a decrease in global levels of H3K27me3 and H3K9me3 following HKMT-I-005 treatment of MDA-MB-231 cells (Supplementary figure 8.3).

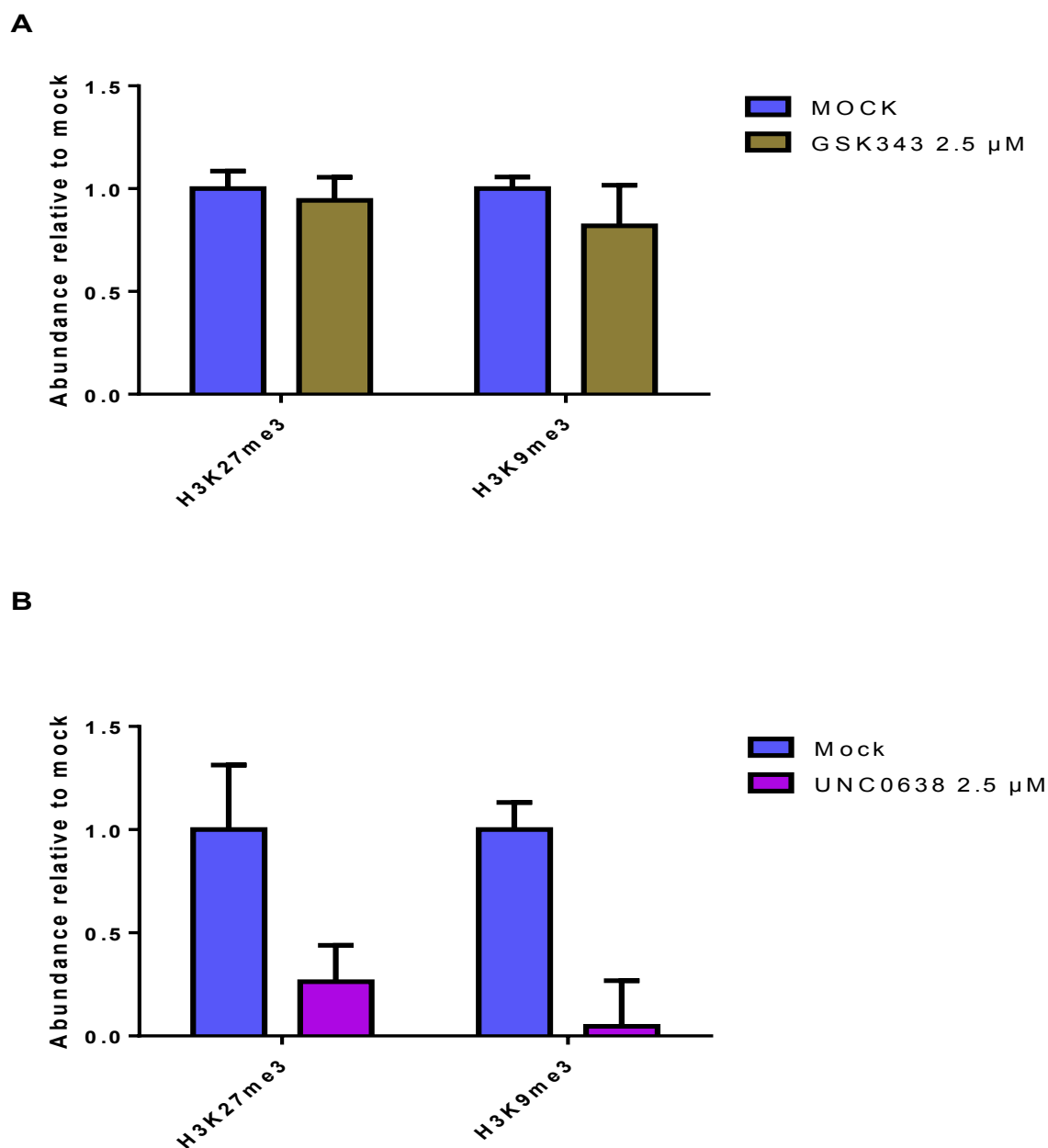


Figure 4.15- ChIP qRT-PCR for H3K27me3 and H3K9me3 at SPINK1 TSS after 24 hour treatment with A) GSK343 B) UNC0638. Error bars SEM of biological replicates ($n \geq 2$), Student's t-test not significant between conditions.

4.6 Summary

In MDA-MB-231 breast cancer cells, dual inhibitors of EZH2 HKMT-I-005, HKMT-I-011, and HKMT-I-022, all showed a capacity to significantly increase the expression levels of genes known to be repressed by EZH2 (4.2). This upregulation was of a consistently more significant nature than that caused by GSK343 or UNC0638 (EZH2 and EHMT2 inhibitors respectively) and this was shown to be a significant difference (in the case of HKMT-I-005 after 24 hours treatment, HKMTI-1-005 upregulated EZH2 silenced genes significantly more than either GSK343 ($p=5.8E-5$) or UNC0638, ($p=1.7E-4$)). TG3184-1, a compound that failed the initial chemical screen for dual inhibitors (Chapter 1, 1.5) showed a capacity to upregulate expression of EZH2 target genes, highlighting the importance of developing a robust panel of biomarkers to enhance the compound selection screen.

Importantly, when EZH2 target genes from a different cell type (MCF-7) were investigated, no significant increase in the expression levels of these genes was observed in the MDA-MB-231 cells treated with the HKMT inhibitors (4.2). This indicates that the targets of EZH2 mediated gene repression vary between cell types, and as such moving forward it will be important to characterise new target gene sets when working in new tissues.

In an effort to address this, a meta-analysis of EZH2 siRNA studies was performed by MRes student **Emma Bell** to find a list of genes showing consistent differential expression after a reduction of EZH2 levels- using this list of EZH2 target genes (Supplementary Table 4.2.7), the dual HKMT inhibitors showed significant upregulation the EZH2 repressed genes identified through the meta-analysis.

Comparison of the treatments impact on expression of identified EZH2 target genes highlighted a great degree of similarity between the dual HKMT and the EHMT2 specific inhibitor UNC0638 (4.3). Both are derived initially from BIX-01294 (Chapter 1, 1.5), and so it is

perhaps unsurprising that they have a similar impact on these target genes, though the dual HKMT inhibitors induce a more specific systematic upregulation of the EZH2 targets than EHMT2 inhibitor UNC0638 does. GSK343, despite showing significant upregulation of EZH2 repressed genes, did not appear to affect these genes in a similar pattern to the dual inhibitors or the EHMT2 specific inhibitor.

Functional annotation enrichment analysis (4.4) highlighted the induction of apoptotic pathways after treatment with the dual HKMT inhibitors or the EHMT2 inhibitor UNC0638- GSK343 showed no significant induction of any apoptotic pathways identified in the MDA-MB-231 cells, in keeping with published literature that EZH2 specific inhibitors have relatively little impact on the proliferation in solid cancers such as breast cancer. The impact of these inhibitors on apoptosis, cell proliferation, and clonogenicity will be further investigated in Chapter 5.

As the capacity of the TG3-184-1 compound to upregulate expression of EZH2 repressed genes illustrated (4.2), pharmacodynamic biomarkers are essential in order to measure the response of a cell type to inhibitors of HKMT like EZH2, which we showed targets different genes depending on cell type- lacking robust biomarkers, potentially potent inhibitors may be wrongly classified as ineffective and not worth pursuing.

Utilising the array data, SPINK1 was identified as a potential biomarker for the reversal of EZH2 mediated silencing (4.5). Further exploration showed that in fact SPINK1 was only upregulated after dual inhibition of EZH2 and EHMT2, which suggests that there may be a subset of genes only upregulated after removal of both of these HKMT.

ChIP qRT-PCR studies confirmed that at loci on target genes FBXO32, KRT17 and SPINK1, treatment with dual HKMT inhibitors leads to a reduction in H3K27me3 and H3K9me3 levels, further supporting the nature of these dual inhibitors and their mechanism of action (global

reduction in H3K27me3 and H3K9me3 levels was also observed after treatment with HKMT-I-005). These preliminary studies were not found to be statistically significant- variability between the chromatin preparations meant that comparison between experiments was not statistically viable. In these experiments (illustrated in Fig 4.12-4.15) the IP values were compared against a mock IP with no antibody. Whilst within experiments this allowed comparison, a lack of internal control for each sample meant that comparing across experiments was very difficult. For future studies, isolating DNA from each sample after the sonication process would allow normalisation against DNA concentration that should allow more robust statistical analyses to be used. In combination with the mock IP, this input DNA would allow further analyses despite differences in chromatin preparation efficiency.

Chapter 5: Effect of dual HKMT inhibition on cancer cell phenotype and cancer stem cells

5.1 Introduction

Having characterised the capacity of dual HKMT inhibitors HKMT-I-005, HKMT-I-011, and HKMT-I-022 to induce re-expression of genes repressed by EZH2 (Chapter 4), the impact of this expression change on proliferation of cancer cell populations was examined. SAM substrate competitive inhibitors of EZH2 have been identified and characterised (Chapter 1, Table 1.2) -their impact on cell proliferation has primarily been characterised in EZH2 mutant lymphoma cells. A panel of cell lines were treated with compound HKMT-I-005 to evaluate the impact of this compound on cell proliferation. The effect of EZH2 inhibition, EHMT2 inhibition, or dual inhibition of EZH2 and EHMT2 on cell proliferation was examined in some of these cell types (5.2).

As discussed (Chapter 1, 1.4), CSC subpopulations have been discovered in multiple cancer types, including breast cancer and ovarian cancer (Chapter 1, Table 1.3). Many of these CSC subpopulations show a reliance on EZH2 expression to maintain their CSC phenotype (e.g. breast and pancreatic cancers⁴⁹, brain cancer⁴⁸, prostate cancer⁹³). This group previously identified the reliance on EZH2 of a CSC-like subpopulation in ovarian cancer cells⁴⁷ - this subpopulation is characterised as overexpressing ABC drug transporters and sustaining chemotherapy resistant growth. A reduction in the levels of EZH2 led to a decrease in this CSC population in IGROV1 ovarian cancer cells.

Inhibition of EZH2, EHMT2, or both was performed on IGROV1 ovarian cancer cells to examine the impact of the dual HKMT inhibitors on CSC activity and self-renewal capacity in this system (5.2).

We showed that EZH2 is linked to poorer RFS and OS in breast cancer (Chapter 3) and that dual HKMT inhibition led to a significant re-expression of EZH2 repressed genes in breast cancer, including significant up-regulation of apoptotic pathways (Chapter 4). EZH2 has been shown to expand the CSC pool in breast cancer through activation of NOTCH1¹³⁸. As high EZH2 expression is linked to the CSC population in breast cancer, the impact of dual HKMT inhibitors on CSC activity, self-renewal capacity, and clonogenicity of MDA-MB-231 cancer cells was studied (5.2) - in addition, the impact of cytotoxic chemotherapy (Cisplatin or Paclitaxel (Taxol)) was compared to the dual HKMT inhibitors and single HKMT inhibitors, as well as combined treatment with HKMT inhibition and these therapies.

This examination of CSC activity and self-renewal capacity was carried out using established sphere forming assays¹²¹. However, sphere-forming assays may not detect quiescent stem cells and sphere-forming assays are not a read-out of *in vivo* stem cell frequency¹³⁹ - as such, in collaboration with **Gillian Farnie** and **Amrita Shergill** at the University of Manchester the impact of dual EZH2 and EHMT2 inhibition on cancer stem cell action in breast cancer cells implanted in immunocompromised mice was examined (5.4). In addition, based upon the data generated combining dual HKMT inhibitors with Paclitaxel in sphere forming models of breast cancer (5.3) a combination treatment of dual HKMT inhibitor and Paclitaxel was investigated using the mouse model.

5.2 Effect of dual HKMT inhibition on cancer cell proliferation

Using dual HKMT inhibitor HKMT-I-005 as an exemplar for its class, proliferation studies were carried out on a number of lymphoma, ovarian cancer, and breast cancer cell lines

(**Materials & Methods: Cell proliferation assay**) after treatment with HKMT-I-005. The IC₅₀ on cell proliferation with HKMT-I-005 treatment is shown, as well as known mutations in EHMT2 and EZH2 (Table 5.1).

Table 5.1- Impact of HKMT-I-005 on cell proliferation of lymphoma, breast cancer, and ovarian cancer cell lines (in collaboration with Anthony Uren, Sarah Kandil, and Elham Shamsaei)

Cell type	Cell Line	IC ₅₀ of cell proliferation (μM)	MUTATIONS	
		HKMT-I-005	EZH2	EHMT2
Lymphoma	SC1	3.71	No mutation	No mutation
	WILL1	5.6	No data	No data
	DOHH2	3.26	No mutation	No mutation
	WSU-FSCLL	3.41	No data	No data
	DB	<1	p.Y646N	No mutation
	SUDHL8	<1	No mutation	No mutation
Ovarian Cancer	A2780	15.96	No mutation	No mutation
	A2780CP	21.21	No data	No data
	PEO23	27.82	No data	No data
	PEO14	22.92	No data	No data
	PEO1	15.45	No mutation	No mutation
	PEO4	29.77	No data	No data
Breast cancer	MDA-MB-231	10.4	No mutation	No mutation
	MCF7	7.7	No mutation	No mutation
	T47D	8.5	No mutation	No mutation
	BT474	2.1	No data	No data
	SKBR3	7.7	No data	No data
Breast epithelial	MCF10A	>15	No data	No data

In lymphoma cell lines, HKMT-I-005 consistently has an IC₅₀ < 6μM. In ovarian cancer a much higher dose (~15-28μM) was needed to see this effect. In breast cancer cell lines, HKMT-

I-005 had an $IC_{50} \leq 10\mu M$, but in the breast epithelial cell line MCF10a the $IC_{50} > 15\mu M$ – these MCF10a cells represent immortalised breast epithelial cells and are used as a ‘normal’ breast endothelial cell model in comparison to the breast cancer cell lines.

One of the most strikingly low IC_{50} s was observed in DB cells, which are characterised as having a point mutation in EZH2 (p.Y646n). Cancers with this mutation are particularly susceptible to EZH2 inhibition (chapter 1, 1.4), so it is encouraging that dual EZH2/EHMT2 inhibitor HKMT-I-005 has this impact. However, a similarly low IC_{50} as observed in the SUDHL8 cell line, which is reported as EZH2 wild-type. This indicates that the impact of EZH2 inhibition is not dependent solely on the EZH2 mutation state of the cell- two public databases of CNV in cancer cell lines were consulted^{120,140} to see if EZH1, EZH2, EHMT1, or EHMT2 showed any increase in copy number in cell lines HKMT-I-005 strongly affected proliferation (Supplementary table 8.12). This data shows SUDHL8, DOHH2, and DB lymphoma cells all show increased copy numbers of EZH1, EZH2, EHMT1, and EHMT2. This data suggests that increased levels of EZH2 can lead to a susceptibility to EZH2 inhibition, be that increase due to mutation (such as in the DB cells) or due to CNV (such as in SUDHL8 cells). In addition, ovarian cancer cell line A2780 showed amplification of copy number of EZH1, EZH2, EHMT1, and EHMT2, and showed one of the lowest IC_{50} values of the ovarian cancer lines ($IC_{50}=15.96\mu M$) – recent studies have shown synthetic lethality between EZH2 inhibition and ARID1A mutation¹⁴¹ and A2780s are characterised as ARID1A mutants which may explain the sensitivity of these cells.

Treatment with EZH2 inhibitor GSK343 or EHMT2 inhibitor UNC0638 was also performed on these lymphoma cell lines (Table 5.2). UNC0638 consistently showed an $IC_{50} \leq 1\mu M$. GSK343 did have a strong impact on the EZH2 mutant cell line DB, but was less efficacious in the other lymphoma cell lines studied.

Table 5.2- Impact of HKMT-I-005, GSK343, and UNC0638 on cell proliferation of lymphoma

Cell Line	IC50 (μ M)			MUTATIONS	
	HKMT-I-005	GSK343	UNC0638	EZH2	EHMT2
SC1	3.71	12.12	1.128	No mutation	No mutation
WILL1	5.6	17.91	<1	No data	No data
DOHH2	3.26	6.15	<1	No mutation	No mutation
WSU-FSCLL	3.41	2.87	<1	No data	No data
DB	<1	<1	<1	p.Y646N	No mutation
SUDHL8	<1	5.11	<1	No mutation	No mutation

Using the MDA-MB-231 cells in which dual inhibition of EZH2 and EHMT2 was characterised (Chapter 4), MRes student **Luke Payne** studied the impact of single inhibitors of EZH2 (GSK343) and EHMT2 (UNC0638) on cell proliferation (Fig.5.1).

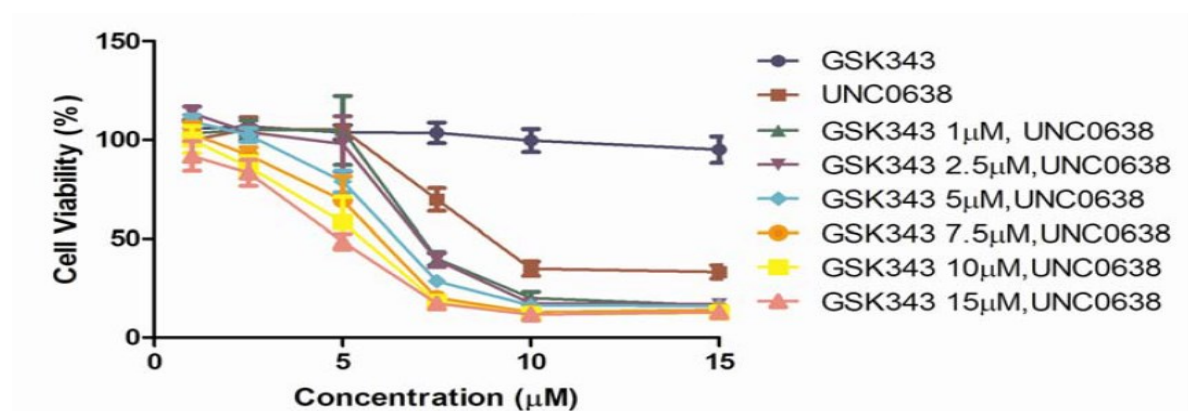


Figure 5.1- MTT assay for cell viability of MDA-MB-231 cells after treatment. MDA-MB-231 cells were seeded in 96 well plates. After 24hrs, increasing doses of GSK343, UNC0638 or combination treatments (1, 2.5, 5, 7.5, 10 and 15 μ M) were added to cells. Control was media with 0.5% DMSO. Cell viability was measured by MTT assay after 48hrs treatment and a 24hr proliferation period. Results are shown from five independent repeats of MTT assays in MDA-MB-231. Error bars represent the mean $\pm$ SEM of five independent repeats.

Treatment with EZH2 inhibitor GSK343 showed no significant reduction in cell viability up to doses of 15 μ M. UNC0638 caused a dose dependent reduction in cell proliferation, with an IC₅₀ of 9.8 μ M. When cells were treated with a combination of both UNC0638 and GSK343, there was a significant increase in growth inhibition. Of particular note is a dose of 5 μ M of both compounds- individually, neither of these inhibitors had a significant impact on cell proliferation at this dose. Combined, they reduce cell viability > 50% (p<0.01)

This supports the theory of combined inhibition of EZH2 and EHMT2 having a stronger impact on EZH2 mediated repression and thus cellular response.

5.3 Effect of dual HKMT inhibition on cancer stem cell activity, self-renewal, and chemosensitivity in *in vitro* models

Having established that HKMT-I-005 is capable of reducing cell proliferation in the main cancer cell population of a number of cell lines, the question as to the efficacy of dual HKMT inhibition on impacting CSC activity and self-renewal was addressed.

As a baseline for comparison, clonogenic assays were performed (**Materials & Methods: Clonogenic assay**) initially in IGROV1 ovarian cancer cells- this model was used in the hope of corroborating the groups earlier published findings⁴⁷ showing EZH2 as vital for CSC activity in this cell line.

Cells were treated with DMSO as a control, HKMT-I-005, HKMT-I-011, GSK343, UNC0638 and the chemotherapeutic agent Paclitaxel.

HKMT-I-005, HKMT-I-011, and UNC0638 all lead to a significant reduction in colony formation of IGROV1 cells at doses as low as 1 μ M (Fig.5.2.A, B, D respectively). Paclitaxel very significantly reduced colony formation in IGROV1 cells, completely halting colony formation at doses >1 μ M (Fig.5.2 E).

EZH2 specific inhibitor GSK343 had no significant impact on colony formation of IGROV1 cells until a dose of 7.5 μ M, and colony formation was >90% of that observed in the control doses after treatment with GSK343 up to 15 μ M (Fig.5.2 C)- statistical significance established by unpaired 2-tailed Student's t-test relative to DMSO control.

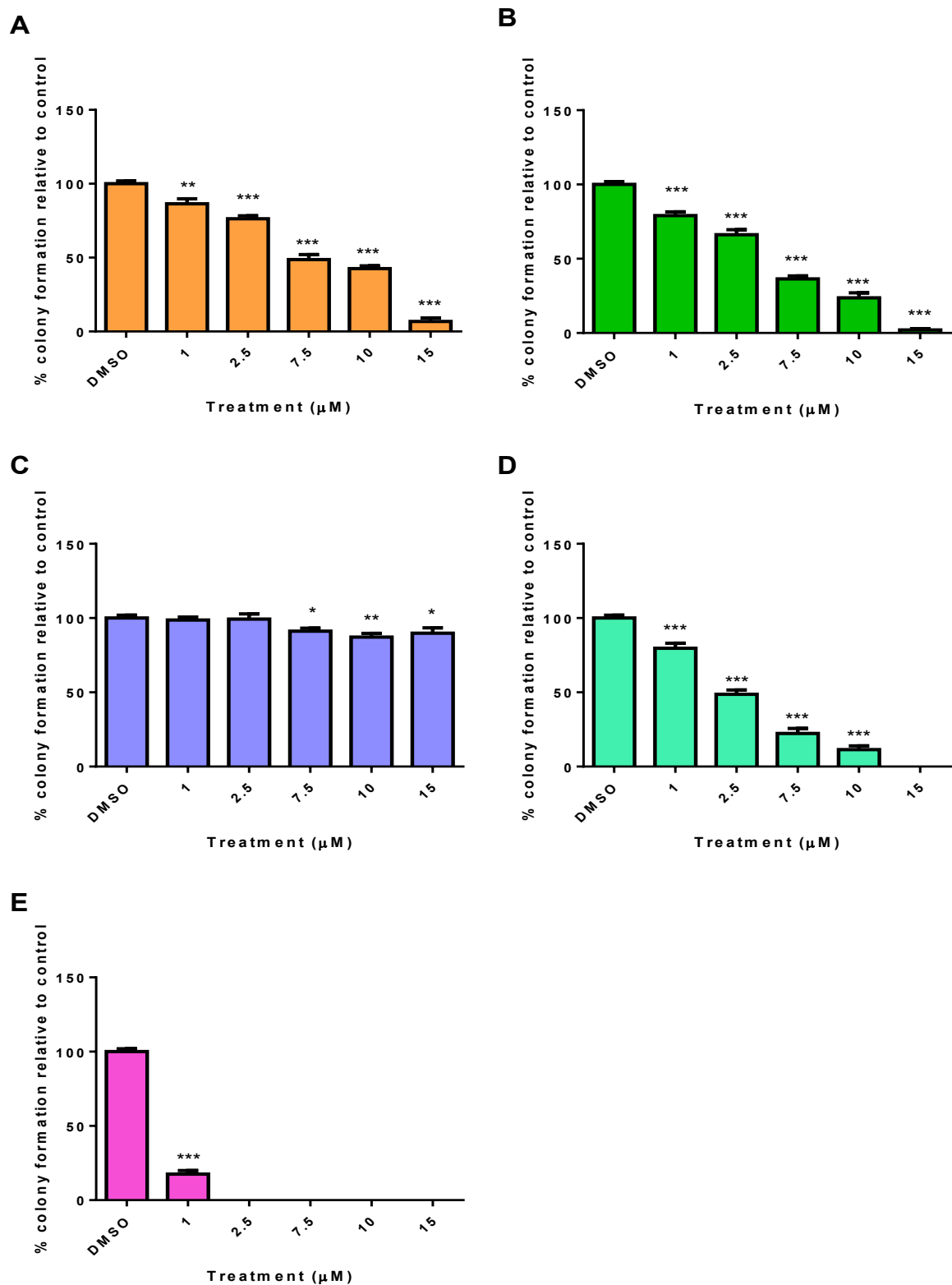


Figure 5.2 Clonogenic activity as measured by colony formation in IGROV1 ovarian cancer cells after treatment with A) HKMT-I-005 B) HKMT-I-011 C) GSK343 D) UNC0638 E) Paclitaxel– statistical significance calculated by Student’s T-test between DMSO control and dose- $p < 0.05 = *$, $p < 0.01 = **$, $p < 0.001 = ***$. Error bars are SEM ($n \geq 3$).

Calculated IC₅₀ doses for these treatments (Table 5.3) interestingly show that GSK343 is far less effective than HKMT-I-005, HKMT-I-011, or UNC0638 at inhibiting colony formation.- this indicates that in these IGROV1 ovarian cancer cells EZH2 specific inhibition is not as capable of reducing clonogenic capacity as dual EZH2/EHMT2 inhibitors, or EHMT2 specific inhibitors. Paclitaxel shows a predictably low IC₅₀ in the reduction of bulk clonogenic capability, as Paclitaxel targets mitotic division and is well known to impact cell proliferation and clonogenicity.

Table 5.3- Clonogenic IC₅₀ of treatments in IGROV1 ovarian cancer cells

	HKMT-I-005	HKMT-I-011	GSK343	UNC0638	PACLITAXEL
IC₅₀ (μM)	4.772	2.8	101.8	0.9349	<0.001

CSC activity was measured in IGROV1 cells by measurement of spheroid formation efficiency (**Materials & Methods: CSC activity and self-renewal capacity**) after treatment with HKMT-I-005, HKMT-I-011, GSK343, UNC0638, and mitotic inhibitor Paclitaxel (statistical significance for this experiment established by unpaired 2-tailed Student's t-test relative to DMSO control).

CSC activity was significantly reduced by as little as 0.1 μM of HKMT-I-005, UNC0638, or GSK343 treatment (Fig. 5.3A) and 0.5 μM of HKMT-I-011, indicating the CSC population in IGROV1 ovarian cancer cells are very susceptible to interference with EZH2 and EHMT2, and the strong response elicited after treatment with GSK343 indicates that this CSC population is more sensitive to the action of EZH2 specific inhibition than the general cell population (as measured by clonogenic assay Fig.5.2).

All of the HKMT inhibitors had IC₅₀ < 0.3 μM in when affecting CSC activity (Table 5.4).

Table 5.4- CSC activity IC₅₀ of treatments in IGROV1 ovarian cancer cells

	HKMT-I-005	HKMT-I-011	GSK343	UNC0638
IC ₅₀ (μM)	0.08483	0.2672	0.6104	0.09534

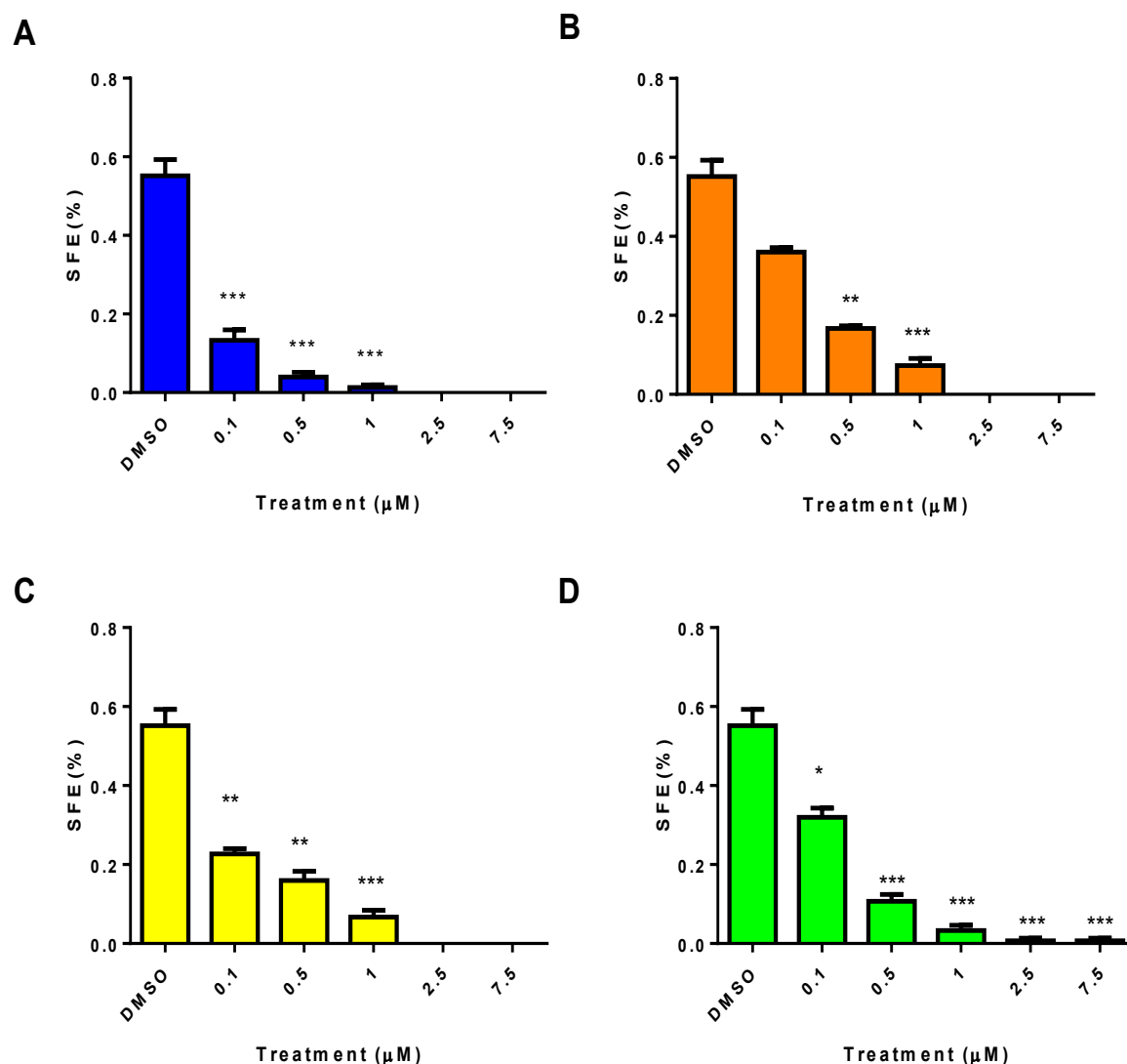


Figure 5.3 CSC activity as measured by spheroid formation efficiency in IGROV1 ovarian cancer cells after treatment with A) HKMT-I-005 B) HKMT-I-011 C) GSK343 D)UNC0638 – statistical significance calculated by Student’s T-test between DMSO control and dose- p<0.05=*, p<0.01=, p<0.001=***. Error bars are SEM (n≥3).**

This data also supports this group's previously published research indicating the CSC population in IGROV1 ovarian cancer cells lines are susceptible to treatment by reduction in EZH2 levels.

This assay was repeated after treatment with mitotic inhibitor Paclitaxel or alkylating agent Cisplatin, as well as Paclitaxel with co-treatment with 1 μ M HKMT-I-005, and Cisplatin with 24 hour pre-treatment with 1 μ M HKMT-I-005 (Fig.5.4).

Paclitaxel shows a significant decrease in CSC activity as measured by SFE (Fig.5.4 A), significant at doses >1 μ M and with a calculated IC₅₀ on CSC activity at ~0.81 μ M (Table 5.4). Cisplatin also significantly decreases IGROV1 CSC activity (Fig.5.4 B) at doses as low as 0.1 μ M with a calculated IC₅₀ of 0.43 μ M.

In the clonogenic assay measuring clonogenic capacity of the cancer cell population the HKMT inhibitors were less efficacious than the traditional chemotherapeutic agent Paclitaxel (Fig.5.2, Table 5.3). When measuring the impact on CSC activity of the IGROV1, the HKMT inhibitors are as potent as or more potent than traditional chemotherapeutic agents Cisplatin and Paclitaxel (Table 5.5).

Table 5.5- CSC activity IC₅₀ of treatments in IGROV1 ovarian cancer cells (including chemotherapy)

	HKMT-I-005	HKMT-I-011	GSK343	UNC0638	PACLITAXEL	PACLITAXEL + 1 μ M HKMT-I-005	CISPLATIN	CISPLATIN + 1 μ M HKMT-I-005
IC₅₀ (μM)	0.08	0.26	0.61	0.09	0.81	>0.01	0.43	0.33

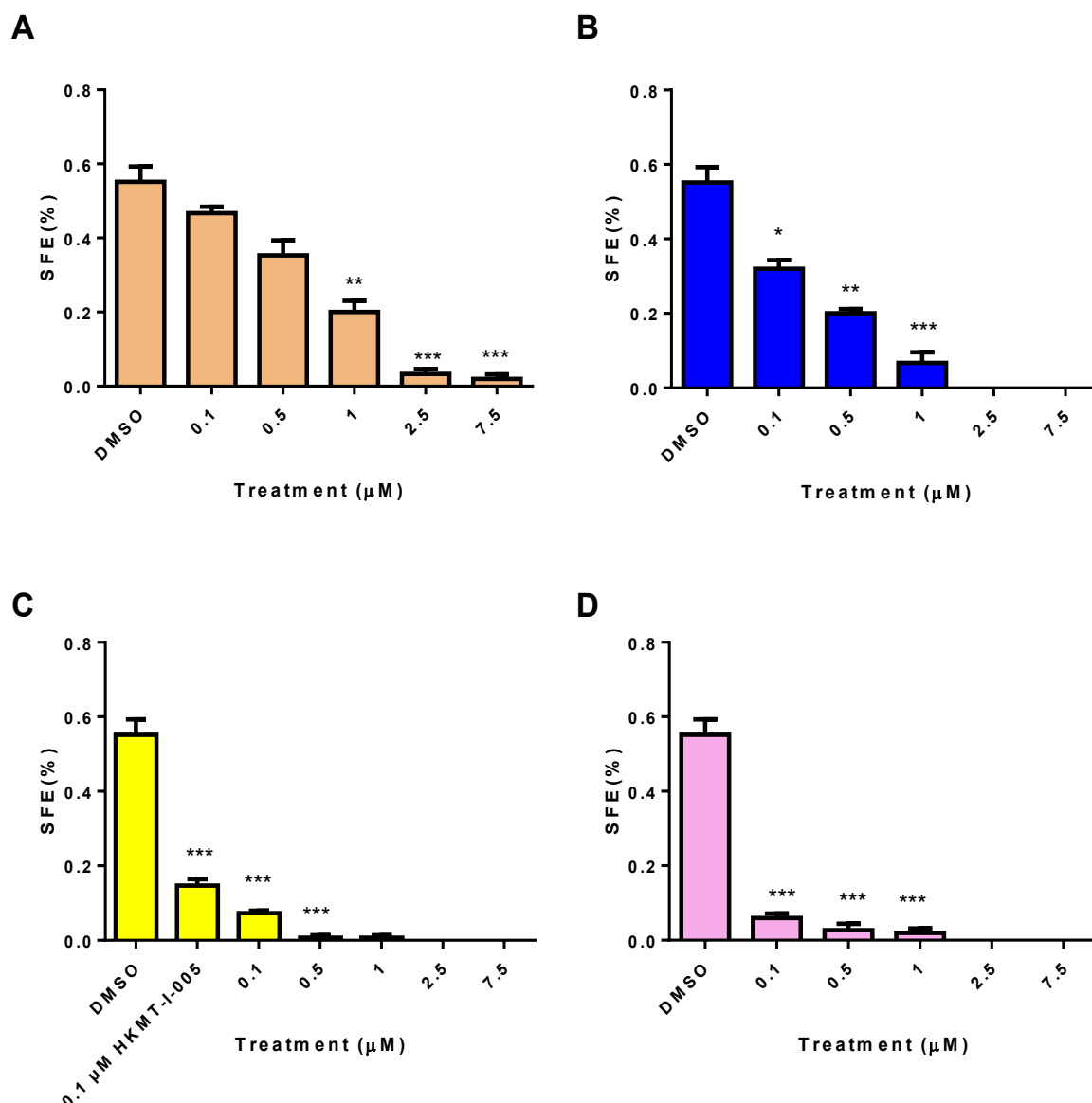


Figure 5.4 CSC activity as measured by spheroid formation efficiency in IGROV1 ovarian cancer cells after treatment with A) Paclitaxel B) Cisplatin C) Paclitaxel +1 μM HKMT-I-005 D) Cisplatin +1 μM HKMT-I-005 – statistical significance calculated by Student’s T-test between DMSO control and dose- $p < 0.05 = *$, $p < 0.01 = **$, $p < 0.001 = ***$. Error bars are SEM ($n \geq 3$).

Notably GSK343 has a very weak impact on clonogenic capacity of the cell population, but a significant strong impact on CSC activity at much lower doses, supporting the theory that the CSC population are reliant on the maintenance of EZH2 levels- all of the HKMT inhibitors had $\text{IC}_{50} < 0.3 \mu\text{M}$ in when affecting IGROV1 CSC activity.

Having corroborated the previous findings that IGROV1 ovarian cancer stem cells were reliant on EZH2 expression to maintain CSC activity, the impact of HKMT inhibitors on the CSC population in the MDA-MB-231 breast cancer cell line was explored.

As a baseline for comparison, clonogenic assays were performed (**Materials & Methods: Clonogenic assay**) in MDA-MB-231 breast cancer cells.

Cells were treated with DMSO as a control, HKMT-I-005, HKMT-I-011, GSK343, UNC0638 and the chemotherapeutic agent Paclitaxel.

HKMT-I-005, HKMT-I-011, and UNC0638 all lead to a significant reduction in colony formation of MDA-MB-231 breast cancer cells at doses as low as 1 μ M (Fig.5.5.A, B, D respectively) and caused a decrease in clonogenic capacity in a dose dependent manner.

Paclitaxel very significantly reduced colony formation in MDA-MB-231 breast cancer cells, with >95% reduction in colony formation at doses >1 μ M (Fig.5.5 E).

EZH2 specific inhibitor GSK343 had significantly reduced colony formation at doses >2.5 μ M, although colony formation was >80% of that seen in the control doses after treatment with GSK343 up to 15 μ M (Fig.5.5 C) - statistical significance in this experiment was established by unpaired 2-tailed Student's t-test relative to DMSO control.

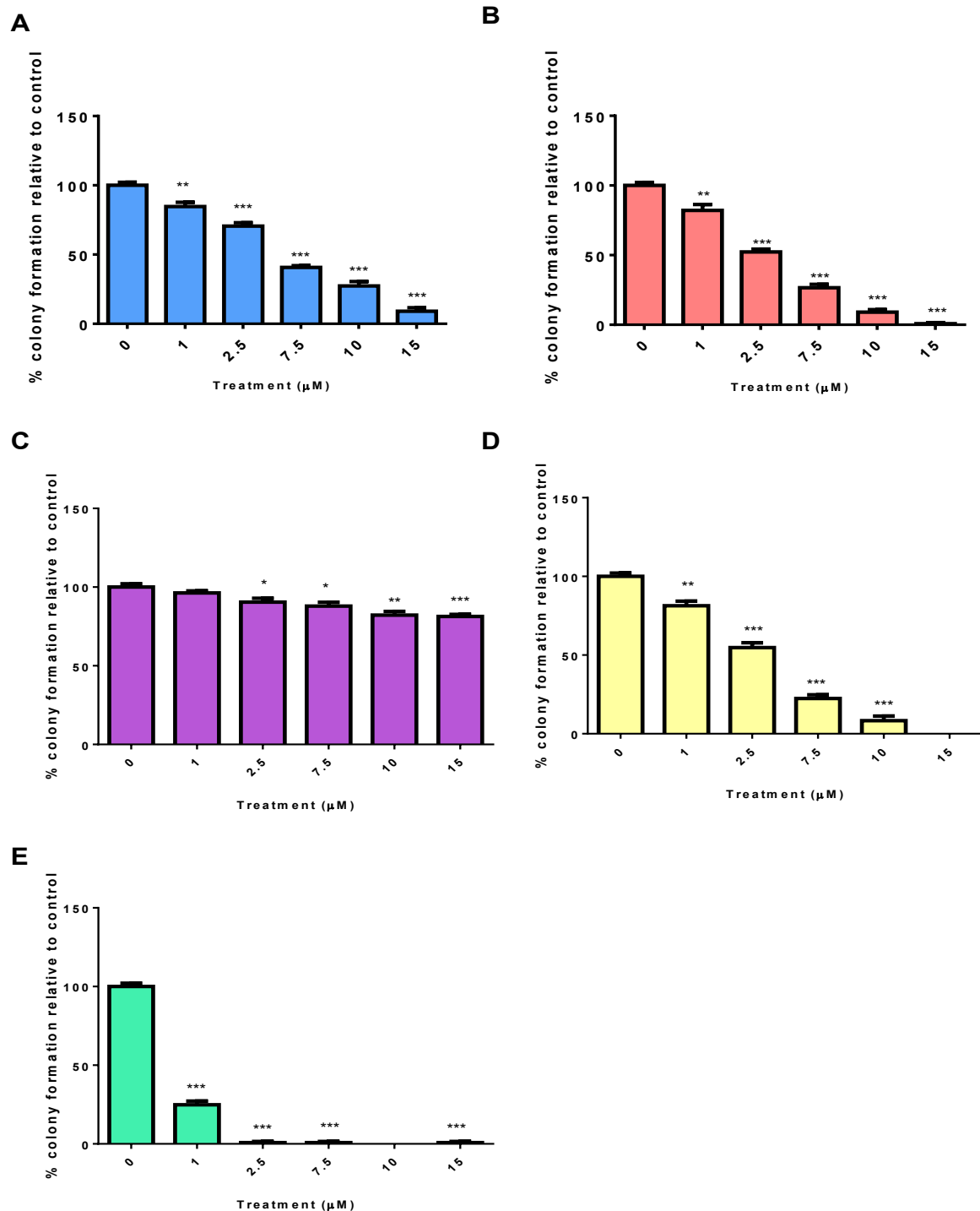


Figure 5.5 Clonogenic activity as measured by colony formation in MDA-MB-231 breast cancer cells after treatment with A) HKMT-I-005 B) HKMT-I-011 C) GSK343 D) UNC0638 E) Paclitaxel– statistical significance calculated by Student’s T-test between DMSO control and dose- $p<0.05=*$, $p<0.01=**$, $p<0.001=***$. Error bars are SEM ($n\geq 3$).

Calculated IC₅₀ doses for these treatments (Table 5.6) show that GSK343 is far less effective than HKMT-I-005, HKMT-I-011, or UNC0638 at inhibiting colony formation, and as in the earlier IGROV1 data Paclitaxel shows a predictably low IC₅₀ in the reduction of bulk clonogenic capability.

Table 5.6- Clonogenic IC₅₀ of treatments in MDA-MB-231 breast cancer cells

	HKMT-I-005	HKMT-I-011	GSK343	UNC0638	PACLITAXEL
IC ₅₀ (μM)	3.34	1.075	76.69	1.015	<0.01

Having established the impact of these inhibitors and agents on the clonogenic capacity of the MDA-MB-231 cell population, the impact on CSC activity was explored (**Materials & Methods: CSC activity and self-renewal capacity**). In comparison to published data using MDA-MB-231 cells in this assay ¹²¹, a similar degree of CSC activity in the MDA-MB-231 cells was observed (as measured by mammosphere formation efficiency (MFE)) under control conditions (MFE~1%), and a similar morphology of resulting mammospheres was observed (Supplementary Figure 8.4).

Treatment with HKMT-I-005 and HKMT-I-011 led to a significant reduction in CSC activity (Fig.5.6 A/B) at doses >1μM. GSK343 showed a significant decrease in CSC activity after 0.5μM (Fig.5.6 C), and UNC0638 showed a significant decrease in MDA-MB-231 CSC activity after treatment with 0.1μM (Fig.5.6 D).

Calculated IC₅₀ values (Table 5.7) indicate that CSC activity is more susceptible than the overall MDA-MB-231 cell population to treatment with HKMT-I-005, GSK343, and UNC0638- HKMT-I-011 showed similar IC₅₀ values for clonogenic capacity and CSC activity in the MDA-MB-231 cells. As observed in the IGROV1 cells, GSK343 shows strong inhibition of CSC activity but is relatively incapable of arresting clonogenic capacity in the total cell population.

Table 5.7- CSC IC50 of treatments in MDA-MB-231 breast cancer cells as measured by MFE

	HKMT-I-005	HKMT-I-011	GSK343	UNC0638
IC50 (μ M)	1.939	5.978	0.2529	1.783

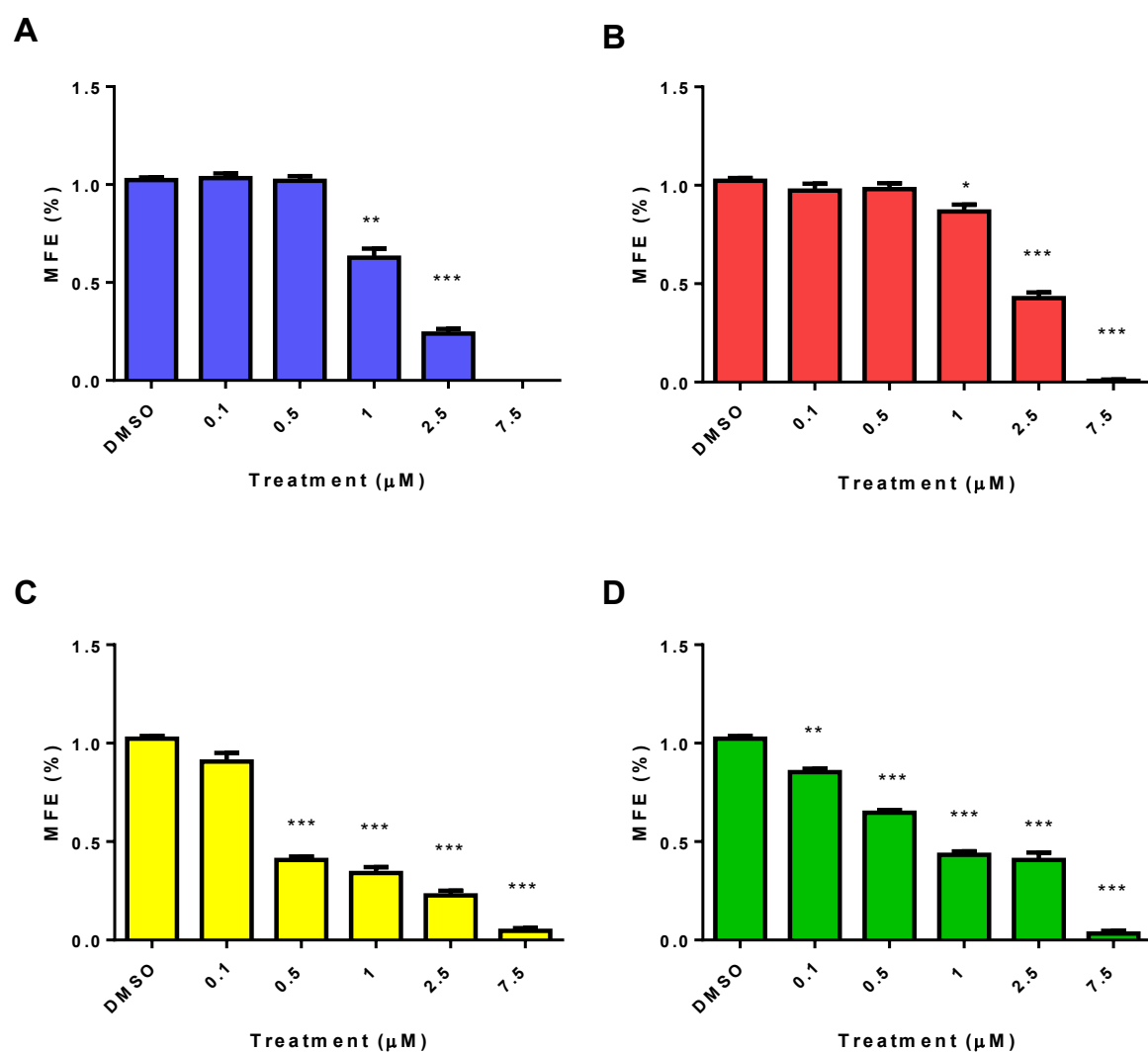


Figure 5.6 CSC activity as measured by mammosphere formation efficiency in MDA-MB-231 breast cancer cells after treatment with A) HKMT-I-005 B) HKMT-I-011 C) GSK343 D)UNC0638 – statistical significance calculated by Student's T-test between DMSO control and dose- $p < 0.05 = *$, $p < 0.01 = **$, $p < 0.001 = ***$. Error bars are SEM ($n \geq 3$).

This assay was repeated after treatment with mitotic inhibitor Paclitaxel or alkylating agent Cisplatin, as well as Paclitaxel with co-treatment with 1 μ M HKMT-I-005, and Cisplatin with 24 hour pre-treatment with 1 μ M HKMT-I-005 (Fig.5.7).

Paclitaxel treatment actually led to an increase in the CSC activity observed in MDA-MB-231 cells (Fig.5.7A). Cisplatin treatment led to a reduction of up to 35% (relative to DMSO control) but this effect appeared to plateau at doses >0.5 μ M and no further decrease in CSC activity was observed up to doses of 7.5 μ M.

When cells were treated with 1 μ M of HKMT-I-005 (which should cause ~50% decrease in CSC activity (Fig 5.5 A)) and Paclitaxel at the same time, a significant decrease in CSC activity was observed (Fig.5.4.5 C), leading to a very significant ($p<0.001$) reduction in CSC activity upon increasing levels of Paclitaxel treatment.

Similarly, whilst Cisplatin treatment plateaued at ~25% decrease in CSC activity relative to control from 0.5-7.5 μ M (Fig.5.7 B), 24 hours of treatment with 1 μ M of HKMT-I-005 prior to cisplatin treatment led to a dose dependent decrease of CSC activity across this same dose range. This preliminary data indicates that HKMT-I-005 treatment may sensitise the CSC population in MDA-MB-231 cells to treatment with conventional chemotherapies Paclitaxel and Cisplatin. Indeed, co-treatment with Paclitaxel and HKMT-I-005 led to a calculated IC₅₀ of 2.697 μ M (Supplementary table 8.13), compared to Paclitaxel treatment alone which did not inhibit MDA-MB-231 CSC activity.

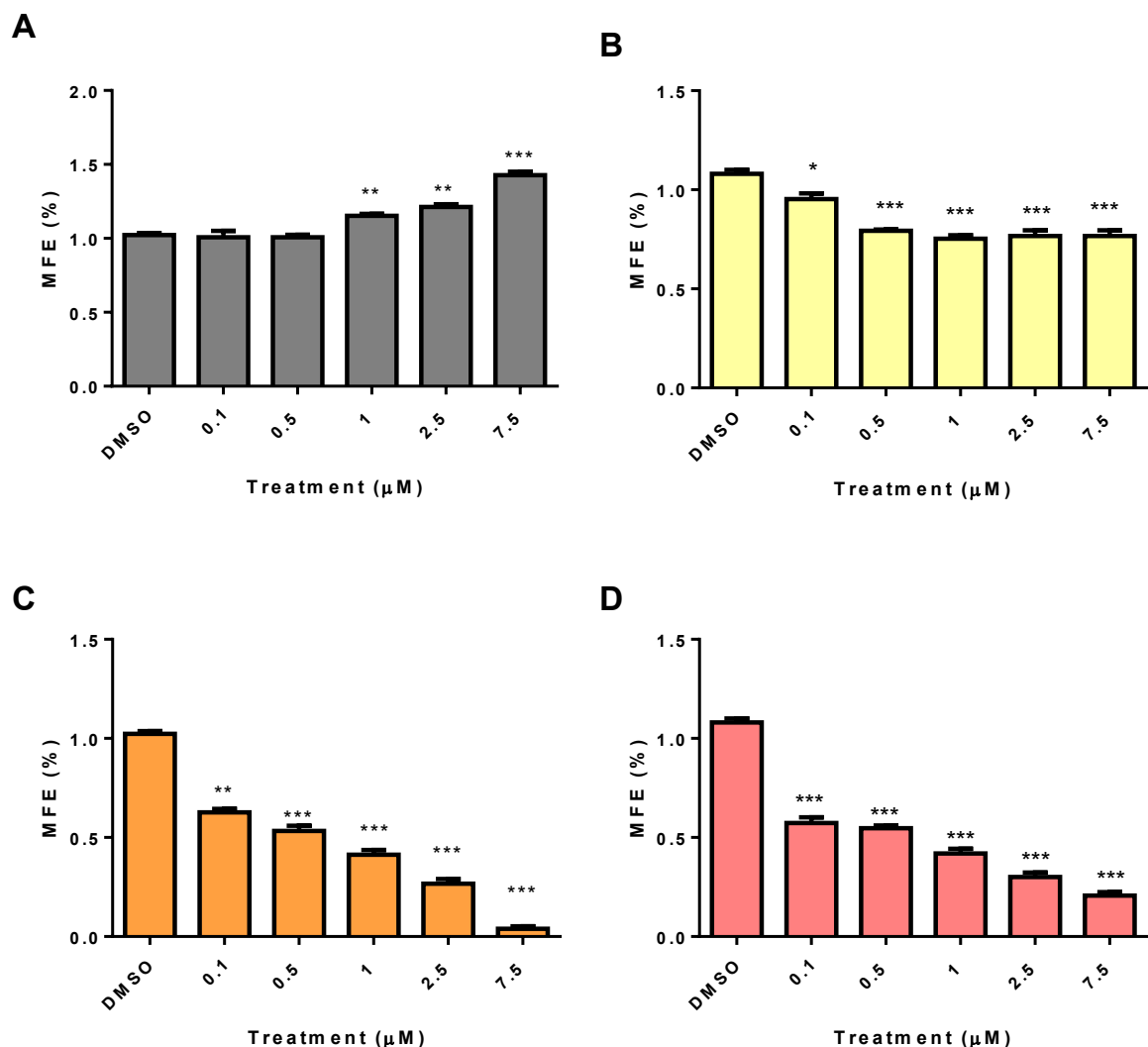


Figure 5.7 CSC activity as measured by mammosphere formation efficiency in MDA-MB-231 breast cancer cells after treatment with A) Paclitaxel B) Cisplatin C) Paclitaxel +1μM HKMT-I-005 D) Cisplatin +1μM HKMT-I-005 – statistical significance calculated by Student’s T-test between DMSO control and dose- $p < 0.05 = *$, $p < 0.01 = **$, $p < 0.001 = ***$. Error bars are SEM ($n \geq 3$).

Preliminary data examining the long term self-renewal capacity of the CSC population in MDA-MB-231 cells was also generated (**Materials & Methods: CSC activity and self-renewal capacity**). This method allows the examination of CSC self-renewal capacity by disaggregating and re-plating 1st generation mammospheres to create a 2nd generation- the impact of treatments during 1st generation on the formation of a 2nd generation indicate the degree of self-renewal capacity present in the CSC cells ¹²¹.

Cells were treated with HKMT inhibitors, chemotherapeutic agents, or a combination of both (Table 5.8) upon the commencement of 1st generation formation- these doses were chosen based upon the IC₅₀ on CSC activity in the case of the HKMT inhibitors, and based upon a dose showing a ~50% reduction in CSC activity after co-treatment of chemotherapeutic agents with HKMT-I-005. Cells were then disaggregated and re-plated (as per **Materials & Methods: CSC activity and self-renewal capacity**) and CSC self-renewal capacity relative to DMSO control was established.

Table 5.8- Treatment of MDA-MB-231 cells

Treatment	Dose (μM)
HKMT-I-005	1
HKMT-I-011	2.5
GSK343	0.25
UNC0638	1.5
Paclitaxel	2.5
Paclitaxel + HKMT-I-005	2.5, co-treatment with 1μM HKMT-I-005
Cisplatin	2.5
Cisplatin + HKMT-I-005	2.5, 24 hours pre-treatment with 1μM HKMT-I-005

Treatment with HKMT-I-005 (either upon plating or prior to plating) at 1μM or HKMT-I-011 at 2.5μM completely ablates the capacity of the CSCs to self-renew (Fig.5.8 A) - UNC0638 and GSK343 also showed a dramatic reduction in CSC self-renewal capacity.

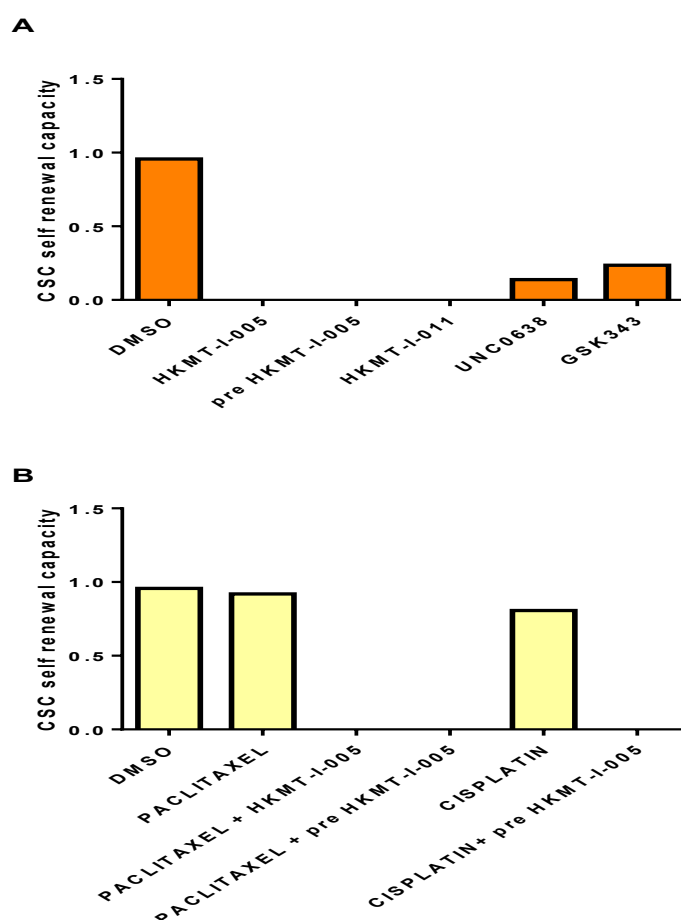


Figure 5.8 CSC self-renewal as measured by 2nd generation mammosphere formation capacity in MDA-MB-231 breast cancer cells after treatment with A) HKMT inhibitors B) Chemotherapeutic agents +HKMT-I-005–(prefix ‘pre’ denotes 24 hour treatment with HKMT-I-005 prior to 1st generation plating) (n=1)

Chemotherapeutic agents Paclitaxel and Cisplatin did not reduce CSC self-renewal capacity at the doses interrogated, though complementary treatment with HKMT-I-005 again completely ablated the CSC self-renewal capacity in the MDA-MB-231 cells (Fig.5.8 B).

EZH2/EHMT2 inhibition significantly reduces CSC activity in MDA-MB-231 cancer cells, and Paclitaxel treatment, whilst efficacious at reducing the clonogenic capacity of these cells, does not reduce this CSC activity. Supplementing Paclitaxel with treatment of HKMT-I-005 appeared to sensitise the CSCs to the action of Paclitaxel.

The mechanism of this sensitisation of CSC cells to Paclitaxel was investigated- using the lists of differentially expressed genes generated during the microarray study after treatment of MDA-MB-231 cells with HKMT-I-005 (**Materials & Methods: Gene expression microarray**), genes known to be involved in the Taxane pathway¹⁴² were compared against lists of differentially expressed genes to see if there was any overlap (Genes related to the Taxane pathway that showed differential expression after HKMT-I-005 treatment shown in Supplementary table 8.14).

In the Taxane pathway, Taxanes such as Paclitaxel block cell division by binding to β -tubulin, stabilizing the microtubules- this leads to cell death. Paclitaxel has been shown to induce BCL2 (which regulates cell death by controlling the mitochondrial membrane permeability) – HKMT-I-005 treatment increases BCL2 expression (Supplementary table 8.14).

Paclitaxel is also linked to expression of Cytochromes P450- a group of enzymes involved in the metabolism of drugs. HKMT-I-005 treatment of MDA-MB-231 cells led to a decrease in CYP3A43 expression, and an increase in CYP1B1 expression (Supplementary table 8.14) - this alteration in the expression of drug metabolising enzymes may explain the sensitisation of CSC cells to Paclitaxel treatment after HKMT-I-005 treatment.

Another group of genes that show altered expression is ABC drug transporters- (Supplementary table 8.14). Here, several of these genes are upregulated, but many more show a decrease in expression after HKMT-I-005 treatment- as ABC transporters are responsible for the transportation of Paclitaxel from the cell¹⁴² and CSC populations have previously been characterised as highly expressing these genes⁴⁷, this is potentially the avenue through which the observed sensitisation is occurring.

Also of note is the decreased expression of several TUBB genes which encode tubulin, the substrate Paclitaxel targets to affect its chemotherapeutic role (Supplementary table 8.14) -

decreasing expression of these genes (and as such altering the ratio of Paclitaxel and its target substrate) may impact the sensitivity of the CSC cells to Paclitaxel.

Based upon this data, further study of combined treatment with Paclitaxel and HKMT-I-005 was taken forward into mouse xenograft models, to interrogate if this effect was replicable *in vivo*.

5.4 Effect of dual HKMT inhibition on cancer stem cell activity, self-renewal, and chemosensitivity in *in vivo* models

In collaboration with **Gillian Farnie** and **Amrita Shergill** at the University of Manchester a series of experiments were completed investigating the impact of HKMT-I-005 in combination with Paclitaxel upon CSC activity using MDA-MB-231 xenografts.

The effect of treatment upon tumour size was investigated (**Materials & Methods: Xenograft study**) after treatment with HKMT-I-005, Paclitaxel, a combination of the two, or a DMSO control. Analysis by Two-way Anova showed no significant difference between these treatments in the fold change in tumour size (Fig.5.9).

Tumours from this study were extracted, disaggregated, and CSC activity was assayed by mammosphere formation efficiency (**Materials & Methods: CSC activity and self-renewal capacity**). The change in mammosphere formation (Fig.5.10) from the cells extracted from treated tumours relative to DMSO treated tumours (n=6 tumours per treatment, with 6 replicates per tumour) - one-way Anova across all samples $p < 0.0001$ indicating a significant difference between these treatment arms.

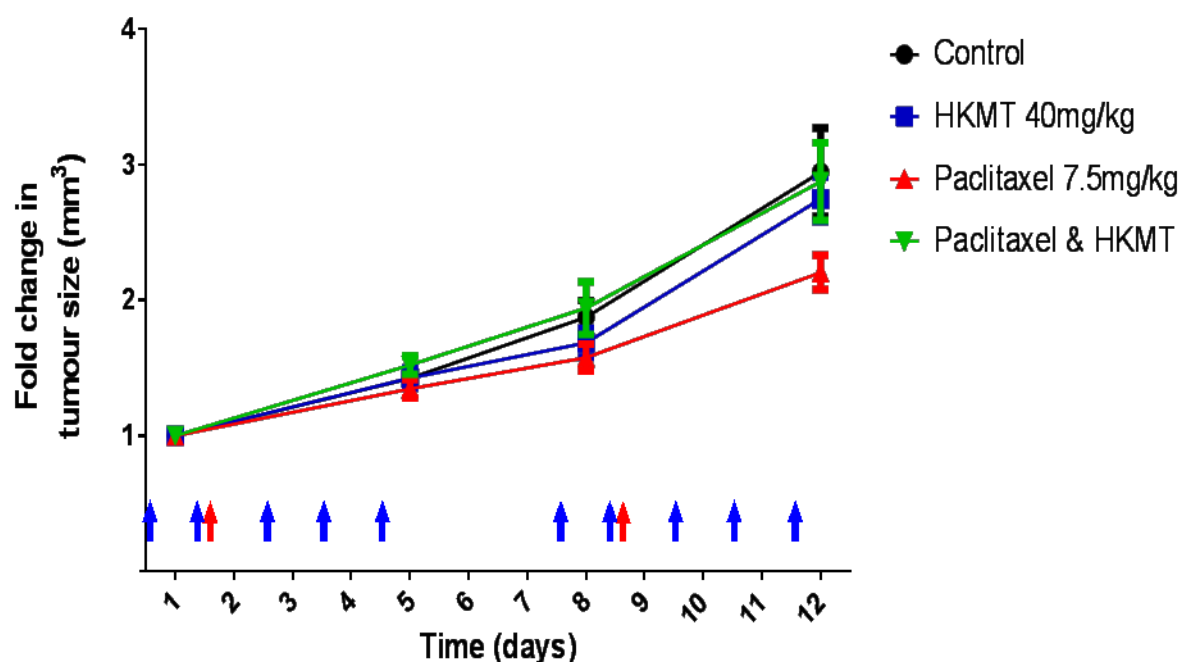


Figure 5.9- Tumour size of MDA-MB-231 xenografts after treatment with DMSO control, HKMT-I-005 (HKMT), Paclitaxel, or Paclitaxel + HKMT-I-005- Experiment in collaboration with Gillian Farnie and Amrita Shergill. N=6, statistical significance calculated by Two-way Anova. Error bars show SEM.

Paclitaxel in combination with HKMT-I-005 significantly reduces CSC activity relative to DMSO control treatments ($p=0.001$), treatment with Paclitaxel alone ($p=0.01$), or treatment with HKMT-I-005 alone ($p=0.01$) – a ~40% decrease in CSC activity as measured by MFE was observed after dual treatment with HKMT-I-005 and Paclitaxel.

Cells were taken from the treated xenograft tumours from the initial experiment and 10 or 5 cells re-injected sub-cutaneously into the flank to create a second generation of tumours (**Materials & Methods: Secondary xenograft culture**). These second-generation tumours received no-further treatment, and tumour size was measured over 7 weeks.

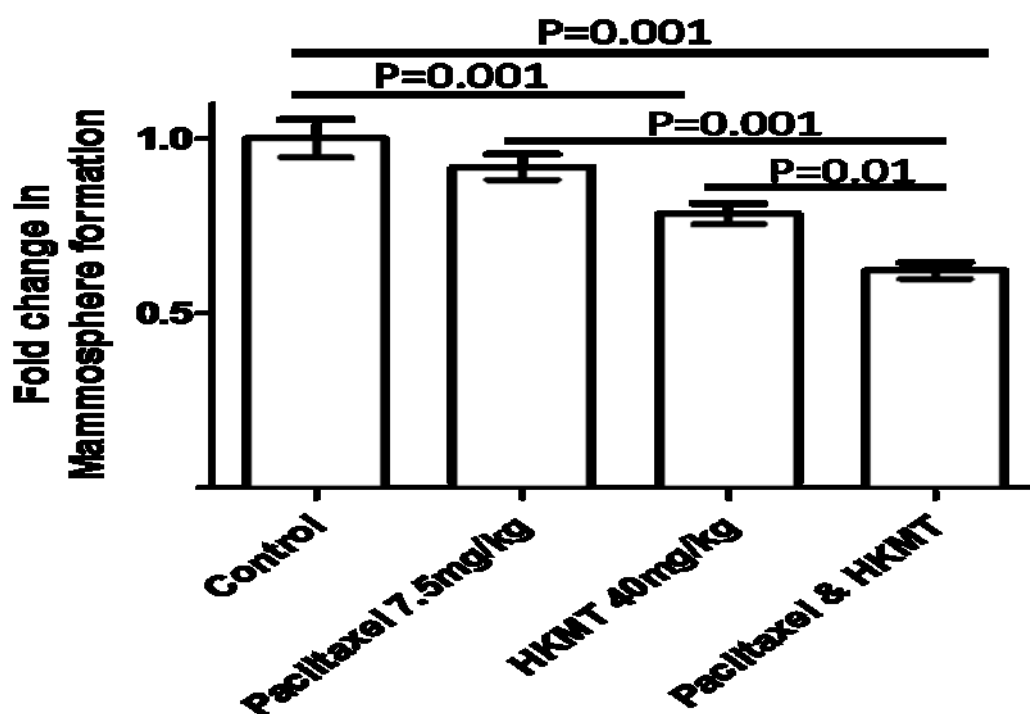


Figure 5.10- Relative change in CSC activity (as measured by MFE) of MDA-MB-231 xenografts after treatment with DMSO control, HKMT-I-005 (HKMT), Paclitaxel, or Paclitaxel + HKMT-I-005 - Experiment in collaboration with Gillian Farnie and Amrita Shergill. n=6 tumours per treatment, with 6 replicates per tumour- one-way Anova across treatments was performed to ascertain statistical significance.. Error bars show SEM.

Upon injection of 10 cells from the initial xenograft study, tumours were formed from cells from each treatment (Fig.5.11 A) - at week 7, Two-way Anova analysis (with Bonferroni multiple comparison) was performed comparing tumour size from cells taken from xenografts which had received different treatments in their first generation- in comparison to DMSO or Paclitaxel treatment alone, Paclitaxel + HKMT-I-005 showed very significantly ($p < 0.0001$) lower tumour volumes, also significantly ($p < 0.01$) lower tumour size than the HKMT-I-005 treatment alone.

Similarly, after injection of 5 cells (Fig.5.11 B), cells treated with Paclitaxel and HKMT-I-005 in combination grew significantly smaller tumours than Paclitaxel ($p < 0.001$), DMSO ($p < 0.001$), or HKMT-I-005 alone ($p < 0.01$).

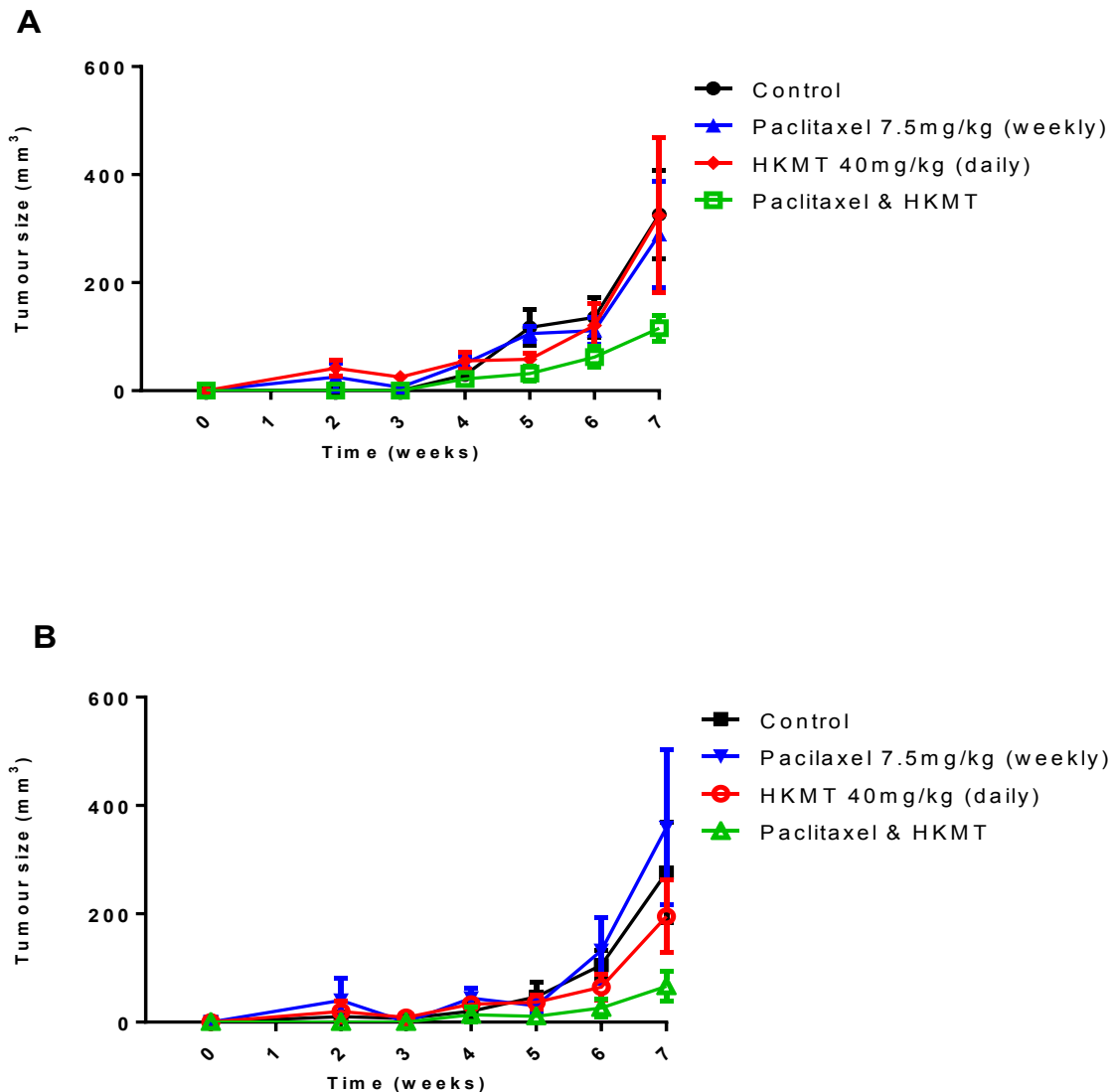


Figure 5.11- Second generation tumour formation after re-injection with A) 10 B) 5 cells from initial tumour study (Fig.5.9) –treatments refer to treatment of 1st generation tumours- 2nd generation tumours received no treatment - Experiment in collaboration with Gillian Farnie and Amrita Shergill. Statistical significance calculated by Two-way Anova ($n \geq 3$), error bars show SEM.

In an attempt to use this data to calculate the approximate number of CSCs in the treated tumours, extreme limiting dilution analysis was carried out (**Materials & Methods: Extreme limiting dilution analysis**) - initially the tumour take rate was established (where anything $< 100 \text{ mm}^3$ is not counted as a tumour growth). Tumour formation is shown at top of the graph (Fig.5.12) where a filled circle denotes tumour growth, and an empty circle signifies no tumour growth.

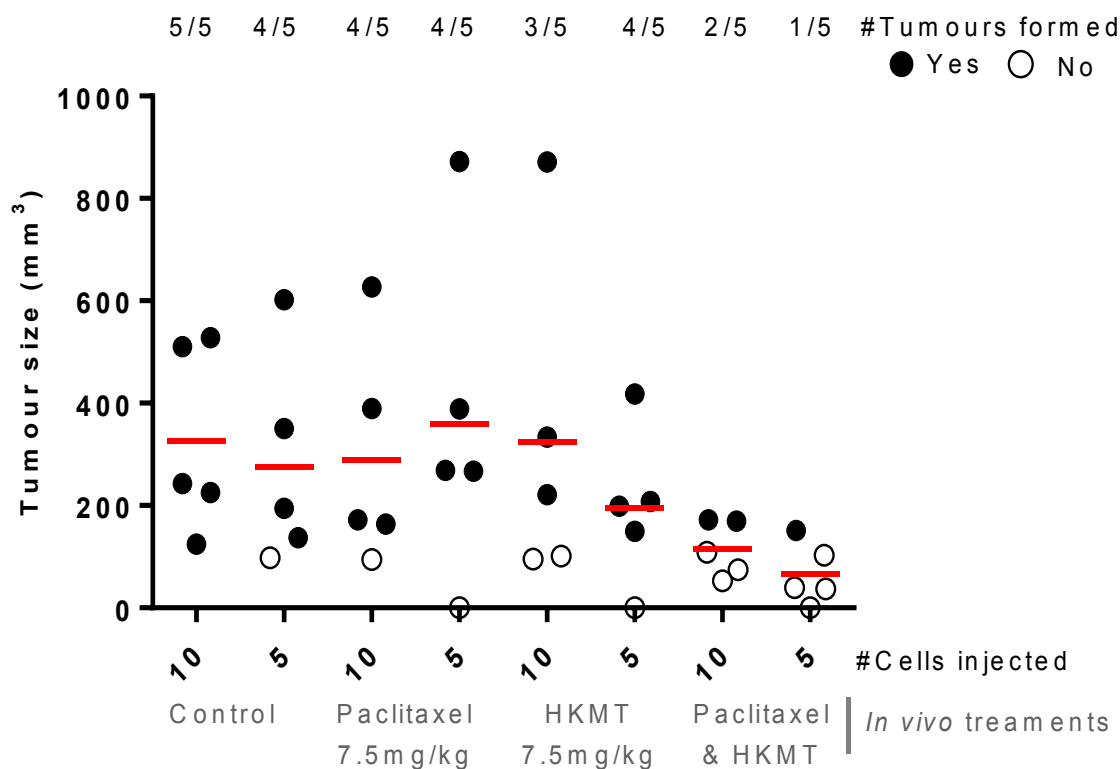


Figure 5.12- Second generation tumour formation after re-injection with 10 or 5 cells from initial tumour study (Fig.5.9) –treatments refer to treatment of 1st generation tumours- 2nd generation tumours received no treatment – any growth <100mm³ was not counted as tumour formation. Experiment in collaboration with Gillian Farnie and Amrita Shergill (n≥4)

As can be seen figuratively, some treatments led to the growth of more tumours in the second generation than others (full tumour take data quantified in supplementary table 8.15).

Using this data, extreme limiting dilution analysis (ELDA) gives an estimated confidence interval for the number of cancer stem cells (Table 5.9).

Table 5.9- ELDA confidence intervals for 1/stem cell frequency

	Lower	Estimate	Upper
Control (DMSO)	7.05	3.23	1.69
Paclitaxel	14.99	6.94	3.38
HKMT-I-005	10.62	5.07	2.59
Paclitaxel & HKMT-I-005	64.57	21.03	7.09

This data indicates that MDA-MB-231 cells re-injected after these treatments, in DMSO treated tumours ~1 in every 3 cells is a CSC; in HKMT-I-005 treated tumour cells ~1 in 7 cells is a CSC, in Paclitaxel treated cells ~1 in 5 cells is a CSC and in cells treated with a combination of Paclitaxel and HKMT-I-005 ~1 in every 21 cells is a CSC. Chi-square testing of all of these groups shows that there is a significant difference in the stem cell frequencies across the groups $P=0.0168$.

Further pairwise tests show the significance of the difference in cancer stem cell frequencies in these different experimental arms (Table 5.10).

Table 5.10- pairwise Chi-Square test between experimental arms CSC frequency

Group 1	Group 2	ChiSq	DF	Pr(>ChiSq)
Control (DMSO)	HKMT-I-005	2.12	1	0.146
Control (DMSO)	Paclitaxel	0.756	1	0.385
Control (DMSO)	Paclitaxel & HKMT-I-005	9.07	1	0.0026
HKMT-I-005	Paclitaxel	0.376	1	0.54
HKMT-I-005	Paclitaxel & HKMT-I-005	2.99	1	0.0837
Paclitaxel	Paclitaxel & HKMT-I-005	5.23	1	0.0221

Paclitaxel and HKMT-I-005 has significantly lower CSC frequency that the DMSO control ($p=0.0026$) or Paclitaxel alone ($p=0.0221$).

5.5 Summary

Numerous inhibitors of EZH2 or the action of EZH2 have been discovered (Chapter 1, Table 1.2), and primarily characterised in their ability to kill Y646n EZH2 mutant lymphoma cells. HKMT-I-005 potently reduces lymphoma cell line proliferation, and also inhibits proliferation in numerous ovarian and breast cancer cell lines (though notably has a much higher IC50 in epithelial breast call line MCF10a than in breast cancer cell line).

HKMT-I-005, GSK343, and UNC0638 all strongly inhibit lymphoma cell line proliferation, supporting the future use of EZH2/EHMT2 inhibitors in these diseases. Especially of note is the

Y646N mutant EZH2 DB cell line, where all three of these inhibitors reduce cell proliferation with and $IC_{50} < 1 \mu M$. Notably, many of these lymphoma lines are EZH2 wild-type, indicating that mutation is not necessary to confer sensitivity to EZH2 inhibition- for example the SUDHL8 lymphoma cell line is EZH2 wild-type, but has an IC_{50} on cell proliferation $< 1 \mu M$ after treatment with HKMT-I-005 or UNC0638.

In MDA-MB-231 breast cancer cells, a combination of EZH2 and EHMT2 inhibition reduces cell proliferation more than inhibition of EZH2 or EHMT2 alone (notably, EZH2 inhibition alone has minimal impact on MDA-MB-231 cell proliferation (5.2)).

With regards to cancer stem cells, in IGROV1 ovarian cancer cells this group previously showed that reduction of EZH2 leads to a reduction in CSC activity- here a reduction in EZH2, EHMT2, or EZH2/EHMT2 leads to a reduction in CSC activity, at notably lower doses than those required to impact colony formation- this indicates the IGROV1 CSC population is more sensitive to HKMT inhibition than the rest of the cell population (5.2/3/4)- notably EZH2 inhibitor GSK343 showed minimal capacity to reduce colony growth, but drastically reduced CSC activity at very low doses in IGROV1 cells.

In MDA-MB-231 breast cancer cells, again GSK343 showed little ability to reduce colony formation, but significantly reduced CSC activity at very low doses. HKMT-I-005, HKMT-I-011, and UNC0638 all reduced colony formation significantly, and also showed a potent reduction in MDA-MB-231 CSC activity- notably, preliminary data shows HKMT-I-005 and HKMT-I-011 both completely eradicated CSC self-renewal capacity. Also, intriguingly, treatment with HKMT-I-005 sensitised MDA-MB-231 cells to Paclitaxel and Cisplatin, both of which showed either no or very little inhibition of CSC activity when used as single agent treatments.

Investigation of previously acquired gene expression data after HKMT-I-005 treatment in the MDA-MB-231 cells (Chapter 4, 4.2) with regards to genes known to be associated with the mechanism of Paclitaxel treatment highlighted several potential pathways (5.3) through which this sensitisation may occur (namely reduced expression of ABC drug transporters and tubulin encoding genes).

These studies together indicate that disruption of EZH2, EHMT2, or EZH2/EHMT2 together leads to a reduction in CSC activity- however; EHMT2 inhibitors and dual EZH2/EHMT2 inhibitors have more of an impact on colony formation in IGROV1/MDA-MB-231 cells than EZH2 specific inhibition, supporting the use of dual inhibitors in these cells.

Taking this data forward in collaboration with Gillian Farnie and Amrita Shergill a series of *in vivo* experiments (5.4) were performed to look at the effect of HKMT-I-005 and Paclitaxel at reducing CSC activity *in vivo*. These experiments showed HKMT-I-005 reducing CSC activity in cells extracted from xenograft studies, and upon secondary implantation these cells treated by HKMT-I-005 and Paclitaxel were less likely to form secondary tumours, and the tumours that did form were smaller than those formed from cells treated with Paclitaxel or HKMT-I-005 alone.

The reportedly specific EHMT2 inhibitor UNC0638 shows similar capacity to inhibit CSC activity and impact cell proliferations as the novel dual HKMT inhibitors (though it did show a significantly lower upregulation of EZH2 repressed genes, Chapter 4, 4.2) – a more detailed comparison of this inhibitor and the novel dual HKMT inhibitors will be performed in Chapter 6.

These studies focused on anchorage independent growth (in the form of mammospheres and spheroids) as measures of CSC activity, as well as xenografts and Limiting Dilution Analysis.

In order to understand more about these findings it will be important to use other methods as well.

One manner of tracking the impact of treatments on the CSC population is through the use of markers for CSCs such as ALDH- numerous isoforms of ALDH have been linked to CSC activity ¹⁴³, and it is routinely used as a marker of CSCs. As such it can be used with experimental set-ups such as Fluorescence-activated cell sorting (FACS) to isolate a sub-population of CSCs ¹⁴⁴.

This capacity to isolate away the CSC from the bulk of cancer cells opens the door to comparisons between these cell types (such as limiting dilution analysis ¹⁴⁵). Additionally, treatment with potential CSC targeting agents can be observed in isolated populations of CSCs or non-CSC or mixed populations, and the results of this kind of experiment may highlight how specific the impact of these drugs are to the CSC setting.

The CSC model is now well established in many systems and the understanding of its complexity is increasing ¹⁴⁶- many protein markers specific to CSC populations (e.g. ¹⁴⁷) and enzymes linked to CSC activity (such as matrix metalloproteinases, which can impact the tumour microenvironment (MMPs) ^{148,149}). Further analysis of the impact of therapies theorised to target CSC populations should include analysis of genes, pathways, and proteins that are linked to the CSC phenotype.

Chapter 6: General discussion

Chapter 6: Discussion

6.1- Introduction

Aberrant EZH2 mediated epigenetic silencing has been observed in multiple cancer types and is linked to negative clinical outcomes and aggressive phenotypes, and based upon published literature it appears that this silencing is supported by the HKMT EHMT2 (summarised in Chapter 1, 1.3). Based upon these observations, dual inhibitors of EZH2 and EHMT2 were developed (Chapter 1, 1.5). At the beginning of this thesis, three aims were stated:

1. Utilise publicly available data to examine the degree to which EZH2/EHMT2 expression, CNV, and mutation status vary between cancer types and within cancer subtypes and patients to establish if stratification by EZH2/EHMT2 expression, CNV or mutations at a patient and disease level is viable
2. Characterise the impact of novel dual HKMT inhibitors on gene expression levels in cancer cell models, and examine how this relates to the chromatin state of target genes with regards to silencing marks H3K27me3 and H3k9me3
3. Examine the effect of dual HKMT inhibition on cancer cell phenotypes linked to HKMT expression (e.g. cancer stem cell activity, cancer cell proliferation, sensitivity to chemotherapeutic treatment)

These aims were based upon the hypothesis that by targeting both EZH2 and EHMT2 simultaneously, a greater reversal of EZH2 mediated epigenetic silencing would occur in

comparison to targeting EZH2 or EHMT2 individually. As such it was theorised that the identified dual HKMT inhibitors should have a stronger impact on HKMT mediated cancer cell phenotypes than individual HKMT inhibition of EZH2 or EHMT2.

To address these aims and assess this hypothesis a series of studies and experiments were performed, presented in this thesis in three chapters, with each chapter enclosing the data supporting one of the primary aims listed above.

This discussion will initially move through these chapters, highlighting significant data, comparing findings to published literature, commenting on study limitations, and positing potential avenues for future work based on the data presented, before moving onto a general discussion of the project and final conclusions and comments.

6.2- Evaluation of EZH2/EHMT2 as targets utilising publicly available data

6.2.1- Discussion

In an attempt to identify cancer types and subtypes that may benefit from dual HKMT inhibition publicly available datasets were interrogated- EZH2 and EHMT2 expression, CNV, and mutation status were all investigated across multiple datasets and cancer types in the hope of identifying targets for dual HKMT inhibition.

What immediately became apparent is that in normal tissue the expression of EZH2 and EHMT2 varies to a large degree across tissue type (Chapter 3, 3.2). Of note is the high expression of EZH2 in ES cells, haematopoietic stem cells, B-cells, T-cells, and most myeloid tissues- EHMT2 is also highly expressed in a number of tissues related to the immune system and haematopoietic system.

This data immediately raises the question of off-target effects. If dual HKMT inhibition may impact cells in the immune and haematopoietic systems then being aware of that as research moves forward is vital- strong off target effects in these systems could spell disaster from a drug development point of view. It has previously been shown that EZH2 specifically constrains differentiation and plasticity of Th1 and Th2 cells¹⁵⁰ and B-cell development¹⁵¹. In the haematopoietic system, EZH2 is known to prevent exhaustion of haematopoietic stem cell¹⁵². These systems are likely to respond to dual EZH2 and EHMT2 inhibition due to their high innate expression of EZH2 and EHMT2, and studying models of the immune and haematopoietic systems would provide indications if this response may be clinically relevant.

In the cancer setting EZH2 showed consistently high expression across numerous cancers relative to matched normal tissues (and EHMT2 also showed high expression in several cancers) – further, the level of EZH2 expression was linked to negative clinical outcomes (such as lymphatic invasion in colon cancer) and in breast cancer linked to poor RFS and OS,

highlighting breast cancer as a potential target for EZH2 inhibition, which supports published data linking EZH2 and EZH2 mediated changes in gene expression with harder to treat subtypes of breast cancer¹⁵³. The high expression of EZH2 and EHMT2 across a multiple cancer types and datasets indicates that inhibition of EZH2 and EHMT2 may impact on numerous cancer types. Where subtype information is available (such as ER-/ER+ breast cancer) the expression of these targets was not always consistent between subtypes. This highlights the need to stratify patient data using available clinical criteria in order to ascertain the best application of potential inhibitors of EZH2 and EHMT2.

Whilst mutation in EZH2 (notably Y641n, which occurs in~21% of DLBCL⁶ can lead to increased expression and sensitivity to EZH2 inhibition, in publicly available datasets mutations of EZH2 appear to be infrequent, never encompassing more than 5% of the cases within a cancer type- those cancer types with EZH2 expression (such as Renal, Ovarian, Brain, Thyroid, Adrenal, Colorectal, Lung, Breast, and Prostate (Fig.3.6) showed few or no reported mutations. As such expression levels of EZH2 may be a preferential indicator of treatment response rather than mutation status in most cases. Similarly, copy number amplification of EZH2 and EHMT2 was examined- this CNV did correlate with expression in some cases, but not all, and the strength of these correlations varied greatly between cancer types and subtypes- as such using CNV data to stratify patients could raise false positive results.

This data highlights the fact that EZH2 and EHMT2 expression consistently correlates across cancer types and subtypes. In addition, high EZH2 expression significantly correlated with negative clinical characteristics and outcomes in all of the datasets studied. Multiple cancer types show negative clinical characteristics and outcomes to be linked to expression of EZH2 and EHMT2. EZH2 and EHMT2 appear to be aberrantly regulated in multiple cancer types, and whilst differences between cancer types and sub-types may alter efficacy of treatment, targeted intervention with dual HKMT inhibitors has the potential to bring about significant clinical

impact. It is clear that expression of EZH2 and EHMT2 strongly positively correlate in numerous settings, further reinforcing the concept of their shared roles and potential redundancy.

The limitations of this large scale public data analysis included obvious bias toward common cancer types- in general, more statistical power is afforded to cancers with more data available, meaning rare cancers/subtypes are less likely to show significant findings. Using data collected from multiple sources always raises the issue of differences in sample and data collection and pre-processing, which could lead to heterogeneity across samples due to technical differences. As well as this, a degree of opacity exists in many public data portals as to what data normalisation and pre-processing techniques have already been applied.

With these caveats are kept in mind, this use of multiple publicly available datasets highlighted several cancers as a potential targets for dual HKMT inhibition (such as colon cancer/kidney renal clear cell) as well as supporting and corroborating the rationale for existing targets (such as breast cancer and ovarian cancer).

6.2.2- Future work

As further patient data becomes publicly available, interrogation of rare cancers and cancer subtypes will gain statistical power and could be interrogated again. In addition, as more data types become available, a greater variety of interrogation is available (e.g. MARMAL-AID, a publicly accessible database of genome-wide methylation data that currently has 88 tissue types¹⁵⁴). Maintaining awareness of new public datasets of different data types may allow interrogation of EZH2 and EHMT2 (and EZH2 targets and EHMT2 targets) in novel tissues or through different data types.

The mutation status, CNV, and differential expression of known targets of EZH2 and EHMT2 could be investigated in similar manner- an example of this would be the recent findings

showing tumours with BRG1 loss-of-function mutations or EGFR gain-of-function mutations confer a sensitivity to topoisomerase II inhibitors in non-small-cell lung cancer ¹⁵⁵, or the recently described synthetic lethality in ovarian cancer with ARID1A mutations, where EZH2 inhibition caused regression of ARID1A mutated tumours *in vivo* ¹⁴¹. Investigating mutations in EZH2 targets like these could highlight other cancer types and subtypes where EZH2 inhibition may be a viable treatment option, despite EZH2 expression/CNV/mutation itself.

6.3- Impact of novel dual HKMT inhibitors on the epigenetic state of cancer cells

6.3.1- Discussion

Using the MDA-MB-231 triple negative breast cancer cells, the impact of novel dual HKMT inhibitors on gene expression was studied. Microarray analysis confirmed that dual HKMT treatment increased expression of genes known to be repressed by EZH2, and that this effect was significantly stronger than the effect observed using an EZH2 specific inhibitor or an EHMT2 specific inhibitor. The novel dual HKMT inhibitors also significantly induced apoptosis related pathways (more than EZH2 inhibition alone).

The reversal of EZH2 mediated silencing was not observed when EZH2 targets were defined using a different breast cancer cell type. This indicates that either the impact of silencing EZH2 on gene expression differs between these cell types significantly, or that in these cells EZH2 repressed a different set of targets. To address this a list of genes consistently repressed by EZH2 across multiple cell types was generated. The dual HKMT inhibitors showed significant and strong upregulation of this list of EZH2 repressed targets, again to a greater degree than GSK343 or UNC0638.

An important aspect of this data was that it suggested there was capacity of a drug which failed the initial compound screen to inhibit EZH2 and increase expression of EZH2 repressed genes. This indicated that the initial compound screen (Chapter 1, 1.5) may be missing compounds that are potentially viable in this project. The identification of novel biomarkers such as SPINK1 should allow the ongoing drug discovery efforts to refine their search for dual inhibitors of EZH2 and EHMT2.

The mechanism for the observed reversal of EZH2 mediated silencing was explored in EZH2 repressed genes FBXO32, KRT17, and SPINK1, all of which showed reduction in H3K27me3 and H3K9me3 levels after treatment with dual inhibitors. Global reduction in these repressive

histone marks was also observed in the MDA-MB-231 cells, but further elucidation of the chromatin landscape around EZH2 repressed genes should be pursued further (see 6.3.2 Future work) to establish if this reversal of silencing is due to changes in the levels of H3K27me3 or H3K9me3.

EHMT2 inhibitor UNC0638 showed a significant capacity to upregulate EZH2 repressed genes- this could be due to two reasons- the first is that as theorised, EHMT2 plays an important supportive role in EZH2 mediated silencing by catalysing H3K27me1 and interacting physically with the PRC2 complex. If this is the case, it would appear that potent EHMT2 inhibition is enough alone to re-express EZH2 silenced genes (though to a lesser extent than dual EZH2/EHMT2 inhibition). The second possibility is that UNC0638 is itself inhibiting EZH2. UNC0638 is based upon the same quinazoline chemical template, and whilst it showed an $IC_{50} > 10 \mu M$ for EZH2 this was in biochemical screens⁹⁶ and it is unclear if EZH2 was measured as a lone substrate, or as part of the PRC2 complex. Certainly it is clear that UNC0638 did not cause expression of SPINK1 to increase and showed a different pattern of expression changes on FBXO32 and KRT17, but this is an issue which should be addressed further (see Future work 6.3.2).

6.3.2- Future work

The questions remaining include the effect of dual HKMT inhibitors on the chromatin landscape at a genome wide-level, and the impact of these dual HKMT inhibitors on EHMT2 repressed genes.

Whilst some EHMT2 siRNA knockdown experiments have been performed in breast cancer¹⁵⁶, the relative paucity of these studies means a meta-analysis to derive targets consistently repressed by EHMT2 was not feasible. By performing EHMT2 siRNA knockdown, EZH2 siRNA knockdown, and a combination siRNA knockdown of EZH2 and EHMT2 and

comparing resulting differential expression with that caused by the inhibitors a fuller picture of the changes in expression caused by dual HKMT inhibition could be gathered.

In conjunction with this, ChIP-seq studies examining the levels of H3K27me3 and H3K9me3 in the MDA-MB-231 cells after EZH2/EHMT2 inhibition would show how much of the differential expression is likely due to alterations in the levels of these repressive marks (as observed on individual target genes), or due to downstream effects. If possible, studying more histone marks such as H3K27me1, H3K4me3, H3K27ac, and H3K9ac would build up a more complete picture of the impact of dual EZH2/EHMT2 inhibition on the chromatin landscape of these cells.

6.4- Effect of dual HKMT inhibition on cancer cell phenotype and cancer stem cells

6.4.1- Discussion

The dual HKMT inhibitor HKMT-I-005 decreased proliferation in a number of breast, ovarian, and lymphoma, cell lines, notably in EZH2 Y641n mutant lymphoma line DB and ovarian cancer cell line A2780, which contains an ARID1A mutation (EZH2 inhibition caused regression of ARID1A mutated tumours *in vivo* ¹⁴¹). This inhibition was not limited to cell lines bearing known mutations conferring sensitivity to EZH2 inhibition- proliferation was decreased in multiple wild-type EZH2 cell lines. This is in contrast with published SAM-competitive EZH2 specific inhibitors, where impact on cell proliferation in wild-type EZH2 solid cancer cells has been limited (Chapter 1, 1.3) and focus has mainly been on the impact on EZH2 Y641n mutant DLBCL cells.

Having established an impact on cell proliferation of EZH2 wild-type cells, the impact of dual HKMT inhibition on the CSC population was investigated- in IGROV1 ovarian cancer cells and MDA-MB-231 breast cancer cells the dual HKMT inhibitors reduced CSC activity, and in the MDA-MB-231 cells showed a reduction in CSC self-renewal capacity and also sensitised the CSC population to Paclitaxel and Cisplatin treatment. The CSC population is a known driver of resistant regrowth ¹⁵⁷. Preliminary data hints that this sensitisation may be due to a decrease in the expression of ABC transporters, but further work is needed to confirm this finding (see Future work 6.4.2).

The use of *in vivo* xenograft models to study the impact of HKMT-I-005 confirmed the capacity of dual HKMT inhibition to decrease CSC activity *in vivo*, and also supported the sensitisation of this population to the action of Paclitaxel.

One of the issues difficult to address in cancer lines or mouse xenograft models is that of cancer cell heterogeneity. Intra-tumoural cellular heterogeneity is well established in many cancer

models ¹⁵⁸ and the theory of clonal evolution combined with CSC theory indicates that at any point there may be multiple clonal sub-population, any of which may contain its own population of CSCs ¹⁴⁶ and addressing tumour heterogeneity with regards to CSC targeted treatment is a question worth addressing (see Future work 6.4.2).

Overall the impact of the dual HKMT inhibitors on the CSC population is pronounced in several cell types, and the biological rationale that EZH2 is essential for maintenance of the CSC population has been established in several cancer models (e.g. glioblastoma ⁴⁸, pancreatic cancer and breast cancer ⁴⁹), indicating dual HKMT inhibition may impact CSC activity in these settings.

6.4.2- Future work

To understand better how the CSC population differs from the cancer cell population within the MDA-MB-231 cells, gene expression microarrays or RNA-seq analysis of the CSC population compared to the cancer cell population in conjunction with ChIP-seq of H3K27me3 and H3K9me3 would help illuminate the mechanisms underlying the CSC population's sensitivity to EZH2 inhibition (these studies would include cells treated with dual HKMT inhibitors, for reasons explained below).

Expansion of the preliminary data on CSC long term self-renewal after HKMT inhibitor treatment to include more replicates and a greater range of treatment doses would allow a greater understanding of the longer term impact of these inhibitors.

The impact of the dual HKMT inhibitors in other CSC populations could be explored, and indeed collaborative work has begun at Imperial College in glioblastoma models with Nelofer Syed and pancreatic models with Fieke Froeling.

One of the key aspects for future exploration is the mechanism by which dual HKMT inhibition appears to sensitise the CSC population to Paclitaxel (observed *in vitro* and *in vivo*). Whilst a

decrease in ABC drug transporters was observed, this was only measured in the general cell population. Understanding how the gene expression and chromatin landscape of the CSC population changes after dual HKMT inhibition could help explain this observed effect.

6.5- General discussion and Conclusions

6.5.1- General discussion

This study has confirmed the initial aims in that dual HKMT inhibitors do appear more efficacious at reversing EZH2 mediated silencing than EZH2 or EHMT2 specific inhibitors. This mechanism of inhibition appears to reduce tumour cell proliferation in multiple models as well as targeting the clinically relevant CSC population.

EZH2 inhibition, however, could be a double edged sword. Whilst increased EZH2 expression is linked to metastasis, the CSC population, and many negative clinical outcomes and phenotypes, a loss of EZH2 expression has also been linked to some negative outcomes.

Loss of function EZH2 mutations have been reported in myelodysplastic syndrome ¹⁵⁹, and EZH2 acts with SUZ12 as tumour suppressor genes in T-cell acute lymphoblastic leukaemia's ¹⁶⁰. In addition somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-cell origin were identified, where this EZH2 Tyr641 is linked to reduced enzymatic activity *in vitro* ¹⁶¹. Recently it was shown that short term inhibition ablated glioblastoma tumour growth in a mouse xenograft model ¹⁶², but when this EZH2 inhibition was sustained EZH2-depleted tumours escaped growth arrest, and became relatively more proliferative than control tumours (though smaller in size).

Moving forward, this published data highlight the importance of vigilance when planning future dosing regimens. As the field of HKMT inhibition moves forward, the differing ways in which different cell populations react to altered HKMT levels is important to keep in mind. Whilst dual inhibition of EZH2 and EHMT2 shows promising results in reducing cancer cell proliferation, CSC activity, and potential sensitisation of the CSC population to chemotherapeutic agents, the effect of these inhibitors on other tissues is as yet unclear, and long term effects may be difficult to model or predict.

In this project the question remains if these HKMT inhibitors are exerting their impact on cell phenotype and gene expression via their theorised targets, or if this is due to off target effects. In order to establish that the dual HKMT inhibitors are acting as hypothesised, several points would need to be experimentally proven:

- The dual HKMT inhibitors decrease the activity of the target HKMT in cells
- This decreased activity leads to an alteration in the epigenetic state and thus gene expression (and potentially further alteration of expression of downstream targets) of target genes
- This alteration in gene expression leads to the observed cellular phenotypes

As this project stands, the proposed dual HKMT inhibitors have been shown to selectively bind to the target HKMTs (Chapter 1, 1.5), and decrease the binding of EHMT2 and EZH2 with the cofactor SAM. Treatment with the dual HKMTi alters the expression of genes known to be regulated by EZH2 in MDA-MB-231 cell lines (Chapter 4, 4.2), and genes identified as EZH2 targets by a consensus pool of targets generated by multiple siRNA and shRNA knockdowns of EZH2.

This altered expression profile has been linked to a decrease in the levels of H3K27me3 and H3K9me3 at target genes SPINK1, KRT17, and FBXO32 (Chapter 4, 4.5) but this data is preliminary and not robustly validated. Many cellular phenotypic impacts have been observed (Chapter 5) after treatment with the dual HKMTi (e.g. decreased proliferation, decreased CSC activity).

Whilst this data indicates these novel compounds may be promising dual HKMT inhibitors, further work is needed- ChIP-seq studies would establish if this alteration in gene expression is primarily being driven by the hypothesised alteration in the epigenetic state. Artificially reducing EZH2 and EHMT2 levels (i.e. using siRNA) alone and in combination and studying

the impact on gene expression, chromatin state, and cellular phenotypes would allow experimental validation of these inhibitors. Until that point, whilst this data suggests these dual HKMT are impacting gene expression and cellular phenotype through the inhibition of EZH2 and EHMT2, it cannot be said for certain that none of the observed responses to treatment are due to off target effects of these drugs.

6.5.2- Conclusions

EZH2 expression is widely deregulated in numerous cancer types, independent of mutation or CNV, and this expression is correlated with negative clinical characteristics and poor clinical outcomes in breast cancer. Novel dual inhibitors of EZH2 and EHMT2 reverse EZH2 mediated gene repression to a greater degree than EZH2 or EHMT2 specific inhibition in breast cancer cells, inducing apoptotic pathways and reducing cell proliferation. These dual HKMT inhibitors inhibited CSC activity in wild-type EZH2 tumour cells (in both breast and ovarian cancer), and in breast cancer dual HKMT inhibitors sensitised the CSC population to treatment with Paclitaxel (*in vitro* and *in vivo*). The reversal of EZH2 mediated gene silencing is an established clinical target- based upon this data; we hypothesise that in certain cancer settings the application of dual HKMT inhibitors rather than EZH2 specific inhibitors may produce beneficial clinical results.

Chapter 7: List of references

1. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generation. *Cell* **144**, 646–674 (2011).
2. Feinberg, A. P. & Vogelstein, B. Hypomethylation of ras oncogenes in primary human cancers. *Biochem. Biophys. Res. Commun.* **111**, 47–54 (1983).
3. Egger, G., Liang, G., Aparicio, A. & Jones, P. a. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429**, 457–63 (2004).
4. Conaway, J. W. Introduction to theme ‘Chromatin, epigenetics, and transcription’. *Annu. Rev. Biochem.* **81**, 61–4 (2012).
5. Dawson, M. a & Kouzarides, T. Cancer epigenetics: from mechanism to therapy. *Cell* **150**, 12–27 (2012).
6. Morin, R. D. *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* **476**, 298–303 (2011).
7. Van Haaften, G. *et al.* Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. *Nat. Genet.* **41**, 521–3 (2009).
8. Cihák, A. Biological Effects of 5-Azacytidine in Eukaryotes. *Oncology* **30**, 405–422 (1974).
9. Kantarjian, H. *et al.* Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. *Cancer* **106**, (2006).
10. Richon, V. M. Cancer biology: mechanism of antitumour action of vorinostat (suberoylanilide hydroxamic acid), a novel histone deacetylase inhibitor. *Br. J. Cancer* **95**, S2–S6 (2006).
11. Nakajima, H., Kim, Y. B., Terano, H., Yoshida, M. & Horinouchi, S. FR901228, a potent antitumor antibiotic, is a novel histone deacetylase inhibitor. *Exp. Cell Res.* **241**, 126–133 (1998).
12. Burgess, D. J. Epigenetics: Chromatin inheritance during DNA replication. *Nat. Rev. Genet.* **13**, 675–675 (2012).
13. Margueron, R. *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**, 762–7 (2009).

14. Kondo, Y. *et al.* Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation. *Nat. Genet.* **40**, 741–50 (2008).
15. Ferrari, K. J. *et al.* Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Mol. Cell* **53**, 49–62 (2014).
16. Dillon, S. C., Zhang, X., Trievel, R. C. & Cheng, X. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol.* **6**, 227 (2005).
17. Miranda, T. B. *et al.* DZNep is a global histone methylation inhibitor that reactivates developmental genes not silenced by DNA methylation. *Mol. Cancer Ther.* **8**, 1579–1588 (2009).
18. Chiang, P. Biological effects of inhibitors of S-adenosylhomocysteine hydrolase. *Pharmacol. Ther.* (1998).
19. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039–43 (2002).
20. Højfeldt, J. W., Agger, K. & Helin, K. Histone lysine demethylases as targets for anticancer therapy. *Nat. Rev. Drug Discov.* **12**, 917–30 (2013).
21. Chen, Y.-H., Hung, M.-C. & Li, L.-Y. EZH2: a pivotal regulator in controlling cell differentiation. *Am. J. Transl. Res.* **4**, 364–75 (2012).
22. Viré, E. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439**, 871–874 (2006).
23. Surani, M. A., Hayashi, K. & Hajkova, P. Genetic and epigenetic regulators of pluripotency. *Cell* **128**, 747–62 (2007).
24. Cao, R., Tsukada, Y.-I. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell* **20**, 845–54 (2005).
25. Mallo, M. & Alonso, C. R. The regulation of Hox gene expression during animal development. *Development* **140**, 3951–63 (2013).
26. He, A. *et al.* PRC2 directly methylates GATA4 and represses its transcriptional activity. *Genes Dev.* **26**, 37–42 (2012).
27. Varambally, S. & Dhanasekaran, S. M. The polycomb group protein EZH2 is involved in progression of prostate cancer. **419**, 388–390 (2002).
28. Kleer, C. G. *et al.* EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11606–11 (2003).
29. Collett, K. *et al.* Expression of enhancer of zeste homologue 2 is significantly associated with increased tumor cell proliferation and is a marker of aggressive breast cancer. *Clin. Cancer Res.* **12**, 1168–74 (2006).

30. Bachmann, I. M. *et al.* EZH2 expression is associated with high proliferation rate and aggressive tumor subgroups in cutaneous melanoma and cancers of the endometrium, prostate, and breast. *J. Clin. Oncol.* **24**, 268–273 (2006).
31. Alford, S. H., Toy, K., Merajver, S. D. & Kleer, C. G. Increased risk for distant metastasis in patients with familial early-stage breast cancer and high EZH2 expression. *Breast Cancer Res. Treat.* **132**, 429–437 (2012).
32. Yehiely, F., Moyano, J. V., Evans, J. R., Nielsen, T. O. & Cryns, V. L. Deconstructing the molecular portrait of basal-like breast cancer. *Trends Mol. Med.* **12**, 537–544 (2006).
33. Raman, J. D. *et al.* Increased expression of the polycomb group gene, EZH2, in transitional cell carcinoma of the bladder. *Clin. Cancer Res.* **11**, 8570–8576 (2005).
34. Rao, Z. Y. *et al.* EZH2 supports ovarian carcinoma cell invasion and/or metastasis via regulation of TGF- β 1 and is a predictor of outcome in ovarian carcinoma patients. *Carcinogenesis* **31**, 1576–1583 (2010).
35. Hu, S. *et al.* Overexpression of EZH2 contributes to acquired cisplatin resistance in ovarian cancer cells in vitro and in vivo. *Cancer Biol. Ther.* **10**, 788–795 (2010).
36. Behrens, C. *et al.* EZH2 protein expression associates with the early pathogenesis, tumor progression, and prognosis of non-small cell lung carcinoma. *Clin. Cancer Res.* **19**, 6556–6565 (2013).
37. Kondo, Y. *et al.* Alterations of DNA methylation and histone modifications contribute to gene silencing in hepatocellular carcinomas. *Hepatol. Res.* **37**, 974–983 (2007).
38. Crea, F. *et al.* EZH2 inhibition: targeting the crossroad of tumor invasion and angiogenesis. *Cancer Metastasis Rev.* **31**, 753–61 (2012).
39. Lee, H. W. & Choe, M. Expression of EZH2 in renal cell carcinoma as a novel prognostic marker. *Pathol. Int.* **62**, 735–741 (2012).
40. Matsukawa, Y. *et al.* Expression of the enhancer of zeste homolog 2 is correlated with poor prognosis in human gastric cancer. *Cancer Sci.* **97**, 484–91 (2006).
41. Yamada, A. *et al.* Aberrant expression of EZH2 is associated with a poor outcome and P53 alteration in squamous cell carcinoma of the esophagus. *Int. J. Oncol.* **38**, 345–353 (2011).
42. Ougolkov, A. V., Bilim, V. N. & Billadeau, D. D. Regulation of pancreatic tumor cell proliferation and chemoresistance by the histone methyltransferase enhancer of zeste homologue 2. *Clin. Cancer Res.* **14**, 6790–6796 (2008).
43. Ma, R. *et al.* Inhibition of GSK 3 β Activity Is Associated with Excessive EZH2 Expression and Enhanced Tumour Invasion in Nasopharyngeal Carcinoma. *PLoS One* **8**, (2013).

44. Min, J. *et al.* An oncogene-tumor suppressor cascade drives metastatic prostate cancer by coordinately activating Ras and nuclear factor-kappaB. *Nat. Med.* **16**, 286–94 (2010).
45. Bracken, A. P. *et al.* EZH2 is downstream of the pRB-E2F pathway , essential for proliferation and ampli ® ed in cancer. **22**, 5323–5335 (2003).
46. Chen, K., Huang, Y. & Chen, J. Understanding and targeting cancer stem cells: therapeutic implications and challenges. *Acta Pharmacol. Sin.* **34**, 732–40 (2013).
47. Rizzo, S. *et al.* Ovarian cancer stem cell-like side populations are enriched following chemotherapy and overexpress EZH2. *Mol. Cancer Ther.* **10**, 325–35 (2011).
48. Suvà, M.-L. *et al.* EZH2 is essential for glioblastoma cancer stem cell maintenance. *Cancer Res.* **69**, 9211–8 (2009).
49. Van Vlerken, L. E. *et al.* EZH2 is required for breast and pancreatic cancer stem cell maintenance and can be used as a functional cancer stem cell reporter. *Stem Cells Transl. Med.* **2**, 43–52 (2013).
50. Yamaguchi, H. & Hung, M.-C. Regulation and Role of EZH2 in Cancer. *Cancer Res. Treat.* **46**, 209–22 (2014).
51. Yap, D. B. *et al.* Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood* **117**, 2451–2459 (2011).
52. Sneeringer, C. J. *et al.* Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 20980–20985 (2010).
53. Tan, J. *et al.* Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev.* **21**, 1050–63 (2007).
54. McCabe, M. T. *et al.* EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* (2012). doi:10.1038/nature11606
55. Knutson, S. K. *et al.* A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells. *Nat. Chem. Biol.* **8**, 890–6 (2012).
56. Knutson, S. K. *et al.* Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. *Proc Natl Acad Sci U S A* **110**, 7922–7927 (2013).
57. Qi, W. *et al.* Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. *Proc. Natl. Acad. Sci.* **109**, 21360–21365 (2012).
58. Verma, S. K. *et al.* Identification of potent, selective, cell-Active inhibitors of the histone lysine methyltransferase EZH2. *ACS Med. Chem. Lett.* **3**, 1091–1096 (2012).

59. Yoo, C. B. & Jones, P. A. Epigenetic therapy of cancer: past, present and future. *Nat. Rev. Drug Discov.* **5**, 37–50 (2006).
60. Cho, H.-S. *et al.* Enhanced Expression of EHMT2 Is Involved in the Proliferation of Cancer Cells through Negative Regulation of SIAH1. *Neoplasia* **13**, 676–684 (2011).
61. Ding, J. *et al.* The histone H3 methyltransferase G9a epigenetically activates the serine-glycine synthesis pathway to sustain cancer cell survival and proliferation. *Cell Metab.* **18**, 896–907 (2013).
62. Wu, H. *et al.* Histone methyltransferase G9a contributes to H3K27 methylation in vivo. *Cell Res.* **21**, 365–367 (2011).
63. Tachibana, M., Sugimoto, K., Fukushima, T. & Shinkai, Y. SET Domain-containing Protein, G9a, is a Novel Lysine-preferring Mammalian Histone Methyltransferase with Hyperactivity and Specific Selectivity to Lysines 9 and 27 of Histone H3. *J. Biol. Chem.* **276**, 25309–25317 (2001).
64. Yoo, K. H. & Hennighausen, L. EZH2 methyltransferase and H3K27 methylation in breast cancer. *Int. J. Biol. Sci.* **8**, 59–65 (2011).
65. Mozzetta, C. *et al.* The histone H3 lysine 9 methyltransferases G9a and GLP regulate polycomb repressive complex 2-mediated gene silencing. *Mol. Cell* **53**, 277–89 (2014).
66. Clarke, M. F. *et al.* Cancer stem cells--perspectives on current status and future directions: AACR Workshop on cancer stem cells. *Cancer Res.* **66**, 9339–44 (2006).
67. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* **3**, 730–737 (1997).
68. Singh, S. K. *et al.* Identification of a cancer stem cell in human brain tumors. *Cancer Res.* **63**, 5821–8 (2003).
69. Ignatova, T. N. *et al.* Human cortical glial tumors contain neural stem-like cells expressing astroglial and neuronal markers in vitro. *Glia* **39**, 193–206 (2002).
70. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3983–8 (2003).
71. Zhang, S. *et al.* Identification and characterization of ovarian cancer-initiating cells from primary human tumors. *Cancer Res.* **68**, 4311–20 (2008).
72. Ricci-Vitiani, L. *et al.* Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111–5 (2007).
73. Malanchi, I. *et al.* Cutaneous cancer stem cell maintenance is dependent on beta-catenin signalling. *Nature* **452**, 650–3 (2008).

74. Prince, M. E. *et al.* Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 973–8 (2007).
75. Eramo, A. *et al.* Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ.* **15**, 504–14 (2008).
76. Li, C. *et al.* Identification of pancreatic cancer stem cells. *Cancer Res.* **67**, 1030–7 (2007).
77. Schatton, T. *et al.* Identification of cells initiating human melanomas. *Nature* **451**, 345–9 (2008).
78. Collins, A. T., Berry, P. a, Hyde, C., Stower, M. J. & Maitland, N. J. Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Res.* **65**, 10946–51 (2005).
79. Beck, B. & Blanpain, C. Unravelling cancer stem cell potential. *Nat. Rev. Cancer* **13**, 727–38 (2013).
80. Grosse-Gehling, P. *et al.* CD133 as a biomarker for putative cancer stem cells in solid tumours: limitations, problems and challenges. *J. Pathol.* **229**, 355–78 (2013).
81. LaBarge, M. A. The difficulty of targeting cancer stem cell niches. *Clin. Cancer Res.* **16**, 3121–9 (2010).
82. Lacerda, L., Pusztai, L. & Woodward, W. A. The role of tumor initiating cells in drug resistance of breast cancer: Implications for future therapeutic approaches. *Drug Resist. Updat.* **13**, 99–108 (2010).
83. Dylla, S. J. *et al.* Colorectal cancer stem cells are enriched in xenogeneic tumors following chemotherapy. *PLoS One* **3**, e2428 (2008).
84. Bao, S. *et al.* Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756–60 (2006).
85. Diehn, M. *et al.* Association of reactive oxygen species levels and radioresistance in cancer stem cells. *Nature* **458**, 780–3 (2009).
86. Agarwal, R. & Kaye, S. B. Ovarian cancer: strategies for overcoming resistance to chemotherapy. *Nat. Rev. Cancer* **3**, 502–16 (2003).
87. Dean, M. ABC transporters, drug resistance, and cancer stem cells. *J. Mammary Gland Biol. Neoplasia* **14**, 3–9 (2009).
88. Chen, J. *et al.* A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* **488**, 522–6 (2012).
89. Sauvageau, M. & Sauvageau, G. Polycomb group genes: keeping stem cell activity in balance. *PLoS Biol.* **6**, e113 (2008).

90. Herrera-Merchan, A. *et al.* Ectopic expression of the histone methyltransferase Ezh2 in haematopoietic stem cells causes myeloproliferative disease. *Nat. Commun.* **3**, 623 (2012).
91. Chang, C.-J. *et al.* EZH2 promotes expansion of breast tumor initiating cells through activation of RAF1- β -catenin signaling. *Cancer Cell* **19**, 86–100 (2011).
92. Burdach, S. *et al.* Epigenetic maintenance of stemness and malignancy in peripheral neuroectodermal tumors by EZH2. *Cell Cycle* **8**, 1991–1996 (2014).
93. Crea, F. *et al.* Pharmacologic disruption of Polycomb Repressive Complex 2 inhibits tumorigenicity and tumor progression in prostate cancer. *Mol. Cancer* **10**, 40 (2011).
94. Nguyen, K. T. *et al.* Strategy to target the substrate binding site of SET domain protein methyltransferases. *J. Chem. Inf. Model.* **53**, 681–91 (2013).
95. Kubicek, S. *et al.* Reversal of H3K9me2 by a Small-Molecule Inhibitor for the G9a Histone Methyltransferase. *Mol. Cell* **25**, 473–481 (2007).
96. Vedadi, M. *et al.* A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. *Nat. Chem. Biol.* **7**, 566–574 (2011).
97. Yun, M., Wu, J., Workman, J. L. & Li, B. Readers of histone modifications. *Cell Res.* **21**, 564–578 (2011).
98. Srimongkolpithak, N., Sundriyal, S., Li, F., Vedadi, M. & Fuchter, M. J. Identification of 2,4-diamino-6,7-dimethoxyquinoline derivatives as G9a inhibitors. *Med. Chem. Commun.* **5**, 1821–1828 (2014).
99. Foletta, V. C., White, L. J., Larsen, A. E., Léger, B. & Russell, A. P. The role and regulation of MAFbx/atrogen-1 and MuRF1 in skeletal muscle atrophy. *Pflugers Arch.* **461**, 325–335 (2011).
100. Cailleau, R., Young, R., Olive, M. & Reeves, W. J. . J. Breast Tumor Cell Lines From Pleural Effusions. *J Natl Cancer Inst* **53**, 661–674 (1974).
101. Bénard, J. *et al.* Characterization of a human ovarian adenocarcinoma line, IGROV1, in tissue culture and in nude mice. *Cancer Res.* **45**, 4970–9 (1985).
102. Chomczynski, P. & Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**, 156–9 (1987).
103. Benita, Y. *et al.* Gene enrichment profiles reveal T-cell development, differentiation, and lineage-specific transcription factors including ZBTB25 as a novel NF-AT repressor. *Blood* **115**, 5376–84 (2010).
104. Smyth, G. in *Bioinforma. Comput. Biol. Solut. Using R Bioconductor* 397–420 (2005).

105. Feichtinger, J., McFarlane, R. J. & Larcombe, L. D. CancerMA: a web-based tool for automatic meta-analysis of public cancer microarray data. *Database (Oxford)*. **2012**, bas055 (2012).
106. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–4 (2012).
107. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pii (2013).
108. Mermel, C. H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
109. Mason, C. Cox proportional hazard models. *UC Berkeley* 1–6 (2005).
110. R Development Core Team, R. R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* **1**, 409 (2011).
111. Györfy, B., Lánckzy, A. & Szállási, Z. Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients. *Endocr. Relat. Cancer* **19**, 197–208 (2012).
112. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, Article3 (2004).
113. Lee, S. T. *et al.* Context-specific regulation of NF- κ B target gene expression by EZH2 in breast cancers. *Mol. Cell* **43**, 798–810 (2011).
114. Tan, J. *et al.* Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. *Genes Dev.* **21**, 1050–63 (2007).
115. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–10 (2002).
116. Breitling, R., Armengaud, P., Amtmann, A. & Herzyk, P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.* **573**, 83–92 (2004).
117. Choi, J. K., Yu, U., Kim, S. & Yoo, O. J. Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics* **19 Suppl 1**, i84–90 (2003).
118. Kamburov, A., Wierling, C., Lehrach, H. & Herwig, R. ConsensusPathDB--a database for integrating human functional interaction networks. *Nucleic Acids Res.* **37**, D623–8 (2009).
119. Nelson, J. D., Denisenko, O. & Bomsztyk, K. Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nat. Protoc.* **1**, 179–85 (2006).

120. Forbes, S. A. *et al.* COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43**, D805–811 (2014).
121. Shaw, F. L. *et al.* A detailed mammosphere assay protocol for the quantification of breast stem cell activity. *J. Mammary Gland Biol. Neoplasia* **17**, 111–7 (2012).
122. Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347**, 70–8 (2009).
123. Cao, R. & Zhang, Y. The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr. Opin. Genet. Dev.* **14**, 155–64 (2004).
124. Xiang, Y. *et al.* JMJD3 is a histone H3K27 demethylase. *Cell Res.* **17**, 850–7 (2007).
125. Sahasrabudhe, A. *et al.* Oncogenic Y641 mutations in EZH2 prevent Jak2/ β -TrCP-mediated degradation. *Oncogene* 1–10 (2014).
126. Kleer, C. G. *et al.* EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 11606–11611 (2003).
127. Holm, K. *et al.* Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns. *Breast Cancer Res* **12**, R36 (2010).
128. Hayashi, A. *et al.* Clinicopathological and prognostic significance of EZH2 expression in upper urinary tract carcinoma. *Virchows Arch.* **464**, 463–71 (2014).
129. Wu, Z. *et al.* Combined aberrant expression of Bmi1 and EZH2 is predictive of poor prognosis in glioma patients. *J. Neurol. Sci.* **335**, 191–196 (2013).
130. He, L.-J. *et al.* Prognostic significance of overexpression of EZH2 and H3k27me3 proteins in gastric cancer. *Asian Pac. J. Cancer Prev.* **13**, 3173–8 (2012).
131. Cao, W. *et al.* EZH2 promotes malignant behaviors via cell cycle dysregulation and its mRNA level associates with prognosis of patient with non-small cell lung cancer. *PLoS One* **7**, e52984 (2012).
132. Lu, Z. *et al.* Histone-lysine methyltransferase EHMT2 is involved in proliferation, apoptosis, cell invasion, and DNA methylation of human neuroblastoma cells. *Anticancer. Drugs* **24**, 484–93 (2013).
133. Chen, M. W. *et al.* H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM. *Cancer Res.* **70**, 7830–7840 (2010).
134. Hudson, T. J. *et al.* International network of cancer genome projects. *Nature* **464**, 993–8 (2010).

135. Bödör, C. *et al.* EZH2 mutations are frequent and represent an early event in follicular lymphoma. *Blood* **122**, 3165–8 (2013).
136. Moore, H. M. *et al.* EZH2 inhibition decreases p38 signaling and suppresses breast cancer motility and metastasis. *Breast Cancer Res. Treat.* **138**, 741–52 (2013).
137. Marchbank, T., Freeman, T. C. & Playford, R. J. Human Pancreatic Secretory Trypsin Inhibitor. *Digestion* **59**, 167–174 (1998).
138. Gonzalez, M. E. *et al.* EZH2 expands breast stem cells through activation of NOTCH1 signaling. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 3098–103 (2014).
139. Pastrana, E., Silva-Vargas, V. & Doetsch, F. Eyes wide open: a critical review of sphere-formation as an assay for stem cells. *Cell Stem Cell* **8**, 486–98 (2011).
140. Korn, J. M. *et al.* Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat. Genet.* **40**, 1253–60 (2008).
141. Bitler, B. G. *et al.* Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. *Nat. Med.* **advance on**, (2015).
142. Oshiro, C. *et al.* Taxane pathway. *Pharmacogenet. Genomics* **19**, 979–83 (2009).
143. Marcato, P., Dean, C. A., Giacomantonio, C. A. & Lee, P. W. K. Aldehyde dehydrogenase: its role as a cancer stem cell marker comes down to the specific isoform. *Cell Cycle* **10**, 1378–84 (2011).
144. Greve, B., Kelsch, R., Spaniol, K., Eich, H. T. & Götte, M. Flow cytometry in cancer stem cell analysis and separation. *Cytometry. A* **81**, 284–93 (2012).
145. O’Brien, C. A., Kreso, A. & Jamieson, C. H. M. Cancer stem cells and self-renewal. *Clin. Cancer Res.* **16**, 3113–20 (2010).
146. Kreso, A. & Dick, J. E. Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275–91 (2014).
147. Ponti, D. *et al.* Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer Res.* **65**, 5506–11 (2005).
148. Li, F., Tiede, B., Massagué, J. & Kang, Y. Beyond tumorigenesis: cancer stem cells in metastasis. *Cell Res.* **17**, 3–14 (2007).
149. Kessenbrock, K., Wang, C.-Y. & Werb, Z. Matrix metalloproteinases in stem cell regulation and cancer. *Matrix Biol.* **44**, 184–190 (2015).
150. Tumes, D. J. *et al.* The polycomb protein Ezh2 regulates differentiation and plasticity of CD4(+) T helper type 1 and type 2 cells. *Immunity* **39**, 819–32 (2013).
151. Su, I.-H. *et al.* Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat. Immunol.* **4**, 124–31 (2003).

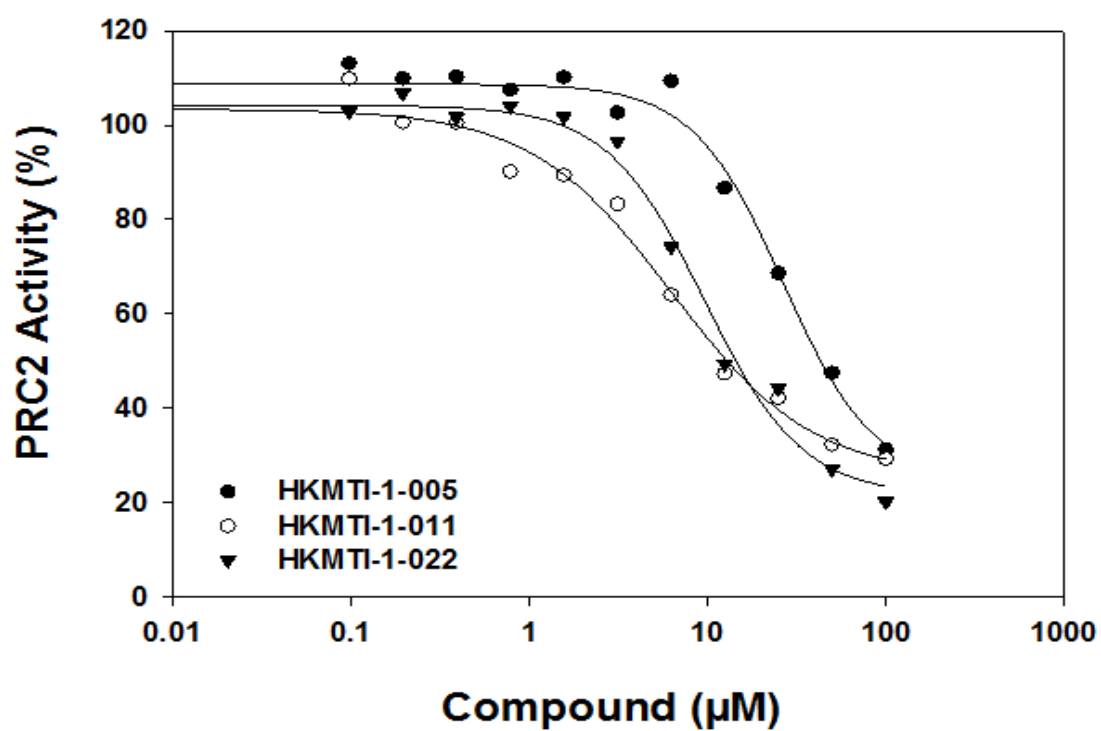
152. Kamminga, L. M. *et al.* The Polycomb group gene Ezh2 prevents hematopoietic stem cell exhaustion. *Blood* **107**, 2170–9 (2006).
153. Jene-Sanz, A. *et al.* Expression of polycomb targets predicts breast cancer prognosis. *Mol. Cell. Biol.* **33**, 3951–61 (2013).
154. Lowe, R. & Rakyan, V. K. Marmalaid—a database for Infinium HumanMethylation450. *BMC Bioinformatics* **14**, 359 (2013).
155. Fillmore, C. M. *et al.* EZH2 inhibition sensitizes BRG1 and EGFR mutant lung tumours to TopoII inhibitors. *Nature advance on*, (2015).
156. Dong, C. *et al.* G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. **122**, (2012).
157. Vinogradov, S. & Wei, X. Cancer stem cells and drug resistance: the potential of nanomedicine. *Nanomedicine (Lond)*. **7**, 597–615 (2012).
158. Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nat. Rev. Cancer* **12**, 323–34 (2012).
159. Sashida, G. *et al.* Ezh2 loss promotes development of myelodysplastic syndrome but attenuates its predisposition to leukaemic transformation. *Nat. Commun.* **5**, 4177 (2014).
160. Ntziachristos, P. *et al.* Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. *Nat. Med.* **18**, 298–301 (2012).
161. Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* **42**, 181–5 (2010).
162. de Vries, N. A. *et al.* Prolonged Ezh2 Depletion in Glioblastoma Causes a Robust Switch in Cell Fate Resulting in Tumor Progression. *Cell Rep.* (2015).

Chapter 8: Supplementary data

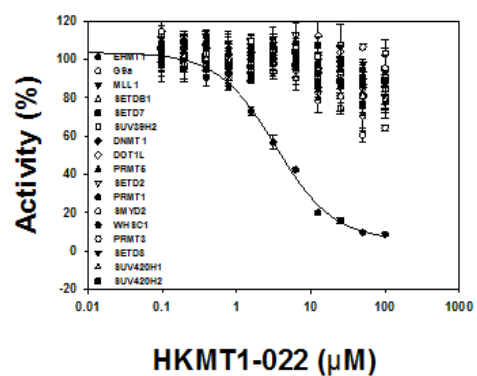
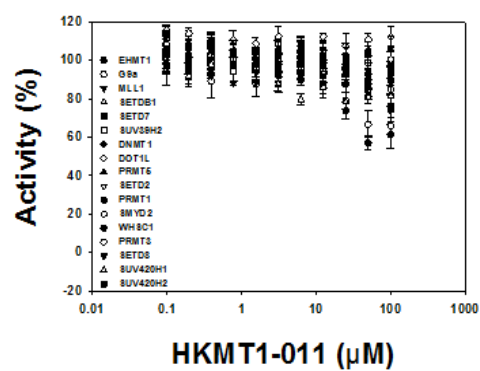
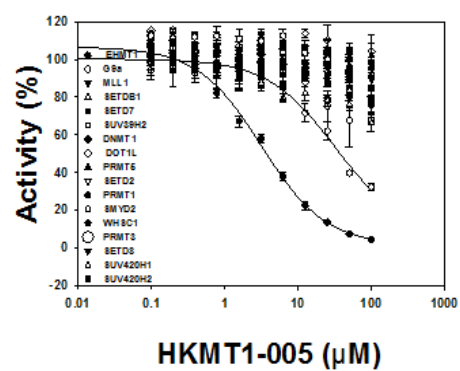
Supplementary Table 8.1 - RT-PCR data for single concentration (10 μ M) dose-RNA levels for target genes are normalised against the housekeeping gene GAPDH and shown as the fold increase compared to the mock treated sample.

Description	Compound	KRT17	FBXO32	JMJD3	EZH2
Hit	HKMTI-1-005	4.05	3.65	3.12	0.63
Hit	HKMTI-1-022	4.28	29.4	11.56	0.21
Hit	HKMTI-1--11	6.95	33.25	6.25	0.22
EHMT1/2i	BIX01294	1.06	3.34	2.7	0.87
EHMT1/2i	UNC0638*	1.1	5.5	3.4	0.4
EZH2i	GSK343	0.9	1.2	1.0	1.0
Negative	HKMTI-1-002	0.66	1.12	1.57	0.86
Negative	HKMTI-1-012	1.32	1.06	0.9	1.38
Negative	HKMTI-1-013	0.78	0.93	0.87	0.13

*UNC0638 treatment at 7.5 μ M, all other compounds given at 10 μ M.



Supplementary Figure 8.1: PRC2 activity following treatment with hit compounds
HKMT-I-005, HKMT-I-011, and HKMT-I-022



Supplementary Figure 8.2- HKMT selectivity screen activity following treatment with hit compound HKMT-I-005

Supplementary Table 8.2: Pearson correlation coefficient of EZH2 expression and target gene expression in normal human tissues with significant correlations highlighted red

EZH2 correlation	RHOQ	SPINK 1	KRT17	JMJD3	EHMT 2	SUZ12	EED	RBBP4
STEM CELLS	- 0.4266	-0.2209	- 0.2638	0.3267 2	0.3559 4	0.7088 2	0.6731 7	0.5800 3
B CELLS	0.3447 9	-0.5828	- 0.6427	0.7631 4	-0.4912	0.6423 3	0.7401 8	- 0.1743
T CELLS	- 0.6813	-0.0676	- 0.3345	0.0061 3	0.8980 5	0.8554 2	0.0050 2	- 0.3329
CNS	-0.201	-0.2526	- 0.1479	- 0.0906	0.2873 7	0.5607 1	0.6000 4	0.6805 3
MUSCLE	- 0.5703	0.9020 4	0.7478 5	0.4801 1	- -0.7503	- 0.6391	- 0.6002	- 0.6007
HEART	- 0.8602	0.0067 2	0.5592 3	- 0.2417	0.8946 7	- 0.1925	- 0.2782	- 0.1877
AIRWAY	-0.89	0.9981 7	0.8346	0.4202 6	0.6540 2	- 0.7721	- 0.2246	- 0.4358
TESTIS	-0.728	0.9539 4	- 0.3251	- 0.5757	0.7567 1	- -0.36	- 0.4854	0.2722 2
ALL DATA	- 0.7827	-0.6906	- 0.7495	0.5871 8	0.7394 7	- 0.1005	- 0.6254	0.6706 4

Supplementary Table 8.3: Pearson correlation coefficient of EHMT2 expression and target gene expression in normal human tissues with significant correlations highlighted red

EHMT2 correlation	RHOQ	SPINK 1	KRT1 7	JMJD3	EZH2	SUZ12	EED	RBBP 4
STEM CELLS	- 0.4291	-0.0911	0.0093 8	- 0.6219	- 0.4912	- 0.6524	- 0.5568	0.6903 9
B CELLS	-0.7138	0.0091 4	- 0.5755	0.0423 8	0.8980 5	0.6881 5	-0.189	- 0.6514
T CELLS	- 0.1852	-0.5352	- 0.2914	- 0.4522	0.2873 7	0.4422 9	- 0.2917	0.5928 8
CNS	0.4560 9	-0.7357	-0.734	- 0.3894	- 0.7503	0.7072 9	0.7763	0.6733 3
MUSCLE	- 0.5417	-0.4407	0.1299 9	- 0.6497	0.8946 7	0.2661	0.1801 5	0.2708 8
HEART	- 0.8157	0.6086 9	0.7824 9	0.6511 6	0.6540 2	- 0.6472	0.099	- 0.2696
AIRWAY	- 0.1285	0.5257 7	- 0.8635	0.0475 5	0.7567 1	0.3002 3	- 0.3011	- 0.2571
TESTIS	- 0.4915	-0.4646	- 0.4104	0.9684 7	0.7394 7	- 0.5759	- 0.1633	0.578
ALL DATA	- 0.2309	-0.3652	- 0.2705	- 0.2149	0.3559 4	0.3947 7	0.2620 6	0.4302

Supplementary table 8.4- Summary of probe IDs analysed for Cox proportional hazard modelling obtained from TCGA

Probe ID	Chromosome	Gene Symbol
A_23_P53216	11	EED
A_23_P53217	11	EED
A_23_P214638	6	EHMT2
A_23_P214639	6	EHMT2
A_24_P303389	6	EHMT2
A_24_P303390	6	EHMT2
A_23_P259641	7	EZH2
A_23_P259643	7	EZH2
A_32_P122579	7	EZH2
A_32_P122580	7	EZH2
NKI_NM_004456	7	EZH2
NM_004456_3_2455	7	EZH2
NM_004456_3_2590	7	EZH2
A_23_P115522	1	JMJD4
A_23_P115523	1	JMJD4
AK026908_1_3458	1	JMJD4
AK026908_1_3596	1	JMJD4
A_23_P422193	X	SUV39H1
A_23_P422195	X	SUV39H1
A_23_P202390	10	SUV39H2
A_23_P202392	10	SUV39H2
A_23_P202394	10	SUV39H2
A_23_P100883	17	SUZ12
A_23_P100885	17	SUZ12
A_24_P873263	17	SUZ12
A_32_P24215	17	SUZ12
A_32_P24223	17	SUZ12
A_32_P4321	17	SUZ12
A_32_P4324	17	SUZ12

Supplementary table 8.5: Significance (p-value) of enrichment of MDA-MB-231 EZH2

target genes

MDA-MB-231 EZH2 targets							
	Drug	Dose (μM)	Time point	Upregulation of EZH2 silenced genes	Downregulation of EZH2 silenced genes	Upregulation of EZH2 activated genes	Downregulation of EZH2 activated genes
Initial Array	GSK343	2.5	24	1.28E-16	1	1.15E-07	1
	GSK343	2.5	48	1.48E-15	1	0.024526	0.975474
	HKMT-I-005	2.5	24	1.32E-40	1	0.019079	0.980921
	HKMT-I-005	2.5	48	0.999997	3.36E-06	0.888697	0.111304
	HKMT-I-011	2.5	24	3.27E-21	1	0.974637	0.025363
	HKMT-I-011	2.5	48	0.070862	0.929139	0.004594	0.995406
	HKMT-I-022	2.5	24	3.81E-37	1	0.996884	0.003116
	HKMT-I-022	2.5	48	0.974032	0.025969	0.139714	0.860287
	TG3-259-1	2.5	24	7.20E-10	1	2.44E-06	0.999998
	TG3-259-1	2.5	48	0.999971	2.92E-05	4.83E-07	1
	UNC0638	2.5	24	3.68E-27	1	0.007298	0.992702
	UNC0638	2.5	48	2.65E-11	1	0.88931	0.11069
	HKMT-I-005	7.5	24	4.53E-43	1	0.999352	0.000648
	HKMT-I-005	7.5	48	0.098298	0.901702	0.999999	9.67E-07
	UNC0638	7.5	24	5.00E-18	1	0.992547	0.007453
	UNC0638	7.5	48	3.07E-10	1	1	6.33E-22
Validation Array	HKMT-I-005	7.5	24	1.73E-41	1	1	2.63E-18
	HKMT-I-005	7.5	48	4.95E-33	1	1	7.75E-27
	HKMT-I-011	2.5	24	1.99E-49	1	1	5.38E-08
	HKMT-I-011	2.5	48	2.43E-16	1	0.990156	0.009844
	TG3-184-1	2.5	24	1.12E-41	1	0.999681	0.000319
	TG3-184-1	2.5	48	3.05E-34	1	0.000986	0.999014

Supplementary table 8.6: Significance (p-value) of enrichment of MCF-7 EZH2 target genes

MCF-7 EZH2 targets					
	Drug	Dose (μM)	Time point	Upregulation of EZH2 silenced genes	Downregulation of EZH2 silenced genes
Initial Array	GSK343	2.5	24	0.095889	0.904113
	GSK343	2.5	48	0.008354	0.991646
	HKMT-I-005	2.5	24	0.025737	0.974264
	HKMT-I-005	2.5	48	0.994348	0.005653
	HKMT-I-011	2.5	24	0.39827	0.601733
	HKMT-I-011	2.5	48	0.488559	0.511444
	HKMT-I-022	2.5	24	0.06862	0.931381
	HKMT-I-022	2.5	48	0.861714	0.138288
	TG3-259-1	2.5	24	0.10086	0.899142
	TG3-259-1	2.5	48	0.876216	0.123786
	UNC0638	2.5	24	0.03815	0.96185
	UNC0638	2.5	48	0.241806	0.758197
	HKMT-I-005	7.5	24	0.467555	0.532448
	HKMT-I-005	7.5	48	0.726186	0.273817
	UNC0638	7.5	24	0.924195	0.075806
	UNC0638	7.5	48	0.586254	0.413749
Validation Array	HKMT-I-005	7.5	24	1	1
	HKMT-I-005	7.5	48	1	1
	HKMT-I-011	2.5	24	1	1
	HKMT-I-011	2.5	48	1	1
	TG3-184-1	2.5	24	1	1
	TG3-184-1	2.5	48	1	1

Supplementary table 8.7: Accession details for EZH2 siRNA array data used to generate meta-analysis targets

Accession no.	Related PubMed ID	Cell line	Platform	No. of EZH2 RNAi:CTRL arrays in Study
E-TABM-128	Citation missing	PC3	Agilent Whole Human Genome Oligo Microarray 012391 G4112A	04:04
GSE12692	19289832	A673	Affymetrix Human Genome U133a Array	03:02
GSE13286	19008416	SKBr3	Agilent Whole Human Genome Microarray 4x44k 014850 G4112F	02:02
GSE13674	19258506	UM-UC-3	Illumina Human-6 V2 Expression Beadchip	03:03
GSE20381	20708159	SKOV3	Illumina Humanht-12 V3.0 Expression Beadchip	04:04
GSE20433	20935635	LNCaP	Affymetrix Genechip Human Genome U133 Plus 2.0	02:02
GSE22209	20348445	HeLa	Rosetta/Merck Human RSTA Custom Affymetrix 2.0 Microarray	05:01
GSE22427	20935635	BJ	Illumina Humanht-12 V4.0 Expression Beadchip	03:03
GSE28501	21532618	OSCC3	Agilent Whole Human Genome Microarray 4x44k 014850 G4112F	02:02
GSE30670	21884980	MDA-MB-231	Illumina Humanref-8 V2.0 Expression Beadchip	03:03
GSE31433	22144423	SKOV3	Agilent Whole	06:03

			Human Genome Microarray 4x44k 014850 G4112F	
GSE36939	22986524	HCC70 and MDA-MB-468	Affymetrix Human Gene 1.0 St Array	04:04
GSE39452	23239736	LNCap and LnCap-abl	Affymetrix Genechip Human Genome U133 Plus 2.0	06:06
GSE41239	23051747	KARPAS-422 and Pfieffer	Affymetrix Genechip Human Genome U133 Plus 2.0	03:03
GSE41610	22966008	Umbilical vein endothelium	Agilent Whole Human Genome Microarray 4x44k 014850 G4112F	03:03
GSE42687	23526793	Mesenchymal stem cells	Phalanx Human Onearray	02:02
GSE6015	16618801	Embryonic fibroblasts	Affymetrix Human Genome U133a Array	03:03
GSE8145	17996646	H16N2 and RWPE	Chinnaiyan Human 20k Hs6	06:06

Supplementary table 8.8: Significance (p-value) of enrichment of meta-analysis EZH2 target genes

		Meta-analysis EZH2 targets					
	Drug	Dose (μM)	Time point	Upregulation of EZH2 silenced genes	Downregulation of EZH2 silenced genes	upregulation of EZH2 activated genes	Downregulation of EZH2 activated genes
Initial Array	GSK343	2.5	24	1.79E-16	1	1.71E-05	0.999983
	GSK343	2.5	48	2.55E-04	0.999745	0.010653	0.989347
	HKMT-I-005	2.5	24	3.65E-26	1	0.005293	0.994708
	HKMT-I-005	2.5	48	0.304096	6.96E-01	0.9927	0.0073
	HKMT-I-011	2.5	24	1.18E-28	1	0.996495	0.003505
	HKMT-I-011	2.5	48	2.43E-15	1	0.673897	0.326104
	HKMT-I-022	2.5	24	3.87E-23	1	0.990441	0.009559
	HKMT-I-022	2.5	48	0.118124	0.881877	0.881083	0.118918
	TG3-259-1	2.5	24	5.28E-03	0.994724	3.96E-07	1
	TG3-259-1	2.5	48	0.356346	6.44E-01	4.11E-03	0.995888
	UNC0638	2.5	24	3.35E-14	1	0.013729	0.986271
	UNC0638	2.5	48	1.64E-07	1	0.691848	0.308154
	HKMT-I-005	7.5	24	2.04E-49	1	0.99515	0.00485
	HKMT-I-005	7.5	48	3.85E-28	1	1	1.46E-11
	UNC0638	7.5	24	2.37E-29	1	0.974202	0.025799
	UNC0638	7.5	48	4.50E-18	1	1	4.09E-15
Validation Array	HKMT-I-005	7.5	24	9.09E-53	1	1	1.03E-24
	HKMT-I-005	7.5	48	1.38E-51	1	1	4.05E-26
	HKMT-I-011	2.5	24	3.18E-44	1	1	2.01E-09
	HKMT-I-011	2.5	48	6.30E-21	1	0.988803	0.011197
	TG3-184-1	2.5	24	2.22E-38	1	0.999991	9.47E-06
	TG3-184-1	2.5	48	1.97E-30	1	6.11E-05	0.999939

Supplementary Table 8.9: Target genes for enrichment analysis

Meta-analysis	
EZH2 silenced genes	EZH2 activated genes
CTSO	EZH2
FKBP14	TMPO
IL6R	USP1
PML	CCNA2
GALNT10	ILF3
OPN3	CCNF
PSAP	NUSAP1
SERP1	SNRPA1
EIF4EBP2	KIF23
DVL3	PSIP1
DNAJB9	TPX2
JARID2	PA2G4
WDFY1	TROAP
ATP6V1G1	TRIP13
ARHGEF3	PRC1
RPS6KA2	ACLY
DAAM1	BUB1
ARMCX3	FOXM1
SESN1	HMGB2
GOSR2	KIF11
MT1G	KIF14
MAN2A1	NEK2
ATP6V1A	PLK1
MT2A	SMC2
PCTP	TOP2A
DYRK2	YWHAE
SERPINE1	MAD2L1
ULK1	CDCA8
SLC20A1	RAD21
KIF1B	PTTG1
RAP2C	UBE2C
BTN3A2	BUB3
IGF1R	MCM6
P2RX5	STMN1
COX7B	STK3
BBX	BIRC5
NEDD4L	CENPE
SURF4	ATAD2
COL4A5	TACC3
SMPD1	CCNG1
BNIP3L	CDKN3
EIF2AK3	PBK
DUSP3	CBX5

ITGA2	MRPS16
ATP6V0D1	GTSE1
SEC24D	CSE1L
CTSC	KIAA0101
PLAG1	SKP2
ANKH	TXNRD1
PTPRE	SYNCRIP
SPRED2	PSME3
PBLD	CENPF
F8	SOCS3
TPM1	PTEN
SERPINE2	HMMR
DGCR2	CDC25C
OSTM1	EXOSC8
CA12	CCNB2
LNPEP	FADS1
CD47	ACAT2
PHLDA1	DUT
HECA	KIF2C
AP1S2	RBM3
KLHL24	MCM4
CASP7	API5
GLB1	DIAPH3
PSKH1	CCNB1
TTC8	TYMS
MT1X	TBCD
ZBTB34	LRP8
NR3C2	SPAG5
COPZ1	CIT
PRDX5	ABCE1
FLRT3	RSRC1
KIAA0226	NQO1
LIPA	WDR76
PAM	CCT6A
CD164	PHC1
MT1E	RAD51C
RNF149	BUB1B
FBXO8	PLK4
SOX4	KIF20A
MGLL	CEP152
JMJD1C	GCAT
COL5A1	WHSC1
NINJ1	CLSPN
MAPKAPK3	ARHGAP11A
PIP5K1C	TMEM48
TMED10	NUP88
HIPK2	PRR11

SGPL1	NUP50
TNFAIP3	TK1
GABARAPL1	DEPDC1
KHNYN	PROSC
LAMP2	SSH1
ELK3	FAM64A
FKBP1A	CENPL
FRAS1	RRM2
DEXI	AURKB
NT5E	ENSA
FAM127A	PSME4
AKAP13	TBCE
CLCN3	FBLN1
TRIM36	CASC5
CEACAM1	GK
KIAA1609	CENPM
APP	KNTC1
TPK1	SNX5
EML1	SMC4
FAM102A	PIF1
TNFRSF10B	PTBP1
ARRDC3	DHFR
BMPR2	SRPK1
IFI44	CCNE2
TULP4	DDX18
IDS	AURKA
RRAS	CDCA2
LMBR1L	ZWILCH
MAP1LC3B	EPB41L4B
TMEM2	PKP4
GEM	POLD3
QSOX1	MCM3
ITGB5	KPNA2
PTPRK	SRR
ARHGAP1	KIF5B
CYCS	RPAIN
KYNU	PCNA
GLG1	ANAPC1
TGFBR2	PRKDC
E2F5	KIF4A
BLCAP	CDCA7
B4GALT1	FBXO5
VPS13B	OIP5
TM7SF3	GOSR1
PLCB4	SMC3
DAZAP2	KIFC1
PDLIM4	HIST1H4C

FN1	PABPC4
LPIN1	ODF2
ZDHHC18	EEF1E1
FAM98A	GPAM
IL1RN	BRCA1
INPP4B	H3F3B
FAM3C	TTF2
ORMDL3	GPKOW
IL1A	MSH2
COL6A1	BARD1
NPAS2	CKAP2
SOCS1	SHMT1
PRRG1	MNS1
CAT	RFC4
AK1	IGFBP5
EDEM3	MELK
MXD4	STIP1
MFAP3	ABCF2
PHF1	PAICS
PPP3CA	SETMAR
ZDHHC3	KPNB1
IFI16	ASPM
C14orf28	DLG3
MLF1	GMNN
ATP6V0E1	SLC39A14
DUSP4	GUCY1B3
ZNF395	YAP1
CDS2	KIF22
IFNGR1	RAD54B
DUSP5	TUBGCP3
TIMP2	PURB
PIK3CA	POLQ
HLA-F	CTNNAL1
APLP2	WDR34
USP12	MDM1
ZFP36L1	CENPI
ANXA4	ACOT7
RIOK3	SORT1
PIK3R2	BCCIP
ZNF177	MTIF2
RHOC	SQLE
KLF6	SF3B1
ADAM10	DLEU1
GRN	ECT2
ICAM1	RANBP1
VTI1A	TCP1
DHRS9	RDX

PARP12	ENPP1
RAB22A	TCEA1
SLC31A1	CEP55
ANXA7	FDFT1
MSI2	YWHAH
TUBB2A	RFC3
NRP1	EIF1AX
SOX9	VRK1
PCMTD1	SRRM1
HMGA2	AKAP12
MT1H	ACACA
TP53INP1	TMEM106C
INPP1	NCAPG
SP110	TTK
TNFRSF21	SPAST
MBNL3	RABGGTB
FNDC3A	ANLN
F2RL2	H1FX
SGSH	FUS
ZBTB20	BCLAF1
GABRE	ERCC6L
ZDHHC9	BCAT2
CYLD	NTHL1
PLEKHB2	CDC7
SP100	SREBF1
FZD8	MTFR1
PLEK2	ABCF1
MBNL1	RRP1B
ATG12	CDCA5
SLK	CDCA3
ABCG1	RBBP9
NAMPT	PCF11
CBLB	HABP4
PLEKHH1	TFRC
SLC35D2	RAB31
CHURC1	PHF17
PEX19	SEPHS1
BMP2K	GCLM
PGAP3	ACTN1
CTSB	PANK2
DLG5	PGK1
ZFAND3	DNAJC8
SCARB2	RNF138
CPA4	RFC5
ANK3	CKLF
GNA12	SEMA3B
KDELRL3	FGFR1OP

F2R	ANP32E
CLDN1	UBTF
HIST2H2BE	ARG2
PGM3	RBM14
RHOQ	ESCO2
IGF2R	PWP1
LACTB	DONSON
TXNIP	CHEK1
SPOCK1	KPNA6
TGM2	HDGF
ZNF616	UPP1
TAX1BP3	ANP32B
ITM2C	CPSF6
MRPS6	CAV1
CPE	BLM
EXOC2	HPSE
MKNK2	TPM2
SLITRK6	NCAPD2
MAPK1	MDC1
AK3	GTPBP4
RASGRP3	MYH9
PRDM1	CDC25A
DDX58	CTBP2
PDE4B	KIF15
SKI	POLR3K
FARP1	GNL3L
RP2	BZW1
GOLPH3L	ADAM15
OPA1	ASNS
SRPX2	GPD2
SVIL	UBE2S
MGST3	TPM4
BMPR1A	CCDC34
ZNF264	DGUOK
DCBLD2	LXN
FAM134A	CEBPZ
MBNL2	MIS18A
SEPN1	SNRPA
PINK1	E2F2
DUSP6	PRKAR2A
CLSTN1	H2AFV
GPR137B	TFAM
CALCOCO2	ABCB10
PPIC	STAG1
RHOBTB3	LIN9
JAZF1	CKS1B
LPP	NUP85

MOV10	DLAT
SLC35F5	WDHD1
PLXDC2	MTMR4
GLRX	CETN3
ENTPD7	GMPS
TMCO1	HMGXB4
PRKAR1A	YARS
ETV1	MYBL2
SNX3	WDR67
DCLK1	DHX33
ID2	MTHFD2
HLA-E	FANCD2
FGFR3	FGFRL1
ANGPTL4	RYBP
ABI1	NEIL3
SLC2A12	EPRS
PELI1	SLC25A15
BIK	DHCR24
TUBB3	GOT1
PNRC1	SMC1A
SH3BGRL3	HMGB1
BTN3A3	NCBP2
COL4A2	PHGDH
NSF	MTHFD1L
RNF144A	CCNH
LAMA4	UHRF1
GCLC	IRS1
GLCCI1	GIN52
SMAP1	RBMX
LIPG	SLC1A5
TRIOBP	ATF1
ATF3	SACS
IREB2	TRAP1
YPEL5	INSIG1
LAMC1	DCK
CASP1	ITGB1
STK38L	MAP2K6
IL13RA1	SHCBP1
HLF	WDR4
CCL2	DTYMK
MVP	KCNK1
ATOX1	NUP93
SEC61B	FTSJ2
GNA11	NDC80
ATP2B1	DAXX
EXT1	AHCTF1
ARFIP1	SLC7A1

DNER	MRPL20
ANTXR2	RQCD1
RAPGEF1	DHX9
CMTM7	NOL7
CYB5R1	JPH1
PRKCE	HNRNPC
PTPRG	SLC43A3
CLCN5	KARS
RCOR3	TADA2A
TNC	CENPA
USP31	METTL7A
ZNF559	RNF6
GNG2	CDC20
SERPINB2	DEK
CREB5	HELLS
ZNF226	NCAPD3
TEP1	UBE3C
PKD2	INTS10
ZNF268	BCAP29
TGFBI	H2AFX
AMPD3	IFIT1
STX7	EXOSC2
APAF1	CENPN
MALL	UPF3B
NMB	LANCL1
SUSD1	ZWINT
BVES	BAX
SMPDL3A	HEATR2
PSD3	WDR36
FUT4	DFFA
UBE2D1	SCNN1A
STC2	H1FO
SIDT2	PIM1
CD58	TFAP2A
ETV5	CDKN2C
NEDD9	CDC6
ZNF136	GPSM2
HCP5	HMGN1
THBS1	FANCI
SMAP2	SLC45A3
CDKN1A	CDCA4
SCG2	MRPS30
PPP6C	MLF1IP
SLC25A37	RBL1
CFL2	RPL23
LRRC8B	DNA2
SYNJ2BP	FAM129A

BACH1	TM4SF1
IRF2	NUP188
PGM2L1	IMP4
PPAP2B	EBP
SMAD7	HMGN2
TMOD2	LMNB1
ZNF211	NFYB
SH3BGR1	ESPL1
RELN	ALDH1B1
TRPC1	KIAA1430
TAPBP	CALD1
ARID5B	NUDT1
FZD3	RHPN2
NFAT5	CPT1A
TOLLIP	CDC14B
TMEM55A	GAS2L3
ATXN1	SRSF1
GRB10	GNB1
KIAA0247	HSPA14
DGKD	UAP1
VAT1	TCF3
DOCK4	LIMA1
TTYH3	ROR1
ZMAT3	MRPS9
MX2	RBMS1
ITSN2	POLR3H
SLC35F2	ROCK1
ETS2	NUP155
ATP7A	CDK1
EFEMP1	EPOR
SEC22A	PTPLB
NCOA2	GTF2F2
PRCP	PSMC3IP
TPD52L2	FANCB
TFDP2	BECN1
QKI	BRIX1
KLHL35	NMU
BIRC3	GRK6
LHFPL2	SGOL1
JUN	TFDP1
TMEM50B	HNRNPH1
NUAK1	OGDH
ALCAM	CDCA7L
RABGAP1L	ACTR6
IGF2BP3	DNM1L
KMO	TIMM10
BTG2	ELOVL6

RPS27	ENO2
SERPINI1	POLD1
TTLL11	CDK5
EGR1	NMT1
PIK3IP1	SRP72
DTX3L	RIF1
DCUN1D3	FASN
HLA-C	EPHA2
POPDC3	HEATR1
SRD5A1	NUP43
RPS27L	CSNK2A1
RNF170	DBF4
ATP6V1H	SF3B2
MYD88	ADD3
MIR22HG	UXT
ZBTB4	DSP
LAMP1	IGFBP3
FUT8	MSH6
SIAE	SUMO3
KRCC1	SMEK2
MSRB3	FEN1
RNF213	SLC16A9
GNG12	GTF3C2
NR1H2	PUS7
TNFRSF11B	HRASLS2
TLE1	ATL2
MT1F	FLOT1
PACS1	RAD51AP1
MNT	MCM5
MAPKAPK2	FXR1
IRF1	ATP1B1
CD81	CD9
NEK6	CDK2
MMP10	TLK1
MT1A	RRM2B
ABLIM3	NSMAF
DSCR3	THOC6
IRF7	PSRC1
SPRY2	FAM83D
SLC6A16	CUL5
PMAIP1	PPA2
TMEM158	SLC7A11
ERAP1	MCM10
LGALS3BP	RCC1
CCNL2	RRM1
ITGA5	NOC2L
C19orf66	MAP2K3

HN1L	PHF10
EDEM1	GLTP
IL24	INCENP
RUNX1	MRPS27
MAP4K3	ADK
RAB18	NUCKS1
ACVR1	DNAJB12
F2RL1	QSER1
DNAJB6	CACYBP
PEX11B	HMGB3
PLD1	RAD51
PMEPA1	RFC2
ANKRD46	C1orf43
ZFHX3	IMPA2
NAPA	RACGAP1
CYB561	GINS4
TNFSF10	PPP2R5A
ATM	KIF18A
ZFAND6	MCM7
HDAC9	HYLS1
TMEM87B	REEP4
COL4A1	TMEM17
COL7A1	CTSL2
FBXW2	RAD54L
NR2F2	WTAP
TNIK	ITPK1
CD63	DDX17
CAMK2D	MRPL39
TTC17	COL9A3
LTBP2	USP48
COL8A1	KLHL7
TGIF2	ADSS
ALG9	NIN

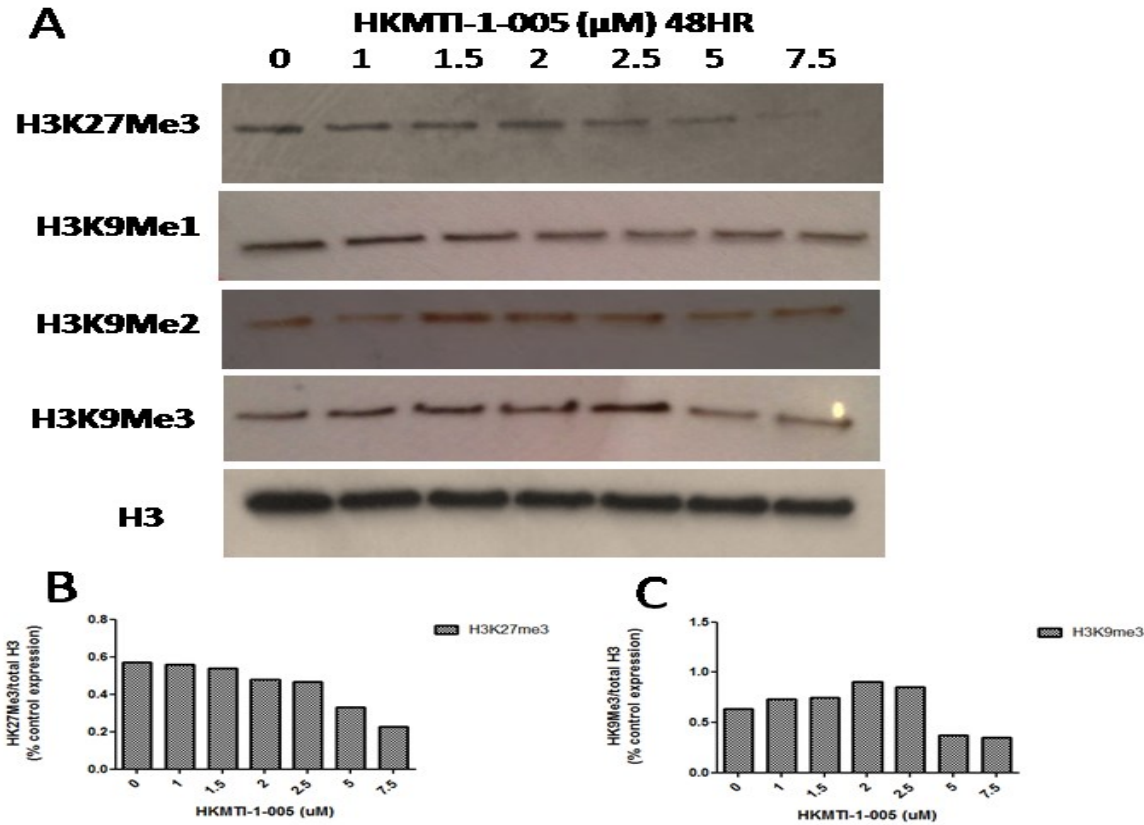
Supplementary Table 8.10: siRNA sequences

Product Name	Target Sequence	Manufacturer/Catalogue#
EHMT2 (G9a)(HS_BAT8_ 1)	ATCGAGGTGATCCGCATGCTA	QIAGEN SI00091189
EHMT2 (G9a)(HS_EHMT2_ 1)	CCTCTTCGACTTAGACAACAA	QIAGEN SI03083241
EZH2(HS_EZH2_ 4)	TTCGAGCTCCTCTGAAGCAAA	QIAGEN SI00063973

EZH2(HS_EZH2_7)	AACCATGTTTACAACTATCAA	QIAGEN SI02665166
-----------------	-----------------------	----------------------

Supplementary Table 8.11: Primers for ChIP-PCR designed using Primer3

Name	forward	reverse	Product
ChIP_KRT17_UTR1	TGGCATTGATGAGTGAGAGG	AGCCGAGAGACATTCTCAA	
ChIP_Ex1_FBXO32	GGGCAGAACTGGGTGAAGAC	CTGAGGTCGCTCACGAAACT	80bp
ChIP_GAPDH	CACCGTCAAGGCTGAGAACG	ATACCCAAGGGAGCCACACC	134bp
ChIP_SPINK1_ChIP	TTGCCTAGTGTGTGATGCAA	GCGAAATCCATGCCTTCTAA	81bp

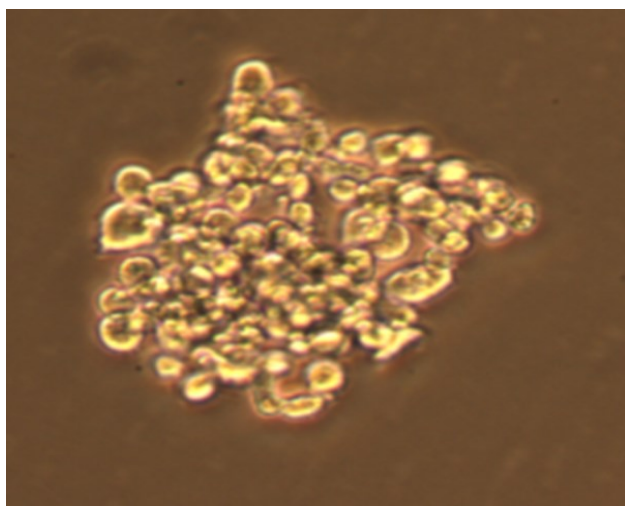


Supplementary Figure 8.3: Western blot analysis by Sarah Kandil of total H3K27Me3, H3K9Me1/2/3/ and H3 histone marks using histone extracts of MDA-MB-231 treated for 48hr with HKMTI-1-005 (0-7.5 μ M). Densitometry analysis, using ImageQuant software, was carried out to assess the H3K27Me3 and H3K9Me3 expression levels relative to total H3 expression levels in the histone extracts

Supplementary Table 8.12- IC50 of cell proliferation after HKMT-I-005 treatment alongside predicted CNV from Broad institute¹⁴⁰ and Sanger institute¹²⁰

Cell type	Cell Line	IC50 cell proliferation (μM) HKMT-I-005	CNV BROAD				CNV SANGER			
			EZH 1	EZH 2	EHMT 1	EHMT 2	EZH 1	EZH 2	EHMT 1	EHMT 2
Lymphoma	SC1	3.71								
	WILL1	5.6								
	DOHH2	3.26					2	3	2	2
	WSU-FSCLL	3.41								
	DB	<1					4	5	2	3
	SUDHL8	<1	2.0	2.9	1.9	2.0				
Ovarian Cancer	A2780	15.96					2	2	2	2
	A2780C P	21.21								
	PEO23	27.82								
	PEO14	22.92								
	PEO1	15.45								
	PEO4	29.77								
Breast cancer	MDA-MB-231	10.4					3	4	3	4

	MCF7	7.7	1.6	2.1	2.0	2.1	3	4	3	3
	T47D	8.5	2.9	1.5	2.9	2.3	4	2	3	3
	BT474	2.1								
	SKBR3	7.7	1.1	2.1	2.1	1.5				
Breast epithelial	MCF10 A	>15								



Supplementary figure 8.4- Representative image of MDA-MB-231 mammosphere at 40x magnification after DMSO control treatment 5 days after beginning of non-adherent culture)

Supplementary table 8.13- CSC activity IC50 of treatments in MDA-MB-231 breast cancer cells (including chemotherapy)

	HKMT-I-005	HKMT-I-011	GSK 343	UNC 0638	PACLITAXEL	PACLITAXEL + 1 μ M HKMT-I-005	CISPLATIN	CISPLATIN + 1 μ M HKMT-I-005
IC50 (μ M)	1.939	5.978	0.2529	1.783	Not calculable	2.697	Not calculable	1.83

Supplementary table 8.14- Genes related to Taxane pathway compared to differentially expressed genes after HKMT-I-005 treatment (Chapter 4, 4.2) - Genes showing a decrease in expression after treatment highlighted **RED, genes showing an increase in expression after treatment highlighted **GREEN****

Description	Gene	Comments
ABC drug transporters	ABCC6	
	ABCB11	
	ABCA1	
	ABCG1	
	ABCC6	
	ABCA13	
	ABCC3	
	ABCG4	
	ABCA7	
	ABCA11P	pseudogene, and is affiliated with the lncRNA
	ABCC2	
	ABCB8	
Cytochrome p450	CYP3A43	
	CYP1B1	
Tubulin-encoding	TUBB8	
	TUBB3	
	TUBB2B	
Kinase inhibitor	CDKN1A	cyclin-dependent kinase inhibitor 1A
Apoptosis regulator	BCL2	Regulates cell death by controlling the mitochondrial membrane permeability

Supplementary table 8.15- Tumour take in second generation following treatment

#Cells injected	#test animals	#Tumours formed	Treatment (1st generation)
10	5	5	Control (DMSO)
5	5	4	Control (DMSO)
10	5	4	Paclitaxel
5	5	4	Paclitaxel
10	5	3	HKMT-I-005
5	5	4	HKMT-I-005
10	5	2	Paclitaxel & HKMT-I-005
5	5	1	Paclitaxel & HKMT-I-005

APPENDIX

Manuscript of Curry et al ‘*Dual EZH2 and EHMT2 histone methyltransferase inhibition increases biological efficacy in breast cancer cells*’

Dual EZH2 and EHMT2 histone methyltransferase inhibition increases biological efficacy in breast cancer cells.

Edward Curry¹, Ian Green¹, Nadine Chapman-Rothe¹, Elham Shamsaei¹, Sarah Kandil¹, Fanny Cherblanc², Luke Payne¹, Emma Bell¹, Thota Ganesh³, Nitipol Srimongkolpithak², Joachim Caron², Fengling Li⁴, Anthony G Uren⁵, James P Snyder⁶, Masoud Vedadi⁴, Matthew J. Fuchter^{2*}, Robert Brown^{1, 7*}.

1. Ovarian Cancer Action Research Centre, Department of Surgery and Cancer, Imperial College London, Hammersmith Hospital Campus, London, W12 0NN, UK.

2. Department of Chemistry, Imperial College London, South Kensington Campus, London SW7 2AZ, UK.

3. Department of Pharmacology, Emory University, Atlanta, GA 30322, USA.

4. Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada

5. MRC Clinical Sciences Centre, Hammersmith Hospital Campus, London W12 0NN, UK

6. Department of Chemistry, Emory University, Atlanta, GA 30322, USA.

7. Section of Molecular Pathology, Institute of Cancer Research, Sutton, SM2 5NG, UK.

*Correspondence should be addressed to R.B. (b.brown@imperial.ac.uk) or M.J.F. (m.fuchter@imperial.ac.uk).

Keywords: epigenetics; breast cancer; histone modifications; gene silencing; chemical probe;

ABSTRACT

Background: Many cancers show aberrant silencing of gene expression and overexpression of histone methyltransferases. The histone methyltransferases (HKMT) EZH2 and EHMT2 maintain the repressive chromatin histone marks H3K27 and H3K9 methylation respectively, which are associated with transcriptional silencing. Although selective HKMT inhibitors reduce levels of individual repressive marks, removal of H3K27me3 by specific EZH2 inhibitors, for instance, may not be sufficient for inducing expression of genes with multiple repressive marks.

Results: We report that gene expression and inhibition of triple negative breast cancer cell growth (MDA-MB-231) are markedly increased when targeting both EZH2 and EHMT2, either by siRNA knockdown or pharmacological inhibition, rather than independently. Indeed, expression of certain genes is only induced upon dual inhibition. We sought to identify compounds which showed evidence of dual EZH2 and EHMT2 inhibition. Using a cell-based assay, based on the substrate-competitive EHMT2 inhibitor BIX01294, we have identified proof-of-concept compounds that induce re-expression of a subset of genes consistent with dual HKMT inhibition. Chromatin immunoprecipitation verified a decrease in silencing marks and an increase in permissive marks at the promoter and transcription start site of re-expressed genes, while Western analysis showed reduction in global levels of H3K27me3 and H3K9me3. The compounds inhibit growth in a panel of breast cancer and lymphoma cell lines with low to sub-micromolar IC₅₀s. Biochemically, the compounds are substrate competitive inhibitors against both EZH2 and EHMT1/2.

Conclusions: We have demonstrated that dual inhibition of EZH2 and EHMT2 is more effective at eliciting biological responses of gene transcription and cancer cell growth inhibition compared to inhibition of single HKMTs, and we report the first dual EZH2- EHMT1/2 substrate competitive inhibitors that are functional in cells.

BACKGROUND

EZH2 along with EED and SUZ12 are the indispensable core components of the polycomb repressive complex (PRC2) responsible for maintenance of the repressive epigenetic mark H3K27me3: trimethylation of lysine 27 of histone 3 [1]. High expression of the histone methyltransferase (HKMT) EZH2, in some cases associated with gene amplification, has been well documented in a variety of cancers [2], [3]. EZH2 over-expression has been linked to poor prognosis [4, 5] and shown to be a marker of aggressive breast cancer [6], associated with difficult to treat basal or triple negative breast cancer [7]. Gene knockdown of EZH2 reduces growth of a variety of tumour cell types [5, 8, 9]. Several groups have reported specific co-factor competitive EZH2 inhibitors [10-16], which have shown a strong capacity to reduce growth of cells expressing mutated forms of EZH2 (such as certain non-Hodgkin's lymphoma, [12]). However, removal of the repressive mark H3K27me3 alone may not always be sufficient for reversal of gene silencing. Indeed, it has been shown that highly specific EZH2 inhibitors require a mutant EZH2 status to inhibit cell growth, being less effective in cells solely expressing wild type EZH2 [5, 8, 9]. Elimination of further repressive methylation marks by inhibition of additional HKMTs may be required to fully realise the epigenetic potential of HKMT inhibitors.

EHMT2 (also known as G9a), and the highly homologous EHMT1 (also known as GLP) are HKMTs partly responsible for mono- and di-methylation of lysine nine of histone 3 (H3K9me1 and H3K9me2 respectively); repressive chromatin marks found on the promoter regions of genes that are often aberrantly silenced in cancer [17]. EHMT2 is over-expressed and amplified in various cancers including leukemia, prostate carcinoma, and lung cancer, with gene knockdown of EHMT2 inhibiting cancer cell growth in these tumour types [18, 19]. BIX-01294 (see Figure 2) was previously identified as an inhibitor of the HKMTs EHMT2 and EHMT1 and subsequent medicinal chemistry studies around the 2,4-diamino-6,7- dimethoxyquinazoline template of BIX-01294 have yielded a number of follow up EHMT2 inhibitors [20-25].

In addition to its role methylating H3K9, EHMT2 has been shown to be able to methylate H3K27 [26, 27]. It has been suggested that this could provide cells with a mechanism to compensate in part for a loss of EZH2 [28]. The picture is further complicated by recent evidence that EHMT2 and EZH2 (via the PRC2 complex) interact physically and share targets for epigenetic silencing [29]. Combining this evidence, it would again suggest that specifically targeting either EZH2 or EHMT2 alone may not be sufficient to reverse epigenetic silencing of genes, but rather combined inhibition may be required. To this end, we have examined the effect of dual EZH2 and EHMT2 gene knock down or enzyme inhibition in breast cancer cells. Consistent with the requirement for removal of both repressive H3K9 and H3K27 methylation marks, we show that dual inhibition of EHMT2 and EZH2 pharmacologically or by SiRNA is necessary for reactivation of certain genes and induces greater inhibition of cell growth than targeting either HKMT alone in triple negative breast cancer MDA-MB-231 cells. Further we have identified proof of concept compounds which are dual (substrate competitive) EZH2-EHMT1/2 inhibitors.

RESULTS

Combined inhibition of EZH2 and EHMT2 is more effective at inducing gene re- expression and inhibiting tumour cell growth than single HKMT inhibition. SiRNA knockdown in the MDA-MB-231 breast tumour cell line was used to examine the effect of combined inhibition of *EZH2* and *EHMT2* expression on epigenetic regulation at select target genes, compared to knockdown of either gene alone in MDA-MB-231 cells (Figure 1A). Knockdown of *EZH2* with two independent SiRNAs induced 2-4 fold increased mRNA levels of *KRT17* and *FBXO32*; genes which are known to be silenced in an EZH2 dependent manner [30]. Knockdown of *EHMT2* (G9a) had limited effects on mRNA levels of these target genes. However, double knockdown of *EZH2* and *EHMT2* had dramatic effects on *SPINK1* mRNA levels; a gene which was not upregulated by silencing of EZH2 or EHMT2 individually. Thus, for at least certain genes, dual reduction in *EZH2* and *EHMT2* levels are necessary to observe marked changes in target gene expression 48h following knockdown.

The effects on gene expression of the selective EZH2 inhibitor GSK343 [10] (Figure 2) and the selective EHMT2 inhibitor UNC0638 [22] (Figure 2) used alone or in combination were also examined using the MDA-MB-231 triple negative breast cancer cell line (Figure 1B). When MDA-MB-231 cells were treated with the EZH2 inhibitor GSK343 at 1-15 μ M for 48h alone there was little change in the mRNA levels of *KRT17*, *FBX032* and *SPINK1* and the H3K27 demethylase *JMJD3* (Figure 1B). UNC0638 at 1-15 μ M for 48h alone showed dose dependent up-regulation of *FBX032* and *JMJD3*, however *KRT17* and *SPINK1* mRNA levels were not significantly altered. However, the combination treatments with GSK343 and UNC0638 showed marked increase in mRNA levels of all the target genes, in contrast to the single agent treatment. Consistent with dual *EZH2/EHMT2* SiRNA knockdown, *SPINK1* has the biggest change in mRNA levels between the single and combination treatments, having a 50-fold increase with the combination treatment.

Next, the effects on cell viability of GSK343 and UNC0638 used alone or in combination were examined (Figure 1C). Treatment alone with GSK343 showed no significant reduction in cell viability up to 15 μ M, while UNC0638 sole treatment caused a dose dependant reduction in cell viability, with a calculated IC₅₀ of 9 μ M. When the cells were treated with both compounds in combination, a marked increase in growth inhibition was observed when compared to single agent treatment using UNC0638 or GSK343 (Figure 1C). This is particularly apparent at a 5 μ M concentration of both compounds, where alone they have no significant effect on reducing cell viability, while in combination they markedly reduce cell viability to >50% (p<0.01).

Analogues of an EHMT2 specific inhibitor can up-regulate EZH2 silenced genes.

Both EZH2 and EHMT1/2, belong to the SET-domain superfamily [31], the catalytic SET-domain being responsible for the methylation of the targeted lysine residues. BIX-01294 has previously been shown, both structurally and biochemically to bind to the substrate (histone) binding pocket of EHMT1/2 [32]. Since protein recognition motifs for histone binding at repressive sites are similar [33] and EHMT2 has been shown to be able to methylate H3K27, in addition to its more common H3K9 target [27], it is likely that there are common aspects to the histone substrate binding pockets of the repressive HKMTs EZH2 and EHMT1/2. We therefore felt it would be feasible to use quinazoline template of BIX-01294 in the discovery of dual (substrate competitive) EZH2-EHMT1/2 inhibitors.

A compound library based on the selective BIX-01294 EHMT2 inhibitor was synthesized and characterised analogously to previously reported methods [20-22, 24, 25, 32] and as described in Supplementary Methods. In light of the previously reported selectivity of this chemical scaffold towards EHMT1/2, the library was primarily examined for compounds showing additional EZH2 inhibitory activity, as defined by re-expression of KRT17 and FBXO32; genes which are known to be silenced in an EZH2 dependent manner [30]. The majority of compounds had little or no effect on both KRT17 and FBXO32 RNA levels. However, we identified three compounds which up-regulate KRT17 and FBXO32 RNA levels. The data for these compounds along with a comparison of the related EHMT2 inhibitors BIX-01294 and UNC0638 and a representative number of negative compounds are shown in Table 1 (for chemical structures see Figure 2). All hit compounds – HKMTI-1-005, HKMTI-1-011, HKMTI-1-022 - showed upregulation of KRT17, FBXO32, and JMJD3 mRNA at a 10⁻⁶ M dose. The reported EHMT2 specific inhibitors BIX-01294 and UNC0638, while being closely related to our hits from a chemical structure perspective, elicit different effects on expression of the target genes. BIX-01294 (Table 1, entry 4) does not up-regulate KRT17, but does up-regulate FBXO32. This is compatible with the observation that FBXO32 is regulated via multiple mechanisms, potentially responding to a variety of factors [34]. An analogous effect is observed for UNC0638 (Table 1, entry 5). The specific EZH2 inhibitor GSK343 has no effect whatsoever on all the target genes studied (Table 1, entry 6) when examined up to 72h following treatment and at concentrations up to 10⁻⁶ M.

To further evaluate the three hit compounds identified, we treated MDA-MB-231 cells for 48h and 72h at various concentrations of compounds (Figure 3A). All hit compounds showed a dose-dependent increase of KRT17, FBXO32, as well as JMJD3 mRNA. Higher doses of certain compounds started to cause cell death, and at these doses, expression of KRT17 was often below the detection limit of low-expressed genes due to cell death.

Chromatin Immunoprecipitation (ChIP) experiments were carried out on treated MDA-MB-231 cells to verify that the detected gene up-regulation is indeed due to chromatin remodelling (Figure 3B). We tested the silencing marks H3K9me3 and H3K27me3 as well as the activating marks H3K4me3, H3K4me2, H3K27ac and H3K9ac. All three compounds showed a clear decrease in repressive chromatin marks (H3K27me3, H3K9me3), and at least in some instances, an increase in permissive marks, at two target genes (Figure 3B). This is consistent with the compounds having dual HKMT inhibitory activity in removing both H3K9me and H3K27me marks, while allowing activating marks to be established at these loci.

Genome-wide changes in gene expression. Agilent microarrays were used to perform gene expression profiling in MDA-MB-231 breast cancer cells after 24 hours of treatment with the hit compound HKMTI-1-005, the EZH2 inhibitor GSK343 [10], and EHMT2 inhibitor UNC0638 [22]. To validate the finding of the initial expression data for the hit compounds, a second microarray experiment was performed on the same platform using HKMTI-1-005 treated MDA-MB-231 cells after 24 hours of treatment. To assess the extent to which our selected analogues - derived from the selective EHMT1/2 inhibitor BIX-01294 - had gained EZH2 inhibitory activity, lists of genes activated or repressed following siRNA knockdown of EZH2 in MDA-MB-231 cells were identified [35] and shown in Supplementary Table S4. These lists of target genes were investigated in the context of genome-wide changes in gene expression following treatment with the compounds. HKMTI-1-005 showed very significant enrichment for upregulation of EZH2 silenced genes (Figure 4A) in both the initial array ($p=4.53 \times 10^{-43}$) and the validation array ($p=1.99 \times 10^{-49}$). GSK343 and UNC0638 also both showed a significant upregulation of EZH2 target genes (Figure 4A) though to a lesser extent than HKMTI-1-005. Indeed analysis of the difference in systematic upregulation showed that HKMTI-1-005 upregulated EZH2 silenced genes significantly more than either GSK343 ($p=5.8 \times 10^{-5}$) or UNC0638, ($p=1.7 \times 10^{-4}$).

The same enrichment tests were repeated using target gene sets identified in an EZH2 siRNA knockdown study in another breast cancer cell line, MCF-7 [30]. Almost no enrichment was observed of this gene set in MDA-MB-231 cells after treatment with any of the compounds (HKMTI-1-005, GSK343 and UNC0638) (Figure 4A), suggesting that EZH2 has cell type specific targets. To investigate this further, we undertook a meta-analysis to identify consensus target genes based on 18 independent EZH2 siRNA studies (details of the meta-analysis are provided in Methods). Encouragingly, treatment of MDA-MB-231 cells with HKMTI-1-005 resulted in highly significant upregulation of these consensus EZH2 repressed genes (Figure 4A). This suggests that key EZH2 target genes that are conserved across a wide range of cell lines are re-expressed upon treatment with our dual HKMT inhibitor. Furthermore, this identifies generally applicable pharmacodynamic biomarkers of EZH2 inhibitors across cell types.

Compound induced changes in H3K9me and H3K27me in cells. The microarray data showed a clear upregulation of the levels of SPINK1 mRNA (a gene previously identified as a target for dual EZH2 and EHMT2 inhibition, see Figure 1) following treatment with HKMTI-

1-005, an observation that was confirmed via qRT-PCR (Figure 4B). These qRT-PCR experiments demonstrated a dose-dependent upregulation of SPINK1 alongside a re-evaluation of the candidate genes (KRT17, FBX032, JMJD3) chosen for the initial compound screen. Furthermore, ChIP-PCR at the SPINK1 transcription start site clearly demonstrated

a reduction in both H3K27me3 and H3K9me3 in MDA-MB-231 cells after treatment with 2.5 μ M HKMT-I-005 (Figure 4C). More broadly, Western analysis showed global levels of H3K27me3 and H3K9me3 are reduced in MDA-MB-231 cells after treatment with HKMTI-1-005 (Figure 4D) and densitometry analysis (Figure 4E) suggests this happens in dose dependent manner. Together these data strongly support the hit compound HKMT-I-005 reduces levels of H3K27me3 and H3K9me3 at concentrations of compound that are less or equivalent to the growth inhibition IC50 concentration for MDA-MB-231 (Table 2).

In order to identify specific pathways being transcriptionally modulated, the microarray data was analysed for enrichment of pathways belonging to each pathway listed on the ConsensusPathDB (CPDB) database [36]. The Benjamini-Hochberg adjusted [37] enrichment p-value estimates for each treatment is given in Supplementary Table S6. Interestingly, genes belonging to the pathway ‘Apoptosis’ displayed a highly significant systematic shifted towards upregulation on treatment with our hit compound(s) at 24hrs ($p < 1E-4$), but not the selective EZH2 (GSK343) or EHMT2 (UNC0638) inhibitor compound ($p=0.42$ and $p=0.30$, respectively). Consistent with induction of apoptosis related genes, hit compound HKMTI-1-005 induces apoptosis in MDA-MB-231 cells in a dose-dependent manner, as measured by Caspase 3/7 activity (Supplementary Figure S1).

Cell growth inhibition induced by HKMT inhibitors. EZH2 inhibitors are reported to be particularly effective at inhibiting cell growth of cell lines with mutant EZH2 [11, 12]. Indeed, the DB lymphoma cell line which has an EZH2 mutation (Y646N, according to the COSMIC database [38]) was observed to be particularly sensitive to the EZH2 inhibitor GSK343 (Table 2). Consistent with the hit compounds having gained EZH2 inhibitory activity, DB cells were also found to be sensitive to HKMTI-1-005. GSK343 was found to be less potent on all the other lymphoma lines, which express wild type EZH2, with anti-proliferative effects observed at μ M concentrations of compounds. This included the cell line SUDLH8, which has amplified and highly expressed wild-type EZH2 (processed data obtained from the Cancer Cell Line Encyclopedia [39]). Interestingly, SUDLH8 is more sensitive to HKMTI-1-005 than the other lymphoma lines with WT EZH2 (Table 2), suggesting that increased sensitivity to this dual inhibitor will not be dependent on cancer cells carrying activating mutations, but perhaps any mechanism of increased dependency on EZH2.

The anti-proliferative effect of HKMTI-1-005 on a small panel of breast cancer cell lines was determined, with IC₅₀ values in the range 2-10 nM (Table 2). All of the cancer breast cell lines examined were found to be more sensitive to HKMTI-1-005 compared to a normal breast epithelial cell line MCF10a. The breast cancer cell line BT-474, which is the cell line most sensitive to HKMTI-1-005 treatment, has the highest relative expression of EZH2, as detected by Western analysis (data not shown).

Hit compounds directly inhibit EZH2 and EHMT1/2 and are substrate competitive inhibitors. We have previously reported the EHMT2 IC₅₀ of HKMTI-1-005, HKMTI-1-011 and HKMTI-1-022 to be 0.10, 3.19, and 0.47 nM respectively [40]. This data was generated using a scintillation proximity assay (SPA) which monitors the transfer of a tritium-labelled methyl group from [³H]S-adenosyl-L-methionine (SAM) to a biotinylated-H3 (1-25) peptide substrate, mediated by EHMT2. A comparable PRC2 enzymatic assay was employed here to assess biochemical inhibitory activity of our hits against EZH2. A trimeric PRC2 complex (EZH2:EED:SUZ12) was employed in this assay, along with a biotinylated-H3 (21-44) peptide substrate. This revealed HKMTI-1-005, HKMTI-1-011 and HKMTI-1-022 to have PRC2 IC₅₀ values of 24, 12 and 16 nM under these assay conditions (see Supplementary Figure S2). Since the peptide substrates used in these assays are poor models for the complex and dynamic structure of the chromatin substrate in cells, and since the only minimal number of PRC2 proteins (EZH2:EED:SUZ12) required for enzymatically active EZH2 were employed in the PRC2 assay, care should be taken in the over interpretation of this in vitro inhibitory data. Nonetheless, we note that both the EHMT2 and PRC2 biochemical potency is comparable to the inhibitory concentrations employed in our cell based assays.

Perhaps more importantly, in accordance with our design rationale, mechanism of inhibition studies on representative hit HKMTI-1-005 revealed it to have a well-defined, peptide substrate competitive mechanism of action (see Supplementary Figure S3), in contrast to all known EZH2 inhibitory chemotypes. Broad screening of our compound library against PRC2 using this assay revealed the IC₅₀ values obtained for all actives to be dependent on peptide substrate concentration (data not shown), further confirming a substrate competitive inhibitory mode for this chemotype.

Finally, a methyltransferase selectivity screen was carried out for the hits on a panel of enzymes including eleven HKMTs, three protein arginine methyltransferases (PRMTs), and one DNA methyltransferase (DNMT) (Supplementary Figure S4). None of the hits had any significant inhibitory activity against these fifteen other methyltransferase targets (up to 100

M), confirming them to be selective for EZH2 and EHMT1/2. Taken together, these data reveal our hit compounds to be dual EZH2 and EHMT1/2 inhibitors with a substrate competitive mechanism of action.

DISCUSSION

It is widely accepted that the installation, maintenance and functional output of epigenetic modifications occur in concert via combinatorial sets of modifications. Therefore removal of specific repressive marks may not alone be sufficient for reversal of gene silencing. Elimination of multiple repressive methylation marks may instead be required to re-express a wider spectrum of genes. Given the complexities of epigenetic regulation and cross-talk between epigenetic regulators, the discovery of inhibitors of epigenetic processes that lead to reversal of epigenetic silencing may be more suited to cell-based methods measuring reactivation of a panel of target genes, rather than cell-free assays that use purified components. Through the use of a breast cancer (MDA-MB-231) cell assay based on the re-expression of epigenetically silenced genes, we report the identification of hit compounds that phenocopy the effects of dual EZH2/EHMT2 pharmacological inhibition and dual SiRNA gene knockdown.

The recently reported specific EZH2 inhibitors are all co-factor competitive, the majority of which have converged to a common chemotype (Figure 2) [10-16]. Conversely, the dual EZH2/EHMT2 inhibitors we here report are substrate competitive. Not only do these represent the first inhibitors uniquely targeting the substrate binding site of EZH2, but also confirm our original hypothesis that the histone binding sites of certain HKMTs are similar [33] and it is therefore possible to discover dual inhibitors targeting this supposedly divergent pocket. Indeed, the results herein suggest there are common aspects to the histone binding pockets of the repressive HKMTs EZH2 and EHMT1/2, different from other HKMTs. Indeed, our selectivity data suggest EZH2 and EHMT1/2 to be the sole HKMT targets of our hit compounds, as does our cell based data. It is interesting that small changes to the chemical structure of these molecules endow our hits with dual activity; something not observed for the structurally related UNC0638. Indeed, quinazoline EHMT2 inhibitors UNC0638 [24], [22] and UNC0642 [25] have been previously shown not to significantly inhibit EZH2 in biochemical assays.

Amplification or overexpression of EZH2 has been observed in a wide range of tumour types [3-8]. Furthermore, it has been proposed that epigenetic dysregulation can be a contributing factor to acquired drug resistance [7, 8, 41]. In cancers, the specific signalling mechanisms that lead to rapid tumour cell proliferation or evasion of drug-induced apoptosis may vary from cell to cell. One of the appeals of epigenetic therapies in cancer is that, rather than trying to target each individual signalling aberration, the target is the means of acquiring aberrant signalling. Therefore, it is hoped that such therapies may fare better in a heterogeneous tumour environment than drugs targeting specific signalling proteins. In this light, we highlight the observation that a set of EZH2 target genes derived from siRNA knock-down in MDA-MB-231 cells was systematically upregulated following treatment of MDA-MB-231 cells with HKMTI-1-005, but not a set of EZH2 targets identified from siRNA knock-down in MCF7 cells. This suggests that the compounds are able to elicit a transcriptional response that is specific to a particular cell line, and thus represent a means of tailoring the response to the targets that are specifically epigenetically repressed in the cancer cells to be treated. However, this fact additionally suggests that it may be difficult to find generally appropriate pharmacodynamic biomarkers indicative of a cellular response to treatment with the compounds. To address this, we carried out a meta-analysis to identify genes with a consistent upregulation following EZH2 knock-down via siRNA across a panel of 18 cell lines. These genes may reflect useful biomarkers for extending the drug screening process into a wider range of cancer cell lines.

Genome-wide expression analysis revealed that genes upregulated upon treatment with HKMTI-1-005 were more enriched for genes silenced by EZH2 than treatment with either the specific EHMT2 inhibitor UNC0638 or the specific EZH2 inhibitor GSK343. It was interesting to note that the EHMT2 inhibitor UNC0638 seemed to be as effective as the specific EZH2 inhibitor GSK343 in terms of specific upregulation of genes silenced by EZH2. This could in part be explained by the fact that EHMT2 has the capacity to methylate H3K27 [26, 27], and that reversal of epigenetic silencing of certain EZH2 targets is dependent on inhibition of EHMT2 [29]. Alternatively, it could be due to differences in the kinetics of the inhibitors that act through different mechanisms, and the fact that genome-wide expression analysis was only carried out within a limited time window.

We also note that the effects observed on gene expression, chromatin marks, and global levels of H3K27me3 and H3K9me3 occur within 24-72h, while some previously reported EZH2 inhibitors only show pharmacodynamic effects at later time points [10, 12, 14-16].

There may be many reasons for these differences, including the mechanism of action of the dual inhibitors, as well as their effects on mRNA levels of EZH2 and the H3K27 demethylase JMJD3. However it should be noted that the kinetics of effects on gene expression we observe with the dual inhibitors are similar to the kinetics of effects on gene expression we observe with double siRNA knockdown of EZH2 and EHMT2. The wealth of cellular data accumulated for our hit compounds, HKMTI-1-005 in particular, argue for direct effects on cells at the target H3K27me and H3K9me modifications at doses of drug less than or equivalent to growth inhibitory doses. Such data includes the specific expression of EZH2 target genes, global histone methylation changes by Western analysis, and local chromatin changes on responsive genes. We also note the increased sensitivity of the mutant EZH2

DB lymphoma cell line to HKMTI-1-005, in accordance with an EZH2 inhibitory effect. Such cellular biological effects are observed at doses of hit compounds less than the *in vitro* biochemical IC₅₀ detected for EZH2. We would argue that the cellular activity is a consequence of dual HKMT activity and so extrapolating from single enzyme IC₅₀ values is difficult. Furthermore, since the *in vitro* biochemical EZH2 activity assay conditions used the minimal number of proteins: (EZH2:EED:SUZ12) and a simple peptide substrate, rather than the complex (and dynamic) *in vivo* target of chromatin, care should be taken in drawing quantitative comparisons with cell-based data.

The hit compounds reported herein represent starting points for the further optimisation of dual EZH2/EHMT2 inhibitors. Indeed, recent reports suggest it is possible to improve the *in vivo* profile of this compound class [25]. While this scaffold has been extensively pursued for selective EHMT1/2 inhibition, further studies are needed to confirm whether it is possible to simultaneously increase potency against both EZH2 and EHMT1/2 and whether it is possible to engineer EHMT1/2 activity out of this scaffold to identify a selective substrate competitive EZH2 inhibitor. Nonetheless, it will continue to be important to ‘repurpose’ existing HKMT inhibitor chemotypes, in light of the low number of validated HKMT inhibitory chemotypes currently available [16]

CONCLUSIONS

Many cancers show aberrant silencing of gene expression and overexpression of histone methyltransferases, including EZH2 and EHMT1/2. We have shown that combined inhibition of EHMT1/2 and EZH2 increases growth inhibition in tumour cells over inhibition of only EHMT1/2 or EZH2, and results in re-expression of silenced genes. We report the first dual EZH2-EHMT1/2 substrate competitive inhibitors and show that they may have greater activity in tumour cells that overexpress wild-type EZH2.

METHODS

qRT-PCR measurements for cell based screening. Following compound treatment of MDA-MB-231 for 48h (in 6-well plates), media was removed and 1.5ml of TRIzol (Invitrogen) was added directly to lyse cells and RNA isolated according to the manufactures instructions. Reverse transcription was done using the SuperScript III First-Strand Synthesis System (Invitrogen) according to the manufactures instructions. Each measurement was done in triplicate, and the List of Primers can be found in Supplementary Table S1. For normalisation we have used GAPDH and RNA pol II. Experiments were also done with the

‘Fast Sybr Green Cell-to-CTTM-Kit’ according to the manufacturer’s instructions (Applied Biosystem). 15,000 cells per 96 well were plated and after 24h treated with compounds at various concentrations.

SiRNA Experiments. SiRNA experiments were carried out on the MDA-MB-231 cell line using Qiagen reagents, according to the manufactures instructions. In brief, cells were seeded at a density of 1×10^5 cells/6 cm well and treated for 48h with siRNAs given in Supplementary Table S2.

Chromatin Immunoprecipitation (ChIP-PCR) assay. ChIP was accomplished using Dynabeads Protein A (Invitrogen) according to [42], except that following the Chelex-DNA purification an additional purification with QIAquick PCR Purification Kit (Qiagen) was carried out, here the ChIP-products were eluted in 50µl and for subsequent qPCR measurements (as described above). The list of Primers can be found in Supplementary Table S3. Results were calculated as a fold increase of the No-antibody control and then normalised to GAPDH (active marks) and beta-globin (inactive marks).

Cell Viability Assay. Lymphoma cells from established lymphoma cell lines were plated at 20,000 cells in 200µl per well in U- bottom 96 well plates in RPMI medium + 20% FCS. 48 hours later cells were resuspended, diluted 10 fold in PBS + propidium iodide (PI), and the concentration of PI negative cells was counted using an Attune flow cytometer with autosampler.

Breast cancer cells from established breast cancer cell lines were seeded at a density of 10000 cells/well in a sterile 96 clear-well plate with 150 µl of DMEM (+10% FCS and 2mM L-Glutamine). Each compound treatment was performed in triplicate for 72h at concentrations of 100nM, 1µM, 5µM, 10µM and 50µM in 100µl of full-medium. After 72h,

20µl of MTT solution (3mg of MTT Formazan, Sigma/1ml PBS) was added to the medium, and incubated for 4h at 37°C in a CO₂-incubator. The MTT-product was solubilised with 100µl DMSO and for 1h incubated in the dark at room-temperature. The optical density was read at 570nm with PHERAstar.

Westerns. MDA-MB-231 cells seeded in 6 well plates at a cell density of 3×10^5 were treated with HKMTI-1-005 (1-7.5uM) for 48hr. Following lysis in Triton Extraction Buffer (TEB: PBS containing 0.5% Triton X 100 (v/v), 1/1000 protease inhibitor) nuclei were re-suspended in 0.2N HCL at a density of 4×10^7 nuclei per ml and incubated over night at 4°C to acid extract the histones, before being centrifuged at 6,500g for 10 minutes at 4°C. Protein concentration was determined using the Bradford assay. H3K27me3, H3K9me3, H3K9me2, H3K9me and total H3 protein expression levels in the histone extract samples were determined using western blot analysis using H3K27me3 (1:1000; Abcam), H3K9me3 (1:1000; Abcam), H3K9me (1:1000) and H3 (1:2000; Abcam) antibodies. After washing the membrane was incubated with a horseradish peroxidase-labelled secondary antibody (1h, room temperature). The membrane was incubated for 1 minute with 5 mL of Pierce ECL Western blotting substrate (Thermo Scientific). Images were captured using Konica Minolta SRX101A Tabletop X-Ray film processor.

Gene Expression Microarrays. Agilent 80k two-colour microarrays were used to profile gene expression changes induced by treatment with drug compounds in MDA MB-231 cells, both at 24h and 48h. In the initial microarray experiment 3 replicates were used for each drug, time combination and in the validation study 4 replicates were used. A separate untreated control sample was used for comparison with each replicate. Sample labelling, array hybridization and scanning were performed by Oxford Gene Technologies, according to manufacturer's instructions. Feature Extracted files were imported into GeneSpring (Agilent) and data was normalised to produce log2 ratios of treated/untreated for each replicate of each drug, time combination.

Statistical Analysis. *Differential Expression.* Normalised log2 gene expression ratios were analysed using LIMMA [43] to obtain empirical Bayes moderated t-statistics for differential expression across the replicates for each drug treatment. After multiple testing adjustment by the Benjamini-Hochberg method, $p < 0.1$ was used to denote significant differential expression in the initial microarray experiment and $p < 0.05$ in the validation experiment. *Enrichment Analysis.* A list of EZH2 targets in MDA MB-231 cells was taken from [35]. Statistical significance of systematic upregulation or downregulation of these targets was evaluated using the 'GeneSetTest' method from the Bioconductor package *limma*. The same method was used to evaluate systematic up- or down-regulation of pathways as annotated in ConsensusPathDB [36]. Further analysis was performed using DAVID [44] for exploration of functional annotation enrichments.

Identification of a set of consensus EZH2-suppressed genes via meta-analysis. A meta-analysis of 18 microarray experiments was carried out as described in Supplementary Methods, resulting in the list of consensus EZH2 target genes given in Supplementary Table S4.

COMPETING INTERESTS

The author(s) declare that they have no competing interests

AUTHORS' CONTRIBUTIONS

EC helped with study design, processed microarray data, performed statistical analysis and drafted the manuscript. IG helped with study design, carried out ChIP, qRT-PCR assays and helped to draft the manuscript. NCR helped with study design, carried out qRT-PCR and ChIP assays, and helped to draft the manuscript. ES carried out qRT-PCR and MTT assays. SK performed Western blots. LP performed qRT-PCR and MTT assays. EB performed microarray meta-analysis to obtain consensus EZH2 targets. FC, TG, NS, JC, JS and MF designed and synthesized the compounds. FL and MV performed the functional HKMT biochemical assays. AGU carried out lymphoma drug sensitivity assays. RB and MF conceived, designed and coordinated the study, and drafted the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

We would like to acknowledge Ovarian Cancer Action and Cancer Research UK for funding (grant C21484/A6944, C536/A13086). IG acknowledges PhD studentship from Imperial Cancer Research UK Centre. SS acknowledges the European Commission for a Marie Curie International Incoming Fellowship (Agreement No. 299857). N.S. was supported by a Royal Thai Government Scholarship and the EPSRC-funded Institute of Chemical Biology Doctoral Training Centre. JC acknowledges support from the ARC. The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Boehringer Ingelheim, the Canada Foundation for Innovation, the Canadian Institutes for Health Research, Genome Canada through the Ontario Genomics Institute [OGI-055], GlaxoSmithKline, Janssen, Lilly Canada, the Novartis Research Foundation, the Ontario Ministry of Economic Development and Innovation, Pfizer, Takeda, and the Wellcome Trust [092809/Z/10/Z].

REFERENCES

1. Cao R, Zhang Y: **The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3.** *Current Opinion in Genetics & Development* 2004, **14**:155-164.
2. Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA, Ghosh D, Sewalt RGAB, Otte AP, Hayes DF, et al: **EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells.** *Proceedings of the National Academy of Sciences* 2003, **100**:11606-11611.
3. Bachmann IM, Halvorsen OJ, Collett K, Stefansson IM, Straume O, Haukaas SA, Salvesen HB, Otte AP, Akslen LA: **EZH2 Expression Is Associated With High Proliferation Rate and Aggressive Tumor Subgroups in Cutaneous Melanoma and Cancers of the Endometrium, Prostate, and Breast.** *Journal of Clinical Oncology* 2006, **24**:268-273.
4. Matsukawa Y, Semba S, Kato H, Ito A, Yanagihara K, Yokozaki H: **Expression of the enhancer of zeste homolog 2 is correlated with poor prognosis in human gastric cancer.** *Cancer science* 2006, **97**:484-491.
5. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienta KJ, Sewalt RGAB, Otte AP, et al: **The polycomb group protein EZH2 is involved in progression of prostate cancer.** *Nature* 2002, **419**:624-629.
6. Collett K, Eide GE, Arnes J, Stefansson IM, Eide J, Braaten A, Aas T, Otte AP, Akslen LA: **Expression of enhancer of zeste homologue 2 is significantly associated with increased tumor cell proliferation and is a marker of aggressive breast cancer.** *Clinical Cancer Research* 2006, **12**:1168-1174.
7. Yehiely F, Moyano JV, Evans JR, Nielsen TO, Cryns VL: **Deconstructing the molecular portrait of basal-like breast cancer.** *Trends in Molecular Medicine* 2006, **12**:537-544.
8. Bracken AP, Pasini D, Capra M, Prosperini E, Colli E, Helin K: **EZH2 is downstream of the pRB- E2F pathway, essential for proliferation and amplified in cancer.** *The EMBO Journal* 2003, **22**:5323-5335.
9. Li H, Cai Q, Godwin AK, Zhang R: **Enhancer of zeste homolog 2 promotes the proliferation and invasion of epithelial ovarian cancer cells.** *Molecular Cancer Research* 2010, **8**:1610-1618.
10. Verma SK, Tian X, LaFrance LV, Duquenne C, Suarez DP, Newlander KA, Romeril SP, Burgess JL, Grant SW, Brackley JA, et al: **Identification of Potent, Selective, Cell-Active Inhibitors of the Histone Lysine Methyltransferase EZH2.** *ACS Medicinal Chemistry Letters* 2012, **3**:1091-1096.
11. McCabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, Liu Y, Graves AP, Iii ADP, Diaz E, et al: **EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations.** *Nature* 2012, **492**:108-112.
12. Knutson SK, Wigle TJ, Warholc NM, Sneeringer CJ, Allain CJ, Klaus CR, Sacks JD, Raimondi A, Majer CR, Song J, et al: **A selective inhibitor of EZH2 blocks H3K27 methylation and kills mutant lymphoma cells.** *Nat Chem Biol* 2012, **8**:890-896.
13. Qi W, Chan H, Teng L, Li L, Chuai S, Zhang R, Zeng J, Li M, Fan H, Lin Y, et al: **Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation.** *Proceedings of the National Academy of Sciences* 2012, **109**:21360-21365.
14. Konze KD, Ma A, Li F, Barsyte-Lovejoy D, Parton T, MacNevin CJ, Liu F, Gao C, Huang X-P, Kuznetsova E, et al: **An Orally Bioavailable Chemical Probe of the Lysine Methyltransferases EZH2 and EZH1.** *ACS Chemical Biology* 2013, **8**:1324-1334.

15. Knutson SK, Warholic NM, Wigle TJ, Klaus CR, Allain CJ, Raimondi A, Porter Scott M, Chesworth R, Moyer MP, Copeland RA, et al: **Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2.** *Proceedings of the National Academy of Sciences* 2013, **110**:7922-7927.
16. Garapaty-Rao S, Nasveschuk C, Gagnon A, Chan Eric Y, Sandy P, Busby J, Balasubramanian S, Campbell R, Zhao F, Bergeron L, et al: **Identification of EZH2 and EZH1 Small Molecule Inhibitors with Selective Impact on Diffuse Large B Cell Lymphoma Cell Growth.** *Chemistry & biology* 2013, **20**:1329-1339.
17. Yoo CB, Jones PA: **Epigenetic therapy of cancer: past, present and future.** *Nat Rev Drug Discov* 2006, **5**:37-50.
18. Cho H-S, Kelly JD, Hayami S, Toyokawa G, Takawa M, Yoshimatsu M, Tsunoda T, Field HI, Neal DE, Ponder BA: **Enhanced expression of EHMT2 is involved in the proliferation of cancer cells through negative regulation of SIAH1.** *Neoplasia (New York, NY)* 2011, **13**:676.
19. Ding J, Li T, Wang X, Zhao E, Choi J-H, Yang L, Zha Y, Dong Z, Huang S, Asara JM: **The Histone H3 Methyltransferase G9A Epigenetically Activates the Serine-Glycine Synthesis Pathway to Sustain Cancer Cell Survival and Proliferation.** *Cell metabolism* 2013, **18**:896-907.
20. Liu F, Chen X, Allali-Hassani A, Quinn AM, Wasney GA, Dong A, Barsyte D, Kozieradzki I, Senisterra G, Chau I, et al: **Discovery of a 2,4-Diamino-7-aminoalkoxyquinazoline as a Potent and Selective Inhibitor of Histone Lysine Methyltransferase G9a.** *Journal of Medicinal Chemistry* 2009, **52**:7950-7953.
21. Liu F, Chen X, Allali-Hassani A, Quinn AM, Wigle TJ, Wasney GA, Dong A, Senisterra G, Chau I, Siarheyeva A, et al: **Protein Lysine Methyltransferase G9a Inhibitors: Design, Synthesis, and Structure Activity Relationships of 2,4-Diamino-7-aminoalkoxy-quinazolines.** *Journal of Medicinal Chemistry* 2010, **53**:5844-5857.
22. Vedadi M, Barsyte-Lovejoy D, Liu F, Rival-Gervier S, Allali-Hassani A, Labrie V, Wigle TJ, DiMaggio PA, Wasney GA, Siarheyeva A, et al: **A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells.** *Nat Chem Biol* 2011, **7**:566-574.
23. Chang Y, Ganesh T, Horton JR, Spannhoff A, Liu J, Sun A, Zhang X, Bedford MT, Shinkai Y, Snyder JP, Cheng X: **Adding a Lysine Mimic in the Design of Potent Inhibitors of Histone Lysine Methyltransferases.** *Journal of Molecular Biology* 2010, **400**:1-7.
24. Liu F, Barsyte-Lovejoy D, Allali-Hassani A, He Y, Herold JM, Chen X, Yates CM, Frye SV, Brown PJ, Huang J, et al: **Optimization of Cellular Activity of G9a Inhibitors 7-Aminoalkoxy- quinazolines.** *Journal of Medicinal Chemistry* 2011, **54**:6139-6150.
25. Liu F, Barsyte-Lovejoy D, Li F, Xiong Y, Korboukh V, Huang X-P, Allali-Hassani A, Janzen WP, Roth BL, Frye SV, et al: **Discovery of an in Vivo Chemical Probe of the Lysine Methyltransferases G9a and GLP.** *Journal of Medicinal Chemistry* 2013, **56**:8931-8942.
26. Tachibana M, Sugimoto K, Fukushima T, Shinkai Y: **SET Domain-containing Protein, G9a, Is a Novel Lysine-preferring Mammalian Histone Methyltransferase with Hyperactivity and Specific Selectivity to Lysines 9 and 27 of Histone H3.** *Journal of Biological Chemistry* 2001, **276**:25309-25317.
27. Wu H, Chen X, Xiong J, Li Y, Li H, Ding X, Liu S, Chen S, Gao S, Zhu B: **Histone methyltransferase G9a contributes to H3K27 methylation in vivo.** *Cell Res* 2011, **21**:365-367.
28. Yoo KH, Hennighausen L: **EZH2 methyltransferase and H3K27 methylation in breast cancer.** *Int J Biol Sci* 2012, **8**:59-65.
29. Mozzetta C, Pontis J, Fritsch L, Robin P, Portoso M, Proux C, Margueron R, Ait-Si-Ali S: **The Histone H3 Lysine 9 Methyltransferases G9a and GLP Regulate Polycomb Repressive Complex 2-Mediated Gene Silencing.** *Molecular Cell* 2014, **53**:277-289.

30. Tan J, Yang X, Zhuang L, Jiang X, Chen W, Lee PL, Karuturi RKM, Tan PBO, Liu ET, Yu Q: **Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells.** *Genes & Development* 2007, **21**:1050-1063.
31. Dillon S, Zhang X, Trievel R, Cheng X: **The SET-domain protein superfamily: protein lysine methyltransferases.** *Genome Biology* 2005, **6**:227.
32. Chang Y, Zhang X, Horton JR, Upadhyay AK, Spannhoff A, Liu J, Snyder JP, Bedford MT, Cheng: **Structural basis for G9a-like protein lysine methyltransferase inhibition by BIX-01294.** *Nat Struct Mol Biol* 2009, **16**:312-317.
33. Yun M, Wu J, Workman JL, Li B: **Readers of histone modifications.** *Cell Res* 2011, **21**:564-578.
34. Foletta V, White L, Larsen A, Léger B, Russell A: **The role and regulation of MAFbx/atrogen-1 and MuRF1 in skeletal muscle atrophy.** *Pflugers Arch - Eur J Physiol* 2011, **461**:325-335.
35. Lee Shuet T, Li Z, Wu Z, Aau M, Guan P, Karuturi RKM, Liou Yih C, Yu Q: **Context-Specific Regulation of NF- κ B Target Gene Expression by EZH2 in Breast Cancers.** *Molecular Cell* 2011, **43**:798-810.
36. Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R: **ConsensusPathDB: toward a more complete picture of cell biology.** *Nucleic Acids Research* 2011, **39**:D712- D717.
37. Benjamini Y, Hochberg Y: **Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.** *Journal of the Royal Statistical Society Series B (Methodological)* 1995, **57**:289-300.
38. Forbes SA, Bindal N, Bamford S, Cole C, Kok CY, Beare D, Jia M, Shepherd R, Leung K, Menzies A, et al: **COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer.** *Nucleic Acids Research* 2011, **39**:D945-D950.
39. Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehar J, Kryukov GV, Sonkin D, et al: **The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.** *Nature* 2012, **483**:603-307.
40. Srimongkolpithak N, Sundriyal S, Li F, Vedadi M, Fuchter MJ: **Identification of 2,4-diamino-6,7-dimethoxyquinoline derivatives as G9a inhibitors.** *MedChemComm* 2014.
41. Brown R, Curry E, Magnani L, Wilhelm-Benartzi CS, Borley J: **Poised epigenetic states and acquired drug resistance in cancer.** *Nat Rev Cancer* 2014, advance online publication.
42. Nelson JD, Denisenko O, Bomsztyk K: **Protocol for the fast chromatin immunoprecipitation (ChIP) method.** *Nat Protocols* 2006, **1**:179-185.
43. Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments.** *Statistical Applications in Genetics and Molecular Biology* 2004, **3**:3.
44. Huang DW, Sherman BT, Lempicki RA: **Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.** *Nat Protocols* 2008, **4**:44-57.

FIGURE LEGENDS

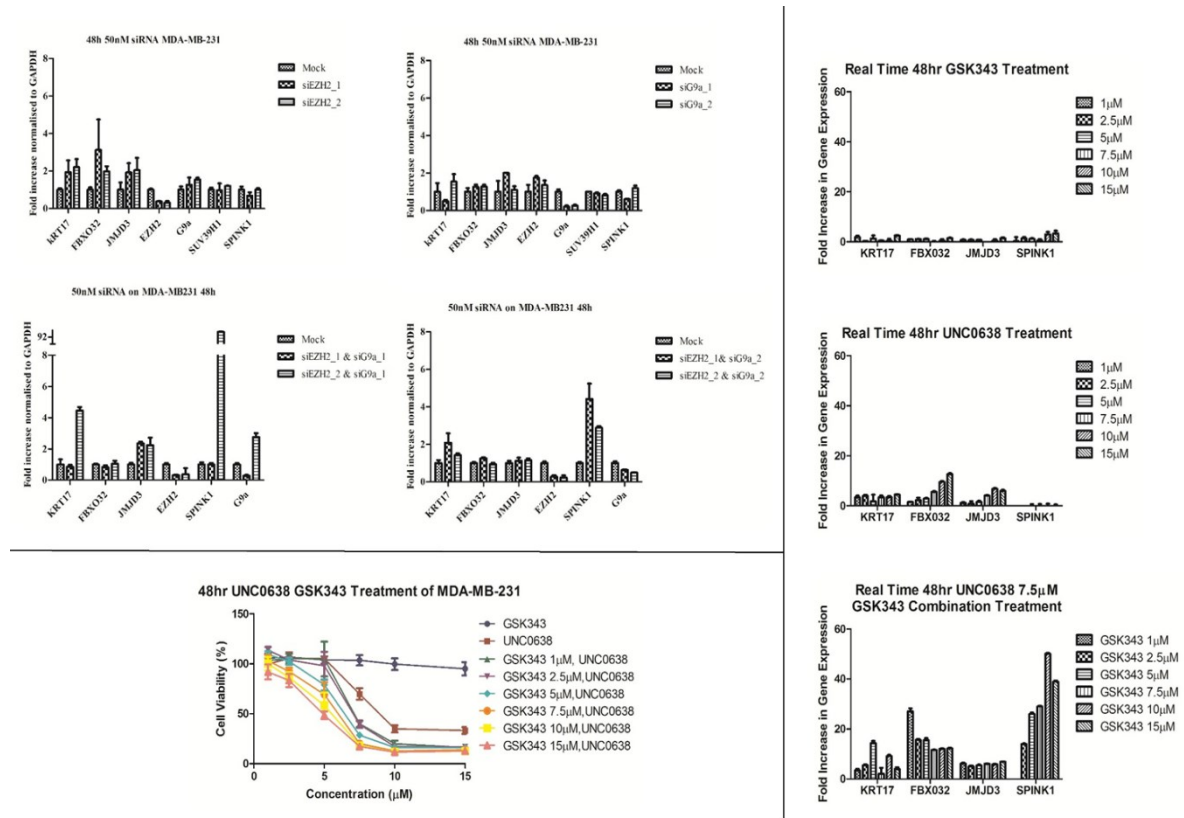


Figure 1 - MTT and mRNA levels in MDA-MB-231 cells after pharmacological inhibition and siRNA knock-down of EZH2 and EHMT2(G9a), individually and in combination.

A) Expression levels of KRT17, FBX032, JMJD3, EZH2, SPINK1 and EHMT2 were measured by qRT-PCR in the MDA-MB-231 cell line 48hrs after transfection with siRNAs targeting EZH2 and EHMT2, both individually and in combination. 2 different siRNAs were used to target each gene, all measurements were normalized to the fold-change (relative to GAPDH) in the mock transfection control. Error bars represent the mean $\pm$ SD of experiment performed in technical triplicate. SPINK1 measurement in right-most figure (dual knock-down) has been truncated for figure. B) Expression levels of KRT17, FBX032, JMJD3 and SPINK1 were measured by qRT-PCR in the MDA-MB-231 cell line treated for 48hr with GSK343, UNC0638, and UNC0638 (at 7.5 μ M) with increasing doses of GSK343. Each group has been compared to the untreated sample following normalisation to GAPDH. Error bars represent the mean $\pm$ SD of experiment performed in technical triplicate. C) MTT assay for cell viability of MDA-MB-231 cells after treatment. MDA-MB-231 cells were seeded in 96 well plates. After 24hrs, increasing doses of GSK343, UNC0638 or combination treatments (1, 2.5, 5, 7.5, 10 and 15 μ M) were added to cells. Control was media with 0.5% DMSO. Cell viability was measured by MTT assay after 48hrs treatment and a 24hr proliferation period. Error bars represent the mean $\pm$ SEM of five independent repeats.

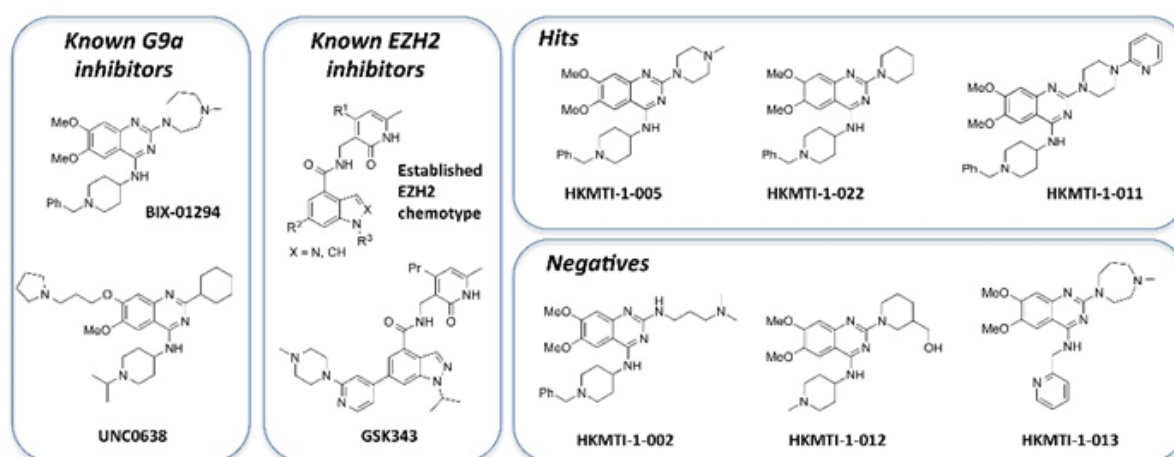


Figure 2 - Chemical structure of Histone Lysine Methyltransferase inhibitors

the untreated sample. B) Sybr green real-time PCR measurement of the FBXO32 transcription start site and KRT17 promoter region following Chromatin Immunoprecipitation, using antibodies to the histone marks shown, of MDA-MB-231 cells treated with 3 selected compounds at 5µM for 72h. Shown are representative examples of a series of ChIP experiments which consistently showed similar changes. The fold difference to the untreated sample is shown. Each IP-value has been determined as the relative increase to the no-antibody control and then normalised to GAPDH levels.

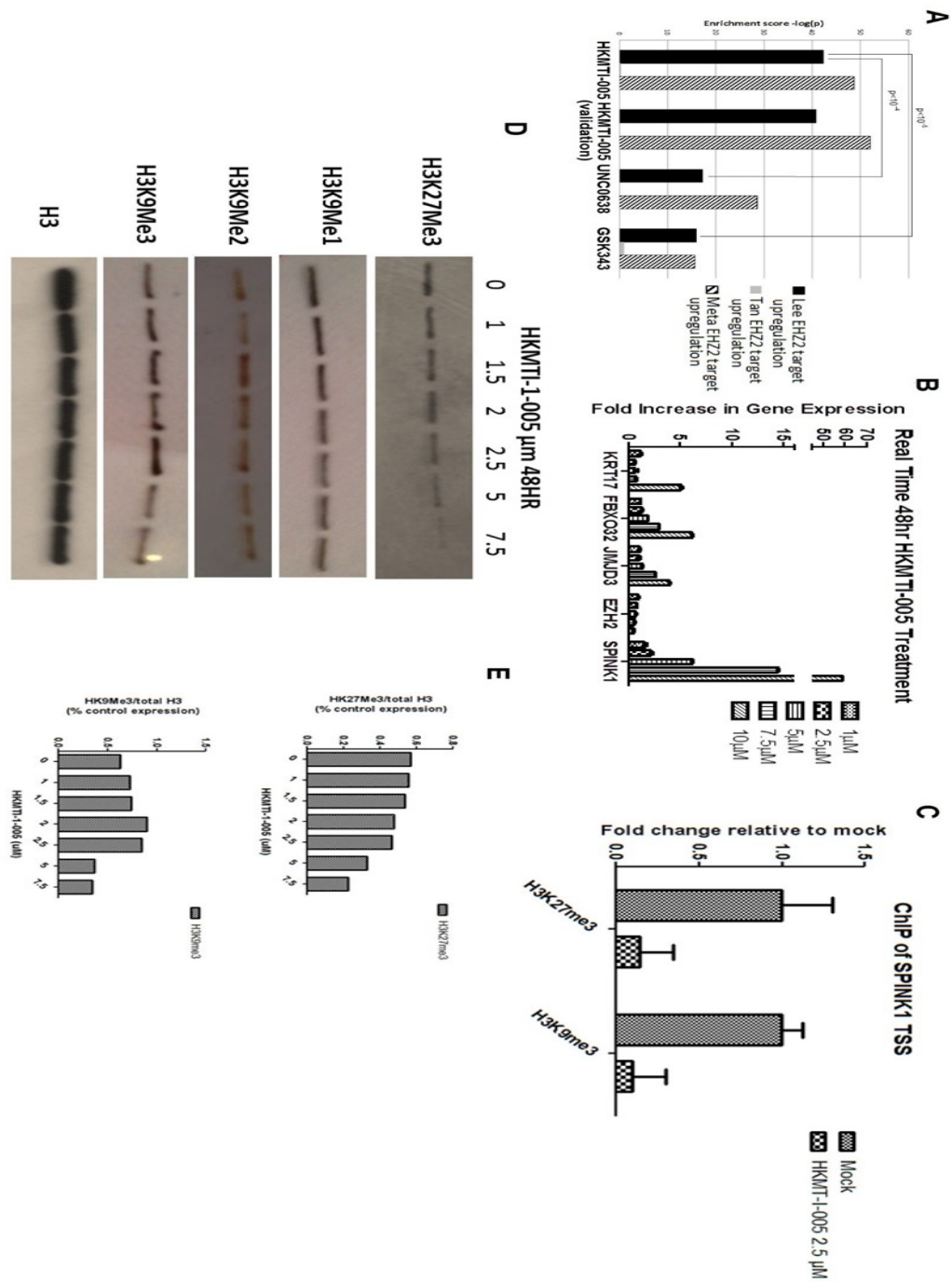


Figure 4 – Compound-induced upregulation of EZH2-repressed target genes

A) Enrichment scores for differential expression of EZH2 targets on treatment with panel of compounds. Enrichment scores are negative logarithm of p-values, such that higher values indicate

more significant enrichment. Left-hand bars show enrichment of targets derived from siRNA knock-down of EZH2 in MDA-MB-231 cell line, middle bars show enrichment of targets derived from siRNA knock-down of EZH2 in MCF7 cell line and right-hand bars show enrichment of targets defined by meta-analysis of 18 independent microarray studies profiling effects of shRNA-mediated EZH2 knock-down in a variety of cell lines. B) Sybr green real-time PCR mRNA level measurement of EZH2 target genes and executing enzymes following a 48h treatment with HKMTI-1-005 at different concentrations of MDA-MB-231 cells C) Sybr green realtime PCR measurement of the SPINK1 transcription start site following Chromatin Immunoprecipitation, using antibodies to the histone marks shown, of MDA-MB-231 cells treated with HKMTI-1-005 or HKMT-I-011 at 2.5 μ M for 24h. Each IP-value has been determined as the relative increase to the no-antibody control and is shown as fold difference relative to the untreated control. D) Western blot showing levels of modified histones, following 48hr treatment with HKMTI-1-005 at different doses. Total H3 levels are shown for comparison. E) Densitometry quantification of Western blot intensity, showing ratio of modified (H3K27me3 top, H3K9me3 bottom) H3 relative to total H3 with increasing dose of HKMTI-1-005 treatment.

ADDITIONAL FILES

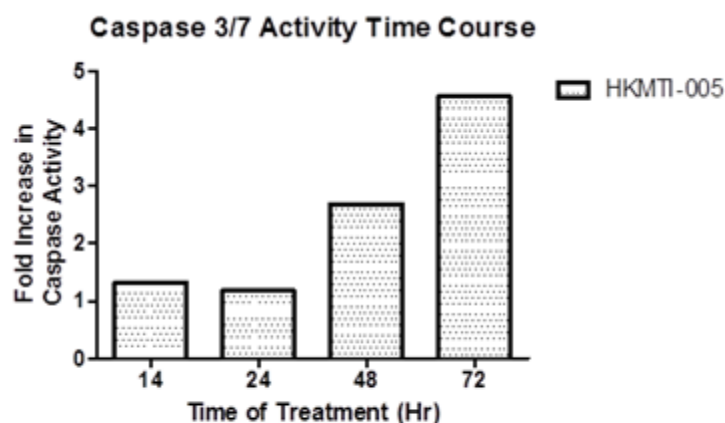


Figure S4

Supplementary Figure 1 – Induction of apoptosis in breast cancer cells by compound treatment

Caspase activity assay shows the increase in Caspase 3/7 following treatment of MDA-MB-

231 cell line with compound HKMTI-1-005. MDA-MB-231 cells were seeded in 96 white walled plates (100 μ l/well) at a density of 5 x 10³ cells per well, then incubated for 24hrs at

37oC, 5% CO₂. The culture media was removed and cells were incubated with culture

media containing 7.5 μ M HKMTI-1-005 compound for 14h, 24h, 48h and 72h. After treatment

Caspase-Glo 3/7 assay kit (Promega) was used as per manufactures instructions. After 1hr the plates

were read on a LUMistar OPTIMA (BMG LABTECH), and values were normalized to DMSO control.

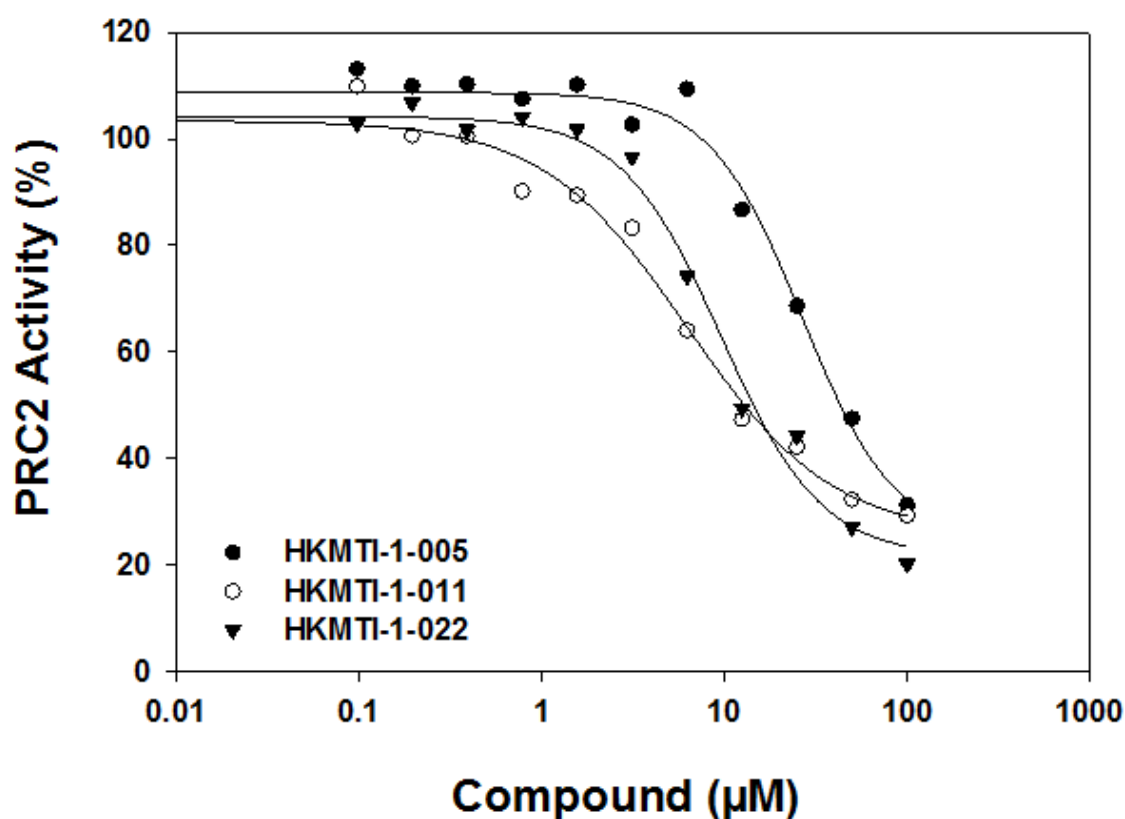
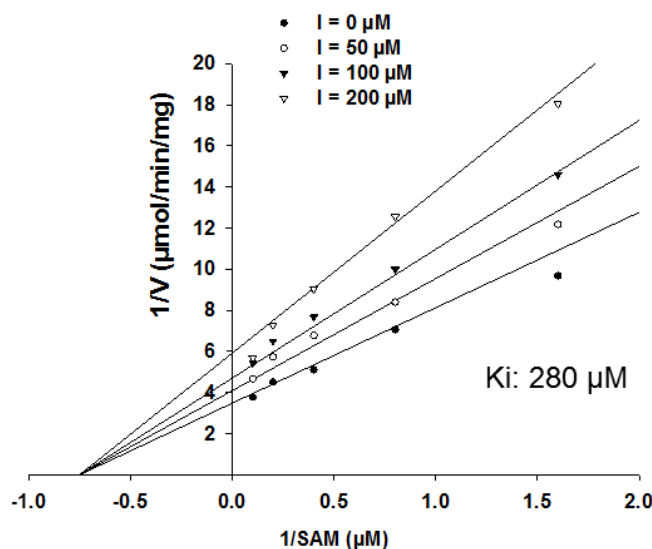


Figure S1

Supplementary Figure 2 - IC₅₀ determination for PRC2 inhibitors.

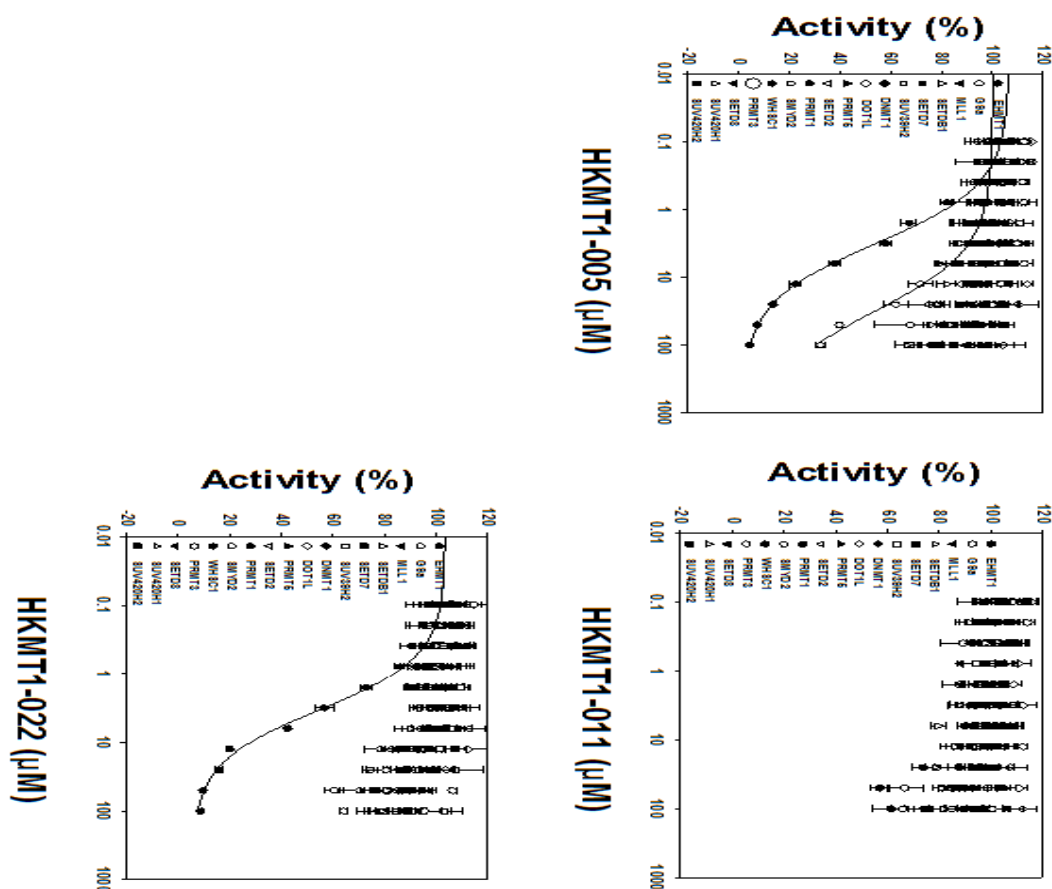
IC₅₀ values were determined for the compounds in triplicate at 0.2 μM of peptide H3 (21-44) and 1 μM of ³H-SAM using 20 nM of EZH2 complex (EZH2:EED:SUZ12) and incubating the reaction mixtures for 1 h at 23°C. To stop the enzymatic reactions, 7.5 M Guanidine hydrochloride was added, followed by 180 μl of buffer (20 mM Tris, pH 8.0), mixed and then transferred to a 96-well FlashPlate (Cat.# SMP103; Perkin Elmer; www.perkinelmer.com). After mixing, the reaction mixtures in Flash plate were incubated for 2 h and the CPM counts were measured using Topcount plate reader ((Perkin Elmer, www.perkinelmer.com). The CPM counts in the absence of compound for each dataset was defined as 100% activity. In the absence of the enzyme, the CPM counts in each dataset was defined as background (0%). The IC₅₀ values were determined using SigmaPlot software and fixing the top and bottom to 100 and 0 respectively.

A)



Supplementary Figure 3 – Mechanism of PRC2 inhibition by HKMTI-1-005.

Inhibition of PRC2 trimeric complex (EZH2:EED:SUZ12) by HKMTI-1-005 (at 0, 50, 100 and 200 μM) at varying concentrations of (A) SAM (from 0.625 to 10 μM) and (B) peptide substrate (0.3 to 5 μM) were assessed by monitoring the incorporation of tritium-labeled methyl group to peptide substrate using SAM2[®] Biotin Capture Membrane from Promega. Lineweaver-Burk plots for kinetic analysis of the inhibition indicates that HKMTI-1-005 is a peptide competitive and SAM noncompetitive PRC2 inhibitor. Peptide concentrations for A and SAM concentration for B were 5 μM and 10 μM respectively. Assays were performed in triplicate. Data were plotted using SigmaPlot, Enzyme Kinetics Module).



Supplementary Figure 4 – Selectivity of HKMTI-1-005, HKMTI-011 and HKMTI-022.

Effects of HKMTI-1-005, HKMTI-1-011, and HKMTI-1-022 on the methyltransferase activity of SUV39H2, SETDB1, SETD8, SUV420H1, SUV420H2, SETD7, SETD2, MLL1 trimeric complex, PRMT1, PRMT3, PRMT5-MEP50 complex, SMYD2, DOT1L, WHSC1 and DNMT1 was assessed by monitoring the incorporation of tritium-labeled methyl group to lysine or arginine residues of peptide substrates by scintillation proximity assay (SPA). Assays were performed in a 20 μ l reaction mixture containing ^{3} H-SAM (Cat.# NET155V250UC; Perkin Elmer; www.perkinelmer.com) at substrate concentrations close to K_m values for each enzyme. Some variations were considered to improve signal-to-noise ratios. Compound concentrations from 50 nM to 50 μ M were used in all selectivity assays. For DNMT1 the dsDNA substrate was prepared by annealing two complementary strands (biotinylated forward strand: B-GAGCCCGTAAGCCCGTTCAGGTCG and reverse strand:

CGACCTGAACGGGCTTACGGGCTC), synthesized by Eurofins MWG Operon. For DOT1L, and WHSC1 (NSD2) a filter-based assay was used. In this assay, 20 μ l of reaction mixtures were incubated at RT for 1 h, 100 μ l of 10% TCA was added, mixed and transferred to filter- plates (Millipore; cat.# MSFBN6B10; www.millipore.com). Plates were centrifuged at 2000 rpm (Allegra X-15R - Beckman Coulter, Inc.) for 2 min followed by 2 additional 10% TCA wash and one ethanol wash. Plates were dried, 70 μ l of MicroO was added and CPM was measured using Topcount plate reader