



**HDACi-induced DNA damage:  
Identifying Potential Endpoints for Safety  
Assessment**

Wenbin Wang

Institute of Cancer and Genetics  
School of Medicine  
Cardiff University

A Thesis Submitted to Cardiff University  
For the Degree of  
Doctor of Philosophy

2016



## **DECLARATION**

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ..... (candidate) Date .....

## **STATEMENT 1**

This thesis is being submitted in partial fulfillment of the requirements for the degree of ..... (insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed ..... (candidate) Date .....

## **STATEMENT 2**

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ..... (candidate) Date .....

## **STATEMENT 3**

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate) Date .....

## **STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS**

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ..... (candidate) Date .....



"Life handed him a lemon,  
As Life sometimes will do.  
His friends looked on in pity,  
Assuming he was through.  
They came upon him later,  
Reclining in the shade  
In calm contentment, drinking  
A glass of lemonade."

- Clarence Edwin Flynn



## **Acknowledgments**

Firstly, I would like to thank my principle supervisor, Professor Simon Reed for giving me the opportunity to undertake this PhD and for his invaluable advice and unwavering support throughout. I would also like to thank my second supervisor, Professor Anthony Lynch for sharing his expertise on aspects of genotox safety assessment and also his hospitality during my visits to GSK.

I would like to thank members of the lab, particularly Dr. Patrick Van Eijk for his invaluable discussions and ideas. I am also grateful to be accompanied by Hamed Aula and Shuvro Prokash Nandi during the course of my studies.

I also wish to extend my thanks to former members of the lab, Dr Shirong Yu, Dr Mark Bennett, Dr Katie Evans, Dr James Powell and Dr Richard Webster for their help and support during the early years of my PhD.

Lastly but more importantly I would like to acknowledge the close support of my parents, family and friends over the past four years.





## Summary

Histone deacetylase inhibitors (HDACi) have been designed to alter the actions of epigenetic modifiers with the aim of 'reprogramming' the epigenome of diseased tissues back to their normal disease-free state. These inhibitors were designed to be non-DNA reactive and therefore considered safe from a genetic toxicology point of view. However, HDACi's have been shown to induce DNA damage in healthy cells through unknown mechanisms, thereby posing significant risks to human health. Studies suggest that HDAC inhibitor-induced DNA damage is partly associated with changes in transcription and replication. Consequently, collisions between these events can result in the formation of DNA lesions and stable DNA:RNA hybrid structures (R-loops), which are implicated in the onset of cancer and various neurological diseases. Therefore, the aims of the current study were to better understand the mechanisms by which HDAC inhibitors may induce DNA damage and to identify potential endpoints for safety assessment:

- Chapter III: Efforts to study the effects of HDAC inhibition through a chemical approach proved unsuccessful in the yeast model organism but identified the HDAC mutant, *rpd3Δ*, showing histone hyper-acetylation compared to the wild type.
- Chapter IV: ChIP-chip was established for the TK6 lymphoblastoid cell line as a genome-wide tool for measuring the genotoxicity of HDAC inhibitors.
- Chapter V: Application of the ChIP-chip method showed that Trichostatin A-induced changes in histone H4 acetylation led to the re-distribution of transcription and replication on chromosome 17 in TK6 cells. This resulted in their co-localisation, suggestive of potential collisions. However, further efforts to determine this by mapping  $\gamma$ H2AX and R-loop formation proved unsuccessful.
- Chapter VI: The yeast genetic mutant *rpd3Δ* was used to mimic the effects of treating with an HDAC inhibitor. The loss of *RPD3* resulted in significantly higher levels of  $\gamma$ H2A, predominantly at telomere regions.

In conclusion, this thesis presents strong evidence to show that Trichostatin A promotes the co-localisation of transcription and replication, suggesting that there is a greater possibility of these processes colliding to form DNA damage.



## Abbreviations

ABE	abasic endonuclease
AID	activation-induced cytidine deaminase
AML	acute myeloid leukemia
AR	androgen receptors
AROS	active regulator of SIRT1
ARS	autonomously replicating sequences
ASF1/SF2	alternative splicing factor 1/pre-mRNA splicing factor 2
BER	base-excision repair
BPTF	bromodomain PDH finger transcription factor
BPTF	Bromodomain PHD finger Transcription Factor
BRCA1	breast cancer early onset 1
BRDs	bromodomains
CAF-1	Chromatin assembly factor
CBPI	cytokinesis-block proliferation index
CDC45	cell division cycle 45
CDC6	cell division cycle 6
CDK2	cyclin-dependent kinase 2
CdLS	Cornelia de Lange syndrome
CDT1	chromatin licencing and DNA replication factor 1
CE	cloning efficiency
CFS	common fragile sites
Chd1	chromatin helicase DNA binding protein 1
ChIP	chromatin immunoprecipitation
CoREST	REST corepressor 1
COX-2	cyclooxygenase-2
CPC	chromosomal passenger complex
CPDs	Cyclo Pyrimidine dimers
CPSF	cleavage and adenylation specificity factor
CR	caloric restriction
CSF1	colony stimulating factor
CSR	class-switch recombination
CstF	cleavage stimulator factor

CtBP	C-terminal binding protein
CTCL	cutaneous T-cell lymphoma
CTD	C-terminal domain
DBC1	deleted in breast cancer 1
DDK	Dbf4/Drf1-dependent CDC7 kinase
DDR	DNA damage response
DHS	DNase I hypersensitive sites
DSB	double strand break
DSIF	DRB sensitivity inducing factor
ECGs	electrocardiograms
EEC	early elongation complex
EFRS	early replicating fragile sites
EMA	ethidium monoazide
ER	oestrogen receptors
ESC	embryonic stem cells
FBS	foetal bovine serum
FDA	Food and Drug Administration
FEN1	flap endonuclease-1
GINS	go-ichi-ni-san
HATs	histone acetyl transferases
HDA1	histone deacetylase 1
HDACs	histone deacetylase
HIF-1 $\alpha$	hypoxia inducible factor 1-alpha
HP1 $\gamma$	heterochromatin protein 1-gamma
HR	homologous recombination
IEGs	immediate-early genes
Ig	immunoglobulin
IL-10	interleukin-10
ING2	Inhibitor of Growth-2
LM-PCR	ligation mediated-polymerase chain reaction
LPS	Lipopolysaccharides
MCM	mini chromosome maintenance
MDC1	mediator of DNA damage checkpoint protein 1
MDM2	mouse double minute 2 homolog

MDS	myelodysplastic syndrome
MEF2	myocyte enhancer factor-2
MEFs	murine embryonic fibroblasts
MiDAC	mitotic deacetylase complex
MMR	mismatch repair
MN	micronuclei
MOA	mode of action
MRN	MRE11-Rad50-Nbs1
MSH2	MutS homolog 2
NAC	N-acetyl cysteine
NCoR	nuclear receptor corepressor
NELF	negative elongation factor
NER	nucleotide excision repair
NFRs	nucleosome free regions
NHEJ	non-homologous end joining
NHL	non-Hodgkin's lymphoma
NODE	Nanog and Oct4 associated deacetylase
NRs	Nuclear receptors
NRSF	neuron-restrictive silencing factor
NSCLC	non-small cell lung cancer
NuRD	nucleosome remodeling HDAC
ORC	origin recognition complex
ORF	open reading frame
PBMCs	peripheral blood mononuclear cells
PBS	phosphate buffered saline
PCNA	proliferating cell nuclear antigen
PH	Pleckstrin Homology
PHD	plant homeodomains
PI	propidium iodide
PIC	pre-initiation complex
PIC	protease inhibitor cocktail
PMSF	phenylmethane sulfonyl fluoride
PRC	polycomb-repressive complex
Pre-RC	pre-replication complex

PTCL	peripheral T-cell lymphoma
PTM	post-translational modification
RCC	Relative cell count
REST	repressor element-1 silencing transcription factor
RFB	replication fork barriers
RFC	replication factor C
RI	replicative-index
RICC	relative increase in cell counts
ROS	reactive oxygen species
RPA	Replication protein A
RPD3	reduced potassium deficiency 3
RT-qPCR	reverse transcription quantitative PCR
RTG	relative total growth
SAHA	suberoylanilide hydroxamic acid
SAP	Sin3 associated proteins
SAPK	stress-activated protein kinase
SGD	<i>Saccharomyces</i> genome database
SHIP1	spermatogenic cell HDAC1 interacting protein 1
SHP	short heterodimer protein
SIN3	Switch INdependent 3
Sir2	silent information regulator 2
SMA	spinal muscular atrophy
SMAR1	scaffold/matrix-associated region binding protein-1
SMRT	Silencing mediator for retinoid and thyroid receptor
TAM	transcription-associated mutagenesis
TAR	transcription-associated recombination
TBL1	transducin $\beta$ -like 1
TBP	TATA-binding protein
TC-NER	transcription-coupled nucleotide excision repair
TERA	telomeric repeat-containing RNA
TES	transcription end site
TFIID	Transcription Factor II D
TGF- $\beta$ 1	transforming growth factor beta 1
TSA	Trichostatin A

TSS	transcription start site
VEGFs	Vascular endothelial growth factors
VPA	valproic acid
WGA	whole genome amplification
XPF	xeroderma pigmentosum F
XPG	xeroderma pigmentosum G
ZBD	Zinc-binding domain
$\gamma$ H2AX	phosphorylated histone H





# Contents

Chapter I - Introduction.....	1
Preface.....	3
1.1 Non-clinical safety assessment .....	4
1.1.1 Core tests.....	4
1.1.2 Supplementary tests .....	5
1.1.3 Cytotoxicity assays .....	6
1.2 Novel genome-wide approaches to study mechanisms.....	6
1.3 Epigenetics .....	7
1.3.1 Regulating the epigenome.....	9
1.4 Histone modifications .....	10
1.4.1 Histone code hypothesis.....	13
1.4.2 Interpreting the code .....	15
1.4.3 Histone modifications & cellular function.....	17
1.4.3.1 Transcription .....	17
1.4.3.2 DNA Replication.....	18
1.4.3.3 DNA Repair .....	19
1.4.4 Histone core modifications .....	22
1.5 Histone deacetylases .....	23
1.5.1 Classical HDACs .....	24
1.5.1.1 Class I .....	24
1.5.1.2 Class II.....	25
1.5.1.3 Class IV.....	27
1.5.2 Sirtuins.....	27
1.5.3 HDAC complexes .....	28
1.5.4 Non-histone targeted effects of HDACs .....	30
1.5.5 Histone deacetylases in cancer and other diseases.....	30
1.6 Chemical inhibitors of HDAC enzymes .....	33
1.6.1 Hydroxamic acids.....	34
1.6.2 Cyclic peptides .....	34
1.6.3 Benzamides .....	36
1.6.4 Aliphatic acids.....	36
1.6.5 HDAC inhibitors in the clinic .....	36

1.6.6	HDAC inhibitor-related side effects .....	38
1.7	Genetic approaches to study HDAC function.....	39
1.8	Evidence for HDAC inhibitor induced DNA damage .....	40
1.8.1	Potential mechanisms of HDACi-induced DNA damage.....	41
1.9	Transcription and replication conflicts as a source of genome instability .....	41
1.9.1	Eukaryotic transcription.....	41
1.9.2	Eukaryotic replication .....	43
1.9.3	Regulating transcription and replication .....	46
1.10	R-loops.....	48
1.10.1	Physiological functions of R-loops .....	48
1.10.2	Mechanisms regulating R-loop formation .....	49
1.10.3	R-loops as a source of genetic and epigenetic instability .....	50
1.10.4	R-loops and double strand breaks as products of transcription and replication conflicts .....	52
1.11	Aims and hypothesis of the present study.....	54
Chapter II - Methods and Materials .....		55
2.1	Mammalian methods.....	57
2.1.1	Mammalian cell culture and treatment.....	57
2.1.2	Freezing and thawing cells.....	57
2.1.3	Counting cells .....	58
2.1.4	Flow cytometry .....	59
2.1.4.1	Apoptosis detection .....	59
2.1.4.2	Histone H2AX phosphorylation and Histone H3K9/14 acetylation.....	59
2.1.4.3	InVitro MicroFlow <sup>®</sup> assay .....	60
2.1.4.4	Cell cycle analysis .....	60
2.1.5	Immunoblot analysis .....	61
2.1.5.1	Whole cell protein extraction.....	61
2.1.5.2	Protein quantification.....	61
2.1.5.3	SDS-Polyacrylamide gel electrophoresis .....	62
2.1.5.4	Western transfer, probing and detection .....	62
2.1.5.5	Membrane stripping.....	63
2.1.6	Chromatin Immunoprecipitation.....	64
2.1.6.1	Chromatin preparation .....	64

2.1.6.2	DNA gel electrophoresis.....	65
2.1.6.3	Immunoprecipitation.....	65
2.1.6.4	Real-time qPCR.....	67
2.1.6.5	Determining optimal antibody conditions for ChIP-chip.....	68
2.1.7	ChIP-chip.....	69
2.1.7.1	Whole-genome amplification.....	69
2.1.7.2	Fluorescent labelling.....	69
2.1.7.3	DNA quantification.....	70
2.1.7.4	Array hybridisation.....	70
2.1.7.5	Array washing, scanning and feature extraction.....	71
2.1.7.6	Data analysis.....	71
2.1.7.7	Mammalian microarray design.....	72
2.1.8	RDIP-chip.....	72
2.2	Yeast methods.....	72
2.2.1	Yeast strains and culture.....	72
2.2.2	Growth analysis.....	73
2.2.3	Whole cell protein extraction.....	73
2.2.4	ChIP-chip.....	73
2.2.4.1	Chromatin preparation.....	73
2.2.4.2	Ligation-mediated PCR (LM-PCR).....	74
2.2.4.3	Yeast microarray design.....	74
2.2.5	DRIP-chip.....	75
2.2.6	Extraction of total RNA.....	75
2.2.7	Preparation of cDNA.....	76
Chapter III	– Using <i>Saccharomyces cerevisiae</i> as a model organism to study the genotoxic effects of HDAC inhibition.....	77
3.1	<i>Saccharomyces cerevisiae</i> as a model organism.....	79
3.2	Histone deacetylase enzymes in yeast.....	79
3.3	Use of HDAC inhibitors in yeast.....	80
3.4	Chapter aims.....	80
3.5	Effect of HDAC inhibitors on the growth of yeast cells.....	81
3.6	Effect of HDAC inhibition on histone acetylation.....	81
3.7	The effects of HDAC inhibition on gene expression.....	85

3.8 Discussion .....	91
----------------------	----

Chapter IV – Establishing the ChIP-chip methodology for TK6 cells as a genome-wide assay to measure the genotoxic effects of HDAC inhibitors ..... 95

4.1 TK6 as a model cell line .....	97
4.2 Use of small molecule inhibitors to determine the status of <i>TP53</i> .....	98
4.3 Challenges of mammalian ChIP-chip .....	98
4.4 Whole human chromosome 17 microarrays .....	99
4.5 Chapter aims .....	100
4.6 Confirming the p53 proficiency of TK6 cells .....	101
4.7 Examining the toxicity of TSA in TK6 cells to establish a suitable treatment regimen .....	104
4.8 Examining the short-term effects of TSA in TK6 cells using induction of $\gamma$ H2AX as a marker of DNA damage .....	105
4.9 Optimising the ChIP-chip assay for TK6 cells .....	110
4.9.1 Effect of formaldehyde fixation on sonication efficiency .....	110
4.9.2 Investigating the effect of isolating nuclei to promote efficient shearing of TK6 cell chromatin .....	110
4.9.3 Assessing the reproducibility of shearing chromatin from TK6 cells .....	113
4.9.4 Examining the ability to enrich for RNA polymerase II (pSer2) using ChIP .....	113
4.9.5 Examining the quality of RNA Pol II (pSer2) ChIP-chip data .....	117
4.9.6 Validating chromosome 17 microarrays of RNA Pol II (pSer2) occupancy in TK6 cells .....	122
4.10 Discussion .....	127

Chapter V – Investigating transcription-replication conflicts as a source of HDAC inhibitor induced DNA damage in TK6 cells ..... 133

5.1 Eukaryotic replication origins .....	135
5.2 Histone acetylation and replication origin firing .....	136
5.3 Chapter aims .....	137
5.4 Measuring the effect of TSA on H4 acetylation, RNA Pol II (pSer2) and ORC1 occupancy on chromosome 17 in TK6 cells .....	138
5.4.1 Examining changes in Histone H4 acetylation in response to TSA .....	138

5.4.2	Examining changes in transcription in response to TSA .....	140
5.4.3	Examining changes in replication initiation in response to TSA .....	140
5.5	Analysing the impact of TSA-induced changes in H4 acetylation on RNA Pol II (pSer2) and ORC1 occupancy on chromosome 17 in TK6 cells .....	146
5.6	Examining TSA-induced co-localisation of transcription and replication on chromosome 17 .....	152
5.7	Determining the potential for DNA damage induction and R-loop formation at sites of TSA-induced transcription and replication co-localisation .....	157
5.8	Discussion .....	158
Chapter VI – A genetic approach to examine DNA damage associated with loss of HDAC function in <i>Saccharomyces cerevisiae</i> .....		163
6.1	Chapter aims.....	165
6.2	Validation of $\gamma$ H2A and R-loop ChIP-chip data .....	166
6.3	Analysis of yeast $\gamma$ H2A and R-loop data .....	170
6.3.1	Effect of <i>RPD3</i> deletion on genome-wide $\gamma$ H2A and R-loop signals in yeast.....	170
6.3.2	Examining the co-localisation of $\gamma$ H2A and R-loop in yeast as a marker of transcription-replication conflicts .....	174
6.3.3	Examining $\gamma$ H2A and R-loops at sites of replisome impedance.....	175
6.3.4	Significance of R-loop enriched genes in <i>rpd3<math>\Delta</math></i> cells.....	176
6.4	Discussion .....	181
Chapter VII – General discussion .....		185
Appendix I – Raw data for Chapter III .....		197
Appendix II – Raw data for Chapter IV .....		199
Appendix III – Raw data for Chapter V.....		209
Appendix IV – Raw data for Chapter VI .....		215
References .....		218



# **Chapter I**

## **Introduction**





## **Preface**

Histone deacetylase (HDAC) inhibitors represent a class of prototypical epigenetic modifying drugs aimed at treating cancer. Recent developments using these compounds, have shown promising results in treating animal models for neurodegeneration, inflammation and metabolic diseases. However, non-clinical safety assessment indicates HDAC inhibitors to elicit genetic damage through unknown mechanisms. While this may not be significant for cancer, the prospect of using these inhibitors for non-life threatening conditions poses a risk to human health. Current testing strategies used, provide little information on the mechanism by which test chemicals exactly generate genetic damage. A better understanding of these mechanisms would allow for thresholds to be set under which certain chemicals may be safe to use. Such initiatives have already begun (Adeleye 2015) based on concepts proposed by AOP (Adverse Outcome Pathways) (Ankley 2010) and TT21C (Toxicity testing in the 21<sup>st</sup> century) (Bhattacharya 2011), envisaging the use of large-scale genome-wide approaches with greater focus on the mechanism of action.

The pharmaceutical industry has made large investments into the development of HDAC inhibitors and interests lie in trying to understand the mechanism of genetic toxicity to better address the risks to human health but also to further the development of next generation HDAC inhibitors. The current study stems from the collaboration between Glaxo-Smith Kline (GSK) Safety Assessment and Cardiff University, supported by funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and GSK Collaborative Awards in Science and Engineering (CASE). The aim is to combine expertise in genetic toxicology from industry with genome-wide technologies available in academia to address the question of how drugs targeting the epigenome, such as HDAC inhibitors, can lead to damage at the genome level.

## 1.1 Non-clinical safety assessment

Safety assessment is carried out by the pharmaceutical industry on all newly developed compounds or ‘test articles’ to screen for potential hazards to human health. A key aspect of this is to assess if the test article can produce changes in the genetic sequence, in which case it is mutagenic, a fundamental event underpinning carcinogenicity. Mutations in germ line cells can lead to hereditary disorders, whereas those in somatic cells can increase the risks of developing cancer. In living cells, various repair mechanisms help to protect the genome from endogenous and exogenous sources of damage. As such, the potential genotoxicity of test articles may not necessarily manifest as mutations. Therefore, genotoxicity testing entails a battery of tests to assess, with greater coverage, the ‘capacity’ of a compound to damage the genome, classing the actions as being mutagenic (change in genetic material), clastogenic (chromosomal breakages) and/or aneugenic (unequal number of chromosomes in daughter cells). These regulatory tests take into account the ‘mode of action’ (MoA), which describes phenotypic changes elicited by test articles at the cellular level that can be readily scored and reliably performed between different laboratories. A positive result in these tests however, does not provide information on the molecular mechanisms of DNA damage.

### 1.1.1 Core tests

The exact compositions of tests used are specific to the industry sector, but generally the core battery consists of a bacterial gene mutation test, followed by an *in vitro* and *in vivo* assay evaluating chromosomal damage. Liver homogenate or S9 fraction is also widely used in conjunction with the tests since some compounds require metabolic activation to become mutagenic.

The Ames test is the most frequently used *in vitro* test for mutagenicity. Strains of *Salmonella typhimurium* are auxotrophic, carrying pre-existing mutations in genes responsible for synthesising histidine. The DNA reactivity of a test compound generates revertants, leading to colony formation in histidine-free medium. Different strains of *S. typhimurium* or in combination with *E. coli* are used to detect base-pair substitutions and frameshift mutations (Tejs 2008).

The chromosome aberration assay uses microscopy to look for gross structural changes in metaphase cells following treatment with a test compound. These changes can be symmetrical (chromosome breaks) or asymmetrical (chromatid breaks).

The micronucleus (MN) assay scores the formation of small membrane enclosed chromosome fragments or entire mis-segregated chromosomes known as micronuclei, formed in the cytoplasm of daughter cells. The *in vitro* MN assay can be seen as an equivalent of the chromosome aberration assay since it is able to detect both clastogens and aneugens. Cytochalasin-B can be used to inhibit cytoplasmic division, allowing identification of divided cells.

Applying the same principles as *in vitro* MN and chromosome aberration tests, the *in vivo* versions of these assays are carried out in rodent models, of which the bone marrow MN is the most common (Raffaella Corvi 2016).

### 1.1.2 Supplementary tests

Supplementary tests can be performed on a case-by-case basis to further support the data gathered from core tests.

The TK assay, carried out in either mouse lymphoma cells (L5178Y) or human-derived lymphoblastoid TK6 cells is used to detect mutagenic and clastogenic events at the heterozygous thymidine kinase locus. Reactive test chemicals causing a forward mutation would give rise to mutant cells ( $TK^{-/-}$ ), deficient in thymidine kinase (Lloyd and Kidd). In the presence of a selective agent such as trifluorothymide (TFT), cells heterozygous for TK are sensitive to the toxic effects of TFT whereas mutants are resistant, able to proliferate and form colonies (Lloyd and Kidd, 2012).

The alkaline comet assay involves electrophoresing cells under alkaline conditions through low-melting point agarose. As the nucleic acid migrates, damaged DNA can be seen progressing in front of the main nuclear body. The assay collectively measures DNA single and double strand breaks, in addition to abasic and alkali liable sites. It provides an alternate measure of clastogenicity that may involve repairable damage but does not provide information on mutagenicity or aneugenicity of a test compound and is therefore recommended as a follow up assay.

### 1.1.3 Cytotoxicity assays

The highest concentration of a test chemical used in any genotoxicity test is limited by its solubility or a threshold of cytotoxicity. While there is no overall consensus, a measure of general cytotoxicity is carried out complementary to the genotoxicity test used with the aim of limiting false positives associated with excessive toxicity (O'Donovan 2012). The recommended cytotoxicity test also differs. For instance, at least 20% reduction in cloning efficiency (CE) or relative total growth (RTG) is suggested for mutation assays, while *in vitro* micronuclei assays suggest up to 50% cytotoxicity measured by relative increase in cell counts (RICC) or proliferation index in the absence of cytochalasin B, otherwise a replicative-index (RI) or cytokinesis-block proliferation index (CBPI) is used in the presence of cytochalasin B (O'Donovan 2012).

## 1.2 Novel genome-wide approaches to study mechanisms

The advent of genome-wide analysis tools such as chromatin immunoprecipitation coupled to microarrays (ChIP-chip) or sequencing (ChIP-seq) has allowed for high-resolution genome-wide mapping of protein-DNA binding, significantly advancing our understanding of their functions. More importantly genomic datasets are rich with information, allowing associations be to drawn leading to a mechanistic insight (Yu et al, 2009), thereby providing a more informative tool in the interests of the current study.

ChIP-chip was originally developed over a decade ago in yeast (Ren B et al, 2000) and has since been used with a number of different organisms ranging from prokaryotes to eukaryotes (Figure 1.1). ChIP-chip links the application of chromatin immunoprecipitation of chemically cross-linked protein-DNA complexes with microarray analysis of the bound DNA. DNA microarrays were originally developed to measure gene expression changes and contain printed features of the genome as short oligonucleotides probes. Resolution therefore is typically limited by the probe converge and length of the hybridising DNA fragment. The resolution of ChIP-seq data on the other hand is not restricted by probe density and sequencing of smaller fragments offers greater coverage and resolution (Ho et al, 2011). With the rapidly reducing costs of sequencing, ChIP-seq has largely replaced ChIP-chip over the past years (Ho et al, 2011). However,

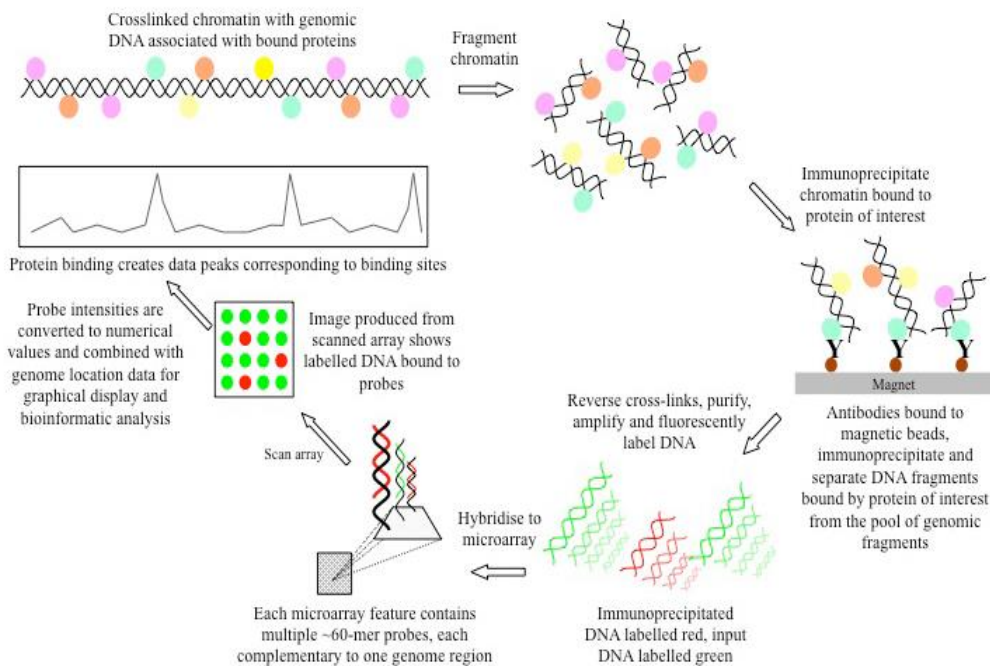
as it stands, ChIP-chip is a well-established technique with extensive bioinformatics tools available and the flexibility in array design offers an economic platform for human studies.

Analysis of ChIP-chip data has long been limited to binary interpretations, entailing the presence or absence of factors at various genomic sites under different experimental conditions, in part owing to the original application of microarray in gene expression analysis (Bennett 2015). Provided these inherent bioinformatic challenges in analysing large datasets, recent efforts from our lab have resulted in a novel method of normalising ChIP-chip data to allow relative comparisons to be made (Bennett 2015). This feature within the ‘Sandcastle’ software (Bennett 2015) allows additional information to be extracted from experiments and is particularly useful when measuring features such as histone acetylation, which tend to take on values within a biologically relevant continuum in response to environmental changes. Therefore, an established ChIP-chip workflow consisting of the necessary bioinformatic tools allows epigenetic changes to be fully examined in response to HDAC inhibitor treatment.

### **1.3 Epigenetics**

The rapidly expanding field of epigenetics emerged more than half a century ago when Conrad Hal Waddington (1905 – 1975) first coined the term ‘*epigenetics*’ in 1942 (Van Speybroeck 2002). Drawing from the long-standing debate surrounding *epigenesis* vs. *preformation* in neoclassical embryology (Van Speybroeck 2002), Waddington described epigenetics in the context of embryonic development, as a process (epigenesis) which utilises the interactions between existing information (preformation, also known as genetics) within the zygote to give rise to a wide array of new characteristics. As suggested in the use of the Greek prefix ‘*epi-*’, simply meaning ‘close to’ or ‘near’. This, in Waddington’s view of Neo-Darwinism portrays:

*“a breach between organism and nature as complete as the Cartesian dualism of mind and matter; an epigenetic consideration of evolution would go some way toward healing it”* (Counce 1958)



**Figure 1.1. Overview of the ChIP-chip workflow.** Protein or modification of interest is chemically cross-linked to DNA and fragmented to an average size of 500bp. Highly specific ChIP-grade antibody are coupled to superparamagnetic beads and used to enrich for target protein or modification of interest. The resulting bound fragments are de-crosslinked and purified to yield DNA (IP) fragments representative of binding location. The DNA is then amplified, differentially labelled and hybridised onto a DNA microarray with non-immunoprecipitated fragmented genomic DNA (IN) as background. The DNA microarray contains printed features of the genome as short oligonucleotide ‘probes’. Competitive binding between IP and IN DNA for probes results in different fluorescent intensities that are scanned and analysed by software. Following data normalisation, high IP/IN ratios likely represent probable binding sites which are determined by peak detection algorithms. Image adapted from (Powell 2014b).

and addresses what many geneticists' at the time conceived as a simplistic relationship between the genotype and the phenotype (Haig 2004). This concept is illustrated in Waddington's classical work on '*The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology*' in the form of an 'epigenetic landscape' (Figure 1.2), a visual metaphor depicting critical stages in a cell's development where 'choices' are made leading to different outcomes.

Since then, the word 'epigenetics' has been broadly assigned to explain seemingly unrelated biological phenomena linked to gene regulation, such as position-effect variegation in *Drosophila* (Wakimoto 1998) and paramutations in maize (Kermicle 1996). In the mid to late 20<sup>th</sup> century, the meaning of 'epigenetics' began to deviate within the field of biology. In the study of morphology, Herring (1993) described it as "*the entire series of interactions among cells and cell products which leads to morphogenesis and differentiation*" (Herring 1993) and in molecular biology terms it refers to "heritable changes in gene function that cannot be explained by changes in DNA sequence", which, to this date, has become the most commonly used definition of 'epigenetics' (Riggs A.D. 1996). Therefore it is not surprising that proper use of the term continues to be the subject of much debate (Bird 2007). This in Haig's view mainly stems from its dual origins (Haig 2004). Apart from the original meaning coined by Waddington, nearly two decades later Nanney adopted the term to describe "*messy*" inherited traits (Nanney 1958). Throughout the course of its history, epigenetics has evolved into a major field in biology, which is still used in a Waddingtonian sense. The study of epigenetics, as it currently stands, focuses on the mechanisms by which information from the ever-changing environment can shape and mold the genetic code to produce different phenotypes without changing its fundamental sequence.

### **1.3.1 Regulating the epigenome**

Following the rise of epigenetics as a prominent field in biology, the study of epigenetics focuses on potential mechanisms by which such information is propagated and stability maintained in cells. To this end, great interest surrounds trying to understand how the epigenome is regulated. Central to this is chromatin itself, the nucleoprotein structure that makes up our chromosomes, which remains a critical factor in regulating DNA accessibility, influencing many key cellular processes such as transcription,

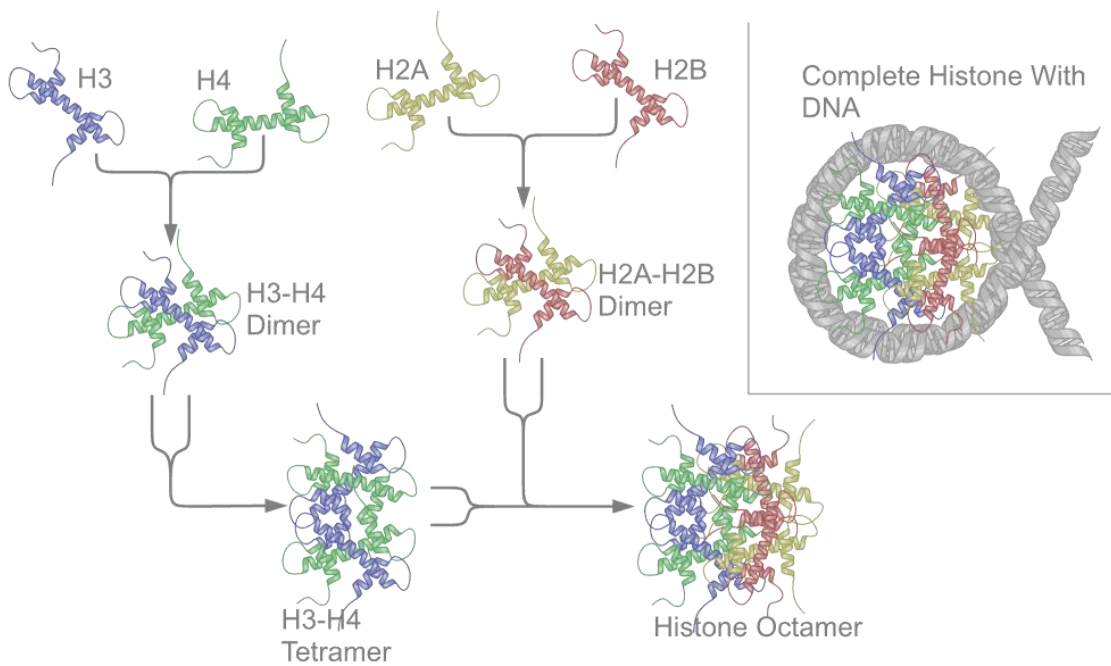
replication and DNA repair (Ehrenhofer-Murray 2004). Covalent modification of DNA bases, post-translational histone modifications and microRNA regulation have been identified as the major mechanisms by which epigenetic information is regulated in living cells (Figure 1.3). Of these, modifications to histone proteins are of particular interest in the current study, as these form a key part of the HDAC inhibitor response.

#### **1.4 Histone modifications**

In eukaryotes, a principal component of chromatin is histones. Roughly 147bp of DNA wrap around these proteins to form the nucleosome core particle (Karolin L 1997). Histones are comprised of two copies of H3, H4, H2A; H2B subunits, forming an octamer (Figure 1.4) (Kornberg R.D. 1999). The importance of histones were largely dismissed in the 1960s, at a time when much of the focus was on the central dogma of biology (Gottesfeld and Mukhopadhyay 2012). Vincent G. Allfrey is largely credited as the father of epigenetics for his pioneering work showing the presence of acetyl-lysine's on the N-terminal of histones H3 and H4 (Vidali G 1968). He further went on to hypothesise that the acetylation of histone tails would neutralise the positive charge, weakening their interaction with DNA (Vidali G 1968). Indeed, later studies have shown this to be the case. The removal of histone tails led to greater DNA accessibility either through nucleosome instability (Ausio 1989) or enhanced nucleosome dynamics (Polach 2000). Histone tail regions are protease sensitive, typically making up 25 - 30% of the total histone mass and protrude away from the chromatin (Wolffe AP 1999). The exposed surface allows for interactions with other proteins. For example, the Sir3 protein (silent information regulator) in yeast binds to deacetylated lysine 16 on the histone H4 tail to bring about transcriptional silencing (Hecht 1995). Our knowledge of the different types of histone tail modifications have greatly expanded with advances in technologies such as Mass-spectroscopy (Wisniewski 2007). Apart from the well-studied association between histone acetylation and transcription (Struhl 1998), to date up to 15 different types of modifications have been described in mammals leading to over 100 possible unique modifications based on their histone locations (Figure 1.5) (Huang 2014).







**Figure 1.4. Assembly of the histone octamer.** Core histone proteins (H3, H4, H2A; H2B) form dimers (H3-H4 and H2A-H2B), which assemble into tetramers. DNA wraps around complete nucleosomes consisting of two histone tetramers. Image from: [en.wikipedia.org](http://en.wikipedia.org).

### 1.4.1 Histone code hypothesis

Early concepts based on prominent histone modifications such as acetylation, methylation, phosphorylation and ubiquitination postulated that with such a diverse network of modifications decorating the histone tail, various mechanisms would exist to ‘read’ and ‘write’ this information at the epigenetic level (Mizzen C 1998). This would undoubtedly place chromatin activity under differential regulatory mechanisms linked to distinct biological processes. For example, histone acetylation and phosphorylation coordinate signalling through independent mechanisms to transiently and rapidly regulate immediate-early genes (IEG’s) demonstrating an additive effect (Thomson 2001). Additionally, phosphorylation of histone H3Ser10 is involved in both transcription and cell division, two processes with opposing chromatin conformations. The dual nature of this modification is highly dependent on neighbouring methylated (Rea 2000) and acetylated (Lo 2000) residues, suggesting that individual histone modifications can be regulated by surrounding modifications. Such evidence supported the notion of histone marks as a ‘language’ which could ultimately be read by the cell, or as Allis and colleagues proposed it – the ‘histone code’ (Strahl BD 2000).

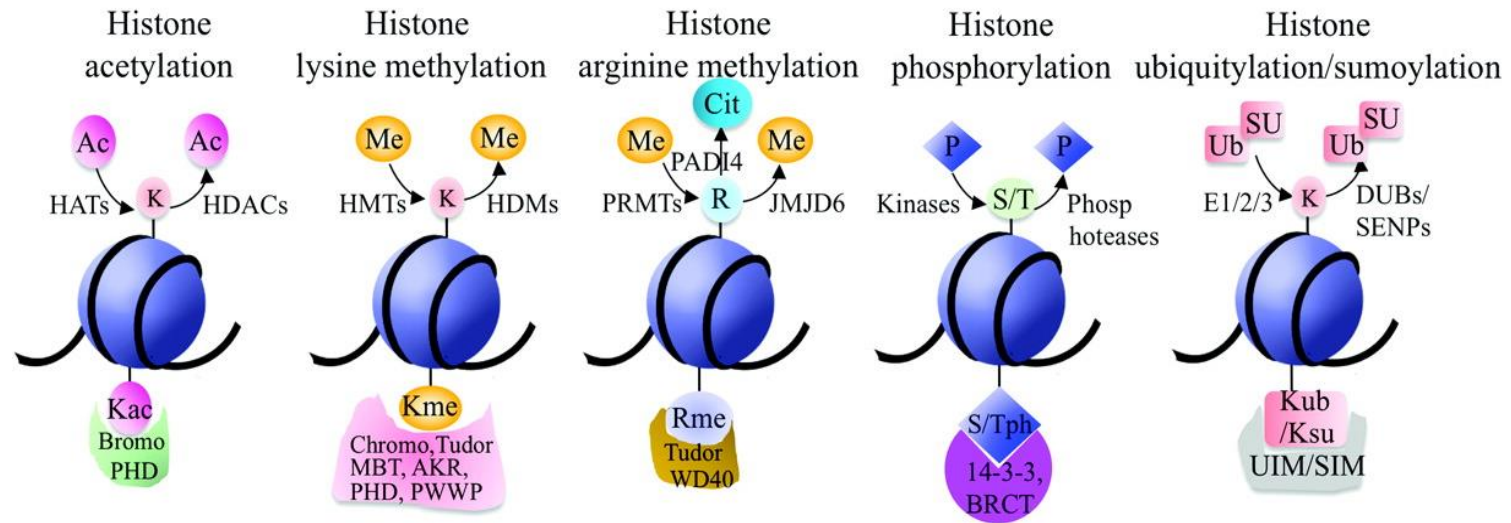
Although the code metaphor remains a popular topic and favours pro-complexity based on various chromatin-modifying proteins with multivalent binding capabilities, some refuted this, showing instead a lack of complexity. Given that acetylation of certain histone residues (ie. H3K9) are found at actively transcribed genes, single and combined lysine to alanine mutations on the tails of histone H3 and H4 exhibit unpredictable and insignificant effect on overall gene expression (Dion 2005). Furthermore, as chromatin modifiers were thought to recognise specific combinations of histone modifications, genome-wide mapping of BPTF (bromodomain PDH finger transcription factor) and Chd1 (chromatin helicase DNA binding protein 1) remodelling factors found additional bindings sites that could not be explained by histone marks alone (Ruthenburg 2011). Instead, it has been proposed that histone modifications may act as allosteric modulators under certain circumstances to regulate the activity of chromatin binding proteins. Noticeably, the removal of such marks has no effect on its binding (Morettini 2011).



### 1.4.2 Interpreting the code

Of the myriad histone post-translational modifications described to date, their presence/absence and biological effects are dictated by chromatin targeting enzymes with the ability to “read”, “write” and “erase” such marks. Regions of DNA must be made more accessible before the onset of transcription. To achieve this, histone-DNA interactions must first be dissolved and this is partly accomplished by proteins such as histone acetyl-transferases (HATs) that help to add (“write”) acetyl groups on histone tails, making it more negatively charged. On the other hand, histone deacetylase (HDAC) enzymes have the opposite effect of removing (“erasing”) these acetyl groups and suppressing the genetic information. While the effects of chromatin “writers” and “erasers” are refractory, “reader” proteins function solely to interpret existing histone modifications and relay biological signals. Acetylation is only one of the many possible histone PTM’s and with each distinct type of mark, unique classes of enzymes exists, responsible for its inscription, interpretation and removal (Figure 1.6). For instance, the enzyme SUV39H1 is responsible for methylating H3K9 (Rea 2000), which in turn is read by HP1 (heterochromatin protein 1) helping to stabilise heterochromatin (Bannister AJ 2001). Acetylation can also occur at the same residue and is read by BRD4 (Chiang 2009), family bromodomain modules, leading to the recruitment of the transcription factor p-TEFb (positive transcription elongation factor) (Brès V 2008).

Over the past decade, many families of conserved protein domains have been identified using techniques such as histone peptide screening (Garske AL 2010), protein microarrays (Nady N 2008) and more recently SILAC-MS (Stable Isotope Labelling by Amino acids in Cell culture) (Vermeulen M 2010), helping to distinguish between specific histone modifications. Once bound, the reader module serves as an intermediary scaffold, signalling the recruitment of downstream proteins involved in cellular processes such as transcription, replication and DNA damage/repair.



**Figure 1.6. Writers, erasers and readers of histone modifications.** Enzymes exist to deposit (write) or remove (erase) histone marks. These modifications are recognised (read) by various proteins through specific domains to signal cellular events. Modified residue(s) with corresponding reader domains shown below histones: PHD – plant homeodomain, MBT - malignant brain tumor, AKR – ankyrin repeats, PWWP - proline-tryptophan-tryptophan-proline, WD40 - tryptophan-aspartic acid dipeptide 40 repeats, BRCT - BRCA1 C-terminal, UIM – ubiquitin-interacting motif; SIM – SUMO-interacting motif. Readers and writer: HATs – histone acetyltransferases, HDACs – histone deacetylases, HMTs – histone methyltransferases, HDMs – histone demethylases, PRMTs – protein arginine methyltransferases, JMJD6 - jumonji domain containing protein 6, DUB –de-ubiquitinating enzymes, SENP - Sentrin/SUMO-specific protease. Histone post-translational modifications: Ac – acetylation, Me – methylation, P – phosphorylation, Ub – ubiquitylation, SU – sumoylation.

### **1.4.3 Histone modifications & cellular function**

Histone modifications are present in abundance throughout the genome, helping to establish regions of euchromatin (relaxed structure) and heterochromatin (compact structure), influencing key DNA-based processes such as transcription, replication and repair locally.

#### **1.4.3.1 Transcription**

Generally speaking, histone acetylation is associated with transcriptional activation, whereas sumoylation, deamination and proline isomerisation are associated with transcriptional repression. Methylation, phosphorylation and ubiquitination can either promote or inhibit transcription depending on the chromatin context. For instance, histone H3K4, K36; K79 methylation is linked with gene activation while histone H3K9, K27, and histone H4K20 methylation is linked with gene repression (Vakoc CR 2005). Apart from correlating various histone modifications with transcriptional state, recent studies demonstrate the presence of heterochromatin protein 1-gamma (HP1 $\gamma$ ) and H3K9 methylation in the coding region of actively transcribed genes (Vakoc CR 2005), highlighting further complexities.

Expression of genes by the transcription machinery requires access to the DNA template, an event that requires extensive chromatin remodelling and can be achieved by enzymes that post-transcriptionally modify histone tails and ATP-dependent remodelling complexes. The presence of histone marks such as H3K9ac correlates universally with active promoters and can be found together with H3K14ac to recruit transcription factors such as TFIID (Transcription Factor II D), a general transcription factor involved in the formation of the RNA polymerase II pre-initiation complex (Agalioti 2002). Phosphorylation of H3S10, S28 and histone H2BS32 were identified in the regulation of proliferation-associated genes. Interestingly, H3S28p has been shown to displace polycomb-repressive complex (PRC), a histone methyltransferase, leading to the demethylation and acetylation of adjacent H3K27 and subsequent transcription (Lau and Cheung 2011). H3S10p is also a well-documented event observed widely in eukaryotes connected to chromatin condensation during mitosis and meiosis (Wei 1999).

### 1.4.3.2 DNA Replication

Given the importance of histones and their ability to shape transcriptional readouts, it has become apparent that the reliable inheritance of chromatin states is equally important as the DNA itself. Various mechanisms have been put forward describing the transmission of histones during DNA replication. An earlier idea, initially proposed by Weintraub *et al.* implicated the segregation of parental histone tetramers into H3-H4 dimers, these parental dimers would then form tetramers with newly synthesised dimers, thereby acting as template for the transmission of histone post-translational modifications (Weintraub H. 1976). In contrast to this, an alternative mechanism could involve a loss of histone PTMs during replication and the re-establishment of these marks on *de novo* histone tetramers by histone modifying enzymes following S-phase (Jackson 1987).

The interactions of histone modifying enzymes with the DNA replication machinery as well as DNA binding proteins have been well described for the propagation of pericentric heterochromatin (Sarraf and Stancheva 2004). Chromatin assembly factor (CAF-1) is a histone chaperone recruited by proliferating cell nuclear antigen (PCNA), a processivity factor for the replicative DNA polymerase  $\delta$ . CAF-1 forms a complex with heterochromatin protein 1 (HP1) and the histone methyltransferase SETDB1, recruited by MBD1 (Sarraf and Stancheva 2004). H3K9 undergoes mono-methylation by SETDB1, followed by di-/tri-methylation by SUV39H (Fujita 2003). The dependence of SETDB1 on MBD1 demonstrates the requirement of methylated CpG-sequences in establishing regions of repressive chromatin following replication. On the other hand, Proteins of the polycomb group (PcG) accumulate and enhance H3K27me3 levels at repressed regions prior to DNA replication and are inherited by dilution, as opposed to *de novo* maturation (Lanzuolo 2011).

Although much debated, recent evidence supports both scenarios. Using stable isotope labelling-based quantitative mass spectrometry and affinity purification Xu *et al.* showed that the bulk of H3.1-H4 tetramers separated in a conserved fashion, although a substantial proportion of H3.3-H4 tetramers took part in splitting events (Xu 2010). Therefore mechanisms of epigenetic inheritance may be unique to regions of eu- and hetero-chromatin, considering that H3.3 is enriched in actively transcribed regions (Xu 2010). Distinct from the single semi-conservative mechanism of DNA replication, reproducing the epigenetic information appears more difficult and may involve several different mechanisms. Such a strategy would ideally allow for flexibility in how “local”



chromatin environments are re-established, which may have significant implications for how a single genome can be interpreted in many ways giving rise to different cell types.

Apart from being able to faithfully reproduce epigenetic information, histone modifications also influence the timing and firing of replication origins. In both eukaryotes and prokaryotes, H3K4 di- and tri-methylation marks are present at origins of replication (Rondinelli 2015). Demethylation of H3K4me3 by KDM5C, a family of Jumonji C (JmjC)-domain containing demethylase, is required for initiating the assembly of the pre-initiation complex on the chromatin, at early origins (Rondinelli 2015). Additionally, monomethylation of H4K20 in mammalian cells by KMT5A, a histone methyltransferase, aids the loading of the pre-replication complex (pre-RC) during origin licensing (Tardat 2010). In budding yeast, acetylation of histone H3 and H4 is required for origin activation during replication (Unnikrishnan 2010).

Particularly, acetylation of H3K56 increases the affinity of histone chaperones CAF-1 (chromatin assembly factor 1) and Rtt106 (regulator of Ty1 transposition) to H3, favouring nucleosome assembly (Sun et al. 2009). Disruptions to H3K56ac affect the completion of DNA replication, sensitising cells to DNA damaging agents that hinder replication fork progression (Wurtele 2012). More so, histone modifications and chromatin remodelling are crucial in facilitating the events leading to the proper repair of DNA damage generated from various sources.

### **1.4.3.3 DNA Repair**

Unlike DNA replication, scheduled to take place in the S phase, the repair of DNA is poised to occur throughout the cell cycle in response to damage. Repair of damaged DNA takes place on the chromatin and as such these events are regulated by histone modifications and remodelling of the chromatin. The initial “access-repair-restore” model proposed by Smerdon *et al.* views chromatin largely as an obstacle to repair (Smerdon 1991), first requiring an “opening” around the site of DNA damage to allow access for repair factors. Interestingly, accumulating evidence suggests the transient formation of heterochromatin is a prominent feature of DNA damage, supported by the recruitment of proteins such as HP1 (Luijsterburg 2009), PRC1, PRC2 (Campbell 2013) and HDAC1, HDAC2 (Miller 2010). On the other hand, failure of chromatin to condense properly leads to improper DNA damage response (DDR) (Campbell 2013). Recently it was shown that

in the context of double strand breaks, HP1 is recruited alongside KAP-1 and SUV39H1 to regions of euchromatin (Ayrapetov 2014). The resulting H3K9me3 lead to activation of ATM (Ayrapetov 2014), a key signalling protein in the DDR. The ensuing phosphorylation of KAP-1 by ATM represents a negative feedback loop, dictating a brief period of chromatin condensation with the release of the HP1/KAP-1/SUV39H1 complex (Ayrapetov 2014). The momentary appearance of repressed chromatin may function to inhibit local transcription and potentially avoid encounters with the repair process (Soria 2012). In light of this, early perturbations in chromatin events was described to effectively “prime” the DNA for subsequent repair (Soria 2012).

Numerous histone modifications have been described in the various stages and pathways of DNA repair. Of these, phosphorylation of Ser139 on histone H2AX ( $\gamma$ H2AX) is the most extensively characterised modification to date affiliated with double-strand break (DSB) repair (Rogakou EP 1998). The formation of  $\gamma$ H2AX in response to DSB constitutes a rapid response, helping to recruit the all-important MRE11-Rad50-Nbs1 (MRN) repair complex via MDC1 binding (mediator of DNA damage checkpoint protein 1). Activation of ATM by MRN further stimulates the phosphorylation of H2AX, amplifying the signal and leading to the formation of “ $\gamma$ H2AX foci”, visible using fluorescent microscopy (Kuo LJ 2008). Microarray studies have also revealed that  $\gamma$ H2AX distributes asymmetrically around the site of DSB, dependent upon transcription (Iacovoni 2010) and cohesion binding (Caron 2012). Interestingly, deficiencies in  $\gamma$ H2AX formation have only minor effects on DSB repair, suggesting that it functions mainly to promote signal retention at sites of DNA damage (Celeste 2003). Nevertheless, given its role in DDR,  $\gamma$ H2AX is widely employed as a biomarker of DNA damage and acts cooperatively with other histone modifications to drive the repair process.

Acetylation of histones plays a key role in the proper repair and recovery of DNA after damage. In yeast, following UV exposure, histone Htz1 enhances GCN5 (General Control Non-derepressible 5) acetyltransferase binding at the sites of UV-induced DNA damage leading to hyperacetylation at H3K9/14. This results in the recruitment of Rad14, a damage recognition protein in the nucleotide excision repair (NER) pathway (Yu 2013). Recent studies also demonstrated that H3K14ac serves as an anchor for the chromatin remodelling complex RSC (Remodels Structure of Chromatin), which functions to relax tightly positioned nucleosomes, assisting the repair of CPD's (Cyclo- Pyrimidine Dimers) formed as a result of UV-damage (Duan and Smerdon 2014).

H3K56ac is an abundant mark present on newly synthesised histones throughout the genome of budding yeast (Simoneau 2015). Its presence signals for chromatin assembly following the completion of DSB repair and DNA replication (Simoneau 2015). Absence of this mark leads to replication stress and the formation of spontaneous DNA damage (Simoneau 2015). Following DSB repair, Rtt109 (regulator of Ty1 transposition) acetylates H3K56 through Asf1 (anti-silencing function) stimulation (Chen 2008). Yeast *asf1* mutants exhibit normal DSB repair and DNA synthesis but lack detectable histones around sites of induced DNA damage (Chen 2008). Additionally, H3K56ac deficiency results in persistent activation of Rad53, a DNA damage checkpoint protein (Chen 2008). Therefore, Asf1 stimulated H3K56ac is not only crucial for restoring chromatin structure around the damaged DNA but more importantly signifies the recovery from checkpoint arrest and re-entry into the cell cycle (Chen 2008). H3K56ac was also identified as a factor along with Rad52 in the selection of sister chromatid recombination as the means of repair for replication-derived DSBs (Munoz-Galvan et al. 2013).

In mammals, BRCA1 (breast cancer early onset 1) and 53BP1 (p53 binding proteins 1) are two key proteins involved in DNA end-resection and act as determinants of homologous recombination (HR) or non-homologous end joining (NHEJ), respectively (Tang et al. 2013). H4K20me2 favours the binding of 53BP1, which is subsequently diminished by H4ac to enhance BRCA1 binding, representing a role of histone PTMs in driving DSB repair choice (Tang et al. 2013). Similarly, the HAT activity of the mammalian chromatin re-modeller Tip60 (60kDa Tat-interactive protein), regulated by H3K9me (Ayrappetov 2014), acetylates both histone H4K16 to promote HR and ATM to phosphorylate H2AX (Ayrappetov 2014). Tip60 can also form a complex with UBC13 (ubiquitin-conjugating enzyme) to acetylate H2AXK5, allowing ubiquitination of H2AXK119 to take place (Ikura et al. 2007). Acetylation and ubiquitination occurs independent of the initial phosphorylation response on histone H2AX and functions in its release as part of the chromatin re-organisation process after DNA repair (Ikura et al. 2007). Furthermore, monoubiquitination of H2A also functions in chromatin remodelling in response to UV-induced lesions (Bergink et al. 2006). In yeast, H2BK123 ubiquitination is shown to be necessary for cell cycle arrest following DNA damage (Giannattasio et al. 2005); highlighting the diversity of modifications and functions associated with histone H2A and its variants in the DNA damage response.

Post-DNA repair involves proper packaging of histones and DNA. Recently, CAF-1 was shown to mediate deposition of newly synthesised H3.1 variants following NER, while additionally acting as a novel regulator in preserving somatic cell identity (Bergink et al. 2006). Also, HIRA (histone chaperone histone regulator A) primes sites of UVC-induced lesions by installing H3.3, serving as a marker for transcriptional recovery after repair (Bergink et al. 2006). Intriguingly, such studies question whether histone turnover contributes to the formation of “damage imprints” as an adaptive mechanism for future DNA damage. Likewise, recovering damaged cells to the pre-damage state is key to maintaining cell identity as deregulation may lead to changes in cell fate such as cancer.

#### **1.4.4 Histone core modifications**

Much of the studied histone modifications to date focused on those solely present on histone tails. This is primarily due to technical limitations of the Edman degradation method, restricting the sequencing of amino acids to less than 30 in practice (Mersfelder and Parthun 2006). However, with the advent of mass-spectroscopy based methods a greater understanding of the histone modifications in the core globular domain was made possible (Mersfelder and Parthun 2006). The contributions of histone tail modifications to DNA-based processes have been characterised. These marks are easily accessible and have been shown to exert their effects mainly through “reader” proteins. Given that DNA wrapped around the nucleosome core particle is in close contact with the histone and that removal of the N-terminal tail has no significant impact on nucleosome stability (Ausio 1989), it is hypothesised that histone core modification may be key determinants of chromatin function.

Core modifications fall into three groups based on their location within the histone structure: (1) solute-accessible surface, (2) lateral surface and (3) histone-histone interface. Those that are solute accessible are predicted to function in a similar manner to tail modifications, given the largely exposed surface area. H3K79 methylation is evolutionarily conserved from yeast to mammals and represents one of the first studied modifications present on the nucleosome face (Li 2005). Methylation of H3K79 is not required for silencing; instead it prevents silencing of euchromatin by excluding the yeast SIR (silent information regulator) protein from these regions (van Leeuwen 2002).

Accessing the portion of DNA that is tightly packaged around histones presents a major challenge. Though, short periods (few milliseconds) of spontaneous unwrapping events have been detected, potentially by remodelling factors, allowing transient exposure of the DNA and modifications on the lateral histone surface (Li 2005). Such unwrapping events are introduced at DNA entry-exit points, regions between the histone tail and core, facilitated by the cooperative phosphorylation and acetylation of surrounding residues (Brehove 2015). The close proximity of lateral residues to DNA is expected to greatly impact histone-DNA interactions, directly affecting chromatin structure and function. Recent finding on H3K112 supports this idea, showing that acetylation of H3K112 alone was sufficient to trigger transcription of chromatin DNA *in vitro*, whereas histone tail modifications were previously incapable of doing so (Tropberger et al. 2013). Other lateral surface modifications discovered also influence nucleosome dynamics and DNA-based processes. Core modifications also extend to those present exclusively between histone-histone boundaries with the potential to regulate histone octamer dynamics.

Mass-spectroscopy screens have identified a number of such modifications but their exact function remains to be characterised (Zhang L 2003). Of these, only H4K91ac carries enough evidence to implicate a role in nucleosome deposition (Li et al. 2008). A mutation in this residue affects the interaction between H3-H4 tetramer and the H2A-H2B dimer, leading to destabilisation of the histone octamer and defects in chromatin assembly (Li et al. 2008). Emerging data on histone core modifications suggest that these are equally important and have diverse functions similar to those readily accessible on the histone N-terminal.

## **1.5 Histone deacetylases**

Histone deacetylase enzymes represent a class of epigenetic erasers with the specific function of removing acetylated histone marks. The growing body of evidence exposing HDACs as potential therapeutic targets for cancer and various other diseases has fuelled interests in these enzymes. Following the first human HDACs identified in the late 1990's (Taunton 1996), current efforts have uncovered 18 different isoforms in mammals. These enzymes are identified as part of an ancient superfamily, sharing a common lineage with prokaryotes. Classification based on sequence homology with their yeast counterparts provided two main groups consisting of four classes (Figure 1.7).

## 1.5.1 Classical HDACs



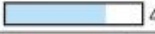











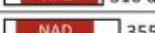

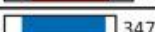

### 1.5.1.1 Class I

The first group of ‘classical’ HDACs (class I, II; IV) requires  $Zn^{2+}$  as a cofactor for its catalytic function and can be inhibited by  $Zn^{2+}$  chelating compounds. Class I consists of HDACs 1, 2, 3 and 8 homologous to yeast *RPD3* (reduced potassium deficiency) (Figure 1.7). Structurally they comprise the highly-conserved deacetylase domain flanked by relatively short carboxyl- and amino-terminals. While expressed ubiquitously in different cell types, they are found primarily in the nucleus where they target histone proteins. Because of this class I HDACs of all, display the most distinctive role in regulating gene expression. Knockout studies in mice demonstrated the importance of class I HDACs in cell survival and proliferation. For instance, HDAC1 knockout mice display developmental retardation and proliferation defects resulting in early embryonic lethality (Lagger 2002), while HDAC2 knockout mice survive development but suffer from a range of cardiac defects (Montgomery et al. 2007). Deletion of both these enzymes in mouse T-cells resulted in a loss of half the total HDAC activity, establishing HDAC1 and HDAC2 as the predominant isoforms (Dovey OM 2013). The relative expression of HDAC1 and HDAC2 was shown to vary significantly with cell type. HDAC1 is found abundantly expressed in rat glial cells whereas HDAC was more predominant in neurons (Broide et al. 2007). Conversely HDAC2 level in Jurket cells were negligible compared to HDAC (Hassig et al. 1998). HDAC3 was shown to be required for the survival of murine embryonic fibroblasts (MEFs) (Bhaskara et al. 2008). More recently a study of patients with loss-of-function HDAC8 mutations revealed phenotypes typical of Cornelia de Lange syndrome (CdLS), a rare genetic congenital malformation disorder (Kaiser et al. 2014). HDAC8 is responsible for deacetylating SMC3, a subunit of the cohesion complex, which helps to create the cohesiveness of chromatin-loaded cohesion (Deardorff et al. 2012). Loading of hyperacetylated SMC3 in the absence of HDAC8 activity results in reduced cohesion binding and transcriptional changes reminiscent of CdLS cell lines (Deardorff et al. 2012).

### 1.5.1.2 Class II

Class II HDACs possess larger amino-terminals and additional catalytic domains. These can be further subdivided into IIa (HDACs 4, 5, 7; 9) and IIb (HDACs 6; 10) sharing homology with yeast HDA1 (histone deacetylase 1) (Figure 1.7). The ability of class IIa HDACs to shuttle between the nucleus and cytoplasm is regulated through its interactions with the 14-3-3 protein. Phosphorylation of serine by kinases on the N-terminal domain facilitates binding of 14-3-3, resulting in a conformational change in the HDAC and its expulsion from the nucleus (Healy, Khan, and Davie 2011). Additionally, class IIa can interact with other factors such as the CtBP (C-terminal binding protein) corepressor and transcription factor MEF2 (myocyte enhancer factor-2) to bring about targeted repression (Zhang et al. 2001). In contrast to class I HDACs, class II HDACs display tissue-specific expression with HDAC4, HDAC5 and HDAC9 being predominantly found in the brain, heart and skeletal muscle (Moresi et al. 2010) and HDAC7 mostly in thymocytes (Dequiedt 2003). Although HDACs are known for their deacetylase functions, biochemical studies have shown that members of the class IIa HDACs lack the conserved tyrosine residue at the catalytic core, instead replaced by a histidine (Lahm et al. 2007). This change leads to more than 1000-fold decrease in deacetylase activity (Lahm et al. 2007). To compensate for this, class IIa HDACs act as adapters recruiting class I HDACs and HP1 to bring about its repressive functions (Zhang, McKinsey, and Olson 2002).

On the other hand, HDAC6 and HDAC10 are solely cytoplasmic deacetylases. HDAC6 contains two catalytic domains and its targets include  $\alpha$ -tubulin, cortactin, chaperones and the interferon receptor IFN $\alpha$ R (Tang et al. 2007). HDAC10 was demonstrated to interact with HDAC3 and possess a leucine-rich C-terminal region, suggesting possible diversity in function (Kobe B 2001). HDAC10 was also recently shown to exclusively deacetylate MSH2 (MutS homolog 2) at Lys73, part of the MutS $\alpha$  or MutS $\beta$  complex involved in the initial DNA mismatch repair (MMR) process (Radhakrishnan et al. 2015). Overexpression and knockdown of HDAC10 in HeLa cells resulted in increases and decreases in DNA MMR activity, respectively (Radhakrishnan et al. 2015).

Class	Isoform	Yeast counterpart	Size	Co-factor	Location	Expression	Catalytic domain
(I)	HDAC1	RPD3	58	Zn <sup>++</sup>	N	Ubiquitous	 482 aa
	HDAC2		59		N		 488 aa
	HDAC3		50		N, C		 428 aa
	HDAC8		44		N		 377 aa
(IIa)	HDAC4	HDA1	120	Zn <sup>++</sup>	N, C	Specific	 1084 aa
	HDAC5		130		N, C		 1122 aa
	HDAC7		110		N, C		 952 aa
	HDAC9		160		N, C		 1011 aa
(IIb)	HDAC6	HDA2	160	Zn <sup>++</sup>	N, C	Specific	 1215 aa
	HDAC10		70		N, C		 (Inactive) 669 aa
(III)	SIRT1	Sir2	120	NAD <sup>+</sup>	N	Variable	 NAD 747 aa
	SIRT2		45		C		 NAD 389 aa
	SIRT3		28		M		 NAD 399 aa
	SIRT4		35		M		 NAD 314 aa
	SIRT5		36		M		 NAD 310 aa
	SIRT6		39		N		 NAD 355 aa
	SIRT7		48		N		 NAD 400 aa
(IV)	HDAC11	RPD3/HDA1	39	Zn <sup>++</sup>	N	Ubiquitous	 347 aa

**Figure 1.7 Mammalian histone deacetylases.** Mammalian HDAC enzymes consist of several isoforms characterised to different classes based on their sequence homology. Similarities in homology are also found in the yeast counterparts. Key co-factors are necessary for enzyme activity with the majority dependent on Zn<sup>2+</sup> whereas sirtuins are dependent on NAD<sup>+</sup>. These exhibits either general or specific expression patterns dependent on the cell/tissue type. Sub-cellular localisation: N – nucleus, C – cytoplasm, M – mitochondria (Shirakawa K 2013).



### 1.5.1.3 Class IV

HDAC11 is the only member of class IV and shares homology with both classes I and II. HDAC11 negatively regulates the gene encoding for interleukin-10 (IL-10) in human antigen-presenting cells (APC), influencing both immune activation and tolerance responses (Villagra et al. 2009). Expression of HDAC11 transcripts were found mainly in the heart, brain, kidney, skeletal muscle and testis (Gao et al. 2002). Interactions with HDAC6 have also been reported in vivo (Gao et al. 2002).

### 1.5.2 Sirtuins

The class III HDACs or sirtuins are most similar to yeast Sir2 (silent information regulator 2). Zn<sup>2+</sup> dependent HDACs are also described as being sensitive to one of the early naturally isolated inhibitors, trichostatin A (TSA), where NAD<sup>+</sup> dependent sirtuins do not display this characteristic. Human sirtuins consists of SIRT1-7, each with its own unique cellular localisation and function. SIRT1 interacts with p53 to regulate cell stress and survival (Villagra et al. 2009). SIRT2 is cytoplasmic and much like HDAC6, targets  $\alpha$ -tubulin. SIRT3, SIRT4 and SIRT5 are present in the mitochondria with a role in energy metabolism (Lombard DB 2011). SIRT6 and SIRT7 bring about transcriptional silencing in regions of heterochromatin (Lombard DB 2011). The activities of these enzymes are not limited to deacetylation alone. In fact, only SIRT1-3 have strong deacetylase activity. SIRT4 and SIRT6 contain ADP-ribosyltransferase activity (Ahuja N 2007), while SIRT5 has a mixture of demalonylation, desuccinylation and deacetylation activities (Du J 2011).

Perhaps the most intriguing aspect of sirtuins is their association with aging. Early studies in yeast showed that up regulation of Sir2 slowed the effects of aging through caloric restriction (CR) (Kaeberlein M1 1999), the same was observed by manipulating orthologs of Sir2 in worms (Tissenbaum HA1 2001). Since caloric restriction can delay the onset of aging-related diseases, such as cancer, hypertension and Alzheimer's, this led to the identification of human sirtuins as new therapeutic targets. However, years of research surrounding this topic resulted in much controversy. While some reports have refuted these claims (Burnett C 2011), more recent findings support the role of sirtuins in CR and aging. Male transgenic mice overexpressing SIRT6 (al 2012), rather than SIRT1

(Daniel Herranz 2010) have increased lifespans. Furthermore, increasing SIRT1 specifically in the brain delays aging in both male and female mice (Sato A 2013).

### 1.5.3 HDAC complexes

Members of class I and II HDACs are commonly found as part of multiprotein complexes. Components of these complexes include DNA/chromatin binding proteins and transcription factors, helping to localise HDACs to chromatin. Aside from the core components, non-catalytic subunits with associated functions in DNA methylation, histone methylation and chromatin remodelling can be recruited to the complex (Silverstein and Ekwall 2005). This allows the cell to assemble the required enzyme conformation at each locus rather than having a specialised enzyme for each, making these complexes extremely flexible in function and modular (Silverstein and Ekwall 2005). Apart from early embryogenesis and brain development, HDAC1 and HDAC2 are functionally redundant, since only mouse models harboring deletions in both genes express a significant phenotype (Montgomery et al. 2007). This may be a consequence of both enzymes being present in the same corepressor complex which include: Sin3 (switch independent 3), NuRD (nucleosome remodelling HDAC) and CoREST (REST corepressor 1) (Figure 1.8) (Kaiser et al. 2014).

The Sin3 complex is highly conserved from yeast to mammals and contains the core subunits: Sin3A/B, HDAC1/2, SAP18, SAP30, RbAp46 and RbAp48 (Figure 1.8). The C-terminal of Sin3 forms the histone deacetylase interaction domain (HID) responsible for recruiting HDAC1/2 (Kaiser et al. 2014) and this interaction is stabilised by the Sin3 associated proteins (SAP) (Zhang et al. 1997). Retinoblastoma proteins (RbAp) are predicted to stabilise the Sin3/nucleosome interactions given their ability to bind histone H4. Like Sin3, the NuRD complex also consists of the same retinoblastoma proteins and additional structural/regulatory proteins such as MTA2 and p66 (Figure 1.8). NuRD is the only complex to comprise of both deacetylase and remodelling activities, provided by the SWI2/SNF2-related ATPase remodelled Mi-2 (Wang and Zhang 2001). Purification of NuRD found the methyl-CpG binding domain proteins MBD2 and MBD3 to occur separately, suggesting they formed distinct complexes with NuRD (Le Guezennec et al. 2006). Since only MBD2 is able to recognise methylated DNA, MBD2/NuRD complexes are mostly found at methylated inactive promoters whereas

MBD3/NuRD complexes are found at unmethylated active promoters (Gunther et al. 2013). The function of these corepressors depend very much on the chromatin context since their recruitment is determined by histone-binding motifs and/or transcription factors. This may also account for the lack of specificity seen in some HDAC1-HDAC2 complexes since each complex may contain a number of these chromatin localising proteins. Sin3 and NuRD are broad acting corepressor complexes while CoREST is exclusive to neurons (Lakowski, Roelens, and Jacob 2006). The CoREST complex, also known as neuron-restrictive silencing factor (NRSF), is recruited to RE1 DNA motifs by the repressor element-1 silencing transcription factor (REST) to regulate aspects of neuronal development (Lakowski, Roelens, and Jacob 2006). Suppression of genes at RE1 sites is achieved through methylation of H3K9 by the LSD1 recruited G9a methyltransferase. Conversely, in embryonic and neural stem cells LSD1 recruits H3K4 methyltransferase to functionally coactivate genes (Adamo A1 2011). In any case, the homo/hetero-dimerisation of HDAC1 and HDAC2 is a key requirement for deacetylase activity. HDAC3 on the other hand, exists within the nuclear hormone corepressor NCoR/SMRT complex along with members of class II HDACs (Figure 1.8) (EA, 2014). Transcriptional repression by NCoR/SMRT is achieved via cooperative interaction of HDAC3 with a combination of class II HDACs and class I HDAC containing complexes (EA, 2014). Mutational studies have shown NCoR (nuclear receptor corepressor) and SMRT (Silencing mediator for retinoid and thyroid receptor) to be crucial for neural and heart development, respectively. TBL1 (transducin  $\beta$ -like 1) is a core interacting protein that mediates corepressor/coactivator switching through ubiquitin activities (Perissi V 2004). Purification of the NCoR subunit also identified several lysine demethylases, namely KDM4A and KDM5C, which support transcriptional repression through histone modifications (Zhang D 2005). Unlike classical HDACs, evidence showing sirtuin-based complexes are lacking although few proteins interactions have been described. The binding of SIRT7 to RNA polymerase 1 is required for transcription of rDNA (Zhang D 2005). AROS (active regulator of SIRT1) stimulates whereas DBC1 (deleted in breast cancer 1) represses SIRT1 deacetylation of p53 (Kim 2008). The interaction of SIRT1 and p53 has important implications for regulating cell survival, metabolism, cell cycle and genome stability. Apart from the better-known chromatin corepressors, several novel complexes have been described. MiDAC (mitotic deacetylase complex) is a distinct cell cycle-associated complex, containing HDAC1 and HDAC2 (Bantscheff M 2011). NODE (Nanog and Oct4 associated deacetylase) is a unique HDAC1/2 complex described in

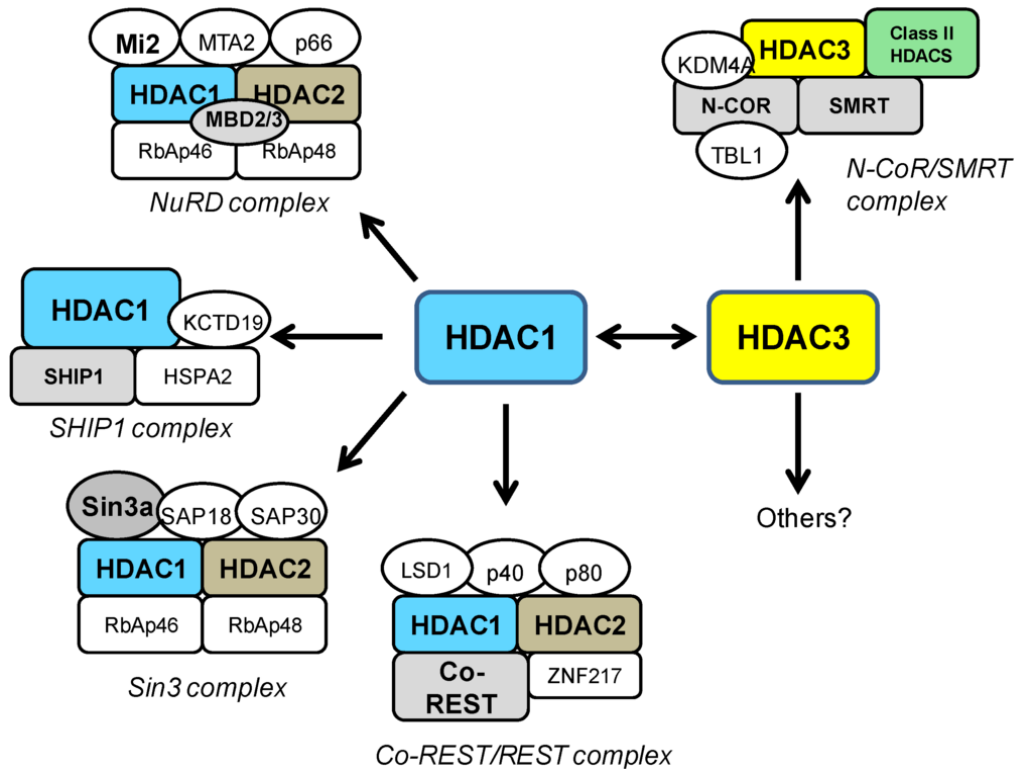
mouse embryonic stem cells (ESC) with a role in differentiation. Recent efforts have also identified SHIP1 (spermatogenic cell HDAC1 interacting protein 1) as a HDAC1 specific complex acting to remodel chromatin during spermatogenesis (Choi et al. 2008).

#### **1.5.4 Non-histone targeted effects of HDACs**

Since these enzymes were first studied in the context of histones they came to be known as histone deacetylases. However, phylogenetic studies in bacteria revealed HDAC enzymes to pre-date the evolution of histones (Gregoretta IV1 2004). Proteomic studies also identified large numbers of acetyl-lysine containing substrates, some of which are the targets of HDACs (Kim SC1 2006). Together these lines of evidence point towards non-histone proteins as the primary target of HDACs and for this reason histone deacetylase and lysine deacetylase (KDAC) are often used interchangeably in literature. cell motility. While some argued for a renaming of the enzyme to better reflect their targets, such notions were deterred to avoid confusion (Allis CD 2007). Nevertheless, acetylation remains one of the most common modifications in eukaryotes and is regulated by HDAC enzymes. The extracellular targets of HDACs include a diverse set of proteins with functions in apoptosis, angiogenesis, cell cycle control, metabolism, and inflammation.

#### **1.5.5 Histone deacetylases in cancer and other diseases**

Considering their profound role in regulating numerous biological processes such as cell development and growth, it is not surprising then that abnormal expression and/or mutations of HDAC enzymes have been identified as a major feature of many human disease, particularly cancer. Studies have reported the broad overexpression of HDAC1, HDAC2 and HDAC3 in a variety of cancers including: colorectal, renal, prostate, gastric, breast and Hodgkin's lymphoma (Adams H 2010; Choi JH 2001) . On the other hand, HDAC8 deregulation has only been described in T-cell lymphoma and childhood neuroblastoma (Oehme II 2009). Since class II HDACs have both proliferative and tumour suppressor functions their roles in cancer development are varied. For instance, overexpression of HDAC4 in glioblastoma cells suppresses the cyclin-dependent kinase



**Figure 1.8. HDAC enzymes function within multiprotein co-repressor complexes.** HDAC1 is a key component of NuRD, SHIP1, Sin3, CoREST/REST co-repressor complexes whereas HDAC2 is found in NuRD, Sin3 and CoREST/REST. HDAC3 has so far been discovered in the N-CoR/SMRT co-repressor complex which also contains class II HDACs. Mi2 – chromatin remodeller, MTA2 - Metastasis-associated gene family member 2, MBD2/3 - methyl CpG binding domain 2/3, RbAp46/48 – Retinoblastoma protein associated protein 46/48, KCTD19 - Potassium channel tetramerisation domain containing 19, SHIP1 – spermatogenic cell HDAC-interacting protein 1, HSPA2 – Heat shock 70kDa protein 2, SAP18/30 – Sin3 associated proteins 18/30, LSD1 – lysine-specific demethylase, ZNF217 – Kruppel-like zinc-finger 217, KDM4A – lysine-specific demethylase, SMRT – Silencing Mediator for Retinoid and Thyroid receptor, N-COR – Nuclear receptor CoRepressor, TBL1 – transducin  $\beta$ -like 1 (EA, 2014).

inhibitor p21(WAF1/Cip1) to enhance proliferation (Mottet D. 2008), whereas downregulation of HDAC4 is correlated with increased levels of VEGF and angiogenesis in chondrosarcoma cells (Xiaojuan Sun 2009). Similarly, HDAC7 is overexpressed in patients with pancreatic cancer but under expressed in mouse models of pro-B acute lymphoblastic leukemias and Burkitt lymphomas (Barneda-Zahonero B 2015). Androgen receptor (AR) signalling leads to the activation of genes responsible for driving the development of prostate cancer and is regarded as the predominant pathway (Gao L 2010). Common interventions involve androgen receptor-ligand disruption or deprivation but such methods eventually succumb to castrate-resistance (Gao L 2010). Though studies connecting deacetylation of HSP90 by HDAC6 and AR stability may provide alternative routes to treatment (Gao L 2010).

Comparably sirtuins appear also to exhibit a dual role in cancer. Overexpression of SIRT1 is reported in the vast majority of cancers such as acute myeloid leukemia (AML), hepatocellular and colorectal cancers, but is downregulated in some, such as gastric cancer (Yang Q 2013). SIRT1 is known to confer drug resistance in cancer cells by increasing the expression of multidrug resistance protein 1 (MDR1), enhancing the rate of detoxification, while also promoting anti-apoptotic factors, DNA repair and acquisition of mutations (Wang Z 2013). SIRT2 is important for cell cycle G<sub>2</sub>/M phase transition and mutations in its active site resulting in a loss of activity is observed in melanomas (Lennerz V 2005). In fibrosarcoma HT1080 cells, SIRT3 and SIRT4 help prevent the onset of apoptosis and promote cell survival by maintaining mitochondrial levels of NAD<sup>+</sup> (Yang H1 2007). More recently, SIRT5 overexpression in non-small cell lung cancer (NSCLC) was identified as a contributing factor in acquiring drug resistance while SIRT6, previously considered a tumour suppressor, upregulates the proinflammatory factor COX-2 (cyclooxygenase-2) in skin cells leading to cell proliferation (Ming M1 2014).

Aberrant changes in HDACs undoubtedly reshape the global acetylation landscape with direct consequences on gene expression. The reported loss of H4K16ac and enhanced H4K20me<sub>3</sub> is a common characteristic of many tumour phenotypes (Fraga MF 2005). Dysregulation of HDACs makes them suitable prognostic markers of cancer progression. Simultaneous overexpression of HDAC5 and HDAC9 in high-risk medulloblastoma is associated with low survival and may serve as a useful indicator for risk stratification. While tissue microarray studies showed that overexpression of HDACs

1-3 correlated with lower survival rates and lymph-node metastasis in patients with gastric cancer (Kanai 2008).

Aside from cancer, HDACs also present a potential therapeutic target for long-term non-life threatening illnesses. Fibrosis is the thickening and scarring of tissue in response to injury and pathological feature of many chronic inflammatory diseases. Particularly, increased HDAC1 and HDAC2 levels observed in fibrosis-related kidney diseases correspond to the production of pro-fibrotic factors, such as CSF-1 (colony stimulating factor 1) (Hashioka 2012). HDAC1/2 knockdown or inhibition significantly reduce CSF-1 levels, providing a means of tackling the inflammatory response during tubule-interstitial injury (Hashioka 2012). Numerous studies in rodent models demonstrate transient histone acetylation as a key event in learning and memory formation, supported by the observation that disruptions to HATs in the mice hippocampus leads to memory impairment, while HDAC inhibition in wild-type mice has the opposite effect of helping to promote synaptic plasticity and learning (Vecsey CG 2007). Such findings form the basis for supporting the development of HDAC targeted therapies for Alzheimer's disease (Hashioka 2012).

The large family of HDAC erasers cover a diverse range of both histone and non-histone targets in the cell and regulate genes and extracellular proteins involved in many key cellular processes. As such, disorders resulting from abnormal changes in these processes often involve or can be remedied by targeting the associated HDAC enzyme, making them prime targets for pharmaceutical intervention.

## **1.6 Chemical inhibitors of HDAC enzymes**

Early studies of the HDAC1 homologue HDLP (histone deacetylase-like protein) in hyperthermophilic bacterium *A. aeolicus* with inhibitors such as trichostatin A (TSA) and suberoylanilide hydroxamic acid (SAHA) (Finnin MS 1999) and the crystal structure of the HDAC8/hydroxamate complex helped to identify the structural interactions between classical HDACs and their inhibitors (Vannini A 2004). The active domains of class I, II and IV HDACs consists of a zinc-binding site, tubular pocket and Asp-His dyads. A collection of both naturally occurring and synthetic inhibitors specific for zinc-dependent HDACs exert their inhibitory effects by disrupting enzyme-metal ion interactions. Consequently, this feature sets apart the class III NAD<sup>+</sup>-dependent sirtuins,

which do not respond to the  $Zn^{2+}$  class of HDAC inhibitors. The pharmacophore of HDAC inhibitors consists of a variable Zinc-binding domain (ZBD) connected via a 'linker' region to either a hydrophobic or aromatic 'surface recognition/cap domain' (Vannini A 2004). Typical classification of inhibitors is based on the structure of the metal-binding group, belonging to one of the major hydroxamic acid, cyclic peptide, aliphatic acid or benzamide groups (Figure 1.9).

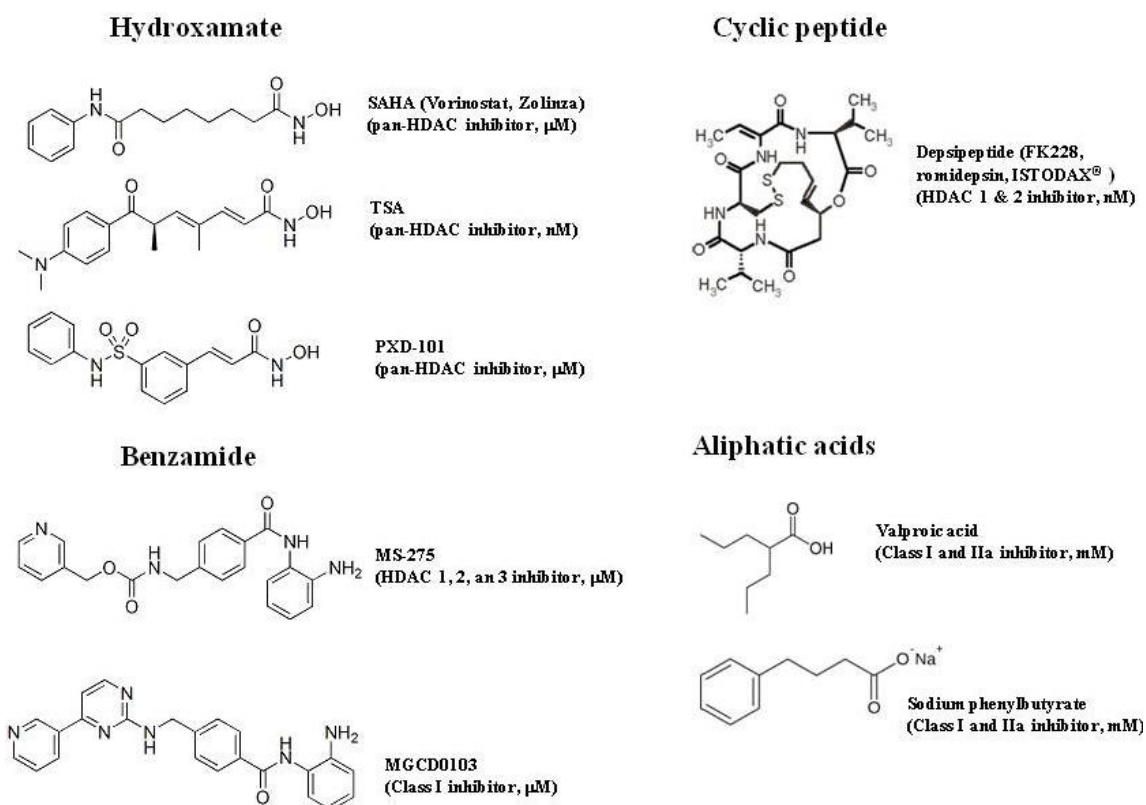
### 1.6.1 Hydroxamic acids

The trichostatin family of compounds are natural hydroxymates originally isolated as antifungals from species of *Streptomyces*. Trichostatin A exhibits inhibitory activities at low nanomolar concentrations *in vitro* and remains one of the most potent pan-HDAC inhibitors available (Figure 1.9). Suberanilohydroxamic acid (SAHA) or vorinostat is structurally similar to TSA and was the first HDAC inhibitor to be approved by the U.S. Food and Drug Administration (FDA) for the treatment of cutaneous T-cell lymphoma (CTCL) (Mann BS1 2007) (Figure 1.9). Vorinostat was marketed under the trade name Zolima<sup>®</sup> by Merck & Co. Inc. in 2006 and is recommended as a follow-up treatment for patients with progressive or persistent forms of CTCL (Mann BS1 2007). More recently, the hydroxamic derivative Belinostat (BELEODAQ<sup>®</sup>, Spectrum Pharmaceuticals Inc.) was approved for the treatment of patients with peripheral T-cell lymphoma (PTCL), a rare and aggressive form of non-Hodgkin's lymphoma (NHL).

### 1.6.2 Cyclic peptides

Depsipeptide (romidepsin, ISODAX<sup>®</sup> Celgene) is the most notable member of the structurally complex class of cyclic peptide inhibitors, also approved for the treatment of CTCL (Figure 1.9). Depsipeptide is administered as a pro-drug and activated intracellularly by the process of reduction, targeting HDAC1 and HDAC2 more favourably than class II enzymes (Furumai R 2002). The complexity of the surface recognition domains in this group of molecules allow for varying iterations of inhibitors which recognise different amino acids on the surface of HDAC isoforms, thus presenting a potential for subclass specificity.





**Figure 1.9. Structural features of the main HDAC inhibitor classes.** Four major classes of HDAC inhibitors with varying potencies, consisting of: 1) hydroxamates, 2) cyclic peptides, 3) benzamides and 4) aliphatic acids. SAHA - Suberoylanilide hydroxamic acid, TSA – trichostatin A. Image from (Bae 2011).

### 1.6.3 Benzamides

Benzamide containing inhibitors show both *in vitro* and *in vivo* anti-cancer activities. Entinostat (MS-275/SNDX-275, Syndax Pharmaceutical Inc.) is a synthetic benzamide derivative specific for class I HDACs (Figure 1.9). Having undergone trials for a range of different cancers it recently achieved the FDA's 'Breakthrough Therapy' status for the treatment of advanced breast cancer in combination with the hormonal therapy drug exemestane (Aromasin<sup>®</sup> Pfizer) (Yardley DA 2013). Mocetinostat (MGCD0103, Mirati Therapeutics) is another compound of this class, targeting class I and IV HDACs (Figure 1.9), demonstrating safe use and anti-leukemic activities in patients with myelodysplastic syndrome (MDS) (Garcia-Manero G 2008).

### 1.6.4 Aliphatic acids

The aliphatic acid group of inhibitors generally targets HDACs with millimolar affinities with the likes of valproic acid (VPA) and sodium phenylbutyrate (Figure 1.9). Generally considered a weak inhibitor of HDACs, valproic acid alone shows little efficacy against cancers but is widely used in the maintenance and treatment of bipolar disorders (Cipriani A 2013).

### 1.6.5 HDAC inhibitors in the clinic

HDAC inhibitors have received the most attention in areas of oncology. Recent reports estimate that at least 20 different HDAC inhibitors were involved in clinical trials for varying types of cancers in the past couple of years. Arguably, these inhibitors received the most clinical success in the treatment of leukaemias and lymphomas while their effectiveness in solid malignancies is rather limited (Slingerland M1 2014). The mechanisms behind these observations remain largely unknown but considering many of the first generation pan-HDAC inhibitors were developed based on their ability to induce tumour differentiation rather than utilising a mechanism-based approach, the resulting low enzyme specificity is thought to underpin the lack of efficacy in solid tumours (Bae 2011). The identification of deregulated HDAC isoforms in cancer subtypes promoted the development of inhibitors aimed at achieving greater efficacy through improved target

specificity. Much of the ensuing second-generation HDAC inhibitors have chemical structures based off agents with demonstrated clinical efficiency such as vorinostat (hydroxymates) and entinostat (benzamides). Of these CHR-3996 is a class I selective inhibitor undergoing trials for solid refractory tumours having successfully completed phase I trials showing encouraging pharmacokinetic, pharmacodynamics and toxicity profiles (Banerji U 2012). Similarly, the phase I trial of chidamide demonstrated preliminary anti-tumour activity in patients with advanced solid tumours or lymphomas, warranting further investigation (Dong M 2012).

While continued efforts are underway to improve the efficacy of HDAC inhibitors in solid tumours, this current shortfall strictly limits the use of these inhibitors as monotherapies in cancer. Instead HDAC inhibitors have shown great potential in improving response rates when synergised with a variety of existing and new anti-cancer agents including platinum-based drugs, DNA cross-linking agents and radiotherapy. The combination of vorinostat with direct DNA interacting agents such as cytarabine and idarubicin resulted in a high overall recovery rate of 85% in patients with acute myelogenous leukemia (AML) or myelodysplastic syndrome (MDS) (Garcia-Manero G 2012). Given the ability of HDAC inhibitors to suppress oestrogen and androgen receptor expression, combined treatments using vorinostat and tamoxifen demonstrated a partial or complete recovery in 19% of patients with hormone-therapy resistant breast cancer (Munster PN 2011).

Yet, long before the molecular targets of these drugs were identified, leading to a turn of attention to oncology, early HDAC inhibitors were used for the treatment of psychiatric and neurodegenerative disorders. Valproic acid (VPA) was licensed as an anticonvulsant for the treatment of seizures in the late 1970's and is still used today, widely for treating migraines, depression and schizophrenia (TR 2003). Knowledge of the effects of HDAC inhibition on gene expression has allowed for a rational approach towards disease targeting. The demonstrated ability of HDAC inhibitors to re-activate the *SMN1* (survival motor neuron 1) gene proved a promising therapeutic against neuromuscular degeneration; as such VPA has shown promising results in trials for spinal muscular atrophy (SMA) (Helmrich, Ballarino, and Tora 2011).

There is mounting evidence to suggest that HDAC inhibitors can provide beneficial anti-inflammatory effects, irrespective of apoptosis or cytotoxicity. Sodium butyrate was recognised as a potent suppressor of cytokine activity in the treatment of acute gout (Cleophas MC 2016). Inflammation is also attributed to the pathogenesis of

hypertensive nephropathy, a condition resulting in kidney damage from sustained high-blood pressure. HDAC6 was identified as a potential therapeutic target in hypertensive mice models using the specific inhibitor tubastatin A, which effectively prevented fibrosis and inflammation (Choi SY 2015). Lipopolysaccharides (LPS) from many species of bacteria are powerful triggers of acute inflammatory reactions as a result of infection or injury. Recent studies demonstrated the ability to alleviate LPS-induced inflammatory responses by specifically targeting class II HDACs. Li *et al.* showed that the HDAC8 specific inhibitor ITF3056, a derivative of the orally active hydroxamate givinostat (ITF2357), significantly reduced LPS-induced cytokines in human peripheral blood mononuclear cells (PBMCs) as well as in mice (Li S 2015). While givinostat has similar anti-inflammatory properties it is highly cytotoxic (Li S 2015). ITF3056 on the other hand, displays no cellular toxicities at concentrations showing reduced inflammation (Li S 2015), suggesting that target specific interactions with HDACs may present safer therapies.

#### **1.6.6 HDAC inhibitor-related side effects**

The most commonly reported clinical toxicities of HDAC inhibitors include nausea, vomiting and fatigue (Subramanian S 2010). Antiemetics are used to combat nausea and vomiting, while fatigue typically subsides following withdrawal of the drug (Subramanian S 2010). Acute decreases in white blood cell and thrombocyte counts have also been reported in addition to cardiac arrhythmias observed from ECGs (electrocardiograms) (Subramanian S 2010). Some of the metabolic side effects range from severe liver toxicities, electrolyte imbalances to minor renal problems (Subramanian S 2010). In extreme cases, patient deaths have been reported from experimental trials of HDAC inhibitors (Subramanian S 2010). A phase II trial of mocetinostat in patients with relapsed or refractory classical Hodgkin's lymphoma resulted in the deaths of four patients, two of which were related to the effects of the drug (Younes A 2011). These side-effects reflect on the toxicity of HDAC inhibitors at the cellular level, with emphasis on developing safer and more effective therapies based on a greater mechanistic understanding.

## 1.7 Genetic approaches to study HDAC function

Aside from clinical interests, HDAC inhibitors have also been extensively used to study HDAC function. The genetics of single celled eukaryotes such as yeast can be easily manipulated compared to that of higher eukaryotes and therefore provides a useful tool for research. The role of HDAC enzymes on histone acetylation has been well studied in yeast using HDAC mutants. Deletions in *RPD3* and/or *HDA1* (homologue of human class I and II HDACs, respectively) have been shown to increase global levels of histone H3 and H4 acetylation in budding yeast (Rundlett SE 1996). This effect was also reproduced in fission yeast treated with the HDAC inhibitor TSA (Ekwall et al. 1997).

Gene specific analysis using chromatin immunoprecipitation showed a high degree of redundancy amongst yeast HDACs in regulating heterochromatin formation at distinct subset of genes. Such arrangements are thought to organise genes so that a collective response can be produced in response to an external stimulus. An example of this can be found in Hda1 targeted HAST (Hda1-affected subtelomeric) domain containing genes, which are activated to regulate growth during stress (Robyr et al. 2002). Deletion in *RPD3* but not *HDA1* increases histone H4K5/K12 acetylation and expression of *IME2*, *SPO13* and *INO1* genes (Rundlett SE 1998). By contrast, Tup1 is a general repressor in yeast and interacts with Hda1 to silence specific genes (Wu J 2001). *HDA1* deletion results in H3/H2B-specific hyperacetylation and expression of *TUP1* controlled genes, including *GAL1* (Wu J 2001).

Genome-wide analysis of fragile sites in yeast using  $\gamma$ H2A, an epigenetic marker of DNA damage, suggest that a large proportion of the  $\gamma$ H2A signal observed corresponds to sites which present barriers to replication forks (Szilard RK 2010). It is known that obstacles to DNA replication can lead to fork collapse and DNA damage. Interestingly, the HDAC mutants *rpd3 $\Delta$*  and *hst1 $\Delta$*  show abolished  $\gamma$ H2A at regulated genes compared to the wild type, suggesting that condensed chromatin imposed by HDAC enzymes may present obstacles to replication (Szilard RK 2010). Similar approaches were later used to examine fragile sites in human cells (Jungmin S 2012). Much of the knowledge surrounding human biology is derived from earlier pioneering work in model organisms such as yeast.

## 1.8 Evidence for HDAC inhibitor induced DNA damage

HDAC inhibitors are highly cytotoxic agents, exerting their effects on a variety of cellular targets leading to cell cycle arrest, apoptosis and differentiation. The observed cytotoxicity was suggested to be in part due to the effects of genotoxicity. The results of several independent studies supported this idea by demonstrating that various HDAC inhibitors tested positive in standard genotoxicity assays. Apicidin and its derivatives (SD-0203 AND SD-2007) induced chromosome aberrations and micronuclei in Chinese hamster ovary (CHO) cells but were not mutagenic in the Ames test (Yoo EJ 2005). A comparison of the effects of sodium butyrate between CHO and mouse (L929) cells showed differences in sensitivity but aneuploidy was observed in both. Additionally sodium butyrate treated cells displayed less intense staining for heterchromatin relative to untreated cells, suggesting potential changes in chromatin structure (Gomez V.P. and Vig B.K. 2002). Trichostatin A (TSA) tested negative in the Ames and positive in the chromosome aberration, micronuclei,  $\gamma$ H2AX and alkaline comet assay in TK6 cells (Olaharski AJ 2006). The genotoxicity of the aforementioned HDAC inhibitors were further substantiated using the non-regulatory GreenScreen<sup>®</sup> HC reporter assay, showing significant increases in fluorescence of the DNA-damage inducible GFP (green fluorescent protein)-tagged *GADD45a* gene (Johnson and Walmsley 2013). Vorinostat showed weak mutagenic effects in Ames, in addition to testing positive for chromosome aberrations in CHO cells and MN in *in vivo* mouse models (Johnson and Walmsley 2013). Drug safety data published by the FDA on the hydroxamates Vorinostat, Belinostat and Panobinostat also indicate mutagenicity (LH 2006). Shen and Kozikowski recently reviewed evidence surrounding the mutagenicity of FDA-approved HDAC inhibitors, proposing that hydroxamates can be converted into unstable isocyanates *in vivo* and react with nucleophilic groups of DNA (Shen S 2016). This effect however has not been reported for TSA or other non-hydroxamate based HDAC inhibitors and therefore the mutagenicity may be limited to a specific structural class. The existing body of evidence present in the literature to date strongly substantiates HDAC inhibitors as genotoxic agents, although the mechanisms behind this are unknown.

### **1.8.1 Potential mechanisms of HDACi-induced DNA damage**

Of the limited evidence surrounding how HDAC inhibitor may generate genetic damage, oxidative DNA damage has been reported in response to Vorinostat in acute myeloid cells (AML) cells, implicated as an important mechanism of cancer cell lethality (Petruccioli 2011). Another HDAC inhibitor LAQ-824, also stimulated the production of reactive oxygen species (ROS) in various human leukaemia cell lines (Rosato RR 2008). In both reports, the addition of the ROS scavenger N-acetyl cysteine (NAC) led to a reduction in  $\gamma$ H2AX levels stimulated by HDAC inhibition but not down to baseline levels equivalent to untreated cells (Petruccioli 2011; Rosato RR 2008). This suggests that ROS only accounts for a proportion of the DNA damage induced by HDAC inhibitors. Mechanistically, HDAC inhibitors downregulate thioredoxin (Bose, Dai, and Grant 2014), a ROS scavenger and depolarise mitochondrial membranes (Rosato RR 2008), which can lead to increased ROS.

Structural changes in chromatin following HDAC inhibitor treatment were implicated in the formation of DNA damage, since early  $\gamma$ H2AX induction was observed in parallel with histone hyperacetylation (Gaymes et al. 2006). Conti and colleagues used single DNA molecule analysis to demonstrate that pharmacological concentration of Vorinostat reduced replication velocity and promoted the activation of dormant origins in human cancer cells (Conti et al. 2010). Additionally, pre-treatment with the transcription inhibitor flavopiridol or the replication inhibitor aphidicolin reduced the  $\gamma$ H2AX signal elicited by Vorinostat, indicating that the DNA damage observed is transcription and replication-dependent (Conti et al. 2010).

## **1.9 Transcription and replication conflicts as a source of genome instability**

### **1.9.1 Eukaryotic transcription**

Syntheses of all protein-coding and non-coding RNAs are carried out by distinct polymerases in eukaryotes. RNA polymerase I specifically transcribes ribosomal RNAs (rRNA), the largest being 28S, 18S and 5.8S. RNA polymerase II on the other hand transcribes protein-coding mRNAs and non-coding ncRNAs, which make up most of the transcriptome. Lastly RNA polymerase III is responsible for transcribing transfer RNA (tRNA) and smaller species of rRNA. All RNA polymerases are confined within the

nucleus and localise to different promoters as multi-complex proteins, consisting up to a dozen unique subunits.

RNA polymerase II (RNA pol II) has received the most attention due to its dominant role in gene expression; in addition its largest subunit contains an extended C-terminal domain (CTD) consisting of YSPTSPS tandem repeats (Geyer 2013). Serine (Ser2, Ser5, Ser7), tyrosine (Tyr1) and threonine (Thr4) residues located within the repeat sequence are liable to phosphorylation, making it a post-transcriptional modification-rich region (Geyer 2013). The CTD does not contribute towards the activity of the polymerase but instead has a role in regulating RNA processing, chromatin structure and docking of various binding proteins (Geyer 2013). Eukaryotic transcription is a highly complex process, orchestrated through a series of stages entailing: pre-initiation, initiation, promoter pausing, elongation and termination (Figure 1.10).

Interaction with transcription factors represents an essential step in initiating transcription. General transcription factors universally bind to TATA-promoter sequences and help in assembling the RNA pol II pre-initiation complex (PIC), with the core components consisting of the TATA-binding protein (TBP) and various transcription factor II proteins (TFIIB, TFIIE, TFIIIF, TFIIH) (Lodish H 2000). Specific transcription factors on the other hand, can bind to regulatory sites to either promote or inhibit transcription. Upon initiation, the ATP-dependent helicase activity of TFIIH unwinds the double stranded DNA (dsDNA) (Guzmán E 1999), and with the aid of TFIIB stabilises the separate DNA strands within the 'transcription bubble'. In addition, the CDK7 (cyclin-dependent kinase) activity of TFIIH phosphorylates Ser5 and Ser7 on the CTD, with phospho-Ser5 being a pre-requisite for recruiting 5'-capping factors (McCracken S 1997). Not all initiating events surmount to active transcription. Short unstable transcripts are generated in cycles of 'abortive initiation' where RNA Pol II terminates prematurely before leaving the promoter and re-engages until its escape. During this process, as the failed transcript is evicted, the RNA Pol II complex reseals and pushes the DNA segment back downstream, whereas in progressive transcription the DNA is pushed upstream. Therefore abortive initiation involves a 'scrunching' mechanism where RNA Pol II remains stationary (Kapanidis AN 2006). Lengthening of the RNA-bound DNA hybrid makes contact with protein loops near the RNA-exit region, the stabilising effect is thought to induce structural changes in the transcription complex to facilitate 'promoter escape' (Saunders A 2006). The complex in this state is also referred to as the early elongation complex (EEC) and exhibits unique pausing at the proximal-promoter. RNA



Pol II pausing is believed to prime genes for a rapid and synchronised response to stimuli (Dao P 2016). A key mechanism coordinating this event involves inhibition of RNA Pol II progression by NELF (negative elongation factor) and DSIF (DRB sensitivity inducing factor) (Saunders A 2006). To alleviate the blockade, phosphorylation of both NELF and DSIF by the cyclin-dependent kinase (CDK9) component of P-TEFb (positive transcription elongation factor) is required (Saunders A 2006).

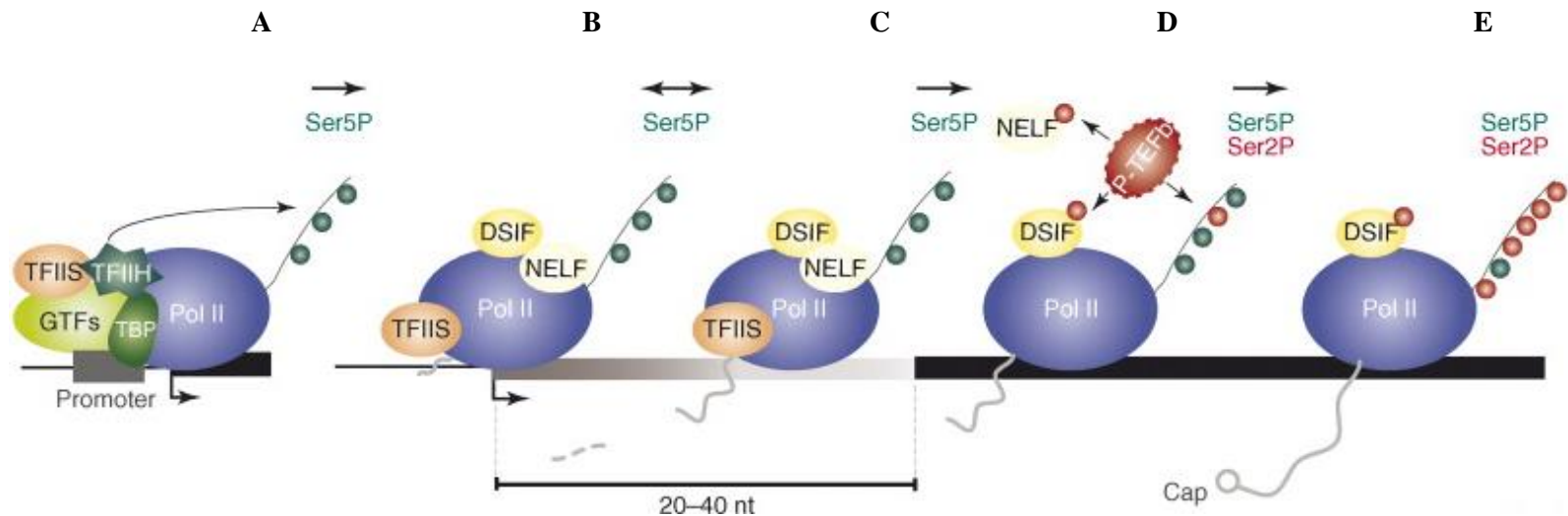
The advancement of RNA Pol II into the elongation phase is signified by the loss of phospho-Ser5 and gain of phosphor-Ser2 (Komarnitsky P 2000). This signal is maintained until the end of transcription and functions to recruit splicing, elongation as well as termination factors (Komarnitsky P 2000). Brief RNA:DNA hybrid interactions help to tether the transcribing RNA Pol II complex to the DNA template, sustaining processivity and preventing early termination (Kireeva ML 2000). Fidelity is maintained during active transcription by ensuring that the correct ribonucleotide triphosphate (rNTP) substrate is incorporated into the extending mRNA chain (Sydow JF 2009). Removal of a mis-incorporated substrate is aided by TFIIS stimulated intrinsic nuclease activity, and occurs only secondary to error detection mechanisms (Sydow JF 2009).

The final step in the transcription process is the termination of RNA Pol II and proceeds through a poly(A)- or Nrd1–Nab3–Sen1-dependent pathway based on available termination factors and 3'-end signals (Kuehner JN 2011). A poly(A)-dependent mechanism is common for all protein-coding genes, during which ser2-phosphorylated CTD attracts the termination complexes CPSF (cleavage and adenylation specificity factor) and CstF (cleavage stimulator factor) to facilitate mRNA cleavage (Kuehner JN 2011). Alternatively, for non-protein coding RNA with different 3'-end processing, the helicase activity of Sen1 coupled with binding proteins Nrd1 and Nab3 is thought to terminate transcription by unwinding the RNA-DNA hybrid (Kuehner JN 2011).

## **1.9.2 Eukaryotic replication**

Accurate duplication of the genetic material is critical to maintaining genome stability and the faithful transmission of information onto successive generations (Figure 1.11). To achieve this level of fidelity, eukaryotic replication is governed by cell cycle checkpoints, defined by the coordinated actions of phase-specific cyclins and their kinases along with other factors. Given the size of genome, time to complete replication

can differ from ~1.4 hours in simple eukaryotes (yeast) to ~24 hours in higher eukaryotes (mammals). To ensure a timely completion, DNA synthesis initiates from multiple positions on each chromosome at defined 'origin' sites. Since only a single duplication of the genome is required per round of cell division, each origin site is given the license to fire once. During G<sub>1</sub>-phase, the factors CDC6 (cell division cycle 6), CDT1 (chromatin licencing and DNA replication factor 1) and MCM2-7 (mini chromosome maintenance) are recruited to origins and form the pre-replication complex (pre-RC). Replication is initiated upon entry into S-phase (Figure 1.11). As the pre-RC is phosphorylation by CDK2 (cyclin-dependent kinase 2) and DDK (Dbf4/Drf1-dependent CDC7 kinase), licensing factors are inactivated effectively limiting the origin to a single firing event. In addition, CDC45 (cell division cycle 45), GINS (go-ichi-ni-san) are loaded onto the complex (Pacek M 2006) to trigger activation of the candidate MCM2-7 ring helicase, which unwinds the double stranded DNA (Takahashi TS 2005). Replication protein A (RPA) assists in the process by attaching to single stranded DNA (ssDNA) and preventing re-annealing and formation of secondary DNA structures that can potentially block replication. DNA polymerase alpha (DNA Pol  $\alpha$ ) follows suit, using its primase activity to generate a short RNA primer, which is further extended by the polymerase. RFC (replication factor C) then recruits the sliding PCNA (Proliferating cell nuclear antigen) clamp and mediates the switch to DNA polymerase delta/epsilon dimer (DNA Pol  $\delta/\epsilon$ ). Due to the 5' to 3' processivity of DNA polymerases, DNA is synthesised continuously on the leading strand by DNA Pol  $\epsilon$  and semi-discontinuously on the lagging strand by DNA Pol  $\delta$ . As new portions of the dsDNA are unwound lagging strand synthesis is continually primed by DNA Pol  $\alpha$  and as a result leaves behind periodic short RNA inserts between DNA (Okazaki fragments). The removal of these obsolete RNA primers is thought to occur through redundant mechanisms involving RNase H2 digestion, DNA Pol  $\delta$ -mediated displacement and other potential enzymes (Balakrishnan L 2013). The remaining 5' overhangs are excised by FEN1 (flap endonuclease-1) and sealed by DNA ligase I. Replisomes proceed bi-directionally away from the origin and terminate when two opposing replication forks converge (Fachinetti D 2010). Besides this, the removal of topological constraints mediated by topoisomerase II and replication fork barriers (RFB) formed by non-nucleosomal proteins also aid termination (Fachinetti D 2010).

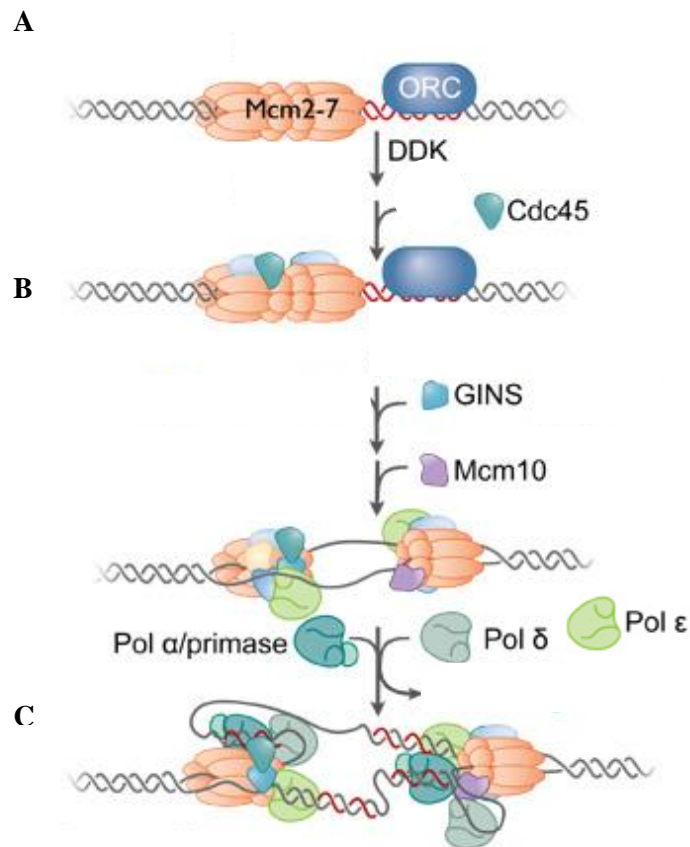


**Figure 1.10. Progression of RNA Polymerase II in eukaryotic transcription.** From *left to right*: **A**) transcription is initiated following the recruitment of general transcription factors and RNA polymerase II to promoter regions. TFIIFH phosphorylates RNA Pol II at Ser5 on the C-terminal tail. **B & C**) prior to elongation within a region of 20 – 40nt, RNA Pol II can undergo abortive initiation or escape, mediated by DSIF and NELF. **D**) kinase activity of p-TEFb phosphorylates NELF to facilitate its removal, DSIF and Ser2 on the tail of RNA Pol II to start elongation. **E**) progressive reduction in Ser5 relative to Ser2 as elongation proceeds until termination. GTF (*light green*) – general transcription factors. TBP (*green*) – TATA-binding protein. TFIIS (*peach*) = transcription factor II A. TFIIFH (*dark green*) – transcription factor II H. Pol II (*blue*) – RNA polymerase II. DSIF (*yellow*) – DRB Sensitivity Inducing Factor. NELF (*light yellow*) – negative elongation factor. P-TEFb (*red*) – positive transcription elongation factor b. Direction of RNA Pol II progression shown as black arrows (Koch F 2008).

### 1.9.3 Regulating transcription and replication

Since both transcription and replication machineries make use of the same template, conflicts between the two are inevitable. The genome is most susceptible to DNA damage during the S-phase because replication requires extensive de-condensation of the chromatin structure and exposes DNA to the harmful effects of endogenous and exogenous factors. In addition to this, topological constraints and DNA-associated proteins, such as those involved in transcription present natural barriers to replication, if unresolved these can lead to fork arrest prompting recombination and spontaneous chromosome rearrangements. Consequently, to protect against such conflicts eukaryotic cells have developed means of coordinating transcription and replication to maintain genome integrity. RFBs are found at gene dense rDNA sites in eukaryotes and are best characterised in yeast. RNA Pol I is highly active in these regions with the continuous transcription of rRNA presenting a challenge to replication. To coordinate this, yeast utilises RFB downstream of rDNA genes to arrest replication forks and prevent collisions with oncoming transcription machinery.

Similar mechanisms involving replication fork pause (RFP) sites have been described at tRNA genes in yeast (Deshpande AM 1996). Besides physical barriers, transcription and replication are compartmentalised on a global scale template (Chakalova L 2005). RNA Pol II is distributed unevenly throughout the genome, organised into foci or 'factories' (Chakalova L 2005). The same can also be said for replication and proposes a model where transcripts or DNA are synthesised by pulling the template through 'immobile' factories as opposed to polymerases moving along 'immobile' stretches of template (Chakalova L 2005). Studies support this view by revealing that the majority (95%) of replication foci do not overlap with transcription foci. Moreover, genome-wide mapping showed that highly transcribed regions exhibit very few replication events (Martin MM 2011). Therefore, by limiting replication to the S-phase and separating regions of transcription with replication, cells are able to limit conflicts in a spatial- and temporal-dependent manner. However, adaptive response to cellular stress requires changes in transcription, which can potentially undermine physiological mechanisms partitioning its coincidence with replication. Recently Hog1, a stress-activated protein kinase (SAPK) in yeast was shown to phosphorylate Mrc1 upon osmotic stress to prevent Cdc45 recruitment



**Figure 1.11. Eukaryotic replisome assembly and activation.** A) putative Mcm2-7 helicase and ORC are loaded onto DNA to start replication. B) Cdc45 associated with Mcm2-7 and DDK-mediated phosphorylation inactivates further licencing. C) GINS and Cdc45 activates the putative Mcm2-7 helicase with Mcm10 coordinating the assembly of the replication fork. D) Pol  $\alpha$ /primase activity primes the DNA template with RNA primers for Pol  $\epsilon$ -driven synthesis on the leading strand and Pol  $\delta$ -driven synthesis on the lagging strand. Newly synthesised strand in shown in red. Mcm2-7 – mini chromosome maintenance 2-7. ORC – origin replication complex. DDK - Dbf4/Drf1-dependent CDC7 kinase. Cdc45 - cell division cycle 45. GINS - go-ichi-ni-san. Mcm10 – mini chromosome maintenance 10. Pol  $\alpha$ /primase – DNA polymerase alpha/primase. Pol  $\epsilon$  – DNA polymerase epsilon. Pol  $\delta$  – DNA polymerase delta. Image from: biology.mit.edu.

therefore delaying replication (Martin MM 2011). Activation of Hog1 is independent of the DNA-damage response and represents a means of preventing transcription and replication conflicts during stress (Martin MM 2011).

Regardless of these efforts, collisions do occur in cells, albeit at a low frequency. Replication sites prone to stress are termed common fragile sites (CFS) with a subset of these identified as early replicating fragile sites (ERFS) (Barlow 2013). ERFS co-occur in regions of actively transcribed genes, rich in CpG and repeat sequences, posing as typical hotspots for transcription and replication collisions (Barlow 2013). Additionally, conflicts are deemed inevitable at human genes greater than 800kb, as these require a longer time to transcribe and overlap with replication in S-phase (Helmrich 2011). As a consequence of these collisions, genome stability becomes severely compromised with the induction of DNA-damage and formation of stable RNA:DNA hybrid structures called R-loops.

## **1.10 R-loops**

### **1.10.1 Physiological functions of R-loops**

R-loops are found abundantly in the genomes of prokaryotes and eukaryotes and perform physiological roles when correctly regulated but can be detrimental to genome integrity when formed unintentionally. R-loop structures are generated by transcription when the RNA transcript threads back on the complementary template DNA and forms a stable interaction displacing the non-template strand as ssDNA. The presence of negative supercoiling behind the progressing RNA polymerase (Drolet 2006), DNA nicks (Roy D 2010) and G-quadruplex structures are factors favouring the formation of R-loops. In particular, C-rich DNA and G-rich RNA interactions constitute a thermodynamically stable structure exceeding that of dsDNA (Roy D 2010).

Physiologically, R-loops affect transcription in a context dependent manner. Their formation at unmethylated CpG islands in human NTERA2 (pluripotent embryonal carcinoma) cells help cap the promoter sequences and protect genes from being silenced by the *de novo* methyltransferase DNMT3B1 (Ginno PA 2012). Phospho-H3S10, a marker of chromatin condensation, is also well correlated with the formation of R-loops (Castellano P.M. 2013). Conversely R-loops and head-to-head antisense transcription are required to maintain expression of the intermediate filament vimentin (Sastre R.B. 2014).

Apart from starting transcription, termination was also shown to involve R-loops at the ends of gene-rich regions (Ginno PA 2013). This was believed to prevent RNA Pol II from running into closely spaced genes by helping to slow the complex through an R-loop mediated anchor. In the case of replication, RNA Pol I pausing in yeast led to R-loop dependent RNA-primed DNA synthesis at the rDNA locus, irrespective of origins (Stuckey R 2015). Furthermore, R-loops are implicated in class-switch recombination (CSR), which alters the isotype of immunoglobulin (Ig) through removal of genes in the heavy chain locus (Roy D 2008). A key part of this process requires single stranded DNA as substrates for the enzyme activation-induced cytidine deaminase (AID) and is provided by the displaced strand following R-loop formation (Roy D 2008).

### **1.10.2 Mechanisms regulating R-loop formation**

While R-loops are important for certain cellular processes, careful regulation is required to prevent these structures from forming outside of their intended context (Figure 1.12). Cells have developed ways of preventing the aberrant formation of R-loops as well as mechanisms to remove them in case they do form. Given that negative supercoiling behind the elongating RNA polymerase provides room for the RNA transcript to thread back onto the complementary DNA strand, topoisomerase I counteracts this by relieving the topology, ensuring that ssDNA is re-sealed (Hage A 2010). Additionally, proper transcript packaging and processing facilitated by RNA biogenesis proteins such as the THO complex (Aguilera 2010) and ASF1/SF2 (alternative splicing factor 1/pre-mRNA splicing factor 2) supplementary to surveillance mechanisms can help to prevent R-loop formation (Gavalda S 2013). In the event that R-loops escape prevention mechanisms, cells employ several enzymes to resolve these structures. The human senataxin helicase and yeast Sen1, help to unwind RNA:DNA hybrids (Mischo HE 2011). Subsequent inactivation of Sen1 in yeast leads to R-loop accumulation (Mischo HE 2011). The specificity of RNase H for degrading RNA was also shown to protect against R-loops (Huertas P 2003). Therefore, controlling the balance of R-loops is important as emerging evidence suggest its accumulation gives rise to genome instability

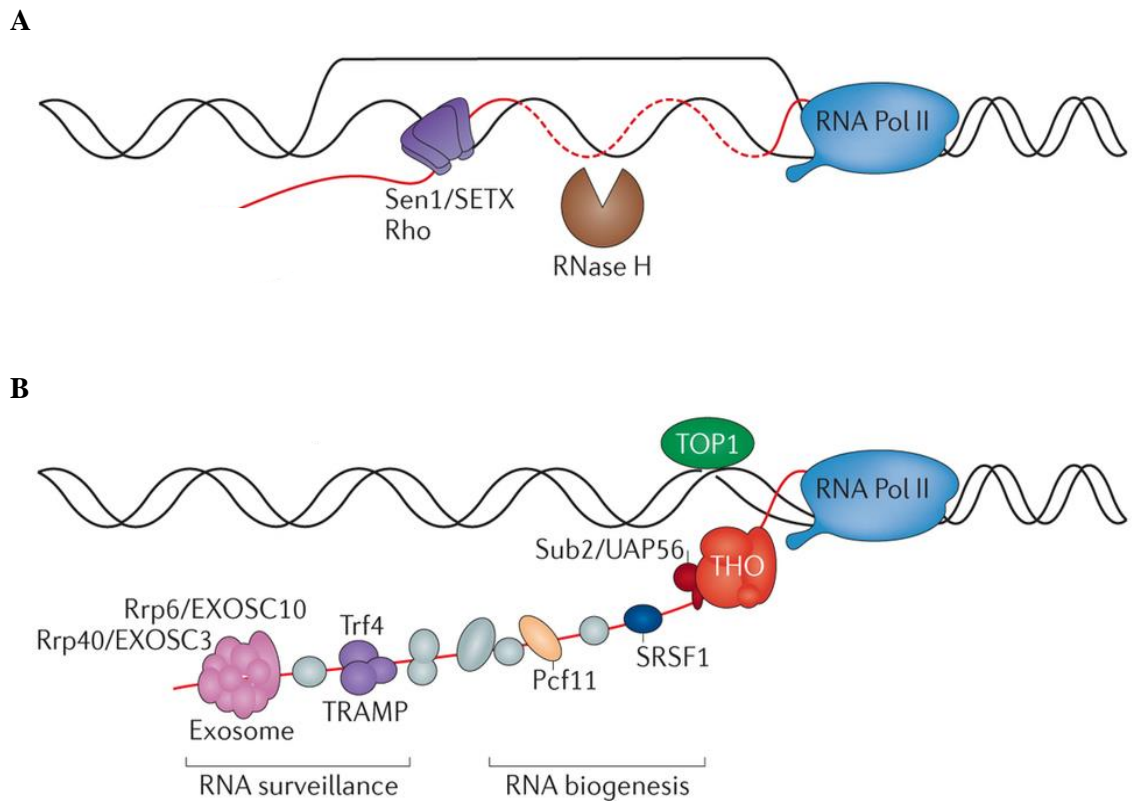
### 1.10.3 R-loops as a source of genetic and epigenetic instability

Co-transcriptional R-loops have been linked to genomic instability arising from recombination, mutations and chromosome segregation and loss. For instance, in B lymphocytes AID can deaminate cytosine to uracil on the ssDNA of R-loops leading to hypermutation. The converted uracil base can also be targeted by components of the base-excision repair (BER) machinery, such as uracil-DNA glycosylase to produce DNA nicks and/or abasic endonuclease (ABE), which gives rise to abasic sites. DNA nicks can generate double strand breaks (DSB), whereas replication past abasic sites leads to incorporation of random nucleotides and hence mutations. Since AID is only expressed in B cells, these mechanisms may not apply in other cell types.

Evidently, defects in proteins that process R-loops can lead to its accumulation and threaten genome stability. Yeast THO-TREX (suppressors of the transcription defects of *hpr1Δ* mutants by overexpression-transcription and export) mutants lacking mRNA maturation and packing capabilities, demonstrate high recombination rates which can be rescued by overexpression of RNase H. This suggests that the build-up of R-loops leads to DNA damage (Huertas P 2003). Similarly, DT40 chicken cells depleted of the mRNA-splicing factor ASF/SF2 lose the ability to suppress R-loops formation and display hypermutation phenotypes (Li X 2005). Recent findings implicate a role for human BRCA1 (breast cancer 1) and BRCA2 (breast cancer 2) in suppressing R-loop formation (Hatchi E 2015; Bhatia V 2014). BRCA1 was shown to direct senataxin to gene termination regions containing R-loops. Disruption of this interaction leads to accumulation of  $\gamma$ H2AX foci and DNA damage (Hatchi E 2015). BRCA2 on the other hand interacts with the TREX-2 mRNA export complex and is required for the processing of R-loops (Bhatia V 2014). Notably BRCA1 and BRCA2 mutations underlie risks of developing human breast cancer.

The structure-specific endonucleases XPF (xeroderma pigmentosum F) and XPG (xeroderma pigmentosum G) belong to the transcription-coupled nucleotide excision repair (TC-NER) pathway, which is involved in the removal of DNA alterations that inhibit the progression of transcribing polymerases. Aside from their main role in TC-NER, more recent studies point towards the XPF and XPG in converting R-loops to DNA damage (Sollier J 2014). It is thought XPF and XPG may cut the RNA:DNA hybrid, in which case single strand breaks are created and can be processed into DSB by incoming





**Figure 1.12. Mechanisms regulating R-loop A) removal and B) formation.** A) Helicase such as SETX (humans), Sen1 (yeast), Rho (bacteria) help to unwind RNA:DNA hybrids whereas RNase H recognise and degrade RNA:DNA hybrids. B) TOP1 relieves negative supercoiling behind the progressing RNA Pol II while RNA biogenesis and surveillance mechanisms help package the transcribed strand to prevent treading-back. SETX/sen1 – senataxin. TOP1 – topoisomerase 1. SRSF1 – serine/arginine-rich splicing factor 1. Pcf11 – cleavage and polyadenylation factor subunit. TRAMP – Trf4–Air2–Mtr4p polyadenylation complex. UAP56 – 56 kDa U2AF65-associated protein (Sub2 in yeast). EXOSC3 – exosome component 3 (Rrp40 in yeast). EXOSC10 – exosome component 10 (Rrp6 in yeast). Image from: (Aguilera 2015).

replication forks or other nucleases (Sollier J 2015). Alternatively, XPF and XPG may cut both the ssDNA and the RNA:DNA hybrid, directly giving rise to DSBs (Sollier J 2015). However, it is still unclear as to whether this mechanism generates inadvertent lesions, or if it is intended to remove detrimental R-loops in circumstances where processing factors are severely compromised and possibly restore damaged DNA through repair pathways.

Patterns of histone modifications could also be affected by transcription-replication conflicts. Repressive chromatin and gene silencing is linked to defects in replication fork progression, as demonstrated by fork stalling promoted by the anticancer drug doxorubicin (Im JS 2014). R-loop formation at mammalian unmethylated CpG promoters (Ginno PA 2012) and transcription termination regions constitute to gene silencing and chromatin condensation, respectively. Abnormal accumulation of R-loops at these sites may overcome their physiological function and lead to a repressed chromatin state. Furthermore, camptothecin-mediated topoisomerase I inhibition promotes R-loop formation and H3K9 dimethylation, resulting in direct suppression of the *FXN* gene (Groh M. et al. 2014). The combination of R-loop and repressive histone mark was thought to drive the onset of Friedreich ataxia (FRDA) in which the *FXN* gene undergoes a GAA-repeat expansion (Groh M. et al. 2014).

#### **1.10.4 R-loops and double strand breaks as products of transcription and replication conflicts**

A more popular view of R-loop induced genome instability arises from the collision between transcription and replication. This is widely supported by the observation that inhibiting replication greatly abolishes R-loop dependent DNA damage (Gan W 2011). In addition, the role of topoisomerase I in ASF/SF2-mediated mRNA processing and its ability to relieve supercoiling are important for preventing the accumulation of R-loops and DSB in S-phase (Tuduri S 2010). Generally, head-on collisions between transcription and replication on the lagging strand present a greater threat to genome stability compared to co-directional collisions on the leading strand (Ivanova D 2015). This is evident in prokaryotes such as *E. coli* where most genes are encoded on the leading strand and inversion of the *rnn* operon to a head-on orientation with replication leads to reduced fork speeds and activation of DNA damage response

(Septenville A.D. 2012). However, no specific patterns of gene orientation have been found in humans (Ivanova D 2015). Evidence from prokaryotes suggests R-loops may either be the cause or the result of transcription and replication collisions. It is likely that R-loop structures extending behind the RNA polymerases may stall oncoming co-directional replication forks, leading to its collapse and DSB formation (Dutta et al. 2011). Alternatively head-on conflicts can lead to formation of R-loops behind the stall RNA polymerase, in which case DNA damage may manifest at the site of collision or the exposed ssDNA (Mirkin and Mirkin 2005).

Inevitably, DNA damage would trigger recombination events potentially resulting in gross chromosomal changes and as such this can lead to transcription-associated recombination (TAR) (Gottipati et al. 2008). The exposed ssDNA in R-loops are also subject to DNA damage, leading to transcription-associated mutagenesis (TAM) (Zlotorynski 2015). Ultimately, disruptions in either transcription or replication can lead to conflicts with disastrous consequences for the genome as well as the epigenome, with strong implications in the development of cancer and neurodegenerative diseases (Brambati A 2015).

## **1.11 Aims and hypothesis of the present study**

Histone deacetylase inhibitors can elicit genotoxic effects through changes in the epigenome. Alteration in cellular replication and transcription in response to these changes are implicated in the resulting genotoxicity. Independent studies have also shown that replication-transcription conflicts can lead to DNA damage and genome instability. In light of our current understanding, this study proposes that HDAC inhibitors induce changes in histone acetylation, which alters the organisation of chromatin, thereby affecting the normal function of processes such as replication and transcription in such a way that the frequency of conflicts between the two increases, leading to DNA damage induction and genome instability.

To test this hypothesis, the current study aims to initially make use of budding yeast as a model organism, for which the ChIP-chip method has previously been established (Teng et al. 2011), to examine the genotoxicity of HDAC inhibitors. Additionally, the ChIP-chip method will be established in the human relevant TK6 lymphoblastoid cell line as a genome-wide genotoxicity. Following this, the established method will be used to map chromatin associated events such as histone acetylation, transcription and replication in TK6 cells in response to HDAC inhibitor treatment. By integrating these datasets together using bioinformatic tools, the question of whether HDAC inhibition can induce increases in the frequency of transcription-replication conflicts will be examined. Finally, the knowledge gained here will be used collectively to address how a mechanistic understanding can be used to identify new endpoints for human safety assessment.

## **Chapter II**

### **Methods and Materials**



## **2.1 Mammalian methods**

### **2.1.1 Mammalian cell culture and treatment**

Human B lymphoblastoid TK6 cells were obtained courtesy of Prof. Anthony Lynch (GlaxoSmithKline, Ware, UK). TK6 cells are derived from the WIL-2 cell line and are heterozygous at the thymidine kinase locus (Skopek et al. 1978). The cell line was maintained in suspension between  $1 - 9 \times 10^5$ /ml in RPMI 1640 + GlutaMAX™ supplement (Gibco, Paisley, UK) containing 10% foetal bovine serum (FBS) (Gibco, Paisley, UK). The cells were passaged approximately every 3 days. Cells nearing confluency were split by removing ~90% of the cell suspension and replacing with fresh media for a final density of  $\sim 1 \times 10^5$ /ml.

Human breast adenocarcinoma MDA-MB-231 (p53 mutant) and prostate adenocarcinoma PC-3 (p53 null) cell lines were obtained courtesy of Dr Zara Poghosyan (Cardiff University, UK) and maintained as adherent cells in DMEM + GlutaMAX™ supplement containing 10% foetal bovine serum (Gibco, Paisley, UK) and F-12K medium + GlutaMAX™ supplement containing 10% foetal bovine serum (Gibco, Paisley, UK), respectively. The cells were maintained between approximately 10% - 80% confluency. Cells were passaged by removing the media and incubating at 37°C in the presence of 0.5% Trypsin-EDTA solution (Gibco, Paisley, UK) for up to 5mins. The cells were then collected by spinning at 1000g for 5mins at room temperature, replacing the trypsin solution with fresh media and reseeding a small proportion of the cells.

All human cell lines were cultured in a copper-lined humidified incubator (Heracell™ 150i, ThermoFisher) at 37°C in the presence of 5% CO<sub>2</sub>. All cell lines are cultured for no more than 6 weeks, following which they are disposed of and new cells are thawed from the liquid nitrogen stock. Exponentially dividing cells were used at an assay dependent-density prior to drug treatment.

### **2.1.2 Freezing and thawing cells**

Cells to be frozen were harvested during the exponential growth phase by centrifuging at 1000g at room temperature for 5mins. For adherent cell lines, the cells were first detached from the flask surface by incubating with 0.5% Trypsin-EDTA solution (Gibco, Paisley, UK) for up to 5mins at 37°C. Freezing mixture was prepared by

mixing Hybri-Max<sup>TM</sup> dimethyl sulfoxide (DMSO) (Sigma-Aldrich Company Ltd, Gillingham, UK) with foetal bovine serum (Gibco, Paisley, UK) in a ratio of 1:4, respectively. Cells were re-suspended in media to a concentration of  $2 \times 10^6$ /ml. 0.5ml of the cell suspension was then gently mixed with an equal volume of the prepared freezing mix and transferred to a 1.5ml cryogenic tube (Nalgene<sup>TM</sup>, Thermo Scientific, Hempstead, UK). Individual cryotubes were placed in a Mr. Frosty<sup>TM</sup> container (Thermo Scientific, Hempstead, UK) with isopropyl alcohol overnight at  $-80^{\circ}\text{C}$  to allow gradual cooling ( $-1^{\circ}\text{C}/\text{min}$ ), before transferring to liquid nitrogen (38K CryoStorage System, Taylor Wharton, US) for long-term storage.

To thaw out new cells, the cryotube was removed from the liquid nitrogen store and immersed halfway in a  $37^{\circ}\text{C}$  water bath (NE1 Clifton, Thermo Scientific, Hempstead, UK), briefly until defrosted. The cryotube was then cleaned in 70% ethanol before transferring the contents to a 15ml tube. 9ml of fresh media was then added slowly in a drop-wise manner and centrifuged at 1000g for 5mins at room temperature to remove the freezing mixture. The cell pellet is then re-suspended in 10ml fresh media and transferred to a T75 flask (Nunc<sup>TM</sup>, Thermo Scientific, Hempstead, UK). The media is replaced following overnight culture to remove any dead cells and residual DMSO.

### 2.1.3 Counting cells

To obtain a reliable estimate of cell numbers, cells in culture were diluted 1:10 in PBS solution and transferred to a sample cup (Beckman Coulter, High Wycombe, UK) before being counted using an automated Vi-CELL<sup>TM</sup> XR system (Beckman Coulter, High Wycombe, UK). Relative cell count (RCC) was used as a measure of general cytotoxicity, expressing the number of cells in the treated sample as a percentage of the number of cells in the untreated control:

$$\frac{\text{no. of cells in treated sample}}{\text{no. of cells in the control sample}} \times 100$$



## **2.1.4 Flow cytometry**

### **2.1.4.1 Apoptosis detection**

Annexin V-FITC binding to externalised phosphatidylserine in combination with propidium iodide (PI) was used to assess for the induction of apoptosis and cell death. The protocol was carried out as specified by the kits' manufacturer (eBioscience, Vienna, Austria) with slight modifications:  $0.4 - 1 \times 10^5$  cells were collected following treatment and washed once in PBS solution before re-suspending in 200 $\mu$ l of 1X Binding buffer containing 5 $\mu$ l anti-Annexin V-FITC dye. The cell suspension was incubated for 10mins at room temperature in the dark before adding 10 $\mu$ l propidium iodide and proceeding to analysis on the flow cytometer (Accuri C6, BD Biosciences, Oxford, UK). Anti-Annexin V-FITC and PI were collected in FL1 and FL3 channels, respectively. A plot of anti-Annexin V-FITC (log x-axis) vs. PI (log y-axis), gated to exclude the debris (FSC vs. SSC), was used to interrogate the percentage of necrotic and apoptotic (early/late) cells in the population by applying a quadrant gate relative to the viable population. Cells staining positive for either Annexin V-FITC, PI or both were scored.

### **2.1.4.2 Histone H2AX phosphorylation and Histone H3K9/14 acetylation**

Changes in histone modifications were measured using protocol outlined by (Smart et al. 2011).  $2 - 4 \times 10^5$  cells were harvested and washed once in PBS solution. The pellet was then re-suspended in 1ml of pre-chilled lysis buffer (10mM Tris HCl pH8.0, 2.5mM MgCl<sub>2</sub>; 320mM sucrose) and incubated on ice for 10mins before centrifuging (1000g, 4°C; 5mins). The nuclei were re-suspended in 100 $\mu$ l PBS solution (+ 1mg/ml BSA) containing an optimised amount of antibody, either anti- $\gamma$ H2AX-FITC (Millipore, Hertfordshire, UK) or anti-acetyl-histone H3-PE (Milli-Mark™, Millipore, Hertfordshire, UK). Anti- $\gamma$ H2AX-FITC (Millipore, Hertfordshire, UK) and anti-acetyl-histone H3-PE (Milli-Mark™, Millipore, Hertfordshire, UK) signals were both collected in the FL1 channel independently.

#### **2.1.4.3 InVitro MicroFlow<sup>®</sup> assay**

The InVitro MicroFlow<sup>®</sup> assay (Litron laboratories, Rochester, USA) was used to measure the induction of micronuclei in response to DNA damage. The assay was conducted in 96-well plates. For each sample, 100µl of pre-warmed media was first added to the well. Next, 100µl of media containing the drug of interest, diluted to 2X the desired concentration, was added to the well designated as the highest drug concentration. The contents of the well were mixed by pipetting and a ½ serial dilution is carried out by transferring 100µl of the mixed media to the next well. 100µl of  $\sim 2 - 2.5 \times 10^4$  cells were added to each well for a final volume of 200µl.

Following drug treatment, the supernatant was removed by centrifuging the 96-well plate at 300g for 5mins. Cells were loosened by gentle tapping and placed on ice for 20mins before further processing. Staining solutions were prepared from components provided in the kit and scaled down based on the number of samples. To stain dead and dying cells, 50µl of Dye A working solution was added to each well and mixed with the cells by pipetting. The lid was removed and the 96-well plate left on wet ice and exposed to a bright light source for 30mins. Following this, cells were washed with 150µl of cold 1X PBS + FBS solution and the supernatant removed by spinning at 300g for 5mins, 4°C. The exposure to light was limited from here on. To stain for total chromatin, cells were re-suspended in 100µl Complete lysis solution 1 and mixed before being placed in a 37°C incubator for 1h. 100µl Complete lysis solution 2 was then added to each well and the samples left overnight at 4°C before analysis. Samples were left to equilibrate at room temperature in the dark for 30mins before being analysed on the flow cytometer (Accuri C6, BD Biosciences, Oxford, UK). The EMA signal was collected in the FL3 channel and the SYTOX Green was collected in the FL1 channel. The data was analysed using the C6 Accuri software (Accuri C6, BD Biosciences, Oxford, UK), collecting 10,000 events per sample. Gating was carried out as described in the manufacturers' protocol (Litron laboratories, Rochester, USA).

#### **2.1.4.4 Cell cycle analysis**

Propidium iodide (PI) was used for cell cycle analysis carried out independently of the *in vitro* Microflow<sup>®</sup> assay (Litron laboratories, Rochester, USA).  $1 \times 10^6$  cells were harvested by centrifuging at 1000g for 5mins at 4°C and washed once in cold PBS

solution. To fix cells, pre-chilled 70% ethanol was added drop-wise while vortexing in between. Samples were left on ice for at least 1h before washing with cold PBS solution (1000g for 5mins at 4°C). The fixed cells are re-suspended in 250µl PBS containing 0.2mg/ml RNase A (ThermoFisher Scientific, Loughborough, UK) and incubated at 37°C for 1h before adding 10µl of 1mg/ml PI (abcam, Cambridge, UK) and analysed on a flow cytometer (Accuri C6, BD Biosciences, Oxford, UK). 10,000 events are collected per sample point using the Accuri C6 software (Accuri C6, BD Biosciences, Oxford, UK) and the data is analysed using FlowJo v10.07 (FLOWJO, LLC, Ashland, USA). Doublets are first eliminated by gating on the healthy population (log FSC vs log SSC), followed by the exclusion of aggregates (plot of PI FL2-A vs FL2-W). This subset of data is then used to generate cell cycle profiles based on the Watson (pragmatic) model (Gaussian fit to G<sub>1</sub> and G<sub>2</sub> phases).

## **2.1.5 Immunoblot analysis**

### **2.1.5.1 Whole cell protein extraction**

To extract total protein content from cells, ~2.5 - 5x10<sup>5</sup> cells were seeded in 6-well plates in a final volume of 2ml. Following treatment, cells were collected and washed once with cold PBS solution and lysed on ice for 10mins in 50µl of RIPA buffer (150mM NaCl, 0.1% SDS, 1% Triton-X 100, 0.5% sodium deoxycholate, 50mM Tris-HCl pH8.0) containing 1X PMSF (Sigma-Aldrich Company Ltd, Gillingham, UK) and 1X PIC (Sigma-Aldrich Company Ltd, Gillingham, UK). In order to improve the lysis of nuclear components, samples are sonicated using a Bioruptor<sup>®</sup> Standard sonication device (Diagenode, Ougrée, Belgium) for 2 cycles (10s on, 30s off) on 'High' settings, then centrifuged at 15,000g for 10mins at 4°C to remove debris. The protein concentration was determined as previously described using the Bradford assay.

### **2.1.5.2 Protein quantification**

Extracted protein samples were quantified using the Bradford assay. A calibration curve was prepared using known concentrations of BSA dilutions against the absorbance (A<sub>590nm</sub>). Up to 5µl of prepared protein extract was diluted to a final volume of 1ml in 1X Bio-Rad protein assay reagent (Bio-Rad laboratories Ltd., Hertfordshire, UK). Samples

are transferred to micro-cuvettes (Thermo Scientific, Hempstead, UK) and the absorbance ( $A_{590\text{nm}}$ ) measured against a reference using a Jenway 6305 spectrophotometer (Bibby Scientific Ltd., Staffordshire, UK).

### **2.1.5.3 SDS-Polyacrylamide gel electrophoresis**

Proteins were resolved using 4 - 20% gradient Bolt® Bis-Tris Plus gels (Invitrogen™, Life Technologies Ltd., Paisley, UK) assembled in a Mini gel tank (Invitrogen™, Life Technologies Ltd., Paisley, UK) with 1X Bolt® MOPS SDS running buffer. 20 - 40µg of whole cell extract was mixed with NuPAGE LDS sample buffer (4X) (Novex™, Life Technologies Ltd., Paisley, UK) and heated for 8mins at 99°C on a Mastercycler nexus GX2 (Eppendorf, Stevenage, UK). Denatured lysates were then loaded sequentially onto the gel, with 5µl of Novex® Sharp pre-stained standard (Novex™, Life Technologies Ltd., Paisley, UK) being loaded separately in the first lane. The gel was then run at 150V for up to 15mins using a Bio-Rad PowerPac 300 (Bio-Rad laboratories Ltd., Hertfordshire, UK). Once complete, the cassette was disassembled and the gel removed and rinsed in dH<sub>2</sub>O prior to transfer.

### **2.1.5.4 Western transfer, probing and detection**

Transfer of proteins from the electrophoresed gel to the PVDF membrane was carried out using an iBlot® 2 Dry blotting system (Invitrogen™, Life Technologies Ltd., Paisley, UK). The gel was placed facing down on the PVDF membrane provided in the iBlot® 2 transfer stack (Invitrogen™, Life Technologies Ltd., Paisley, UK) and covered with a filter paper soaked in dH<sub>2</sub>O. A roller was used to remove any trapped bubbles before placing the remaining cathode stack and sponge on top, completing the transfer stack. The blotting system was run using the manufacturers' protocol based on the proteins' molecular weight (Invitrogen™, Life Technologies Ltd., Paisley, UK).

After completing the transfer, the PVDF membrane is removed from the stack and rinsed in dH<sub>2</sub>O. Probing of the membrane was carried out using an iBind™ Western device (Invitrogen™, Life Technologies Ltd., Paisley, UK). Firstly, the working iBind solution was prepared by adding 6ml of the 5X iBind™ buffer (Novex™, Life Technologies Ltd., Paisley, UK) with 23.7ml of dH<sub>2</sub>O and 300µl of 300X additive

(Novex™, Life Technologies Ltd., Paisley, UK). The solution was then mixed by vortexing. An optimised concentration of primary and secondary antibody was prepared in 2ml of the working iBind solution and mixed by pipetting. To probe the membrane, an iBind™ card (Novex™, Life Technologies Ltd., Paisley, UK) was placed in the iBind™ Western device (Novex™, Life Technologies Ltd., Paisley, UK) and soaked with 5ml of the working iBind solution. 1ml of the working iBind solution was pipetted on the centre of the card and then PVDF membrane placed on top, with the protein side facing down. The iBind device is closed and the antibody solutions were added to the wells of the cassette sequentially as outlined in the manufacturers' instructions (Life Technologies Ltd., Paisley, UK). The blot was left to run by lateral flow for at least 3h at room temperature or overnight at 4°C.

After probing, the PVDF membrane was removed from the iBind Western device (Novex™, Life Technologies Ltd., Paisley, UK) and rinsed in dH<sub>2</sub>O. The membrane was then incubated for 5mins in 1ml of SuperSignal™ West Dura Extended Duration Substrate (Thermo Scientific, Hempstead, UK) containing a 1:1 mixture of luminol enhancer and stable peroxide buffer. Chemiluminescent signals from the blot were detected using Labworks software v4.6 (Labworks Inc., Costa Mesa, USA) connected to a BioSpectrum 610 Imaging System with an OptiChemi 610 camera (UVP Inc., Upland, USA).

#### **2.1.5.5 Membrane stripping**

To re-probe the blot, the PVDF membrane is stripped in 10ml of Restore™ PLUS Western Stripping buffer (Thermo Scientific, Hempstead, UK) for 15mins at room temperature. The membrane is then washed once in dH<sub>2</sub>O and the efficiency of the stripping checked using SuperSignal™ West Dura Extended Duration Substrate (Thermo Scientific, Hempstead, UK) as aforementioned.

## 2.1.6 Chromatin Immunoprecipitation

### 2.1.6.1 Chromatin preparation

Chromatin must first be extracted from the cell and sheared to a range of fragment sizes before it can be used in immunoprecipitation reactions. TK6 cells are cultured in T175 flasks (NUNC™, Thermo Scientific, Hempstead, UK) to a density of  $\sim 5 \times 10^5$  cells/ml in 100ml RPMI 1640 before being treated with DMSO (solvent control) or 400nM Trichostatin A (HDACi treated sample) for 6hrs. Cells from 4 - 8 flasks are harvested per treatment condition (1000g for 5mins at room temperature) and re-suspended in the same volume of PBS solution in a 2L erlenmeyer flask at  $\sim 5 \times 10^5$  cells/ml. The flasks are then placed on an Innova™ 2100 bench top orbital shaker (New Brunswick Innova™, Fisher Scientific, Loughborough, UK) set at 80rpm to continuously mix the cells during fixation at room temperature. For treated cells, Trichostatin A is added to the PBS at a final concentration of 400nM, to maintain the effects of the inhibitor during handling. DNA associated proteins are cross-linked by diluting a stock solution of 37% formaldehyde (Sigma-Aldrich Company Ltd, Gillingham, UK) in the cell/PBS solution to a final 1% (v/v) for 8mins. The reaction is quenched by adding 2.5M stock glycine to a final concentration of 125mM, for 5mins. Following this, the fixed cells are allowed to cool on ice for 10mins before being transferred to 50ml conical tubes (Falcon™, Fisher Scientific, Loughborough, UK) and collected (1000g for 5mins at 4°C). Subsequent steps are all carried out on ice or at 4°C, to preserve protein activity. Cells are washed 3x in cold PBS solution (1000g for 5mins at 4°C) to remove traces of the formaldehyde fixative. To efficiently shear chromatin from fixed TK6 cells, an additional lysis step is necessary to remove cytoplasmic/non-fixed components, this was carried out by lysing  $\sim 1 \times 10^7$  cells per 1ml ChIP lysis buffer (10mM Tris-HCl pH8, 10mM NaCl, 0.5% NP-40) on ice for 10mins. After centrifuging (1000g for 5mins at 4°C) and removing the supernatant, the 'nuclear' pellet obtained is finally re-suspended in 330 $\mu$ l ChIP shearing buffer (50mM Tris-HCl pH8, 10mM EDTA pH8, 1% SDS) per  $\sim 1 \times 10^7$  cells in a standard 1.5ml microfuge tube (Fisher Scientific, Loughborough, UK). In order to shear the chromatin to an average fragment size of 500bp, samples were sonicated using a Bioruptor® Standard sonication device (Diagenode, Ougrée, Belgium) with a NESLAB Digital Plus RTE7 circulating water bath (ThermoScientific, Loughborough, UK) set to 4°C. Six samples are loaded into the tube holder each time and sonicated for 7

cycles of 30s on/off using 'High' settings. Following this, the samples are spun at 15,000g for 10mins at 4°C to remove any debris resulting from the shearing. The chromatin prepared from the same treatment is then pooled together and mixed by vortexing; 50µl of the sample is transferred to a new tube and de-crosslinked overnight at 65°C in a water bath. This sample is purified using the PureLink® PCR purification kit (Invitrogen™, Life Technologies Ltd., Paisley, UK) and the extent of the shearing assessed using gel electrophoresis. The remaining chromatin is aliquoted into lots of 1ml in 1.5ml Eppendorf® Protein LoBind tubes (Fisher Scientific, Loughborough, UK), snap-frozen in liquid nitrogen and stored at -80°C.

#### **2.1.6.2 DNA gel electrophoresis**

The fragment distribution of sheared chromatin is routinely checked after each preparation using 1.2% mini agarose gels containing 0.2µg/ml ethidium bromide. The gels are electrophoresed at 80V for 50mins in TBE buffer (100mM Tris-base, 100mM Boric acid, 20mM EDTA pH8). The image is assessed under UV light using a BioSpectrum 610 Imaging System (UVP Inc., Upland, USA).

#### **2.1.6.3 Immunoprecipitation**

To estimate the binding locations of proteins/histone marks on a genomic scale, the corresponding DNA fragments are isolated using highly specific antibodies against the target of interest. This immunoprecipitation (IP) reaction is typically scaled to use larger amounts of starting chromatin and antibody to satisfy the requirements for whole genome amplification using the WGA2 kit (Sigma-Aldrich Company Ltd, Gillingham, UK). The amount of material used greatly depends on the abundance of the target, typically for rare events such as γH2AX, 1.2x10<sup>8</sup> cell equivalent of chromatin is required for a single IP, however as few as 3x10<sup>7</sup> cell equivalent of chromatin is enough for studying histone acetylation. The optimal amount of antibody was determined by performing a titration experiment.

To perform immunoprecipitation using chromatin prepared from TK6 cells, the ChIP-validated antibody was first bound to superparamagnetic Dynabeads® (Invitrogen™, Life Technologies Ltd., Paisley, UK) from the same species that the

antibody is raised in. 50µl per sample of Dynabeads<sup>®</sup> are washed 3x in PBS solution (+ 1mg/ml BSA), each time the beads are allowed to collect at the side of the tube by placing it in a DynaMag<sup>™</sup>-Spin magnet (Life Technologies Ltd., Paisley, UK) for 1min. The beads are re-suspended again in the same volume of PBS solution (+ 1mg/ml BSA) and transferred to 1.5ml Eppendorf<sup>®</sup> Protein LoBind tubes (Fisher Scientific, Loughborough, UK). The optimised amount of antibody is added to the beads and incubated for 60mins at 30°C in a Thermomixer<sup>®</sup> comfort (Fisher Scientific, Loughborough, UK) set to 1,300rpm. Based on the manufacturers specifications, 50µl of Dynabeads<sup>®</sup> have a capacity to bind to 0.4 - 4µg of antibody. For larger amounts of antibody, the amount of dynabeads used is scaled accordingly. TSA treated and untreated chromatin was defrosted on ice and centrifuged at 15,000g for 2mins at 4°C. This step is essential to removing the bulk of the SDS present in the ChIP shearing buffer. While SDS is important for shearing the chromatin, it negatively affects the ability of antibodies to immunoprecipitate their targets. To counteract any residual SDS remaining in the chromatin sample, 500µl of ChIP dilution buffer (50mM Tris-HCl pH8, 150mM NaCl, 2mM EDTA pH8, 1% Triton-X, 0.01% SDS) containing 1% Triton-X100 is added to 1ml of chromatin to sequester the SDS. Additionally, fresh phenylmethane sulfonyl fluoride (PMSF) and protease inhibitor cocktail (PIC) (Sigma-Aldrich Company Ltd, Gillingham, UK) are added to 1X concentration. At this stage 50µl of chromatin is kept as the 'input' (IN) sample, representing the non-immunoprecipitated background. Once the antibody has been coupled to the Dynabeads<sup>®</sup>, 2 further washes in PBS solution (+ 1mg/ml BSA) are carried out to remove any unbound antibody. 50 - 100µl of antibody/Dynabead<sup>®</sup> mixture (antibody dependent) is then added to 1.5ml of chromatin and further incubated in a thermomixer for 3hrs at 21°C at 1,300rpm.

After the incubation is complete, the protein of interest and the corresponding DNA fragment will be bound to the antibody/Dynabead<sup>®</sup> complex. A series of wash steps are then carried out for 3mins each on a Revolver<sup>™</sup> platform (Labnet International Inc., Edison, USA) in 1ml of buffer and collected using a DynaMag<sup>™</sup>-Spin magnet as before. The IP samples are first washed 1x in low salt solution (20mM Tris-HCl pH8, 150mM NaCl, 2mM EDTA pH8, 1% Triton-X, 0.1% SDS), followed by 3x in high salt solution (20mM Tris-HCl pH8, 500mM NaCl, 2mM EDTA pH8, 1% Triton-X, 0.1% SDS) and 1x LiCl solution (250mM LiCl, 0.5% NP-40, 0.5% Na-deoxycholate, 10mM Tris-HCl pH8, 1mM EDTA pH8). A final wash in TE buffer (10mM Tris-HCl pH8, 1mM EDTA pH8) was carried out to remove residual salts from the wash buffers before eluting in



125µl ChIP elution buffer (10mM Tris-HCl pH8, 1mM EDTA, 0.5% SDS) on a thermomixer set at 65°C, 1300rpm for 25mins. The eluted IP sample is separated from the Dynabeads® using the DynaMag™-Spin magnet and transferred to a new 1.5ml tube. To de-crosslink the protein from the DNA, both the IP and IN sample is left in a 65°C water bath overnight. The DNA is then purified using PureLink® PCR purification kit (Life Technologies Ltd., Paisley, UK), eluted in 50µl E1 buffer and stored at -20°C.

#### **2.1.6.4 Real-time qPCR**

Real-time quantitative PCR was used to measure enrichment in the immunoprecipitated sample relative to the input. Using SYBR Green chemistry and primers for genes where the protein/histone modification is known to be present/absent, the enrichment between treated and untreated samples can be compared. For examining background levels of a signal, the comparison was made between IP's carried out using the antibody specific for the target of interest and a matching IgG of the same species. This method was used to determine the success of the IP reaction prior to proceeding to the microarrays. 10x dilutions of the IP and IN samples were made. 4µl of the IP and IN sample was diluted in 36µl of ddH<sub>2</sub>O. For standards, a 10x serial dilution (1/10 – 1/100,000) was prepared using 5µl IN in 45µl ddH<sub>2</sub>O. iTaq™ Universal SYBR® Green Supermix (Bio-Rad laboratories Ltd., Hertfordshire, UK) was prepared by adding forward and reverse primers at a final concentration of 1mM from a 100mM stock. The solutions were mixed by vortexing and 5µl of the diluted DNA (IN, IP or standards) was mixed with 5µl of SYBR Green/primer mix in a Hard-Shell® 96-well semi-skirted PCR plate (Bio-Rad laboratories Ltd., Hertfordshire, UK). ddH<sub>2</sub>O was used as a non-template control (NTC) and samples were loaded in triplicates. The volumes specified are for the analysis of 2 gene loci, typically one positive and one negative, on a single 96-well PCR plate. Samples runs were performed on a CFX Connect™ Real-Time PCR detection system (Bio-Rad laboratories Ltd., Hertfordshire, UK) following the cycling protocol documented by the manufacturer. The data was then analysed using CFX Manager™ Software v3.1, sample wells within triplicates displaying greater than 0.5 Ct variation was removed. Melt curves were checked for non-specific amplicons.

#### 2.1.6.4.1 Primer design and optimisation

Primers for real-time qPCR were designed based on the human genome assembly hg19 (UCSC Genome Browser) (Appendix, IV). Potential primers with an optimal size of 22 - 27bp, GC content of 40 - 60% and  $T_m$  between 65 - 70°C were generated using Primer 3 v0.4 based on the target sequence. For real-time qPCR of ChIP samples, primers yielding 100 - 200bp amplicons were selected and validated *in silico* (UCSC Genome Browser). Primers were synthesised on a 0.01 $\mu$ mol scale, using HPSF purification (Eurofins Genomics, Ebersberg, Germany) and reconstituted in ddH<sub>2</sub>O to 100mM. To efficiently amplify regions of interest, optimisation was first carried out to identify parameters under which the primer would perform the best.  $T_m$  was determined by running a temperature gradient, values displaying low  $C_t$  without non-specific amplification were deemed suitable. Next, a qPCR run at a defined  $T_m$  across a range of template concentrations (10ng - 16pg) was used to assess the reaction efficiency. Fully optimised primers showed: 1) efficiency and sensitivity:  $E > 90\%$ , 2) specificity: no non-specific amplicons; 3) reproducibility: triplicates are within  $0.5C_t$ .

#### 2.1.6.5 Determining optimal antibody conditions for ChIP-chip

Optimal antibody concentration for chromatin immunoprecipitation was determined empirically by carrying out a small-scale titration reaction, typically using 0 - 5 $\mu$ g of antibody with 500 $\mu$ l of chromatin. The binding levels relative to the IgG control was assessed using qPCR with primers for regions of expected high (positive control loci) and low (negative control loci) enrichment.

For proteins where the binding location in the genome is unknown based on literature searches, a post-IP western was carried out instead to examine whether if the antibody is ChIP-grade. To do this, a titration was performed as mentioned above, except the beads are re-suspended in 50 $\mu$ l of RIPA buffer after the series of washes in salt solution. 20 $\mu$ l of 4X NuPAGE LDS sample buffer (Novex<sup>TM</sup>, Life Technologies Ltd., Paisley, UK) was added and heated at 99°C for 8mins. The beads were separated from the IP sample with a DynaMag<sup>TM</sup>-Spin magnet; 1 $\mu$ l was loaded for western blot analysis and the presence of the immunoprecipitated protein examined against the IgG control.

## **2.1.7 ChIP-chip**

### **2.1.7.1 Whole-genome amplification**

The yield of DNA from ChIP is too low for microarray analysis. To overcome this, samples are exponentially amplified using the GenomePlex<sup>®</sup> Complete Whole Genome Amplification kit (WGA2) (Sigma-Aldrich Company Ltd, Gillingham, UK). The result is a ~500-fold increase in DNA that is representative of the original IP sample. The method was performed as outlined in the manufacturer's instructions with slight modifications. 46µl of the IP sample was concentrated to ~11µl using a Savant<sup>™</sup> ISS-100 Speed Vac System (Fisher Scientific, Loughborough, UK). The IN sample concentration is quantified on a NanoDrop and diluted to ~2ng/µl. For the preparation of the OmniPlex<sup>®</sup> library, 2µl of 1X Library Preparation Buffer and 1µl of Library Stabilisation Buffer are added to 11µl of either the diluted IN or concentrated IP sample. The reaction is mixed by vortexing and heated at 95°C for 2mins on a Mastercycler nexus GX2 (Eppendorf, Stevenage, UK). 1µl of Library Preparation Enzyme was added and mixed before returning to incubate following the manufacturers' instructions.

Following this, samples are amplified by PCR with the addition of 7.5µl of 10X Amplification Master Mix, 47.5µl of ddH<sub>2</sub>O and 5µl of WGA Polymerase to the 15µl reaction and thermocycled. The samples are then purified using the PureLink<sup>®</sup> PCR purification kit (Invitrogen<sup>™</sup>, Life Technologies Ltd., Paisley, UK) and eluted in 50µl of TE buffer supplied with the BioPrime<sup>®</sup> Total Genomic Labelling System (Life Technologies Ltd., Paisley, UK). The purity and yield of the amplified ChIP DNA is then quantified using the NanoDrop and stored at -20°C.

### **2.1.7.2 Fluorescent labelling**

For two colour ChIP-microarrays, IP and IN samples are labelled with two different fluorescent dyes and hybridised onto a single array. The ratios of fluorophore intensities (IP/IN) are a measure of the relative signal enrichment. Sample labelling was carried out using the BioPrime<sup>®</sup> Total Genomic Labelling System (Life Technologies Ltd., Paisley, UK). Briefly, 2µg of amplified IN and IP samples are used as starting material for the reaction. The volume was made up to 22µl using TE Buffer and 25µl of 2X Alexa Flour<sup>®</sup> 3 was added to the IN and 25µl Alexa Flour<sup>®</sup> 5 added to the IP. Exposure

to light was minimised to reduce photo-induced signal degradation, particularly for Alexa Flour<sup>®</sup> 5. The reaction was mixed by pipetting and incubated on a thermocycler at 95°C for 5mins and immediately cooled to 4°C. 3µl of highly concentration Exo-Klenow Fragment (40U/µl) was added and further allowed to incubate at 37°C for 2hrs. Once the labelling is complete, samples are purified using the PureLink<sup>®</sup> PCR purification kit (Invitrogen<sup>™</sup>, Life Technologies Ltd., Paisley, UK) as mentioned previously.

### **2.1.7.3 DNA quantification**

A NanoDrop 1000 Spectrophotometer (Thermo Scientific, Hempstead, UK) was used to measure the yield and purity of DNA after PCR amplification. After referencing with the appropriate reagent, absorbance was measured at 230nm, 260nm and 280nm. For pure DNA, a 260/280nm ratio of ~1.8 and a 230/260nm ratio of 2.0 - 2.2 is expected. The degree of labelling and yield can also be measured by using the 'Microarray' function within the NanoDrop software v3.8.1. In addition to the 260/280nm ratio, the excitation/emission maxima, 555/565nm and 650/670nm, are measured as an indication of Alexa Flour<sup>®</sup> 3 and Alexa Flour<sup>®</sup> 5 incorporation, respectively. Successfully labelled IN samples show significantly greater 555/565nm ratio compared to 650/670nm and the opposite is true for IP samples. For a starting amount of 2µg, ≥7µg of labelled DNA is observed.

### **2.1.7.4 Array hybridisation**

Sample hybridisation was carried out as described in Agilent's ChIP-on-Chip protocol v11.2. For 2x400K arrays, 5µg of correspondingly labelled IN and IP samples are combined and made up to 79µl in ddH<sub>2</sub>O. 25µl of human Cot-1 DNA<sup>®</sup> (1mg/ml) (Life Technologies Ltd., Paisley, UK), 26µl of 10X aCGH Blocking Agent (Agilent Technologies Ltd., Cheshire, UK) and 130µl of 2X HI-RPM Hybridisation Buffer (Agilent Technologies Ltd., Cheshire, UK) are added sequentially and heated on a thermocycler at 95°C for 3mins and 37°C for 30mins. To load the samples onto the array, a matching gasket slide is placed on top of a SureHyb Microarray Hybridisation Steel Chamber (Agilent Technologies Ltd., Cheshire, UK). 250µl of the mixture is pipetted by slowly dragging across centre of the gasket slide, making sure to avoid the rubber lining.

The microarray slide is the gently lowered on top, with the Agilent barcode facing up. The chamber is sealed and placed in a 65°C Microarray Hybridisation Oven (Agilent Technologies Ltd., Cheshire, UK) at 20rpm for 40hrs.

#### **2.1.7.5 Array washing, scanning and feature extraction**

Once the hybridisation is complete, the array/gasket sandwich is first removed from the steel chamber and disassembled in Array Wash Buffer 1 (300ml of 20X SSPE, 250µl of 20% sarcosine in 1L ddH<sub>2</sub>O) using a pair of plastic tweezers. Next, the array slides are placed in a holder and moved to a new glass container with 200ml of fresh Array Wash Buffer 1 and incubated at room temperature for 5mins with gentle stirring. To prevent signal degradation, the amount of light exposure is minimised by covering the glass container with aluminium foil. The arrays are then moved to a new container for a further 5mins wash with Array Wash Buffer 2 (3ml of 20X SSPE in 1L ddH<sub>2</sub>O), pre-warmed to 31°C. Following this, the arrays are dried quickly using a Slide Spinner (Labnet International Inc., Edison, USA) and stored in an aluminium wrapped slide holder until ready for scanning. Washed arrays are scanned on Agilent type C Scanner (Agilent Technologies Ltd., Cheshire, UK) following the recommended parameters (Agilent Technologies Ltd., Cheshire, UK).

The scanned array produces a high-resolution TIFF (tagged image file format) file containing the relative intensities of IN (green emission) and IP (red emission) for each feature. This data is extracted for quantification as a numerical value (IP/IN ratio) using the Feature Extraction Software v10.7 (Agilent Technologies Ltd., Cheshire, UK).

#### **2.1.7.6 Data analysis**

Microarray dataset generated in this study are analysed using the Sandcastle R-package, developed specifically for making comparisons between multiple linked ChIP-chip experiments (Bennett et al. 2015). The main attraction of the software is the ability to compare relative changes in binding following a novel procedure, which normalises the datasets to a common background. In addition, scripts for loading, plotting, enrichment and peak detection are also presented. An example of the theory behind the normalisation process is demonstrated using ChIP-chip data generated in Chapter IV.

### 2.1.7.7 Mammalian microarray design

Microarrays used in the established TK6 cell ChIP-chip method are custom designed in Agilent's' 2x400K format (design ID: AMADID:074126) using the SureDesign database (Agilent Technologies Ltd., Cheshire, UK). These contain 2 printed microarrays per slide with 400,000 features representing the entire human chromosome 17. Such economic design provides optimal genomic coverage of an entire chromosome whilst ensuring good data resolution (81.195Mbp/414043 probes ~196bp spacing between each probe).

### 2.1.8 RDIP-Chip

R-loops in TK6 cells were examined following the published RDIP-Chip protocol with minor modifications (Nadel et al. 2015). Briefly, genomic DNA was extracted using the salting out method (Miller, Dykes, and Polesky 1988). Samples are then treated with RNase A under high salt conditions (500mM NaCl) to digest only single and double stranded RNA, as these types of RNA were previously shown to interfere with the proper detection of R-loops (Zhang Z.Z. 2015). 5µg of sonicated DNA was immunoprecipitated with 10µg of the anti-DNA:RNA hybrid [S9.6] antibody (Kerafast, Boston, USA). Subsequent steps using the IP'd DNA were performed with the established ChIP-chip workflow as described above.

## 2.2 Yeast methods

### 2.2.1 Yeast strains and culture

The *S. cerevisiae* strains BY4742 (*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) and BY4742 *rpd3Δ* (*MATa, his3Δ1, leu2Δ0, lys2Δ0, ura3Δ0, rpd3Δ::KanMX4*) were cultured in sterilised yeast extract peptone dextrose (YPD) media (1% BactoYeast extract, 2% BactoPeptone; 2% Dextrose) at 30°C, 180rpm in an Infors HT multitron incubator (Infors HT, Basel, Switzerland).

Yeast cultures were maintained at -80°C (in 15% glycerol) for long term-storage and as streaked colonies on YPD agar plates for short-term use. Pre-cultures were prepared by inoculating a single colony from agar plates and culturing overnight in 20ml

YPD. A pre-determined volume of pre-culture is added to 50 – 200ml of YPD and grown overnight to logarithmic phase ( $\sim 2 \times 10^7$  cells/ml) before treatment or harvesting.

### **2.2.2 Growth analysis**

The growth of yeast cells was monitored by measuring the optical density at 600nm using a Jenway 6305 spectrophotometer (Bibby Scientific Ltd., Staffordshire, UK). To monitor growth, yeast cells were cultured to  $OD_{600} = 0.1$  and the change in  $OD_{600}$  recorded every hour. Drug treatments were administered at the start of the exponential phase, as determined from the growth curve.

### **2.2.3 Whole cell protein extraction**

To prepare whole cell lysates, yeast cells are grown to a density of  $2 - 3 \times 10^6$  cells/ml in 50ml and harvested by centrifuging at 3000g for 5mins. Cells are washed once in PBS and re-suspended in 300 $\mu$ l yeast lysis buffer (30mM HEPES pH7.4, 50mM KCl; 10 % Glycerol). An equal volume of glass beads is then added and the tube is vortexed at maximum speed for 20mins at intervals of 1min on ice. Following this, the supernatant was separated by centrifuging at 10,000g for 15mins and stored at  $-80^\circ\text{C}$  before use. Total protein concentrations were measured as previous described using the Bradford assay.

### **2.2.4 ChIP-chip**

The yeast ChIP-chip protocol follows the same structure as the established TK6 ChIP-chip protocol, differing only in chromatin preparation, whole genome amplification and the arrays used.

#### **2.2.4.1 Chromatin preparation**

To extract chromatin from yeast, 200ml of exponentially growing cells ( $2 - 3 \times 10^6$  cells/ml) were harvested at 3000g for 5mins and cross-linked in 1% formaldehyde and phosphate buffered saline (PBS) solution for 10mins, with shaking at room temperature. A final concentration of 125mM glycine was used to stop the reaction and the cell pellet

was washed 3x in PBS before being re-suspended in 500µl of FA/SDS buffer (50mM HEPES-KOH pH7.5, 150mM NaCl, 1mM EDTA, 1% Triton-X, 0.1% Sodium deoxycholate, 0.1% SDS) containing 1X PIC and 1X PMSF (Sigma-Aldrich Company Ltd, Gillingham, UK) in a 2ml tube. Equal volume of glass beads are added and vortexed at maximum speed for 20mins at intervals of 5min on ice. To ensure the greatest recovery of chromatin, 2ml tubes were pierced at the bottom and placed in a 15ml falcon tube. The supernatant was collected by centrifuging at 2000g for 2mins.

Subsequently, whole cell chromatin was fragmented to an average size of 400 – 500bp using a Bioruptor<sup>®</sup> Standard sonication device (Diagenode, Ougrée, Belgium) for 7 cycles (10s on, 30s off) on ‘High’ settings. Chromatin was recovered by centrifuging at 13,000g for 15mins and 50µl de-crosslinked overnight at 65°C to assess the sonication efficiency by gel electrophoresis. To preserve chromatin, samples were snap-frozen using liquid nitrogen and stored at -80°C until use.

#### **2.2.4.2 Ligation-mediated PCR (LM-PCR)**

Typically, for yeast, 200 – 500µl of chromatin is sufficient for a single ChIP reaction and washes are performed in the same manner described for mammalian ChIP. The immunoprecipitated (IP) or background DNA (IN) from yeast is amplified using the ligation-mediated PCR method, exactly as described in Agilents’ ChIP-on-chip protocol v11.3 (Agilent Technologies Ltd., Cheshire, UK) and consists of (i) creating blunt ends, (ii) adaptor ligation and (iii) PCR. Briefly, the blunt end mix is added to both the IP and the diluted IN DNA separately and incubated in a thermocycler at 12°C for 20mins. The DNA is phenol/chloroform extracted once and dissolved in ddH<sub>2</sub>O. Following this, adapters are ligated to the fragmented DNA by an overnight incubation step. Two rounds of PCR are performed with the DNA and then purified to check yield. Reagents and condition are provided in Agilents’ ChIP-on-chip protocol v11.3 (Agilent Technologies Ltd., Cheshire, UK).

#### **2.2.4.3 Yeast microarray design**

Yeast microarrays are custom designed in Agilents’ 4x44K format (design ID: AMADID:038550) using the SureDesign database (Agilent Technologies Ltd., Cheshire, UK). These contain 4 printed microarrays per slide with 44,000 features representing the



entire yeast genome; with an average probe spacing of 275bp (12.1Mbp/44,000 ~275bp spacing between each probe).

### **2.2.5 DRIP-Chip**

For mapping the locations of DNA:RNA hybrids (R-loops) in the yeast genome, the DRIP-Chip protocol published by Chan and colleagues was adapted and used here (Chan et al. 2014). Primarily, DNA instead of chromatin is prepared from yeast cells, since R-loops are DNA-associated and easier to prepare compared to chromatin. Additionally, the established LM-PCR method is used to amplify ChIP'd material instead of using T7-DNA Polymerase method.

DNA was extracted from yeast cells using the 'Bust n Grab' protocol with slight modifications (Harju, Fedosyuk, and Peterson 2004). 50ml of logarithmically growing yeast cells are harvested and suspended in 2ml tubes containing 1ml lysis buffer (10mM Tris-HCl pH8.0, 1mM EDTA pH 8.0, 100mM NaCl, 1% Triton X-100, 1% SDS). The tubes are submerged in liquid nitrogen for 2mins and immediately transferred to a 95°C water bath. This procedure was repeated again before extracting the DNA with phenol/chloroform and dissolving the DNA in 50µl TE buffer. Single and double-stranded RNA were then removed by incubating with RNase A (Zhang Z.Z. 2015). Sonication was performed using 4 cycles to fragment the DNA to 400 – 500bp as with chromatin preps. Immunoprecipitation was performed by coupling 10µg of the anti-DNA:RNA hybrid [S9.6] antibody (Kerafast, Boston, USA) to 80µl of anti-mouse IgG Dynabeads® (Invitrogen™, Life Technologies Ltd., Paisley, UK) for 3 hours at 4°C. 3µg of the fragmented yeast genomic DNA was then added and incubated overnight at 4°C on a rotor. The resulting immunoprecipitated DNA was washed following the published protocol (Harju, Fedosyuk, and Peterson 2004), labelled and hybridised onto microarrays in the same way as ChIP-chip, outlined in Agilents' ChIP-on-chip protocol v11.3 (Agilent Technologies Ltd., Cheshire, UK).

### **2.2.6 Extraction of total RNA**

The hot acid phenol method was used to extract RNA from intact yeast cells. 10ml of exponentially growing yeast cells are harvested by centrifuging at 4000g for 3mins.

The cell pellet is washed once with ice-cold PBS, transferring to a 1.5ml tube. Cells are pelleted by centrifuging for 10s at 10,000g and the PBS removed. Following this, cells are re-suspended in 400µl of TES buffer (10mM Tris-HCl pH7.5, 10mM EDTA; 0.5% SDS) and 400µl of acid phenol is added and mixed by vortexing. The sample is then incubated at 65°C for 45mins. After this, the tube is placed on ice for 5mins and centrifuged at 10,000g for 5mins at 4°C. The supernatant is transferred to a new tube and phenol chloroform extraction was performed, twice using 400µl of acid phenol and once with 400µl chloroform. The top aqueous layer is transferred to a new tube and the RNA precipitated by adding 40µl of 3M sodium acetate pH5.2 and 1ml of ice-cold ethanol. The sample is centrifuged at 10,000g for 5mins at 4°C and washed once with 70% ethanol and air-dried. Finally, the pellet is re-suspended in 50µl ddH<sub>2</sub>O and the yield measured using a NanoDrop 1000 Spectrophotometer (Thermo Scientific, Hempstead, UK).

### **2.2.7 Preparation of cDNA**

Complementary DNA (cDNA) is prepared using 50ng of the extracted total RNA following protocols outlined in the iScript cDNA synthesis kit (Bio-Rad Laboratories Ltd., Hertfordshire, UK).

## **Chapter III**

# **Using *Saccharomyces cerevisiae* as a model organism to study the genotoxic effects of HDAC inhibition**



### 3.1 *Saccharomyces cerevisiae* as a model organism

Classical molecular genetics helped to introduce much of the basic concepts surrounding cellular organisms (i.e. replication origin, promoters) using simple bacterial systems such as *E. coli*. These rapidly growing prokaryotes are highly manipulable, both biochemically and genetically, making them favourable tools in recombinant DNA technology (Botstein and Fink 1988). This feature allowed researchers to associate gene structure with protein and biological function. While bacteria share many fundamental properties of other organisms (i.e. genes, mRNA; ribosomal protein synthesis) they lack features of higher order eukaryotes such as subcellular organelles, nuclear compartment and alternative splicing (Botstein and Fink 1988). Complex multicellular organisms present inherent problems such as genetic tractability, slow growth rates, large genome and cell specialisation, thereby prohibiting the direct application of classical bacterial molecular genetics (Botstein and Fink 1988). Given these differences, bacteria undoubtedly did not fit the role of a model organism for higher eukaryotes, a role, which instead could be filled by yeast.

*Saccharomyces cerevisiae*, commonly known as budding or baker's yeast is a simple eukaryotic organism carrying the same level of genetic manipulability as bacteria, but shares many similarities in its metabolism and cellular pathways as higher eukaryotes. Sequencing revealed the genome of *S. cerevisiae* to consist of approximately 12 million base pairs on 16 chromosomes coding for around 6000 genes (Goffeau A 1996).

### 3.2 Histone deacetylase enzymes in yeast

Yeast histone deacetylases consists of the Zn<sup>2+</sup> dependent (Rpd3, Hda1; Hos1-3) and the NAD<sup>+</sup> dependent (Sir2, Hst1-4) family of enzymes (Grunstein 2003). Microarray studies examining the effect of HDAC mutants on global histone acetylation identified Rpd3 and Hda1 as the primary HDACs in yeast, responsible for regulating numerous promoters throughout the genome (Robyr et al. 2002). Hos1 and Hos3, on the other hand, are largely responsible for deacetylating the locus of ribosomal DNA (Robyr et al. 2002), while Hos2 associated binding, in contrast to the actions of other HDACs, is required for gene activation (Wang, Kurdistani, and Grunstein 2002). The SIR silencing complex is responsible for reinforcing heterochromatin states around telomeric regions and the silent-mating type loci in yeast (Rusche, Kirchmaier, and Rine 2002). Sir2, a component

of this complex, deacetylates H4K16 to facilitate Sir3-dependent recruitment of Sir4 (Rusche, Kirchmaier, and Rine 2002). Sir4 further promotes the binding of Sir2 and in turn forms a positive feedback loop repressing regional transcription (Rusche, Kirchmaier, and Rine 2002). Particularly, *sir2* mutants displayed a loss in hypoacetylation in regions near the telomere compared to wild type (Kimura, Umehara, and Horikoshi 2002).

### **3.3 Use of HDAC inhibitors in yeast**

Histone deacetylase inhibitors have been used in yeast to study the effect on transcription, demonstrating similar changes in gene expression profile as the HDAC mutant *rpd3Δ* (Bernstein, Tong, and Schreiber 2000). The hydroxamate Trichostatin A (TSA) has been shown to alter transcription in yeast in a concentration dependent manner from 10nM to 31.25μM (Bernstein, Tong, and Schreiber 2000; Wan et al. 2011), while members of the short-chain fatty acids, such as sodium butyrate and valproic acid, have previously been shown to disrupt cell growth at concentrations of 60mM and 100mM, respectively (Nguyen et al. 2011; Mitsui et al. 2005). The demonstrated effect of HDAC inhibitors in yeast should allow for comparable observations to be made in this study.

### **3.4 Chapter aims**

With the ChIP-chip method previously established in yeast (Teng et al. 2011), the aim of this chapter is to use a genome-wide approach to model the DNA damaging effects of HDAC inhibitors. To achieve this, an HDAC inhibitor will be selected for use which demonstrates an effect at the epigenetic level in yeast cells. The effect of the drug on genome-wide acetylation along with markers of transcription and replication as well as phospho-H2A, a marker of DNA damage, will be examined. Comparison of these results are expected to provide further insights into whether HDAC inhibition may lead to an increase in the frequency of transcription-replication conflicts and DNA damage, central to the hypothesis presented in this thesis. Results obtained in the yeast model are aimed at providing a greater insight into the possible actions of HDAC inhibitors in the ensuing humans studies.

### **3.5 Effect of HDAC inhibitors on the growth of yeast cells**

Studies in the literature have reported varying levels of success in demonstrating the effects of HDAC inhibitors in yeast cells (Carmen, Rundlett, and Grunstein 1996; Alonso and Nelson 1986; Vu Q.A 1987). Therefore, the initial aim is to utilise an HDAC inhibitor at concentrations previously reported which showed an effect in yeast (as described above in section 3.3). As such, the exact type or class of pan-inhibitor used here is unimportant.

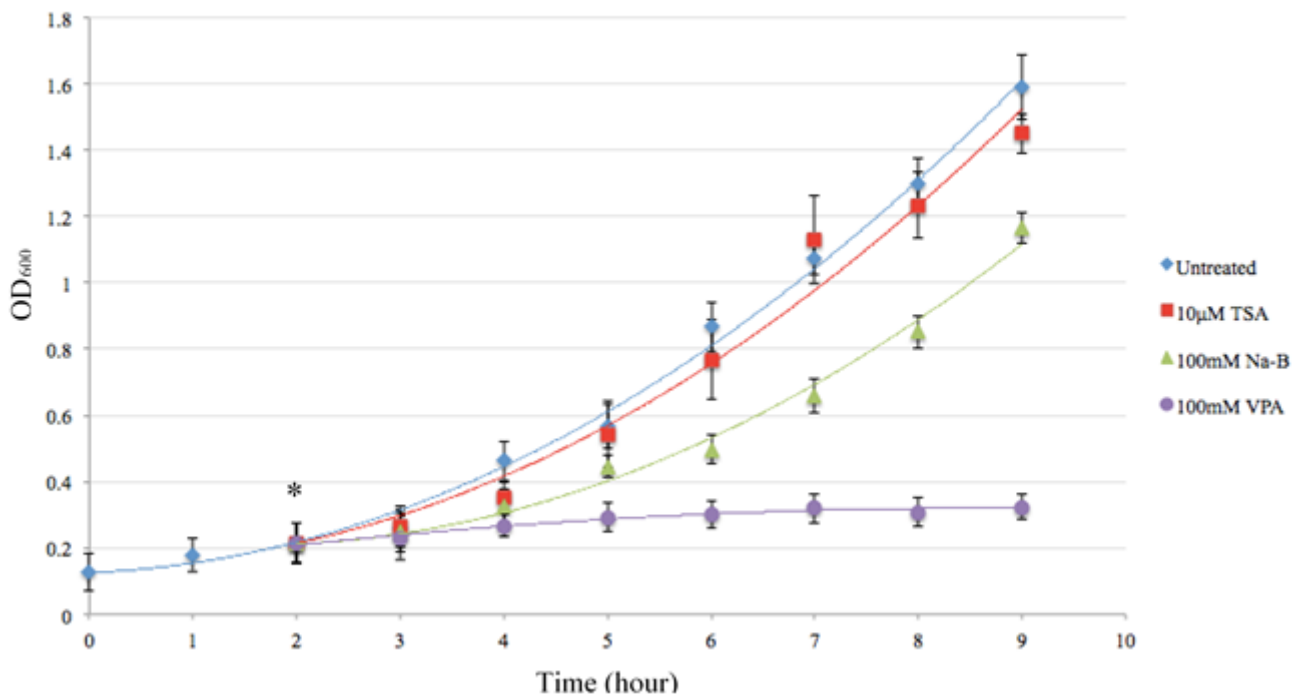
To first examine the response of *S. cerevisiae* BY4742 (budding yeast) strain to the various HDAC inhibitors, the effect on cell growth was measured as an indicator of general toxicity. HDAC inhibitor compounds were selected based on their use in literature. Concentrations used here were the highest reported for each compound where possible, to ensure that an effect could be observed. By comparison, HDAC inhibitor concentrations in yeast are several magnitudes higher than those used in mammalian cell lines, possibly to overcome barriers to uptake presented by the cell wall. Wild-type cells were independently treated with HDAC inhibitors at the beginning of the exponential growth phase (Figure 3.1, asterisk) and changes in optical density ( $OD_{600}$ ) were monitored every hour for 7 hours.

Following treatment, 100mM Valproic acid (VPA) completely inhibited cell growth. Sodium butyrate (Na-B) treatment has an intermediate effect on cell growth at 100mM (Figure 3.1). Trichostatin A (TSA) on the other hand, showed less of an inhibitory effect on cell growth, although the concentration used here is significantly lower in comparison with the other inhibitors (Figure 3.1). Together these results show that treating yeast cells with a range of HDAC inhibitors impedes cell growth.

### **3.6 Effect of HDAC inhibition on histone acetylation**

Having demonstrated the cytotoxicity of HDAC inhibitors in BY4742 cells, the effect on histone acetylation was examined. The yeast histone deacetylases *RPD3* and *HDA1* are Trichostatin A-sensitive and deletion of these HDACs leads to global and site-specific increase in histone H3 and H4 acetylation. Therefore, similar observations are expected when treating yeast cells with HDAC inhibitors. To test this, cells were treated with the HDAC inhibitors under the conditions previously used in the growth assay (Figure 3.1). Independent cultures of log phase cells ( $OD_{600} = 0.4$ ) were incubated in the

Effect of HDAC inhibitors on *S. cerevisiae* growth rate

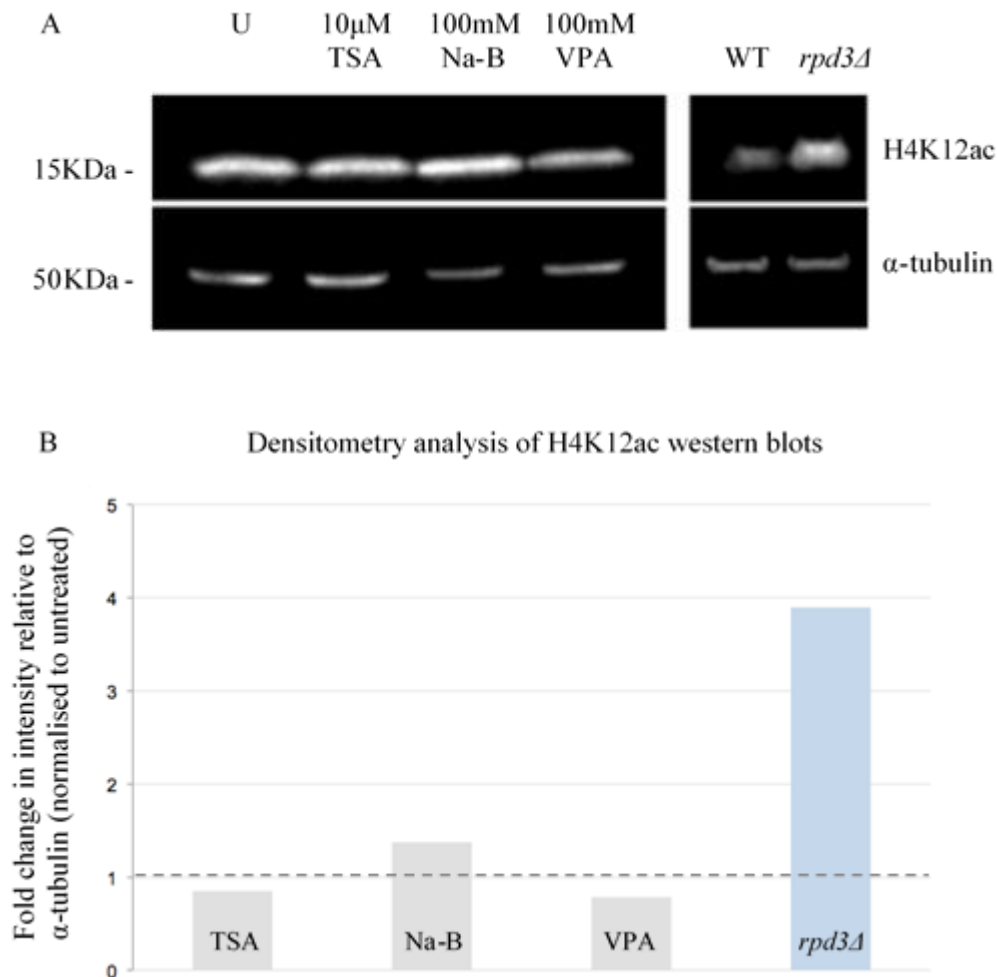


**Figure 3.1. HDAC inhibitors inhibit the growth of yeast cells.** Cells entering the exponential growth phase (\*) were treated with the selected HDAC inhibitors and OD<sub>600</sub> measurements taken every hour are plotted over the course of 7 hours. Untreated (*blue*), Red –TSA (*red*) – Trichostatin A, Na-B (*green*) – sodium butyrate; VPA (*purple*) – valproic acid. Error bars as standard error of the mean, n = 3. (Raw data available in Appendix, I).



presence of each HDAC inhibitor for 1h and whole cell lysates were prepared for western blotting. The effect of HDAC inhibitor treatment on global H4K12ac was measured and normalised against  $\alpha$ -tubulin using densitometry (Figure 3.2B). In comparison to the positive *rpd3 $\Delta$*  control, which showed a 4-fold higher H4K12ac than the wild type, treatment with HDAC inhibitors did not show any significant effect on global H4K12ac compared to *rpd3 $\Delta$*  (Figure 3.2B).

Given that the level of H4K12ac in *rpd3 $\Delta$*  was only 4-fold higher over the wild type, it is conceivable that HDAC inhibitors produced a more subtle effect on histone acetylation in comparison to the HDAC mutant. To examine this, changes in histone acetylation at specific loci sensitive to either *RPD3* or *HDA1* deletion was measured using quantitative-PCR (qPCR) in response to the HDAC inhibitor sodium butyrate. The ability of sodium butyrate to inhibit *RPD3* and *HDA1* was assessed by monitoring increases in H4K12 and H3K9 acetylation at the *INO1* and *GALI* gene, respectively. To do this, chromatin was first prepared by harvesting and fixing exponentially growing yeast cells in formaldehyde. Fixed cells were then lysed in buffer and chromatin was fragmented to an average size of 500bp using an ultrasound sonicator (Chapter II, 2.2.5.1). The protein-DNA crosslinks were reversed from the DNA by de-crosslinking overnight at 65°C and the protein removed by addition of pronase (Sigma-Aldrich Company Ltd, Gillingham, UK). The extent of fragmentation was checked on an agarose gel following electrophoresis of the purified DNA sample (Chapter II, 2.1.6.1). Of the different cycles tested, 4 x 30s alternating on/off at 'High' settings was determined as the optimal cycle number to achieve the desired fragment size of 500bp, which is required for efficient immunoprecipitation (Figure 3.3). Next, 0 - 5 $\mu$ g of the anti-H4K12ac and anti-H3K9ac antibody was titrated with 200 $\mu$ l of sonicated chromatin. The right arm telomeric region of chromosome VI (TEL06R) was used as a negative control locus, since these regions are hypoacetylated in wild type yeast (Suka et al. 2001). The ability of each antibody to enrich for its target at the positive loci (anti-H4K12ac at *INO1*; anti-H3K9ac at *GALI*) compared to the negative loci (TEL06R) demonstrates its suitability to distinguish the acetylation status of these loci (Figure 3.4A). From the titration experiment, 2 $\mu$ g was identified as the optimal amount of anti-H4K12ac, showing 15-fold higher enrichment at *INO1* locus compared to TEL06R (Figure 3.4A). Likewise, 3 $\mu$ g was identified as the optimal amount of anti-H3K9ac, showing 35-fold higher enrichment at *GALI* locus compared to TEL06R (Figure 3.4A). Having prepared chromatin from wild type yeast and shown that both antibodies work for ChIP, the effect of sodium butyrate treatment



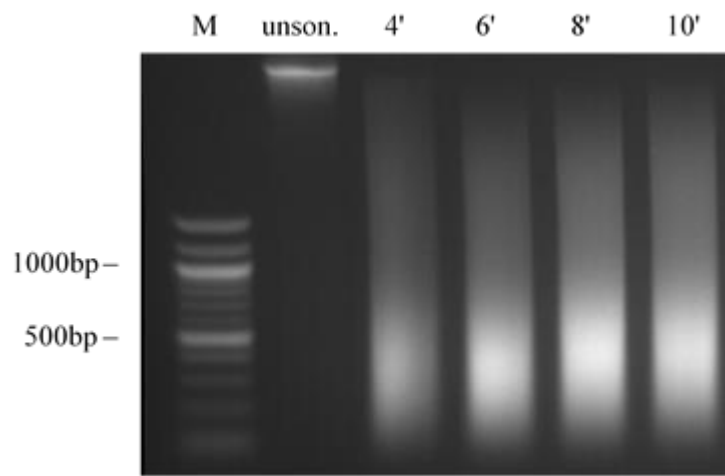
**Figure 3.2. Treating yeast cells with HDAC inhibitors has no effect on global H4K12 acetylation.** **A:** changes in global H4K12 acetylation was examined by probing whole cell extracts of untreated, *rpd3 $\Delta$*  and HDAC inhibitor treated cells with anti-H4K12ac antibody. Detection of  $\alpha$ -tubulin was carried out as a normalising control. **B:** fold-change in H4K12ac signal calculated as a ratio over  $\alpha$ -tubulin and normalised to the untreated sample. TSA – Trichostatin A, Na-B – sodium butyrate; VPA – valproic acid. Band intensities are quantified using densitometry (Image J). n = 2. (Raw data available in Appendix, I).

was then examined. Figure 3.4B shows that in response to 100mM sodium butyrate no increase in H4K12ac at the *INO1* gene and H3K9ac at the *GALI* gene is observed, relative to the untreated cells. Together with the results of the western blot analysis this indicates that the HDAC inhibitors tested have no significant effect on histone H4K12 and H3K9 acetylation.

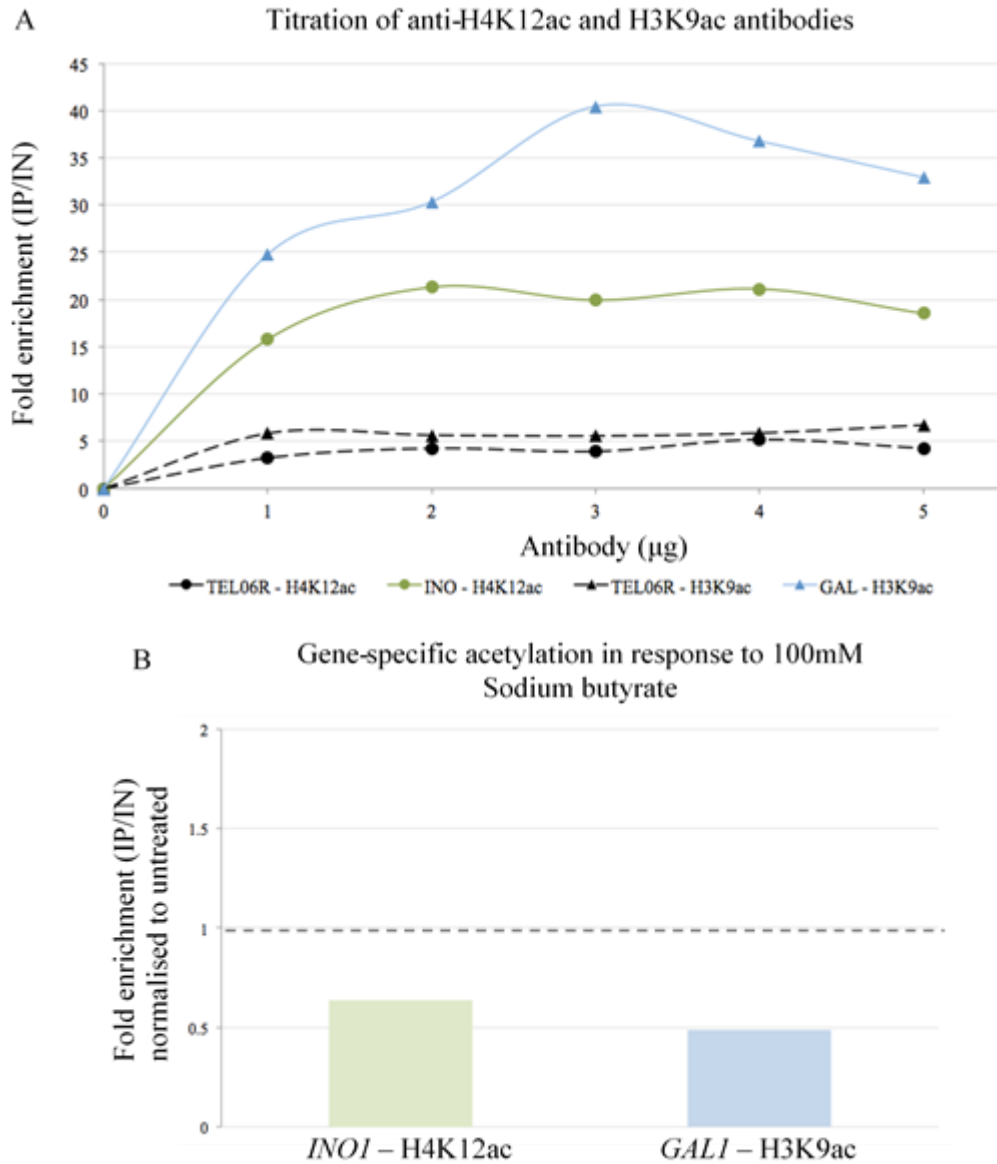
### 3.7 The effects of HDAC inhibition on gene expression

Another major hallmark of HDAC inhibitors is their ability to induce changes in gene expression due to changing histone acetylation. While treatment with an array of HDAC inhibitors appeared to have no significant effect on histone H3K9 and H4K12 acetylation levels, efforts to further examine the effect on gene expression were made by attempting to reproduce the results of a more recently published finding. Wan *et al.* used DNA microarrays to identify global changes in transcription in response to TSA in a yeast strain of the same genetic background (BY4742). This effectively allows relevant comparisons to be made with the current study and for this reason TSA will be used here onwards. Select genes showing varying degrees of fold change were chosen from the results of the microarray study conducted by Wan *et al.* as the objective of this study (*FRE1*, *FRE7* and *CTR1*) and normalised to the same reference gene (*ACT1*). Additionally, the experiment was performed exactly as described, using the same primer sequences and treatment regimen (10 $\mu$ M TSA, 1h in YPD under standard yeast culture conditions) (Wan et al. 2011).

In order to measure changes in these gene transcripts, total RNA was first isolated from log phase cells treated with 10 $\mu$ M TSA for 1h. As an initial assessment of yield and purity, the absorbance of samples was measured on a NanoDrop spectrophotometer (Thermo Scientific, Hempstead, UK) at 230, 260 and 280nm. Nucleotides absorb at 260nm, while contaminating protein and carbohydrates absorb at 230nm and 280nm, respectively. Therefore, the ratios of these are commonly used to assess sample purity. In addition, uracil has a higher absorbance at 260nm compared to thymine and so an  $A_{260/280}$  ratio of  $\sim 2$  is generally considered pure for RNA, whereas a ratio of  $\sim 1.8$  is considered pure for DNA. The  $A_{260/230}$  is also expected to be higher than the  $A_{260/280}$  for samples free of residual contamination. Moreover, the integrity of extracted RNA in the sample is indicated by the presence of two distinct 26S and 18S ribosomal RNA bands when examined by denaturing gel electrophoresis. Total RNA was extracted from yeast cells



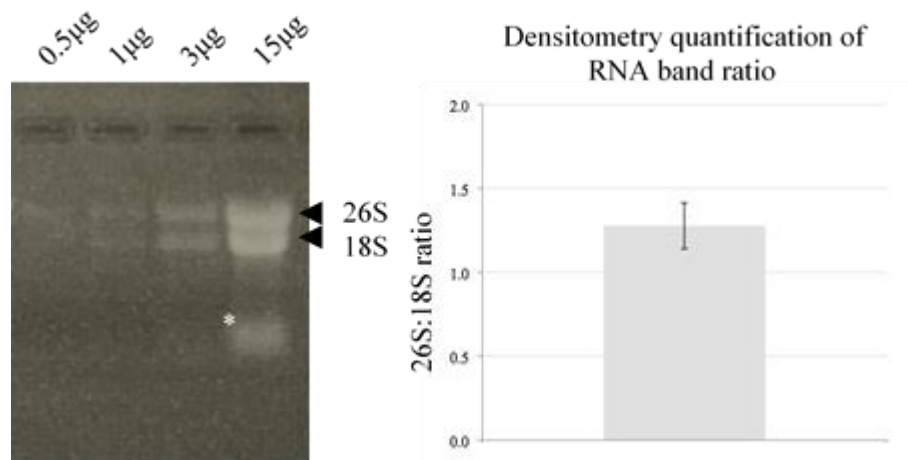
**Figure 3.3. Sonicated whole cell chromatin prepared from wild type yeast cells.** Whole cell chromatin extracted from wild type yeast cells are fragmented using an ultrasound sonicator at 4, 6, 8 and 10 cycles of 30s on/off. Samples were de-crosslinked and DNA resolved by gel electrophoresis showing the fragment sizes achieved at each cycle. unson – unsonicated chromatin. Representative image shown.



**Figure 3.4. Treating yeast cells with sodium butyrate does not lead to an increase in histone acetylation at *INO1* and *GAL1* genes.** **A:** titration using 0 - 5µg of anti-H4K12ac and anti-H3K9ac antibodies, showing enrichment at positive loci, *INO1* and *GAL1*, relative to the negative control loci, TEL06R. n = 1. **B:** no hyperacetylation detected at predicted loci in response to 100mM sodium butyrate. n = 2. (Raw data available in Appendix, I).

using the hot acid phenol method (Chapter II, 2.2.7), demonstrating high yield, an  $A_{260/280}$  above 2 (Table 3.1) and intact RNA bands (Figure 3.5). Apart from the presence of these bands, the ratio (derived from the nucleic acid size) is also commonly used as an indicator of integrity. The theoretical ratio between 26S (~3.8kbp) and 18S (~2kbp) is ~1.9 (3.8/2) (Agilent Application Note 2005). To investigate this, different amounts of RNA were loaded on a denaturing gel and the intensities of the two bands clearly observed bands at 3µg was selected for quantification (Figure 3.5). The average 26S:18S ratio quantified from 3 independent extractions was estimated to be ~1.3, slightly lower than the expected ratio (Figure 3.5). A smaller band/smear is visible below the 18S band when larger amounts of the sample are loaded (Figure 3.5), possibly representing smaller RNA species or degraded RNA. Though, a lack of smearing between the 26S and 18S bands indicates an overall high RNA integrity within the extract (Figure 3.5). Together these results show that RNA of high quality can be effectively isolated from yeast cells which can be used subsequently to generate complementary DNA (cDNA).

In order to perform reverse transcription quantitative PCR (RT-qPCR) analysis of these transcripts by exploiting the ability of DNA polymerase to exponentially amplify DNA, mRNA is first converted into cDNA through the process of reverse transcription (Chapter II, 2.2.8). During this process, multiple copies of cDNA are produced from the mRNA templates, yielding ~900ng/µl of cDNA from a starting concentration of ~25ng/µl total RNA (Table 3.1). Moreover, a shift in the  $A_{260/280}$  ratio from ~2 to ~1.8 indicated the presence of mostly DNA in the sample. Next, the optimal annealing/extension temperature for each set of primers (*FRE1*, *FRE7*, *CTR1* and *ACT1*) was calculated by running a temperature gradient (50 - 65°C) and melt curves were assessed for any non-specific amplification. The efficiency of the qPCR reaction was calculated to be > 90% for all primers across a template concentration ranging from 16pg - 10ng. The relative abundance of each transcript was then measured using RT-qPCR and normalised to *ACT1* (Chapter II, 2.1.6.4). Following treatment with 10µM TSA, the mRNA levels of *FRE1*, *FRE7* and *CTR1* showed no significant change relative to *ACT1* (Figure 3.6B). These results are also not comparable with the findings reported by Wan *et al.*, where *FRE1* showed a 2.5-fold decrease in mRNA levels after TSA treatment (Figure 3.6A). To investigate if a higher dose may lead to greater changes in transcript levels, cells were treated with 50µM TSA. However, increasing the TSA dose by 5-fold also did not result in a further decrease in transcript levels (Figure 3.6B). Therefore, the experiment conducted here to measure and compare the effect of treating yeast cells with TSA on the



**Figure 3.5. Successful extraction of total RNA from yeast.** Successful isolation of total RNA using the hot acid phenol method. Denaturing gel electrophoresis of extracted total RNA, showing the expected 26S (~3.8kbp) and 18S (~2kbp) yeast ribosomal bands. The average 26S:18S ratio estimated at 1.3 using densitometry for 3 biological replicates.

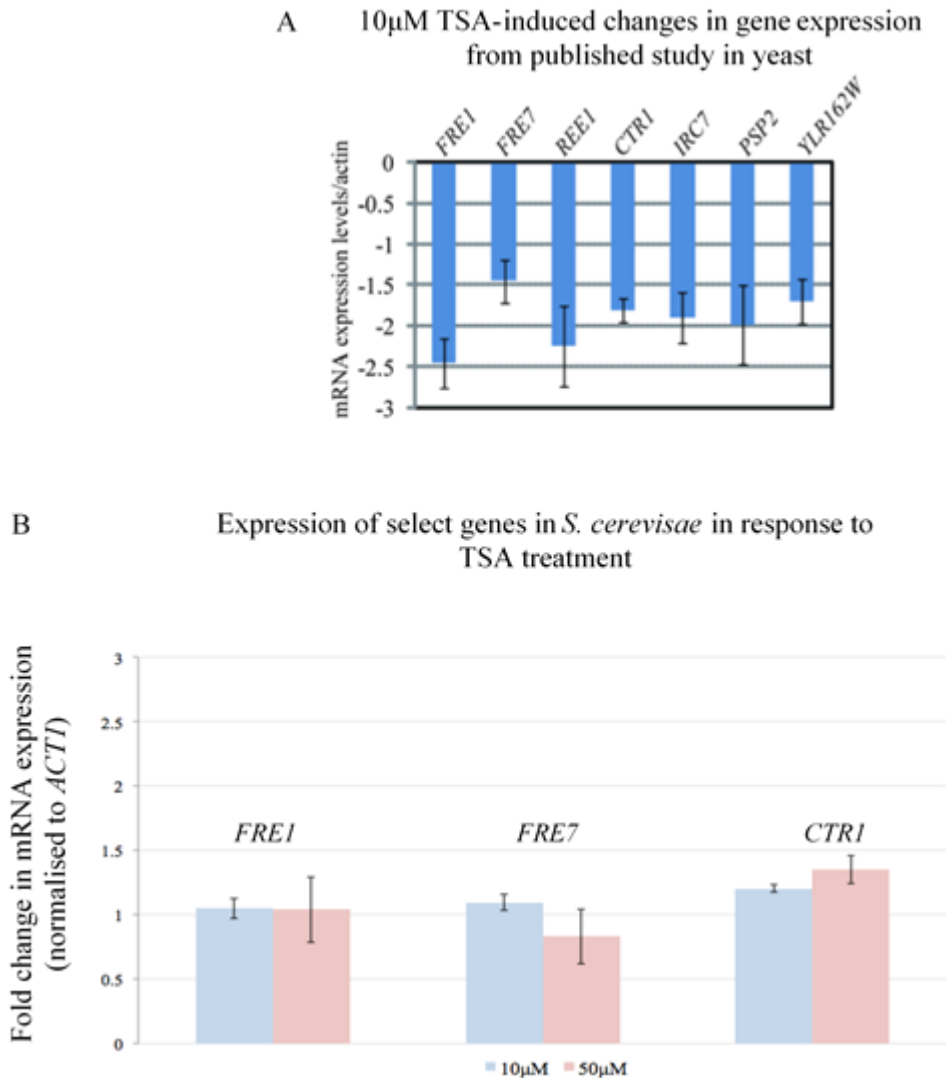
**Pre-RT:**

Sample	ng/ $\mu$ l	A <sub>260</sub>	A <sub>280</sub>	A <sub>260/280</sub>	A <sub>260/230</sub>
Vehicle control (DMSO)	402.29	10.057	4.652	2.16	2.45
10 $\mu$ M TSA	389.28	9.732	4.544	2.14	2.42

**Post-RT:**

Sample	ng/ $\mu$ l	A <sub>260</sub>	A <sub>280</sub>	A <sub>260/280</sub>	A <sub>260/230</sub>
Vehicle control (DMSO)	904.54	22.614	12.315	1.84	2.44
10 $\mu$ M TSA	896.94	22.423	12.318	1.82	2.42

**Table 3.1. Spectrophotometer sample readings before and after reverse-transcription.** *Pre-RT*: RNA extracts prepared using hot acid phenol method for untreated and 10 $\mu$ M TSA treated cells. A<sub>260/280</sub> ~2 indicates the presence of RNA. *Post-RT*: spectrophotometer readings following reverse transcription using 500ng total RNA shows an A<sub>260/280</sub> ~1.8 indicating the successful conversion of RNA to cDNA.



**Figure 3.6. Effect of TSA on the transcription of select genes is inconsistent with published findings.** **A:** Yeast genes showing the greatest change in mRNA expression in response to 10 $\mu$ M TSA from the microarray study as validated by Wan *et al* (Wan *et al.* 2011). **B:** Changes in *FRE1*, *FRE7* and *CTR1* mRNA, normalised to *ACT1* mRNA, in response to 10 $\mu$ M TSA measured in the current study using methods described by Wan *et al.* Additionally 50 $\mu$ M TSA was also examined. Error bars as standard error of the mean, n = 3. (Raw data available in Appendix, I).



expression of certain genes reported in the literature did not find evidence for any significant changes to gene expression. Overall, these results show that in the current study, HDAC inhibitors have no measurable effect on chromatin-associated events, thereby lacking evidence to support their use in the yeast model.

### **3.8 Discussion**

With the aim of trying to understand how HDAC inhibitors induced DNA damage, the initial aim was to conduct genome-wide measurements of its effects on histone acetylation, transcription, replication and the DNA damage response in the yeast model organism.

A panel of HDAC inhibitors was selected based on published work citing their use in yeast to study their potential to illicit DNA damage. Short-chain fatty acids such as sodium-butyrate and valproic acid both inhibit the growth of yeast cells (Nguyen et al. 2011), while the hydroxamic acid Trichostatin A has been shown to affect mRNA expression in yeast (Wan et al. 2011). To demonstrate the effects of selected HDAC inhibitors in the commonly used BY4742 yeast strain, their ability to affect normal cell growth was measured as an indicator of general cytotoxicity (Figure 3.1). The ability of selected HDAC inhibitors to inhibit cells growth shown here indicates that these compounds are cytotoxic in yeast and is consistent with studies done in human cell lines where they are described as potent cytotoxic agents (Bose, Dai, and Grant 2014). The yeast cell wall has been known to pose a barrier towards entry for many chemicals, given the initial observations of cytotoxicity this does not appear to be a problem for the HDAC inhibitors tested here.

In the main interest of trying to understand how epigenetic changes can lead to genetic damage, the effect of HDAC inhibition on histone acetylation was investigated. Both *in vitro* and *in vivo* studies have shown that a major hallmark of HDAC inhibitors is their ability to elicit hyperacetylation of lysine residues on histone tails (Bose, Dai, and Grant 2014). The yeast HDACs Rpd3 and Hda1 are described as Trichostatin A-sensitive and *RPD3* mutants show the greatest increase in acetylation at H4K5 and K12. Therefore, HDAC inhibitor treatment is expected to inhibit the activities of both Rpd3 and Hda1, leading to an increase in histone acetylation. However, treating yeast cells with HDAC inhibitors did not show a higher level of global H4K12 acetylation measured by western blotting (Figure 3.2). It is possible that the effects of drug treatment are much subtler

compared to the *RPD3Δ* strain. Therefore a more sensitive approach was taken, using chromatin immunoprecipitation (ChIP) and qPCR, to study the *INO1* gene, which is hyperacetylated at H4K12 in *RPD3* mutants (Rundlett et al. 1998) and similarly *GALI*, hyperacetylated at H3K9 in *HDA1* mutants (Wu J 2001). ChIP experiments using the anti-H4K12ac and anti-H3K9ac antibodies (Figure 3.3) showed enrichment at the expected *INO1* and *GALI* genes, respectively (positive control loci), compared to the telomeric region (TEL06R), which is hypoacetylated in wild type cells (negative control loci) (Suka et al. 2001) (Figure 3.4A). However, comparing H4K912ac at *INO1* and H3K9ac at *GALI* between untreated and 100mM sodium butyrate treated cells showed no increase in acetylation at these sites for two biological repeats (Figure 3.4B). In both yeast and human cells, HDAC mutants and HDAC inhibitor treatment share common features such as slower/inhibited growth and global increase in histone acetylation. While HDAC inhibition impeded yeast cell growth (Figure 3.1), an increase in histone acetylation was not observed (Figure 3.2 and Figure 3.4). It is unclear to what extent, if at all, Rpd3 and Hda1 are inhibited by the concentration of drugs used here since the affinities of these drugs for yeast HDAC enzymes have not been reported. Histone hyperacetylation is a well-known phenomenon linked to HDAC inhibition and has been described in fission yeast treated with Trichostatin A (Ekwall et al. 1997). It is possible that under the treatment conditions presented here, increases in global histone acetylation are subtle and difficult to reproduce when western blots are analysed by densitometry. While ChIP is a more sensitive technique by comparison, histone acetylation at *INO1* and *GALI* has not been studied previously in response to sodium butyrate and may therefore be different from that observed in *RPD3* and *HDA1* mutants. Reports also cite the induction of histone acetylation when spheroplasts (yeast cells with the cell wall removed), are treated with 10µM TSA (Carmen, Rundlett, and Grunstein 1996). It is unclear if removal of the yeast cell wall was required for detecting the chromatin-associated effects of TSA in this case, but generating spheroplasts severely compromises cell integrity and was therefore not performed here in this study. Interestingly, another study did not find any changes in histone acetylation when 2µM TSA was added to whole cell budding yeast, isolated yeast nuclei, spheroplasts or even crude cell extracts (Alonso and Nelson 1986). An increase in histone acetylation was only noted when partially purified histone deacetylases were treated with TSA (Alonso and Nelson 1986). The authors suggest that the organisation of HDACs in multi-protein complexes may obscure the drug-binding site and only revealed following purification (Alonso and Nelson 1986). Similar studies also support this

showing that the short chain fatty acid *n*-butyrate has no effect on deacetylase activity in yeast *in vitro* (Vu Q.A 1987). It is possible that such contradictory observations stem from a combination of using different yeast strains as either whole cells or manipulated in some way. Additionally, the techniques used to detect changes in response to HDAC inhibition vary, ranging from radiolabelling of HDAC substrates to immunostaining.

In light of this, focus was turned towards a more recent study showing changes in the transcriptomic profile elicited by Trichostatin A (Wan et al. 2011). Since the experiment was conducted using the same yeast strain (BY4742) the results of their study would be comparable. The experiment outlined by Wan *et al.* was repeated as described and the expression of *FRE1*, *FRE7*, *CTR1* and *ACT1* mRNA in response to 10 $\mu$ M TSA was selected for analysis in this study (Figure 3.6A) (Wan et al. 2011). Following the extraction of RNA (Figure 3.5) and conversion to cDNA (Table 3.1), qPCR analysis showed that the results of the published data could not be reproduced here in this study (Figure 3.6). When normalised to the same *ACT1* control the levels of *FRE1*, *FRE7* and *CTR1* did not show a considerable change (Figure 3.6B). Additionally, increasing the dose of TSA to 50 $\mu$ M also had no effect on the expression of these genes (Figure 3.6B). Interestingly, genes selected for qPCR validation in the published study represented those that displayed the greatest change in response to TSA from microarray analysis (Wan et al. 2011). Amongst these *FRE1* showed a 2.7-fold decrease in mRNA relative to *ACT1* and was the greatest difference reported (Figure 3.6A), suggesting that in their studies the effects of 10 $\mu$ M TSA on mRNA expression is very subtle. Arguably, for such subtle changes in expression, genes measured here can be normalised to the geometric means of several internal control genes, which may improve the sensitivity (Carmen, Rundlett, and Grunstein 1996). However, collectively the data presented here suggests that while HDAC inhibition can inhibit the growth of budding yeast, no significant changes in global or gene specific histone acetylation is observed. Additionally, changes in mRNA expression presented in published data could not be reproduced. Therefore, under the conditions of the current study, of the drugs tested, it is possible that HDAC inhibitor-induced cytotoxicity is a result of its actions in the cytoplasm and that these compounds are unable to exert its effects on chromatin.

Even though chemical means of inhibiting HDAC enzymes proved unsuccessful, the fact that *rpd3 $\Delta$*  cells show a higher level of histone acetylation compared to the wild type, allows for an alternative genetic approach to be taken using yeast to address some of the issues raised in human studies (Chapter V).



## **Chapter IV**

**Establishing the ChIP-chip methodology for TK6 cells  
as a genome-wide assay to measure the genotoxic  
effects of HDAC inhibitors**



#### **4.1 TK6 as a model cell line**

As chemical inhibition of HDACs in yeast proved technically challenging, parallel work undertaken to establish the ChIP-chip method in TK6 cells is reported in this chapter. The aims are to set the foundations for examining the mechanisms of HDAC inhibitor-induced DNA damage in mammalian cells by focusing on optimising and validating the ChIP-chip assay for the human TK6 B lymphoblastoid cell line. Given the role played by HDACs in regulating chromatin architecture, it is likely that changes in cellular processes that rely heavily on histone acetylation, may be the cause of genome instability. Therefore, to question how epigenetic mechanisms can affect these changes in humans, existing ChIP-chips developed in the lab for yeast needs to be established in a relevant human cell line.

While many different types of cell lines exist, for the purposes of examining the genotoxic effects of HDAC inhibitors, it is important to use cells capable of displaying a normal DNA damage response through the expression of a fully functioning p53 pathway. Surmounting evidence has also stressed the importance of using p53 competent cells to reduce ‘false positive’ results in genotoxicity testing (Fowler et al. 2012). The immortalised human TK6 B lymphoblastoid cell line stably expresses wild-type p53 (Xia et al. 1995) and was originally derived from Epstein-Bar transformed WI-L2 spleen lymphoblasts (Levy, Virolain.M, and Defendi 1968).

Further to this, the pan-HDAC inhibitor Trichostatin A has demonstrated genotoxic effects in TK6 cells (Olaharski et al. 2006). These include: (i) early induction of  $\gamma$ H2A, (ii) increase in chromosome breaks and (iii) micronuclei formation (Olaharski et al. 2006). Such effects are widely described in other cell lines and are characteristic of HDAC inhibitors (Yoo and Lee 2005). Together with the described p53-proficiency and its application in genotoxicity testing, these features make TK6 cells an ideal candidate cell line for use in this study.

## 4.2 Use of small molecule inhibitors to determine the status of *TP53*

The tumour suppressor protein p53 is a crucial determinant of genetic integrity, preventing the proliferation of cells that may be carrying mutations as a result of DNA damage and stress. Under physiological conditions, p53 is constantly expressed in cells, which consequently upregulates the E3 ubiquitin ligase Mdm2 (mouse double minute 2 homologue) (Momand et al. 1992). Mdm2, in complex with the transcriptional co-activator p300/CBP, polyubiquitinates p53 for proteasomal degradation (Momand et al. 1992). Therefore, p53 is very unstable in unstressed cells owing to the negative autoregulatory mechanism. Cells harbouring mutant p53 lose the ability to maintain the feedback loop and as a result display high levels of inactive p53. Small molecule inhibitors of Mdm2 such as nutlin-3 were originally developed to reactivate p53 in human cancers still retaining wild type p53 (Shangary and Wang 2009). Its inhibitory effect leads to apoptosis and is exclusively dependent on the presence of wild type p53. Based on this, clinical proposals have been put forward for using nutlin-3 as a fast and efficient means of screening cells with *TP53* mutations, evaluating the induction of apoptosis as the end point (Pozzo et al. 2013). Therefore, this method can similarly be used in this study to validate the *TP53* status of TK6 cells.

## 4.3 Challenges of mammalian ChIP-chip

ChIP-chip has been utilised in a variety of different organisms to estimate the genome-wide locations of chromatin-associated events. While general methods have been published (Lee, Johnstone, and Young 2006), these are only sufficient to describe certain protocols. The differences in cell type and sonication equipment used, requires that optimal conditions are established for the specific cell line used. In addition, this process can be particularly challenging for mammalian cells due to the complexity of the human genome (Lee, Johnstone, and Young 2006). Several key steps such as crosslinking, sonication and antibody washes require most attention, as these variables tend to be crucial for the experiment to succeed. Protein-DNA cross-linking using formaldehyde (FA) for too long can result in protein aggregation and denaturation while insufficient fixing can lead to inefficient protein capture (Lee, Johnstone, and Young 2006). Arguably, the fixation step can be avoided when studying histone modifications, since histone proteins are tightly bound to DNA. However, it is recommended to introduce FA cross-linking as



some genomic regions can undergo nucleosome rearrangements during chromatin preparation in the absence of FA. For microarray analyses, chromatin should ideally be fragmented to a range of 200-1000bp with an average of 500bp, to provide the optimal signal resolution. Sonication greatly depends on the type of equipment, cell type, cell density, chromatin concentration and volume. Under-sonicated chromatin consists of large DNA fragments which do not hybridise well to the array, on the other hand, over-sonicating chromatin can damage epitopes. Following immunoprecipitation, washing the superparamagnetic bead-bound antibody is necessary to remove non-specific binding, though the number of washes and salt concentration used is entirely dependent on the antibody. The success of a ChIP assay is frequently determined by measuring enrichment at specific gene(s) where the target of interest is known to be bound (positive control loci) along with a region where it's not bound (negative control loci) using qPCR. Consequently, troubleshooting a failed ChIP experiment can be challenging due to a lack of endpoints and contributions from multiple factors.

#### **4.4 Whole human chromosome 17 microarrays**

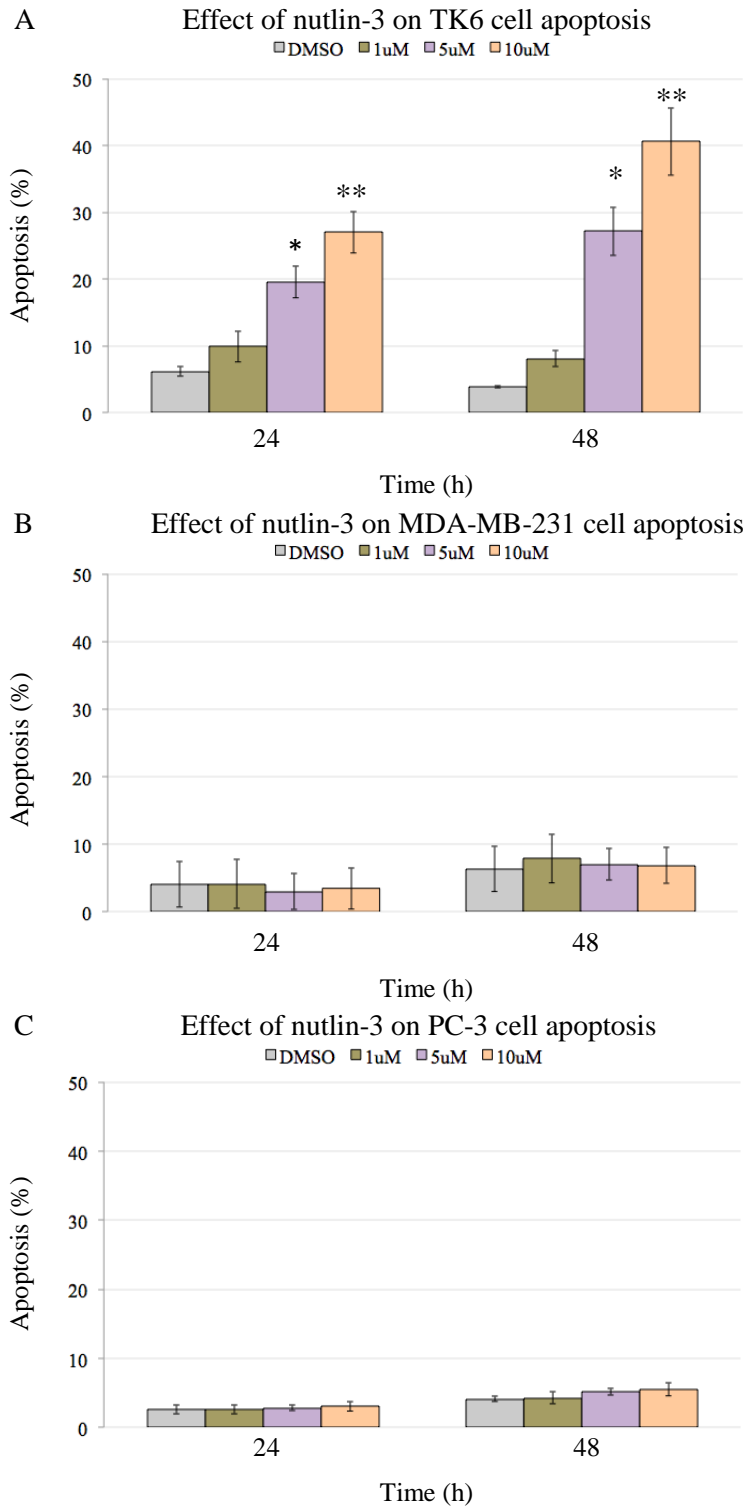
With the intention of measuring biological events such as histone acetylation, transcriptional elongation and replication that span large regions of chromatin, the use of arrays with features representing continuous sections of the genome is preferred. In addition, the Agilent array platform has proven performance in past studies conducted in yeast (Powell 2014b) and forms a seamless workflow with the 'Sandcastle' R-package. For these reasons, Agilent SurePrint G3 2x400K arrays were custom designed to cover whole human chromosome 17 (hg19) (Chapter II, 2.1.7.7). The inclusion of *TP53*, located on the short arm of chromosome 17 (17p13.1), would also serve as a useful control gene in this context. Juan and colleagues presented evidence for the involvement of HDAC enzymes in suppressing p53 transcription (Juan et al. 2000). Later studies conducted by Sonnemann and others confirmed this by detecting an increase in p53 expression, normally triggered in response to DNA and/or cellular stress, in the presence of the HDAC inhibitor Vorinostat (Sonnemann et al. 2014), which is a synthetic TSA derivative. Therefore, based on the expected increase at *TP53* following HDAC inhibition and the 'known' pattern of RNA Pol II (pSer2) occupancy (Hsin JP 2012), mapping transcription across chromosome 17 can be used as a means of validating the ChIP-chip assay in TK6 cells.

## **4.5 Chapter aims**

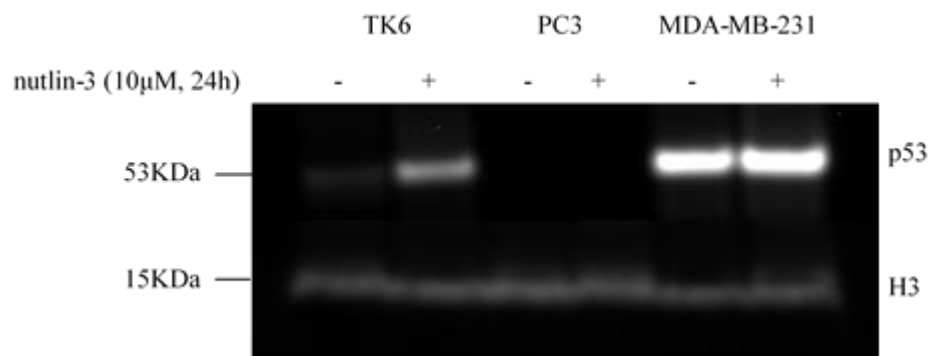
In the interests of understanding how HDAC inhibitors generate DNA damage for the purposes of human risk assessment, this chapter aims to establish and validate the ChIP-chip assay in a suitable human cell line, such as TK6 lymphoblasts. Continuing on from the previous yeast chapter where attempts to demonstrate the effects of HDAC inhibition were largely unsuccessful, it will therefore be important to show some of the primary effects of HDAC inhibitors in the newly introduced human cell line. The outcome of this chapter will be an optimised human cell line-specific ChIP-chip protocol that can then be applied to make the necessary measurements to test the main hypothesis set out by the thesis.

#### 4.6 Confirming the p53 proficiency of TK6 cells

Since the focus lies in trying to understand how HDAC inhibitors induce DNA damage, it is important that cells can respond to DNA damage in a normal manner. Therefore, the *TP53* status of TK6 cells used in this study was first assessed by measuring apoptosis in response to nutlin-3. For comparison, the p53-mutant MDA-MB-231 breast cancer cell line (Hui et al. 2006) and the p53-null PC3 prostate cancer cell line (Scott, Earle, and Gumerlock 2003) were selected as negative controls. Apoptosis was measured using Annexin-V (Chapter II, 2.1.4.1), which has a strong Ca<sup>2+</sup>-dependent affinity for phosphatidylserine residues externalised on the plasma membrane during the onset of programmed cell death. When treated with increasing concentrations of nutlin-3 (1µM, 5µM; 10µM) for up to 48 hours, a significant increase in the number of cells expressing Annexin-V was observed in TK6 cells (P < 0.05, 5µM; P < 0.01, 10µM), indicating the onset of apoptosis (Figure 4.1A). In contrast to this, neither MDA-MB-231 nor PC3 cells showed noticeable changes in Annexin-V (Figure 4.1B; C). To further examine changes in the expression of p53, lysates prepared from cells treated with 10µM nutlin-3 for 24h were probed for p53 protein levels by immunoblotting (Chapter II, 2.1.5). An increase in p53 protein expression was observed only in TK6 cells and absent from PC3 cells lacking p53 (Figure 4.2). MDA-MB-231 cells expressed high levels of mutant p53 and were insensitive to the effects of nutlin-3 (Figure 4.1). Together these results demonstrate that TK6 cells used here are p53-proficient.



**Figure 4.1. Nutlin-3 promotes apoptosis in TK6 cells.** **A:** TK6 cells, with that of control cell lines **B:** MDA-MB-231 (p53 mutant) and **C:** PC3 (p53 null). Percentage of cells in population undergoing apoptosis quantified using flow cytometry, gating on populations staining positive for Annexin-V (early apoptosis) and Annexin-V + PI (late apoptosis). Statistical analysis (one-way ANOVA with Turkey’s Post-hoc analysis) was performed using SPSS v22.0. n = 3. \*P < 0.05, \*\*P < 0.01. Error bars as SEM. (Raw data in Appendix, II).



**Figure 4.2. Immunoblot of p53 expression in select cell lines in response to nutlin-3.** The expression of p53 was probed from whole cell extracts of cells treated with 10 $\mu$ M nutlin-3 for 24h. Histone H3 was used as a loading control.

#### **4.7 Examining the toxicity of TSA in TK6 cells to establish a suitable treatment regimen**

Since previous attempts to use HDAC inhibitors in yeast proved unsuccessful, it is important to ensure that the effects of HDAC inhibition can be reproduced in the mammalian TK6 cell line. The toxicity of HDAC inhibitors are most pronounced following a 24h treatment, with cells showing a dose-dependent increase in DNA damage and cell cycle arrest (Olaharski et al. 2006). Particularly, Olaharski and others noted that the pan-HDAC inhibitor Trichostatin A induced a significant increase in micronuclei formation in addition to the number of cells arrested in G<sub>1</sub> phase (Olaharski et al. 2006). While TSA proved ineffective in yeast cells due to various complications, it has been extensively used in a number of different mammalian cell lines with demonstrated effects (Mogal and Abdulkadir 2006; Finnin MS 1999; Cong et al. 2013; Hrebackova et al. 2009). Therefore, the aim was to first use an HDAC inhibitor such as TSA which has previously been used in TK6 cells (Olaharski et al. 2006) and demonstrate its genotoxic effects here in this study. Furthermore, this would also allow the establishment of a treatment regimen to study the mechanisms of DNA damage induction by HDAC inhibitors.

To do this, micronuclei (MN) formation and cell cycle changes were measured using the InVitro MicroFlow® assay (Chapter II, 2.1.4.3), which has comparable performance to the conventional micronuclei assay carried using microscopy (Collin J.E. et al.). Using the InVitro MicroFlow® assay, the percentage of cells in the population exhibiting micronuclei can be determined through SYTOX Green staining. Ethidium monoazide (EMA) binds to DNA of lysed cells and serves as an indicator of cell death and in combination with relative cell count (RCC) was used as indicators of cytotoxicity. Firstly, the InVitro MicroFlow® assay was validated using the aneugenic compound Vinblastine (Collin J.E. et al.). TK6 cells treated with increasing concentrations of Vinblastine for 24h showed a significant increase ( $P < 0.05$ ) in the percentage of cells positive for MN for the two top doses (Figure 4.3A). In addition, the highest level of cytotoxicity observed did not exceed 30%. This confirms Vinblastine as a genotoxic agent and successfully validates the assay.

Next, to examine ability of TSA to induce MN formation in TK6 cells as an indication of its genotoxicity, cells were subjected to increasing concentrations of the drug over a 24h period. TK6 cells treated with 313nM and 625nM TSA showed a significant increase in MN formation ( $P < 0.05$ ,  $P < 0.01$ , respectively) (Figure 4.3A).

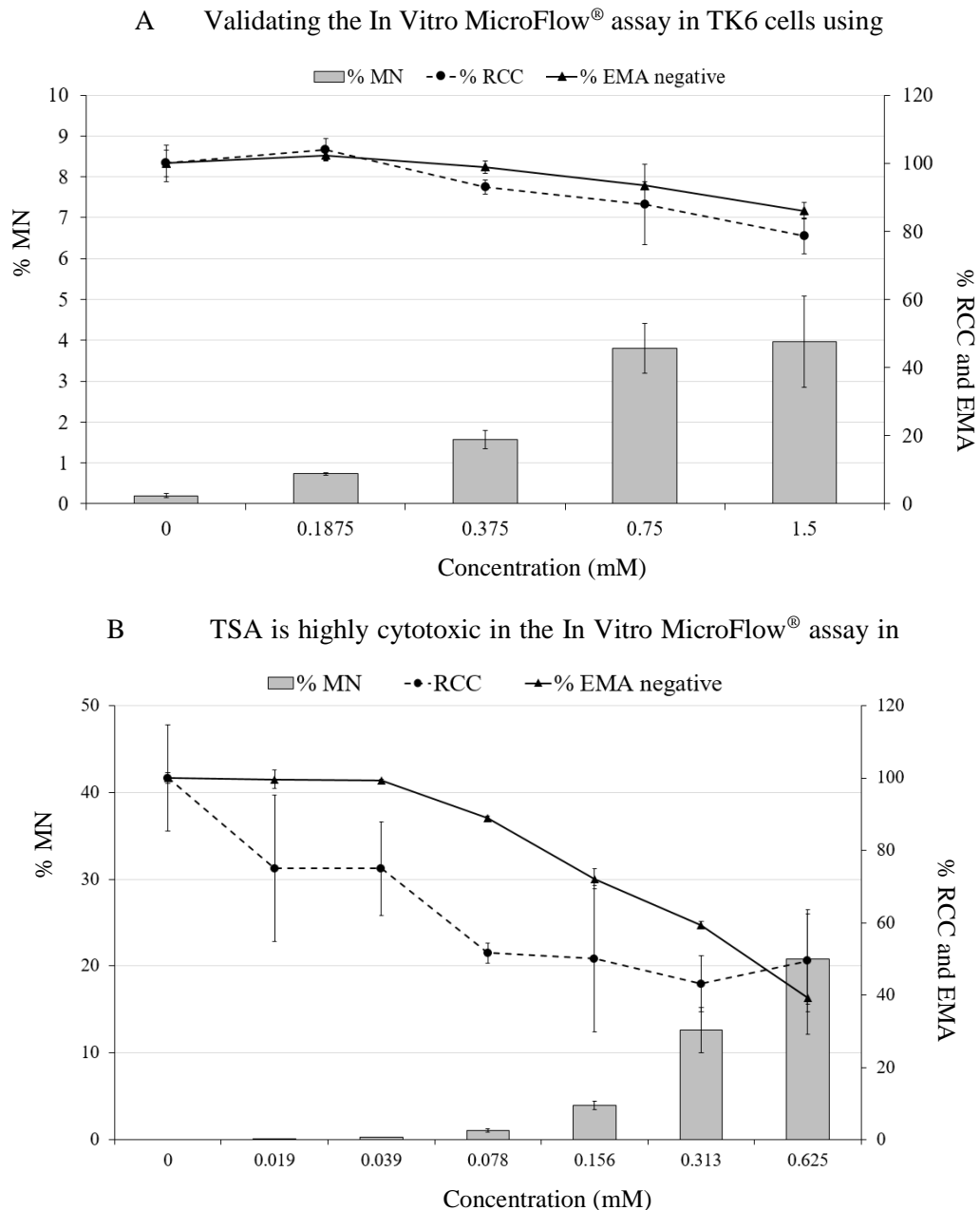
However, these doses also display a high level of cytotoxicity as shown by the reduction in RCC (~50%) and EMA negative cells (~60%) (Figure 4.3A). This shows that concentrations of TSA displaying genotoxicity occurs at the acceptable limit of cytotoxicity ( $55 \pm 5\%$ ) ('OECD (2014). Test No. 487: In Vitro Mammalian Cell Micronucleus Test '). Concurrently, analysis of the same EMA data showed that a significant proportion of TK6 cells were arrested in  $G_{0/1}$  with increasing concentration of TSA ( $P < 0.01$ ) (Figure 4.4). Nevertheless, these results show that because TSA demonstrates borderline genotoxicity, measured using the InVitro MicroFlow® assay, it would not be suitable to use such a treatment regimen for studying DNA damage induction. Therefore, to avoid this problem, a shorter treatment time point will be used.

#### **4.8 Examining the short-term effects of TSA in TK6 cells using induction of $\gamma$ H2AX as a marker of DNA damage**

Olaharski and others also reported in their studies that TSA was able to induce phosphorylation of histone H2AX Serine 139 ( $\gamma$ H2AX), an epigenetic marker of DNA damage (Kuo LJ 2008), in TK6 cells within hours of treatment (Olaharski et al. 2006). Additionally, this coincided with the increase in histone acetylation, a hallmark of HDAC inhibitors (Zhou et al. 2000). Therefore, to examine these early effects, TK6 cells were treated with a fixed dose of TSA (400nM) across a range of time points leading up to 24h, with changes in histone H3 lysine 9 acetylation (H3K9ac) and  $\gamma$ H2AX measured using flow cytometry (Chapter II, 2.1.4.2).

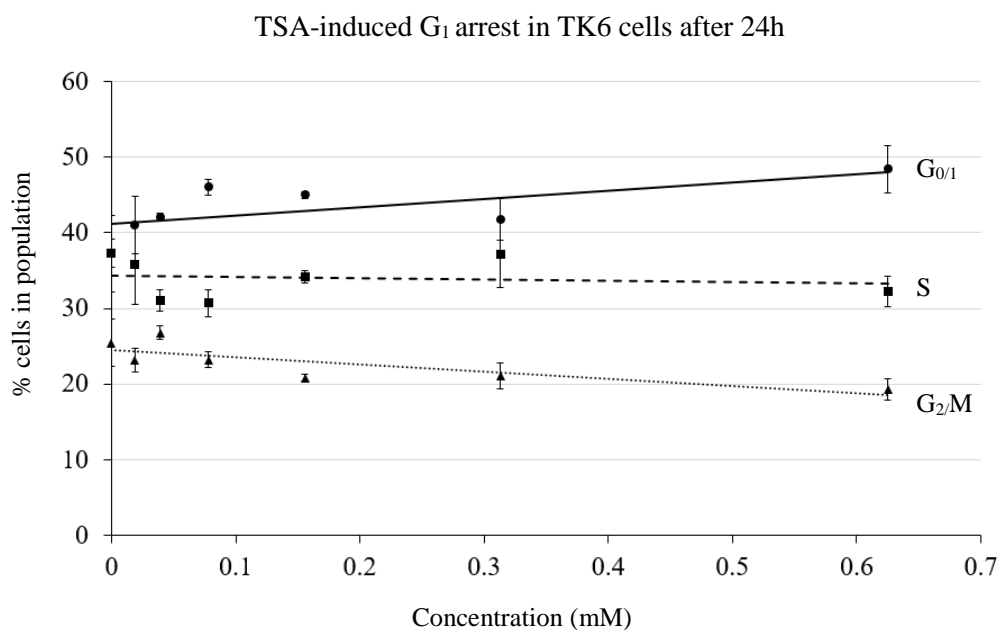
Following 400nM TSA treatment, TK6 cells displayed an increase in H3K9ac up to 12h and decreased at 24h, with the H3K9ac response being most significant at 6h ( $P < 0.05$ ) and 12h ( $P < 0.01$ ) (Figure 4.5A). In parallel with this, is a significant time-dependent increase in  $\gamma$ H2AX up to 24h ( $P < 0.01$ ) (Figure 4.5B). This shows that TSA is able to provoke a significant effect on both histone acetylation and  $\gamma$ H2AX in TK6 cells 6h following treatment, which is consistent with those reported previously in the literature (Leea J.H 2010; Olaharski et al. 2006; Zhou et al. 2000).

Next, to examine the extent of cytotoxicity induced over the 6h exposure period, TK6 cells were treated with TSA up to 400nM and changes in the Annexin-V levels, an early marker of apoptosis, and cell cycle  $G_{0/1}$  arrest were measured using flow cytometry (Chapter II, 2.1.4.1 and 2.1.4.4). In comparison to the vehicle control, TK6 cells treated

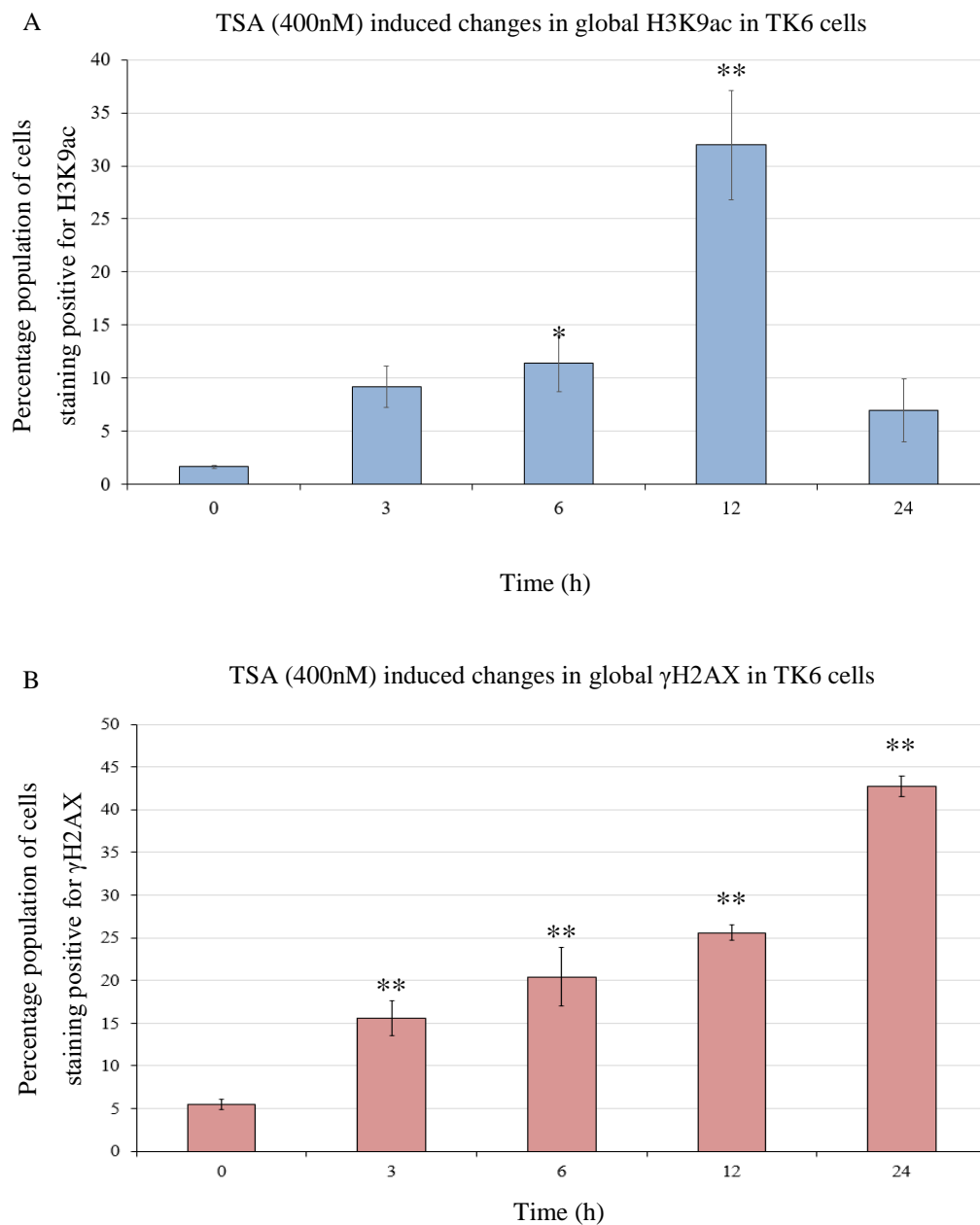


**Figure 4.3. Trichostatin A shows high levels of cytotoxicity in TK6 cells following 24h treatment as measured using the In Vitro MicroFlow<sup>®</sup> assay. A:** TK6 cells treated with Vinblastine for 24h was used to validate the In Vitro MicroFlow<sup>®</sup> assay. Significant increase in MN-induction at higher doses of 0.6mM and 1.2mM showing low levels of cytotoxicity (RCC and EMA negative) confirms that Vinblastine tests positive in the micronucleus flow assay, consistent with published findings (Collin J.E. et al.). **B:** TK6 cells treated with Trichostatin A for 24h tested using the In Vitro MicroFlow<sup>®</sup> assay shows a significant increase in MN-induction at 0.313mM and 0.625mM which coincides with high levels of cytotoxicity ( $\geq 50\%$  RCC or EMA negative). DMSO used as a vehicle control. Statistical analysis (Dunnnett's multiple comparison test) was performed using SPSS v22.0 on the %MN DATA. n = 3. \*P < 0.05, \*\*P < 0.01. Error bars as SEM. (Raw data in Appendix, II).

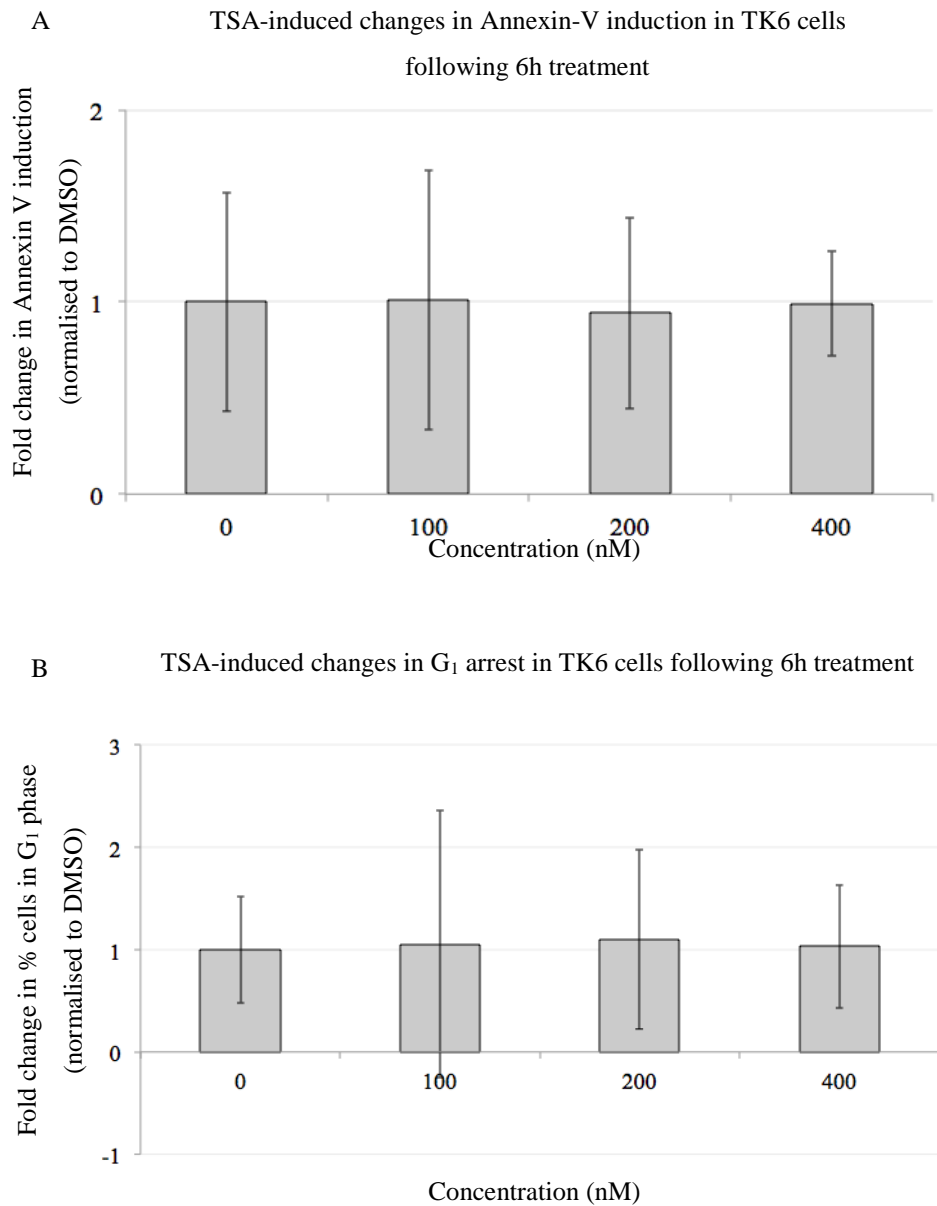




**Figure 4.4. Trichostatin A induces G<sub>1</sub> cell cycle arrest in TK6 cells following 24h treatment.** EMA data from flow assay used to assess changes in cell cycle profile in response to TSA, showing an increase in G<sub>1</sub> arrest  $P < 0.01$  (Trend analysis). DMSO used as a vehicle control.  $n = 3$ , Error bars as SEM. (Raw data in Appendix, II).



**Figure 4.5. TSA increases global H3K9 acetylation and H2AX phosphorylation in TK6 cells.** Nuclei from TK6 cells treated with TSA were extracted and stained with either **A:** anti-H3K9ac-PE or **B:** anti- $\gamma$ H2AX-FITC conjugated antibodies and analysed by flow cytometry. Percentage population of cells staining positive for either antibody is shown. Statistical analysis (Dunnnett's multiple comparison test) was performed using SPSS v22.0. \* $P < 0.05$ , \*\* $P < 0.01$ . Error bars as SEM,  $n = 3$ . (Raw data in Appendix, II).



**Figure 4.6. Treating TK6 cells with 400nM TSA for 6h does not trigger early apoptosis or G<sub>1</sub> arrest.** Results of flow cytometry analysis showing **A**: no significant difference in percentage of cell population staining positive for Annexin V-FITC ( $P > 0.9$ ), as a marker of apoptosis. Forward and side scatter gating was performed on DMSO treated cells. **B**: propidium iodide (PI) staining was used to examine changed in cell cycle profile and showed no significant increase in G<sub>1</sub> phase cells in response to TSA ( $P > 0.4$ ). Fold change calculated by normalising against the DMSO control. Statistical analysis (Dunnett's multiple comparison test) was performed using SPSS v22.0. Error bars as SE,  $n = 3$ . (Raw data in Appendix, II).

with TSA for 6 hours showed no significant increase in Annexin-V up to a dose of 400nM ( $P > 0.9$ ) (Figure 4.6A). Similarly, no significant increase in  $G_1$  arrest was observed under the same conditions ( $P > 0.4$ ) (Figure 4.6B). Together these results show that  $\gamma$ H2AX detected at 6h using 400nM TSA is independent of cytotoxicity. Taken together, these results show that treating TK6 cells with 400nM TSA for 6h displays indications genotoxicity in the absence of cytotoxicity and therefore represents a suitable treatment regimen to be used for studying the mechanisms of HDAC inhibitor-induced DNA damage using genome-wide approaches.

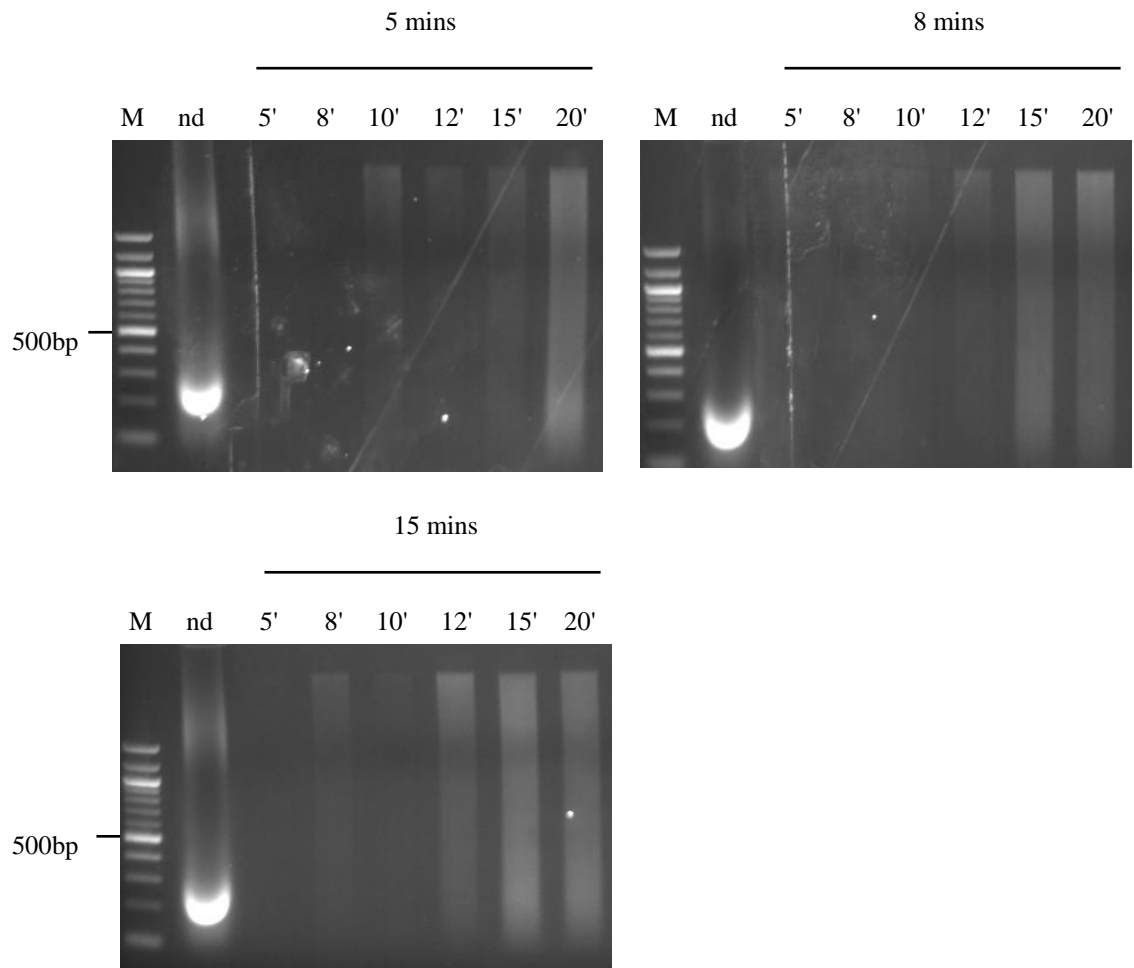
## **4.9 Optimising the ChIP-chip assay for TK6 cells**

### **4.9.1 Effect of formaldehyde fixation on sonication efficiency**

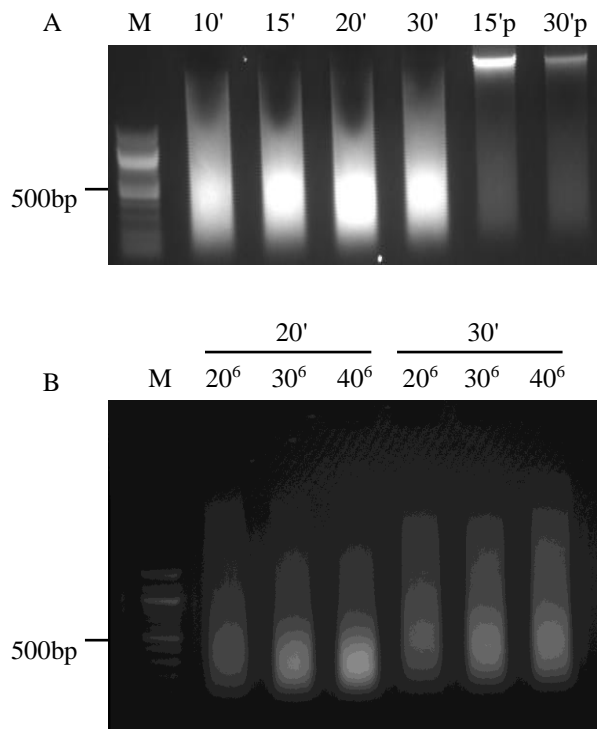
Having demonstrated a clear cellular response to TSA, the next step was to optimise sonication conditions in which chromatin from TK6 cells can be reliably fragmented to an average size of 500bp. To test if fixation time affects sonication efficiency, untreated TK6 cells were fixed in formaldehyde for 5, 8 and 15mins before being sonicated at different cycles. Under these conditions the distribution of de-crosslinked chromatin appears to be even at the higher cycles, ranging from genomic to 100bp (Figure 4.7). Fixation time therefore does not appear to affect sonication but rather, other factors are preventing the chromatin from being sheared effectively.

### **4.9.2 Investigating the effect of isolating nuclei to promote efficient shearing of TK6 cell chromatin**

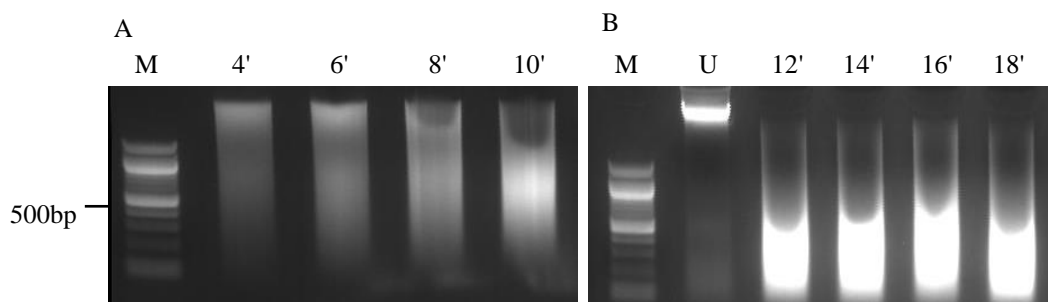
Inefficient cell lysis and a high cell density during sonication can potentially inhibit efficient chromatin shearing. To address the possibility that cells are not lysed properly prior to sonication, a 'two-step' lysis was adopted, which is sometimes used in commercial kits as opposed to the generally reported single-step lysis that uses a shearing buffer (Chapter II, 2.1.6). The results show a clear improvement in shearing from 10 – 30 cycles, producing a dense collection of fragments just under 500bp (Figure 4.8A). The remaining pellet for samples sonicated for 15 and 30 cycles following high-speed centrifugation shows that most of the chromatin remains in the supernatant, with negligible loss (Figure 4.8A).



**Figure 4.7. Effects of formaldehyde fixation time on shearing efficiency.** TK6 cells were fixed in formaldehyde for 5, 8 and 15mins. Chromatin was then extracted and sonicated from 5 to 20 cycles (30s on/off per cycle), before being de-crosslinked at 65°C. DNA fragment sizes were analysed by gel electrophoresis using 1.2% agarose gels. M – NEB 100bp ladder; nd – non-decrosslinked sample.



**Figure 4.8. Cytoplasmic pre-clearing is required to efficiently shear chromatin from varying densities of TK6 cells.** Following changes to the protocol, separating the cytoplasmic and nuclear lysis into two steps, **A**: chromatin can be successfully sheared using a range of cycles ranging from 10 to 30. Pellets gathered from samples sonicated for 15 (15'p) and 30 (30'p) cycles show negligible loss of fragmented chromatin. **B**: chromatin from TK6 cells can be sheared in different densities at a fixed volume of 100 $\mu$ l per 1.5ml tube. M – NEB 100bp ladder.



**Figure 4.9. 10 cycles yield the optimal fragment length at a density of  $4 \times 10^6$  cells per tube.** Chromatin sonicated using the  $4 \times 10^6$  cells per tube for **A**: 4 – 10 cycles and **B**: 12 – 18 cycles. M – NEB 100bp ladder, U – unsonicated de-crosslinked sample.

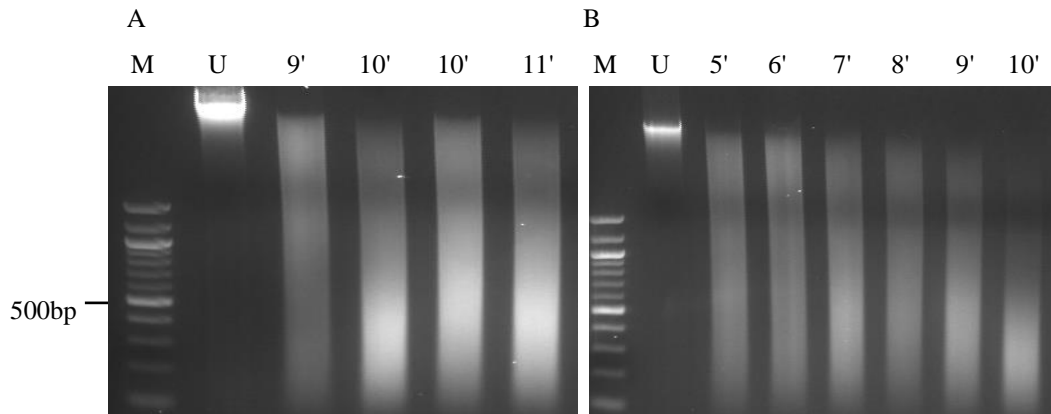
Provided that the lysis conditions are established, the maximum cell density allowed per tube in a fixed volume of buffer was investigated. Identifying the highest cell density that can be used would allow for greater numbers of cells to be processed each time. For 2, 3 and 4 million cells sonicated for 20 and 30 cycles, successful shearing was observed up to the highest density of cells tested (Figure 4.8B). To identify the optimal cycle that would yield an average fragment size of 500bp, different cycles were tested with a density of  $4 \times 10^6$  cells per 100 $\mu$ l of shearing buffer. A decrease in fragment size was observed with increasing cycles, with the average desired fragment size achieved at 10 cycles with no further reduction in fragment size when using more than 12 cycles (Figure 4.9).

#### **4.9.3 Assessing the reproducibility of shearing chromatin from TK6 cells**

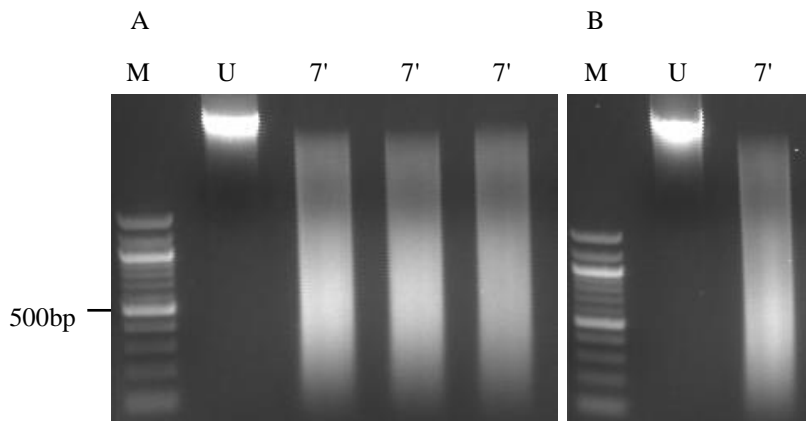
Further attempts to reproduce the shearing observed at 10 cycles however resulted in inconsistencies (Figure 4.10). To rectify this, the cell density used during sonication was reduced from  $4 \times 10^6$  to  $3 \times 10^6$ , which previously was also shown to be viable (Figure 4.8B). Since the cell density parameter has been altered, sonication cycle was re-optimised (Figure 4.10B). Under the newly established conditions,  $3 \times 10^6$  cells sonicated for 7 cycles produced the desired average fragment length and these results can be reliably reproduced between both technical (Figure 4.11A) and biological (Figure 4.11B) repeats.

#### **4.9.4 Examining the ability to enrich for RNA polymerase II (pSer2) using ChIP**

Having successfully identified a robust means of fragmenting chromatin from TK6 cells to the optimal size required for ChIP, the ability to immunoprecipitate proteins of interest using the prepared material was assessed by enriching for RNA polymerase II phospho-serine 2 (RNA Pol II pSer2). The house keeping genes *GAPDH* and *RPL10* were selected as positive control loci where RNA Pol II is known to bind (Aithal MG 2015). On the other hand, *GADI* and *MYOD1* were used as negative controls. *GADI* encodes for glutamic acid decarboxylase that is responsible for the production of the inhibitory neurotransmitter GABA ( $\gamma$ -Aminobutyric acid) and is expressed only in neurons. Similarly, *MYOD1* is required for muscle cell differentiation and therefore not expressed in TK6 cells. These two genes are often silenced in commonly used cell lines and serve as suitable negative control genes for RNA Pol II occupancy (Sigma-Aldrich 2006).



**Figure 4.10. Sonicating with a high cell density affects reproducibility.** A: two independent repeats at 10 cycles using  $4 \times 10^6$  cells per tube shows variability in shearing. B: sonication cycles were re-optimised using a lower cell density ( $3 \times 10^6$  per 100 $\mu$ l per 1.5ml tube). M – NEB 100bp ladder, U – unsonicated de-crosslinked sample.

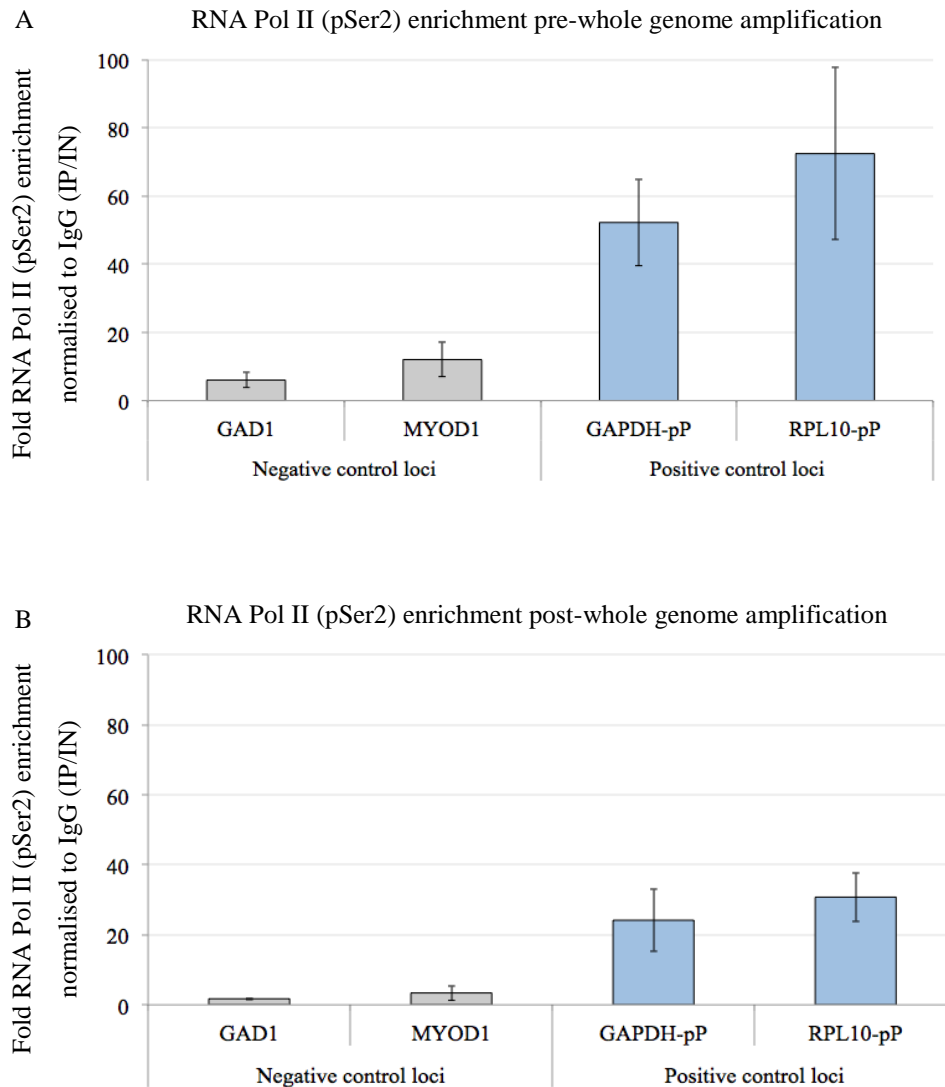


**Figure 4.11. Reducing cell density significantly improves the reproducibility of sonication.** Chromatin from TK6 cells can be reliably sheared to the optimal length using 7 cycles at a density of  $3 \times 10^6$  per tube, demonstrating reproducibility between **A**: technical and **B**: biological repeats. M – NEB 100bp ladder, U – unsonicated de-crosslinked sample.



RNA Pol II (pSer2) immunoprecipitation using chromatin from untreated TK6 cells was carried out as described (Chapter II, 2.1.6) using the optimal antibody amount identified from a titration experiment (Appendix, II). Fold enrichment of RNA Pol II (pSer2) at each gene was calculated by dividing the IP (immunoprecipitated DNA) by the input (IN) (non-immunoprecipitated background) and normalised against the IgG control (raised in the same species but lacking antigen specificity). An average 50 and 70-fold enrichment in RNA Pol II (pSer2) was detected at *GAPDH* and *RPL10* (Figure 4.12A). In contrast, the negative control loci *GADI* and *MYOD1* showed less enrichment in comparison (Figure 4.12A), indicating that the immunoprecipitation experiment was successful in enriching for RNA Pol II (pSer2) at the expected genes and that the conditions for performing chip on TK6 derived chromatin were fully optimised.

Following this, amplification of the ChIP DNA is required to generate enough material for microarray hybridisation. This is commonly done through PCR-based methods. Though the exponential nature of the technique can lead to bias if the ChIP'd DNA is of low quantity (< 10ng) (Sigma-Aldrich 2006). Since it is difficult to reliably quantify low nanogram quantities of DNA typically yielded from ChIP experiments, the ability to retain high RNA Pol II (pSer2) enrichment at positive control loci over negative control loci was investigated to determine if PCR amplification significantly alters representation of the original ChIP DNA. RNA Pol II (pSer2) immunoprecipitated samples were amplified following instructions provided by the GenomePlex<sup>®</sup> WGA2 kit (Chapter II, 2.1.7.1) and qPCR was performed to evaluate enrichment at the same control genes. Post-whole genome amplification analysis showed an overall reduction in RNA Pol II (pSer2) enrichment (Figure 4.12B). Particularly, a 2-fold decrease (50 to 25) at *GAPDH* and a 2.3-fold decrease (70 to 30) at *RPL10* was observed (Figure 4.12). Additionally, the negative control loci also showed a drop in enrichment. However, RNA Pol II (pSer2) enrichment is still higher at the positive compared to the negative control genes. This suggests that even though the overall enrichment level may be less, the relative enrichment is still effectively maintained and therefore RNA Pol II (pSer2) occupancy patterns observed from microarray experiments would still reflect the original ChIP sample.



**Figure 4.12. Whole genome amplification retains relative RNA Pol II (phospho-Ser2) enrichment at select genes.** **A:** Results of qPCR showing RNA Pol II (phospho-Ser2) enrichment at genes where occupancy expected (*GAPDH* and *RPL10*) versus gene where it's not expected (*GAD1* and *MYOD1*).  $n = 4$ . **B:** following amplification using the WGA2 kit (Sigma-Aldrich) the relative RNA Pol II (phospho-Ser2) enrichment is maintained between positive and negative control genes.  $n = 3$ . Error bars and SEM. (Raw data in Appendix, II).

#### 4.9.5 Examining the quality of RNA Pol II (pSer2) ChIP-chip data

Next, to examine RNA Pol II (pSer2) occupancy in greater detail on chromosome 17 in TK6 cells, IN and IP samples were differentially labelled with fluorescent dyes and hybridised onto the custom designed DNA microarrays (Chapter II, 2.1.7.2). After 48h, the microarrays are removed and scanned in an Agilent Type-C scanner. The fluorescent intensities of each probe on the array is captured and returned as an image file. To quantify this as relative IN and IP ratios, the image was analysed using the Agilent Feature Extraction software (Chapter II, 2.1.7.5). The final output is a large datafile containing the raw IP/IN  $\log_2$  ratios of each probe in the array along with their identifiers (ie. Probe ID, ProbeName; GeneName).

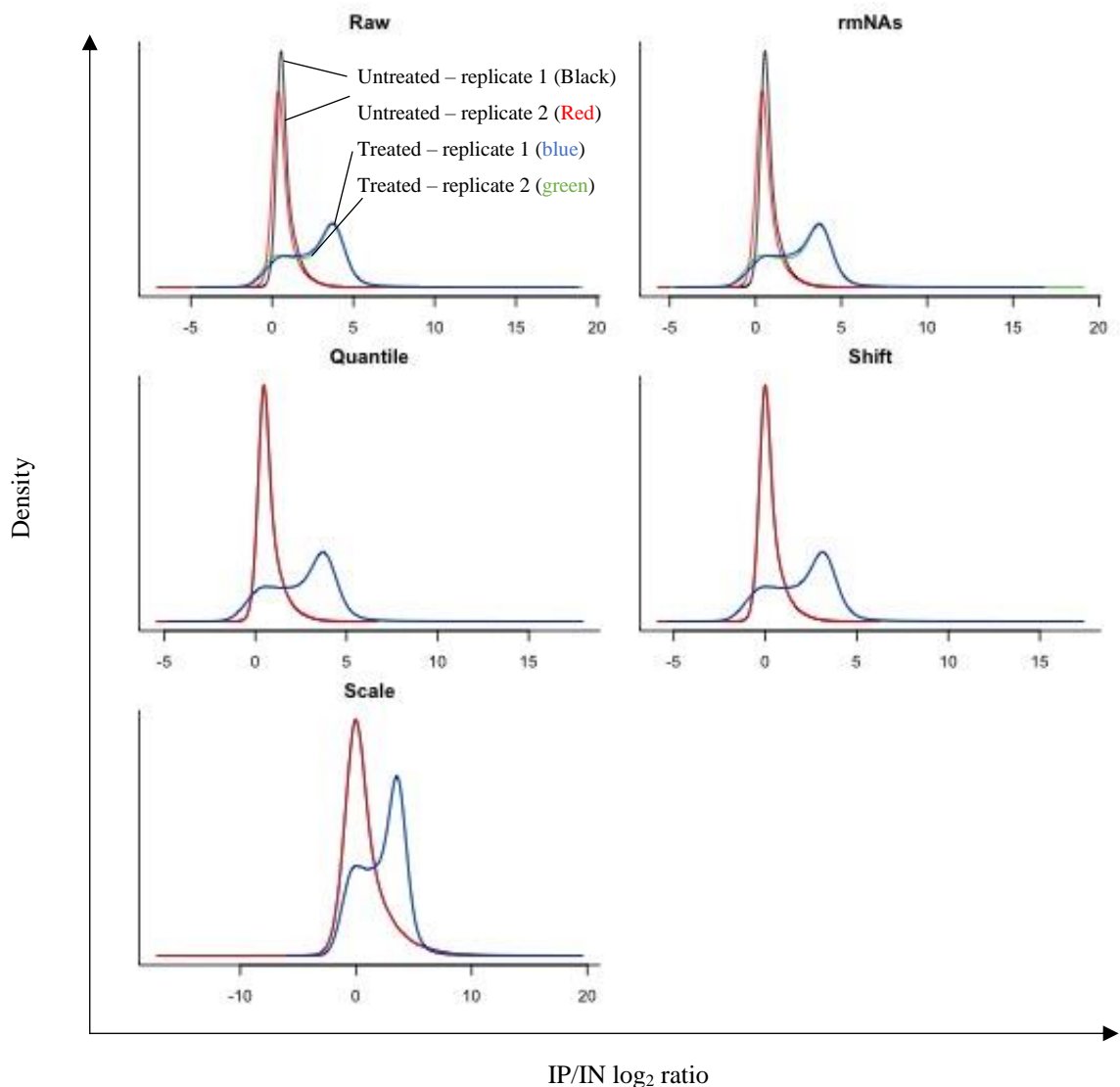
Before further analysis can take place, the raw data must first be normalised. This is important as it minimises technical bias while allowing relative comparisons of biological differences to be made. To do this, the raw data file is loaded into Sandcastle, which is a software developed in the statistical programming language R, specifically for the purposes of normalising and analysing microarray data from two or more related experiments (Bennett 2015). Figure 4.13 illustrates the normalisation process implemented by 'Sandcastle' for the raw RNA Pol II (pSer2) array data generated from untreated and TSA treated cells. The raw data is visualised as a kernel density plot of IP/IN  $\log_2$  ratios (x), verses density (y). This type of plot is an effective way of representing the overall distribution of the data. The two biological repeats of RNA Pol II (pSer2) array data from untreated cells are shown in red and black, while those from TSA treated cells are shown in green and blue, with highly reproducible data overlapping each other (Figure 4.13, *raw*). The next step is to remove redundant probe values in each dataset, for instance those that fail to generate an IP/IN  $\log_2$  ratio (Figure 4.13, *rmNAs*).

A key assumption of the normalisation process is that the background sub-population (non-enriched array data) follows a standard normal distribution with the enriched sub-population (enriched array data) displaying higher than average IP/IN  $\log_2$  ratios and therefore tends towards the right-tail of the distribution (Bennett 2015). The normalisation process takes two stages, the first being quantile normalisation which imposes the same distribution on each of the biological repeats (Figure 4.13, *Quantile*). Secondly, the distribution of each dataset is shifted so that the centres of the densest distribution (background sub-population) aligns on the  $x = 0$  axis (Figure 4.13, *Shift: dashed line*). In this case, as the RNA Pol II (pSer2) data gathered from untreated cells

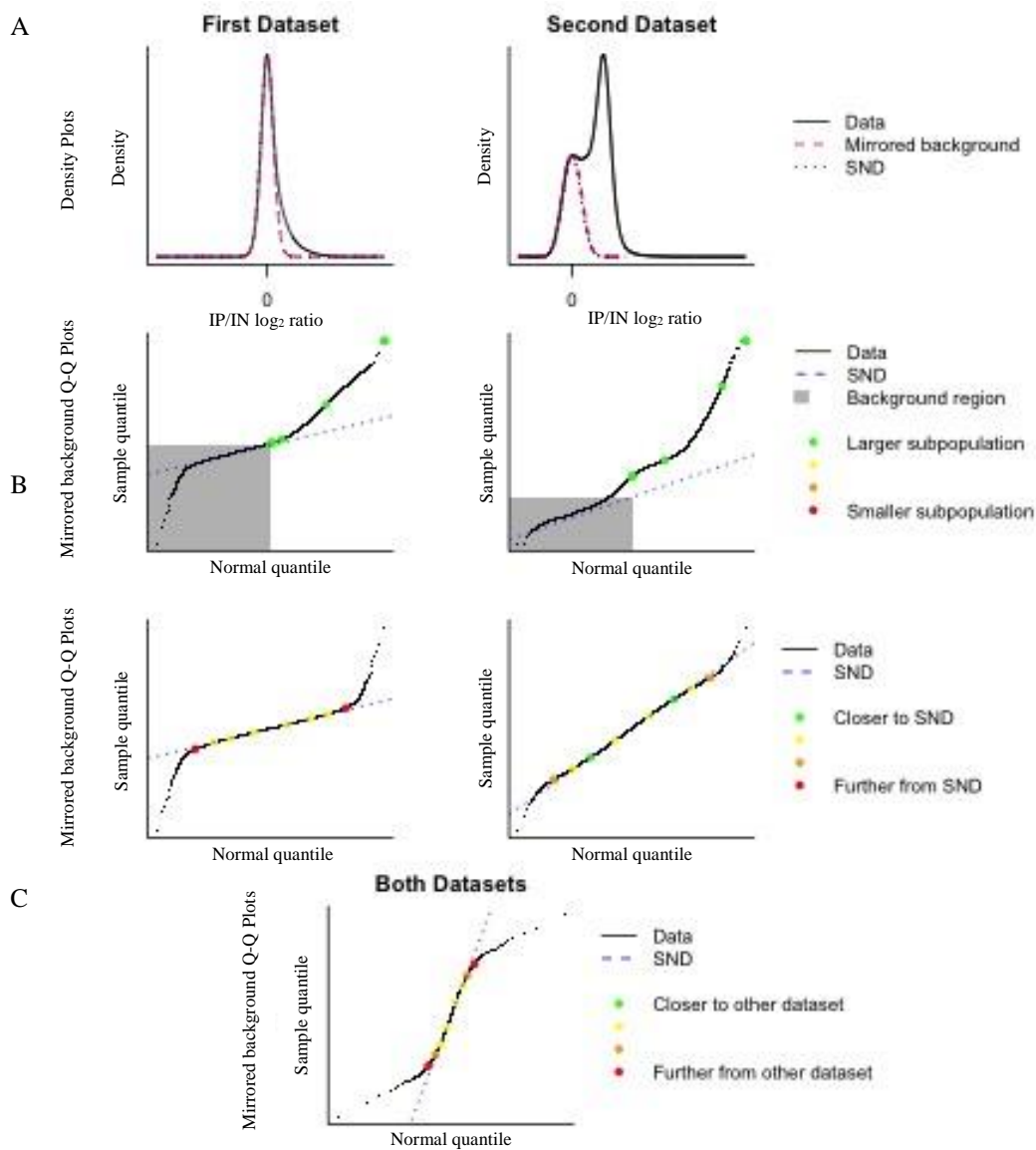
are normalised with the data from TSA treated cells, the background sub-population of all four datasets are centred on  $x = 0$  (Figure 4.13, *Shift: dashed line*). Lastly scaling is used to control for variability arising in the microarray workflow (Figure 4.13, *Scale*), for instance differences in labelling and hybridisation efficiencies, array artefacts, changes in the environment. This is done by using all the probe values in the estimated background to adjust the remaining probe values so that the data follows a standard normal distribution (SND) with mean = 0 and standard deviation = 1 (Bennett 2015).

The ability of any two or more sets of data to be normalised together depends entirely on their background sub-populations being similar to that of a SND. Should this not be the case then comparisons would be restricted to binary rather than relative changes in signal, limiting the potential of the data. Therefore, to assess if the array data generated in untreated and TSA treated cells can be normalised together, a post-normalisation test is carried out to visually evaluate if: 1) the averaged data for a single treatment group satisfies the assumptions for normalisation, 2) the averaged data between two treatment groups satisfy the assumptions for normalisation.

To test the first assumption using the RNA Pol II (pSer2) array data, the quantiles of the data from untreated and TSA treated cells are compared independently against the SND as a Q-Q plot. Such plots are commonly used to check if the data collected follows a normal distribution. Figure 4.14A shows the normalised data for each treatment group with the mirrored background sub-population and the SND. In each case the left-hand side ( $x < 0$ ) follows a SND, whereas the right-hand side ( $x > 0$ ) deviates from the SND by a varying extent. This is therefore reflected in the Q-Q plots as the data that match the SND fall on the  $x = y$  line and those values that do not deviate significantly from the  $x = y$  line (Figure 4.14B). Typically, any data deviating from the SND on the right-hand side have higher IP/IN  $\log_2$  ratios and represent biological signal. Moreover, since a greater deviation is observed in the data collected from TSA treated cells compared to that of untreated cells, this provides an early indication that there is an increase in RNA Pol II (pSer2) signal following TSA treatment (Figure 4.14B). Next, only the mirrored background sub-population data is compared to the SND in the Q-Q plot and since the majority of this data lie on the line, showing that the mirrored background sub-population is similar to the SND, the first assumption for normalisation are met (Figure 4.14C). Lastly, to test the second assumption and see if the RNA Pol II (pSer2) data from untreated and TSA treated cells can be normalised together, a Q-Q plot of the mirrored background sub-population data of each is compared. Figure 4.14D shows that since the

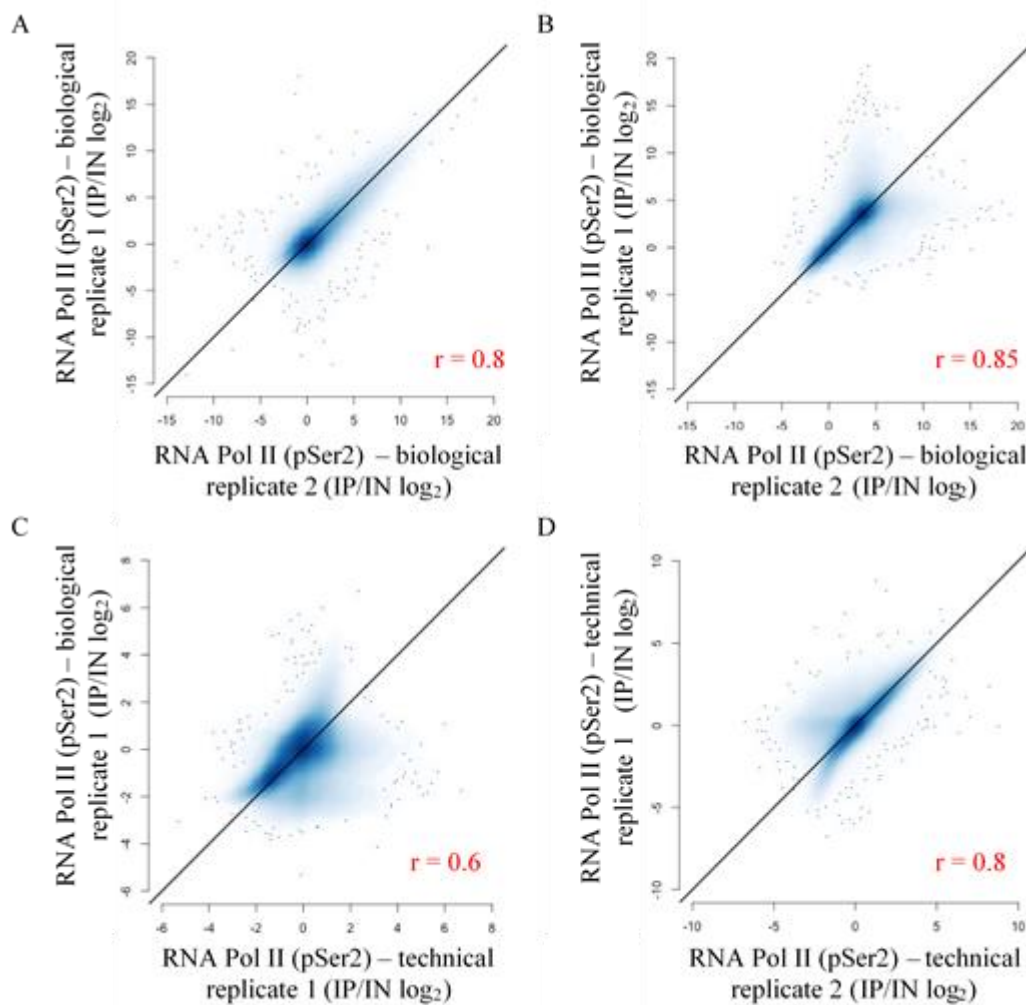


**Figure 4.13. Normalisation procedure for RNA Pol II (pSer2) microarray data.** Two biological replicates showing raw data from: untreated cells – *black* and *red*; TSA treated cells – *blue* and *green*. *Raw* – kernel densities of two raw biological repeats for each treatment condition shown in the same colour. *rmNAs* – raw data after removing probes with missing values. *Quantile* – biological repeats from the same treatment condition are quantile normalised together. *Shift* – aligns data from both treatment conditions to the central point of the background population. *Scale* – the standard deviation for the estimated background sub- population is adjusted to 1 for both datasets.



**Figure 4.14. RNA Pol II (pSer2) microarray data meets the assumptions for normalisation.**

First dataset – from untreated cells. Second dataset – from TSA treated cells. **A:** density plots of the mirrored background approximating to the standard normal distribution (SND). **B:** density plots shown in A represented as Q-Q plots. Data in the background (shaded) region should closely follow the SND, whereas deviations of the right tail represents the enriched sub-population. **C:** Q-Q plot of only the mirrored background data. **D:** Q-Q plot of the mirrored background for both datasets. Little/no departure from the SND indicates that the two datasets are suitable to be normalised together.



**Figure 4.15. Results of RNA Pol II (pSer2) chromosome 17 arrays are reproducible.** Normalised IP/IN log<sub>2</sub> ratios between two biological repeats of **A**: untreated TK6 cells and **B**: TSA treated (400nM, 6h) TK6 cells shown as smooth scatter plots. A high correlation was observed between biological repeats. The effect of atmospheric ozone on reproducibility demonstrated in **C**: technical replicates hybridised on different days and **D**: technical replicates hybridised on the same day. Pearson's correlation calculated using functions in the base R package. x = y line shown in black.

background sub-populations of these two datasets are similar, they can effectively be normalised together allowing further in-depth analysis of the relative changes. Together these results show that successfully immunoprecipitated targets in human TK6 cells, such as RNA Pol II (pSer2) can be reliably mapped using the established ChIP-chip protocol.

Following this, correlation between replicates were analysed to determine if the results of the human microarrays are reproducible (Figure 4.15). For both experimental conditions tested, biological replicates showed a high degree of correlation ( $r > 0.8$ ) (Figure 4.15A; B). To question the technical reproducibility of the arrays, labelled DNA from the first TSA treated sample was re-hybridised onto new arrays. Doing so showed that the technical replicates display a lower correlation (Figure 4.15C). As expected, microarray data tend to correlate better when post-hybridisation washes are carried out simultaneously (Figure 4.15A; B). To confirm this, an additional technical repeat using chromatin from TSA treated cells was performed. Matching samples were then hybridised and washed together, resulting in a higher correlation (Figure 4.15D), suggesting that external factors may contribute to the variability of the readout from the arrays. Together these results show that successfully immunoprecipitated targets, such as RNA Pol II (pSer2) occupancy, in human TK6 cells can be reliably mapped using microarrays.

#### **4.9.6 Validating chromosome 17 microarrays of RNA Pol II (pSer2) occupancy in TK6 cells**

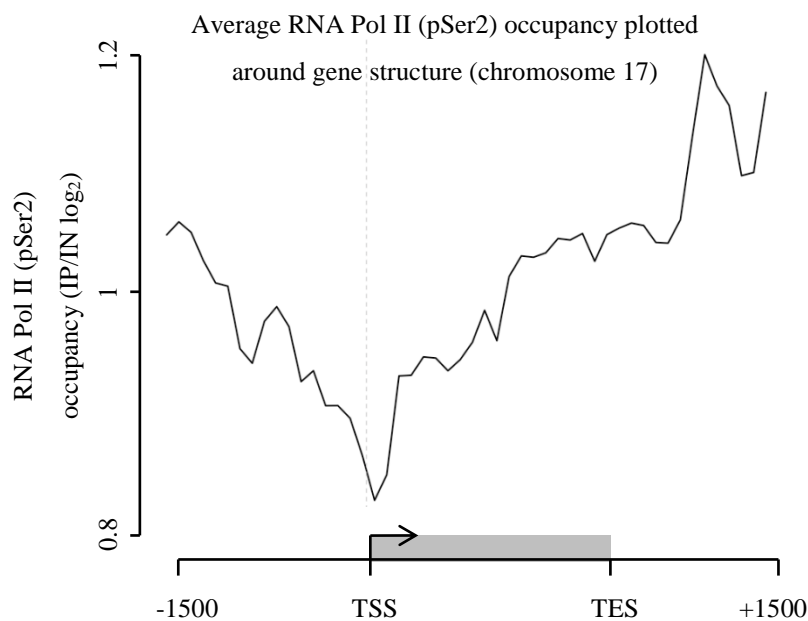
RNA polymerase II serine 5 and serine 2 phosphorylation are conserved marks of transcriptional initiation and elongation, respectively (Hsin JP 2012). RNA Pol II is phosphorylated at Serine 2 (pSer2) along the open reading frame (ORF) and ends of transcribed genes whereas Serine 5 is phosphorylated primarily at transcription start sites (TSS) (Hsin JP 2012). Given the distinct pattern of distribution, the average RNA Pol II (pSer2) occupancy was plotted +/- 1500bp relative to the ORFs of all genes on chromosome 17 for untreated TK6 cells (Figure 4.16). The observed pattern is consistent with what was expected (Hsin JP 2012), showing lower occupancy at the TSS compared to the ORF and downstream of the transcription end site (TES) (Figure 4.16).

To assess the impact of treating with TSA on RNA Pol II (pSer2) enrichment at *TP53* gene, genome plots of untreated and TSA treated data are presented, showing an overall higher level of RNA Pol II (pSer2) binding following TSA treatment (Figure 4.17A). The change in RNA Pol II (pSer2) enrichment at *TP53* gene due to the effect of

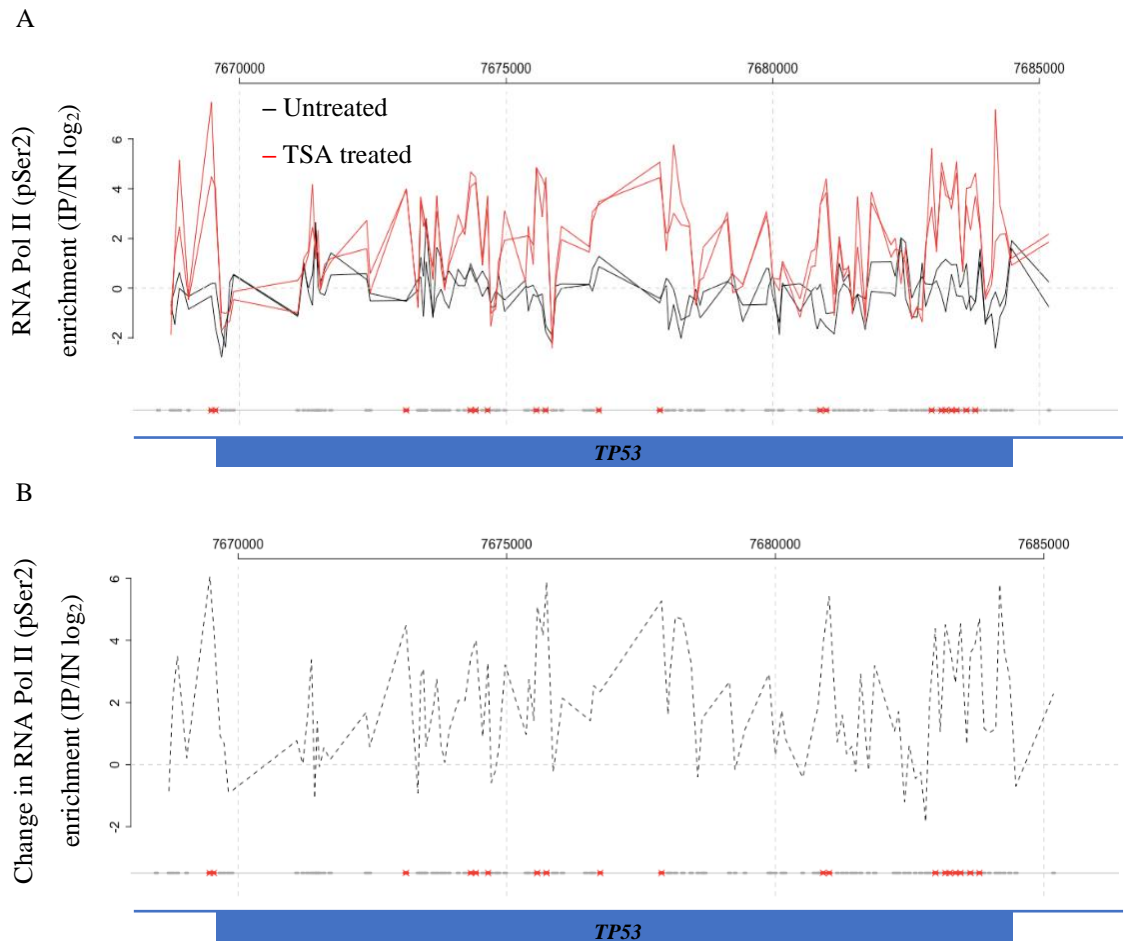


TSA is illustrated by subtracting untreated data from treated data (Figure 4.17B). The relative increase in ‘elongating’ RNA Pol II (pSer2) at *TP53* observed here is consistent with results presented by Sonnemann and others (Sonnemann et al. 2014) and may signify increased transcription of p53 as a result of genetic stress and or DNA damage. An alternative means of validating these results would be to make comparisons with those of published data available from the NCBI GEO repository (Barrett et al. 2013). To this end, chromosome 17 genes enriched for RNA Pol II in HepG2 (Hepatocellular carcinoma), H9 (pluripotent embryonic stem cells), U87 (glioblastoma astrocytoma) and TK6 cells were identified and the significance of the overlap examined using the hypergeometric test. The Venn diagram (Figure 4.18) shows a significant overlap ( $p < 0.01$ ) in genes enriched for RNA Pol II between all cell lines, suggesting that the results observed in TK6 cells are consistent with previous observations.

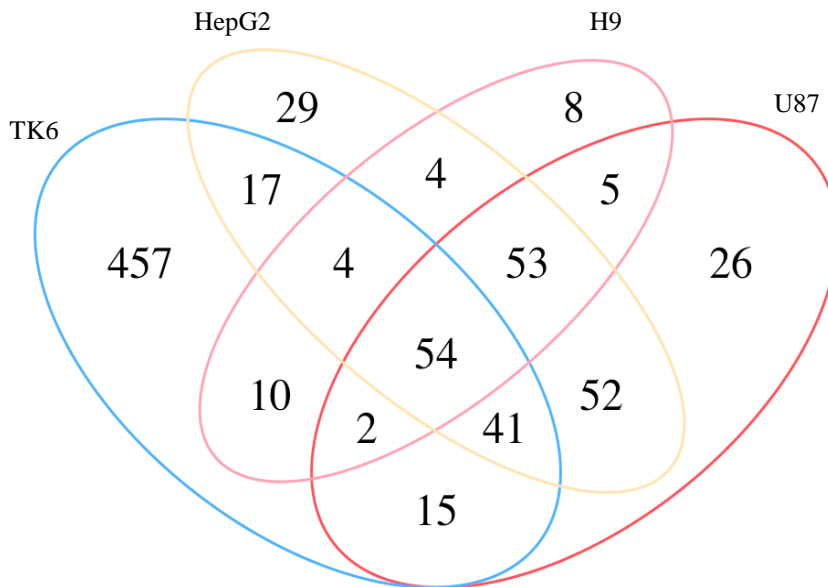
The phosphorylated state of RNA Pol II was demonstrated to reflect mRNA levels (Odawara et al. 2011), therefore changes in RNA Pol II (pSer2) measured following TSA treatment in TK6 cells are expected to correlate with changes in gene expression. However, a lack of suitable data complicates comparisons by introducing caveats such as different cell lines, HDAC treatment time and doses used. This was revealed by attempts to compare the TK6 RNA Pol II (pSer2) occupancy data generated here with expression data from THP-1 cells (human monocytic cell line, treated with 300nM TSA for 150mins) and HUVECs (human umbilical vein endothelial cell, treated with 500nM TSA for 24h) (Barrett et al. 2013), where a significant correlation was only observed for down regulated genes ( $P < 0.05$ , FDR  $< 0.05$ ) (Appendix, II). These results show that use of several methods to validate the RNA Pol II (pSer2) microarray data have been successful, demonstrating the establishment of the ChIP-chip protocol for TK6 cells.



**Figure 4.16. Average RNA Pol II (pSer2) enrichment around all genes on chromosome 17 displays the expected biological patterns.** The average of two biological repeats of RNA Pol II (pSer2) occupancy plotted +/- 1500 bp of the open reading frame (grey box) for all genes on chromosome 17 in untreated TK6 cells. Enrichment is pronounced within transcribed regions and downstream of TES. Total of 1932 protein coding and non-coding genes on chromosome 17 (hg19). RNA Pol II (pSer2) enrichment corresponds to 6.85% of total probes. TSS – transcription start site, TES – transcription end site. n = 2.



**Figure 4.17. TSA promotes RNA Pol II (pSer2) occupancy at *TP53* in TK6 cells.** **A:** genome plot showing RNA Pol II (pSer2) enrichment in two biological repeats of untreated and TSA treated (400nM, 6h) samples. **B:** the effect of TSA treatment shown, calculated from subtracting the untreated data away from the TSA treated data. Probes enriched in the treated sample are shown in red along the x-axis.



**Figure 4.18. Chromosome 17 genes enriched for RNA Pol II (pSer2) correlate with those identified in different cell lines.** Venn diagram showing significant overlaps ( $P < 0.01$ ) between chromosome 17 genes enriched for: RNA Pol II (pSer2) in TK6 cells, un-phosphorylated RNA Pol II in H9 cells and RNA Pol II (pSer5) in HepG2 and U87 cells. Publicly available data retrieved from the NCBI GEO repository under the series GSE39312 for H9 cells and GSE18499 for HepG2 and U87 cells. Genes in TK6 cells were considered enriched for RNA Pol II (pSer2) when two or more peaks were detected within the open reading frame. Enriched peaks for data retrieved from GEO were presented in BED (browser extensible data) format. The coordinates were converted to UCSC gene names using the table browser function from UCSC Genome Browser (available at <http://genome-euro.ucsc.edu/>). Overlaps were identified based on matching gene names using a web-based tool (available at <http://bioinformatics.psb.ugent.be/webtools/Venn/>) and drawn using R-based 'VennDiagram' package v1.6.17. Statistical analysis (Hypergeometric test) was performed using the R-based 'stats' package v3.2.4.

#### 4.10 Discussion

The aim of this chapter was to establish the ChIP-chip method for TK6 cells as a genome-wide tool for assessing the mechanisms of genotoxicity imposed by HDAC inhibitors. Initially, it was important to validate the wild type status of p53 in the TK6 cells used in this study for the purposes of confirming a normal DNA damage response following TSA treatment. This was performed by taking advantage of the unique effects of nutlin-3, an MDM2 inhibitor, which only induces apoptosis in cells containing wild-type p53; evident in the marked increase in Annexin-V observed only in TK6 cells compared to either PC-3 cells lacking p53 or MDA-MB-232 cells with mutant p53 (Figure 4.1). Although unrelated tumour cell lines were used here to control for the lymphoblast cell line, the outcomes of the assay are primarily dependent on the *TP53* status rather than the cell type. The detection of Annexin-V using flow cytometry also allowed for simple and quick method of determining p53 status in different cell lines. These results are further supported with the detection of p53 protein expression (Figure 4.2).

Since the use of HDAC inhibitors proved unsuccessful in yeast, attempts to demonstrate its effects in human cells focused on TSA, a pan-HDAC inhibitor which has published genotoxic effects in TK6 cells (Olaharski et al. 2006). In addition, it was also important to establish a treatment regimen from the initial dose-response studies which would be used to examine the mechanisms of DNA damage induction by TSA using ChIP-chip. The capability of HDAC inhibitors to induce DNA damage has been studied in genotoxicity assays measuring its effects at 24h. Particularly for TSA, a significant increase in micronuclei formation and cells arrested in G1 is observed at 24h, following its treatment in TK6 cells (Olaharski et al. 2006). To reproduce this, the InVitro MicroFlow® assay was first validated using the known genotoxic compound Vinblastine (Figure 4.3A). However, treating TK6 cells with increasing concentrations of TSA for 24h showed a borderline genotoxic response for the top doses as the cytotoxicity response (both RCC and EMA negative cells) was near the limit of 50% (Figure 4.3B). Therefore, the results of the InVitro MicroFlow® assay did not conclusively demonstrate the genotoxicity of TSA in TK6 cells at 24h. While an alternative approach would be to conduct the same experiment using the conventional microscopy assay, given the highly cytotoxic nature of HDAC inhibitors it is possible that longer treatment times may result in cytotoxicity-induced genotoxicity. This can potentially lead to confounding results

when studying the mechanisms of TSA that directly induce DNA damage. In light of this, attention was directed at using shorter treatment times. Since the micronuclei endpoint requires cell division to take place, a shorter TSA treatment would entail a prolonged incubation period for cell division to occur. During this time, it is perceivable that some of DNA damage resulting from drug treatment may be repaired, and so a treatment regimen based on this would not be suitable for analysis on the microarrays. Although DNA repair inhibitor can be added to prevent this, since the type of DNA damage elicited by HDAC inhibitors are unknown combined with the possibility that inhibitors themselves can be toxic to the cells would only complicate the results of the assay. Therefore, instead of using the micronuclei endpoint as a marker of DNA damage for short TSA treatments,  $\gamma$ H2AX, which is another well-characterised epigenetic marker of DNA damage was used (Kuo LJ 2008). Moreover, the ability of TSA to induce early (within hours) induction of histone acetylation and  $\gamma$ H2AX in TK6 cells were also described by Olaharski and others, measured using western blotting (Olaharski et al. 2006). To show these effects in the current study, flow cytometry (Chapter II, 2.1.4.2) was used as a faster alternative assay to western blotting. TK6 cells treated with a fixed dose of 400nM TSA showed a significant increase in global H3K9ac, which coincided with significant increases in global  $\gamma$ H2AX at 6h and 12h (Figure 4.5). To examine if the  $\gamma$ H2AX response is due to TSA and independent of cytotoxicity, markers of early apoptosis and changes in cell cycle were monitored using flow cytometry (Chapter II, 2.1.4.1; 2.1.4.4). No significant increase in either annexin-v or cells arrested in G<sub>1</sub> were observed (Figure 4.6) and therefore collectively shows that TK6 cells treated with 400nM TSA for 6h displays an effect at the epigenetic level, inducing markers of DNA damage that is independent of cytotoxicity. These observations are also consistent with those published showing the ability of HDAC inhibitors to induce early DNA damage and histone acetylation (Leea J.H 2010; Olaharski et al. 2006; Zhou et al. 2000). Under these conditions, the treatment regimen established will be used to study the mechanisms of TSA-induced DNA damage using the genome-wide method ChIP-chip.

Successful application of ChIP-chip for TK6 cells entails optimisation in key areas of the workflow such as chromatin preparation, immunoprecipitation and post-DNA amplification. To first optimise the conditions for preparing chromatin from TK6 cells a series of parameters such as fixation time, cycle number and cell densities were examined. Gel electrophoresis was used to assess the impact on chromatin shearing following changes made to each parameter (Figure 4.7 – 4.11). Of particular importance was the

proper use of a lysis buffer (Figure 4.8) and the appropriate cell density (Figure 4.10) which proved critical in shearing chromatin from TK6 cells to the desired size, an important factor in achieving a good resolution of the signal on the microarrays. Such extensive modifications to the standard protocol may be due to the small size of lymphoblastoid cells having ‘coarser’ chromatin with greater tendencies to aggregate (Kumar 2010). Blood-derived suspension cells such as monocytes and T/B lymphocytes have also been reported as being ‘difficult to shear’ (Diagenode). In conclusion, efforts to optimise the ChIP protocol for use with TK6 cells were successful and are described fully in Chapter II 2.1.7.

In eukaryotes, RNA polymerase II (pSer2) elongation is well studied and found at housekeeping genes, with a distinct distribution around actively transcribed genes, primarily occupying the open reading frame and 3-prime ends of genes. For this reason, the chromatin immunoprecipitation procedure for TK6 cells was assessed by using ChIP-qPCR of RNA Pol II (pSer2) enrichment at gene loci where it is known to be present. RNA Pol II (pSer2) enrichment at positive control genes in comparison to negative control genes, demonstrates that the ChIP protocol can detect abundant protein-DNA binding events (Figure 4.12A) and that these relative differences can still be observed following whole genome amplification (Figure 4.12B). This data also provided early assurances of the validity of the amplification method, since amplifying small amounts of DNA could potentially introduce amplification bias. For amplification of human ChIP DNA, the LM-PCR method used for yeast was substituted with the GenomePlex<sup>®</sup> Complete Whole Genome Amplification kit (WGA2) from Sigma-Aldrich. While both methods seek to amplify DNA representative of the starting material, LM-PCR ligates the same linker sequences to all fragments and PCR amplifies using linker specific primers (Agilent). WGA2 on the other hand, uses a set of universal oligonucleotide adapters, annealing randomly and uniformly to all fragments, generating an OmniPlex ‘library’ which is then PCR amplified (Sigma-Aldrich 2006). Previous investigation in the lab concluded that LM-PCR performed better using yeast samples while WGA2 performed better with human sample (Powell 2014a). O’Geen and colleagues also reported the superior signal to noise ratio observed using the WGA2 kit over LM-PCR for amplifying human DNA in a study using microarray to detect rare transcription factor binding events (O’Geen et al. 2006). For these reasons the WGA2 kit was used in this study for subsequent microarray work with TK6 cells.

For untreated and TSA treated (400nM, 6h) TK6 cells, quality controls were carried out to ensure that the data can be normalised together using Sandcastle (Figure 4.14). Additionally, a linear relationship for RNA Pol II (pSer2) occupancy was shown for each of the two biological repeats on chromosome 17 (Figure 4.15), highlighting the reproducibility of the arrays. The finding that post-hybridisation washes of the arrays accounts for the majority of variability detected in technical rather than biological replicates (Figure 4.15) suggested that other experimental actors were accountable. It has been noted that this can be caused by the fluctuating levels of ozone (O<sub>3</sub>) which can degrade cyanine 5 (Cy5)-labelled IP samples. This issue has also been addressed in gene expression arrays studies employing the same two-dye system using cyanine 3 and cyanine 5 (Branham et al. 2007). To rectify this in the present study, biological repeats of the same ChIP experiment were performed separately from different cell batches but entered the hybridisation workflow on the same day and improvements to reproducibility were evident in the increase in correlation (Figure 4.15C; D).

To validate the patterns of RNA Pol II (pSer2) observed on chromosome 17 arrays, the average signal distribution was plotted around gene structure, which showed the expected pattern of RNA Pol II (pSer2) enrichment inside and at the ends of genes (Figure 4.16). In addition, changes in RNA Pol II (pSer2) activity following TSA treatment at *TP53* indicated an increase in occupancy (Figure 4.17), consistent with previously predicted (Juan et al. 2000) and observed (Sonnemann et al. 2014) data. Comparing genes enriched for RNA Pol II (pSer2) on chromosome 17 detected in TK6 cells with those in 3 other cell lines showed a significant overlap ( $P < 0.01$ ), suggesting a similar set of genes may be transcribed. A greater number of genes showing RNA Pol II (pSer2) enrichment were identified in TK6 cells by comparison (Figure 4.18). This could be a reflection of the different data normalisation, peak calling algorithms and antibodies used. Affymetrix tiling arrays were used to generate ‘initiating’ RNA Pol II (pSer5) data for U87 and HepG2 cells (Barrett et al. 2013) analysed using the ‘Model-based Analysis of Tiling arrays’ (MAT) algorithm where probe values are standardised by cross-referencing sequence and copy number on each array, avoiding the need for sample normalisation (Johnson 2006). In contrast, non-phosphorylated RNA Pol II was mapped in H9 cells using Nimblegen arrays and data processed using the NimblegenScan software (Barrett et al. 2013). Along with the protocol presented in this chapter and data analysis using the novel ‘Sandcastle’ approach, these disparities are likely to result in the



differences in enrichment observed. Nevertheless, there is a statistically significant overlap in genes regarded as enriched for RNA Pol II between the different cell lines.

The results of this chapter demonstrate that while the use of HDAC inhibitors were ineffective in budding yeast, their chromatin associated effects can be clearly seen in the human TK6 lymphoblastoid cell line. Following this, a dose of TSA was selected for microarray analysis which showed the presence of DNA damage in the absence of cytotoxicity. Finally, a working ChIP-chip protocol was established for TK6 cells and validated using RNA Pol II (pSer2) through a number of different methods that showed a strong agreement with the expected outcomes. This allows for further measurements of changes in histone acetylation and replication to be made in response to TSA at a high resolution on chromosome 17 in the succeeding chapter, with focus on examining the mechanisms of HDACi-induced DNA damage.



## **Chapter V**

### **Investigating transcription-replication conflicts as a source of HDAC inhibitor induced DNA damage in TK6 cells**



## 5.1 Eukaryotic replication origins

The coordinated firing of replication origins is key to ensuring stability and proper duplication of the genome. In eukaryotes, the origin recognition complex (ORC) is required for initiating DNA replication and consists of six highly conserved subunits (ORC1 – 6) (Duncker BP 2009). Its recruitment precedes the formation of the pre-replication complex (pre-RC) and serves to load additional factors such as the putative Mcm2-7 helicase (Duncker BP 2009). In budding yeast, ORC shows a high affinity for consensus 11bp ‘A-elements’ known as autonomously replicating sequences (ARSs) in addition to 10 – 15bp ‘B-elements’ (Bell 2002). Conserved basic residues (Lys-362 and Arg-367) on the ORC1 subunit was recently shown to function in origin recognition (Kawakami et al. 2015). This is in contrast to fission yeast in which ORC recognises AT-rich sequences lacking a conserved motif (Clyne and Kelly 1995). In humans, ORC1 is ubiquitinated and dissociates from chromatin following entry into S-phase (Li and DePamphilis 2002), while other ORC subunits (2 – 5) remain intact. In this way, the binding of ORC1 is thought to confer selection of active origins (DePamphilis ML 2006). Recent work by Dellino and colleagues supported this idea by showing that ORC1 sites identified in HeLa cells using ChIP-seq correlated with actively replicating DNA, measured by repli-seq (Dellino et al. 2013). Additional efforts aimed at characterising replication origins in greater detail involve a range of genome-wide approaches such as the sequencing of short nascent DNA strands (SNS-seq) (Besnard et al. 2012), transcription bubbles (bubble-seq) (Mesner et al. 2013), Okazaki fragments (OK-seq) (Kahli P.N. 2016) and *in vitro* initiating sites (ini-seq) (Langley A.R. 2016). Nevertheless, a lack of correlation between datasets generated thus far using the various methods paints an obscure picture of replication initiation in humans. Additionally, it appears that origin specification in human cells do not depend on a consensus sequence compared to yeast (Vashee et al. 2003). A study comparing different cell lines showed that replication origins are not randomly distributed but rather utilised at different frequencies and may be defined by unique replication programs established during development for that particular cell type (Besnard et al. 2012). On top of this, studies in human and hamster cells indicate that ORC can be re-established at different sites following each round of cell division (Dimitrova and Gilbert 1999; Gerhardt et al. 2007), further complicating their identification. In this respect, initiation of replication is described to occur within

specific ‘zones’, estimated to range from 2.6 – 150kb or greater (Kahli P.N. 2016; Crosetto et al. 2013)

The early pioneering work of Huberman and Riggs, combining pulse labelling and DNA radiography, identified around 40,000 potential origins sites in the human genome (Huberman J.A. 1968). Current data from short nascent DNA strand sequencing places the estimate at around 120,000 to 250,000 (Mesner et al. 2013; Kahli P.N. 2016). Studies show that the numbers of established origins are far in excess of those used at any given time, giving rise to ‘dormant’ and ‘active’ origins. Dormant origins remain silent in normally dividing cells and can be activated in response to replication stress (Yekezare, Gomez-Gonzalez, and Diffley 2013). Such mechanisms help compensate for the timely completion of replication during S-phase. While the exact mechanisms of origin utilisation are still unknown, emerging evidence suggests that DNA structure; such as ssDNA and G-quadruplex (G4) motifs (tertiary DNA structures that form in guanine-rich regions) may have a role in origin selection (Hoshina et al. 2013). Additionally, replication initiation correlates with regions of open chromatin (Mesner et al. 2013), which favours the formation of G4 structures and is commonly characterised by susceptibility to DNase I cleavage (DNase I hypersensitive sites, DHS) or nucleosome free regions (NFRs) (Dellino et al. 2013). The latter is also commonly found at the promoters of actively transcribed genes, in comparison to actively transcribed regions downstream, which tend to be relatively accessible (Livaja R. 2010).

## **5.2 Histone acetylation and replication origin firing**

There is a strong body of evidence linking histone acetylation/deacetylation to the organisation and firing of origins, through shaping of the chromatin landscape. Early replicating sites tend to correlate with euchromatin, whereas late replicating sites correlate with heterochromatin (Lygerou 2012). A more open structure of chromatin is likely to be more accessible to factors involved in replisome assembly. Since histone acetylation is a well-known post-translational modification linked to chromatin relaxation, changes in acetylation have been shown to affect origin activity. In yeast, the association of ORC with histone acetyltransferase Hat1p, which acetylates H4K5 and K12, provides a novel link between acetylation status and replication (Suter et al. 2007). Histone hyperacetylation is also essential in maintaining replication, as demonstrated in yeast

where mutations to bulk lysines on histone H3 and H4 tails severely compromised cell growth (Unnikrishnan, Gafken, and Tsukiyama 2010).

HDAC enzymes have also been shown to affect the distribution of replication origins. *RPD3* deletions in yeast alter replication timing by promoting the early loading of initiation factors and origin firing (Vogelauer et al. 2002). Similar evidence is also presented in human studies, where G<sub>1</sub> synchronised HeLa cells cultured in the presence of the HDAC inhibitor TSA progressed through S-phase faster than non-TSA treated cells, suggesting that late-firing origins may have been activated earlier (Kemp et al. 2005). Moreover, replication timing of the  $\beta$ -globin gene can be altered by promoting/suppressing HDAC activities locally (Kemp et al. 2005).

### **5.3 Chapter aims**

Published evidence shows that increases in histone acetylation induced by HDAC inhibitors alters transcription profiles and the organisation of replication sites in the genome, implicating these events in the generation of DNA damage (Chapter I, 1.8). To test the possibility that these changes may promote transcription-replication conflicts as a source of DNA damage in human cells, work in the current chapter will employ the ChIP-chip method established in TK6 cells (Chapter IV) to examine at high resolution, TSA-induced changes in markers of histone acetylation (Histone H4 acetylation), transcription (RNA Pol II phospho-Serine 2), replication (Origin replication complex 1) and extend this to markers of DNA damage such as  $\gamma$ H2AX and R-loops.

## **5.4 Measuring the effect of TSA on H4 acetylation, RNA Pol II (pSer2) and ORC1 occupancy on chromosome 17 in TK6 cells**

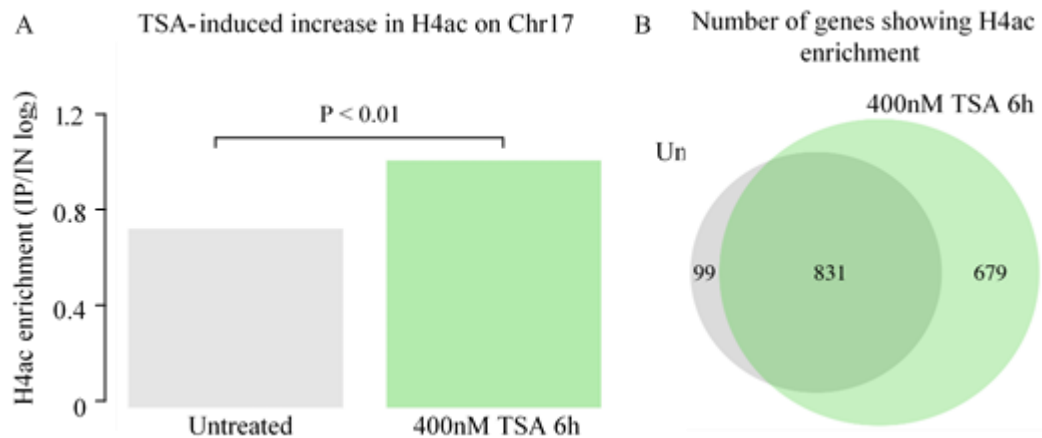
Using the previously identified dose of TSA (400nM, 6h), showing indications of DNA damage in the absence of cytotoxicity, ChIP-chip data is firstly generated to compare the overall effect that TSA has on histone H4 acetylation, transcriptional elongation and replication initiation on chromosome 17 in TK6 cells.

### **5.4.1 Examining changes in Histone H4 acetylation in response to TSA**

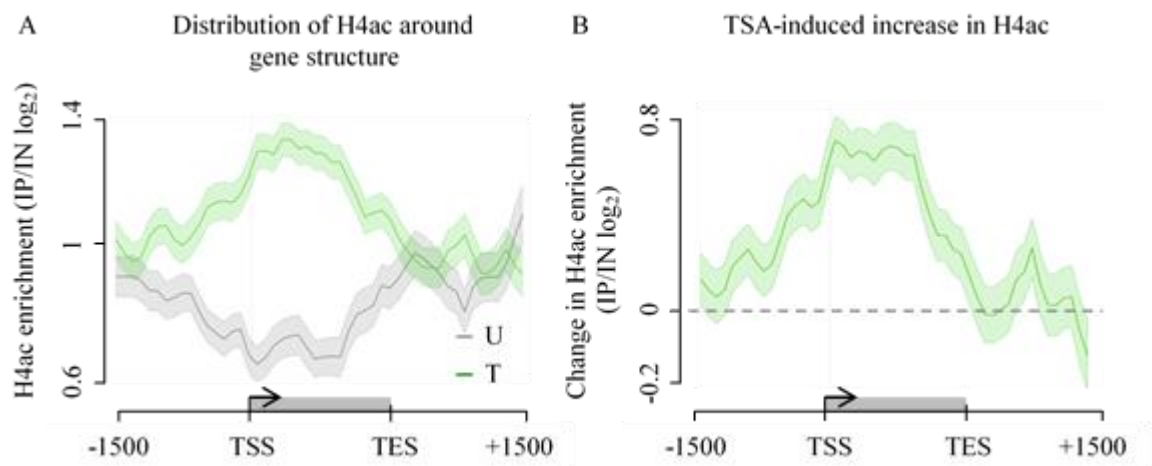
HDAC inhibitors increases global levels of histone acetylation, which is thought to promote a more open chromatin structure allowing association of transcription and replication factors. Therefore, it is expected that transcription and replication may be more active in regions of high acetylation. To study this in greater detail, changes in histone H4 acetylation were measured on chromosome 17 using previously designed microarrays (Chapter II, 2.1.7.7) as a marker of euchromatin, in response to the pre-determined TSA treatment (Chapter IV, 4.8). Different amounts of the anti-pan H4 acetylation antibody (K5, K8, K12, K16) were firstly titrated with chromatin prepared from TK6 cells to identify the optimal amount and to show that the antibody is suitable for use in ChIP (Appendix, III). ChIP-chip was carried out using established conditions (Chapter IV) with untreated and TSA treated (400nM, 6h) TK6 cells. Quality control was carried out as before (Chapter IV, 4.9.5) to show that the histone H4 acetylation (H4ac) array data from untreated and TSA treated cells can be normalised together (Appendix, III).

To first examine the overall effect of treating with TSA, the average H4ac signal was compared between untreated and TSA treated cells. To demonstrate this, firstly the mean IP/IN  $\log_2$  ratios are calculated for each condition and plotted as a histogram to compare the net change. Figure 5.1A shows that TK6 cells treated with TSA exhibit significantly higher level of H4ac on chromosome 17 ( $P < 0.01$ ). Next, to gain a better idea of the extent of the increase observed, the number of gene regions enriched for H4ac before and after TSA treatment was compared. The venn diagram shows as a result of TSA treatment, a large number of genes are hyperacetylated on histone H4 (Figure 5.1B). To examine where the increase in H4ac signal occurs in these gene regions, the average H4ac enrichment measured before and after TSA treatment was plotted +/- 1500bp





**Figure 5.1. TSA increases histone H4 acetylation on chromosome 17.** **A:** average H4ac enrichment calculated from normalised array data from untreated (mean = 0.73, SEM = 0.002) and TSA treated (mean = 1, SEM = 0.002) cells showing a significant increase in H4ac following treatment ( $P < 2.2 \times 10^{-16}$ , Welch's t-test). Small SEMs are due to the large sample size (~400,000) and are omitted from the plot.  $n = 3$ . **B:** a large number of genes show H4ac enrichment in response to TSA treatment. Diagrams drawn using R base package v3.2.4 and the 'VennDiagram' package v1.6.17.



**Figure 5.2. TSA increases histone H4 acetylation in the ORFs of genes in chromosome 17.** **A:** pattern of H4ac plotted around gene structure in untreated (*U*: grey) and TSA treated (400nM, 6h) TK6 cells (*T*: green). **B:** TSA increases H4ac primarily around ORFs of genes, calculated by subtracting untreated data away from TSA treated data. SEM plotted as error shapes around the average of 3 biological replicates.

around the open reading frame (ORF) of genes on chromosome 17. The result is a profile plot showing that in untreated TK6 cells, H4ac is enriched upstream of the TSS and downstream of the TES. However, following TSA treatment, an increase in H4ac is observed in the ORFs of genes (Figure 5.2A). To identify the net change in distribution of H4ac, the data from untreated cells was subtracted from the data generated from TSA treated cells, showing that TSA promotes H4 hyperacetylation mainly in the ORFs of genes on chromosome 17 (Figure 5.2B).

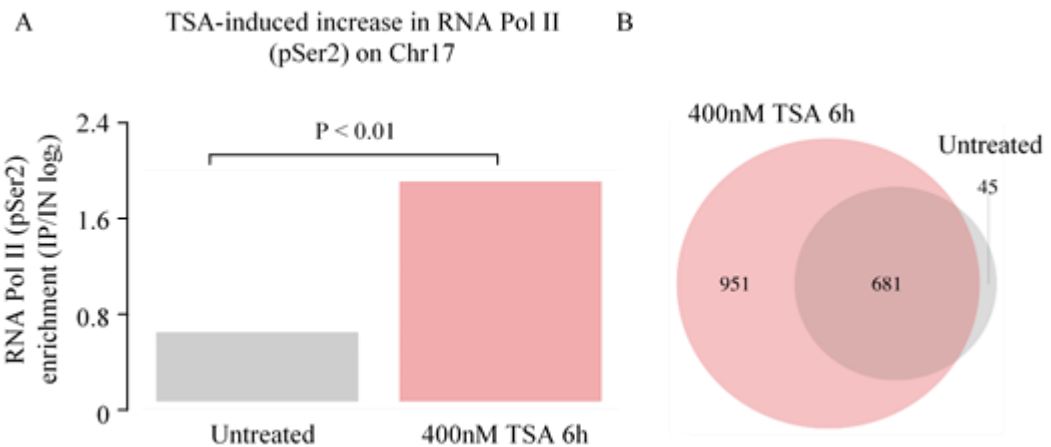
#### **5.4.2 Examining changes in transcription in response to TSA**

Following this, to inspect how TSA treatment affects normal transcription in TK6 cells, RNA Pol II (pSer2) ChIP-chip data generated previously to validate the established technique (Chapter IV) was examined here in greater detail. Comparison of the data shows that TSA significantly increases the levels of RNA Pol II (pSer2) enrichment in treated TK6 cells ( $P < 0.01$ ) (Figure 5.3A) and is detected in a large number of genes previously un-enriched in untreated TK6 cells (Figure 5.3B). The enrichment of RNA Pol II (pSer2) is higher overall following TSA treatment (Figure 5.4A), with the greatest increase observed up- and down-stream of the TSS and TES, respectively (Figure 5.4B).

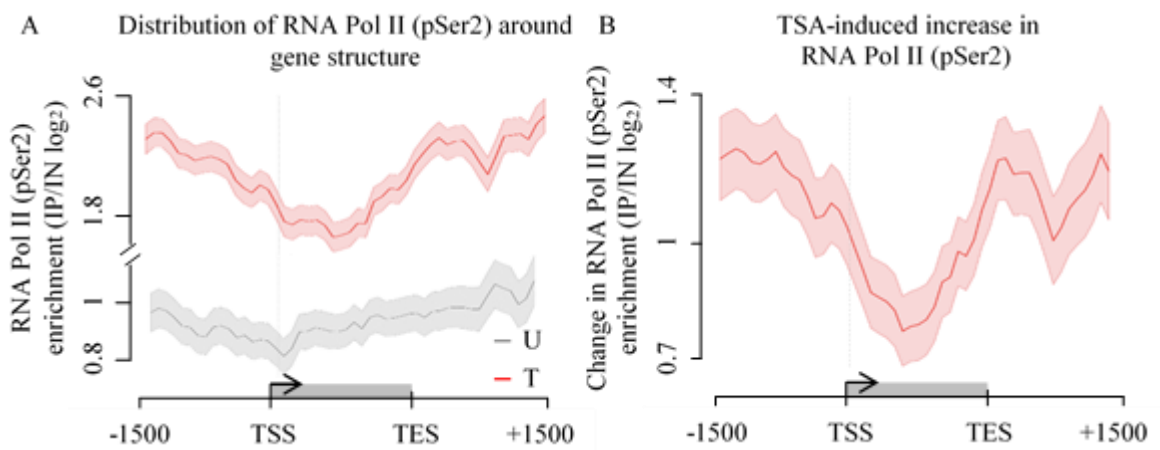
#### **5.4.3 Examining changes in replication initiation in response to TSA**

To characterise replication, ORC1 was selected to mark zones of replication initiation (Duncker BP 2009). Since replication origins have not been identified in TK6 cells and these reportedly vary between different cell types, the specificity of the anti-ORC1 antibody was tested by performing a western blot of the denatured ChIP sample (Chapter II, 2.1.6.5). The observed single band at the expected molecular weight indicates that the antibody is highly specific for ORC1, justifying its use in ChIP-chip (Appendix, III).

Quality control was performed on the ORC1 ChIP-chip data as before to show that the normalisation procedure could be applied to the data from untreated and TSA treated cells (Appendix III). Since previous attempts by Dellino and colleagues reported difficulties in enriching for ORC1 using whole cell chromatin (Dellino et al. 2013), the validity of the data generated here was first assessed by comparing replicates of the ORC1



**Figure 5.3. TSA increases RNA Pol II (pSer2) enrichment on chromosome 17.** **A:** average RNA Pol II (pSer2) enrichment calculated from normalised data from untreated (mean = 0.7, SEM = 0.003) and TSA treated (mean = 2.21, SEM = 0.003) cells showing a significant increase following treatment ( $P < 2.2 \times 10^{-16}$ , Welch's t-test). Small SEMs are due to the large sample size (~400,000) and are omitted from the plot.  $n = 3$ . **B:** a larger number of genes show RNA Pol II (pSer2) enrichment in response to TSA treatment. Diagrams drawn using R base package v3.2.4 and the 'VennDiagram' package v1.6.17.



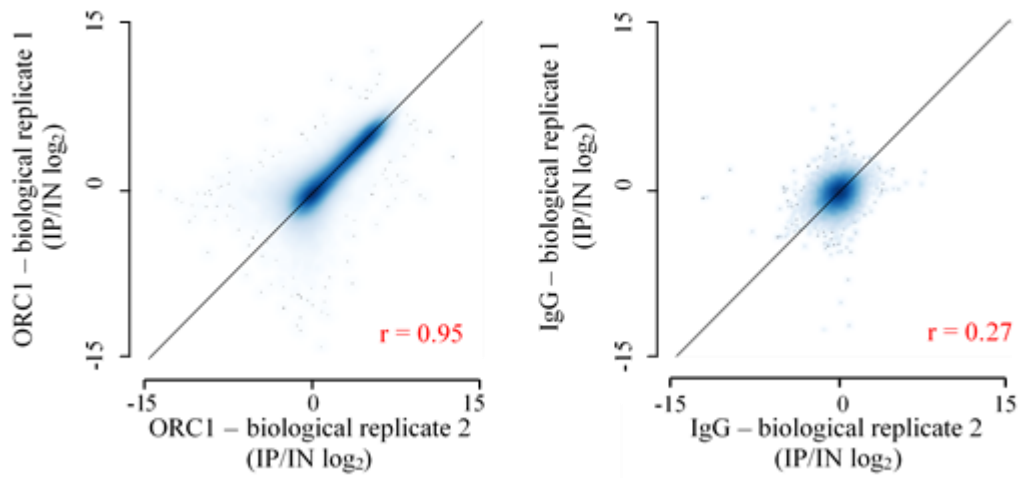
**Figure 5.4. TSA increases RNA Pol II (pSer2) enrichment outside the ORFs of genes on chromosome 17.** **A:** pattern of RNA Pol II (pSer2) enrichment plotted around gene structure in untreated (*U*: grey) and TSA treated (400nM, 6h) TK6 cells (*T*: red). **B:** TSA increases RNA Pol II (pSer2) enrichment primarily upstream of TSS and downstream of TES, calculated by subtracting untreated data away from TSA treated data. Standard deviation error shapes plotted around the average of 2 biological replicates.

microarray data from untreated TK6 cells with that of a negative control (background binding data) generated using a species-specific anti-IgG antibody (Figure 5.5). Consequently, biological replicates of ORC1 microarray data showed a high level of correlation (Pearson's correlation  $r = 0.95$ ), compared to the biological replicates of IgG microarray data (Pearson's correlation  $r = 0.27$ ) (Figure 5.5), which failed to detect peaks and do not display any pattern of enrichment (Figure 5.6). This suggests that the ORC1 microarray data generated here is highly reproducible and is representative of putative ORC1 sites found on chromosome 17 in TK6 cells.

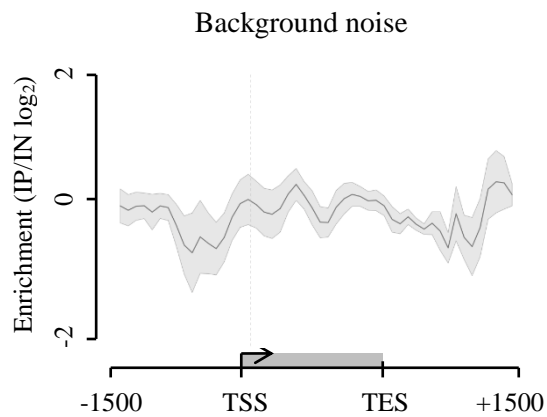
Chromatin structure is known to play an important role in defining how transcription and replication are organised in cells. Relaxed chromatin regions and G-quadruplex structures have emerged as reliable predictors of replication initiation zones (Hertsch R.H. 2016; Kahli P.N. 2016) To examine if the replicating zones identified here map to these locations, ORC1 microarray data is plotted around DNase I hypersensitive sites from the ENCODE database identified in human K562 (myelogenous leukemia) cells (Dunham et al. 2012), which represent regions of open chromatin susceptible to DNase I cleavage, in addition to G4-sites recently mapped by sequencing in HaCat (human keratinocyte) cells (Hertsch R.H. 2016). Figure 5.7 shows that the ORC1 data generated from untreated TK6 cells map to both DNase I hypersensitive sites and G4 sites, in comparison to a set of randomly generated sites showing no particular pattern. This demonstrates that ORC1 sites identified in untreated TK6 cells correlate with open chromatin regions and G4 sites, which are predictive of replication initiation zones.

Closer inspection of the normalised data shows that in untreated TK6 cells replication initiating zones are distributed downstream of the TSS, in the ORFs of genes (Figure 5.8A). This is consistent with previously published studies also showing replication origins to be enriched downstream of the TSS (Dellino et al. 2013; Martin et al. 2011). However, in response to TSA treatment, these initiating zones are altered and re-distributed upstream of the TSS and downstream of the TES (Figure 5.8B). The net effect is a small but significant increase in ORC1 occupancy on chromosome 17 of TK6 cells treated with TSA (Figure 5.9).

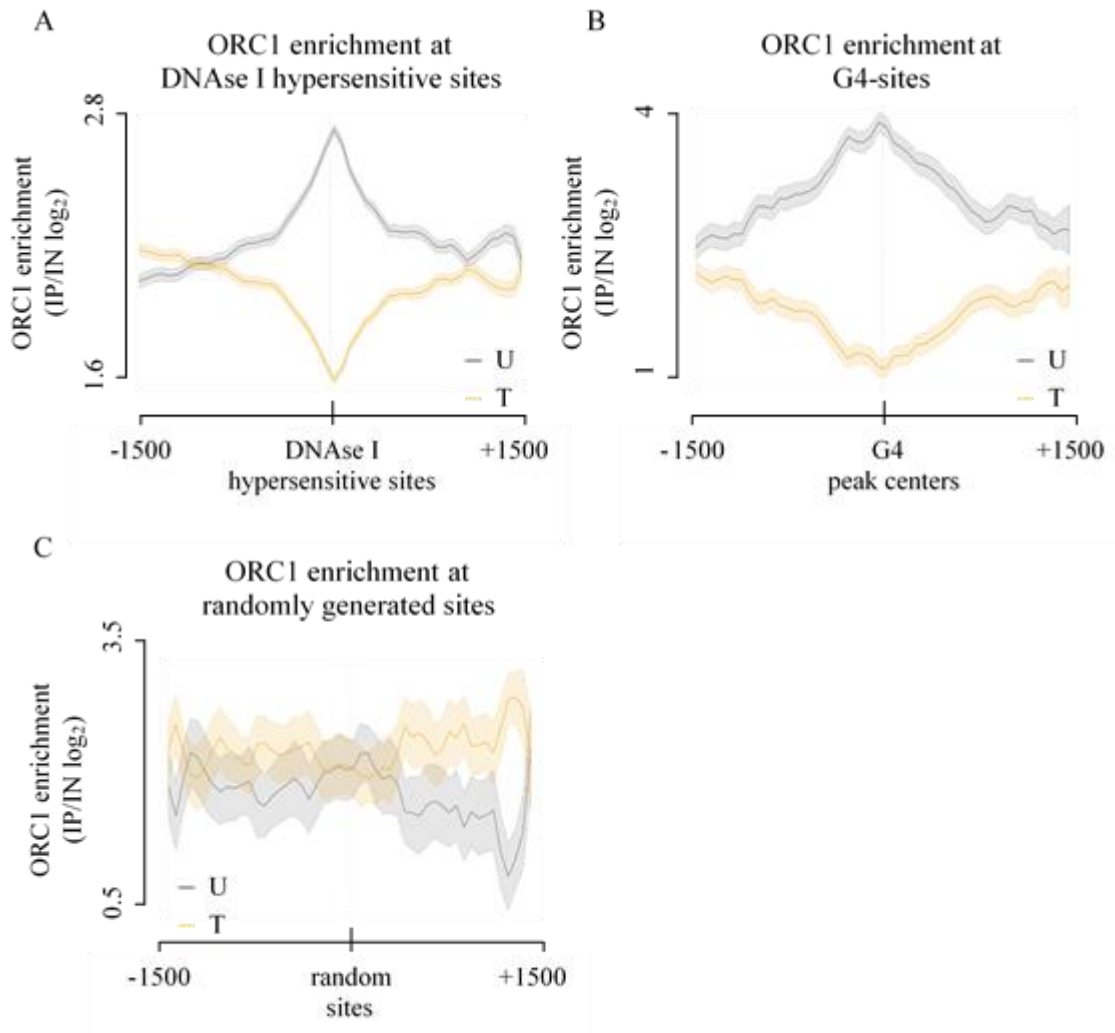
In summary, these results show that treating TK6 cells with the pan-HDAC inhibitor TSA alters the distribution, as well as significantly increasing the levels of histone H4 acetylation, RNA Pol II (pSer2) and ORC1 enrichment on chromosome 17.



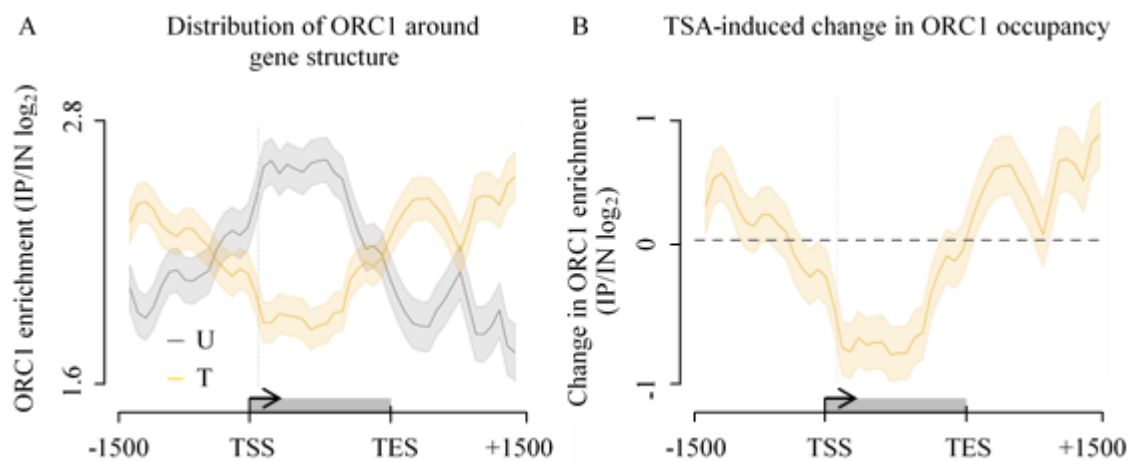
**Figure 5.5. Biological replicates of ORC1 microarray data display a strong positive correlation compared to background IgG microarray data.** A: smooth scatter plot comparing two biological repeats of ORC1 microarray data generated in untreated TK6 cells, showing a strong positive correlation. B: Smooth scatter plot comparing two biological repeats of data generated using a species-specific IgG antibody as background control, showing a weak positive correlation. Pearson's correlation calculated using functions in the base R package.  $x = y$  line shown in black.



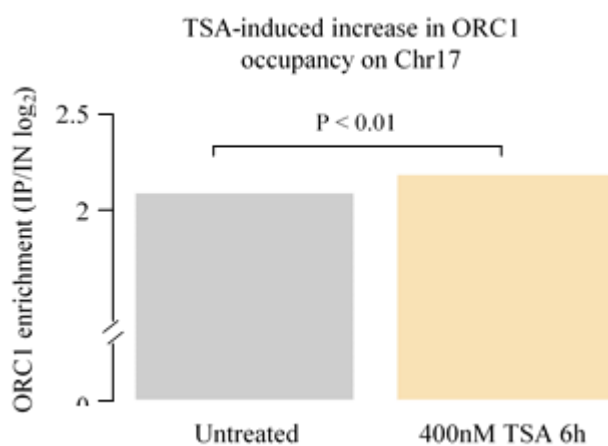
**Figure 5.6. Background IgG microarray data does not display enrichment.** Normalised IgG data plotted around gene structure.



**Figure 5.7. ORC1 occupancy on chromosome 17 in untreated TK6 cells consistently map to sites predictive of replication initiation.** ORC1 data from untreated (*U*: grey) and TSA treated (*T*: yellow) TK6 cells plotted around **A**: DNase I hypersensitive sites identified in K562 cells (ENCODE: ENCSR527IGO) **B**: G4-sites identified in HaCat cells (NCBI GEO: GSE76688) and **C**: list of randomly generated site on chromosome 17 ( $n = 1000$ ) using the BEDTools extension in Galaxy (Afgan et al. 2016). SEM plotted around the average of 3 biological repeats.



**Figure 5.8. TSA treatment re-distributes ORC1 occupancy away from the ORFs of genes on chromosome 17.** **A:** pattern of ORC1 occupancy plotted around gene structure in untreated (*U*: grey) and TSA treated (400nM, 6h) TK6 cells (*T*: yellow). **B:** TSA reduces ORC1 occupancy within the ORFs of genes, calculated by subtracting the array data from untreated TK6 cells away from TSA treated TK6 cells. SEM plotted as error shapes around the average of 3 biological replicates.



**Figure 5.9. TSA treatment increases ORC1 occupancy on chromosome 17.** Average ORC1 enrichment calculated from normalised array data generated from untreated (mean = 2.11, SEM= 0.004) and TSA treated (mean = 2.3, SEM = 0.003) cells showing a significant increase in ORC1 occupancy following treatment ( $P < 2.2 \times 10^{-16}$ , Welch's t-test). Small SEMs are due to the large sample size (~400,000) and are omitted from the plot.  $n = 3$ . Diagram drawn using R base package v3.2.4.

## 5.5 Analysing the impact of TSA-induced changes in H4 acetylation on RNA Pol II (pSer2) and ORC1 occupancy on chromosome 17 in TK6 cells

An increase in histone acetylation is associated with higher levels of transcriptional and replication activity in cells. Therefore, to examine how changes in H4ac induced by TSA treatment might alter transcription and replication on chromosome 17 in TK6 cells their correlations will be examined. Scatter plots are a useful way of examining the overall change/relationship between two sets of array data. When the normalised array data is plotted in the form of a scatter plot, each probe value in the first dataset is matched with the corresponding value in the second dataset. In this way, the correlation of two datasets can be estimated using the Pearson's coefficient ( $-1 < r < 1$ ). A low  $r$  value would indicate that a weak relationship, whereas a high  $r$  value would indicate a strong relationship. Also,  $r$  can either be positive or negative, with positive representing a tendency of one variable to increase as the other increases and vice versa. Additionally, by smoothing the scatter plot, the distribution of data points on the graph can be easily discerned by observing the colour density. Darker regions (Figure 5.10) represent high concentration of data points whereas lightly coloured or white regions represent little or no data points respectively. Following a general analysis, the relationship shown in smoothed scatterplots can be examined in detail by plotting the same data in the context of a collection of binding sites in the form of a profile plot.

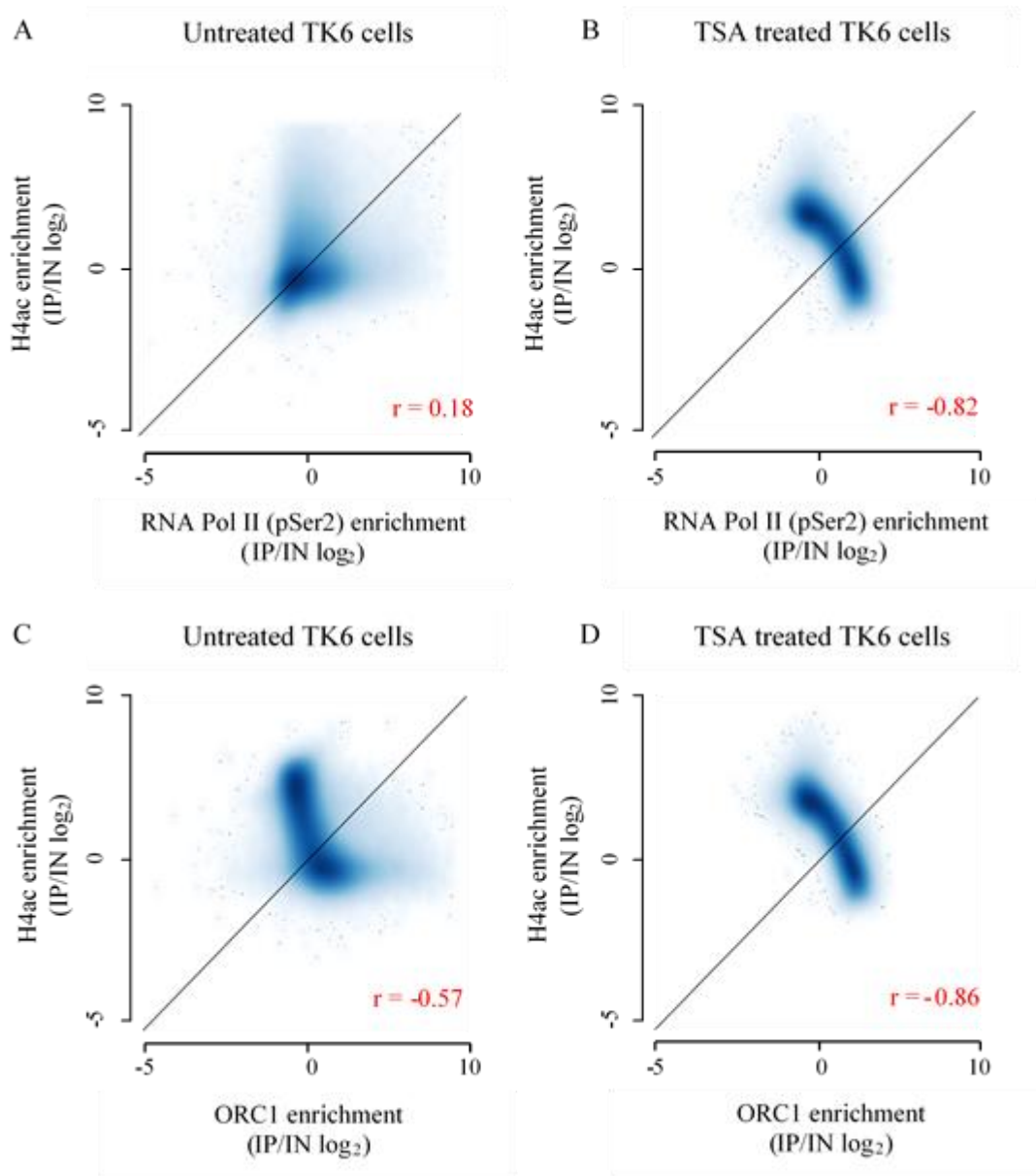
Following this method of analysis, to examine how TSA-induced changes in H4ac affect cellular processes such as transcription and replication on chromosome 17 in TK6 cells, correlations between the normalised H4ac data are compared to the normalised data of RNA Pol II (pSer2) and ORC1 independently in untreated and TSA treated conditions (Figure 5.10). In untreated TK6 cells, the H4ac data shows a weak positive correlation with RNA Pol II (pSer2) (Pearson's correlation  $r = 0.18$ ) and a negative correlation with ORC1 (Pearson's correlation  $r = -0.57$ ) (Figure 5.10A; C). Following TSA treatment, H4ac shows a strong negative correlation with both RNA Pol II (pSer2) and ORC1 (Pearson's correlation  $r < -0.8$ ) (Figure 5.10B; D). This suggests that TSA-induced changes in H4ac do not associate directly with the changes in RNA Pol II (pSer2) and ORC1 occupancy on chromosome 17. This is also consistent with the general changes in distributions observed previously, where TSA-induced H4ac hyper-acetylation around the ORFs of genes on chromosome 17 (Figure 5.2) is accompanied by enrichment of RNA Pol II (pSer2) (Figure 5.4) and ORC1 (Figure 5.6) outside of the ORFs.



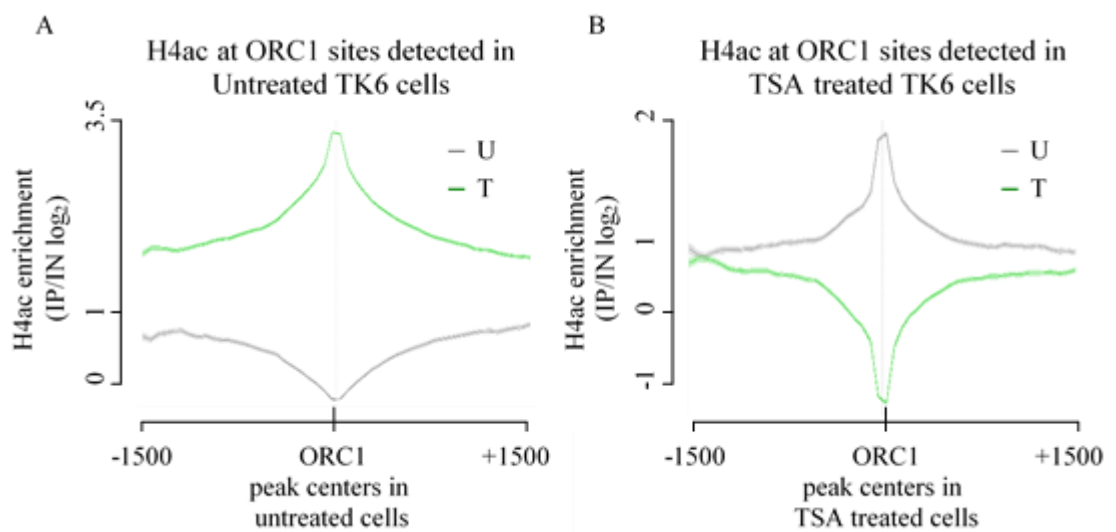
To investigate this in greater detail, the H4ac data from both untreated and TSA treated cells were plotted around ORC1 peak centres identified in untreated TK6 cells (Figure 5.11A). The results show that in untreated cells, H4ac is enriched around ORC1 peaks but not at ORC1 peaks (Figure 5.11A, *grey line*). However, in response to TSA, an overall increase in H4ac was observed which is most pronounced at ORC1 peak centres (Figure 5.11A, *green line*). Similarly, to examine the impact of TSA-induced changes H4ac on RNA Pol II (pSer2) occupancy, H4ac data from untreated and TSA treated cells were plotted in relation to RNA Pol II (pSer2) peaks identified in untreated and TSA treated cells, respectively. In contrast to the relationship between H4ac and ORC1, for RNA Pol II (pSer2) peaks detected in untreated cells, H4ac measured in untreated cells are mostly enriched at the centres of these peaks (Figure 5.12A, *grey line*). Conversely, following TSA treatment this pattern is inverted, resulting in the re-distribution of H4ac away from these sites (Figure 5.12A, *green line*). Comparing with RNA Pol II (pSer2) peaks detected in TSA treated cells, in the absence of TSA these sites show a high level of H4ac enrichment (Figure 5.12B, *grey line*). Subsequently, the reciprocal is observed in response to TSA and suggests that a loss of H4ac from these sites associates with the accumulation of RNA Pol II (pSer2) (Figure 5.12B, *green line*). Taken together, these results show that the TSA-induced changes H4ac correlates with the re-distribution of RNA Pol II (pSer2) and ORC1 observed following TSA treatment.

As evidence in the literature point towards the involvement of NFRs at promoter regions over transcribed regions (Livaja R. 2010), RNA Pol II (pSer2) data from untreated and TSA treated cells were plotted around the same DHS used earlier to examine ORC1 occupancy (Figure 5.7A). The results show that RNA Pol II (pSer2) is not enriched at these sites (Figure 5.13, *grey line*), confirming the published reports (Livaja R. 2010).

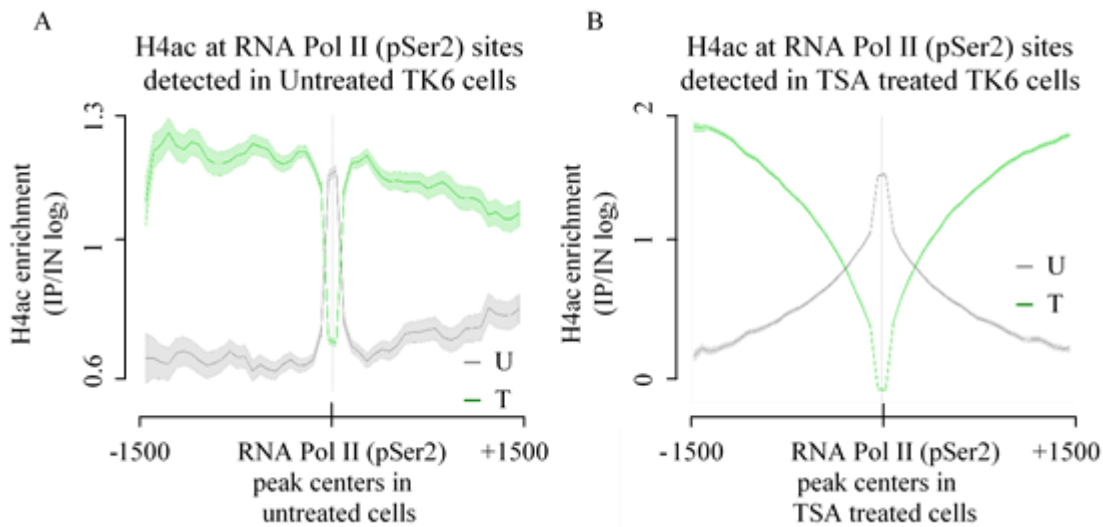
To gain an idea of where changes in transcription and replication are taking place in relation to the TSA-induced changes in H4ac, RNA Pol II (pSer2) and ORC1 enrichment were both plotted independently around genes showing H4 hyper-acetylation and H4 hypo-acetylation following TSA treatment (Figure 5.14). As a result, RNA Pol II (pSer2) and ORC1 are highly enriched at genes showing H4 hypo-acetylation in response to TSA (Figure 5.14A; B). By contrast, at genes showing H4 hyper-acetylation in response to TSA, RNA Pol II (pSer2) and ORC1 show reduced occupancy in the region of ORFs (Figure 5.14C; D). Interestingly, closer examination shows that TSA promotes the co-enrichment of RNA Pol II (pSer2) and ORC1 downstream of the TES at H4 hyper-



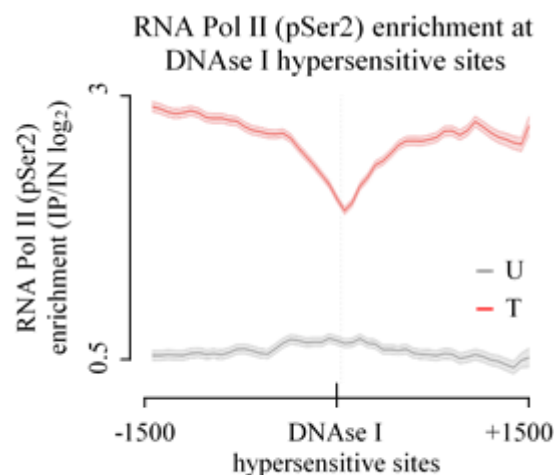
**Figure 5.10. TSA-induced changes in H4ac inversely correlates with transcription and replication.** Smooth scatterplots showing: **A:** a weak positive correlation between H4ac and RNA Pol II (pSer2) in untreated TK6 cells. **B:** a strong negative correlation between H4ac and RNA Pol II (pSer2) in TSA treated TK6 cells. **C and D:** H4ac and ORC1 do not correlate in both untreated and TSA treated TK6 cells, respectively.



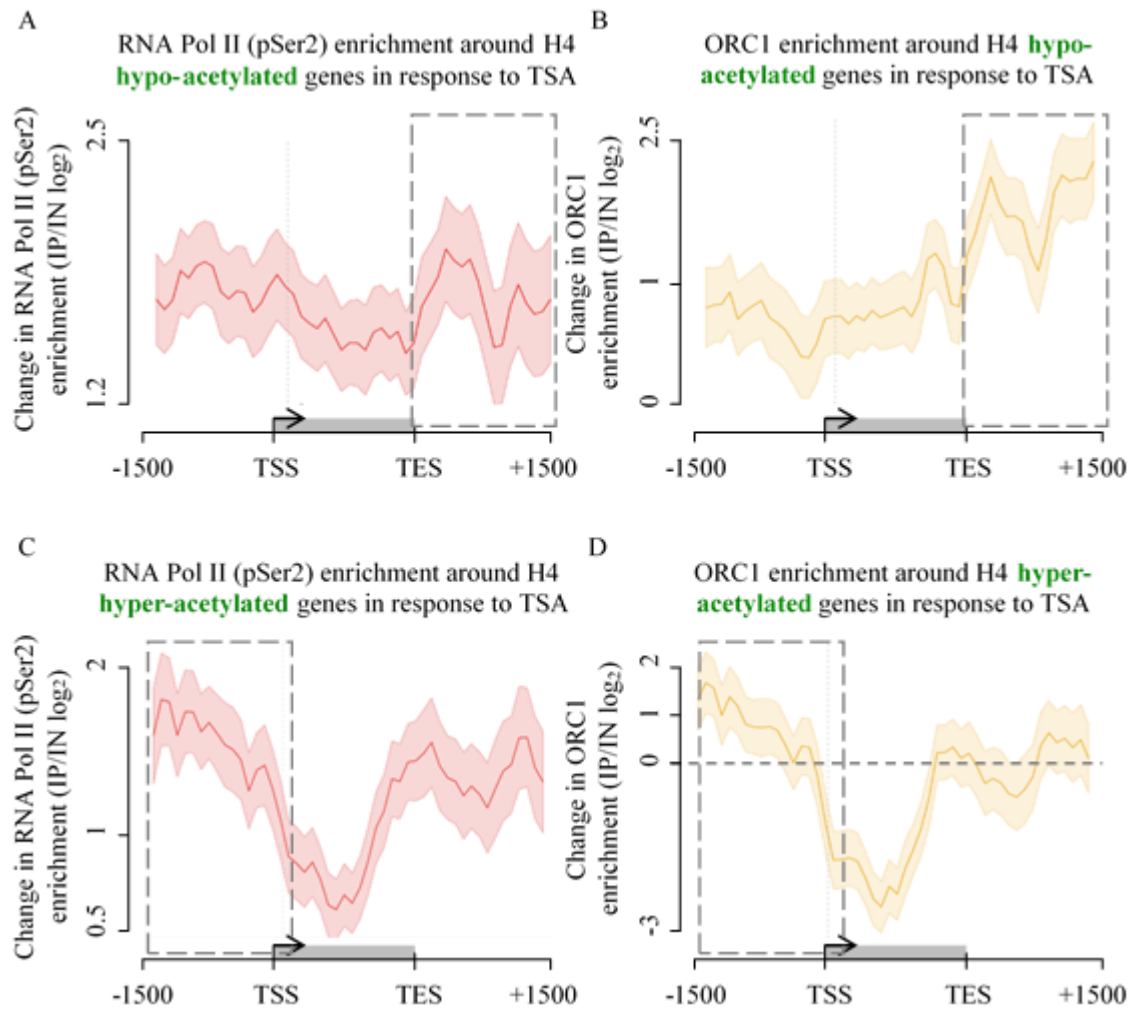
**Figure 5.11. TSA induces changes in H4ac at ORC1 sites in untreated and TSA treated TK6 cells.** **A:** H4ac data from untreated (*U*: grey) and TSA treated (*T*: green) TK6 cells plotted around ORC1 peaks identified in untreated TK6 cells, showing an increase in H4ac at these sites following TSA treatment and **B:** H4ac data from untreated (*U*: grey) and TSA treated (*T*: green) TK6 cells plotted around ORC1 peaks identified in TSA treated TK6 cells, showing decrease in H4ac at these sites following TSA treatment. SEM plotted as error shapes around the average of 3 biological replicates.



**Figure 5.12. TSA induces changes in H4ac at RNA Pol II (pSer2) sites in untreated and TSA treated TK6 cells.** **A:** H4ac data from untreated (*U*: grey) and TSA treated (*T*: green) TK6 cells plotted around RNA Pol II (pSer2) peaks identified in untreated TK6 cells, showing a decrease in H4ac at these sites following TSA treatment and **B:** H4ac data from untreated (*U*: grey) and TSA treated (*T*: green) TK6 cells plotted around RNA Pol II (pSer2) peaks identified in TSA treated TK6 cells, showing an increase in H4ac at these sites following TSA treatment. SEM plotted as error shapes around the average of 3 biological replicates.



**Figure 5.13. RNA Pol II (pSer2) occupancy does not correlate with DNase I hypersensitive sites.** RNA Pol II (pSer2) data from untreated (*U*: grey) and TSA treated (*T*: red) TK6 cells plotted around DNase I hypersensitive sites showing a lack of enrichment. DNase I hypersensitive sites identified in K562 cells were downloaded from the ENCODE database (ENCSR527IGO). Standard deviation error shapes plotted around the average of 2 biological replicates.



**Figure 5.14. TSA promotes specific-patterns of RNA Pol II and ORC1 co-enrichment at hyper- and hypo-H4ac genes.** **A** and **B**: TSA-induced H4 hypo-acetylated genes show an increase in RNA Pol II and ORC1, primarily downstream of the TES. **C** and **D**: TSA-induced H4 hyper-acetylated genes show an increase in RNA Pol II and ORC1 upstream of the TSS. TSS – transcription start site. TES – transcription end site. Standard deviation error shapes plotted around the average of 2 biological replicates of RNA Pol II data. SEM plotted as error shapes around the average of 3 biological replicates of ORC1 data.

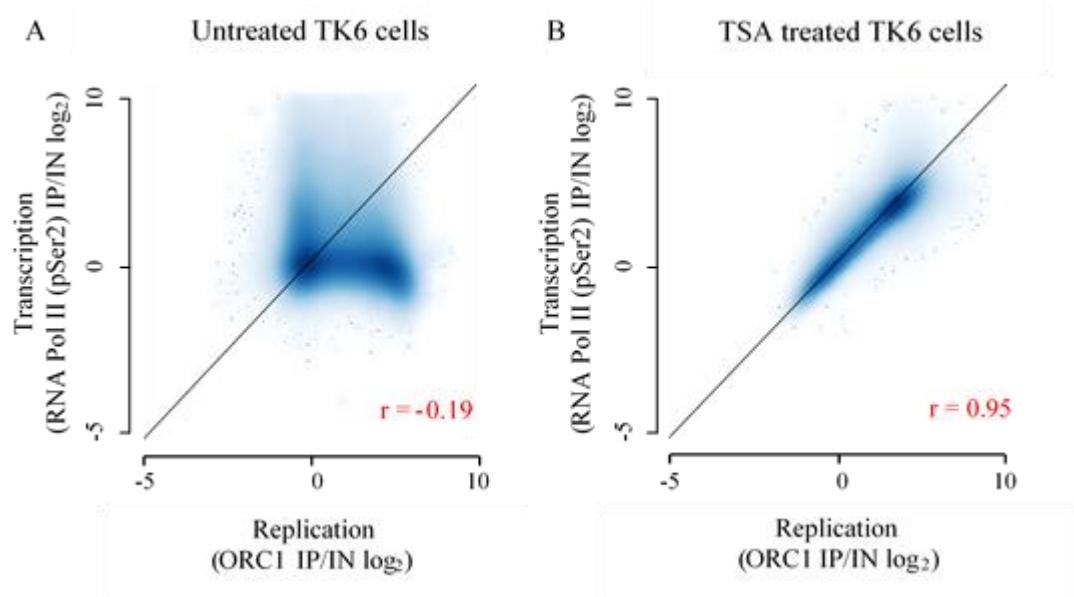
acetylated genes (Figure 5.14C; D, *dashed boxes*), whereas at H4 hypo-acetylated genes, they are both co-enriched upstream of the TSS (Figure 5.14C; D, *dashed boxes*). The results of the analysis therefore show that TSA can potentially promote the co-localisation of transcription and replication in certain regions, dependent on the hyper- or hypo-H4ac status of genes present on chromosome 17 in TK6 cells.

## **5.6 Examining TSA-induced co-localisation of transcription and replication on chromosome 17**

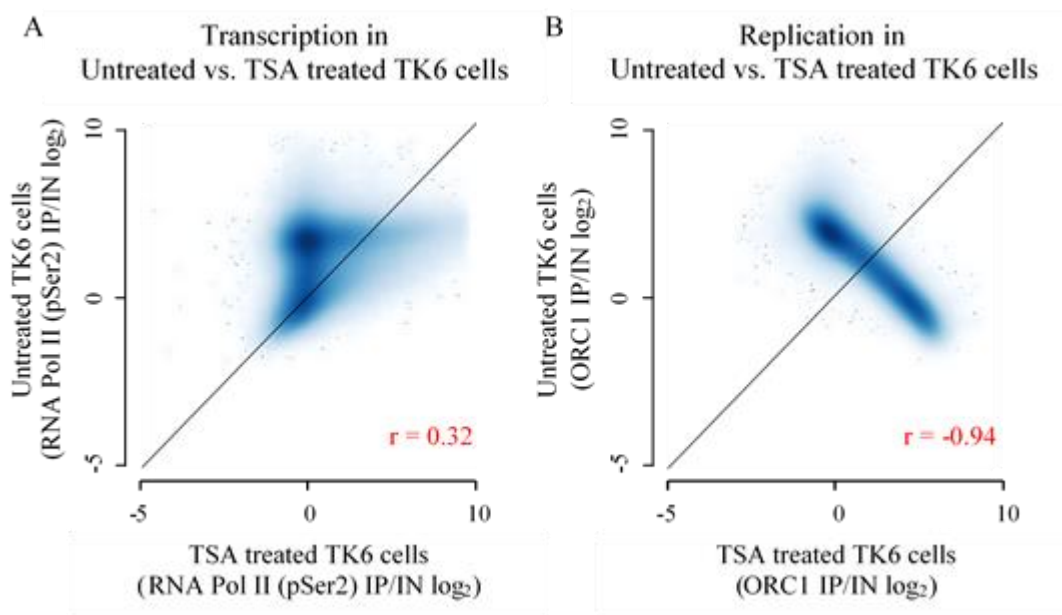
To determine if the predicted co-enrichment of RNA Pol II (pSer2) and ORC1 observed in response to TSA-induced changes in H4 acetylation represents an increase in co-localisation of transcription and replication events on chromosome 17, the correlations between RNA Pol II (pSer2) and ORC1 microarray data was compared between untreated and TSA treated cells (Figure 5.15). In untreated TK6 cells, a low correlation was observed (Pearson's correlation  $r = -0.19$ ), suggesting that RNA Pol II (pSer2) and ORC1 do not occupy the same genomic regions (Figure 5.15A). However, after treating TK6 cells with TSA, a strong positive correlation was observed instead (Pearson's correlation  $r = 0.81$ ), suggesting that RNA Pol II (pSer2) and ORC1 now occupy the same genomic regions (Figure 5.15B). Therefore, this shows that TSA promotes the co-localisation of transcription and replication on chromosome 17 in TK6 cells.

To estimate the extent to which TSA affects the distribution of transcription on chromosome 17 in TK6 cells, the correlation between RNA Pol II (pSer2) array data was compared before and after TSA treatment. In comparison to a negligible effect, which would be represented by a Pearson's correlation of 0, comparing RNA Pol II (pSer2) microarray data shows a weak positive correlation (Pearson's correlation  $r = 0.32$ ), indicating that TSA has a minor effect in altering the occupancy of RNA Pol II (pSer2) on chromosome 17 (Figure 5.16A). The same analysis performed with ORC1 microarray data shows a strong negative correlation between untreated and TSA treated conditions (Pearson's correlation  $r = -0.94$ ), indicating that TSA significantly alters the distribution of ORC1 and that their locations on chromosome 17 in treated TK6 cells is drastically different compared to untreated TK6 cells (Figure 5.16B).

To confirm these observations, both the normalised ORC1 data from untreated and treated cells are plotted around putative sites of ORC1 (peaks) detected in the untreated cells. In response to TSA, ORC1 is lost from its original sites found in untreated

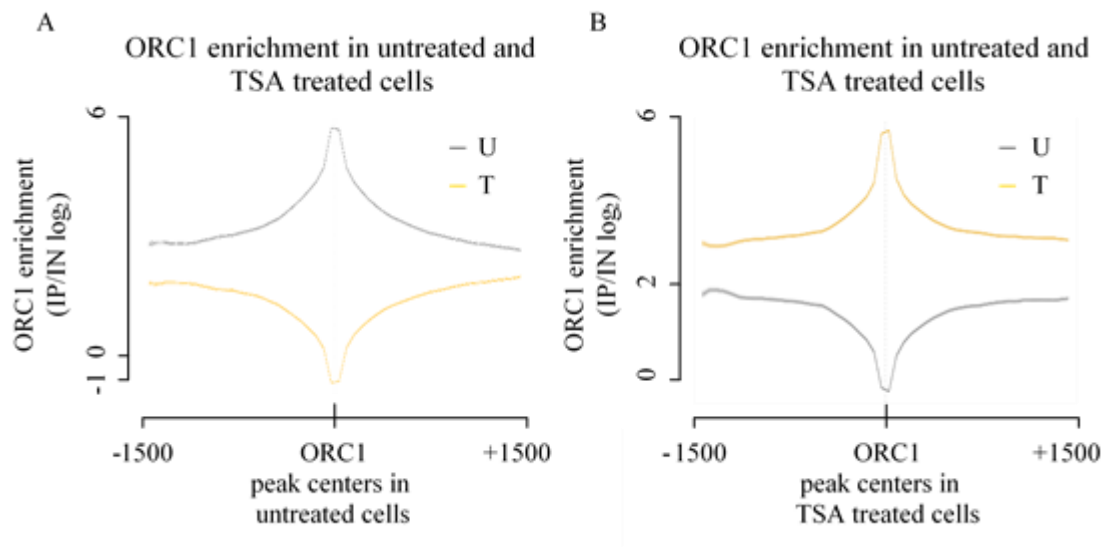


**Figure 5.15. TSA promotes the co-localisation of transcription and replication on chromosome 17 in TK6 cells.** **A:** Smooth scatter of RNA Pol II (pSer2) data plotted against ORC1 data generated from untreated TK6 cells showing a lack of correlation (Pearson's  $r = -0.19$ ). **B:** Smooth scatter of RNA Pol II (pSer2) data plotted against ORC1 data generated from TSA treated TK6 cells showing a strong positive correlation (Pearson's  $r = 0.95$ ). Average normalised data plotted,  $n = 2$  for RNA Pol II (pSer2);  $n = 3$  for ORC1.

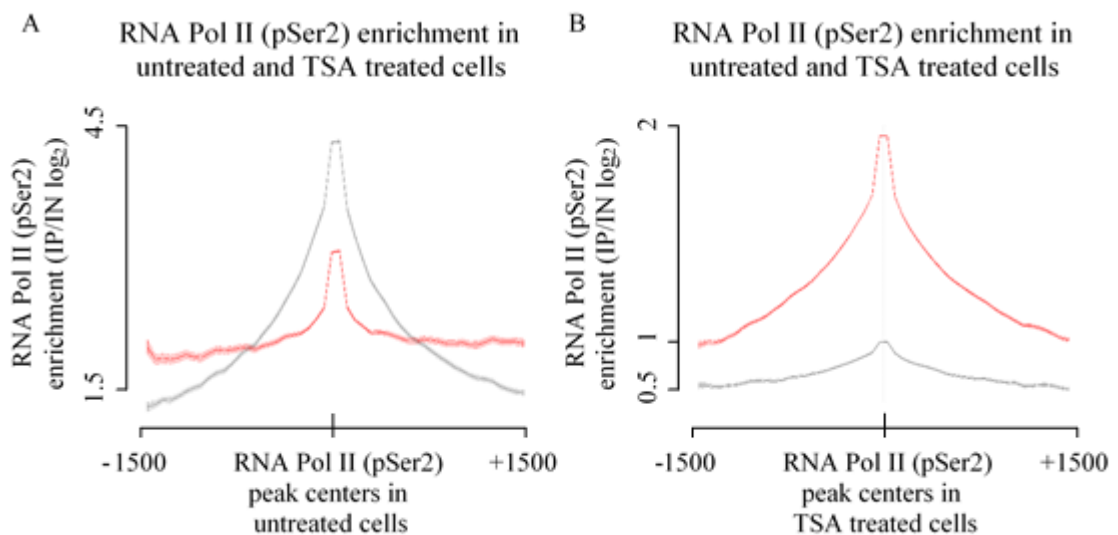


**Figure 5.16. TSA promotes the re-distribution of transcription and replication on chromosome 17 in TK6 cells.** **A:** Smooth scatter of untreated vs. TSA treated RNA Pol II (pSer2) data showing a weak correlation (Pearson's  $r = 0.32$ ). **B:** Smooth scatter of untreated vs. TSA treated ORC1 data showing a strong negative correlation (Pearson's  $r = -0.94$ ). Average normalised data plotted,  $n = 2$  for RNA Pol II (pSer2);  $n = 3$  for ORC1.

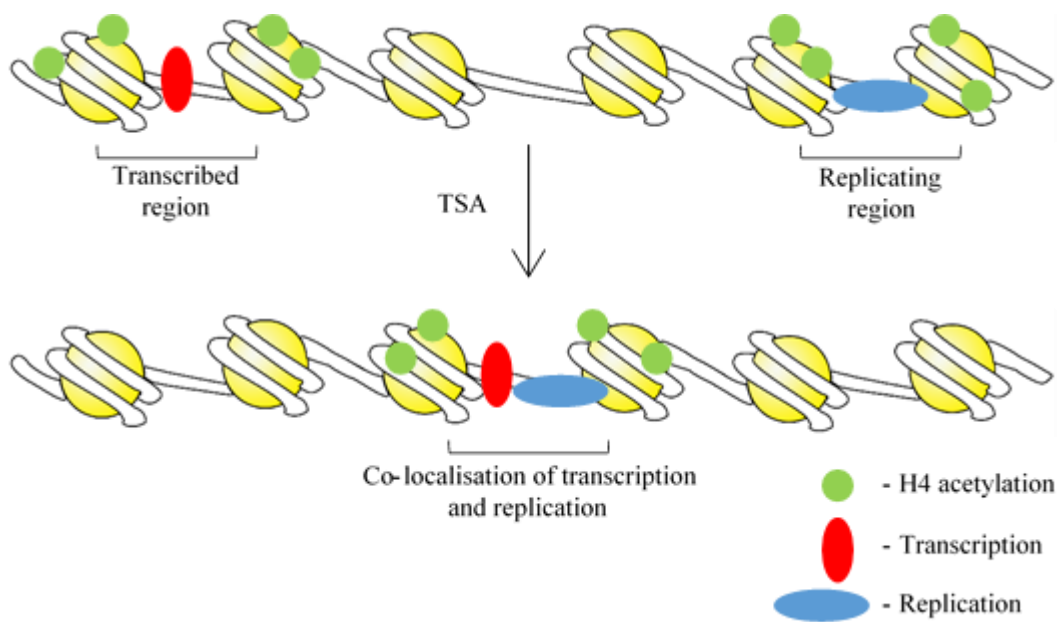




**Figure 5.17. TSA significantly alters the distribution of ORC1 on chromosome 17 in TK6 cells.** **A:** ORC1 data from untreated (*U*: grey) and TSA treated cells (*T*: yellow) plotted around ORC1 peaks identified in untreated cells, showing a loss of ORC1 occupancy at these sites in response to TSA. **B:** ORC1 data from untreated (*U*: grey) and TSA treated cells (*T*: yellow) plotted around ORC1 peaks identified in treated cells, showing a gain of ORC1 occupancy at these sites in response to TSA. SEM plotted as error shapes around the average of 3 biological replicates.



**Figure 5.18. TSA alters the distribution of RNA Pol II (pSer2) on chromosome 17 in TK6 cells.** **A:** RNA Pol II (pSer2) data from untreated (*U*: grey) and TSA treated cells (*T*: red) plotted around RNA Pol II (pSer2) peaks identified in untreated cells, showing a partial loss of RNA Pol II (pSer2) occupancy at these sites in response to TSA. **B:** RNA Pol II (pSer2) data from untreated (*U*: grey) and TSA treated cells (*T*: red) plotted around RNA Pol II (pSer2) peaks identified in treated cells, showing a gain of RNA Pol II (pSer2) occupancy at these sites in response to TSA. SEM plotted as error shapes around the average of 2 biological replicates.



**Figure 5.19. Proposed model of TSA-induced co-localisation of transcription and replication.** *Top:* in untreated TK6 cells, transcription and replication processes are spatially separated into different regions of the chromosome. *Bottom:* In response to TSA treatment, changes in H4 acetylation disrupts the spatial organisation and promotes the co-localisation of transcription and replication processes to the same chromatin region.

TK6 cells (Figure 5.17A). Doing the same for putative sites of ORC1 detected in TSA treated cells, shows a gain of ORC1 occupancy at these sites (Figure 5.17B). Likewise, TSA promotes a partial loss of RNA Pol II (pSer2) enrichment at its original sites in untreated TK6 cells (Figure 5.18A) and a gain of RNA Pol II (pSer2) enrichment elsewhere on chromosome 17 (Figure 5.18B). Overall, these are consistent with the results observed in Figure 5.15, demonstrating that TSA promotes the co-localisation of transcription and replication events on chromosome 17 (Figure 5.19) and as a result, may lead to an increase in the probability of these two processes colliding with one another, potentially resulting in DNA damage and the formation of R-loops.

### **5.7 Determining the potential for DNA damage induction and R-loop formation at sites of TSA-induced transcription and replication co-localisation**

The discovery that TSA promotes the co-localisation of transcription and replication activity presents the possibility that these may conflict and lead to the formation of DNA damage. To measure this, changes in phosphorylation of histone H2AX ( $\gamma$ H2AX) in response to TSA was used as a marker of DNA double strand breaks (similar to the work described in Chapter III). ChIP performed with the anti- $\gamma$ H2AX antibody using chromatin from untreated and TSA treated TK6 cells demonstrated enrichment following treatment at both *RPL10* and *GAD1* genes, showing that the antibody is suitable for ChIP. However, examining the reproducibility of  $\gamma$ H2AX microarray data shows a low level of correlation between three biological repeats of  $\gamma$ H2AX in untreated and TSA treated cells (Appendix, III). Additionally, plotting the  $\gamma$ H2AX data measured in untreated cells around the ORFs of genes showed no enrichment (Appendix, III), similar to background levels previously observed in the IgG control (Figure 5.6). This shows that no  $\gamma$ H2AX signal was detected from the ChIP-chip experiments.

Since  $\gamma$ H2AX is expected to be less abundant than any of the signals measured beforehand, the lack of enrichment may be caused by insufficient amounts of antigen or antibody. To address this, two further biological repeats of  $\gamma$ H2AX ChIP-chip was performed using greater amounts of chromatin ( $10^8$  instead of  $10^7$  cell equivalent) and anti- $\gamma$ H2AX antibody (20 $\mu$ g instead of 8 $\mu$ g) per experiment. Doing so however, does not

improve the sensitivity of the ChIP-chip experiment and as before, no enrichment was observed (Appendix, III).

On the other hand, the occupancy of RNA Pol II and ORC1 upstream of TSS and downstream of TES regions coincides with regions of known R-loop formation (Chapter I, 1.10.1). While titration of the anti-DNA-RNA hybrid [S9.6] antibody demonstrated enrichment, a lack of correlation was observed between three independent replicates and similar to  $\gamma$ H2AX ChIP-chip data, no enrichment was detected (Appendix, III). These results show that attempts were unsuccessful in mapping  $\gamma$ H2AX and R-loop formation to the same genomic locations using the current ChIP-chip protocol established in TK6 cells.

## 5.8 Discussion

Studies reported in this chapter aimed to examine if TSA may alter cellular processes such as transcription and replication in a way that would increase the frequency of potential collisions between them. Using the established ChIP-chip method for TK6 cells reported previously (Chapter IV), the initial steps focused on examining the effect of TSA on: 1) H4ac, a marker of euchromatin, 2) ORC1, a marker of replication initiation and 3) RNA Pol II (pSer2), a marker of transcriptional elongation. Comparisons showed that overall, TSA treatment led to a significant increase in H4ac, RNA Pol II (pSer2) and ORC1 enrichment on chromosome 17 of TK6 cells (Figures 5.1, Figure 5.3 and Figure 5.9). Plotting the data in the context of gene structure showed that this increase in signal was detected in ORFs for H4ac (Figure 5.2), in contrast to RNA Pol II (pSer2) (Figure 5.4) and ORC1 (Figure 5.8) which tend to occupy regions around ORFs.

Previous reports highlighted difficulties in being able to enrich for ORC1 binding from whole cell chromatin (Dellino et al. 2013). However, comparing the replicates of ORC1 data generated here and the replicates of background binding data (IgG negative control) demonstrated a strong reproducibility between ORC1 data, relative to the IgG data (Figure 5.5). Additionally, plotting the ORC1 data around gene structure, showed distinct enrichment (Figure 5.8) above the background levels detected in IgG data (Figure 5.6). This shows that ORC1 occupancy can be successfully mapped using the ChIP-chip method presented here. The reported association of open chromatin and G4 structures with replication initiating zones promoted the analysis of ORC1 data at these sites (Hoshina et

al. 2013). Plotting ORC1 enrichment from both untreated and TSA treated TK6 cells around DNase I hypersensitive sites (DHS identified in K562 cells, ECODE: ENSR527IGO), used to denote open regions of chromatin potentially free of nucleosomes and at G4 sites (identified in HaCat cells, NCBI GEO: GSE76688), showed that ORC1 sites in untreated TK6 cells correlate with both these features (Figure 5.7A; B). The broad nature of these peaks suggests that only a subset of ORC1 sites detected in this study co-localise with DHS and G4 structures. This could simply be due to DHS and G4 sites playing a partial role in defining initiation zones, or more likely that these sites vary from one type of cell to another but still share a common set of sites. It is also possible that some of these G4 sites overlap with DHS, since G4 mostly form in open chromatin regions (Hertsch R.H. 2016). Conversely, following TSA treatment there is a loss of ORC1 at these sites (Figure 5.7A; B), indicating a redistribution in occupancy.

In an attempt to understand the relationship between TSA-induced changes in H4ac with that of transcription and replication; the correlation between data from untreated and TSA treated cells were compared for an idea of the strength of the relationship. Firstly, to assess transcription, RNA Pol II (pSer2) data plotted against H4ac data from untreated cells showed a weak positive correlation (Pearson's correlation  $r = 0.18$ ) (Figure 5.10A). Following TSA treatment, RNA Pol II (pSer2) and H4ac become inversely correlated (Pearson's correlation  $r = -0.82$ ) (Figure 5.10B). By the same token, comparison of ORC1 and H4ac data showed that in untreated cells these two signals inversely correlated with one another, becoming more so upon TSA treatment (Figure 5.10C; D). Therefore, these results suggest that in regions of chromosome 17 showing H4ac after TSA, there is at the same time a loss of RNA Pol II (pser2) and OCR1 occupancy at these sites.

Surprisingly such observations appear contradictory to the proposed role of histone acetylation in promoting transcription and replication, as supported by evidence from the literature (Chapter V, 5.2). However, a closer analysis of the patterns of H4ac at RNA Pol II (pSer2) and ORC1 sites found in untreated verses TSA treated cells suggests that a possibility for the observed inverse correlation may be the result of RNA Pol II (pSer2) and ORC1 accumulating within nucleosome free regions, since the lack of histones would account for a lack of H4ac at these sites (Figure 5.11; 5.12). An alternative explanation is that these sites contain non-acetylated H4 histones. Although given the strong association of ORC1 data in untreated TK6 cells with DNase I hypersensitive sites

(Figure 5.7A), this is strongly suggestive of NFRs rather than non-acetylated histone H4 (Figure 5.11A).

Examining the relationship between H4ac and RNA Pol II (pSer2), showed that RNA Pol II (pSer2) sites identified in untreated cells displayed a high level of H4ac enrichment (Figure 5.12A). As evidence in the literature point towards the involvement of NFRs at promoter regions over transcribed regions (Livaja R. 2010), this could explain the enriched H4ac pattern observed here. To support this, RNA Pol II (pSer2) data from untreated cells plotted around the same DHS do not show particular patterns of enrichment (Figure 5.13), suggesting that NFRs are not essential for RNA Pol II (pSer2) occupancy. Additionally, an inversion of the H4ac pattern to a 'V-shape' at the sites of RNA Pol II (pSer2) occupancy found in TSA treated cells suggests the possibility of nucleosome depletion (Figure 5.11A). Hence, further experiments mapping changes in nucleosome positioning in TK6 cells using the same TSA treatment would shed light on the current speculation. In summary, examination of the relationships between the variables measured here is strongly suggestive and supports the idea that TSA promotes the redistribution of RNA Pol II (pSer2) and ORC1 to potentially NFRs, and while H4ac is absent at these site, it is highly enriched in the surrounding regions.

Although TSA treatment promotes an overall increase in H4ac on chromosome 17, a small number of genes however, show lower levels of H4ac in response to TSA (Figure 5.1). Therefore, to examine the change in RNA Pol II (pSer2) and ORC1 patterns at H4 hyper- and hypo-acetylated genes following treatment, the data was plotted around gene structure. Overall, both RNA Pol II (pSer2) and ORC1 enrichment appear to be higher downstream of TES at H4 hypo-acetylated genes (Figure 5.14A; B) and also higher upstream of TSS at H4 hyper-acetylated genes (Figure 5.14C; D). While the regions extending away from the gene body may or may not display the same level of H4ac, they are nevertheless related to those genes. Therefore, irrespective of the H4ac levels at TSS and TES's, these results indicate a scenario where RNA Pol II (pSer2) and ORC1 may co-localise outside of the ORFs.

To investigate this prospect, the correlation between RNA Pol II (pSer2) and ORC1 data was compared in untreated versus TSA treated cells (Figure 5.15). Evidently the increase in correlation from untreated (Pearson's correlation  $r = -0.19$ ) to TSA treated cells (Pearson's correlation  $r = 0.95$ ) strongly indicates that transcription and replication events co-localise to the same regions on chromosome 17 following treatment. Additionally, the impact of TSA on transcription and replication was demonstrated by

comparing the correlations between each dataset before and after treatment (Figure 5.16). The results showed that TSA had a significant effect on ORC1 occupancy, with almost a complete re-distribution of replicating sites (Figure 5.15B). This is consistent with the published work of Conti and colleagues showing that HDAC inhibition promotes activation of dormant origins (Conti et al. 2010). Therefore, it is possible that the ORC1 changes here reflect a shift in activity in origins sites. On the other hand, TSA had a less pronounced effect on RNA Pol II distribution (Figure 5.15A). Detailed analysis of the changes in RNA Pol II (pSer2) and ORC1 at sites identified in untreated and TSA treated cells also support the changes in correlation observed (Figure 5.17; 5.18).

Taken together, these results show that TSA promotes the localisation of transcription and replication to the same regions of chromosome 17, prompting the analysis of potential conflicts as a mechanism of generating DNA damage. However, attempts to map  $\gamma$ H2AX in TK6 cells proved unsuccessful in detecting enrichment. This was demonstrated by the fact that biological replicates did not correlate strongly and showed no enrichment when plotted in relation to gene structure (Appendix, III), even though  $\gamma$ H2AX enrichment was detected by ChIP at different sites in response to TSA (Appendix, III). It is possible that the ChIP DNA was sufficient for qPCR analysis but below the required amount for whole genome amplification. To address this, larger amounts of chromatin and antibody were used, yet no improvement was observed, suggesting that other factors are limiting its detection of the microarrays (Appendix, III). Possibilities to improve this include optimising the post-IP washes to remove potential background, or the use of an alternative anti- $\gamma$ H2AX antibody.

The pattern of RNA Pol II (pSer2) and ORC1 co-accumulation at upstream TSS and downstream TES regions presents evidence to suggest that if transcription-replication conflicts do occur then R-loop formation is likely, since coincidentally these regions favour the formation of R-loop under physiological settings (Sollier J 2015). Yet, attempts to map R-loops in TK6 cells using RDIP-chip resulted in the same outcome as  $\gamma$ H2AX (Appendix, III). Titration of the anti-DNA-RNA [S9.6] antibody at positive control loci showed a small but distinguishable enrichment (5-fold) over the negative control loci, suggesting that the antibody is suitable for ChIP (Appendix, III). Since both  $\gamma$ H2AX and R-loop signals are predicted to be less abundant than H4ac, RNA Pol II (pSer2) and ORC1, it is possible that the currently established ChIP-chip protocol for TK6 cells lacks the sensitivity to detect these and require further optimisation.





## **Chapter VI**

**A genetic approach to examine DNA  
damage associated with loss of HDAC  
function in *Saccharomyces cerevisiae***



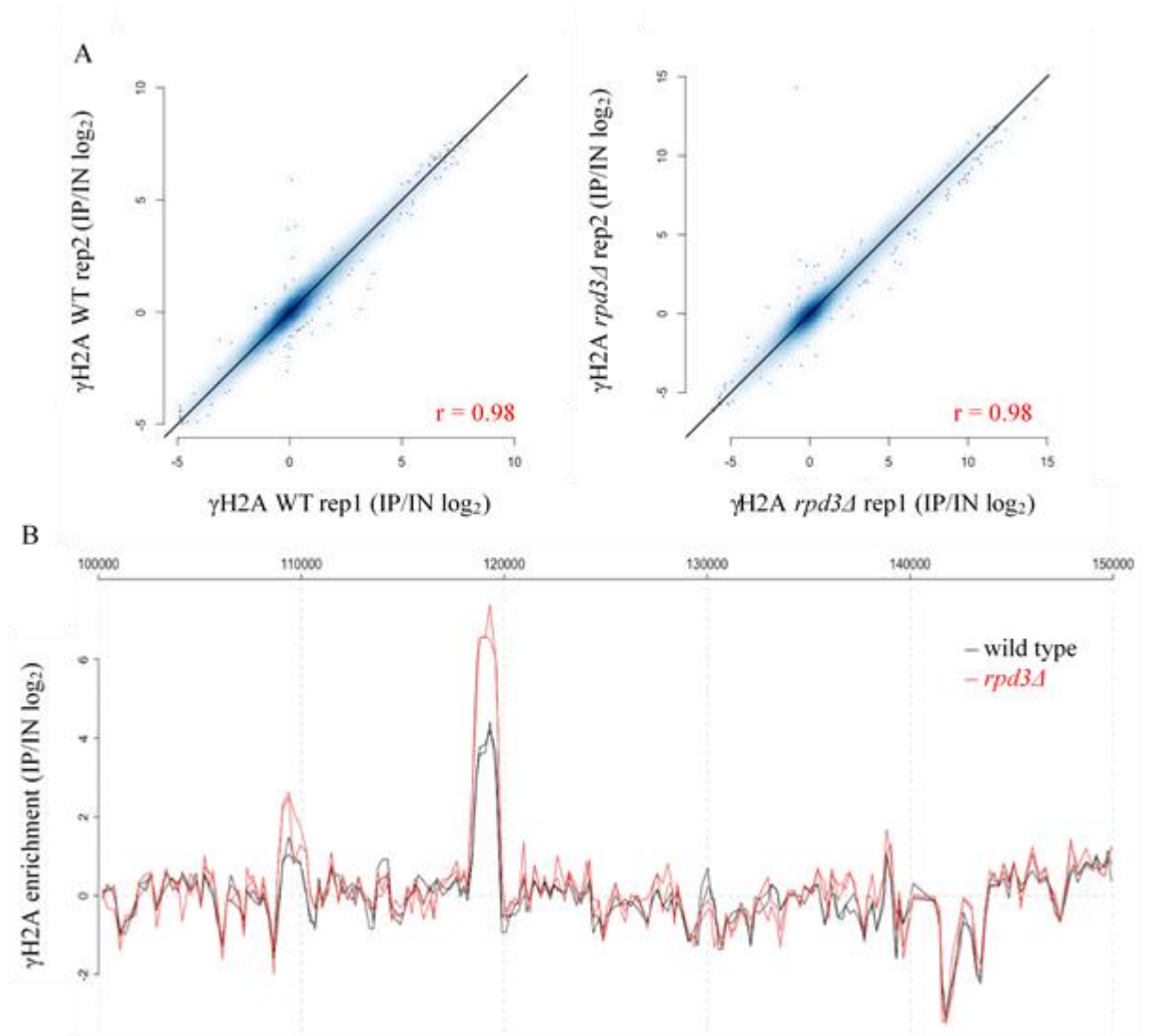
## 6.1 Chapter aims

Given the technical difficulties in trying to measure  $\gamma$ H2AX and R-loops by ChIP-Chip in TK6 cells, studies in this chapter seek instead to model these in yeast. The previously identified *rpd3 $\Delta$*  strain (Chapter III) will be used here to mimic the effects of treating with an HDAC inhibitor, and  $\gamma$ H2A (ortholog of  $\gamma$ H2AX) and R-loops will be mapped to investigate if a loss in HDAC function leads to the possibility of an increased level of transcription-replication collisions. The results of these findings may help to support the observations made in TK6 cells.

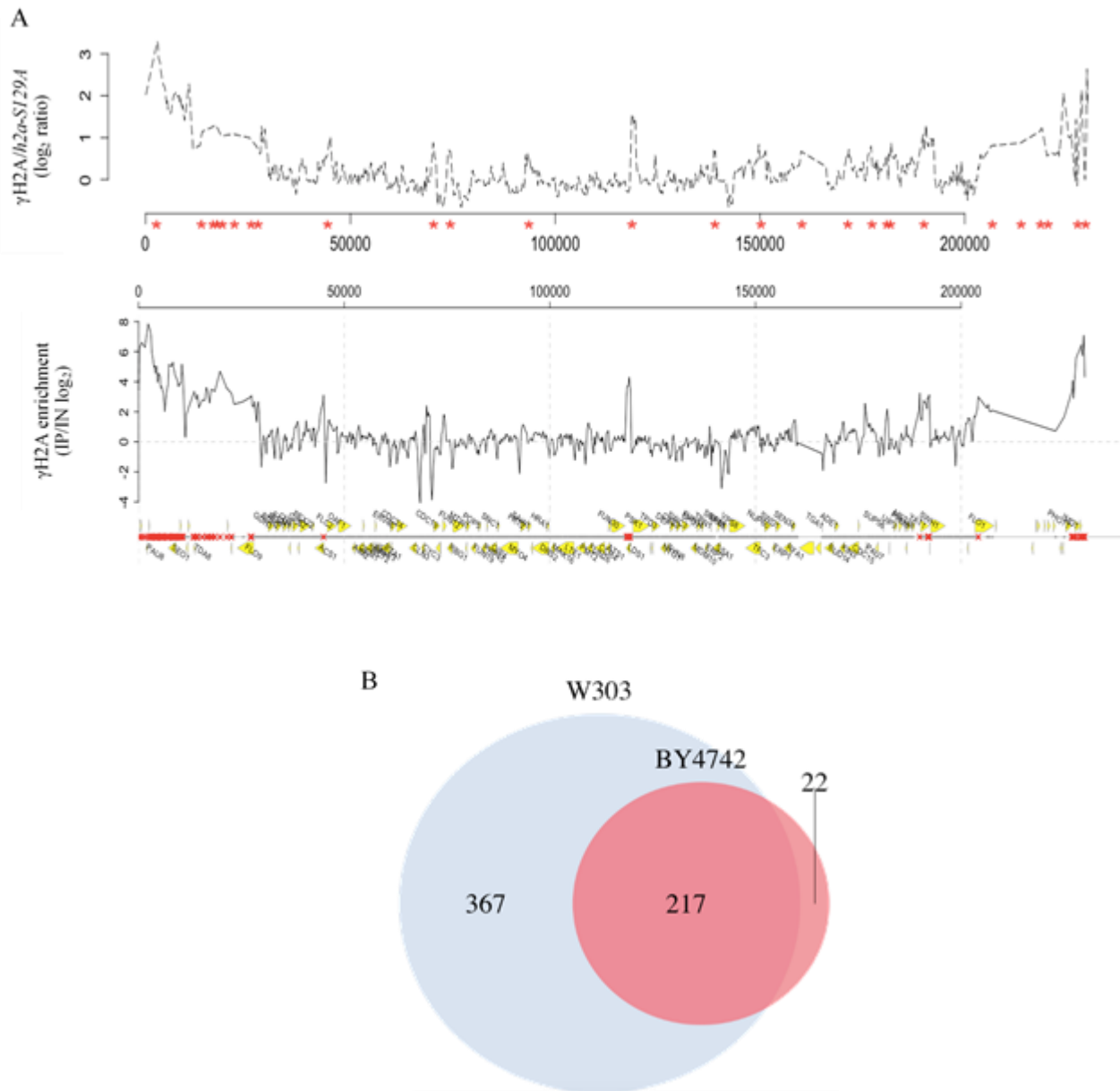
## 6.2 Validation of $\gamma$ H2A and R-loop ChIP-chip data

To map genome-wide  $\gamma$ H2A enrichment, chromatin was prepared using previously optimised conditions (Chapter III) and immunoprecipitated with the anti-histone H2A (phospho-S129) ChIP-grade antibody. For R-loop microarrays, genomic DNA was extracted from cells and digested with RNase A under high salt conditions to remove single and double stranded RNA. R-loops are then enriched using the anti-DNA-RNA hybrid [S9.6] antibody (Chan et al. 2014). For both targets, the ChIP reaction was performed with genetic material obtained from wild type and *rpd3 $\Delta$*  yeast cells. Immunoprecipitated DNA was then amplified and labelled prior to array hybridisation. For this study, *S. cerevisiae* whole genome arrays are used (Chapter II, 2.2.4.3).

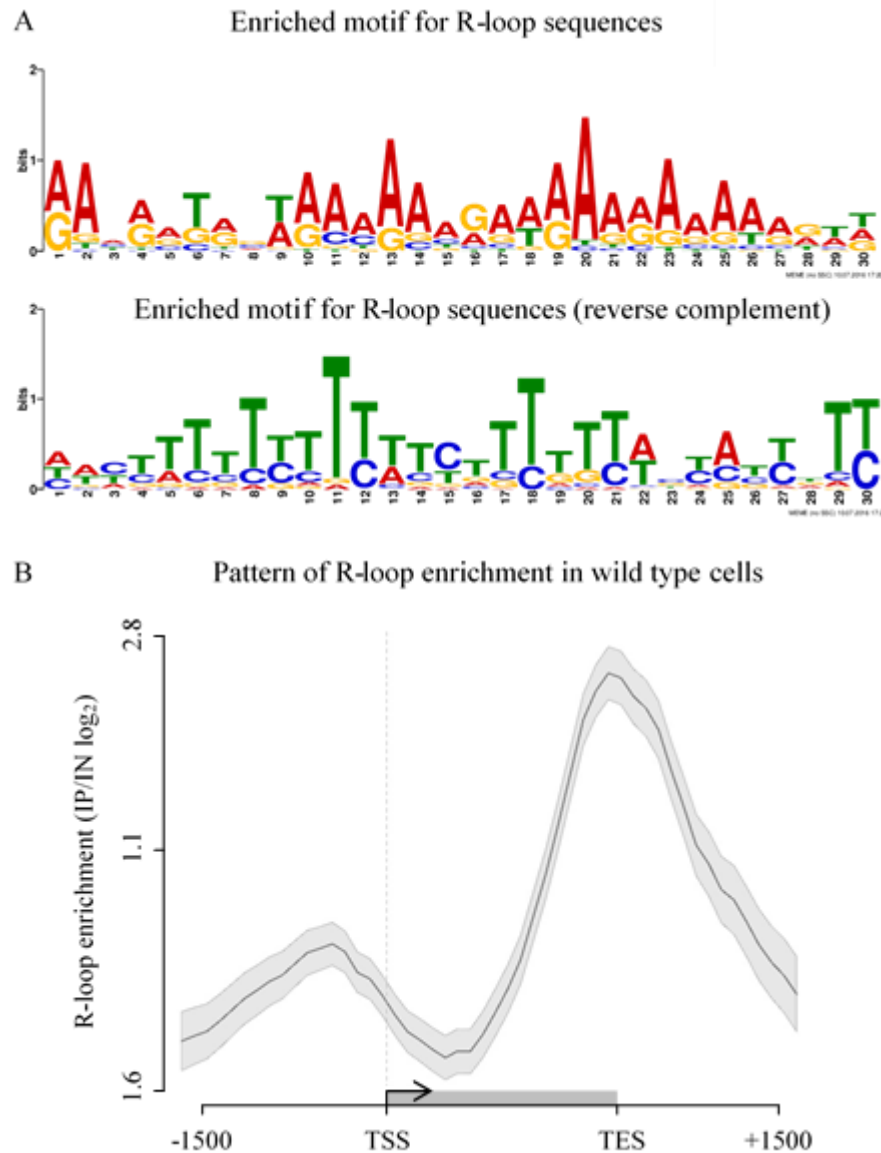
Following data normalisation and quality control (Appendix, IV), biological replicates were first compared to examine reproducibility. A high correlation (Pearson's  $r > 0.9$ ) was observed between the two biological replicates generated in both WT and *rpd3 $\Delta$*  strains measuring  $\gamma$ H2A (Figure 6.1A) and R-loops (Appendix, IV). This is also evident when examining data plotted along a small section of chromosome 1 (100kbp - 150kbp) (Figure 6.1B), demonstrating that these results are highly reproducible in yeast. To validate the results of the  $\gamma$ H2A microarray, normalised data from the wild type strain was compared to  $\gamma$ H2A data mapped in W303 strains published by Szilard and colleagues (Szilard et al. 2010). Visual comparison of  $\gamma$ H2A for chromosome 1 shows a very similar pattern of enrichment between the two datasets (Figure 6.2A). Furthermore, a comparison of all the  $\gamma$ H2A peaks detected throughout the yeast genome showed that the majority of  $\gamma$ H2A peaks identified in this study match those previously identified by Szilard and colleagues (Szilard et al. 2010) (Figure 6.2B). C-rich DNA and G-rich RNA are important factors that favour the formation of stable DNA:RNA hybrid structures (Skourti S.K., Kamieniarz G.K., and Proudfoot N. J. 2014). Recent evidence has also shown that R-loops are prevalent throughout the genome and are preferentially found at CpG-containing promoters (Ginno et al. 2012) and G-rich transcription termination regions (Skourti-Stathaki, Proudfoot, and Gromak 2011). To validate the R-loop microarray data generated here, motif enrichment analysis was carried out in MEME-ChIP (Ma, Noble, and Bailey 2014) using sequences of identified R-loop peaks in the wild type strain. The motif with the highest statistical significance (Figure 6.3A *Top panel*:  $E = 2.1 \times 10^{-27}$ ) is defined by the greatest number of sequences that express the highest logarithmic ratio of a given motif compared to a null background model (Ma et al., 2014). From the list of



**Figure 6.1. Biological repeats of yeast  $\gamma$ H2A microarray data show high reproducibility. A:** smoothed scatter plots showing the Pearson's correlation coefficient between two biological repeats for wild type (*left*) and *rpd3Δ* (*right*). **B:** normalised array data plotted along a section of chromosome 1 (100 – 150kb). Array probes are shown in grey along the annotation line with enriched regions in red.  $n = 2$ .



**Figure 6.2.  $\gamma\text{H2A}$  microarray data generated in this study is consistent with published microarray data.** **A:**  $\gamma\text{H2A}$  peaks detected in wild type W303 strain from published findings (Chan et al. 2014) (*top*, dashed line), compared to  $\gamma\text{H2A}$  peaks detected in the current study using wild type BY4742 strain (*bottom*, solid line), plotted along the full length of yeast chromosome 1. **B:** majority of  $\gamma\text{H2A}$  peaks identified in the yeast genome (*red circle*) overlap with those identified in the W303 strain (*blue circle*) (Chan et al. 2014). Peak matching analysis conducted using ChIPpeakAnno v3.4.6 in R with maxgap = 1000bp.



**Figure 6.3. R-loop microarray data generated in this study is characteristic of DNA:RNA hybrid structures.** **A:** motif-enrichment analysis detects high GA-skew in the RNA component and a high CT-skew in the DNA component of R-loop peak sequences in the wild type strain (Skourti-Stathaki and Proudfoot 2014) ( $E$ -value =  $2.1 \times 10^{-27}$ ). Analysis was carried out using meme-chip scripts (Ma, Noble, and Bailey 2014) in Perl v5.22.1. **B:** R-loops detected in the wild type strain plotted in the context of gene structure ( $\pm 15$ kb up and downstream of the open reading frame, grey box) showing higher distribution at promoter and transcription end sites. Standard deviation error shapes plotted around the average of two biological repeats. TSS – transcription start site. TES – transcription end site.  $n = 2$ .

peaks analysed, the RNA component is highly GA-rich (Figure 6.3A: *Top panel*), whereas the corresponding template DNA is CT-rich (Figure 6.3A: *Bottom panel*), suggesting that these peaks do indeed display sequences that can contribute to the described thermostability of R-loop structures (Lynda Ratmeyer 1994). Plotting the normalised wild type strain data in the context of gene structure (+/- 1500bp up and downstream of the open reading frame) for all genes in the yeast genome, showed elevated levels of enrichment in the promoter and transcription end sites (TES) (Figure 6.3B), which is consistent with previous findings (Ginno et al. 2012). Together these results demonstrate that both genome-wide  $\gamma$ H2A and R-loop microarray data generated in this study agrees with published findings, warranting further analysis.

### 6.3 Analysis of yeast $\gamma$ H2A and R-loop data

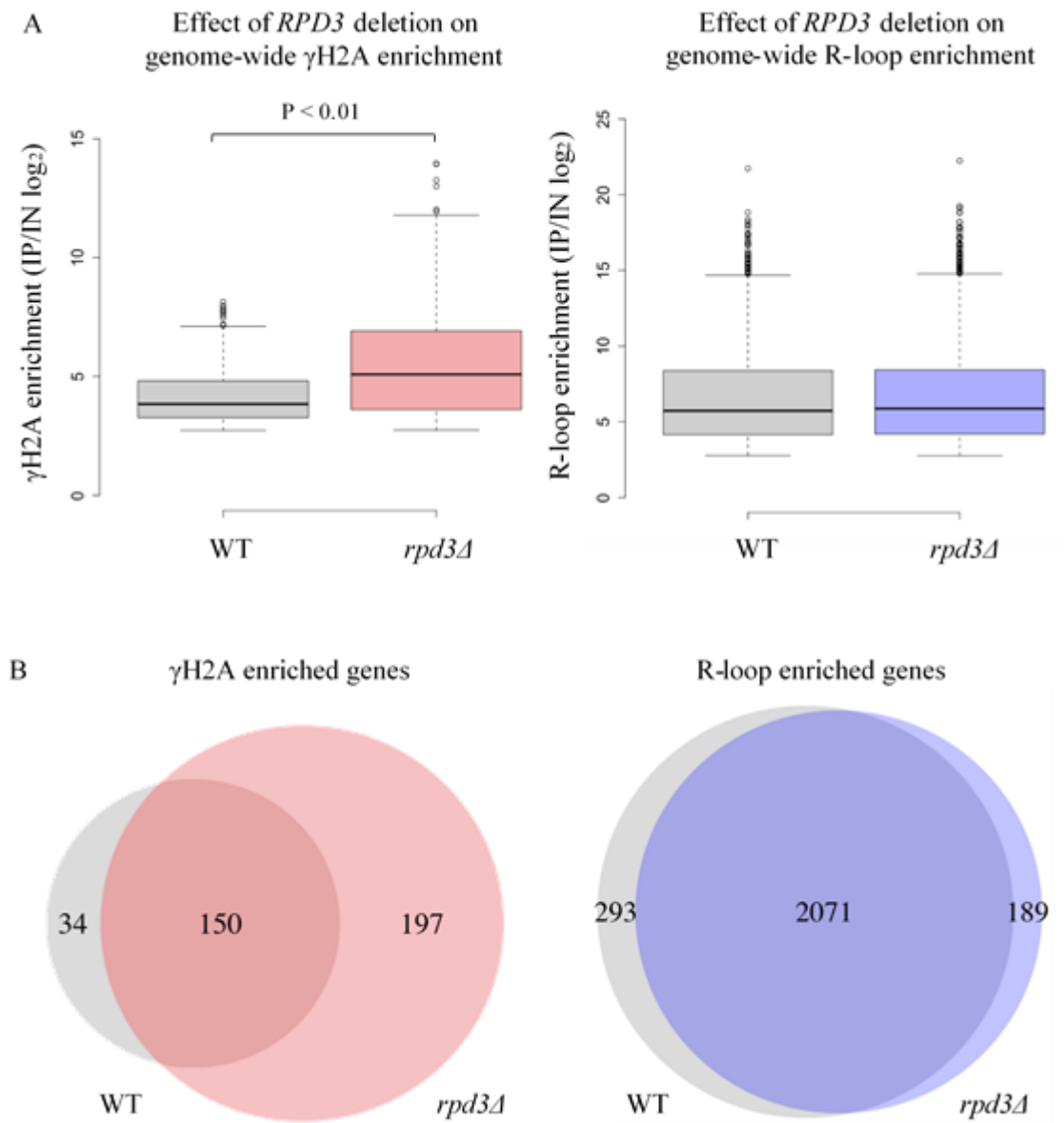
Having normalised and validated the microarray data, the following section aims to examine firstly the overall effect of *RPD3* deletion on  $\gamma$ H2A and R-loops in the yeast genome, followed by analysis of possible co-localisation between the two as a potential indicator of transcription-replication conflicts.

#### 6.3.1 Effect of *RPD3* deletion on genome-wide $\gamma$ H2A and R-loop signals in yeast

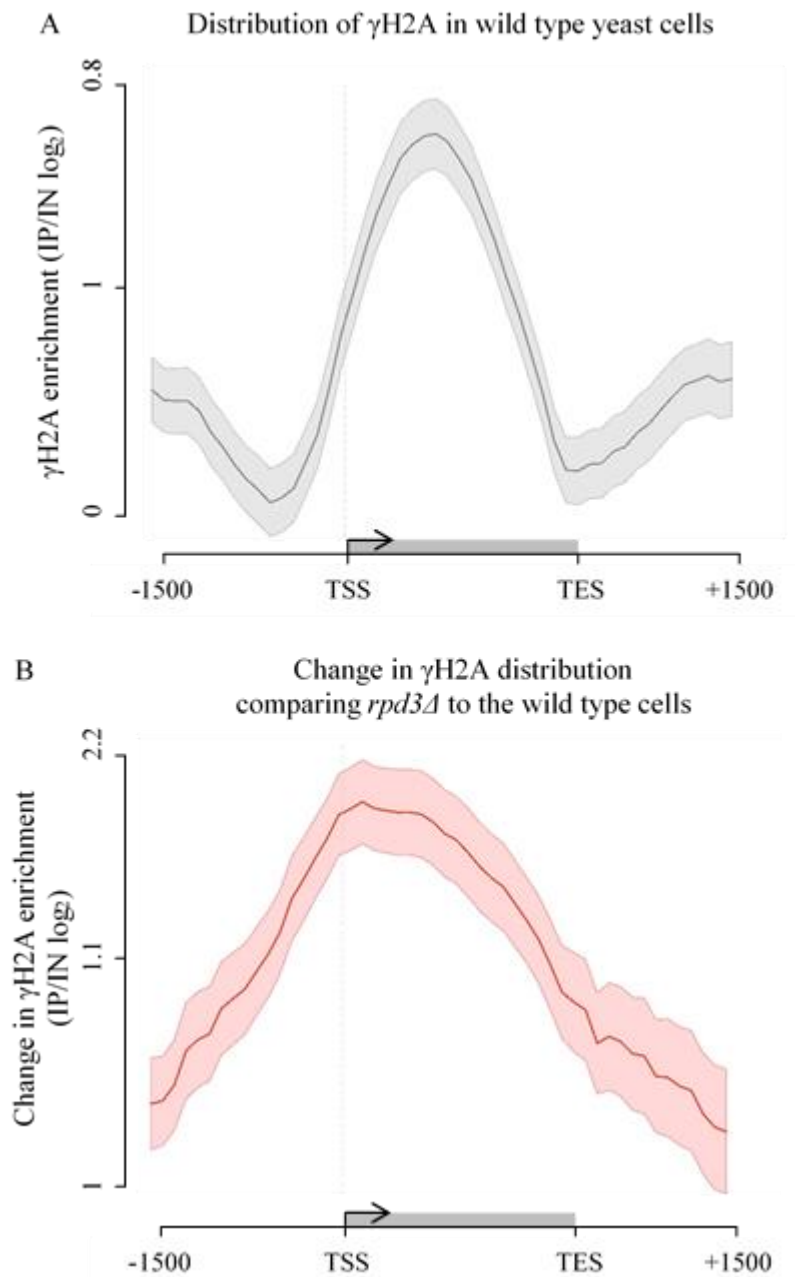
To investigate if the *RPD3* deletion significantly affected  $\gamma$ H2A and R-loop formation in yeast, the net change between wild type and mutant cells was examined. Comparing the IP/IN  $\log_2$  ratios of peaks detected, *rpd3 $\Delta$*  strain showed a significant increase in  $\gamma$ H2A ( $P < 2.2 \times 10^{-16}$ ) relative to the wild type, on the other hand, no significant difference in R-loop formation was observed ( $P = 0.148$ ) (Figure 6.4A). These differences are also highlighted when examining enriched genes. Approximately half the genes showing  $\gamma$ H2A enrichment (Figure 6.4B:  $197/347 = 57\%$ ) are unique in *rpd3 $\Delta$* , whereas the majority of R-loop enriched genes remain the same in both the wild type and *rpd3 $\Delta$*  strains (Figure 6.4B).

To demonstrate where in the genome these changes in both these signals occur, the relative difference in enrichment was calculated by subtracting the wild type data

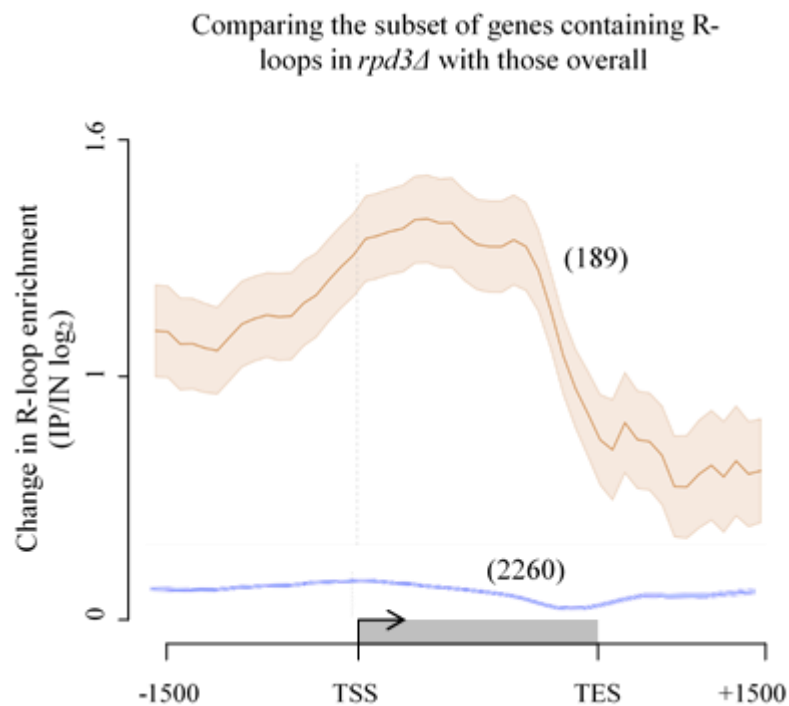




**Figure 6.4. *RPD3* deletion significantly increases genome-wide  $\gamma$ H2A enrichment with little effect on R-loop formation.** **A:** *left* – significant increase in global  $\gamma$ H2A enrichment in the *rpd3Δ* strain compared to the wild type ( $P < 2.2 \times 10^{-16}$ , Welch’s t-test); *right* – no significant difference observed in R-loop formation between the *rpd3Δ* and the wild type strains ( $p = 0.148$ , Welch’s t-test). **B:** Overlap of genes enriched for  $\gamma$ H2A (*left*) and R-loops (*right*) in the wild type compared to the *rpd3Δ* strain. Peak enrichment values used to calculate relative increase/decrease. Diagrams drawn using R base package v3.2.4 and the ‘VennDiagram’ package v1.6.17.



**Figure 6.5.** *rpd3Δ* cells display an increase in  $\gamma$ H2A enrichment primarily in ORFs compared to the wild type. Profile plots around gene structure averaged for all genes in the yeast genome showing: **A:** distribution of  $\gamma$ H2A signal in the wild type strain (*grey*). **B:** change in  $\gamma$ H2A enrichment, calculated by subtracting wild type data away from *rpd3Δ* data (*red*). Standard deviation error shapes plotted around the average of two biological repeats.



**Figure 6.6. R-loops identified in *rpm3Δ* cells compared to the wild type are enriched in the ORFs of genes.** Profile plots around gene structure averaged for all genes in the yeast genome showing the change in overall R-loop enrichment, calculated by subtracting wild type data away from *rpm3Δ* data (blue, 2260 genes) and the change in R-loop enrichment for a subset of genes identified only in *rpm3Δ* cells (dark orange, 189 genes). Standard deviation error shapes plotted around the average of two biological repeats.

away from the *rpd3Δ* data. The change in  $\gamma$ H2A signal plotted in the context of gene structure can be compared to the pattern observed in the wild type to identify regions showing changes in enrichment when *RPD3* is lost (Figures 6.5B; 6.6). Comparison of  $\gamma$ H2A profiles show that the significant overall increase observed in Figure 6.4A occurs predominantly within the open reading frames (ORFs) of genes (Figure 6.5). The lack of a significant change in R-loop formation between the strains is shown in Figure 6.6, where the overall profile lies around 0, reflecting the results observed in Figure 6.4A. However, for a small subset of genes exclusively enriched for R-loops in *rpd3Δ* cells, a significant increase is noted in the ORFs (as the IP/IN  $\log_2$  ratio is  $>1$ , equating to a relative fold change of  $>2$ ). Interestingly, this suggests that R-loops identified in the subset of genes in *rpd3Δ* cells may co-localise with  $\gamma$ H2A, given that they are both enriched within the ORFs (Figure 6.5B; 6.6).

### **6.3.2 Examining the co-localisation of $\gamma$ H2A and R-loop in yeast as a marker of transcription-replication conflicts**

Since  $\gamma$ H2A and R-loop formation are the predicted outcomes of transcription and replication collisions (Skourti-Stathaki and Proudfoot 2014), the co-localisation of both signals might be expected under these circumstances. Although previous results demonstrated an overall increase in  $\gamma$ H2A with no significant change in R-loop formation, closer examination shows a similar number of genes enriched for  $\gamma$ H2A (197) and R-loops (189) only in the *rpd3Δ* strain (Figure 6.4B). In addition, genes enriched for  $\gamma$ H2A in the wild type show higher levels of  $\gamma$ H2A in *rpd3Δ* (Figure 6.1B), whereas R-loops in the wild type and *rpd3Δ* strains remain mostly unchanged (Figure 6.4A), apart from those appearing only in the *rpd3Δ* strain (Figure 6.4B). To identify genes showing an increase in both  $\gamma$ H2A and R-loop enrichment following *RPD3* deletion, the average peak enrichment ( $\log_2$  IP/IN) was first calculated for each gene in both wild type and *rpd3Δ* strains. Comparing the two lists, genes in *rpd3Δ* showing a 2-fold relative increase in the average peak enrichment (IP/IN) over the wild type was selected. As a result, 320 genes are identified, showing a 2-fold or greater level of  $\gamma$ H2A enrichment in *rpd3Δ* compared to the wild type and 189 genes identified as showing a 2-fold or greater level of R-loop enrichment in *rpd3Δ* compared to the wild type. By matching both lists, only 9 genes show an increase (2-fold or greater) in both  $\gamma$ H2A and R-loop (Figure 6.7). However,

such a small number of genes do not represent a statistically significant overlap ( $P = 0.73$ ), suggesting that co-localisation of the two signals do not appear abundantly throughout the genome.

To investigate if these regions might represent potential sites where transcription and replication machineries may conflict, the bidirectional spreading of  $\gamma$ H2A across large genomic regions, a well-described phenomenon linked to DNA double strand breaks, was examined (Szilard et al. 2010). For an estimate of  $\gamma$ H2A spread, the genomic range between  $\gamma$ H2A peaks lying within 1kb of each other was calculated and considered part of the same  $\gamma$ H2A signal. From the data, the average estimated  $\gamma$ H2A signal (spread) is similar in both wild type (mean = 873bp) and *rpd3 $\Delta$*  (mean = 924bp) strains (Figure 6.8), with no significant difference ( $P = 0.354$ , Welch's t-test). Although the average probe spacing on the microarray is calculated to be ~273bp, the actual distribution varies, with gene rich regions abundantly represented while other regions are less populated with probes. This is because regions containing repetitive DNA sequences cannot be accurately represented and is particularly apparent at telomeres and sub-telomeres, which lack representation on the arrays. Even so,  $\gamma$ H2A is highly enriched in these areas, covering ~30kb on all chromosome ends in the yeast genome of wild type cells (Figure 6.9). This phenomenon was also reported in previous studies in yeast (Szilard et al. 2010) and human cell lines (Siow et al. 2012). Intriguingly, *rpd3 $\Delta$*  significantly elevates the  $\gamma$ H2A signal observed in the wild type at telomere and sub-telomeric regions of all yeast chromosomes, although R-loops do not appear to show the same level of change (representative data for chromosome I shown in Figure 6.9). Also, of the 9 genes showing co-occurrence of  $\gamma$ H2A and R-loops, 4 are found within telomere regions, which are known to exhibit high levels of replication stress (Szilard et al. 2010).

### **6.3.3 Examining $\gamma$ H2A and R-loops at sites of replisome impedance**

Work by Szilard *et al.* demonstrated that in addition to telomeres,  $\gamma$ H2A map to other sites susceptible to replication fork stalling/collapse, such as replication origins, tRNA genes, long terminal repeats (LTRs), a small number of coding-genes and rDNA sequences (Szilard et al. 2010). To examine how *rpd3 $\Delta$*  affects  $\gamma$ H2A and R-loop formation at some of these sites, microarray data is plotted (+/- 1500bp) around the starting positions of known replication origins, tRNA genes and LTRs (Figure 6.10). The

results show that  $\gamma$ H2A in wild type cells are enriched at replication origins and LTRs but not at tRNA genes. In comparison with *rpd3 $\Delta$*  cells,  $\gamma$ H2A levels are highly elevated and retain the same pattern of distribution (Figure 6.10, *left*). R-loops are enriched at replication origins and tRNA genes but not at LTRs although the level and pattern of R-loop enrichment in both wild type and *rpd3 $\Delta$*  cells remain the same. Interestingly, R-loop enrichment at replication origins sites coincides with high levels of  $\gamma$ H2A (Figure 6.10A), which suggests that in wild type cells, transcription may be interfering with replication at a subset of common fragile sites located close to replication origins.

#### 6.3.4 Significance of R-loop enriched genes in *rpd3 $\Delta$* cells

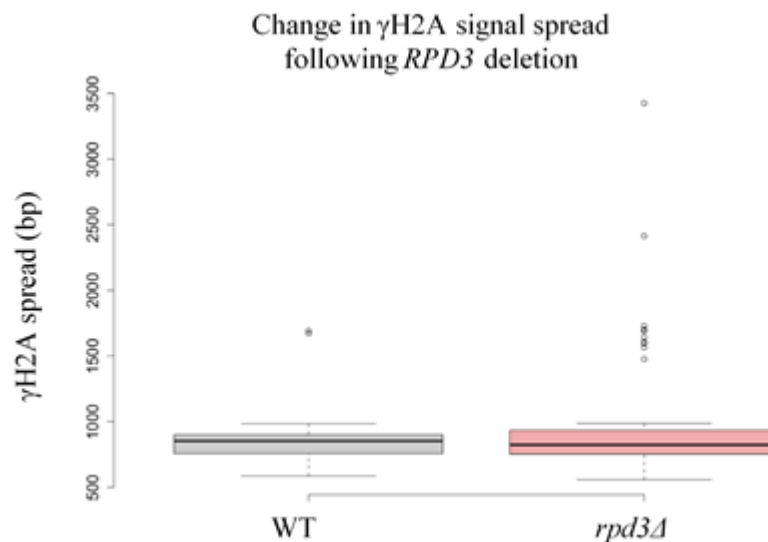
While R-loop formation is not as greatly affected by the loss of *RPD3* as  $\gamma$ H2A, a small group of genes (189) show R-loop enrichment only in *rpd3 $\Delta$*  cells Figure 6.7. To examine the significance of this, the molecular processes associated with these genes were surveyed by carrying out a Gene Ontology search using the *Saccharomyces* genome database (SGD) (Siow et al. 2012). Consequently, an over-representation in genes coding for small nucleolar RNA (snoRNA)-binding proteins was observed ( $P < 0.01$ ) (Figure 6.11). The products of these genes are constituents of small nucleolar RNA-protein complexes (snoRNP), which function in pre-ribosomal RNA (rRNA) processing. As genes in this group display an accumulation of R-loops in ORFs (Figure 6.6), this pattern is distinct from the physiological accumulation of R-loops at CpG promoter (Ginno et al. 2012) and transcription termination sites (Skourti-Stathaki, Proudfoot, and Gromak 2011) (Figure 6.3B) and may have a negative effect on cellular processes such as transcription and possibly replication. It is therefore possible that the expression of these snoRNA-binding proteins may be suppressed.

Overall, these results comparing  $\gamma$ H2A and R-loop formation in wild type and *rpd3 $\Delta$*  cells present evidence to suggest that the loss of *RPD3* does not significantly increase the prevalence of transcription-replication conflicts in the yeast genome. Instead, a significant increase in  $\gamma$ H2A observed throughout the genome is likely to reflect elevated levels of replication stress.

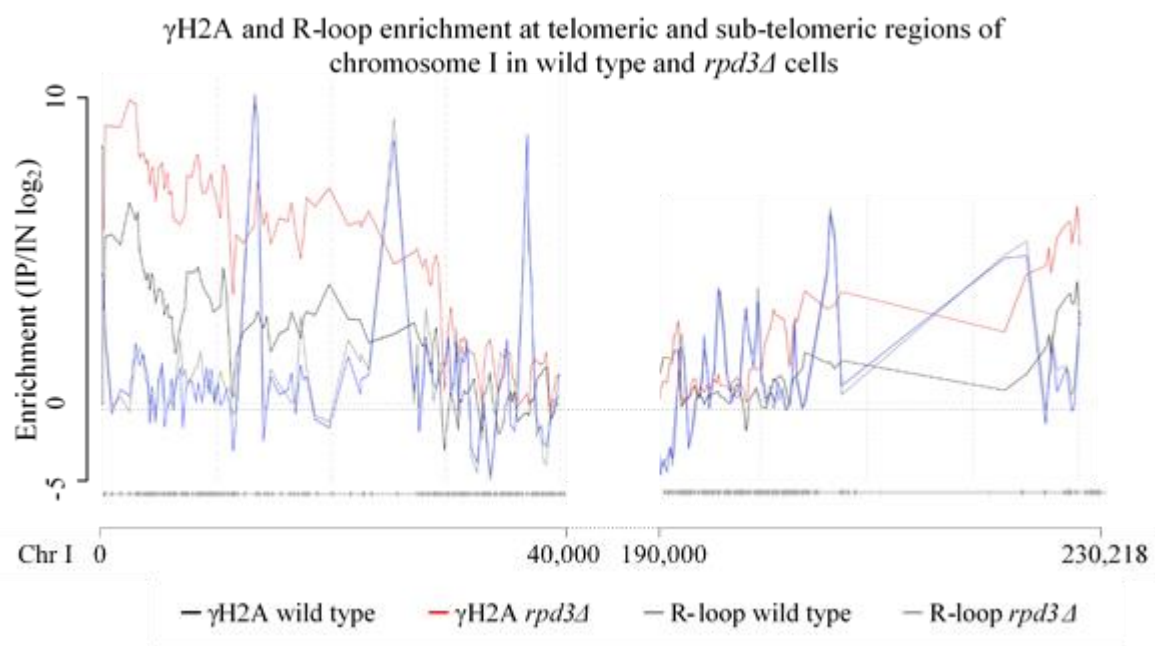
Deletion of *RPD3* results in an increase in both  $\gamma$ H2A and R-loop enrichment at 9 genes



**Figure 6.7.  $\gamma$ H2A and R-loops co-localise in the region of 9 genes as a result of *RPD3* deletion.** Calculating the difference between *rpd3 $\Delta$*  and wild type data identified 320 genes and 189 genes showing an increase in  $\gamma$ H2A and R-loop formation, respectively. Within these genes only 9 displayed a significant increase in both  $\gamma$ H2A and R-loops but do not represent a statistically significant overlap ( $P = 0.73$ ). Diagram drawn using R-based ‘VennDiagram’ package v1.6.17. Statistical analysis (Hypergeometric test) was performed using the R-base package v3.2.4.

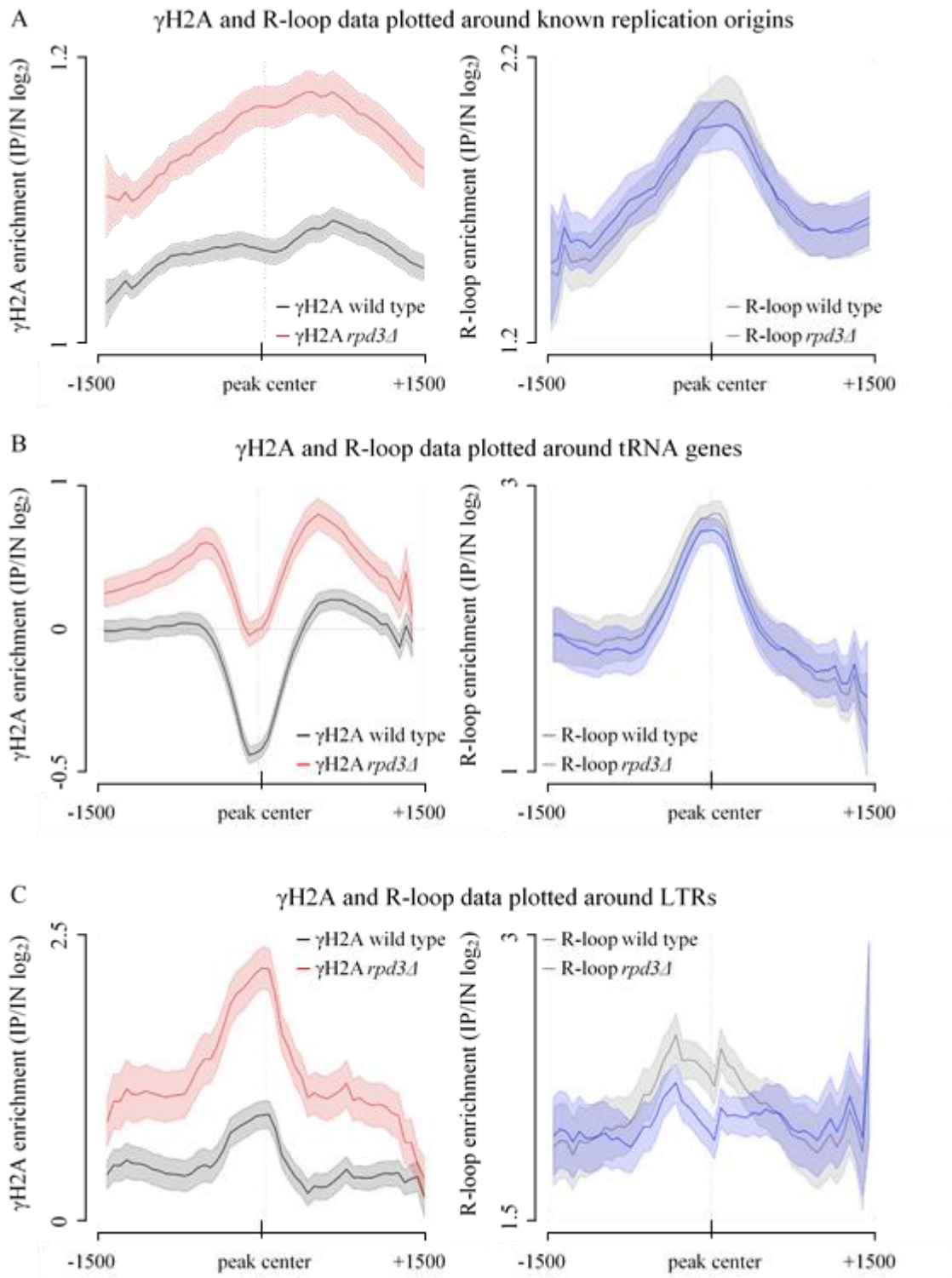


**Figure 6.8.  $\gamma$ H2A detected as a result of *RPD3* deletion do not resemble DNA-damage induced  $\gamma$ H2A signals.** Boxplot showing the distribution of  $\gamma$ H2A peak ranges for wild type (mean = 873bp) and *rpd3 $\Delta$*  (mean = 924bp). Diagram drawn using the R-base package v3.2.4.

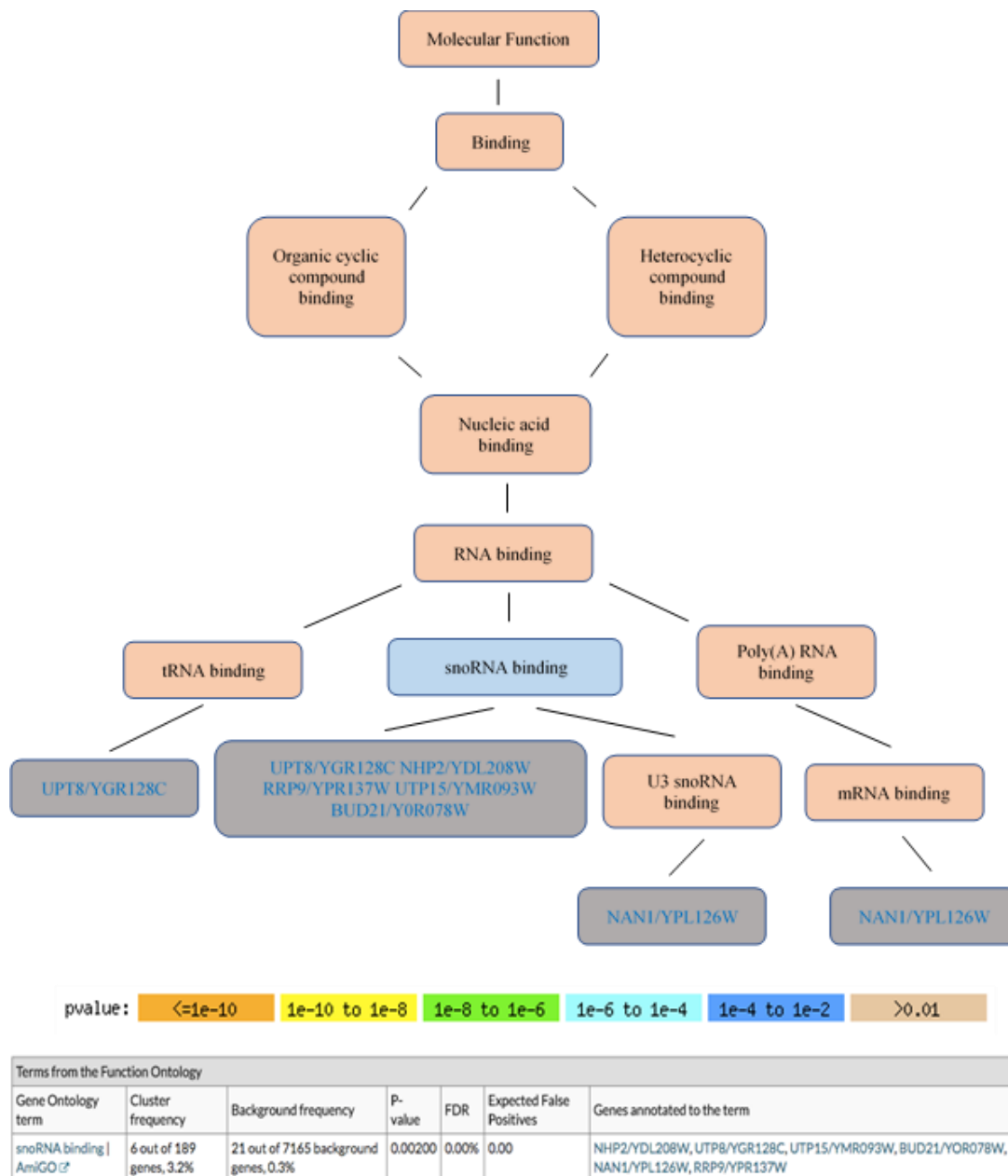


**Figure 6.9.  $\gamma$ H2A and R-loops plotted at telomeric and sub-telomeric ends of yeast chromosome 1 in wild type and *rpd3 $\Delta$*  cells.** Section of a genome plot showing  $\gamma$ H2A and R-loop formation on the ends of yeast chromosome 1 (from 0 – 40kb and 190kb – 230,218bp). Average of two biological repeats shown.





**Figure 6.10.  $\gamma$ H2A and R-loop microarray data plotted around sites known to impede replication fork progression.** **A:** Replication origins, **B:** tRNA genes and **C:** long terminal repeats (LTRs) are locations in the yeast genome identified to potentially inhibit replication forks, leading to cell stress. Starting coordinates for consensus replication origins in wild type yeast are obtained from OriDB (DNA replication origin database) (Siow et al. 2012). Starting coordinates for tRNA genes and LTRs are obtained from YeastMine (Balakrishnan et al. 2012). Standard deviation error shapes plotted around the average of two biological repeats.



**Figure 6.11. Genes exhibiting R-loop formation only in *rpd3Δ* cells are highly over-represented for those encoding for snoRNA-binding proteins.** R-loop formation is detected in 189 genes exclusive to *rpd3Δ* but not in wild type cells. Gene Ontology analysis (GO term) of molecular function was performed in SGD (Cherry et al. 2012) using the 189 identified genes, showing a significant over-representation of genes coding for small nucleolar RNA binding (snoRNA) proteins: 6 of 21 (P = 0.002).

## 6.4 Discussion

Using genome-wide analysis tools to identify associations between the products of transcription-replication conflicts, such as  $\gamma$ H2A and R-loops the yeast *RPD3* HDAC mutant, provides a useful way of identifying if and where these events occur throughout the genome, affording potential evidence to support a mechanism by which HDAC inhibitors could induce DNA damage.

While earlier chemical means of inhibiting HDAC enzymes proved unsuccessful (Chapter III), an alternative genetic approach was taken to make use of an HDAC deletion strain to mimic the effects of treating yeast cells with an HDAC inhibitor. As *RPD3* most resembles mammalian class I nuclear HDAC enzymes, the BY4742 strain harbouring an *RPD3* deletion was selected. To examine if the loss of *RPD3* would lead to an increase in the frequency of transcription and replication conflicts as predicted with HDAC inhibitors in TK6 cells, markers of these events are mapped throughout the yeast genome using ChIP-chip.  $\gamma$ H2A is an epigenetic marker of DNA double strand breaks and R-loop structures are formed co-transcriptionally. Both of these have been associated with DNA damage arising from conflicts between transcription and replication and therefore are of interest in this study.

Since target enrichment is key to ChIP-chip and depends almost exclusively on the availability and quality of the antibody, 'ChIP-grade' antibodies were preferentially selected based on published studies referencing their use. In this case, both commercially available anti- $\gamma$ H2A and anti-DNA:RNA-hybrid (R-loop) antibodies have been used in microarray studies, warranting their application (Chan et al. 2014; Stirling et al. 2012). Several changes were made to the DRIP-Chip protocol described by Chan *et al.* for mapping R-loops in yeast to reflect developments made surrounding the methodology since its publication. Firstly, since DNA-bound proteins are not the targets of DRIP-Chip, DNA is used instead of chromatin, which is easier to prepare and quantify. Secondly, RNase A digest was carried out at high salt concentrations (400mM NaCl) to remove only single and double stranded RNA but not RNA present in DNA:RNA hybrids prior to immunoprecipitation, as this was previously shown to result in false positive signals (Zheng Z Zhang 2015). Next, a high degree of reproducibility was shown between the two biological repeats (Pearson's  $> 0.9$ ) of  $\gamma$ H2A (Figure 6.1) and R-loop (Appendix, IV). The data was then validated against previously published work.  $\gamma$ H2A enrichment detected in this study followed a similar pattern to those reported by Szilard and

colleagues (Figure 6.2) (Szilard et al. 2010), while R-loop peak sequences are purine-rich and displayed the expected distribution around gene structure (Figure 6.3). Together these observations establish that the data generated is of sufficient quality for further analysis.

For an indication of the overall effect of *rpd3Δ* on  $\gamma$ H2A and R-loop formation in yeast, the net change was examined by comparing peak enrichment between wild type and *rpd3Δ* cells. Figure 6.4 and Figure 6.5 show that *RPD3* deletion significantly increases in the genome-wide  $\gamma$ H2A signal primarily within the open reading frames of genes, maintaining the pattern of distribution. However, the effect of *RPD3* deletion does alter the global levels R-loops (Figure 6.4). Instead, a small number of genes identified in *rpd3Δ* cells express R-loops that are not found in wild type cells and these preferentially accumulate over the open reading frames of genes (Figure 6.6), similar to the observed  $\gamma$ H2A signal (Figure 6.5). This raises the possibility that a small number of regions may show a co-localisation of the two signals, indicative of transcription-replication conflicts.

To investigate this idea, genes showing a 2-fold or greater enrichment in  $\gamma$ H2A or R-loops were compared (Figure 6.7). As a result, a very small subset of genes (9) shows both  $\gamma$ H2A and R-loops enrichment but proved statistically insignificant ( $P = 0.73$ ) (Figure 6.7). To examine whether  $\gamma$ H2A signal exhibited by deletion of *RPD3* are reminiscent of DNA-damage induced  $\gamma$ H2A signals, which can cover regions in excess of 10kb (Szilard et al. 2010; Chan et al. 2014), the average spread was compared. To calculate  $\gamma$ H2A spreading and to account for the uneven probe distribution on the arrays,  $\gamma$ H2A peaks within 1kb were considered to be part of the same signal. Doing so shows no significant increase in  $\gamma$ H2A spreading comparing *rpd3Δ* with wild type cells (Figure 6.8). Relaxing the peak windows beyond 1kb was not considered since  $\gamma$ H2A peaks separated by 2kb or greater may represent two independent signals. However, telomeric and sub-telomeric regions, typically covering ~30kb on either end of each chromosome, show high levels of continuous  $\gamma$ H2A enrichment. These regions are not picked up in the analysis since they contain numerous repeat sequences not represented by array features, resulting in ‘gaps’ in the data, but are apparent upon visual inspection (Figure 6.9). Interestingly these patterns observed in the wild type strain, consistent with previous reports (Szilard et al. 2010), show significant increase upon *RPD3* deletion, which may reflect elevated levels of endogenous stress owing to changes in replication and possibly transcription. Yeast telomeres are hotspots for replication stress as identified by high levels of  $\gamma$ H2A in previous studies (Szilard et al. 2010). In addition,  $\gamma$ H2A is found at

sites with the potential to pose as barriers to replisome progression (Szilard et al. 2010). Further analysis of this shows that  $\gamma$ H2A is highly elevated at replication origins and LTRs in the wild type and that the loss of *RPD3* further promotes  $\gamma$ H2A at these sites (Figure 6.10). Examination of R-loops show that in wild type cells, the observed R-loop enrichment at replication origins co-localise with high levels of  $\gamma$ H2A (Figure 6.10A), suggesting that transcription may be interfering at these sites prone to replication fork stalling and collapse. The same origins sites in wild type cells are enriched for  $\gamma$ H2A in *rpd3 $\Delta$*  cells, in contrast to R-loops between the two strains. Evidence in yeast and human suggests that inhibiting/deleting HDACs places cells under replication stress and as a result dormant origins are activated (Mantiero et al. 2011; Conti et al. 2010). It would therefore be interesting to examine if the 9 genes where  $\gamma$ H2A and R-loops co-localise upon *RPD3* deletion map to dormant origins, in which case it would represent HDAC liable fragile sites. However, the locations of replication origins activated following *RPD3* deletion have not yet been reported. While previous studies also reported  $\gamma$ H2A enrichment at tRNA genes, this was only observed in *rpd3 $\Delta$*  but not in wild type cells in the current study.

Given the emerging implication of R-loops in a range of neurological disorders and cancer, the molecular function of genes showing R-loop formation only in the *rpd3 $\Delta$*  strain was examined. The GO term analysis identified a significant over-representation of genes whose products are involved snoRNA-binding (Figure 6.11). Additionally, the pattern of R-loop accumulation in these genes is more pronounced in the ORF (Figure 6.6) and raises the possibility that these genes may be suppressed, given the heterochromatin state enforced by R-loops (Castellano-Pozo et al. 2013). The role of snoRNAs focuses mainly on applying modifications to pre-rRNA, guided by proteins forming an RNA/protein complex known as the small nucleolar ribonucleoproteins (snoRNPs). Coincidentally Meskauskas and colleagues highlighted the importance of the repressive yeast HDAC complex Sin3-Rpd3 in aiding early ribosome biogenesis, as mutations in *RPD3* results in rRNA processing delays (Meskauskas et al. 2003). The authors speculate that exclusion of histone deacetylase activity silences the transcription of snoRNAs (Meskauskas et al. 2003), though evidence for this is lacking. Instead in the context of evidence presented in this chapter, a possible epigenetic-driven-R-loop based mechanism, leading to the suppression of snoRNA-binding proteins may similarly diminish the activity of snoRNP, thereby affecting rRNA processing. Additional experiments to confirm R-loop formation at these loci in combination with repressive

histone marks and gene expression studies would help to clarify these underlying mechanisms.

Together, the results of the  $\gamma$ H2A and R-loop microarray analysis between wild type and *rpd3 $\Delta$*  yeast cells suggest transcription and replication collisions to be largely absent, following deletion of *RPD3*, at least in regions of the genome where R-loops can be mapped (non-repetitive sites). However, such events may be prevalent at telomeres and sub-telomeres of chromosomes which exhibit high levels of  $\gamma$ H2A spanning the full length of these regions. While the precise role of  $\gamma$ H2A at telomeric and sub-telomeric regions is not entirely understood, Seo *et al.* found that in lymphocytes, sub-telomeres are more prone to endogenous stress than to external DNA-damaging sources such as  $\gamma$ -irradiation and is suggested to mirror transcription and replication induced stress in cycling cells (Jungmin Seo 2012). Furthermore Kim *et al.* suggested that these regions could contain transient double strand breaks (Kim et al. 2007). Therefore, it would be interesting to determine if the  $\gamma$ H2A observed here, which tend to be in excess of 30kb, constitute to DSB induced by transcription-replication conflicts. Chan and others also detected high levels of R-loops at telomeres in yeast using genome wide tiling arrays (Chan et al. 2014). Intriguingly, 4 of the 9 genes identified in this study where  $\gamma$ H2A and R-loops co-localise (Figure 6.7) are found on the ends of chromosomes. Given the limitation of not being able to map events in regions containing repetitive DNA, it is reasonable to assume based on the work by Chan and others that R-loops could be highly enriched at telomeres in the cells used in this study, thereby implicating transcription as a barrier to replication (Chan et al. 2014). Further work mapping proteins involved in transcription, replication, double strand break repair factors such as MRE11 as well as R-loops using tiling arrays could help determine if a loss of HDAC function significantly affects these processes giving rise to the formation of double strand breaks.

## **Chapter VII**

### **General discussion**





The current study aimed to address aspects of drug safety concerned with efforts to utilise epigenetic-targeting HDAC inhibitors for non-life threatening illnesses, following substantial evidence from genotoxicity studies indicating their potential to elicit genetic damage. Central to this is the basic need for a better understanding of the biological mechanisms underpinning these genotoxic effects. Only then can we begin to identify more informative endpoints that could be used to assess risk.

To approach the problem at hand, it is first important to recognise the possible connections between HDAC inhibitors and the onset of DNA damage. To this end, a review of the literature surrounding the effects of HDAC inhibition led to identification of several potential mechanisms of DNA damage induction, namely cytoplasmic effects leading to the generation of reactive oxygen species and nuclear effects resulting in the dysregulation of existing cellular processes to generate damage. With interests in trying to understand the more direct connection between the epigenome and the genome, focus was directed towards examining the nuclear effects of HDAC inhibitors, particularly on chromatin and how these changes may lead to the formation of damage at the genetic level. Chromatin structure plays an important role in regulating transcription and replication, two key processes involved in the expression and propagation of genetic information stored within cells, and because of this, they both compete for the same DNA template. Interestingly, studies have shown that collisions between transcription and replication are inevitable, occurring in prokaryotes as well as eukaryotes (Helmrich, Ballarino, and Tora 2011; Brambati A 2015; Mirkin and Mirkin 2005). The result of such collisions is the presumed formation of DNA double strand breaks along with co-transcriptional by-products such highly stable RNA:DNA hybrid structures, otherwise known as R-loops (Stathaki K.S. 2014). Surmounting evidence in the past few years suggests that while R-loops have regulatory roles within cells, their locations in the genome are closely regulated by a range of mechanisms. Aberrant formation of R-loops on the other hand, can be detrimental to the cell if not resolved properly (Sollier and Cimprich 2015). Not only can R-loops impede the passage of proteins which required access to the DNA but the exposed single stranded DNA is highly susceptible to mutations. As a consequence, aberrant R-loops can lead to various pathological outcomes such as the onset of cancer and certain neurodegenerative diseases (Groh M. and Gromak N. 2014; Groh M. et al. 2014). Therefore, based on what is currently known, the work presented in this thesis aims to investigate the possibility that HDAC inhibitors may elicit DNA damage through changes in the chromatin, resulting in higher incidences of

transcription-replication collisions throughout the genome. If shown to be the case, then a combination of DNA damage and R-loops may lead to genome instability and predispose cells to a disease prone state.

To best answer this question, ChIP-chip was selected as the method of choice for several reasons. Firstly, the technique allows genome-wide measurements of both epigenetic (histone modifications) and genetic (DNA-binding proteins) targets of interest in a ‘top-down’ approach. This is a key advantage since the sites where transcription and replication collisions may take place following HDAC inhibitor treatment are unknown, so it is important to survey large regions of the genome in a non-biased manner. Secondly, bioinformatic tools developed within the lab present a novel method of data normalisation, thus allowing relative comparisons to be made between data gathered from untreated and HDAC inhibitor treated cells. This is important since biological signals such as histone acetylation are present at a basal level in cells and may increase/decrease after HDAC inhibition. In the absence of proper normalisation, interpretation of the data would strictly be limited to binary states, either present or absent (Bennett 2015). Last but not more importantly, the ability to combine and analyse several ‘layers’ of genomic data allows the information-rich data to be broken-down, revealing the unique relationships between the seemingly complex biology and is used as the basis of the mechanistic study presented here.

ChIP-chip, nevertheless still has its drawbacks, namely lacking in resolution compared to ChIP-seq. ChIP-chip also suffers from signal saturation, where each probe on the array has a limited binding capacity. By contrast, ChIP-seq sequences all the fragments present in the sample and therefore does not possess the same limitations. While ChIP-seq is rapidly replacing ChIP-chip as the genome-wide method of choice, Agilent’s offer of cost-efficient custom array designs along the aforementioned advantages make ChIP-chip the ideal method for this project. Additionally, since the aim here is mainly concerned with identify changes in the distribution and patterns of biological signals rather than their precise locations within the genome, the resolution of the microarrays are more than adequate for this requirement.

To initially gain an idea of the genome-wide effects of HDAC inhibition, namely to compare changes in histone acetylation and markers of transcription, replication and DNA damage, the decision was taken to study these in yeast. This presented several advantages, mainly because the ChIP-chip protocol has already been established in yeast (Yu et al. 2011), and since yeast cells grow much faster than mammalian cells yeast array

data can be generated much more quickly than human array data. Therefore, the results of the yeast study were expected to shed light on potential effects that may be observed in human cells following HDAC inhibitor treatment. Early attempts aimed to show the key effects of HDAC inhibitors in yeast cells prior to the use of microarrays. As such the type of inhibitor and doses tested were chosen based on those published in the literature. However, several attempts to measure the effects of HDAC inhibition on histone acetylation and transcription proved unsuccessful (Figure 3.2 – Figure 3.6). Although reduction in cell growth was observed, this may be due to HDAC inhibitors having an effect on the integrity of the cell wall (Figure 3.1). Hence the results of these studies show that in the current study, HDAC inhibitors failed to have any effect at on the chromatin. One possibly is that the yeast cell wall posed as a barrier, limiting the uptake of the drug. A potential solution to this would be to digest the yeast cell wall and treat the resulting spheroplasts with HDAC inhibitors. However, removal of the cell wall would severely compromise cell integrity and the notion of treating ‘sick’ cells with toxic HDAC inhibitors may lead to complications in the microarray data. For this reason, the spheroplast method was not pursued. On top of this, reports in the literature describe varying level of success when trying to use HDAC inhibitors in yeast. The reasons for such discrepancies are unknown, however the observation of histone hyperacetylation in the *RPD3* HDAC mutant strain would provide an alternative genetic approach to study the effects of HDAC inhibition (Chapter VI).

Work in Chapter IV aimed to facilitate the transmission from yeast to mammalian cells by establishing the ChIP-chip method in TK6 cells. The outcome would be the development of a genome-wide tool for measuring the genotoxicity of HDAC inhibitors in a human relevant cell line. For this purpose, it was important to use a cell line that could respond normally to DNA damaging agents. To confirm this, the *TP53* status of TK6 cells used in this study was examined by measuring p53 protein expression and the apoptotic response to nutlin-3. Since the western blot analysis showed p53 expression in both TK6 and MDA-MB-231 cells, to show that the p53 protein found in TK6 cells were active wild-type compared to the inactive mutant found in MDA-MB-231 cells, an apoptosis assay was carried out following treatment with nutlin-3. Nutlin-3 has been shown to induce apoptosis only in cell lines harbouring fully functional wild-type p53 (Pozzo et al. 2013), and is similarly used here to show that the p53 expressed in TK6 cells is indeed functional (Figure 4.1 and Figure 4.2). Next, as attempts to show the effects of HDAC inhibitor treatment in yeast were unsuccessful, it was therefore important to show

these effects in the TK6 cell line prior to progressing onto the microarrays. Since the classical HDAC inhibitor TSA has been previously used in TK6 cells (Olaharski et al. 2006) and that much of the existing genotoxic measurements on HDAC inhibitors were made following a 24h treatment (Johnson and Walmsley 2013; Yoo EJ 2005; Gomez V.P. and Vig B.K. 2002; Olaharski et al. 2006), the aim was to confirm some of these observations under similar conditions in the current study to show that TSA is indeed genotoxic in the TK6 cells used here. To do this, the InVitro MicroFlow® assay was used to measure micronuclei induction as a genotoxic endpoint. Although the assay was validated with the known genotoxic compound Vinblastine as a proof of performance, the result of treating TK6 cells for 24h with TSA showed borderline genotoxicity, given the high levels of cytotoxicity observed (Figure 4.3 and Figure 4.4). The InVitro MicroFlow® assay measures micronuclei formation by using two fluorescent dyes with different staining properties, followed by a series of gating to identify the population of micronuclei 'events'. While this commercial assay has been well validated by the manufacturer (Collin J.E. et al.), it is nevertheless an indirect means of measurement. Therefore, an alternative method would be to measure micronuclei using the more traditional microscopy approach. Given the highly cytotoxic nature of HDAC inhibitors in general, the decision was taken not to use 24h treatments to avoid the possibility of observing cytotoxicity-induced genotoxicity (i.e. DNA damage resulting from apoptosis), which can complicate the actual genotoxic effect of the drug. To take advantage of the fact that drug-induced genotoxicity commonly manifests earlier, TK6 cells were exposed to TSA for shorter periods of time. To examine the genotoxic response elicited by TSA immediately following each time point,  $\gamma$ H2AX was used as an epigenetic marker of DNA damage and measured using flow cytometry. Since a significant increase in both global H3K9ac and  $\gamma$ H2AX was observed in TK6 cells after 6h treatment with TSA, further assessment of annexin-V and cell cycle changes showed a lack of cytotoxicity, leading to the identification of a suitable treatment regimen for studying the genotoxic mechanisms using microarrays (Figure 4.5 and Figure 4.6). The main reason for using flow cytometry is that it presents a rapid method of confirming many of the characteristic effects of HDAC inhibitors, such as cytotoxicity, histone hyperacetylation and phosphorylation, which have all been well-described in the literature (Bose, Dai, and Grant 2014).

Following this, the ChIP-Chip method was successfully established for TK6 cells. This involved a series of optimisation steps to ensure that chromatin from TK6 cells could

be fragmented to the desired size (Figure 4.7 – Figure 4.11). Enzymatic digestion can also be used here to produce chromatin fragments, but sonication is used as it tends to produce a more random population of fragments. Further immunoprecipitation of RNA Pol II (pSer2) showed that the targets of interest could be enriched from the chromatin and this population of DNA fragments (IP) can be PCR amplified with little loss in representation (Figure 4.12). Since PCR-based methods of amplification tend to generate bias with increasing cycles, one possible method would be to use isothermal amplification which overcomes this limitation by amplifying DNA in a linear fashion. The result would typically be more representative of the original sample with a greater signal/noise ratio. Its performance here however, remains to be tested, since the commercial WGA2 kit still provides a reliable signal on the microarray, as shown by the binding of RNA Pol II (pSer2) matching the expected patterns (Figure 4.17) (Hsin JP 2012). Although tools presented in Sandcastle allow normalisation of data from linked experiments, this feature depends solely on the array data having a background-subpopulation which closely follows a standard normal distribution (Bennett 2015). If the data lacks this feature, then the interpretation would be limited to the presence/absence of proteins at distinct binding sites (Bennett 2015). Therefore, a series of quality controls are always carried out to ensure that the data generated in this study meet the requirements for Sandcastle normalisation (Figure 4.14). Using different approaches to validate the RNA Pol II (pSer2) array data (Figure 4.15 – Figure 4.18), showed that the ChIP-chip method established for TK6 cells could be used to reliably map other targets.

Taking this forward, Chapter V describes the work that applies the established ChIP-Chip method to further examine changes in histone H4 acetylation as a marker of euchromatin and ORC1, as a marker of replication initiating zones, on chromosome 17 in response to TSA treatment. The aim was to investigate whether TSA could potentially increase the frequency of collisions between transcription and replication, as a source of DNA damage induction. Although the current study utilises chromosome 17 arrays, the observations made here are expected hold true throughout the genome, since these events are common to all chromosomes. Comparing these data separately showed that TSA promotes changes in the distribution patterns of H4ac (Figure 5.1 and Figure 5.2), RNA Pol II (pSer2) (Figure 5.3 and Figure 5.4) and ORC1 (and Figure 5.8; Figure 5.9). To investigate if there is a potential relationship between H4ac and RNA Pol II (pSer2) occupancy, similarly H4ac and ORC1 occupancy, the correlations between these sets of data were plotted for untreated and TSA treated cells. Interestingly, the analysis showed

that TSA treatment resulted in a strong negative relationship between H4ac and RNA Pol II (pSer2), as well as ORC1 (Figure 5.10 – Figure 5.12). This suggested that in regions of chromosome 17 showing higher levels of H4ac as a result of TSA treatment, would display a loss/decrease in both RNA Pol II (pSer2) and ORC1 occupancy. This observation initially appears to contradict what is widely reported in the literature where higher levels of histone acetylation favour transcription and replication initiation (Kemp et al. 2005). To explain this observation, it is possible that a lack of H4ac may represent either a region occupied by a deacetylated nucleosome or a region which is not occupied by a nucleosome. Comparisons of the ORC1 data with published DHS and G4 data, indicative of nucleosome free regions, show that ORC1 is enriched at these sites (Figure 5.7), suggesting that the lack of H4ac observed may well represent a region deprived of nucleosome. This would also fit in with the idea of euchromatin, since nucleosome free regions would be more accessible to proteins. This proposition would suggest that while global measurements associate higher levels of histone acetylation with increased transcription and replication activity, high resolution approaches show that histone acetylation may in fact be higher in the surrounding regions around transcription and replication rather than specifically at those regions (Figure 5.2A, Figure 5.4A and Figure 5.8A). Although the analysis here supports this view, further experiments mapping nucleosome positions on chromosome 17 in untreated and TSA treated cells would help to clarify these findings.

The changes in ORC1 occupancy observed following TSA treatment are consistent with published observations (Conti et al. 2010). The study used DNA fibre analysis, showed that treating cancer cells with the HDAC inhibitor vorinostat for 4 hours, promoted a slowdown in replication fork speed, concomitant with the induction of  $\gamma$ H2AX and activation of dormant origins (Conti et al. 2010). In relation to the current study, TK6 cells treated with TSA for 6 hours display a significant change in the distribution of ORC1 on chromosome 17 (Chapter V, 5.4.3). This could indicate that new sites of ORC1 binding may similarly represent dormant origin activation. DNA fibre analysis could be carried out on stretches of chromosome 17 in response to TSA to support this view.

To examine the data gathered so far in the context of the hypothesis, the relationship between transcription and replication was examined. The analysis showed that in untreated TK6 cells, transcription and replication processes are spatially separated from each other in different regions on chromosome 17 (Figure 5.15A). However,

following TSA treatment, this spatial organisation is lost, causing transcription and replication processes to be localised to the same regions on chromosome 17 (Figure 5.15B). As a result, it is reasonable to predict that an increase in transcription-replication col-localisation may lead to higher frequencies of collision events taking place between the two (Figure 5.19). However, attempts to demonstrate by mapping  $\gamma$ H2AX and R-loops formation in response to TSA proved unsuccessfully, possibly owing to a lack of sensitivity of the current ChIP-chip assay. Hence, future studies aimed at optimising the assay for the purposes of measuring  $\gamma$ H2AX and R-loops would help to determine if sites of potential collisions are indeed associated with DNA damage.

It should also be noted that a major limitation of using an integrative approach with ChIP-chip data is that relationships between variables are merely inferred. In this case, it is only reasonable to conclude that TSA-induced loss of H4ac is associated with a gain in both RNA Pol (pSer2) and ORC1 occupancy, but does not provide the evidence to indicate a loss of H4ac causes these changes. To make this type of conclusion, further studies abrogating the ability of H4 to undergo acetylation following TSA treatment would demonstrate the role of H4ac in altering transcription and replication.

The adaptability of ChIP-Chip between different organisms opens up the possibility of using budding yeast as a model organism to address the shortfalls of Chapter V. The observed hyper-acetylation in *RPD3* mutants (Chapter III) merits an alternative genetic approach to mimic the effects of treating with a chemical HDAC inhibitor. Therefore, Chapter VI mapped and compared genome-wide  $\gamma$ H2A (ortholog of  $\gamma$ H2AX) and R-loop formation between wild type and *rpd3 $\Delta$*  cells for a possible indication of the predicted transcription-replication conflicts observed in TK6 cells. The key results of this chapter showed that, *rpd3 $\Delta$*  cells displayed a significant increase in genome-wide  $\gamma$ H2A enrichment compared to wild type cells (Figure 6.4A), similar to the measured global  $\gamma$ H2AX increase in TK6 cells. However, R-loop formation did not change significantly in response to treatment (Figure 6.4A). Comparison of both datasets identified  $\gamma$ H2A and R-loop co-localisation within the regions of 9 genes as a result of *RPD3* deletion. However, estimating the peak distribution of the  $\gamma$ H2A signal suggested that those observed in *rpd3 $\Delta$*  cells do not reflect the spreading of the  $\gamma$ H2A signal normally associated with double strand break formation. Instead,  $\gamma$ H2A is highly enriched across the sub-telomeric and telomeric regions on all chromosomes in wild type cells (Figure 6.9) and is consistent with published findings (Chen et al. 2013). Interestingly, in *rpd3 $\Delta$*  cells these same regions show higher levels of  $\gamma$ H2A enrichment, compared to those

detected in wild type cells. Genome-wide mapping of  $\gamma$ H2A has been reported previously as a means of identifying fragile sites (Chen et al. 2013), owing to DNA damage or replication fork stalling. Specifically,  $\gamma$ H2A enrichment was reported at telomeres, which could suggest that these regions are prone to replication stress due to their highly repetitive primary structure making is difficult to replicate and prone to recombination. Other studies have shown that telomeric  $\gamma$ H2A is coupled to downstream heterochromatin imposed by the Sir3 HDAC enzyme (Kitada et al. 2011) and that Rpd3 is required to limit the actions of Sir proteins at telomeres (Ehrentraut et al. 2010). More recently in fission yeast,  $\gamma$ H2A was shown to protect against replication fork collapse when cells faced difficulties in loading PCNA (Mejia-Ramirez et al. 2015). Taken together, these data suggest that, loss of Rpd3 function may result in redundancy, allowing Sir proteins to establish larger regions of heterochromatin at telomeres. This may pose as a barrier to replication, leading to an increase in the frequency of stalled forks, resulting in the elevated levels of  $\gamma$ H2A required to maintain these from collapsing. Further work aimed at mapping histone methylation as marker of heterochromatin may provide evidence to support this idea. Additionally, previous genome-wide studies mapping R-loops in yeast using tiling arrays also reported high levels of enrichment at sub-telomeric regions and telomeres of chromosomes (Chan et al. 2014). A major limitation of the microarrays used in this study is that the locations of a target, which binds to regions of the genome containing repetitive sequences, cannot be reliably determined and therefore are not represented by probes on the arrays. This potentially explains why R-loops were not detected at the ends of chromosomes in the current study, as these tend to contain repetitive elements. The high levels of R-loops at telomeres in wild type cells reported by Chan and colleagues may be the result of RNA Polymerase II transcribed telomeric repeat-containing RNAs (TERA), which forms DNA:RNA hybrids (R-loops) with the template strand (Cusanelli E. and Chartrand P. 2015). Therefore, it would be interesting to see from future studies if R-loops in these regions are elevated following a loss of Rpd3 function using tiling arrays and explore the possibility of conflicts with replication in these regions. Overall, the results of Chapter VI fail to present evidence in support of the claims made in Chapter V using a yeast HDAC mutant. However, since these observations were made using a single HDAC mutant, these effects may be less severe compared to a strain lacking multiple HDAC enzymes. Moreover, the differences between model organisms, create several caveats in the comparison of the yeast mutant data with data derived from TSA treated TK6 cells. As such, TSA-induced transcription-



replication conflicts within the human context remains a possibility, which requires further study. As the development of new HDAC inhibitor molecules stresses isoform specificity to target non-life threatening ailments, it would be interesting to compare the effects of class-specific HDAC inhibitors with those of TSA and see if similar effects are elicited. This would help in trying to address the possible risks posed by these drugs when used for non-life threatening diseases.

The results of the current study present strong evidence to support the view that the pan-HDAC inhibitor TSA promotes the co-localisation of transcription and replication events on chromosome 17 in TK6 cells. As a consequence of this, it is reasonable to predict that the chances of transcription-replication collisions occurring may be higher and therefore act as a potential mechanism of generating DNA damage and R-loops. This may also partly explain why these drugs which were design to target the epigenome can generate damage at the genomic level.

The findings in this thesis has several implications within the field of safety assessment which may be in the interests of GSK. First and foremost, there is a potential for the mechanistic understanding gained here to be translated into a unique genotoxicity assay for risk assessment. As our understanding of R-loops has increased a great deal over the last few years, and it is becoming clear that aberrant accumulation of R-loops can lead to cancer and neurological diseases (Stathaki K.S. 2014; Santos-Pereira and Aguilera 2015; Sollier and Cimprich 2015). Since R-loops are susceptible to mutations, further work identifying a correlation between R-loop formation and mutations in the genetic sequence with sites of transcription and replication collisions identified in the present study would provide mechanistic evidence to support R-loops as an endpoint for risk assessment. Alternatively, it is conceivable that genomic changes in R-loop formation could be a universal indicator of genotoxicity, induced either by direct or indirect effects of drug treatment. This could also be analysed. These findings can be translated into routine assays, coupled with high-throughput screening. For instance, cells can be stained using a fluorescently-tagged R-loop antibody following drug treatment and the signal measured using flow cytometry. On top of this thresholds may be set to indicate levels of R-loops which represent a genotoxic response. Examples of existing toxicity assay that utilise a similar approach include the GreenScreen® HC assay (Cyprotex PLC) and the ToxTracker® assay (Toxys).

A second potential application of these findings is the use of CHIP-chip described here for TK6 cells as a genome-wide assay for measuring DNA damage and repair.

Optimisations to the protocol would allow  $\gamma$ H2AX to be mapped at high-resolutions across the genome. Existing genotoxicity assays either examine endpoints at the global level (i.e. gross chromosomal changes or  $\gamma$ H2AX using flow cytometry or western blotting) or at specific loci (i.e. Pig-a assay). A high-resolution approach offers several advantages, including the ability to examine events at many different genomic loci that would otherwise be refractory to detection by conventional assays. Additionally, by normalising the data using Sandcastle, the loss of  $\gamma$ H2AX following the removal of the drug can be used as an indication of repair. This may serve as an alternative endpoint for risk assessment, as this feature is currently lacking in the tests available. It is important to note that  $\gamma$ H2AX induction is often used as the gold standard for indicating the presence of DNA damage (Kuo LJ 2008). However, emerging evidence has shown that  $\gamma$ H2AX also has a role in protecting replication forks from collapse at fragile sites (Szilard RK 2010; Turinetto V. and Giachino C. 2015). Based on this, it may be unreliable to interpret a higher global level of  $\gamma$ H2AX as a sole indication of DNA damage. Therefore a high-resolution approach would help to map and distinguish between the sites of replication stress and DNA damage by way of examining the different  $\gamma$ H2AX profiles, since  $\gamma$ H2AX associated with replication stress tend to be more localised compared to the more diffuse  $\gamma$ H2AX signal associated with double strand breaks (Szilard RK 2010).

In conclusion, the work in this thesis has shown that genome-wide approaches such as ChIP-chip can be an effective tool for unravelling the potential mechanisms of DNA damage induction by a class of epigenome-targeting drugs. It has also provided a novel understanding of how interfering with the epigenome in ways which we do not fully comprehend, can potentially lead to negative effects in the genome.

## Appendix I – Raw data for Chapter III

Raw data for: Figure 3.1

Time	Untreated				
	1	2	3	Mean	SEM
0	0.087	0.265	0.188	0.18	0.05153963
1	0.112	0.32	0.211	0.21433333	0.06006755
2	0.187	0.342	0.246	0.25833333	0.04516759
3	0.381	0.576	0.431	0.463	0.05847602
4	0.441	0.67	0.584	0.565	0.06678573
5	0.734	0.986	0.879	0.866	0.07302131
6	1.14	1.103	0.978	1.074	0.04901134
7	1.41	1.162	1.331	1.301	0.07314597
8	1.74	1.407	1.62	1.589	0.09737043
9	1.74	1.68	1.58	1.667	0.04666667

Time	10 $\mu$ M TSA				
	1	2	3	Mean	SEM
2	0.211	0.112	0.32	0.214	0.06006755
3	0.25	0.165	0.38	0.265	0.06251666
4	0.36	0.271	0.429	0.353	0.04573231
5	0.732	0.387	0.51	0.543	0.10095048
6	0.936	0.535	0.831	0.767	0.12005601
7	1.24	0.867	1.282	1.130	0.13189179
8	1.137	1.13	1.431	1.233	0.09918725
9	1.46	1.34	1.55	1.450	0.06082763

Time	100mM VPA				
	1	2	3	Mean	SEM
2	0.32	0.112	0.211	0.214	0.06006755
3	0.342	0.118	0.227	0.229	0.06467096
4	0.331	0.223	0.242	0.265	0.0332883
5	0.376	0.234	0.264	0.291	0.04321008
6	0.381	0.242	0.282	0.302	0.04131317
7	0.391	0.238	0.33	0.320	0.04446847
8	0.389	0.233	0.301	0.308	0.04515652
9	0.392	0.266	0.312	0.323	0.03681183

Time	100mM Na-B				
	1	2	3	Mean	SEM
2	0.211	0.32	0.112	0.214	0.06006755
3	0.284	0.343	0.128	0.252	0.06413614
4	0.331	0.411	0.231	0.324	0.05206833
5	0.467	0.487	0.38	0.445	0.03284475
6	0.522	0.554	0.413	0.496	0.04267838
7	0.73	0.683	0.56	0.658	0.05068311
8	0.946	0.831	0.776	0.851	0.05008326
9	1.253	1.093	1.149	1.165	0.04687572

Raw data for: Figure 3.2

	Average of two quantifications		H4K12-ace/	
	a-tubulin	H4K12-ace	a-tubulin	Normalised to untreated
Untreated	31138.43	41113.087	1.32	1.00
10µM TSA	32422.48	36353.652	1.12	0.85
100mM Na-B	22030.723	39960.187	1.81	1.37
100mM VPA	27357.137	28292.187	1.03	0.78
WT	45503.907	27022.401	0.59	1.00
rpd3-delta	44342.401	102638.08	2.31	3.90

Raw data for: Figure 3.4

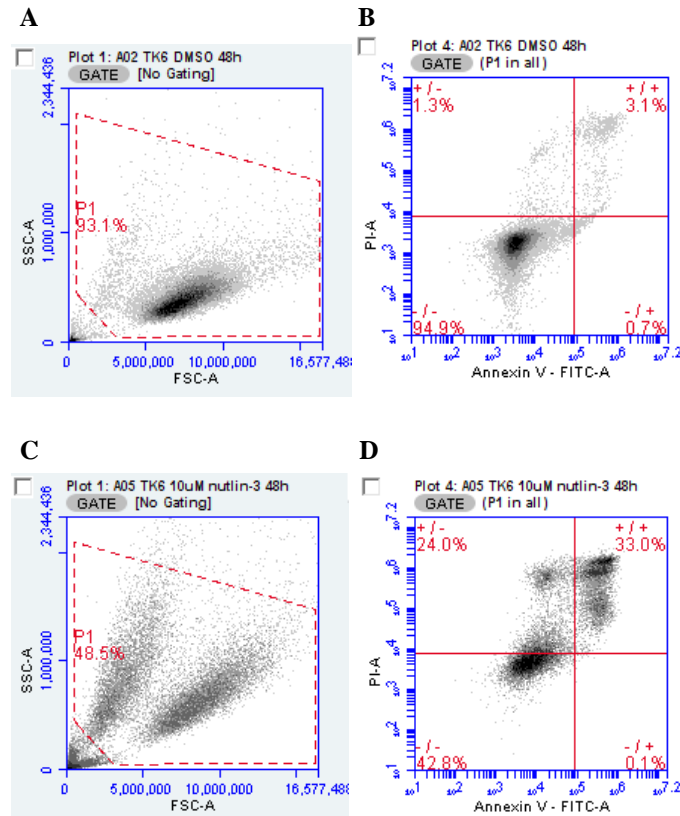
	Average of two H4K12ac			Normalised to DMSO
	TEL06R	INO1	INO1/TEL06R	
DMSO	4.25	21.306	5.01	1.00
na-but	5.619	17.939	3.19	0.64
	Average of two H3K9			
	TEL06R	GAL1	GAL1/TEL06R	
DMSO	5.562	40.455	7.27	1.00
na-but	7.988	28.323	3.55	0.49

Raw data for: Figure 3.6

	FRE1					SEM
	1	2	3	Ave	Fold	
DMSO	1.00029252	1.03451532	1.15099347	1.06193377	1	
10uM	0.79751201	1.00087782	1.0551341	0.95117465	0.89570053	0.0784115
50uM	1.298	0.53291628	1.294	1.04163876	0.98088863	0.25436386
	FRE7					
	1	2	3	Ave	Fold	
DMSO	1.02920879	0.9941639	0.99738657	1.00691976	1	
10uM	1.03576228	1.02782418	1.2235707	1.09571905	1.08818905	0.06396688
50uM	0.63308828	0.60852779	1.259	0.83353869	0.82781044	0.21284877
	CTR1					
	1	2	3	Ave	Fold	
DMSO	1.22847947	1.15313435	1.14353265	1.17504882	1	
10uM	1.16724359	1.2610872	1.18622902	1.20485327	1.02536443	0.02864613
50uM	1.567	1.20966421	1.274	1.3502214	1.14907685	0.10996893

## Appendix II – Raw data for Chapter IV

Example data for: Figure 4.1



**Detecting apoptosis in TK6 cells following 48h treatment with nutlin-3. DMSO vehicle control: A – B and 10 $\mu$ M nutlin-3 48h: C – D. A & C:** Gating to exclude events associated with debris (P1). **B & D:** Applying a quadrant gate in a plot of PI vs. Annexin-V in order to distinguish: health cells (bottom left), dead cells (top left), dead and dying cells (top right); cells undergoing apoptosis (bottom right). PI – propidium iodide used to stain DNA associated with dead cells. Annexin-V used as a marker of early apoptosis.

Raw data for: Figure 4.1

		TK6					
Batch	1		2		3		
	24	48	24	48	24	48	
U	2.3	4.1	7.2	3.6	4.3	4.7	
DMSO	5.8	3.8	7.5	3.5	5.1	4.3	
1uM	7.4	6.8	14.5	10.6	7.8	6.9	
5uM	15.9	23.8	24.1	34.4	18.8	23.3	
10uM	21.5	33.1	32.2	50.2	27.3	38.5	

		TK6			
Average	24	SEM	48	SEM	
	U	4.60	1.42	4.13	0.32
DMSO	6.13	0.71	3.87	0.23	
1uM	9.90	2.30	8.10	1.25	
5uM	19.60	2.40	27.17	3.62	
10uM	27.00	3.09	40.60	5.05	

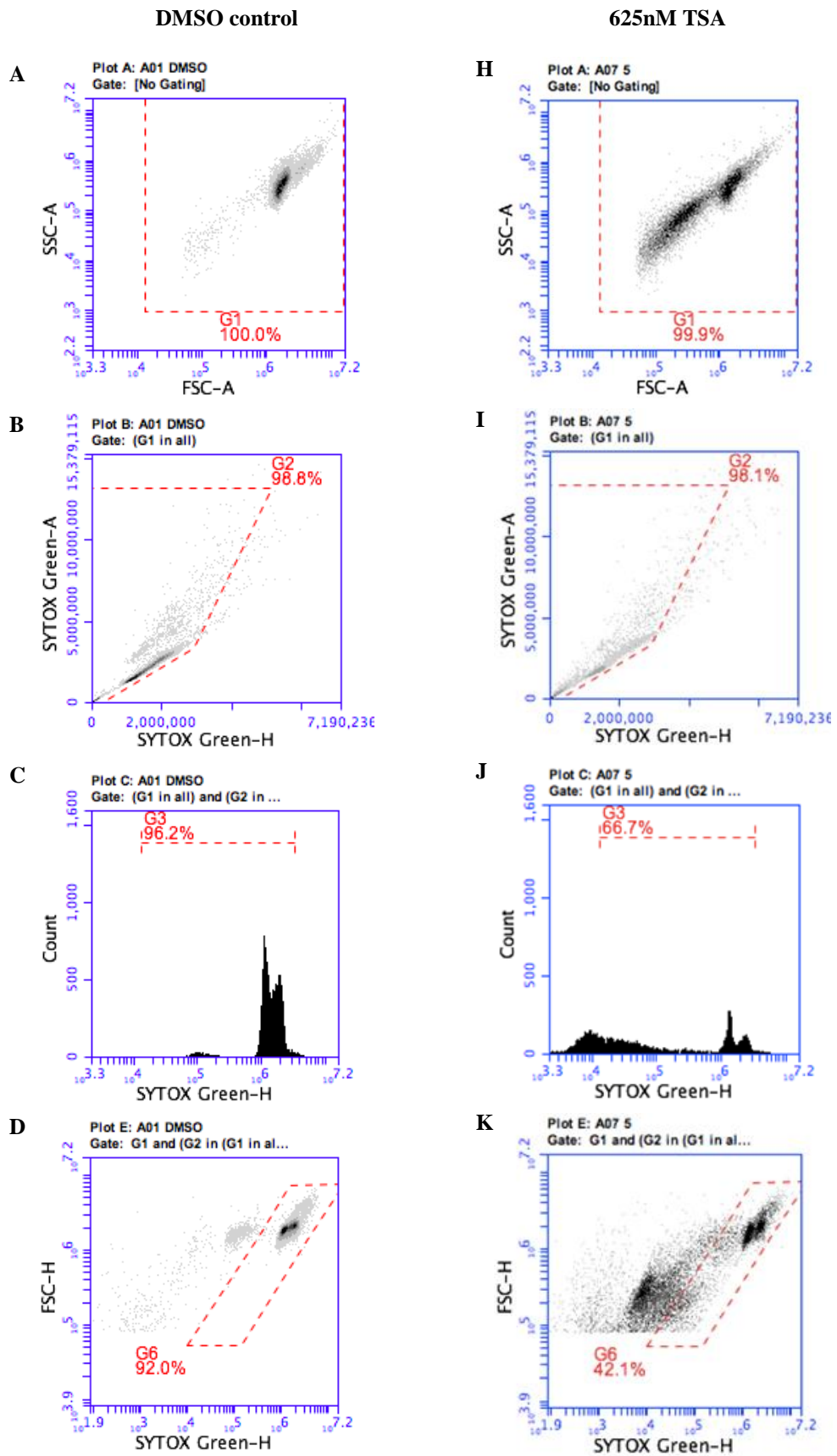
		MDA-MB-231					
Batch	1		2		3		
	24	48	24	48	24	48	
U	0.5	12.1	0.2	2.6	10.2	5.2	
DMSO	0.7	13.1	0.7	3.4	10.8	2.6	
1uM	0.7	15	0.2	4.9	11.5	3.7	
5uM	0.5	11.6	0.2	5.1	8.2	4.2	
10uM	0.4	12	0.4	3.2	9.6	5.2	

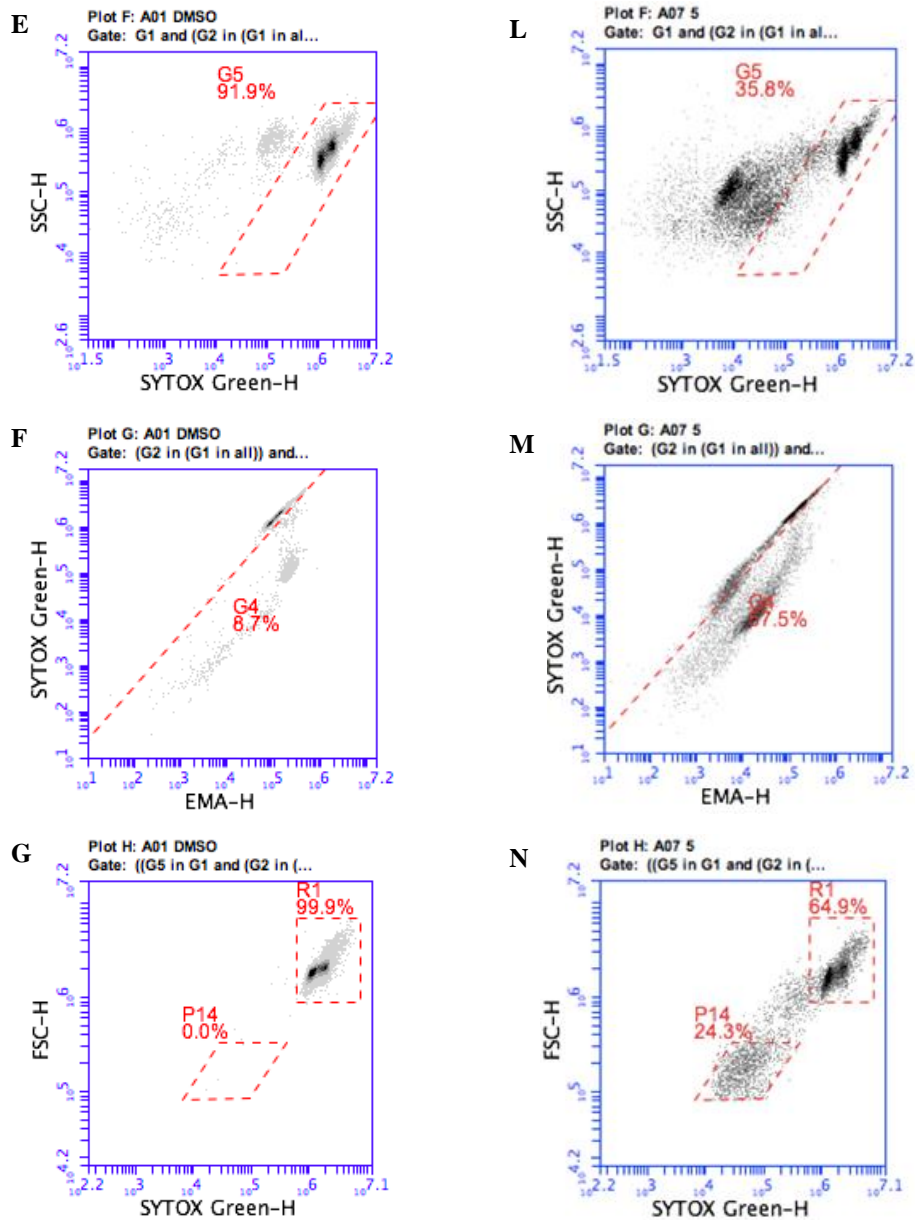
		MDA-MB-231			
Average	24	SEM	48	SEM	
	U	3.63	3.28	6.63	2.83
DMSO	4.07	3.37	6.37	3.37	
1uM	4.13	3.69	7.87	3.58	
5uM	2.97	2.62	6.97	2.33	
10uM	3.47	3.07	6.80	2.66	

		PC-3					
Batch	1		2		3		
	24	48	24	48	24	48	
U	1.6	3.6	2.6	5.2	2.4	4.7	
DMSO	1.4	3.9	2.8	2.9	3.6	2.8	
1uM	1.4	5.4	2.6	4.8	3.8	2.6	
5uM	2.1	4.6	3.7	4.7	2.7	6.1	
10uM	1.7	5.4	3.5	4.6	3.9	7.8	

		PC-3			
Average	24	SEM	48	SEM	
	U	2.20	0.31	4.50	0.47
DMSO	2.60	0.64	3.20	0.35	
1uM	2.60	0.69	4.27	0.85	
5uM	2.83	0.47	5.13	0.48	
10uM	3.03	0.68	5.93	0.96	

Example data for: Figure 4.3





**Gating procedure involved in the InVitro MicroFlow® assay for the representative samples: DMSO vehicle control (A – G) and 625nM TSA (H – N). A & H: Gating (G1) to select all events in plot. B & I: Doublet discrimination (G2). C & J: All events above  $10^4$  are selected (G3) to ensure the inclusion of MN events, while events below  $10^4$  corresponding to small fragmented DNA and are omitted. D & K: The exclusion of events (DNA) falling outside a defined region associated with dead/dying cells based on forward scatter (G6). E & L: The exclusion of events (DNA) falling outside a defined region associated with dead/dying cells based on side scatter (G5). F & M: Events falling below G4 are EMA +. G & N: MN falling within a distinct region (P14) can be discerned from the nuclei population (R1). SYTOX Green – stains all DNA, EMA – stains DNA associated with dead cells.**



Raw data for: Figure 4.3

% EMA -						
Dose (uM)	1	2	3	Average	N to 100	SEM
0	91.3	87.6	92.7	90.53	100	1.52
0.019	85.2	92.6	92.8	90.20	99.63181149	2.50
0.039	88.8	90.2	90.5	89.83	99.22680412	0.52
0.078	81.7	80.1	79.9	80.57	88.99116348	0.57
0.156	59.9	69.1	66.8	65.27	72.09131075	2.76
0.313	52.7	55.6	53.1	53.80	59.42562592	0.91
0.625	42.5	48.1	15.6	35.40	39.10162003	10.03

% MN					
Dose (uM)	1	2	3	Average	SEM
0	0	0.1	0.1	0.07	0.033
0.019	0.1	0.2	0.1	0.13	0.033
0.039	0.4	0.4	0.4	0.40	0.000
0.078	1.6	0.9	0.9	1.13	0.233
0.156	3.9	3.2	4.9	4.00	0.493
0.313	17.9	9.9	10.3	12.70	2.603
0.625	24.3	10.6	27.7	20.87	5.226

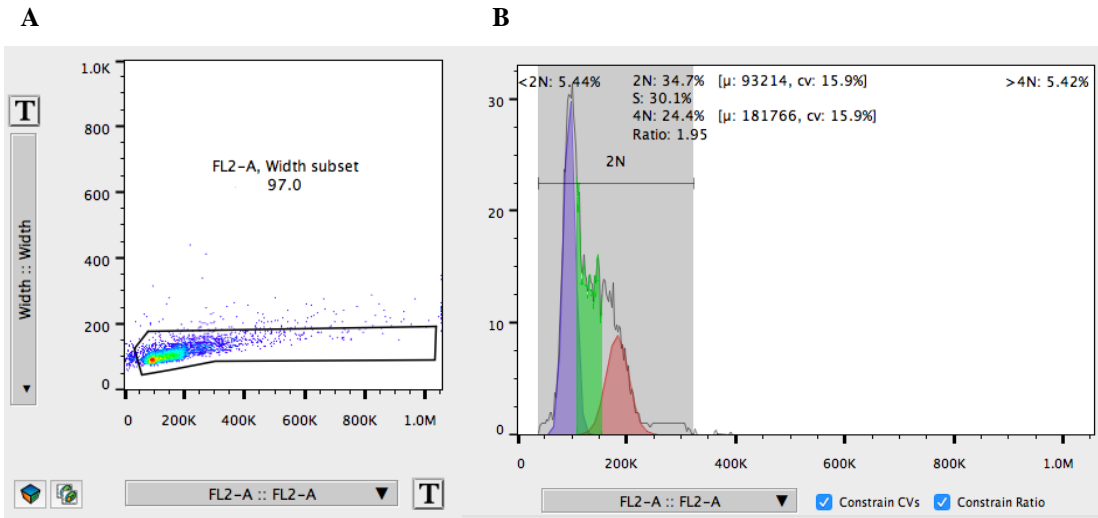
% RCC					
Dose (uM)	1	2	3	Average	% RCC
0	0.8	0.49	0.59	0.63	100.00
0.019	0.63	0.48	0.3	0.47	75.00
0.039	0.58	0.46	0.37	0.47	75.00
0.078	0.34	0.31	0.32	0.32	51.60
0.156	0.44	0.25	0.25	0.31	50.00
0.313	0.3	0.28	0.23	0.27	43.09
0.625	0.24	0.39	0.3	0.31	49.47

ChIP antibodies used in this study

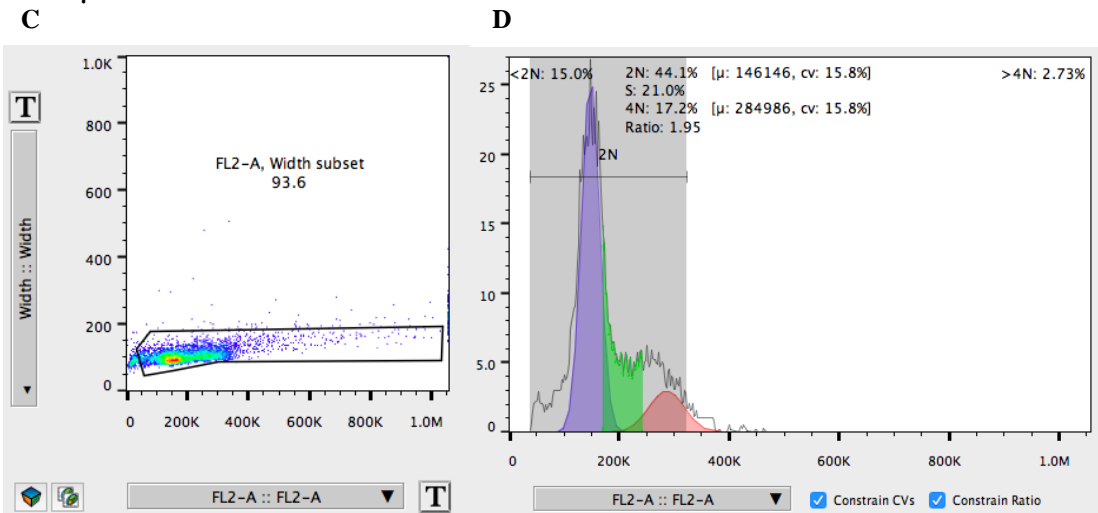
Antibody	Manufacturer	Product code	Species reactivity
anti-H4K12ac	Merk Millipore	07-595	Yeast
anti-H3K9ac	Merk Millipore	ABE18	Human/Yeast
anti-RNA Pol II CTD repeat YSPTSPS (phospho S2)	abcam	ab5095	Human/Yeast
anti-OCR1	abcam	ab60	Human
anti-H4K5ac, K8ac, K12ac, K16ac	Diagenode	C15410024	Human
anti-gamma H2A.X (phospho S139)	abcam	ab2893	Human
anti-Histone H2A (phospho S129)	abcam	ab15083	Yeast
anti-DNA-RNA Hybrid [S9.6]	Kerafast	ENH001	Human/Yeast

Example data for: Figure 4.4

**DMSO control**



**0.625 $\mu$ M TSA**



**Cell cycle analysis using EMA data from the InVitro MicroFlow® assay. A + B:** Data from DMSO vehicle control. **C + D:** Data from 0.625 $\mu$ M TSA. **A + C:** Exclusion of aggregated cells or doublets. **B + D:** Cell cycle profiles fitted using the Watson model, with constrains applied CV's and ratio.

Raw data for: Figure 4.4

replicate 1			
Concentration (uM)	G0/1	S	G2/M
0	41.13	27.31	31.56
0.019	47.56	26.95	25.49
0.039	42.89	29.77	27.34
0.078	48.05	28.54	23.41
0.156	44.50	34.45	21.05
0.313	36.03	45.73	18.23
0.625	44.66	33.29	22.05
replicate 2			
Concentration (uM)	G0/1	S	G2/M
0	34.70	30.10	24.40
0.019	41.03	35.09	23.88
0.039	42.43	29.54	28.03
0.078	44.38	34.20	21.42
0.156	44.65	35.32	20.02
0.313	43.46	35.80	20.74
0.625	44.10	21.00	17.20
replicate 3			
Concentration (uM)	G0/1	S	G2/M
0	35.54	41.34	23.12
0.019	34.46	45.30	20.24
0.039	41.03	33.94	25.03
0.078	45.74	29.40	24.85
0.156	45.87	32.63	21.49
0.313	45.14	30.71	24.15
0.625	54.56	28.34	17.10

Concentration (uM)	Mean			SEM		
	G0/1	S	G2/M	G0/1	S	G2/M
0	37.12	32.92	26.36	1.93	5.06	3.14
0.019	41.02	35.78	23.20	3.78	5.31	1.55
0.039	42.12	31.08	26.80	0.56	1.43	0.91
0.078	46.06	30.71	23.23	1.07	1.76	1.00
0.156	45.01	34.14	20.85	0.43	0.79	0.43
0.313	41.55	37.41	21.04	2.80	4.41	1.71
0.625	47.77	27.54	18.79	3.12	2.02	1.44

Raw data for: Figure 4.5

H3Kac					
	3h			Mean	SEM
DMSO	2.9	1.7	1.1	1.90	0.49
400nM	7.4	7.1	13	9.17	0.12
6h					
DMSO	1.1	0.8	1.7	1.20	0.12
400nM	6.3	15.3	12.5	11.37	3.67
12h					
DMSO	1.9	2.2	1.4	1.83	0.12
400nM	38.3	35.9	21.7	31.97	0.98
24h					
DMSO	1.1	2.1	1.7	1.63	0.41
400nM	8.9	1.1	10.9	6.97	3.18

γH2AX					
	3h			Mean	SEM
DMSO	5.10	4.50	4.90	4.83	0
400nM	12.00	19.10	15.50	15.53	2.05
6h					
DMSO	7.20	2.60	4.20	4.67	1.35
400nM	25.00	13.80	22.50	20.43	3.39
12h					
DMSO	4.60	10.80	6.40	7.27	1.84
400nM	24.30	27.40	25.10	25.60	0.93
24					
DMSO	5.50	4.00	5.70	5.07	0.54
400nM	41.20	45.10	41.90	42.73	1.20

Raw data for: Figure 4.6

Apoptosis						
Dose	1	2	3	Mean	Normalised to 0	SEM
0	3	4.9	4.4	4.10	1	0.57
0.1	3.9	3.1	5.4	4.13	1.01	0.67
0.2	3.1	3.7	4.8	3.87	0.94	0.50
0.4	3.9	3.7	4.6	4.07	0.99	0.27

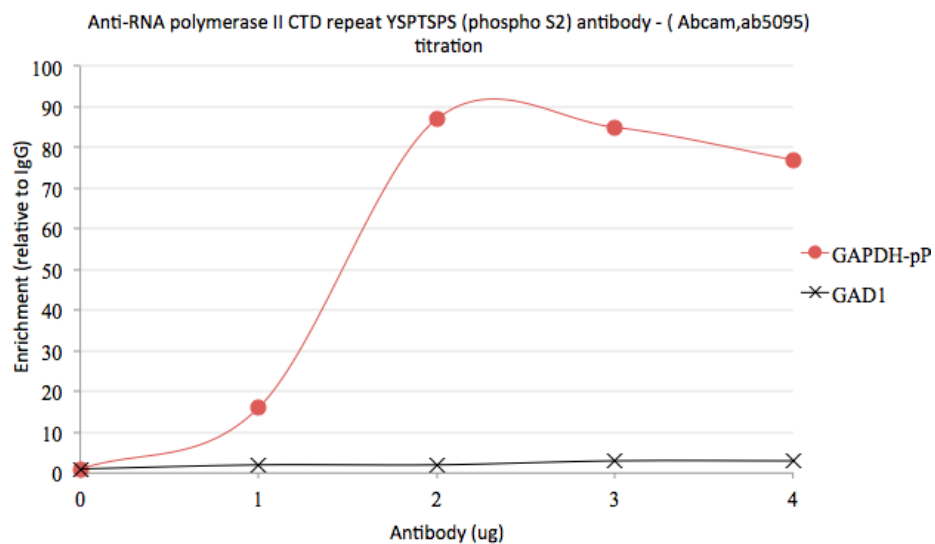
G1 Arrest						
	1	2	3	Mean	Normalised to U	SEM
DMSO	31.9	30.9	30.1	30.97	1.00	0.52
100	30.2	32.8	34.7	32.57	1.05	1.30
200	32.8	35.7	33.5	34.00	1.10	0.87
400	31.6	31.1	33.1	31.93	1.03	0.60

Raw data for: Figure 4.12

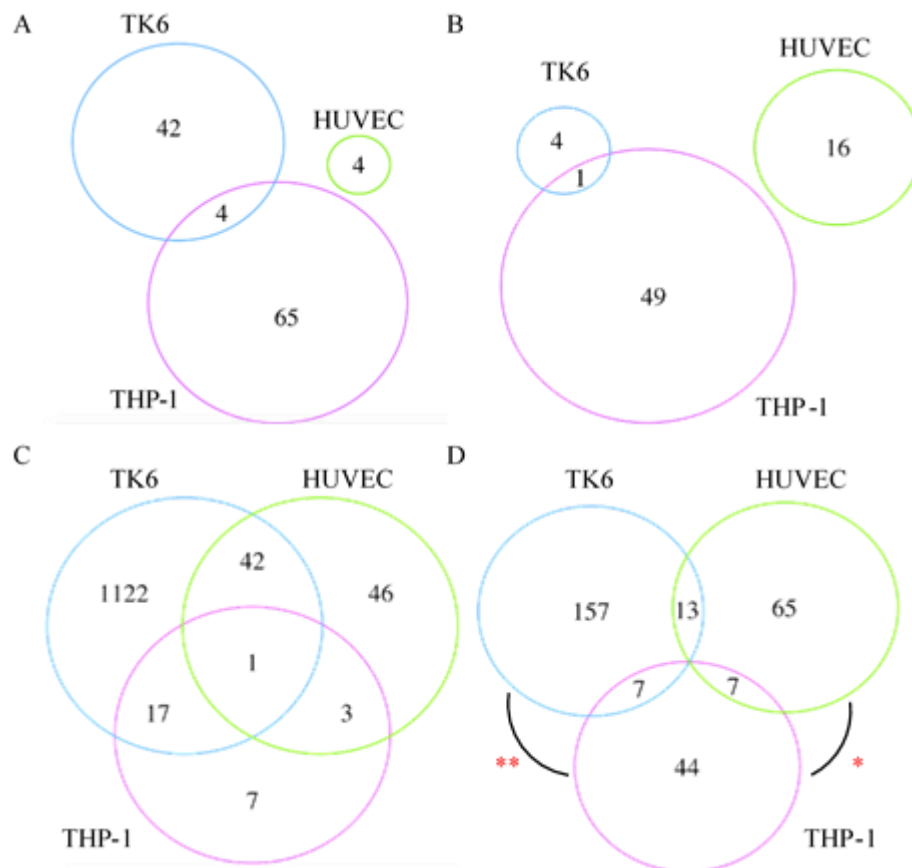
	pre-WGA2 (IP/IN)				Mean	SEM
	1	2	3	4		
GAD1	4.83	5	10.23	3	5.77	1.56
MYOD1	3.79	20.34	21.02	3	12.04	4.99
GAPDH-pP	74.58	72.54	23.80	38	52.23	12.66
RPL10-pP	69.11	73.39	135.21	12	72.43	25.17

	post-WGA2 (IP/IN)			Mean	SEM
	1	2	3		
GAD1	2.20	1.59	1.39	1.73	0.24
MYOD1	7.54	2.16	0.49	3.40	2.13
GAPDH-pP	41.96	13.45	17.10	24.17	8.96
RPL10-pP	43.66	28.72	19.92	30.77	6.93

RNA Pol II (pSer2) titration:



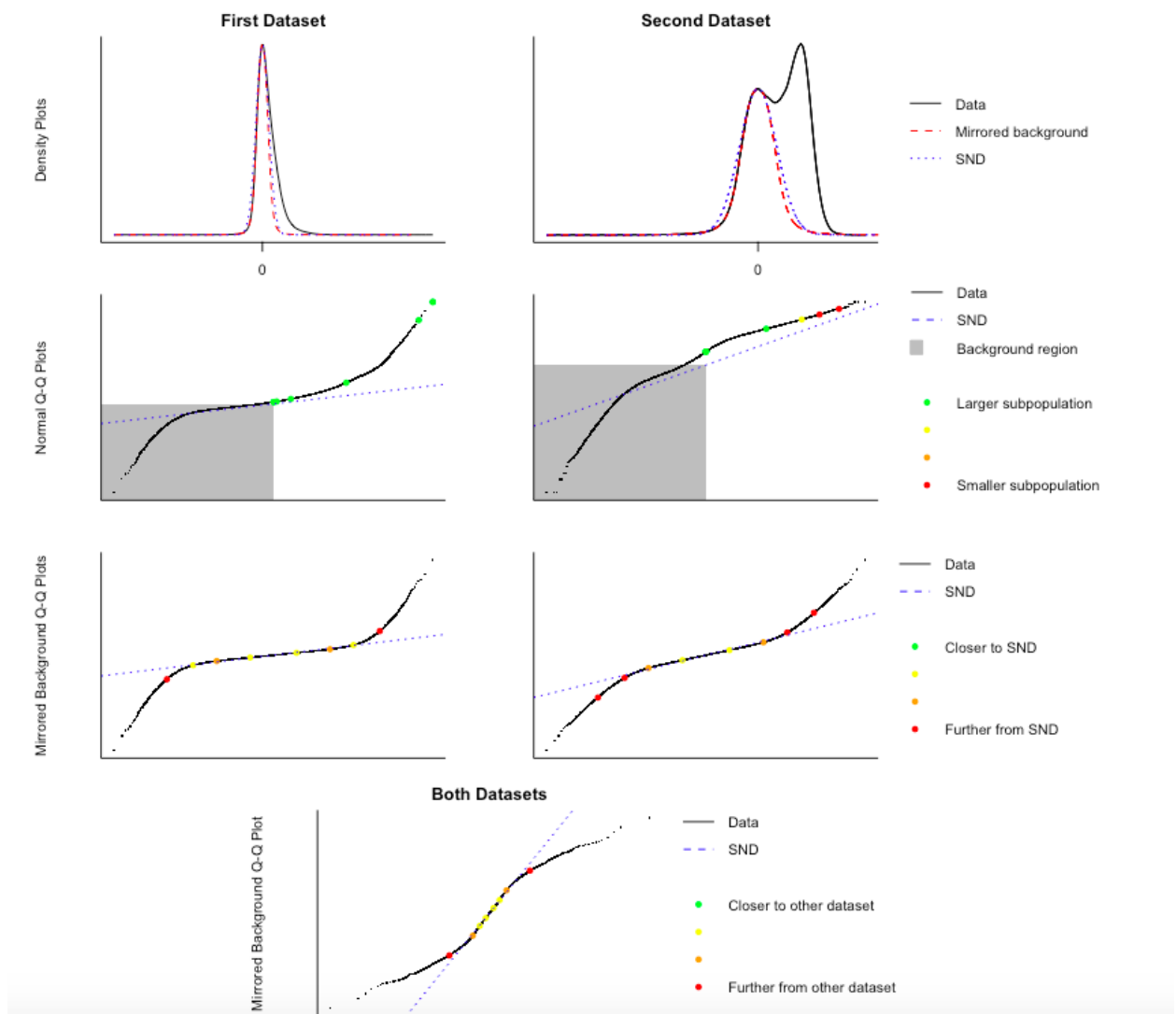
**Titration of the anti-RNA Pol II (pSer2) antibody from 0 – 4µg against a fixed volume of chromatin.** Target enrichment was examined at the proximal promoter region of *GAPDH* and genic region of *GAD1*.



**Comparison of changes in RNA Pol II (phosphor-Ser2) at genes on chromosome 17 in response to TSA with published gene expression data.** Venn diagrams examining similarities between genes showing an increase in RNA Pol II (phosphor-Ser2) occupancy and expression at **A**: FDR < 0.01 and **C**: FDR < 0.05 and similarly those showing a decrease at **B**: FDR < 0.01 and **D**: FDR < 0.05. For ChIP-Chip data, the average log<sub>2</sub> IP/IN ratio for each gene is calculated for normalised untreated and TSA treated replicates. A t-test is performed on a gene-to-gene basis and corrected for multiple testing using the Benjamini & Hochberg method, otherwise known as the false discovery rate (FDR). The means before and after treatment are compared to classify genes as either showing an increase or decrease in relative signal. The same analysis is performed online using GEO2R with publicly available data from GEO, for THP-1 cells treated with 300nM TSA for 150mins (GSE36323) and HUVEC cells treated with 500nM TSA for 24h (GSE5856). Overlaps were identified based on matching gene names using a web-based tool (available at <http://bioinformatics.psb.ugent.be/webtools/Venn/>) and drawn using R-based ‘VennDiagram’ package v1.6.17. Statistical analysis (Hypergeometric test) was performed using the R-based ‘stats’ package v3.2.4. \*P < 0.05, \*\*P < 0.01.

## Appendix III – Raw data for Chapter V

### Normalisation assumptions for H4ac data



**H4ac microarray data meets the assumptions for normalisation.** First dataset – from untreated cells. Second dataset – from TSA treated cells. *First row:* density plots of the mirrored background approximating to the standard normal distribution (SND). *Second row:* density plots shown in A represented as Q-Q plots. Data in the background (shaded) region should closely follow the SND, whereas deviations of the right tail represents the enriched sub-population. *Third row:* Q-Q plot of only the mirrored background data. *Fourth row:* Q-Q plot of the mirrored background for both datasets. Little/no departure from the SND indicates that the two datasets are suitable to be normalised together.

Correlation values between biological replicates of H4ac and ORC1 microarray data

H4ac Untreated			
	1	2	3
1		0.61	0.59
2			0.64
3			

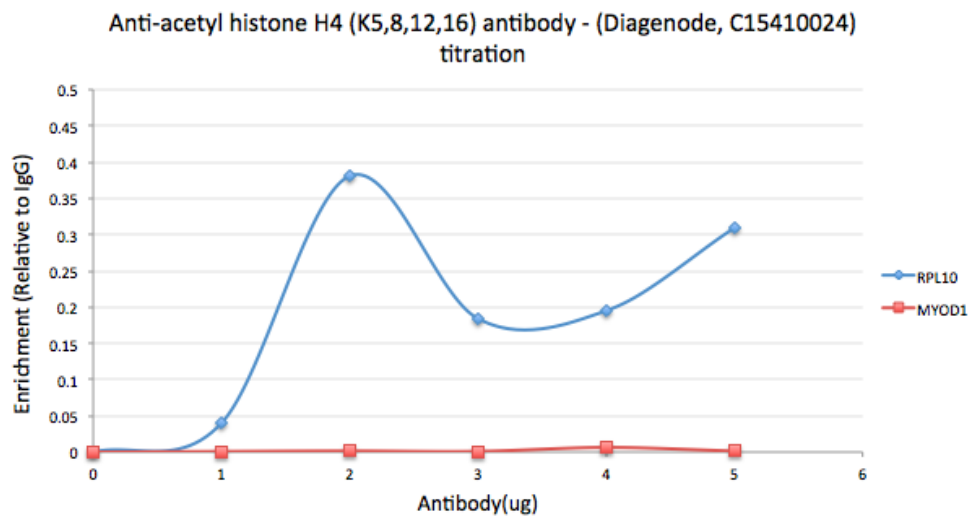
ORC1 Untreated			
	1	2	3
1		0.95	0.95
2			0.95
3			

H4ac TSA treated			
	1	2	3
1		0.92	0.91
2			0.92
3			

ORC1 TSA treated			
	1	2	3
1		0.92	0.9
2			0.91
3			

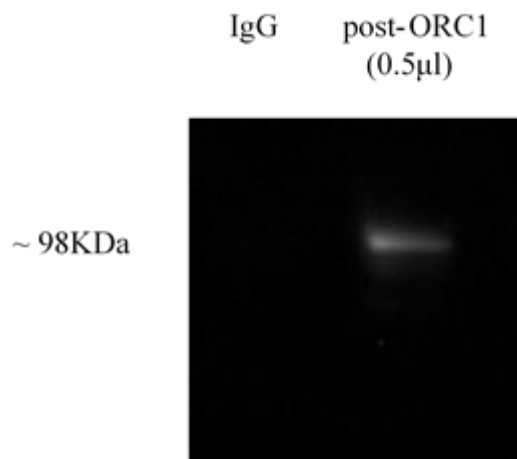
H4ac titration:



Titration of the anti-pan H4ac antibody from 0 – 5µg against a fixed volume of chromatin.

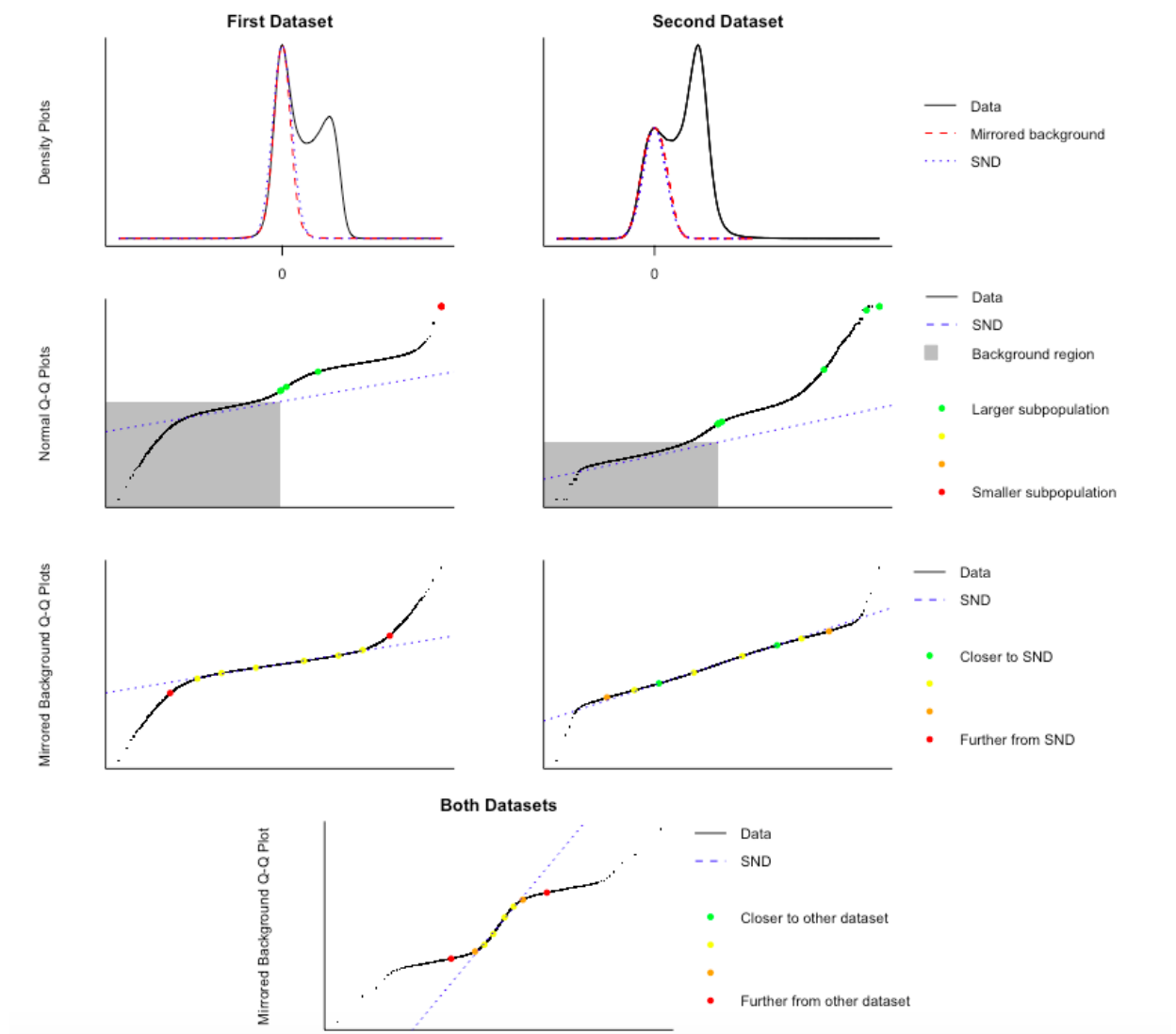
Target enrichment was examined at the genes *RPL10* and *MYOD1*.

ORC1 post-IP western to check for antibody specificity:



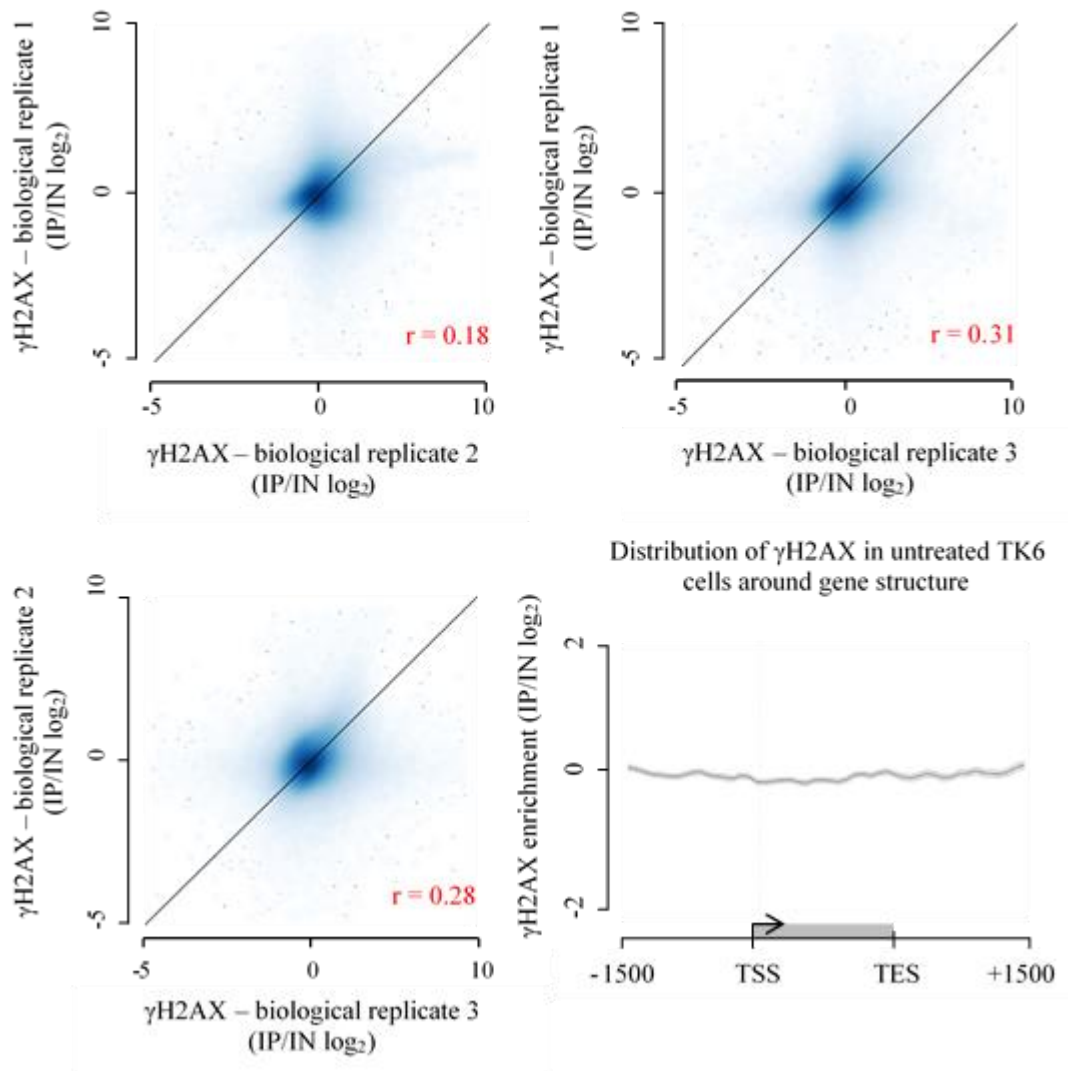


## Normalisation assumptions for ORC1 data

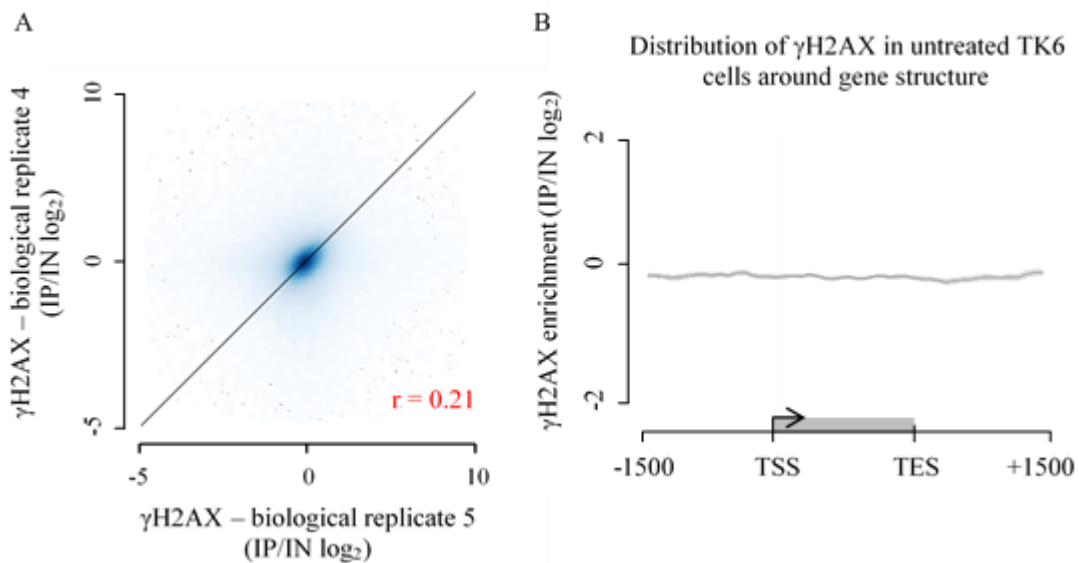


**ORC1 microarray data meets the assumptions for normalisation.** First dataset – from untreated cells. Second dataset – from TSA treated cells. *First row:* density plots of the mirrored background approximating to the standard normal distribution (SND). *Second row:* density plots shown in A represented as Q-Q plots. Data in the background (shaded) region should closely follow the SND, whereas deviations of the right tail represents the enriched sub-population. *Third row:* Q-Q plot of only the mirrored background data. *Fourth row:* Q-Q plot of the mirrored background for both datasets. Little/no departure from the SND indicates that the two datasets are suitable to be normalised together.

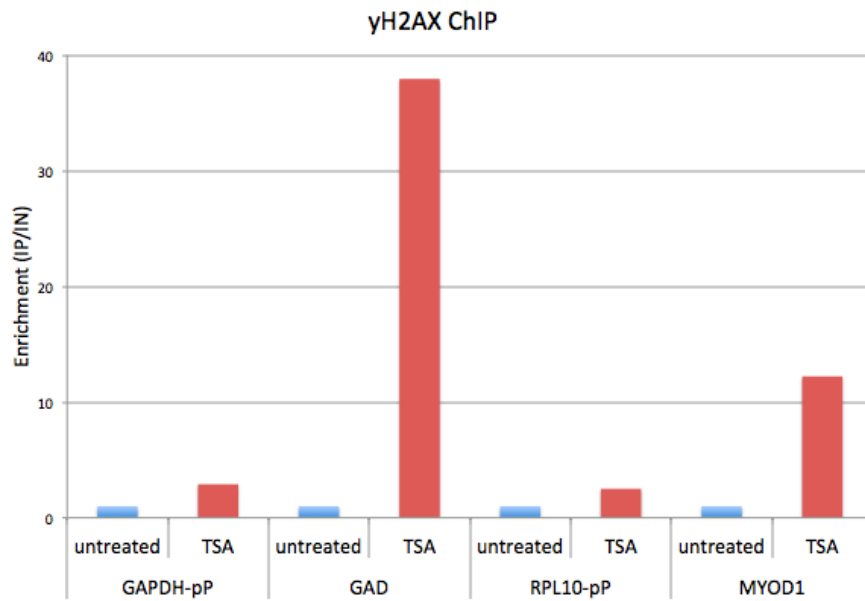
## Biological replicates of $\gamma$ H2AX



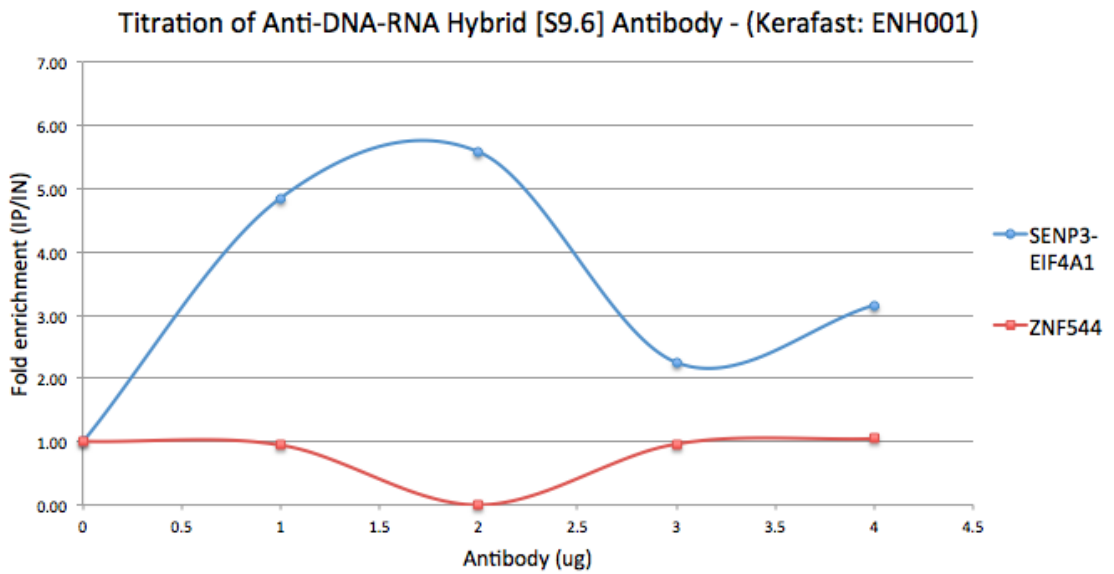
## Biological replicates of $\gamma$ H2AX – following changes to protocol



$\gamma$ H2AX enrichment following TSA treatment measured by ChIP-qPCR  
 - Anti-gamma H2A.X (phospho S139) antibody - ChIP Grade (ab2893)

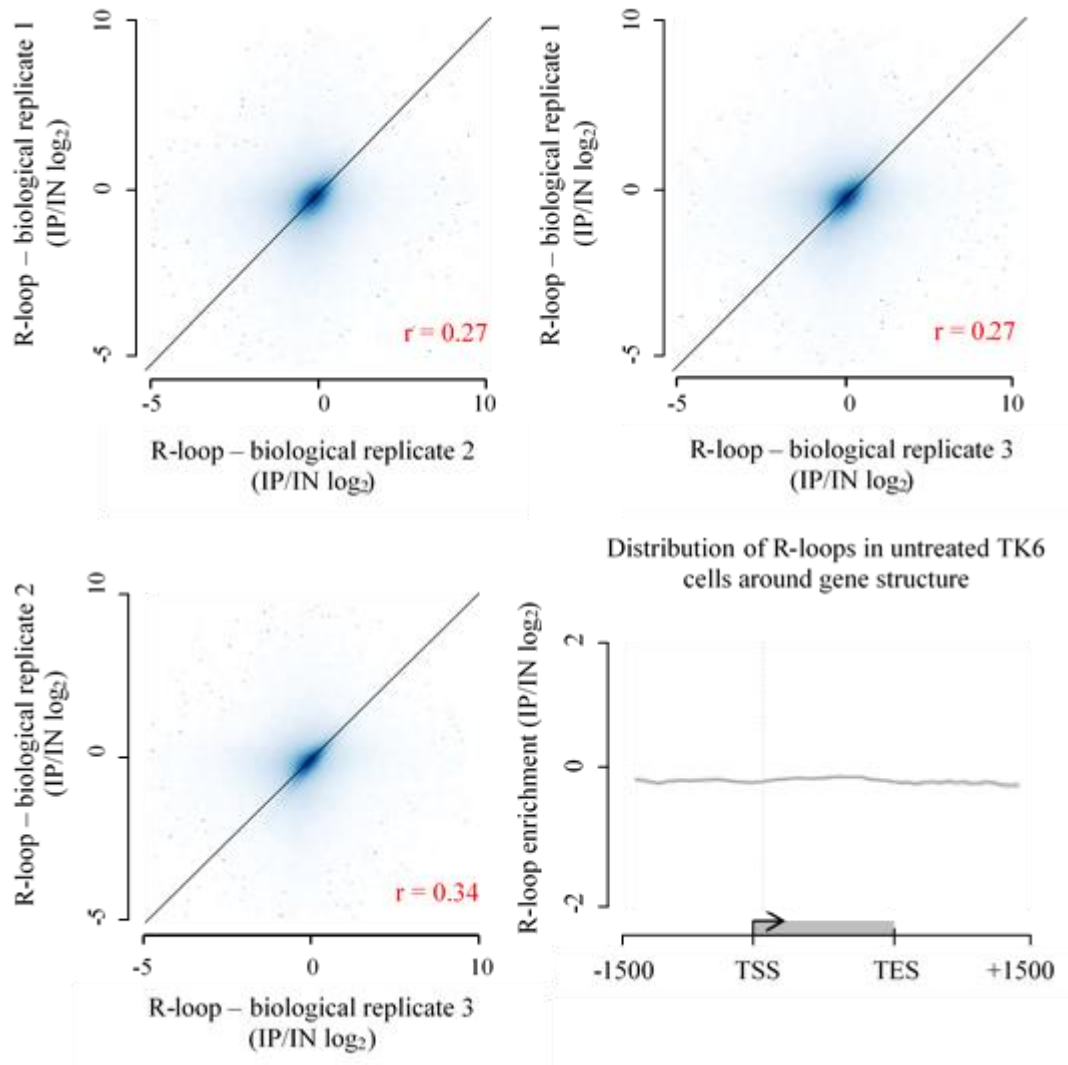


R-loop titration:



**Titration of the anti-DNA-RNA hybrid antibody from 0 – 4  $\mu$ g against a fixed volume of chromatin. Target enrichment was examined at the genes *SENP3-EIF4A1* and *ZNF544*.**

## Biological replicates of R-loop enrichment

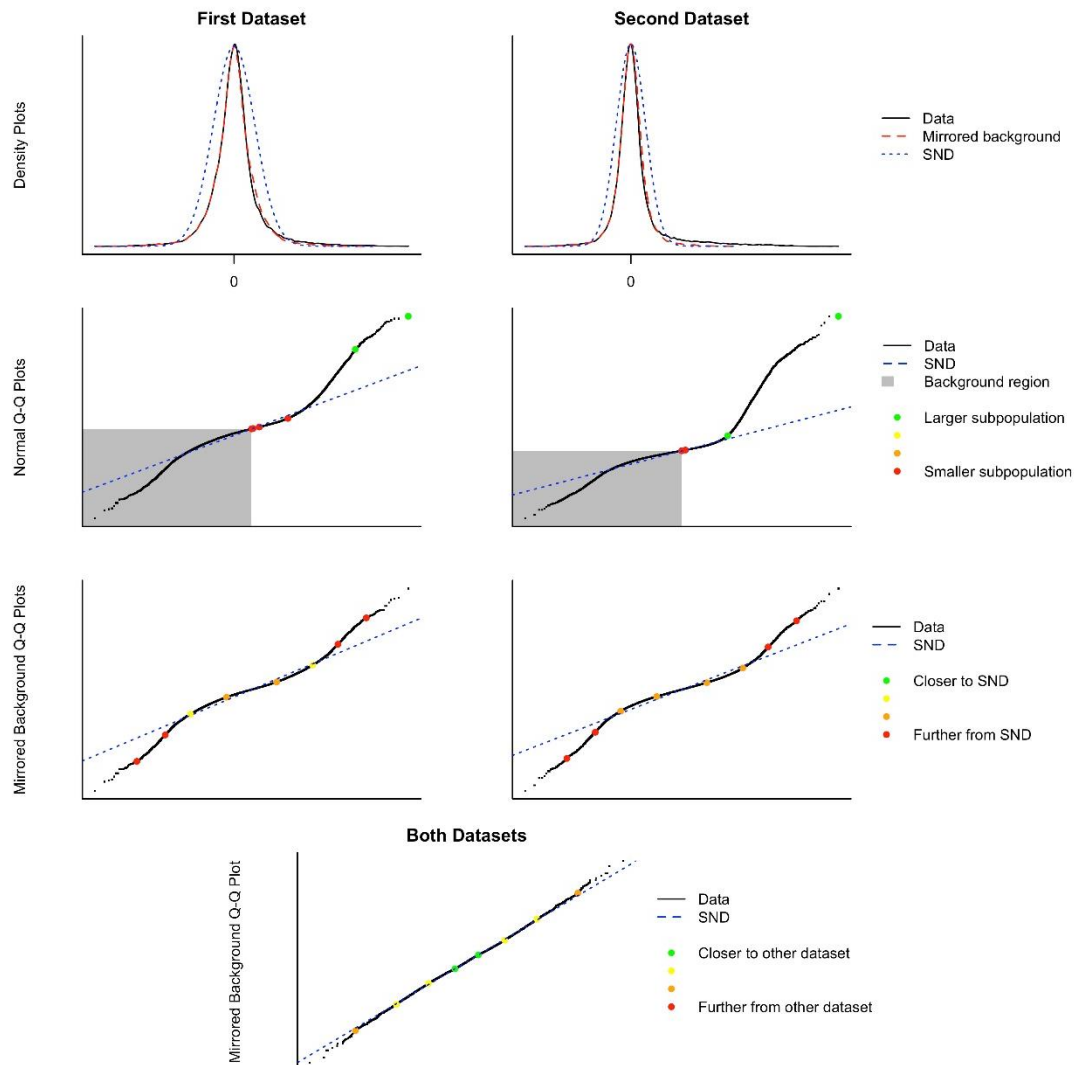


## Primers used

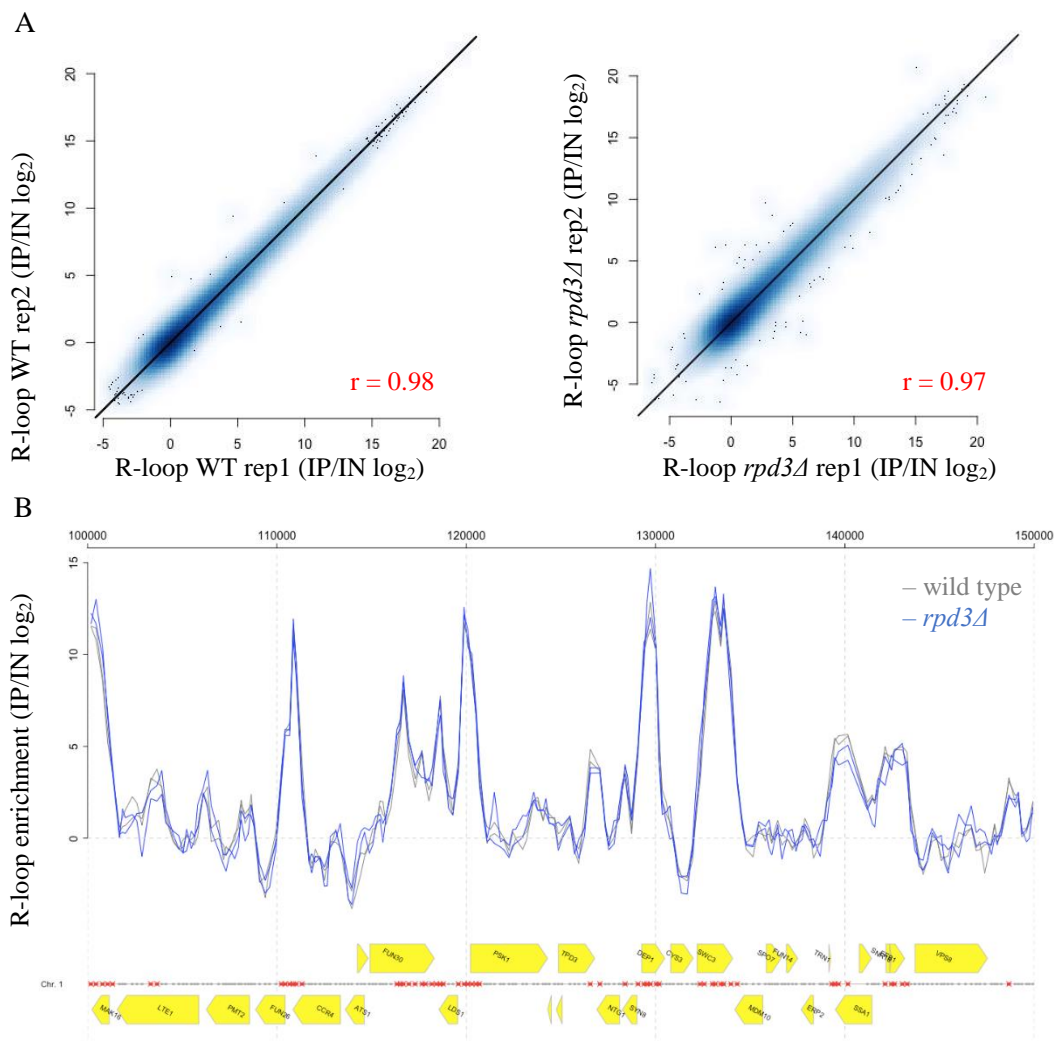
GAPDH-pP - F	TAC TAG CGG TTT TAC GGG CG
GAPDH-pP - R	TCG AAC AGG AGG AGC AGA GAG CGA
MYOD1 - F	ATG GAG CTA CTG TCG CCA C
MYOD1 - R	GAG TGC TCT TCG GGT TTC AG
GAD1 - F	GCT AAG AAC GGT GAG GAG CA
GAD1 - R	GAG AGC TCC AAG TTG AAG CC
RPL10-pP - F	ACC CGT CTT CGA CAG GAC T
RPL10-pP - R	GGA ACG GAA GAC GAG AAC AG

## Appendix IV – Raw data for Chapter VI

### Normalisation assumptions for $\gamma$ H2A data

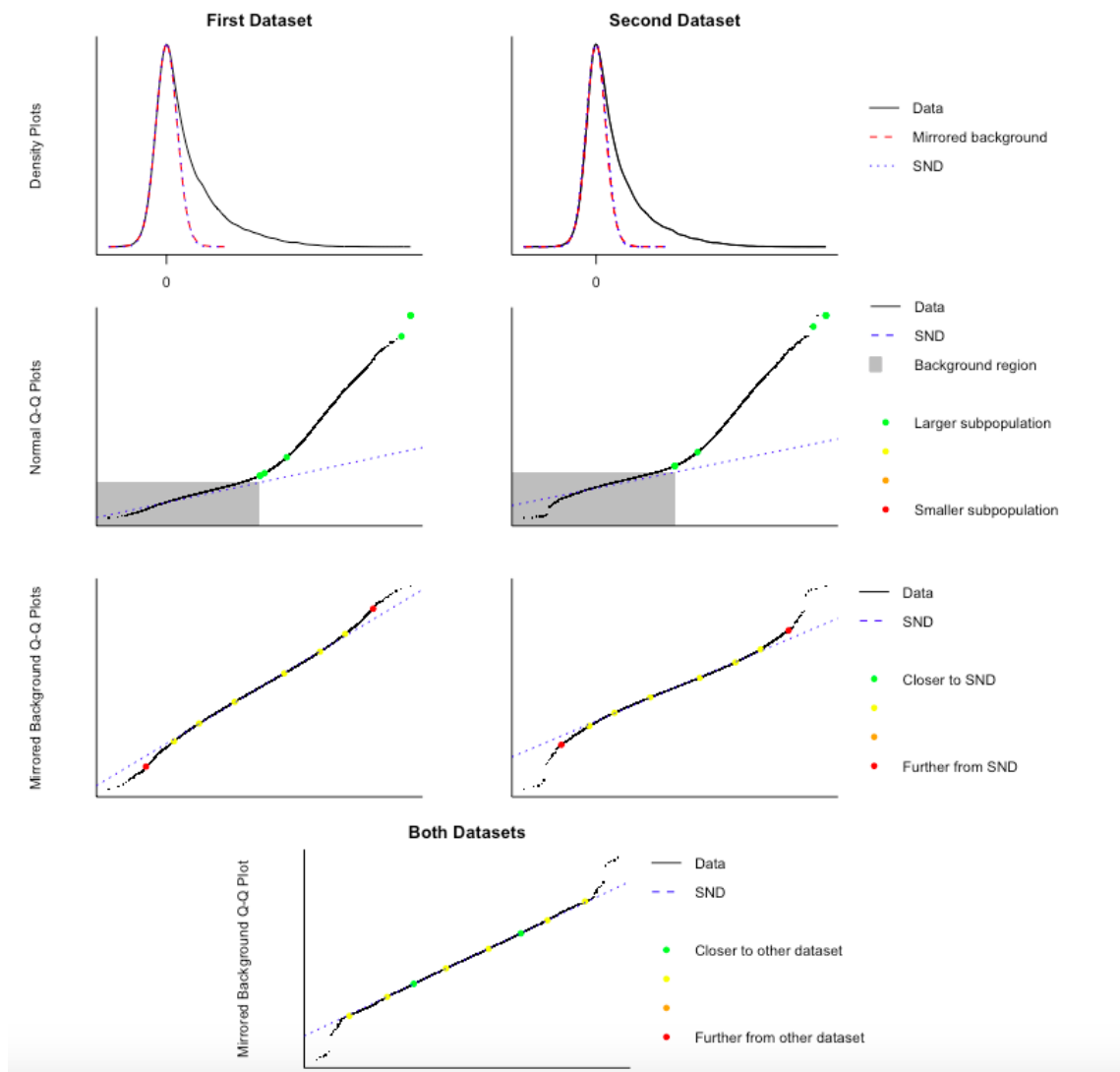


**$\gamma$ H2A microarray data meets the assumptions for normalisation.** First dataset – from wild type strain. Second dataset – from *rpd3 $\Delta$*  strain. *First row*: density plots of the mirrored background approximating to the standard normal distribution (SND). *Second row*: density plots shown in A represented as Q-Q plots. Data in the background (shaded) region should closely follow the SND, whereas deviations of the right tail represents the enriched sub-population. *Third row*: Q-Q plot of only the mirrored background data. *Fourth row*: Q-Q plot of the mirrored background for both datasets. Little/no departure from the SND indicates that the two datasets are suitable to be normalised together.



**Biological repeats of yeast R-loop microarray data show high reproducibility.** **A:** smoothed scatter plots showing the Pearson's correlation coefficient between two biological repeats for wild type (*left*) and *rpd3Δ* (*right*). **B:** normalised array data plotted along a section of chromosome 1 (100 – 150kb). Array probes are shown in grey along the annotation line with enriched regions in red. The positions of genes are represented as yellow boxes.

## Normalisation assumptions for R-loop data



**R-loop microarray data meets the assumptions for normalisation.** First dataset – from wild type strain. Second dataset – from *rpd3Δ* strain. *First row:* density plots of the mirrored background approximating to the standard normal distribution (SND). *Second row:* density plots shown in A represented as Q-Q plots. Data in the background (shaded) region should closely follow the SND, whereas deviations of the right tail represents the enriched sub-population. *Third row:* Q-Q plot of only the mirrored background data. *Fourth row:* Q-Q plot of the mirrored background for both datasets. Little/no departure from the SND indicates that the two datasets are suitable to be normalised together.

## References

- Adamo A1, Sesé B, Boue S, Castaño J, Paramonov I, Barrero MJ, Izpisua Belmonte JC. 2011. 'Adamo A1, Sesé B, Boue S, Castaño J, Paramonov I, Barrero MJ, Izpisua Belmonte JC.', *nat cell bio*, 6: 652-9.
- Adams H, Fritzsche FR, Dirnhofer S, Kristiansen G, Tzankov A. 2010. 'Class I histone deacetylases 1, 2 and 3 are highly expressed in classical Hodgkin's lymphoma.', *Expert Opin Ther Targets.*, 6: 577-84.
- Adeleye, Y., Andersen, M., Clewell, R., Davies, M. Denta, M. Edwards, S. Fowler, P. Malcomber, S. Nicol, B. Scott, A. Scott, S. Sun, B. Westmoreland, C. White, A. Zhang, Q. Carmichael, P. L. 2015. 'Implementing Toxicity Testing in the 21st Century (TT21C): Making safety decisions using toxicity pathways, and progress in a prototype risk assessment', *Toxicology*, 332: 102-11.
- Afgan, E., D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Cech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, B. Gruning, A. Guerler, J. Hillman-Jackson, G. Von Kuster, E. Rasche, N. Soranzo, N. Turaga, J. Taylor, A. Nekrutenko, and J. Goecks. 2016. 'The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update', *Nucleic Acids Research*, 44: W3-W10.
- Agalioti, T.Chen, G. Y.Thanos, D. 2002. 'Deciphering the transcriptional histone acetylation code for a human gene', *Cell*, 111: 381-92.
- Agilent. 'ChIP-on-chip Analysis', ver11.3.
- Agilent Application Note. 2005. 'Stringent RNA quality control using the Agilent 2100 bioanalyzer', *Publication number 5989-1086EN*.
- Aguilera, José M. Santos-Pereira & Andrés. 2015. 'R loops: new modulators of genome dynamics and function', *Nature Reviews Genetics*, 16: 583-97.
- Aguilera, Sonia Jimeno Andrés. 2010. 'The THO complex as a key mRNP biogenesis factor in development and cell differentiation', *j biol*, 1: 6.
- Ahuja N, Schwer B, Carobbio S, Waltregny D, North BJ, Castronovo V, Maechler P, Verdin E. 2007. 'Regulation of insulin secretion by SIRT4, a mitochondrial ADP-ribosyltransferase.', *J Biol Chem.*, 46: 33583-92.
- Aithal MG, Rajeswari N. 2015. 'Validation of Housekeeping Genes for Gene Expression Analysis in Glioblastoma Using Quantitative Real-Time Polymerase Chain Reaction', *Brain Tumor Res Treat*, 1: 24-9.
- Alonso, W. R., and D. A. Nelson. 1986. 'A NOVEL YEAST HISTONE DEACETYLASE - PARTIAL CHARACTERIZATION AND DEVELOPMENT OF AN ACTIVITY ASSAY', *Biochimica Et Biophysica Acta*, 866: 161-69.
- Ankley, G. T. Bennett, R. S. Erickson, R. J. Hoff, D. J. Hornung, M. W. Johnson, R. D. Mount, D. R. Nichols, J. W. Russom, C. L. Schmieder, P. K. Serrano, J. A. Tietge, J. E. Villeneuve, D. L. 2010. 'ADVERSE OUTCOME PATHWAYS: A CONCEPTUAL FRAMEWORK TO SUPPORT ECOTOXICOLOGY RESEARCH AND RISK ASSESSMENT', *Environmental Toxicology and Chemistry*, 29: 730-41.
- Ausio, J., Dong, F., Vanholde, K. E. 1989. 'USE OF SELECTIVELY TRYPSINIZED NUCLEOSOME CORE PARTICLES TO ANALYZE THE ROLE OF THE HISTONE TAILS IN THE STABILIZATION OF THE NUCLEOSOME', *Journal of Molecular Biology*, 206: 451-63.
- Ayrapetov, M. K.Gursoy-Yuzugullu, O.Xu, C.Xu, Y.Price, B. D. 2014. 'DNA double-strand breaks promote methylation of histone H3 on lysine 9 and transient formation of repressive chromatin', *Proceedings of the National Academy of Sciences of the United States of America*, 111: 9169-74.
- Bae, Hyun-Jung Kim and Suk-Chul. 2011. 'Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs', *Am J Transl Res*, 3: 166 - 79.
- Balakrishnan L, Bambara RA. 2013. 'Okazaki Fragment Metabolism', *cold Spring Harb Perspect bio*, 2.
- Balakrishnan, R., J. Park, K. Karra, B. C. Hitz, G. Binkley, E. L. Hong, J. Sullivan, G. Micklem, and J. M. Cherry. 2012. 'YeastMine-an integrated data warehouse for *Saccharomyces cerevisiae* data as a multipurpose tool-kit', *Database-the Journal of Biological Databases and Curation*.
- Banerji U, van Doorn L, Papadatos-Pastos D, Kristeleit R, Debnam P, Tall M, Stewart A, Raynaud F, Garrett MD, Toal M, Hooftman L, De Bono JS, Verweij J, Eskens FA. 2012. 'A phase I pharmacokinetic and pharmacodynamic study of CHR-3996, an oral class I selective histone deacetylase inhibitor in refractory solid tumors.', *Clin Cancer Res*, 9: 2687-94.



- Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, Kouzarides T. 2001. 'Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain', *Nature*, 1: 120-4.
- Bantscheff M, Hopf C, Savitski MM, Dittmann A, Grandi P, Michon AM, Schlegl J, Abraham Y, Becher I, Bergamini G, Boesche M, Delling M, Dümpelfeld B, Eberhard D, Huthmacher C, Mathieson T, PoECKel D, Reader V, Strunk K, Sweetman G, Kruse U, Neubauer G, Ramsden NG, Drewes G. 2011. 'Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes.', *nat biotechnol*, 3: 255-65.
- Barlow, Jacqueline H. 2013. 'Identification of Early Replicating Fragile Sites that Contribute to Genome Instability', *Cell*, 152: 620–32.
- Barneda-Zahonero B, Collazo O1, Azagra A1, Fernández-Duran I1, Serra-Musach J2, Islam AB3, Vega-García N4, Malatesta R4, Camós M4, Gómez A5, Román-González L1, Vidal A6, López-Bigas N7, Villanueva A8, Esteller M9, Parra M1. 2015. 'The transcriptional repressor HDAC7 promotes apoptosis and c-Myc downregulation in particular types of leukemia and lymphoma.', *Cell Death & Disease*, 6: e1635.
- Barrett, T., S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. G. Zhang, C. L. Robertson, N. Serova, S. Davis, and A. Soboleva. 2013. 'NCBI GEO: archive for functional genomics data sets-update', *Nucleic Acids Research*, 41: D991-D95.
- Bell, Stephen P. 2002. 'The origin recognition complex: from simple origins to complex functions', *Genes & Dev*, 16: 659-72.
- Bennett, M., K. E. Evans, S. R. Yu, Y. M. Teng, R. M. Webster, J. Powell, R. Waters, and S. H. Reed. 2015. 'Sandcastle: software for revealing latent information in multiple experimental ChIP-chip datasets via a novel normalisation procedure', *Scientific Reports*, 5.
- Bergink, S., F. A. Salomons, D. Hoogstraten, T. A. M. Groothuis, H. de Waard, J. Wu, L. Yuan, E. Citterio, A. B. Houtsmuller, J. Neefjes, I. H. J. Hoeijmakers, W. Vermeulen, and N. P. Dantuma. 2006. 'DNA damage triggers nucleotide excision repair-dependent monoubiquitylation of histone H2A', *Genes & Development*, 20: 1343-52.
- Bernstein, B. E., J. K. Tong, and S. L. Schreiber. 2000. 'Genomewide studies of histone deacetylase function in yeast', *Proceedings of the National Academy of Sciences of the United States of America*, 97: 13708-13.
- Besnard, E., A. Babied, L. Lapasset, O. Milhavet, H. Parrinello, C. Dantec, J. M. Marin, and J. M. Lemaitre. 2012. 'Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs', *Nature Structural & Molecular Biology*, 19: 837-44.
- Bhaskara, S., B. J. Chyla, J. M. Amann, S. K. Knutson, D. Cortez, Z. W. Sun, and S. W. Hiebert. 2008. 'Deletion of histone deacetylase 3 reveals critical roles in S phase progression and DNA damage control', *Molecular Cell*, 30: 61-72.
- Bhattacharya, S. Zhang, Q. Carmichael, P. L. Boekelheide, K. Andersen, M. E. 2011. 'Toxicity Testing in the 21(st) Century: Defining New Risk Assessment Approaches Based on Perturbation of Intracellular Toxicity Pathways', *Plos One*, 6.
- Bird, Adrian. 2007. 'Perceptions of epigenetics', *Nature*, 447: 396-98.
- Bose, P., Y. Dai, and S. Grant. 2014. 'Histone deacetylase inhibitor (HDACI) mechanisms of action: Emerging insights', *Pharmacology & Therapeutics*, 143: 323-36.
- Botstein, D., and G. R. Fink. 1988. 'YEAST - AN EXPERIMENTAL ORGANISM FOR MODERN BIOLOGY', *Science*, 240: 1439-43.
- Brambati A, Colosio A, Zardoni L, Galanti L, Liberi G. 2015. 'Replication and transcription on a collision course: eukaryotic regulation mechanisms and implications for DNA stability.', *Frontiers in genetics*, 6.
- Branham, W. S., C. D. Melvin, T. Han, V. G. Desai, C. L. Moland, A. T. Scully, and J. C. Fuscoe. 2007. 'Elimination of laboratory ozone leads to a dramatic improvement in the reproducibility of microarray gene expression measurements', *Bmc Biotechnology*, 7.
- Brehove, M., Wang, T., North, J., Luo, Y., Dreher, S. J., Shimko, J. C., Ottesen, J. J., Luger, K., Poirier, M. G. 2015. 'Histone Core Phosphorylation Regulates DNA Accessibility', *Journal of Biological Chemistry*, 290: 22612-21.
- Brès V, Yoh SM, Jones KA. 2008. 'The multi-tasking P-TEFb complex', *Curr Opin Cell Biol.*, 20: 334-40.
- Broide, R. S., J. M. Redwine, N. Aftahi, W. Young, F. E. Bloom, and C. J. Winrow. 2007. 'Distribution of histone deacetylases 1-11 in the rat brain', *Journal of Molecular Neuroscience*, 31: 47-58.
- Burnett C, Valentini S, Cabreiro F, Goss M, Somogyvári M, Piper MD, Hoddinott M, Sutphin GL, Leko V, McElwee JJ, Vazquez-Manrique RP, Orfila AM, Ackerman D, Au C, Vinti G, Riesen M,

- Howard K, Neri C, Bedalov A, Kaeberlein M, Soti C, Partridge L, Gems D. 2011. 'Absence of effects of Sir2 overexpression on lifespan in *C. elegans* and *Drosophila*.', *Nature*, 477: 482-5.
- Campbell, S.Ismail, I. H.Young, L. C.Poirier, G. G.Hendzel, M. J. 2013. 'Polycomb repressive complex 2 contributes to DNA double-strand break repair', *Cell Cycle*, 12: 2675-83.
- Carmen, A. A., S. E. Rundlett, and M. Grunstein. 1996. 'HDA1 and HDA3 are components of a yeast histone deacetylase (HDA) complex', *Journal of Biological Chemistry*, 271: 15837-44.
- Caron, P.Aymard, F.Iacovoni, J. S.Briois, S.Canitrot, Y.Bugler, B.Massip, L.Losada, A.Legube, G. 2012. 'Cohesin Protects Genes against gamma H2AX Induced by DNA Double-Strand Breaks', *Plos Genetics*, 8.
- Castellano-Pozo, M., J. M. Santos-Pereira, A. G. Rondon, S. Barroso, E. Andujar, M. Perez-Alegre, T. Garcia-Muse, and A. Aguilera. 2013. 'R Loops Are Linked to Histone H3 S10 Phosphorylation and Chromatin Condensation', *Molecular Cell*, 52: 583-90.
- Celeste, A.Fernandez-Capetillo, O.Kruhlak, M. J.Pilch, D. R.Staudt, D. W.Lee, A.Bonner, R. F.Bonner, W. M.Nussenzweig, A. 2003. 'Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks', *Nature Cell Biology*, 5: 675-U51.
- Chakalova L, Emmanuel Debrand1, Jennifer A. Mitchell1, Cameron S. Osborne1 & Peter Fraser. 2005. 'Replication and transcription: Shaping the landscape of the genome', *Nature Reviews Genetics*, 6: 669-77.
- Chan, Y. J. A., M. J. Aristizabal, P. Y. T. Lu, Z. L. Luo, A. Hamza, M. S. Kobor, P. C. Stirling, and P. Hieter. 2014. 'Genome-Wide Profiling of Yeast DNA: RNA Hybrid Prone Sites with DRIP-Chip', *Plos Genetics*, 10.
- Chen, C. C.Carson, J. J.Feser, J.Tamburini, B.Zabaronick, S.Linger, J.Tyler, J. K. 2008. 'Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair', *Cell*, 134: 231-43.
- Chen, Wei-Ta, Amir Alpert, Courtney Leiter, Fade Gong, Stephen P. Jackson, and Kyle M. Miller. 2013. 'Systematic Identification of Functional Residues in Mammalian Histone H2AX', *Molecular and Cellular Biology*, 33: 111-26.
- Cherry, J. M., E. L. Hong, C. Amundsen, R. Balakrishnan, G. Binkley, E. T. Chan, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. R. Engel, D. G. Fisk, J. E. Hirschman, B. C. Hitz, K. Karra, C. J. Krieger, S. R. Miyasato, R. S. Nash, J. Park, M. S. Skrzypek, M. Simison, S. Weng, and E. D. Wong. 2012. 'Saccharomyces Genome Database: the genomics resource of budding yeast', *Nucleic Acids Research*, 40: D700-D05.
- Chiang, CM. 2009. 'Brd4 engagement from chromatin targeting to transcriptional regulation: selective contact with acetylated histone H3 and H4.', *F1000 Biol Rep*, 1.
- Choi, E., C. Han, I. Park, B. Lee, S. Jin, H. Choi, D. H. Kim, Z. Y. Park, E. M. Eddy, and C. Cho. 2008. 'A Novel Germ Cell-specific Protein, SHP1, Forms a Complex with Chromatin Remodeling Activity during Spermatogenesis', *Journal of Biological Chemistry*, 283: 35283-94.
- Choi JH, Kwon HJ, Yoon BI, Kim JH, Han SU, Joo HJ, Kim DY. 2001. 'Expression profile of histone deacetylase 1 in gastric cancer tissue.', *Jpn J Cancer Res*: 1300-4.
- Choi SY, Ryu Y1, Kee HJ2, Cho SN1, Kim GR1, Cho JY1, Kim HS3, Kim IK4, Jeong MH5. 2015. 'Tubastatin A suppresses renal fibrosis via regulation of epigenetic histone modification and Smad3-dependent fibrotic genes.', *Vascul Pharmacol*, 72: 130-40.
- Cipriani A, Reid K, Young AH, Macritchie K, Geddes J. 2013. 'Valproic acid, valproate and divalproex in the maintenance treatment of bipolar disorder', *Cochrane Database Syst Rev.*, 10.
- Cleophas MC, Crişan TO1, Lemmers H1, Toenhake-Dijkstra H1, Fossati G2, Jansen TL3, Dinarello CA4, Netea MG1, Joosten LA1. 2016. 'Suppression of monosodium urate crystal-induced cytokine production by butyrate is mediated by the inhibition of class I histone deacetylases.', *Ann Rheum Dis.*, 3: 593-600.
- Clyne, R. K., and T. J. Kelly. 1995. 'Genetic analysis of an ARS element from the fission yeast *Schizosaccharomyces pombe*', *Embo Journal*, 14: 6348-57.
- Collin J.E., Patricia C. Ellis, Angela T. White, Antonia E.G. Booth, Claire E. Moore, Mark Burman, Robert W. Rees, and Anthony M. Lynch. 'Evaluation of the Litron In Vitro MicroFlow® Kit for the flow cytometric enumeration of micronuclei (MN) in mammalian cells', *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 654: 76-81.
- Cong, P. Q., K. J. Zhu, Q. Q. Ji, H. J. Zhao, and Y. S. Chen. 2013. 'Effects of trichostatin A on histone acetylation and methylation characteristics in early porcine embryos after somatic cell nuclear transfer', *Animal Science Journal*, 84: 639-49.
- Conti, C., E. Leo, G. S. Eichler, O. Sordet, M. M. Martin, A. Fan, M. I. Aladjem, and Y. Pommier. 2010. 'Inhibition of Histone Deacetylase in Cancer Cells Slows Down Replication Forks, Activates Dormant Origins, and Induces DNA Damage', *Cancer Research*, 70: 4470-80.

- Counce, S. J. 1958. 'The Strategy of the Genes', *Yale J Biol Med*, 30: 470-71.
- Crosetto, N., A. Mitra, M. J. Silva, M. Bienko, N. Dojer, Q. Wang, E. Karaca, R. Chiarle, M. Skrzypczak, K. Ginalski, P. Pasero, M. Rowicka, and I. Dikic. 2013. 'Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing', *Nature Methods*, 10: 361-+.
- Cusanelli E., and Chartrand P. 2015. 'Telomeric repeat-containing RNA TERRA: a noncoding RNA connecting telomere biology to genome integrity', *Front Genet*, 6.
- Daniel Herranz, Maribel Muñoz-Martin, Marta Cañamero, Francisca Mulero, Barbara Martinez-Pastor, Oscar Fernandez-Capetillo & Manuel Serrano. 2010. 'Sirt1 improves healthy ageing and protects from metabolic syndrome-associated cancer.', *nature comm.*
- Dao P, Wojtowicz D1, Nelson S2, Levens D3, Przytycka TM1. 2016. 'Ups and Downs of Poised RNA Polymerase II in B-Cells', *PLoS Comput Biol.*, 4.
- Deardorff, M. A., M. Bando, R. Nakato, E. Watrin, T. Itoh, M. Minamino, K. Saitoh, M. Komata, Y. Katou, D. Clark, K. E. Cole, E. De Baere, C. Decroos, N. Di Donato, S. Ernst, L. J. Francey, Y. Gyftodimou, K. Hirashima, M. Hullings, Y. Ishikawa, C. Jaulin, M. Kaur, T. Kiyono, P. M. Lombardi, L. Magnaghi-Jaulin, G. R. Mortier, N. Nozaki, M. B. Petersen, H. Seimiya, V. M. Siu, Y. Suzuki, K. Takagaki, J. J. Wilde, P. J. Willems, C. Prigent, G. Gillesen-Kaesbach, D. W. Christianson, F. J. Kaiser, L. G. Jackson, T. Hirota, I. D. Krantz, and K. Shirahige. 2012. 'HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle', *Nature*, 489.
- Dellino, G. I., D. Cittaro, R. Piccioni, L. Luzi, S. Banfi, S. Segalla, M. Cesaroni, R. Mendoza-Maldonado, M. Giacca, and P. G. Pelicci. 2013. 'Genome-wide mapping of human DNA-replication origins: Levels of transcription at ORC1 sites regulate origin selection and replication timing', *Genome Research*, 23: 1-11.
- DePamphilis ML, Blow JJ, Ghosh S, Saha T, Noguchi K, Vassilev A. 2006. 'Regulating the licensing of DNA replication origins in metazoan', *Curr Opin Cell Biol.*, 3: 231-9.
- Deshpande AM, Newlon CS. 1996. 'DNA replication fork pause sites dependent on transcription.', *Science*, 5264: 1030-3.
- Diagenode. 'The Ultimate Guide for Chromatin Shearing Optimization with Bioruptor® Standard and Plus'.
- Dimitrova, D. S., and D. M. Gilbert. 1999. 'The spatial position and replication timing of chromosomal domains are both established in early G1 phase', *Molecular Cell*, 4: 983-93.
- Dion, M. F. Altschuler, S. J. Wu, L. F. Rando, O. J. 2005. 'Genomic characterization reveals a simple histone H4 acetylation code', *Proceedings of the National Academy of Sciences of the United States of America*, 102: 5501-06.
- Dong M, Ning ZQ, Xing PY, Xu JL, Cao HX, Dou GF, Meng ZY, Shi YK, Lu XP, Feng FY. 2012. 'Phase I study of chidamide (CS055/HBI-8000), a new histone deacetylase inhibitor, in patients with advanced solid tumors and lymphomas.', *Cancer Chemother Pharmacol.*, 6: Cancer Chemother Pharmacol.
- Dovey OM, Foster CT, Conte N, Edwards SA, Edwards JM, Singh R, Vassiliou G, Bradley A, Cowley SM. 2013. 'Histone deacetylase (HDAC) 1 and 2 are essential for normal T cell development and genomic stability in mice', *Blood*, 121: 1335-44.
- Drolet, M. 2006. 'Growth inhibition mediated by excess negative supercoiling: The interplay between transcription elongation, R-loop formation and DNA topology', *mol microbiol*, 3: 723-30.
- Du J, Zhou Y, Su X, Yu JJ, Khan S, Jiang H, Kim J, Woo J, Kim JH, Choi BH, He B, Chen W, Zhang S, Cerione RA, Auwerx J, Hao Q, Lin H. 2011. 'Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase.', *Science*, 334: 806-9.
- Duan, M. R., and M. J. Smerdon. 2014. 'Histone H3 Lysine 14 (H3K14) Acetylation Facilitates DNA Repair in a Positioned Nucleosome by Stabilizing the Binding of the Chromatin Remodeler RSC (Remodels Structure of Chromatin)', *Journal of Biological Chemistry*, 289: 8353-63.
- Duncker BP, Chesnokov IN, McConkey BJ. 2009. 'The origin recognition complex protein family.', *Genome Biology*, 3: 214.
- Dunham, I., et al. Encode Project Consortium. 2012. 'An integrated encyclopedia of DNA elements in the human genome', *Nature*, 489: 57-74.
- Dutta, D., K. Shatalin, V. Epshtein, M. E. Gottesman, and E. Nudler. 2011. 'Linking RNA Polymerase Backtracking to Genome Instability in E. coli', *Cell*, 146: 533-43.
- Ehrenhofer-Murray, AE. 2004. 'Chromatin dynamics at DNA replication, transcription and repair.', *Eur J Biochem*, 271: 2335-49.
- Ehrentraut, S., J. M. Weber, J. N. Dybowski, D. Hoffmann, and A. E. Ehrenhofer-Murray. 2010. 'Rpd3-dependent boundary formation at telomeres by removal of Sir2 substrate', *Proceedings of the National Academy of Sciences of the United States of America*, 107: 5522-27.

- Ekwall, K., T. Olsson, B. M. Turner, G. Cranston, and R. C. Allshire. 1997. 'Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres', *Cell*, 91: 1021-32.
- Fachinetti D, Bermejo R, Cocito A, Minardi S, Katou Y, Kanoh Y, Shirahige K, Azvolinsky A, Zakian VA, Foiani M. 2010. 'Replication Termination at Eukaryotic Chromosomes Is Mediated by Top2 and Occurs at Genomic Loci Containing Pausing Elements', *mol cell*, 4: 595-605.
- Finnin MS, Donigian JR, Cohen A, Richon VM, Rifkind RA, Marks PA, Breslow R, Pavletich NP. 1999. 'Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors.', *Nature*, 401: 188-93.
- Fowler, P., K. Smith, J. Young, L. Jeffrey, D. Kirkland, S. Pfuhler, and P. Carmichael. 2012. 'Reduction of misleading ("false") positive results in mammalian cell genotoxicity assays. I. Choice of cell type', *Mutation Research-Genetic Toxicology and Environmental Mutagenesis*, 742: 11-25.
- Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Pérez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, Esteller M. 2005. 'Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer.', *nat genet*, 4: 391-400.
- Fujita, N. Watanabe, S. Ichimura, T. Tsuruzoe, S. Shinkai, Y. Tachibana, M. Chiba, T. Nakao, M. 2003. 'Methyl-CpG binding domain 1 (MBD1) interacts with the Suv39h1-HP1 heterochromatic complex for DNA methylation-based transcriptional repression', *Journal of Biological Chemistry*, 278: 24132-38.
- Furumai R, Matsuyama A, Kobashi N, Lee KH, Nishiyama M, Nakajima H, Tanaka A, Komatsu Y, Nishino N, Yoshida M, Horinouchi S. 2002. 'FK228 (depsipeptide) as a natural prodrug that inhibits class I histone deacetylases.', *Cancer Research*, 17: 4916-21.
- Gan W, Guan Z, Liu J, Gui T, Shen K, Manley JL, Li X. 2011. 'R-loop-mediated genomic instability is caused by impairment of replication fork progression', *Genes & Dev*, 19: 2041-56.
- Gao L, Alumkal J. 2010. 'Epigenetic regulation of androgen receptor signaling in prostate cancer.', *Epigenetics*, 2: 100-4.
- Gao, L., M. A. Cueto, F. Asselbergs, and P. Atadja. 2002. 'Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family', *Journal of Biological Chemistry*, 277: 25748-55.
- Garcia-Manero G, Assouline S, Cortes J, Estrov Z, Kantarjian H, Yang H, Newsome WM, Miller WH Jr, Rousseau C, Kalita A, Bonfils C, Dubay M, Patterson TA, Li Z, Besterman JM, Reid G, Laille E, Martell RE, Minden M. 2008. 'Phase 1 study of the oral isotype specific histone deacetylase inhibitor mgcd0103 in leukemia.', *Blood*, 4: 981-9.
- Garcia-Manero G, Tambaro FP, Bekele NB, Yang H, Ravandi F, Jabbour E, Borthakur G, Kadia TM, Konopleva MY, Faderl S, Cortes JE, Brandt M, Hu Y, McCue D, Newsome WM, Pierce SR, de Lima M, Kantarjian HM. 2012. 'Phase II Trial of Vorinostat With Idarubicin and Cytarabine for Patients With Newly Diagnosed Acute Myelogenous Leukemia or Myelodysplastic Syndrome.', *J Clin Oncol.*, 18: 2204-10.
- Garske AL, Oliver SS, Wagner EK, Musselman CA, LeRoy G, Garcia BA, Kutateladze TG, Denu JM. 2010. 'Combinatorial profiling of chromatin binding modules', *nat chem bio*, 6: 283-90.
- Gavaldá S, Gallardo M, Luna R, Aguilera A. 2013. 'R-loop mediated transcription-associated recombination in trf4Δ mutants reveals new links between RNA surveillance and genome integrity', *Plos One*, 6: e65541.
- Gaymes, T. J., R. A. Padua, M. Pla, S. Orr, N. Omidvar, C. Chomienne, G. J. Mufti, and F. V. Rassool. 2006. 'Histone deacetylase inhibitors (HDI) cause DNA damage in leukemia cells: A mechanism for leukemia-specific HDI-dependent apoptosis?', *Molecular Cancer Research*, 4: 563-73.
- Gerhardt, J., S. Jafar, M. P. Spindler, E. Ott, and A. Schepers. 2007. 'Identification of new human origins of DNA replication by an origin-trapping assay (vol 26, pg 7731, 2006)', *Molecular and Cellular Biology*, 27: 788-88.
- Geyer, Dirk Eick and Matthias. 2013. 'The RNA polymerase II carboxy-terminal domain (CTD) code', *Chem Rev*, 11: 8456-90.
- Giannattasio, M., F. Lazzaro, P. Plevani, and M. M. Falconi. 2005. 'The DNA damage checkpoint response requires histone H2B ubiquitination by Rad6-Bre1 and H3 methylation by Dot1', *Journal of Biological Chemistry*, 280: 9879-86.
- Ginno, P. A., P. L. Lott, H. C. Christensen, I. Korf, and F. Chedin. 2012. 'R-Loop Formation Is a Distinctive Characteristic of Unmethylated Human CpG Island Promoters', *Molecular Cell*, 45: 814-25.

- Ginno PA, Lim YW, Lott PL, Korf I, Chédin F. 2013. 'GC skew at the 5' and 3' ends of human genes links R-loop formation to epigenetic regulation and transcription termination', *Genome Research*, 10: 1590-600.
- Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, Louis EJ, Mewes HW, Murakami Y, Philippsen P, Tettelin H, Oliver SG. 1996. 'Life with 6000 genes.', *Science*, 5287: 563-7.
- Gomez V.P., and Vig B.K. 2002. 'Differential induction of numerical chromosome changes by sodium butyrate in two transformed cell lines', *Cancer Genet Cytogenet.*, 1: 56-63.
- Gottesfeld, J., and R. Mukhopadhyay. 2012. 'Vincent Allfrey's Work on Histone Acetylation', *Journal of Biological Chemistry*, 287: 2270-71.
- Gottipati, P., T. N. Cassel, L. Savolainen, and T. Helleday. 2008. 'Transcription-associated recombination is dependent on replication in mammalian cells', *Molecular and Cellular Biology*, 28: 154-64.
- Gregoretta IV1, Lee YM, Goodson HV. 2004. 'Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis.', *J mol biol*, 1: 17-31.
- Groh M., and Gromak N. 2014. 'Out of Balance: R-loops in Human Disease', *Plos Genetics*, 10.
- Groh M., Michele M. P., Martins R.W, and Gromak N. 2014. 'R-loops Associated with Triplet Repeat Expansions Promote Gene Silencing in Friedreich Ataxia and Fragile X Syndrome', *Plos Genetics*.
- Grunstein, Siavash K. Kurdistani and Michael. 2003. 'Histone acetylation and deacetylation in yeast', *Nature Reviews Molecular Cell Biology*, 4: 276-84.
- Gunther, K., M. Rust, J. Leers, T. Boettger, M. Scharfe, M. Jarek, M. Bartkuhn, and R. Renkawitz. 2013. 'Differential roles for MBD2 and MBD3 at methylated CpG islands, active promoters and binding to exon sequences', *Nucleic Acids Research*, 41: 3010-21.
- Guzmán E, Lis JT. 1999. 'Transcription factor TFIIF is required for promoter melting in vivo.', *mol cell biol*, 8: 5652-8.
- Hage A, French SL, Beyer AL, Tollervey D. 2010. 'Loss of Topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis', *Genes & Dev*, 14: 1546-58.
- Hagood, James S. 2014. 'Beyond the Genome: Epigenetic Mechanisms in Lung Remodeling', *Physiology*, 29: 177-85.
- Haig, D. 2004. 'The (dual) origin of epigenetics', *Cold Spring Harbor Symposia on Quantitative Biology*, 69: 67-70.
- Harju, S., H. Fedosyuk, and K. R. Peterson. 2004. 'Rapid isolation of yeast genomic DNA: Bust n' Grab', *Bmc Biotechnology*, 4.
- Hashioka, Sadayuki; Klegeris, Andis; McGeer, Patrick L. 2012. 'Histone deacetylase modulates the proinflammatory and -fibrotic changes in tubulointerstitial injury.', *JOURNAL OF NEUROINFLAMMATION*, 9: 113.
- Hassig, C. A., J. K. Tong, T. C. Fleischer, T. Owa, P. G. Grable, D. E. Ayer, and S. L. Schreiber. 1998. 'A role for histone deacetylase activity in HDAC1-mediated transcriptional repression', *Proceedings of the National Academy of Sciences of the United States of America*, 95: 3519-24.
- Healy, S., D. H. Khan, and J. R. Davie. 2011. 'Gene Expression Regulation Through 14-3-3 Interactions with Histones and HDACs', *Discovery Medicine*, 59: 349-58.
- Hecht, A. Laroche, T. Strahlbolsinger, S. Gasser, S. M. Grunstein, M. . 1995. 'HISTONE H3 AND H4 N-TERMINI INTERACT WITH SIR3 AND SIR4 PROTEINS - A MOLECULAR-MODEL FOR THE FORMATION OF HETEROCHROMATIN IN YEAST', *Cell*, 80: 583-92.
- Helmrich, A., M. Ballarino, and L. Tora. 2011. 'Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes', *Molecular Cell*, 44: 966-77.
- Herring, S. W. 1993. 'FORMATION OF THE VERTEBRATE FACE - EPIGENETIC AND FUNCTIONAL INFLUENCES', *American Zoologist*, 33: 472-83.
- Hertsch R.H., Dario Beraldi, Stefanie V Lensing, Giovanni Marsico, Katherine Zyner, Aled Parry, Marco Di Antonio, Jeremy Pike, Hiroshi Kimura, Masashi Narita, David Tannahill & Shankar Balasubramanian. 2016. 'G-quadruplex structures mark human regulatory chromatin', *nature letter*, 48: 1267-72.
- Hoshina, S., K. Yura, H. Teranishi, N. Kiyasu, A. Tominaga, H. Kadoma, A. Nakatsuka, T. Kunichika, C. Obuse, and S. Waga. 2013. 'Human Origin Recognition Complex Binds Preferentially to G-quadruplex-preferable RNA and Single-stranded DNA', *Journal of Biological Chemistry*, 288: 30161-71.
- Hrebackova, Jana, Jitka Poljakova, Tomas Eckschlagler, Jan Hrabeta, Pavel Prochazka, Svatopluk Smutny, and Marie Stiborova. 2009. 'Histone deacetylase inhibitors valproate and trichostatin A are toxic to neuroblastoma cells and modulate cytochrome P450 1A1, 1B1 and 3A4 expression in these cells', *Interdisciplinary toxicology*, 2: 205-10.

- Hsin JP, Manley JL. 2012. 'The RNA polymerase II CTD coordinates transcription and RNA processing', *Genes & Dev*, 19: 2119-37.
- Huang, H. Sabari, B. R. Garcia, B. A. Allis, C. D. Zhao, Y. M. 2014. 'SnapShot: Histone Modifications', *Cell*, 159: 458-+.
- Huberman J.A., Riggs A.D. 1968. 'On the mechanism of DNA replication in mammalian chromosomes', *Journal of Molecular Biology*, 32: 327.
- Huertas P, Aguilera A. 2003. 'Cotranscriptionally formed DNA:RNA hybrids mediate transcription elongation impairment and transcription-associated recombination', *mol cell*, 3: 711-21.
- Hui, L., Y. Zheng, Y. Yan, J. Bargonetti, and D. A. Foster. 2006. 'Mutant p53 in MDA-MB-231 breast cancer cells is stabilized by elevated phospholipase D activity and contributes to survival signals generated by phospholipase D', *Oncogene*, 25: 7305-10.
- Iacovoni, J. S. Caron, P. Lassadi, I. Nicolas, E. Massip, L. Trouche, D. Legube, G. 2010. 'High-resolution profiling of gamma H2AX around DNA double strand breaks in the mammalian genome', *Embo Journal*, 29: 1446-57.
- Ikura, Tsuyoshi, Satoshi Tashiro, Akemi Kakino, Hiroki Shima, Naduparambil Jacob, Ravindra Amunugarna, Kristine Yoder, Shunsuke Izumi, Isao Kuraoka, Kiyoji Tanaka, Hiroshi Kimura, Masae Tkura, Shuichi Nishikubo, Takashi Ito, Akihiko Muto, Kiyoshi Miyagawa, Shunichi Takeda, Richard Fishel, Kazuhiko Igarashi, and Kenji Kamiya. 2007. 'DNA damage-dependent acetylation and ubiquitination of H2AX enhances chromatin dynamics', *Molecular and Cellular Biology*, 27: 7028-40.
- Im JS, Keaton M, Lee KY, Kumar P, Park J, Dutta A. 2014. 'ATR checkpoint kinase and CRL1 TRCP collaborate to degrade ASF1a and thus repress genes overlapping with clusters of stalled replication forks', *Genes & Dev*, 8: 875-87.
- Ivanova D, Toni Taylor<sup>1</sup>, †, Sarah L. Smith<sup>1</sup>, Juachi U. Dimude<sup>1</sup>, Amy L. Upton<sup>2</sup>, Mana M. Mehrjouy<sup>3</sup>, Ole Skovgaard<sup>4</sup>, David J. Sherratt<sup>2</sup>, Renata Retkute<sup>5</sup> and Christian J. Rudolph<sup>1,\*</sup>. 2015. 'Shaping the landscape of the Escherichia coli chromosome: replication-transcription encounters in cells with an ectopic replication origin', *NAR*, 16: 7865 - 77.
- Jackson, V. 1987. 'Deposition of newly synthesized histones: new histones H2A and H2B do not deposit in the same nucleosome with new histones H3 and H4.', *Biochemistry*, 26: 2315-25.
- Johnson, D., and R. Walmsley. 2013. 'Histone-deacetylase inhibitors produce positive results in the GADD45a-GFP GreenScreen HC assay', *Mutation Research-Genetic Toxicology and Environmental Mutagenesis*, 751: 96-100.
- Johnson, W. Evan. 2006. 'Model-based analysis of tiling-arrays for ChIP-chip', *pnas*, 103: 12457-62.
- Juan, L. J., W. J. Shia, M. H. Chen, W. M. Yang, E. Seto, Y. S. Lin, and C. W. Wu. 2000. 'Histone deacetylases specifically down-regulate p53-dependent gene activation', *Journal of Biological Chemistry*, 275: 20436-43.
- Jungmin S, Sang Cheol Kim, Heun-Sik Lee, Jung Kyu Kim, Hye Jin Shon, Nur Lina Mohd Salleh, Kartiki Vasant Desai, Jae Ho Lee, Eun-Suk Kang, Jin Sung Kim, and Jung Kyoong Choi. 2012. 'Genome-wide profiles of H2AX and  $\gamma$ -H2AX differentiate endogenous and exogenous DNA damage hotspots in human cells', *Nucleic Acids Research*.
- Kaeberlein M1, McVey M, Guarente L. 1999. 'The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms.', *Genes & Dev*, 19: 2570-80.
- Kahli P.N., d'Aubenton-Carafa Y2, Jaszczyszyn Y2, Shen Y2, Silvain M2, Thermes C2, Chen CL2, Hyrien O1. 2016. 'Replication landscape of the human genome', *nat commun*, 11: 10208.
- Kaiser, F. J., et al. Consortium CareRare Canada, and Geno Univ Washington Ctr Mendelian. 2014. 'Loss-of-function HDAC8 mutations cause a phenotypic spectrum of Cornelia de Lange syndrome-like features, ocular hypertelorism, large fontanelle and X-linked inheritance', *Human Molecular Genetics*, 23: 2888-900.
- Kanai, Yae. 2008. 'Overexpression of HDACs: a prognostic marker for gastric cancer identified by tissue microarray.', *lancet*, 9: 91-93.
- Kapanidis AN, Margeat E, Ho SO, Kortkhonjia E, Weiss S, Ebright RH. 2006. 'Initial Transcription by RNA Polymerase Proceeds Through a DNA-Scrunching Mechanism', *Science*, 5802: 1144-7.
- Karolin L, Armin W. Mäder, Robin K. Richmond, David F. Sargent & Timothy J. Richmond. 1997. 'Crystal structure of the nucleosome core particle at 2.8 Å resolution', *Nature*, 389: 251-60.
- Kawakami, H., E. Ohashi, S. Kanamoto, T. Tsurimoto, and T. Katayama. 2015. 'Specific binding of eukaryotic ORC to DNA replication origins depends on highly conserved basic residues', *Scientific Reports*, 5.
- Kemp, M. G., M. Ghosh, G. Q. Liu, and M. Leffak. 2005. 'The histone deacetylase inhibitor trichostatin A alters the pattern of DNA replication origin activity in human cells', *Nucleic Acids Research*, 33: 325-36.

- Kermicle, Jerry L. 1996. 'Epigenetic Silencing and Activation of a Maize r Gene', *Cold Spring Harbor Laboratory Press*: 267-87.
- Kim, J. A., M. Kruhlak, F. Dotiwala, A. Nussenzweig, and J. E. Haber. 2007. 'Heterochromatin is refractory to gamma-H2AX modification in yeast and mammals', *Journal of Cell Biology*, 178: 209-18.
- Kim, Ja-Eun. 2008. 'DBC1 is a negative regulator of SIRT1.', *Nature*, 451: 583-86.
- Kim SC, Sprung R, Chen Y, Xu Y, Ball H, Pei J, Cheng T, Kho Y, Xiao H, Xiao L, Grishin NV, White M, Yang XJ, Zhao Y. 2006. 'Substrate and functional diversity of lysine acetylation revealed by a proteomics survey.', *Mol Cell.*, 4: 607-18.
- Kimura, A., T. Umehara, and M. Horikoshi. 2002. 'Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing', *Nature Genetics*, 32: 370-77.
- Kireeva ML, Komissarova N, Waugh DS, Kashlev M. 2000. 'The 8-nucleotide-long RNA:DNA hybrid is a primary stability determinant of the RNA polymerase II elongation complex.', *J Biol Chem*, 9: 6530-6.
- Kitada, T., T. Schleker, A. S. Sperling, W. Xie, S. M. Gasser, and M. Grunstein. 2011. 'gamma H2A is a component of yeast heterochromatin required for telomere elongation', *Cell Cycle*, 10: 293-300.
- Kobe B, Kajava AV. 2001. 'The leucine-rich repeat as a protein recognition motif.', *Curr Opin Struct Biol*, 11: 725-32.
- Koch F, Jourquin F, Ferrier P, Andrau JC. 2008. 'Genome-wide RNA polymerase II: not genes only!', *Trends Biochem Sci.*, 6: 265-73.
- Komarnitsky P, Eun-JC, Buratowsk S. 2000. 'Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription', *Genes & Dev*, 19: 2452-60.
- Kornberg R.D., Lorch Y. 1999. 'Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome', *Cell*, 98: 285-94.
- Kuehner JN, Pearson EL, Moore C. 2011. 'Unravelling the means to an end: RNA polymerase II transcription termination', *nat rev mol cell bio*, 5: 283-94.
- Kumar, Vinay; Abbas, Abul K.; Fausto, Nelson; Aster, Jon C. 2010. 'Robbins and Cotran Pathologic Basis of Disease.', *Philadelphia: Saunders*.
- Kuo LJ, Yang LX. 2008. 'Gamma-H2AX - a novel biomarker for DNA double-strand breaks', *in vivo*, 22: 305-9.
- Lagger, G.O'Carroll, D.Rembold, M.Khier, H.Tischler, J.Weitzer, G.Schuettengruber, B.Hauser, C.Brunmeir, R.Jenuwein, T.Seiser, C. 2002. 'Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression', *Embo Journal*, 21: 2672-81.
- Lakowski, B., I. Roelens, and S. Jacob. 2006. 'CoREST-like complexes regulate chromatin modification and neuronal gene expression', *Journal of Molecular Neuroscience*, 29: 227-39.
- Langley A.R., Stefan Gräf, James C. Smith and Torsten Krude. 2016. 'Genome-wide identification and characterisation of human DNA replication origins by initiation site sequencing (ini-seq)', *Nucleic Acids Research*.
- Lanzuolo, C.Lo Sardo, F.Diamantini, A.Orlando, V. 2011. 'PcG Complexes Set the Stage for Epigenetic Inheritance of Gene Silencing in Early S Phase before Replication', *Plos Genetics*, 7.
- Lau, P. N. I., and P. Cheung. 2011. 'Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing', *Proceedings of the National Academy of Sciences of the United States of America*, 108: 2801-06.
- Le Guezennec, X., M. Vermeulen, A. B. Brinkman, W. A. M. Hoeijmakers, A. Cohen, E. Lasonder, and H. G. Stunnenberg. 2006. 'MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties', *Molecular and Cellular Biology*, 26: 843-51.
- Lee, T. I., S. E. Johnstone, and R. A. Young. 2006. 'Chromatin immunoprecipitation and microarray-based analysis of protein location', *Nature Protocols*, 1: 729-48.
- Leea J.H, M. L. Choya, L. Ngoa, S. S. Fosterb, and Paul A. Marks. 2010. "Histone deacetylase inhibitor induces DNA damage, which normal but not transformed cells can repair." In, 14639-44. PNAS.
- Lennerz V, Fatho M, Gentilini C, Frye RA, Lifke A, Ferel D, Wölfel C, Huber C, Wölfel T. 2005. 'The response of autologous T cells to a human melanoma is dominated by mutated neoantigens.', *Proc Natl Acad Sci U S A*, 44: 16013-8.
- Levy, J. A., Virolain.M, and V. Defendi. 1968. 'HUMAN LYMPHOBLASTOID LINES FROM LYMPH NODE AND SPLEEN', *Cancer*, 22: 517-&.
- Li, C. J., and M. L. DePamphilis. 2002. 'Mammalian Orc1 protein is selectively released from chromatin and ubiquitinated during the S-to-M transition in the cell division cycle', *Molecular and Cellular Biology*, 22: 105-16.
- Li, G., Levitus, M., Bustamante, C., Widom, J. 2005. 'Rapid spontaneous accessibility of nucleosomal DNA', *Nature Structural & Molecular Biology*, 12: 46-53.

- Li, Q., H. Zhou, H. Wurtele, B. Davies, B. Horazdovsky, A. Verreault, and Z. G. Zhang. 2008. 'Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly', *Cell*, 134: 244-55.
- Li S, Fossati G, Marchetti C, Modena D, Pozzi P, Reznikov LL, Moras ML, Azam T, Abbate A, Mascagni P, Dinarello CA. 2015. 'Specific inhibition of histone deacetylase 8 reduces gene expression and production of proinflammatory cytokines in vitro and in vivo.', *J Biol Chem*, 4: 2368-78.
- Livaja R., M. Rando, O. J. 2010. 'Nucleosome positioning: How is it established, and why does it matter?', *Developmental Biology*, 339: 258-66.
- Lo, W. S. Trievel, R. C. Rojas, J. R. Duggan, L. Hsu, J. Y. Allis, C. D. Marmorstein, R. Berger, S. L. 2000. 'Phosphorylation of serine 10 in histone H3 is functionally linked in vitro and in vivo to Gcn5-mediated acetylation at lysine 14', *Molecular Cell*, 5: 917-26.
- Lodish H, Berk A, Zipursky SL, . 2000. 'The RNA polymerase II preinitiation complex', *Molecular Cell Biology*. 4th edition. New York.
- Lombard DB, Tishkoff DX, Bao J. 2011. 'Mitochondrial sirtuins in the regulation of mitochondrial activity and metabolic adaptation.', *Handb Exp Pharmacol*, 206: 163-88.
- Luijsterburg, M. S. Dinant, C. Lans, H. Stap, J. Wiernasz, E. Lagerwerf, S. Warmerdam, D. O. Lindh, M. Brink, M. C. Dobrucki, J. W. Aten, J. A. Foustari, M. I. Jansen, G. Dantuma, N. P. Vermeulen, W. Mullenders, L. H. F. Houtsmuller, A. B. Verschure, P. J. van Driel, R. 2009. 'Heterochromatin protein 1 is recruited to various types of DNA damage', *Journal of Cell Biology*, 185: 577-86.
- Lygerou, Z. 2012. 'Control over DNA replication in time and space', *Febs Journal*, 279: 12-12.
- Lynda Ratmeyer, Ravi Vinayak, Yi Yi Zhong, Gerald Zon, W. David Wilson. 1994. 'Sequence specific thermodynamic and structural properties for DNA RNA duplexes', *Biochemistry*, 33: 5298-304.
- Ma, W. X., W. S. Noble, and T. L. Bailey. 2014. 'Motif-based analysis of large nucleotide data sets using MEME-ChIP', *Nature Protocols*, 9: 1428-50.
- Mann BS1, Johnson JR, Cohen MH, Justice R, Pazdur R. 2007. 'FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma.', *oncologist*, 10: 1247-52.
- Mantiero, D., A. Mackenzie, A. Donaldson, and P. Zegerman. 2011. 'Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast', *Embo Journal*, 30: 4805-14.
- Martin, M. M., M. Ryan, R. Kim, A. L. Zakas, H. Q. Fu, C. M. Lin, W. C. Reinhold, S. R. Davis, S. Bilke, H. F. Liu, J. H. Doroshov, M. A. Reimers, M. S. Valenzuela, Y. Pommier, P. S. Meltzer, and M. I. Aladjem. 2011. 'Genome-wide depletion of replication initiation events in highly transcribed regions', *Genome Research*, 21: 1822-32.
- McCracken S, Fong N, Rosonina E, Yankulov K, Brothers G, Siderovski D, Hessel A, Foster S, Shuman S, Bentley DL. 1997. '5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II.', *Genes & Dev*, 24: 3306-18.
- Mejia-Ramirez, E., O. Limbo, P. Langerak, and P. Russell. 2015. 'Critical Function of gamma H2A in S-Phase', *Plos Genetics*, 11.
- Mersfelder, E. L., and M. R. Parthun. 2006. 'The tale beyond the tail: histone core domain modifications and the regulation of chromatin structure', *Nucleic Acids Research*, 34: 2653-62.
- Meskauskas, A., J. L. Baxter, E. A. Carr, J. Yasenachak, J. E. G. Gallagher, S. J. Baserga, and J. D. Dinman. 2003. 'Delayed rRNA processing results in significant ribosome biogenesis and functional defects', *Molecular and Cellular Biology*, 23: 1602-13.
- Mesner, L. D., V. Valsakumar, M. Cieslik, R. Pickin, J. L. Hamlin, and S. Bekiranov. 2013. 'Bubble-seq analysis of the human genome reveals distinct chromatin-mediated mechanisms for regulating early- and late-firing origins', *Genome Research*, 23: 1774-88.
- Miller, K. M. Tjeertes, J. V. Coates, J. Legube, G. Polo, S. E. Britton, S. Jackson, S. P. 2010. 'Human HDAC1 and HDAC2 function in the DNA-damage response to promote DNA nonhomologous end-joining', *Nature Structural & Molecular Biology*, 17: 1144-U15.
- Miller, S A, D D Dykes, and H F Polesky. 1988. 'A simple salting out procedure for extracting DNA from human nucleated cells.', *Nucleic Acids Research*, 16: 12-15.
- Ming M1, Han W1, Zhao B1, Sundaresan NR2, Deng CX3, Gupta MP4, He YY5. 2014. 'SIRT6 promotes COX-2 expression and acts as an oncogene in skin cancer.', *Cancer Research*, 20: 5925-33.
- Mirkin, E. V., and S. M. Mirkin. 2005. 'Mechanisms of transcription-replication collisions in bacteria', *Molecular and Cellular Biology*, 25: 888-95.
- Mischo HE, Gómez-González B, Grzechnik P, Rondón AG, Wei W, Steinmetz L, Aguilera A, Proudfoot NJ. 2011. 'Yeast Sen1 helicase protects the genome from transcription-associated instability', *mol cell*, 1: 21-32.
- Mitsui, K., D. Nakagawa, M. Nakamura, T. Okamoto, and K. Tsurugi. 2005. 'Valproic acid induces apoptosis dependent of Yca1p at concentrations that mildly affect the proliferation of yeast', *Febs Letters*, 579: 723-27.



- Mizzen C, Kuo MH, Smith E, Brownell J, Zhou J, Ohba R, Wei Y, Monaco L, Sassone-Corsi P, Allis CD. 1998. 'Signaling to chromatin through histone modifications: how clear is the signal?', *Cold Spring Harb Symp Quant Biol*, 63: 469-81.
- Mogal, A., and S. A. Abdulkadir. 2006. 'Effects of Histone Deacetylase Inhibitor (HDACi); Trichostatin-A (TSA) on the expression of housekeeping genes', *Molecular and Cellular Probes*, 20: 81-86.
- Momand, J., G. P. Zambetti, D. C. Olson, D. George, and A. J. Levine. 1992. 'THE MDM-2 ONCOGENE PRODUCT FORMS A COMPLEX WITH THE P53 PROTEIN AND INHIBITS P53-MEDIATED TRANSACTIVATION', *Cell*, 69: 1237-45.
- Montgomery, R. L., C. A. Davis, M. J. Potthoff, M. Haberland, J. Fielitz, X. X. Qi, J. A. Hill, J. A. Richardson, and E. N. Olson. 2007. 'Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility', *Genes & Development*, 21: 1790-802.
- Mottet D, S Pirotte, V Lamour, M Hagedorn, S Javerzat, A Bikfalvi, A Bellahcène, E Verdin, V Castronovo. 2008. 'HDAC4 represses p21(WAF1/Cip1) expression in human cancer cells through a Sp1-dependent, p53-independent mechanism.', *Oncogene*, 2: 243-56.
- Munoz-Galvan, S., S. Jimeno, R. Rothstein, and A. Aguilera. 2013. 'Histone H3K56 Acetylation, Rad52, and Non-DNA Repair Factors Control Double-Strand Break Repair Choice with the Sister Chromatid', *Plos Genetics*, 9.
- Munster PN, Thurn KT, Thomas S, Raha P, Lacevic M, Miller A, Melisko M, Ismail-Khan R, Rugo H, Moasser M, Minton SE. 2011. 'A phase II study of the histone deacetylase inhibitor vorinostat combined with tamoxifen for the treatment of patients with hormone therapy-resistant breast cancer.', *British Journal of Cancer*, 12: 1828-35.
- Nadel, J., R. Athanasiadou, C. Lemetre, N. A. Wijetunga, P. O. Broin, H. Sato, Z. D. Zhang, J. Jeddeloh, C. Montagna, A. Golden, C. Seoighe, and J. M. Grealley. 2015. 'RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships', *Epigenetics & Chromatin*, 8.
- Nady N, Min J, Kareta MS, Chédin F, Arrowsmith CH. 2008. 'SPOT on the chromatin landscape? Histone peptide arrays as a tool for epigenetic research.', *Trends Biochem Sci*, 33: 305-13.
- Nanney, D. L. 1958. 'EPIGENETIC CONTROL SYSTEMS', *Proc Natl Acad Sci U S A*, 44: 712-17.
- Nguyen, L. N., L. C. L. Lopes, R. J. B. Cordero, and J. D. Nosanchuk. 2011. 'Sodium butyrate inhibits pathogenic yeast growth and enhances the functions of macrophages', *Journal of Antimicrobial Chemotherapy*, 66: 2573-80.
- Noble, Denis. 2015. 'Conrad Waddington and the origin of epigenetics', *ournal of Experimental Biology*, 218: 816-18.
- O'Donovan, M. 2012. 'A critique of methods to measure cytotoxicity in mammalian cell genotoxicity assays', *Mutagenesis*, 27: 615-21.
- O'Geen, H., C. M. Nicolet, K. Blahnik, R. Green, and P. J. Farnham. 2006. 'Comparison of sample preparation methods for ChIP-chip assays', *Biotechniques*, 41: 577-80.
- Odawara, J., A. Harada, T. Yoshimi, K. Maehara, T. Tachibana, S. Okada, K. Akashi, and Y. Ohkawa. 2011. 'The classification of mRNA expression levels by the phosphorylation state of RNAPII CTD based on a combined genome-wide approach', *Bmc Genomics*, 12.
- 'OECD (2014). Test No. 487: In Vitro Mammalian Cell Micronucleus Test '. *OECD Publishing, Paris*.
- Oehme II, Deubzer HE, Wegener D, Pickert D, Linke JP, Hero B, Kopp-Schneider A, Westermann F, Ulrich SM, von Deimling A, Fischer M, Witt O. 2009. 'Histone deacetylase 8 in neuroblastoma tumorigenesis.', *Clin Cancer Res*, 1: 91-9.
- Olaharski, A. J., Z. Y. Ji, J. Y. Woo, S. Lim, A. E. Hubbard, L. P. Zhang, and M. T. Smith. 2006. 'The histone deacetylase inhibitor trichostatin A has genotoxic effects in human lymphoblasts in vitro', *Toxicological Sciences*, 93: 341-47.
- Pacek M, Tutter AV, Kubota Y, Takisawa H, Walter JC. 2006. 'Localization of MCM2-7, Cdc45, and GINS to the Site of DNA Unwinding during Eukaryotic DNA Replication', *mol cell*, 4: 581-7.
- Perissi V, Aggarwal A, Glass CK, Rose DW, Rosenfeld MG. 2004. 'A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors.', *Cell*, 4.
- Petrucelli, L. A. Dupere-Richer, D. Pettersson, F. Retrouvey, H. Skoulikas, S. Miller, W. H. 2011. 'Vorinostat Induces Reactive Oxygen Species and DNA Damage in Acute Myeloid Leukemia Cells', *Plos One*, 6: 11.
- Polach, K. J. Lowary, P. T. Widom, J. 2000. 'Effects of core histone tail domains on the equilibrium constants for dynamic DNA site accessibility in nucleosomes', *Journal of Molecular Biology*, 298: 211-23.

- Powell, James. 2014a. 'Measuring DNA damage and associated epigenetic changes genome-wide in cells following exposure to platinum analogue chemotherapeutic drugs, *institute of Cancer and Genetics, School of Medicine, Cardiff University* 2014b.
- Pozzo, F., M. Dal Bo, N. Peragine, R. Bomben, A. Zucchetto, F. M. Rossi, M. Degan, D. Rossi, A. Chiarenza, A. Grossi, F. Di Raimondo, F. Zaja, G. Pozzato, P. Secchiero, G. Gaidano, G. Del Poeta, G. Zauli, R. Foa, A. Guarini, and V. Gattei. 2013. 'Detection of TP53 dysfunction in chronic lymphocytic leukemia by an in vitro functional assay based on TP53 activation by the non-genotoxic drug Nutlin-3: a proposal for clinical application', *Journal of Hematology & Oncology*, 6: 11.
- Radhakrishnan, R., Y. X. Li, S. Y. Xiang, F. H. Yuan, Z. G. Yuan, E. Telles, J. Fang, D. Coppola, D. Shibata, W. S. Lane, Y. B. Zhang, X. H. Zhang, and E. Seto. 2015. 'Histone Deacetylase 10 Regulates DNA Mismatch Repair and May Involve the Deacetylation of MutS Homolog 2', *Journal of Biological Chemistry*, 290: 22795-804.
- Raffaella Corvi, Federica Madia 2016. 'In vitro genotoxicity testing – Can the performance be enhanced?', *Food and Chemical Toxicology*.
- Rea, S., Eisenhaber, F., O'Carroll, N., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., Jenuwein, T. 2000. 'Regulation of chromatin structure by site-specific histone H3 methyltransferases', *Nature*, 406: 593-99.
- Riggs A.D., Robert A. Martienssen, Vincenzo E.A. Russo. 1996. 'Epigenetic Mechanisms of Gene Regulation', *Cold Spring Harbor Laboratory Press*, 32: 1.
- Robyr, D., Y. Suka, I. Xenarios, S. K. Kurdistani, A. Wang, N. Suka, and M. Grunstein. 2002. 'Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases', *Cell*, 109: 437-46.
- Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. 1998. 'DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139', *J Biol Chem.*, 273: 5858-68.
- Rondinelli, B., Schwerer, H., Antonini, E., Gaviraghi, M., Lupi, A., Frenquelli, M., Cittaro, D., Segalla, S., Lemaître, J. M., Tonon, G. 2015. 'H3K4me3 demethylation by the histone demethylase KDM5C/JARID1C promotes DNA replication origin firing', *Nucleic Acids Research*, 43: 2560-74.
- Rosato RR, Almenara JA, Maggio SC, Coe S, Atadja P, Dent P, Grant S. 2008. 'Role of histone deacetylase inhibitor-induced ROS and DNA damage in LAQ-824/fludarabine antileukemic interactions', *mol cancer ther*, 10: 3285-97.
- Roy D, Yu K, Lieber MR. 2008. 'Mechanism of R-Loop Formation at Immunoglobulin Class Switch Sequences', *mol cell biol*, 1: 50-60.
- Roy D, Zhang Z, Lu Z, Hsieh CL, Lieber MR. 2010. 'Competition between the RNA transcript and the nontemplate DNA strand during R-loop formation in vitro: a nick can serve as a strong R-loop initiation site.', *mol cell biol*, 1: 146-59.
- Rundlett, S. E., A. A. Carmen, N. Suka, B. M. Turner, and M. Grunstein. 1998. 'Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3', *Nature*, 392: 831-35.
- Rundlett SE, Carmen AA, Kobayashi R, Bavykin S, Turner BM, Grunstein M. 1996. 'HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription.', *Proc Natl Acad Sci U S A.*, 25: 14503-8.
- Rundlett SE, Carmen AA, Suka N, Turner BM, Grunstein M. 1998. 'Transcriptional repression by UME6 involves deacetylation of lysine 5 of histone H4 by RPD3', *Nature*, 392: 831-5.
- Rusche, L. N., A. L. Kirchmaier, and J. Rine. 2002. 'Ordered nucleation and spreading of silenced chromatin in *Saccharomyces cerevisiae*', *Molecular Biology of the Cell*, 13: 2207-22.
- Santos-Pereira, J. M., and A. Aguilera. 2015. 'R loops: new modulators of genome dynamics and function', *Nature Reviews Genetics*, 16: 583-97.
- Sarraf, S. A., and I. Stancheva. 2004. 'Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly', *Molecular Cell*, 15: 595-605.
- Sastre R.B., Marta Solera, Cristina Oliveira-Mateosa, Anna Portelaa, Catia Moutinhoa, Sergi Sayolsa,1, Alberto Villanuevab, Manel Estellera,c,d,2, and Sonia Guila,2. 2014. 'Head-to-head antisense transcription and R-loop formation promotes transcriptional activation', *pnas*, 112: 5785–90.
- Satoh A, Brace CS, Rensing N, Cliften P, Wozniak DF, Herzog ED, Yamada KA, Imai S. 2013. 'Sirt1 extends life span and delays aging in mice through the regulation of Nk2 homeobox 1 in the DMH and LH.', *cell metab*, 3: 416-30.
- Saunders A, Leighton J. Core & John T. Lis. 2006. 'Breaking barriers to transcription elongation', *Nature Reviews Molecular Cell Biology*, 7: 557-67.
- Scott, S. L., J. D. Earle, and P. H. Gumerlock. 2003. 'Functional p53 increases prostate cancer cell survival after exposure to fractionated doses of ionizing radiation', *Cancer Research*, 63: 7190-96.

- Septenville A.D., Stéphane Duigou, Hasna Boubakri, Bénédicte Michel 2012. 'Replication Fork Reversal after Replication–Transcription Collision', *Plos Genetics*, 8.
- Shangary, S., and S. M. Wang. 2009. 'Small-Molecule Inhibitors of the MDM2-p53 Protein-Protein Interaction to Reactivate p53 Function: A Novel Approach for Cancer Therapy', *Annual Review of Pharmacology and Toxicology*, 49: 223-41.
- Shen S, Kozikowski AP. 2016. 'Why Hydroxamates May Not Be the Best Histone Deacetylase Inhibitors-What Some May Have Forgotten or Would Rather Forget?', *ChemMedChem*: 15-21.
- Shirakawa K, Chavez L, Hakre S, Calvanese V, Verdin E. 2013. 'Reactivation of latent HIV by histone deacetylase inhibitors', *Trends Microbiol*, 21: 277-85.
- Sigma-Aldrich. 2006. 'Application note. Amplification of genome-representative DNA from limited sources with GenomePlex® WGA technology for use in genetic alterations studies', *Nature Methods*, 3.
- Silverstein, R. A., and K. Ekwall. 2005. 'Sin3: a flexible regulator of global gene expression and genome stability', *Current Genetics*, 47: 1-17.
- Simoneau, A. Delgosaie, N. Celic, I. Dai, J. B. Abshiru, N. Costantino, S. Thibault, P. Boeke, J. D. Verreault, A. Wurtele, H. 2015. 'Interplay Between Histone H3 Lysine 56 Deacetylation and Chromatin Modifiers in Response to DNA Damage', *Genetics*, 200: 185-U382.
- Siow, C. C., S. R. Nieduszynska, C. A. Muller, and C. A. Nieduszynski. 2012. 'OriDB, the DNA replication origin database updated and extended', *Nucleic Acids Research*, 40: D682-D86.
- Skopek, T. R., H. L. Liber, B. W. Penman, and W. G. Thilly. 1978. 'ISOLATION OF A HUMAN LYMPHOBLASTOID LINE HETEROZYGOUS AT THYMIDINE KINASE LOCUS - POSSIBILITY FOR A RAPID HUMAN CELL MUTATION ASSAY', *Biochemical and Biophysical Research Communications*, 84: 411-16.
- Skourti-Stathaki, K., and N. J. Proudfoot. 2014. 'A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression', *Genes & Development*, 28: 1384-96.
- Skourti-Stathaki, K., N. J. Proudfoot, and N. Gromak. 2011. 'Human Senataxin Resolves RNA/DNA Hybrids Formed at Transcriptional Pause Sites to Promote Xrn2-Dependent Termination', *Molecular Cell*, 42: 794-805.
- Skourti S.K., Kamieniarz G.K., and Proudfoot N. J. 2014. 'R-loops induce repressive chromatin marks over mammalian gene terminators', *Nature*, 516.
- Slingerland M1, Guchelaar HJ, Gelderblom H. 2014. 'Histone deacetylase inhibitors: an overview of the clinical studies in solid tumors.', *Anticancer Drugs*, 2: 140-9.
- Smart, D. J., K. P. Ahmed, J. S. Harvey, and A. M. Lynch. 2011. 'Genotoxicity screening via the gamma H2AX by flow assay', *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis*, 715: 25-31.
- Smerdon, MJ. 1991. 'DNA repair and the role of chromatin structure.', *Curr Opin Cell Biol*, 3: 422-8.
- Sollier J, Cimprich KA. 2015. 'Breaking bad: R-loops and genome integrity', *Trends Cell Biol*, 9: 514-22.
- Sollier J, Stork CT1, García-Rubio ML2, Paulsen RD1, Aguilera A2, Cimprich KA3. 2014. 'Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability.', *mol cell*, 6: 777-85.
- Sonnemann, J., C. Marx, S. Becker, S. Wittig, C. D. Palani, O. H. Kramer, and J. F. Beck. 2014. 'p53-dependent and p53-independent anticancer effects of different histone deacetylase inhibitors', *British Journal of Cancer*, 110: 656-67.
- Soria, G. Polo, S. E. Almouzni, G. 2012. 'Prime, Repair, Restore: The Active Role of Chromatin in the DNA Damage Response', *Molecular Cell*, 46: 722-34.
- Stathaki K.S., Proudfoot N. J. 2014. 'A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression', *Genes & Development*, 28: 1384-96.
- Stirling, P. C., Y. J. A. Chan, S. W. Minaker, M. J. Aristizabal, I. Barrett, P. Sipahimalani, M. S. Kobor, and P. Hieter. 2012. 'R-loop-mediated genome instability in mRNA cleavage and polyadenylation mutants', *Genes & Development*, 26: 163-75.
- Strahl BD, Allis CD. 2000. 'The language of covalent histone modifications', *Nature*, 403: 41-5.
- Struhl, Kevin. 1998. 'Histone acetylation and transcriptional regulatory mechanisms', *Genes & Dev*, 12: 599-606.
- Stuckey R, García-Rodríguez N1, Aguilera A1, Wellinger RE. 2015. 'Role for RNA:DNA hybrids in origin-independent replication priming in a eukaryotic system', *Proc Natl Acad Sci U S A*, 18: 5779-84.
- Subramanian S, Susan E. Bates, John J. Wright, Igor Espinoza-Delgado and Richard L. Piekarz 2010. 'Clinical Toxicities of Histone Deacetylase Inhibitors', *Pharmaceuticals (Basel)*, 3: 2751-67.
- Suka, N., Y. Suka, A. A. Carmen, J. S. Wu, and M. Grunstein. 2001. 'Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin', *Molecular Cell*, 8: 473-79.

- Sun, Y. L., X. F. Jiang, Y. Xu, M. K. Ayrapetov, L. A. Moreau, J. R. Whetstine, and B. D. Price. 2009. 'Histone H3 methylation links DNA damage detection to activation of the tumour suppressor Tip60', *Nature Cell Biology*, 11: 1376-U273.
- Suter, B., O. Pogoutse, X. Guo, N. Krogan, P. Lewis, J. F. Greenblatt, J. Rine, and A. Emili. 2007. 'Association with the origin recognition complex suggests a novel role for histone acetyltransferase Hat1p/Hat2p', *Bmc Biology*, 5.
- Sydow JF, Cramer P. 2009. 'RNA polymerase fidelity and transcriptional proofreading', *Curr Opin Struct Biol*, 6: 732-9.
- Szilard, R. K., P. E. Jacques, L. Laramée, B. Cheng, S. Galicia, A. R. Bataille, M. Yeung, M. Mendez, M. Bergeron, F. Robert, and D. Durocher. 2010. 'Systematic identification of fragile sites via genome-wide location analysis of gamma-H2AX', *Nature Structural & Molecular Biology*, 17: 299-U62.
- Takahashi TS, Wigley DB, Walter JC. 2005. 'Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase', *Trends Biochem Sci*, 8: 437-44.
- Tang, J. B., N. W. Cho, G. F. Cui, E. M. Manion, N. M. Shanbhag, M. V. Botuyan, G. Mer, and R. A. Greenberg. 2013. 'Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination', *Nature Structural & Molecular Biology*, 20: 317-25.
- Tang, X. L., J. S. Gao, Y. J. Guan, K. E. McLane, Z. L. Yuan, B. Ramratnam, and Y. E. Chin. 2007. 'Acetylation-dependent signal transduction for type I interferon receptor', *Cell*, 131: 93-105.
- Tardat, M. Brustel, J. Kirsh, O. Lefevbre, C. Callanan, M. Sardet, C. Julien, E. 2010. 'The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells', *Nature Cell Biology*, 12: 1086-U82.
- Taunton, J. Hassig, C. A. Schreiber, S. L. 1996. 'A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p', *Science*, 272: 408-11.
- Tejs, Sebastian. 2008. 'The Ames test: a methodological short review', *ENVIRONMENTAL BIOTECHNOLOGY*, 4: 7 - 14.
- Teng, Y. M., M. Bennett, K. E. Evans, H. Zhuang-Jackson, A. Higgs, S. H. Reed, and R. Waters. 2011. 'A novel method for the genome-wide high resolution analysis of DNA damage', *Nucleic Acids Research*, 39: 6.
- Thomson, S. Clayton, A. L. Mahadevan, L. C. 2001. 'Independent dynamic regulation of histone phosphorylation and acetylation during immediate-early gene induction', *Molecular Cell*, 8: 1231-41.
- Tissenbaum HA1, Guarente L. 2001. 'Increased dosage of a sir-2 gene extends lifespan in *Caenorhabditis elegans*.', *Nature*, 410: 227-30.
- TR, Henry. 2003. 'The history of valproate in clinical neuroscience.', *Psychopharmacol Bull.*, 2: 5-16.
- Tropberger, P., S. Pott, C. Keller, K. Kamieniarsz-Gdula, M. Caron, F. Richter, G. H. Li, G. Mittler, E. T. Liu, M. Buhler, R. Margueron, and R. Schneider. 2013. 'Regulation of Transcription through Acetylation of H3K122 on the Lateral Surface of the Histone Octamer', *Cell*, 152: 859-72.
- Tuduri S, et al. 2010. 'Topoisomerase I suppresses genomic instability by preventing interference between replication and transcription', *Nature Cell Biology*, 12: 1110-22.
- Turinetto V., and Giachino C. 2015. 'Multiple facets of histone variant H2AX: a DNA double-strand-break marker with several biological functions', *Nucleic Acids Res*, 43: 2489-98.
- Unnikrishnan, A., P. R. Gafken, and T. Tsukiyama. 2010. 'Dynamic changes in histone acetylation regulate origins of DNA replication', *Nature Structural & Molecular Biology*, 17: 430-U68.
- Vakoc CR, Mandat SA, Olenchock BA, Blobel GA. 2005. 'Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin', *Mol. Cell*, 19: 381-91.
- van Leeuwen, F., Gafken, P. R., Gottschling, D. E. 2002. 'Dot1p modulates silencing in yeast by methylation of the nucleosome core', *Cell*, 109: 745-56.
- Van Speybroeck, L. 2002. 'From epigenesis to epigenetics - The case of C. H. Waddington', *From Epigenesis to Epigenetics: the Genome in Context*, 981: 61-81.
- Vannini A, Volpari C, Filocamo G, Casavola EC, Brunetti M, Renzoni D, Chakravarty P, Paolini C, De Francesco R, Gallinari P, Steinkühler C, Di Marco S. 2004. 'Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor.', *Proc Natl Acad Sci U S A.*, 42: 15064-9.
- Vashee, S., C. Cvetic, W. Y. Lu, P. Simancek, T. J. Kelly, and J. C. Walter. 2003. 'Sequence-independent DNA binding and replication initiation by the human origin recognition complex', *Genes & Development*, 17: 1894-908.
- Vecsey CG, Hawk JD, Lattal KM, Stein JM, Fabian SA, Attner MA, Cabrera SM, McDonough CB, Brindle PK, Abel T, Wood MA. 2007. 'Histone deacetylase inhibitors enhance memory and synaptic plasticity via CREB:CBP-dependent transcriptional activation.', *J Neurosci.*, 23: 6128-40.

- Vermeulen M, Eberl HC, Matarese F, Marks H, Denissov S, Butter F, Lee KK, Olsen JV, Hyman AA, Stunnenberg HG, Mann M. 2010. 'Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers', *Cell*, 142: 967-80.
- Vidali G, Gershey EL, Allfrey VG. 1968. 'Chemical studies of histone acetylation. The distribution of epsilon-N-acetyllysine in calf thymus histones', *J Biol Chem*, 243: 6361-6.
- Villagra, A., F. Cheng, H. W. Wang, I. Suarez, M. Glozak, M. Maurin, D. Nguyen, K. L. Wright, P. W. Atadja, K. Bhalla, J. Pinilla-Ibarz, E. Seto, and E. M. Sotomayor. 2009. 'The histone deacetylase HDAC11 regulates the expression of interleukin 10 and immune tolerance', *Nature Immunology*, 10: 92-100.
- Vogelauer, M., L. Rubbi, I. Lucas, B. J. Brewer, and M. Grunstein. 2002. 'Histone acetylation regulates the time of replication origin firing', *Molecular Cell*, 10: 1223-33.
- Vu Q.A, Zhang DE, Chreoneos ZC, Nelson DA. 1987. 'Polyamines inhibit the yeast histone deacetylase', *Febs Letters*, 220: 79-83.
- Wakimoto, B. T. 1998. 'Beyond the nucleosome: Epigenetic aspects of position-effect variegation in *Drosophila*', *Cell*, 93: 321-24.
- Wan, Y. K., W. M. Chen, J. Xing, J. Y. Tan, B. Li, H. Y. Chen, Z. X. Lin, J. H. Chiang, and R. A. Saleem. 2011. 'Transcriptome profiling reveals a novel role for trichostatin A in antagonizing histone chaperone Chz1 mediated telomere anti-silencing (vol 585, pg 2519, 2011)', *Febs Letters*, 585: 3146-46.
- Wang, A., S. K. Kurdistani, and M. Grunstein. 2002. 'Requirement of Hos2 histone deacetylase for gene activity in yeast', *Science*, 298: 1412-14.
- Wang, H. B., and Y. Zhang. 2001. 'Mi2, an auto-antigen for dermatomyositis, is an ATP-dependent nucleosome remodeling factor', *Nucleic Acids Research*, 29: 2517-21.
- Wang Z, Chen W. 2013. 'Emerging Roles of SIRT1 in Cancer Drug Resistance.', *genes cancer*, 3: 82-90.
- Wei, Y. Yu, L. L. Bowen, J. Gorovsky, M. A. Allis, C. D. 1999. 'Phosphorylation of histone H3 is required for proper chromosome condensation and segregation', *Cell*, 97: 99-109.
- Weintraub H., Abraham Worcel, Bruce Alberts. 1976. 'A model for chromatin based upon two symmetrically paired half-nucleosomes', *Cell*, 9: 409-17.
- Wisniewski, J. R. Zougman, A. Kruger, S. Mann, M. 2007. 'Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue', *Molecular & Cellular Proteomics*, 6: 72-87.
- Wolffe AP, Hayes JJ. 1999. 'Chromatin disruption and modification.', *Nucleic Acids Research*, 27: 711-20.
- Wu J, Suka N, Carlson M, Grunstein M. 2001. 'TUP1 Utilizes Histone H3/H2B-Specific HDA1 Deacetylase to Repress Gene Activity in Yeast', *mol cell*, 1: 117-26.
- Wurtele, H. Kaiser, G. S. Bacal, J. St-Hilaire, E. Lee, E. H. Tsao, S. Dorn, J. Maddox, P. Lisby, M. Pasero, P. Verreault, A. 2012. 'Histone H3 Lysine 56 Acetylation and the Response to DNA Replication Fork Damage', *Molecular and Cellular Biology*, 32: 154-72.
- Xia, F., X. Wang, Y. H. Wang, N. M. Tsang, D. W. Yandell, K. T. Kelsey, and H. L. Liber. 1995. 'ALTERED P53 STATUS CORRELATES WITH DIFFERENCES IN SENSITIVITY TO RADIATION-INDUCED MUTATION AND APOPTOSIS IN 2 CLOSELY-RELATED HUMAN LYMPHOBLAST LINES', *Cancer Research*, 55: 12-15.
- Xiaojuan Sun, Lei Wei, Qian Chen, and Richard M. Terek. 2009. 'HDAC4 represses vascular endothelial growth factor expression in chondrosarcoma by modulating RUNX2 activity.', *J Biol Chem*, 33: 21881-90.
- Xu, M. Long, C. Z. Chen, X. Z. Huang, C. Chen, S. Zhu, B. 2010. 'Partitioning of Histone H3-H4 Tetramers During DNA Replication-Dependent Chromatin Assembly', *Science*, 328: 94-98.
- Yang H1, Yang T, Baur JA, Perez E, Matsui T, Carmona JJ, Lamming DW, Souza-Pinto NC, Bohr VA, Rosenzweig A, de Cabo R, Sauve AA, Sinclair DA. 2007. 'Nutrient-Sensitive Mitochondrial NAD+ Levels Dictate Cell Survival.', *Cell*, 6: 1095-107.
- Yang Q, Wang B, Gao W, Huang S, Liu Z, Li W, Jia J. 2013. 'SIRT1 is downregulated in gastric cancer and leads to G1-phase arrest via NF-κB/Cyclin D1 signaling.', *mol cancer res*, 12: 1497-507.
- Yardley DA, Ismail-Khan RR, Melichar B, Lichinitser M, Munster PN, Klein PM, Cruickshank S, Miller KD, Lee MJ, Trepel JB. 2013. 'Randomized Phase II, Double-Blind, Placebo-Controlled Study of Exemestane With or Without Entinostat in Postmenopausal Women With Locally Recurrent or Metastatic Estrogen Receptor-Positive Breast Cancer Progressing on Treatment With a Nonsteroidal Aromatase Inhibitor.', *J Clin Oncol.*, 17: 2128-35.
- Yekezare, M., B. Gomez-Gonzalez, and J. F. X. Diffley. 2013. 'Controlling DNA replication origins in response to DNA damage - inhibit globally, activate locally', *Journal of Cell Science*, 126: 1297-306.

- Yoo, E. J., and B. M. Lee. 2005. 'Comparative mutagenicity of apicidin and apicidin derivatives (SD-0203 and SD-2007), histone deacetylase inhibitors', *Journal of Toxicology and Environmental Health-Part a-Current Issues*, 68: 2097-109.
- Younes A, Oki Y, Bociek RG, Kuruvilla J, Fanale M, Neelapu S, Copeland A, Buglio D, Galal A, Besterman J, Li Z, Drouin M, Patterson T, Ward MR, Paulus JK, Ji Y, Medeiros LJ, Martell RE. 2011. 'Mocetinostat for relapsed classical Hodgkin's lymphoma: an open-label, single-arm, phase 2 trial', *lancet*, 13: 1222-8.
- Yu, S. R., Y. M. Teng, R. Waters, and S. H. Reed. 2011. 'How Chromatin Is Remodelled during DNA Repair of UV-Induced DNA Damage in *Saccharomyces cerevisiae*', *Plos Genetics*, 7: 9.
- Yu, Y. C. Deng, Y. B. Reed, S. H. Millar, C. B. Waters, R. 2013. 'Histone variant Htz1 promotes histone H3 acetylation to enhance nucleotide excision repair in Htz1 nucleosomes', *Nucleic Acids Research*, 41: 9006-19.
- Zhang D, Yoon HG, Wong J. 2005. 'JMJD2A is a novel N-CoR-interacting protein and is involved in repression of the human transcription factor achaete scute-like homologue 2 (ASCL2/Hash2). *Mol Biol Cell*, 15: 6404-14.
- Zhang L, Eugeni EE, Parthun MR, Freitas MA. 2003. 'Identification of novel histone post-translational modifications by peptide mass fingerprinting', *Chromosoma.*, 112: 77-86.
- Zhang, Y., R. Iratni, H. ErdjumentBromage, P. Tempst, and D. Reinberg. 1997. 'Histone deacetylases and SAP18, a novel polypeptide, are components of a human Sin3 complex', *Cell*, 89: 357-64.
- Zhang Z.Z., Nicholas R Pannunzio, Chih-Lin Hsieh, Kefei Yu and Michael R Lieber. 2015. 'Complexities due to single-stranded RNA during antibody detection of genomic rna:dna hybrids', *BMC research notes*, 8.
- Zhou, Q., Z. K. Melkounian, A. Lucktong, M. Moniwa, J. R. Davie, and J. S. Strobl. 2000. 'Rapid induction of histone hyperacetylation and cellular differentiation in human breast tumor cell lines following degradation of histone deacetylase-1', *Journal of Biological Chemistry*, 275: 35256-63.
- Zlotorynski, Eytan. 2015. 'Transcriptional mutagenesis by R-loops', *Nature Reviews Molecular Cell Biology*, 17.