Molecular insights into host gene regulation by Epstein-Barr virus nuclear antigen EBNA3C

Jens Sebastian Kalchschmidt

Thesis submitted for the degree of Doctor of Philosophy

Molecular Virology

Department of Medicine

Imperial College London

May 2016

Abstract

Most adults are infected by Epstein-Barr virus (EBV), which establishes persistent infection in B cells. EBV nuclear antigen 3C (EBNA3C) is a latent viral protein that co-regulates many cellular genes, but little is known about the mechanisms by which EBNA3C acts as repressor or activator of gene expression. This study aimed to further explore molecular mechanisms that underlie EBNA3C host gene regulation. For this, four EBNA3C target genes – three repressed (*COBLL1*, *ADAM28* and *ADAMDEC1*) and one induced (*AICDA*) – were selected based on a previous exon microarray on EBV-recombinant lymphoblastoid cell lines (LCL) carrying conditional EBNA3C (3CHT).

Using EBV recombinants to infect primary B cells, EBNA3C was shown to be the main regulator of all four genes. Surprisingly, regulation by EBNA3C over orders of magnitude followed highly exponential activation or repression profiles and required ~30 days after infection. This could be reproduced efficiently in the 3CHT LCLs by activating EBNA3C, proving the utility of these cells to replicate EBNA3C gene regulation. Analysis of chromatin immunoprecipitation (ChIP) coupled to deep-sequencing identified EBNA3C binding to distal regulatory elements at all four genes and detailed ChIP analysis of histone modifications and cellular factors was performed. Unexpectedly, recruitment and/or stabilisation of the DNA-sequence binding factor RBPJ was observed at the EBNA3C binding sites only when EBNA3C was functional. This challenges existing models of how RBPJ functions in EBV-regulated gene expression. EBNA3C failed to regulate all four genes when it is unable to bind to RBPJ. Recruitment of a polycomb protein, generally linked to gene repression, was observed at all four genes irrespective of whether repression or activation by EBNA3C occurred. Changes to histone modifications at these loci correlated well with gene expression and indicate a two-step mechanism for EBNA3C-mediated repression and suggest the activation is linked to a very complex pattern of chromatin looping.

Declaration of Originality

I hereby declare that all work presented in this thesis has been produced by me, except where clearly stated in the text.

Jens Sebastian Kalchschmidt

May 2016

Copyright Declaration

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

Acknowledgments

I would like to thank my supervisors Professor Martin Allday and Dr Kostas Paschos for their constant guidance, advice and support. I am deeply grateful for everything that you taught me during my PhD both scientifically and personally. I am certain that it will help me in the future and I will always treasure this experience – thank you so much!

I wish to thank Gillian Parker for always providing help whatever the matter and for ensuring that the lab runs smoothly, Dr Goedele Maertens for her help with the SUZ12 protein purification, Dr Quentin Bazot for his help with the chromatin conformation capture technique, Dr Malte Paulsen for his advice and help with flow cytometry and Dr Rob White for general advice and his help with virus production.

I would also like to thank everybody from the Allday and White laboratories for the great time that we had in and outside the lab. In particular, Agnieszka Szymula and Guiyi Ho for always sticking together through thick and thin from the beginning of our PhDs. I will miss all the funny conversations we had and I hope that we will see each other in a year's time at Royal Albert Hall! Adam Gillman, Christine Styles, Mohammed Ba Abdullah, Dr Quentin Bazot and Dr Richard Palermo for joining in later and for contributing to all those funny conversations. Adam, Christine and Mohammed good luck for the rest of your PhDs! I also like to thank everybody in the Section of Virology and I will never forget the Friday pubs.

I am grateful to the Wellcome Trust for their sponsorship and I like to thank all organisers and students of the PhD programme Molecular and Cellular Basis of Infection, particularly Bevin, Luke, Rebecca and Thomas. I would also like to thank Rachael, Natalie and both Eds for the fun times living together and Etuna, Thomas, Malte and Remi for the great times outside the lab and on the climbing walls.

Last but not least, I would like to thank my family, in particular my parents and sister, for always supporting and encouraging me. Being far away it not always easy, but you will always be close to my heart.

Table of Contents

Abstract **2**

Declaration of Originality **3**

Copyright Declaration **4**

Acknowledgments **5**

Table of Contents **6**

List of figures **12**

List of tables **16**

List of abbreviations **17**

1 Chapter I: Introduction **24**

1.1 B cell biology **24**

1.1.1 Immunoglobulins 24

1.1.2 B cell development 25

1.1.3 T cell dependent B cell activation 28

1.1.4 Germinal centre reaction 29

1.1.5 Activation-induced cytidine deaminase 32

1.1.5.1 Mechanisms underlying AID-induced mutations 33

1.1.5.2 Somatic hypermutation (SHM) and class switch recombination (CSR) 35

1.1.5.3 Regulation of AID activity 38

1.1.5.3.1 Transcriptional regulation of AID 39

1.1.5.3.2 Post-transcriptional regulation of AID 41

1.1.5.3.3 Compartmentalisation and post-translational regulation of AID 42

1.1.5.3.4 Targeting of AID 42

1.2 Epstein-Barr virus **44**

1.2.1 Overview 44
1.2.2 Establishment of latent infection 44
1.2.2.1 In vivo infection 44
1.2.2.2 In vitro infection 47
1.2.3 Latent gene transcription from the circular EBV genome 48
1.2.4 Latent viral proteins 49
1.2.4.1 EBV nuclear antigens 49
1.2.4.1.1 EBNA1 49
1.2.4.1.2 EBNA2 49
1.2.4.1.3 EBNA3A, EBNA3B, EBNA3C 51
1.2.4.1.4 EBNA-LP 56
1.2.4.2 Latent membrane proteins 57
1.2.4.2.1 LMP1 57
1.2.4.2.2 LMP2A and LMP2B 58
1.2.5 EBV-associated lymphomas 58
1.2.5.1 B cell lymphomagenesis 59
1.2.5.2 Burkitt's lymphomas 59
1.2.5.3 Hodgkin lymphomas 61
1.2.5.4 Immunodeficiency-associated lymphoproliferative disease 61
1.3 Gene regulation 62
1.3.1 Overview 62
1.3.2 Positioning and higher order chromatin structure of chromatin 62
1.3.2.1 Chromosome territories 63
1.3.2.2 Topologically associating domains 64
1.3.2.3 Long-range chromatin interactions 64
1.3.3 Chromatin and chromatin modifications 66
1.3.3.1 DNA methylation 67
1.3.3.2 Histone modifications and modifiers 68
1.3.3.2.1 Polycomb-group proteins 69

1.3.4 Control of gene expression by chromatin interactions and modifications 74

1.4 Background to this project 77

1.4.1 Notch signalling pathway 77

1.4.2 Gene regulation by EBNA2 and EBNA3 family 78

1.4.3 Available recombinant EBV 83

1.4.4 Identification of EBNA3C target genes 86

1.5 Aims of this PhD project 87

2 Chapter II: Materials and Methods 88

2.1 Solutions and buffers 88

2.2 Cell culture 93

2.2.1 Cell lines and culture conditions 93

2.2.2 Freezing and thawing cells 95

2.2.3 Harvesting cells 95

2.2.4 Time-course experiments 96

2.2.4.1 EBNA3C-conditional LCL 96

2.2.4.2 EBNA2-conditional LCL 96

2.3 Validation of HEK293 EBV producing cell lines and virus production 97

2.3.1 Validation of HEK293 EBV producing cell lines by episomal rescue 97

2.3.1.1 Extraction of EBV BACs from HEK293 virus producers 97

2.3.1.2 Transformation of recombinant EBV BACs into E. coli 98

2.3.1.3 Extraction of EBV BACs from E. coli 98

2.3.1.4 Restriction enzyme digestion 98

2.3.1.5 Pulsed-field gel electrophoresis 99

2.3.2 Virus production 99

2.3.2.1 Transfection and harvest 99

2.3.2.2 Virus titration 100

2.4 Isolation and infection of primary B cells 100

2.4.1 Isolation of primary B cells 100

2.4.2 Infection of primary B cells with recombinant EBV 100

2.5 Reverse transcription quantitative PCR (RT-qPCR) 101

2.6 Western blotting 102

2.7 Chromatin immunoprecipitation (ChIP) 105

2.8 Transient reporter assays 108

2.8.1 Cloning of transient reporter vectors 108

2.8.2 Cloning of RBPJ binding mutant (BM) EBNA3C into expression vector 111

2.8.3 Electroporation and transient reporter assays 111

3 Chapter III: Gene repression by EBNA3C 113

3.1 Introduction 113

3.2 Results 114

3.2.1 Validation of EBNA3C repressed genes in EBV-infected primary B cells 114

3.2.2 Recapitulation of EBNA3C-mediated gene repression in EBNA3C-conditional LCL 119

3.2.3 Comparison between primary B and EBNA3C-conditional LCL 122

3.2.4 Molecular mechanism of EBNA3C-mediated gene repression 124

3.2.4.1 EBNA3 binding at the COBLL1 and the ADAM28/ADAMDEC1 loci 124

3.2.4.2 Changes of epigenetic histone marks correlate with gene expression levels 126

3.2.4.3 Factor recruitment to the COBLL1 or ADAM28/ADAMDEC1 loci 130

3.2.5 Recapitulation of repression in transient reporter assays 134

3.2.6 Repression is dependent on interaction between RBPJ and EBNA3C 139

3.3 Summary 141

3.4 Discussion 143

3.4.1 Recruitment of RBPJ to chromatin requires functional EBNA3C 143

3.4.2 Two step mechanism of EBNA3C-mediated gene repression 144

3.4.2.1 First step: Loss of activation-associated histone marks 144

3.4.2.2 Second step: Recruitment of polycomb proteins and deposition of repressive histone marks 145
3.4.3 Simultaneous recruitment of PRC1 and PRC2 proteins to two distinct loci 146
3.4.4 Comparisons with EBNA3A-mediated repression 147
3.4.5 Biological role of *ADAM28*, *ADAMDEC1* and *COBLL1* – why are they repressed by EBV? 148
3.4.6 Remarks 152
3.4.7 Conclusion 152

4 Chapter IV: Gene activation by EBNA3C 153

4.1 Introduction 153

4.2 Results 153

4.2.1 Validation of *AICDA* as an EBNA3C-activated gene in primary B cells infected with EBV 153
4.2.2 Recapitulation of EBNA3C-mediated gene activation of *AICDA* in EBNA3C-conditional LCL 155
4.2.3 Comparison between primary B and EBNA3C-conditional LCL 157
4.2.4 Molecular mechanism of EBNA3C-mediated gene activation 158
4.2.4.1 EBNA3C-binding at regulatory regions of AICDA 158
4.2.4.2 Changes of epigenetic histone marks control expression levels 160
4.2.4.3 Factor recruitment to the AICDA locus 163
4.2.5 Induction of *AICDA* is dependent on interaction between RBPJ and EBNA3C 166
4.2.6 EBNA2 represses and LMP1 induces expression of AID 167
4.2.7 Transient reporter assays in an attempt to recapitulate regulation of *AICDA* by EBNA3C and EBNA2 171

4.3 Summary 173

4.4 Discussion 175

4.4.1 Dynamic recruitment of RBPJ during EBNA3C-mediated induction of *AICDA* 175

4.4.2 Role of BMI1 (PRC1) and p300 in EBNA3C-mediated gene activation 176
4.4.3 Regulation of *AICDA* by EBNA2 and LMP1 179
4.4.4 Why does EBV activate AID expression and what at are the consequences for the EBV-infected B cells? 182
4.4.5 Conclusion 187

5 Chapter V: General discussion of gene repression and activation by EBNA3C 189

5.1 Similarities between EBNA3C-mediated repression and activation 190
5.1.1 Kinetics of EBNA3C-mediated gene regulation 190
5.1.2 Recruitment of RBPJ is dependent on presence of EBNA3C 193
5.2 Differences between EBNA3C-mediated repression and activation 196
5.3 Open questions 197
5.3.1 How is EBNA3C recruited to chromatin and which factors are important? 197
5.3.2 Why does EBNA3A bind to the COBLL1 peak and the regulatory regions of AICDA although it is not involved in their regulation? 198
5.3.3 How are PRC1 and PRC2 proteins recruited? 200
5.3.3.1 Default mechanism 200
5.3.3.2 Direct recruitment 201
5.3.4 What determines whether a locus is repressed or activated by EBNA3C? 203
5.3.4.1 Distinct forms of repressive or activating EBNA3C complexes 203
5.3.4.2 Sequence of the EBNA3-binding site 203
5.3.4.3 Locus specific features or factors 204
5.4 Final conclusions 204

6 References 206

7 Appendix 250

List of figures

Figure 1.1: Schematic overview of immunoglobulin heavy (IgH) and light (IgL) chains as part of the B cell receptor on the cell surface or as secreted antibody. 25

Figure 1.2: B cell development and overview of $V_H(D)J_H$ and V_L-J_L recombination. 27

Figure 1.3: Germinal centre reaction after antigen-specific B cell activation. 31

Figure 1.4: Mutations caused by activation-induced cytidine deaminase (AID). 34

Figure 1.5: Somatic hypermutation (SHM) and class switch recombinantion (CSR) induced by activation-induced cytidine deaminase (AID). 37

Figure 1.6: Activity of AID is controlled on multiple layers. 38

Figure 1.7: *In vivo* infection with EBV in the germinal centre model of persistence. 46

Figure 1.8: *In vitro* infection with EBV leads to the establishment of continuously proliferating lymphoblastoid cell lines (LCL). 47

Figure 1.9: Latent protein-coding gene transcription from circular EBV genome. 48

Figure 1.10: Overview of the EBNA3 family of proteins and selected interactions between EBNA3C and cellular factors. 53

Figure 1.11: Positioning and higher order chromatin structures in the nucleus. 65

Figure 1.12: Chromatin structure and modifications. 66

Figure 1.13: Overview of histone modifications. 68

Figure 1.14: Composition of mammalian PCR1 and PRC2 complexes. 71

Figure 1.15: Regulation of gene expression. 76

Figure 1.16: Model of gene regulation by EBNA2 and EBNA3C. 79

Figure 1.17: Overview of recombinant EBV available and production of infectious EBV virions. 85

Figure 2.1: Luciferase vectors cloned during this study. 110

Figure 3.1: Verification of HEK293 virus producer cell lines. 115

Figure 3.2: Repression of *ADAM28*, *ADAMDEC1* and *COBLL1* in primary B cells infected with recombinant EBV. 117

Figure 3.3: Western blot analysis of LCL established from primary B cell infections. 118

Figure 3.4: Repression of *ADAM28*, *ADAMDEC1* and *COBLL1* in EBNA3C-conditional LCL. 120

Figure 3.5: HT does not repress EBNA3C-target genes in non-conditional 3CKO LCL. 121

Figure 3.6: Comparative gene expression analysis between newly EBV-infected primary B cells and 3CHT LCL after activation of EBNA3C. 123

Figure 3.7: EBNA3A and EBNA3C binding at *COBLL1* and *ADAM28/ADAMDEC1*. 125

Figure 3.8: Changes in epigenetic marks at the *COBLL1* and *ADAM28/ADAMDEC1* loci after activation of EBNA3C during the 3CHT A13 time-course. 127

Figure 3.9: Correlation of changes of epigenetic marks at the TSS of *ADAM28*, *ADAMDEC1* and *COBLL1* with gene expression. 129

Figure 3.10: Factor binding to the *COBLL1* and *ADAM28/ADAMDEC1* loci during the 3CHT A13 time-course. 131

Figure 3.11: Binding of SUZ12 to the TSS of *COBLL1* and BMI1 and RBPJ to *ADAM* peak and *COBLL1* peak after activation of EBNA3C. 132

Figure 3.12: Western blot analysis of 3CHT A13 time-course. 133

Figure 3.13: Transient reporter assays of *ADAM28* constructs – *ADAM* peak acts as an enhancer and repression by EBNA3A and EBNA3C is dependent on *ADAM* peak and RBPJ. 137

Figure 3.14: Transient reporter assays of *COBLL1* constructs – *COBLL1* peak acts as an enhancer and repression by EBNA3C is dependent on *COBLL1* peak and RBPJ. 138

Figure 3.15: Ability of EBNA3C to bind to RBPJ is essential for the repression of *ADAM28*, *ADAMDEC1* and *COBLL1*. 140

Figure 3.16: Two-step model of EBNA3C-mediated gene repression. 142

Figure 4.1: Induction of *AICDA* by EBNA3C in primary B cells infected with recombinant EBV. 154

Figure 4.2: Induction of *AICDA* during the 3CHT A13 time-course and comparison across various cell lines. 156

Figure 4.3: HT does not induce AID in non-conditional 3CKO LCL. 157

Figure 4.4: Comparison of *AICDA* induction between newly infected primary B cells and 3CHT LCL after activation of EBNA3C. 157

Figure 4.5: EBNA3C binding at the *AICDA* locus. 159

Figure 4.6: Changes in epigenetic marks across the *AICDA* locus during the 3CHT A13 time-course. 161

Figure 4.7: Correlation between changes of histone marks H3K4me3 and H3K9ac at the TSS of *AICDA* and gene expression. 162

Figure 4.8: Factor binding to the *AICDA* locus during the 3CHT A13 time-course. 164

Figure 4.9: Binding of BMI1, p300 and RBPJ to *AICDA* region IV and VI after activation of EBNA3C. 165

Figure 4.10: Ability of EBNA3C to bind to RBPJ is essential for the activation of *AICDA*. .. 166

Figure 4.11: Repression of AID by EBNA2 in EREB 2-5 cells. 168

Figure 4.12: Global EBNA2, EBNA3 and LMP1 protein levels and occupancy of EBNA2 at the *AICDA* locus during the 3CHT A13 time-course. 170

Figure 4.13: Attempt to recapitulate regulation of *AICDA* in transient reporter assays. 172

Figure 4.14: Model of gene activation by EBNA3C at AID genomic locus. 174

Figure 4.15: Higher order chromatin structure of the *AICDA* locus in LCL GM12878. 178

Figure 4.16: NFκB binding at *AICDA* locus in GM12878. 181

Figure 4.17: Are there different routes to EBV latency and what is the role of AID? 184

Figure 5.1: Gene expression analysis of EBNA3C-regulated host genes. 192

Figure 5.2: Binding of RBPJ and BMI1 to EBNA3C repressed and activated genes. 195

Figure 7.1: Gene expression analysis in the 3CHT C19 time-course. 251

Figure 7.2: Western blot analysis of the 3CHT C19 time-course. 252

Figure 7.3: ChIP analysis of the *COBLL1* locus during the 3CHT C19 time-course. 253

Figure 7.4: ChIP analysis of the *ADAM28-ADAMDEC1* locus during the 3CHT C19 time-course. 254

Figure 7.5: No changes to EBNA2 occupancy at *ADAM* peak and *COBLL1* peak during the 3CHT A13 time-course. 255

Figure 7.6: ChIP analysis of the *AICDA* locus during the 3CHT C19 time-course. 257

Figure 7.7: Occupancy of SUZ12 across the *AICDA* locus during the 3CHT A13 time-course. 258

Figure 7.8: SHM and clonal development during 3CHT A13 time-course. 259

Figure 7.9: EBNA3A and EBNA3C binding sites on *ADAM28/ADAMDEC1* locus are co-occupied by multiple transcription factors. 261

Figure 7.10: EBNA3A and EBNA3C binding sites on *COBLL1* locus are co-occupied by multiple transcription factors. 263

Figure 7.11: EBNA3A, EBNA3C and EBNA2 binding sites on *AICDA* locus are co-occupied by multiple transcription factors. 265

List of tables

Table 1.1: Top 10 genes repressed by EBNA3C 86

Table 1.2: Top 10 genes induced by EBNA3C 87

Table 2.1: Cell lines used in this project 94

Table 2.2: Sequences of RT-qPCR primers 102

Table 2.3: Composition of resolving and stacking gels for SDS-PAGE 103

Table 2.4: Primary antibodies for western blotting 104

Table 2.5: Secondary antibodies for western blotting 104

Table 2.6: Antibodies used for ChIP 106

Table 2.7: Sequences of ChIP primers 107

Table 2.8: Sequences of primers used to clone transient reporter vectors 109

Table 2.9: Vectors used in transient reporter assays 112

Table 7.1: Overview of restriction enzyme digestion fragments for recombinant EBV BACs used in this study. 250

List of abbreviations

1°B	primary B cell
3CHT	EBNA3C-conditional cell line
3D	three-dimensional
ADAM	a disintegrin and metalloprotease
ADAM28	ADAM family member 28 (MDC-L, eMDCII and TECADAM)
ADAMDEC1	ADAM-like decysin 1
AICDA	gene encoding activation-induced cytidine deaminase (AID)
AID	activation-induced cytidine deaminase (encoded by AICDA)
ALAS1	5'-aminolevulinate synthase 1
APE1	apurinic/apyridimic endonuclease 1
APOBEC	apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like
APS	ammonium persulfate
ATR	ataxia telangiectasia and Rad3 related
AUTS2	autism susceptibility gene 2
BAC	bacterial artificial chromosome
BCL2	B cell lymphoma 2 protein
BCL2L11	gene encoding BCL2-interacting mediator (BIM)
BCL6	B cell lymphoma 6 protein
BER	base excision repair
BIM	BCL2-interacting mediator
BL	Burkitt's lymphoma
BM	binding mutant
bp	base pair
C/EBP	CCAAT-enhancer binding protein
CART	C-terminal activation regions
CBX	chromobox-domain family
CD	cluster of differentiation
CD40L	CD40 ligand
CDK	cyclin-dependent kinase

CDKN2A	gene encoding cyclin-dependent kinase inhibitor 2A ($p16^{INK4A}$)
cDNA	complementary DNA
CDR	complementarity determining region
ChIA-PET	chromatin interaction analysis by paired-end tag sequencing
ChIP	chromatin immunoprecipitation
ChIP-qPCR	chromatin immunoprecipitation followed by qPCR
ChIP-Seq	chromatin immunoprecipitation followed by deep-sequencing
CIR	CBF1-interacting co-repressor
CML	chronic myelogenous leukemia
COBL	Cordon-bleu
COBLL1	Cordon-bleu like 1
Cp	C promoter
CpG	Cytosine-phosphate-guanine dinucleotide
cPRC1	canonical PRC1 complex
CSR	class switch recombination
CTAR	C terminal activation region
CtBP	C-terminal binding protein
CTBP2	C terminal binding protein 2
CTCF	CCCTC-binding factor
CtIP	CtBP-interacting protein
CTL	cytotoxic T cells
CXCR4	C-X-C chemokine receptor type 4
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
DSB	double strand break
DTT	dithiothreitol
DZ	dark zone
E. coli	*Escherichia coli*
EBER	Epstein-Barr virus-encoded small RNA
EBF	early B cell factor

eBL	endemic Burkitt's lymphoma
EBNA	Epstein-Barr nuclear antigen
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetraacetic acid
EED	embryonic ectoderm development
eEF1A	eukaryotic elongation factor 1
EGF	epidermal growth factor
EGR-1	early growth response 1
ENCODE	ENCyclopedia Of DNA Elements
ePICh	end-targeting proteomics of isolated chromatin fragments
ESC	embryonic stem cell
Exo1	exonuclease 1
EZH	enhancer of zeste
FBS	foetal bovine serum
FDC	follicular dendritic cells
FISH	fluorescence *in situ* hybridisation
GAPHD	glyceraldehyde 3-phosphate dehydrogenase
GC	germinal centres
GFP	green fluorescent protein
GNB2L1	guanine nucleotide binding protein beta polypeptide 2-like 1
GRU	green Raji unit
H	histone
H3K27ac	histone 27 lysine 27 acetylation
H3K27me3	histone 27 lysine 27 trimethylation
H3K4me3	histone 3 lysine 4 trimethylation
H3K9ac	histone 3 lysine 9 acetylation
HAT	histone acetyltransferase
HBV	hepatitis B virus
HCV	hepatitis C virus
HDAC	histone deacetylase
Hi-C	chromosome conformation capture technique

HIV	human immunodeficiency virus
HL	Hodgkin lymphoma
HOTAIR	HOX transcript antisense RNA
HOX	homeobox gene cluster
HoxC4	Homeobox C4
HPH	human polyhomeotic homolog family
HPV	human papilloma virus
HRP	horseradish peroxidase
HRS	Hodgkin-Reed-Sternberg cells
HSC	haematopoietic stem cells
HSP90	heat-shock protein 90 kDa
HT	4-hydroxytamoxifen
HTLV-1	human T lymphotropic virus 1
Id	inhibitor of differentiation
Ig	immunoglobulin
Ig-α	signaling molecule of BCR (also known as CD79a)
Ig-β	signaling molecule of BCR (also known as CD79b)
IGF2	insulin-like growth factor 2
IgH	immunoglobulin heavy chain
IgL	immunoglobulin light chain
IL-4	interleukin-4
IP	immunoprecipitation
IPTG	isopropyl β-D-1-thiogalactopyranoside
IRF	interferon regulatory factor
ITAM	immunoreceptor tyrosine-based activation motifs
kb	kilo base pairs
kDa	kilo Daltons
KO	knock-out
KSHV	Kaposi's sarcoma associated herpes virus
LCL	lymphoblastoid cell line
LID	lipofectin/integrin-targeting peptide/DNA

LMP	latent membrane protein
LP	leader protein
LZ	light zone
MACS	model-based analysis of ChIP-Seq
MAPK	mitogen-activated protein kinase
MHCII	major histocompatibility complex II
miR	microRNA
MMR	mismatch repair
mRNA	messenger RNA
NCoR	nuclear receptor co-repressor
ncPRC1	non-canonical PRC1 complex
NF-κB	nuclear factor κB
NICD	Notch intracellular domain
OriP	origin of replication
P-Rb	phosphorylated retinoblastoma protein
PAGE	polyacrylamid gel electrophoresis
PBMC	peripheral blood mononuclear cells
PBS	phosphate-buffered saline
PBS-T	phosphate-buffered saline containing 0.1% (v/v) Tween-20
PcG	polycomb group
PCGF	polycomb group factor
PCR	polymerase chain reaction
PI3-K	phosphatidylinositol 3-kinase
PMSF	phenylmethanesulfonyl fluoride
PRC	polycomb repressive complex
PTLD	post-transplant lymphoproliferative disease
PTM	post-translational modification
Qp	Q promoter
qPCR	quantitative PCR
RAG	recombination activating gene
RAM	RBPJ-associated molecule

Rb	retinoblastoma protein
RBPJ	recombination binding protein-J (aka CBF1, Su(H), and Lag 1)
Rev	revertant
RING1	ring finger protein 1
RIPA buffer	radioimmunoprecipitation assay buffer
RLU	relative luciferase units
RNA	ribonucleic acid
RPA	replication protein A
RPA70	replication protein A 70
RPMI	Roswell Park Memorial Institute medium
RT	reverse transcription
RT-qPCR	reverse transcription quantitative polymerase chain reaction
RUNX	RUNT-related transcription factor
RYBP	RING1/YY1 binding protein
S region	switch region
sBL	sporadic Burkitt's lymphoma
SD	standard deviation
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHARP	SMRT/HDAC1-associated repressor protein (aka MINT or SPEN)
SHM	somatic hypermutation
shRNA	short-hairpin RNA
SL chain	surrogate light chain
SMRT	silencing mediator of retinoid and thyroid hormone receptors
Spt5	suppressor of Ty5 homolog
SSB	single strand break
ssDNA	single stranded DNA
STAT6	signal transducers and activators of transcription 6
SUZ12	suppressor of zeste 12
TAD	topologically associating domains
TAF	TATA-binding protein associated factor

TAP	tandem affinity purification tag (2x Strep and 1x Flag)
TBE	Tris-Borate-EDTA
TBP	TATA box-binding protein
TCR	T cell receptor
TE	Tris-EDTA
TEMED	tetramethylethylenediamine
TET	ten eleven translocation
TF	transcription factor
T_{FH} cell	follicular T helper cell
TFII	transcription factor II
TGF-β	tumour growth factor β
Th cell	T helper cell
TP53	tumour protein 53
TPA	12-O-tetradecanoylphorbol-13-acetate
TR	terminal repeats
TSS	transcription start site
UNG	uracil-DNA-glycosylase
UTR	untranslated region
v/v	volume by volume
w/v	weight by volume
Wp	W promoter
wt	wild type
YAF2	YY1-asociated factor 2

1 Chapter I: Introduction

This study investigates the role of Epstein-Barr virus nuclear antigen 3C (EBNA3C) in the regulation of host genes expression. Epstein-Barr virus (EBV) biology is tightly associated with B cells and although EBV infects other cell types, expression of EBNA3C appears to be exclusive to EBV-infected B cells. Therefore, first of all, general B cell biology will be introduced followed by an introduction of EBV and mechanisms of gene regulation.

1.1 B cell biology

1.1.1 Immunoglobulins

B cells form the main part of the humoral immune response of the adaptive immune system. They are specialised cells that express immunoglobulins (Ig), special molecules able to bind to an almost unlimited number of antigens. Ig consist of two heavy chains (IgH) and two light chains (IgL) comprised of different combinations of numerous variable (V), diversity (D, IgH only) and joining (J) segments together with a constant (C) segment. The IgH chain consists of a N-terminal variable region (V_H, D and J_H) and a C-terminal constant (C_H) region (Figure 1.1A). The C_H segment determines the Ig class, C_μ (IgM), C_δ (IgD), C_γ (IgG), C_α (IgA) and C_ε (IgE), with each class having different effector fuctions. IgL chains also contain a N-terminal variable region (V_L and J_L), but no D segment and a shorter C-terminal constant region (Figure 1.1B). Furthermore, there are two different isoforms of IgL chains (κ and λ). The variable regions of the combined IgH and IgL chains are responsible for antigen binding and especially three complementarity determining regions (CDR 1-3) influence the antigen specificity. CDR1 and CDR2 are located within the V_H and V_L segments, whereas CDR3 is spanning the different $V_H(D)J_H$ and V_L-J_L segments [reviewed in (Hwang et al., 2015)]. There are two forms of Ig, transmembrane Ig (Figure 1.1C) form the B cell receptor (BCR) together with Ig-α (CD79a) and Ig-β (CD79b), which contain immunoreceptor tyrosine-based activation motifs (ITAMs) that can be phosphorylated by tyrosine protein kinases, and secreted Ig (Figure 1.1D), also known as antibody.

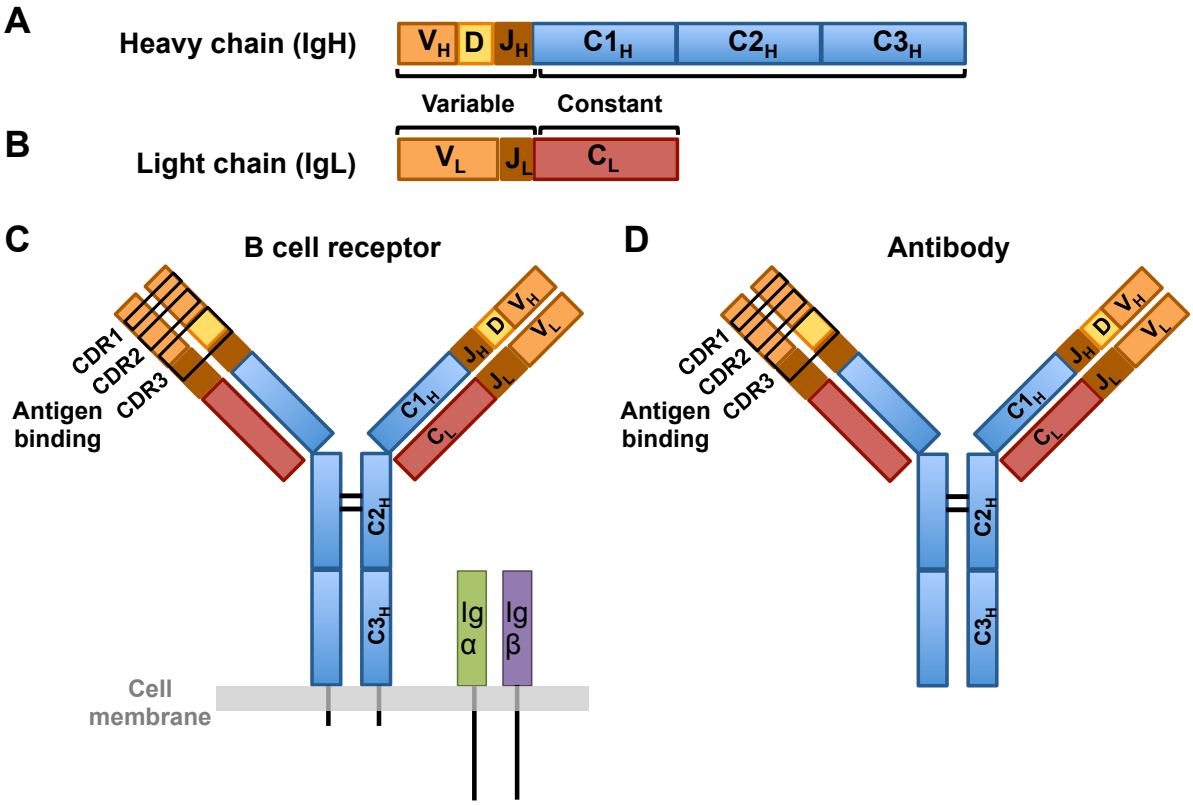


Figure 1.1: Schematic overview of immunoglobulin heavy (IgH) and light (IgL) chains as part of the B cell receptor on the cell surface or as secreted antibody.

(**A**) Schematic overview of the IgH chain consisting of variable (V_H), diversity (D), joining (J_H) and constant (C_H) segments. Different combinations of individual V_H, D and J_H segments form the variable N-terminal part and C_H1-3 the constant C-terminal part. (**B**) Schematic overview of the IgL chain consisting of variable (V_L), joining (J_L) and constant (C_L) segments. Different combinations of individual V_L and J_L segments form the variable N-terminal part and C_L the constant C-terminal part. (**C**) B cell receptors are formed by two IgL and two IgH chains anchored in the cell membrane and the co-receptor transmembrane proteins Ig-α and Ig-β. Disulphide bridges link the two IgH chains and antigen-binding specificity is determined predominantly by the complementarity determining regions (CDR) 1-3 of the variable region of the combined IgH and IgL chains. (**D**) As in C, but for secreted antibodies that do not contain a transmembrane domain of IgH.

1.1.2 B cell development

During adult life, B cells are continuously produced in the bone marrow where they originate from $CD34^+$ haematopoietic stem cells (HSC). HSCs give rise to multipotent myeloid/lymphoid progenitors that further differentiate into common lymphoid progenitors, the progenitors of both B and T cells (Figure 1.2A).

The B cell lineage begins in pro-B cells with the expression of enzymes of the recombination machinery, such as recombination activating gene 1 [RAG1, (Schatz et al., 1989)] and RAG2 (Oettinger et al., 1990), that catalyse the rearrangement of the IgH locus by first joining one D segment to one J_H segment (DJ_H rearrangement) followed by one V_H segment resulting in $V_H(D)J_H$ rearrangement (Figure 1.2B). In large pre-B cells, the rearranged IgH locus is expressed as C_μ isotype alongside a surrogate light (SL) chain (Sakaguchi and Melchers, 1986; Kudo and Melchers, 1987; Karasuyama et al., 1990; Tsubata and Reth, 1990) allowing the formation of pre-BCR on the cell surface in the case of productive IgH rearrangement (Figure 1.2A). Unproductive IgH rearrangement does not result in pre-BCR formation and additional rounds of $V_H(D)J_H$ recombination are performed until productive rearrangement is achieved from either of the two IgH alleles or no more rearrangements are possible, in which case the pro-B undergoes apoptosis. Productive rearrangement results in signalling through the pre-BCR [review on BCR signalling (Dal Porto et al., 2004)], resulting in down regulation of recombination enzymes to prevent continued $V_H(D)J_H$ rearrangement (Grawunder et al., 1995; Galler et al., 2004), allelic exclusion of the second IgH locus ensuring expression of a single IgH allele (Kitamura and Rajewsky, 1992) and proliferation (Karasuyama et al., 1994; Hess et al., 2001).

Pre-BCR signalling also down regulates expression of the surrogate light chain (Parker et al., 2005) resulting in a cessation of proliferation and differentiation into resting small pre-B cells. Small pre-B cells re-express RAG1 and RAG2 (Grawunder et al., 1995) that are now targeted to generate IgL V_L-J_L rearrangements [(Constantinescu and Schlissel, 1997), Figure 1.2C]. The newly rearranged IgL locus is expressed and tested for compatibility with the previously rearranged IgH chain to assemble as functional IgM on the surface of the immature B cell. In the case of incompatibility, additional rounds of IgL V_L-J_L rearrangements occur on either allele of the κ- or λ-L chain loci until productive rearrangement is achieved and the immature B cell exits the bone marrow and migrates to

the spleen to become a mature B cell expressing predominantly IgD and less IgM (Figure 1.2A).

A number of checkpoints exist to ensure that only non-self-reactive B cells end up as mature B cells [reviewed in (Melchers, 2015)]. An enormous pool of B cells each containing an individual B cell receptor able to recognise almost any possible antigen is created through the combinatorial joining of a large number of individual segments, *e.g.* for IgH alone there are at least 44 functional V_H genes, 23 functional D_H genes and 6 functional J_H genes (Watson and Breden, 2012), the junctional diversity through insertion of additional nucleotides in between the rearranged segments (Alt and Baltimore, 1982) together with the pairing of one light and one heavy chain.

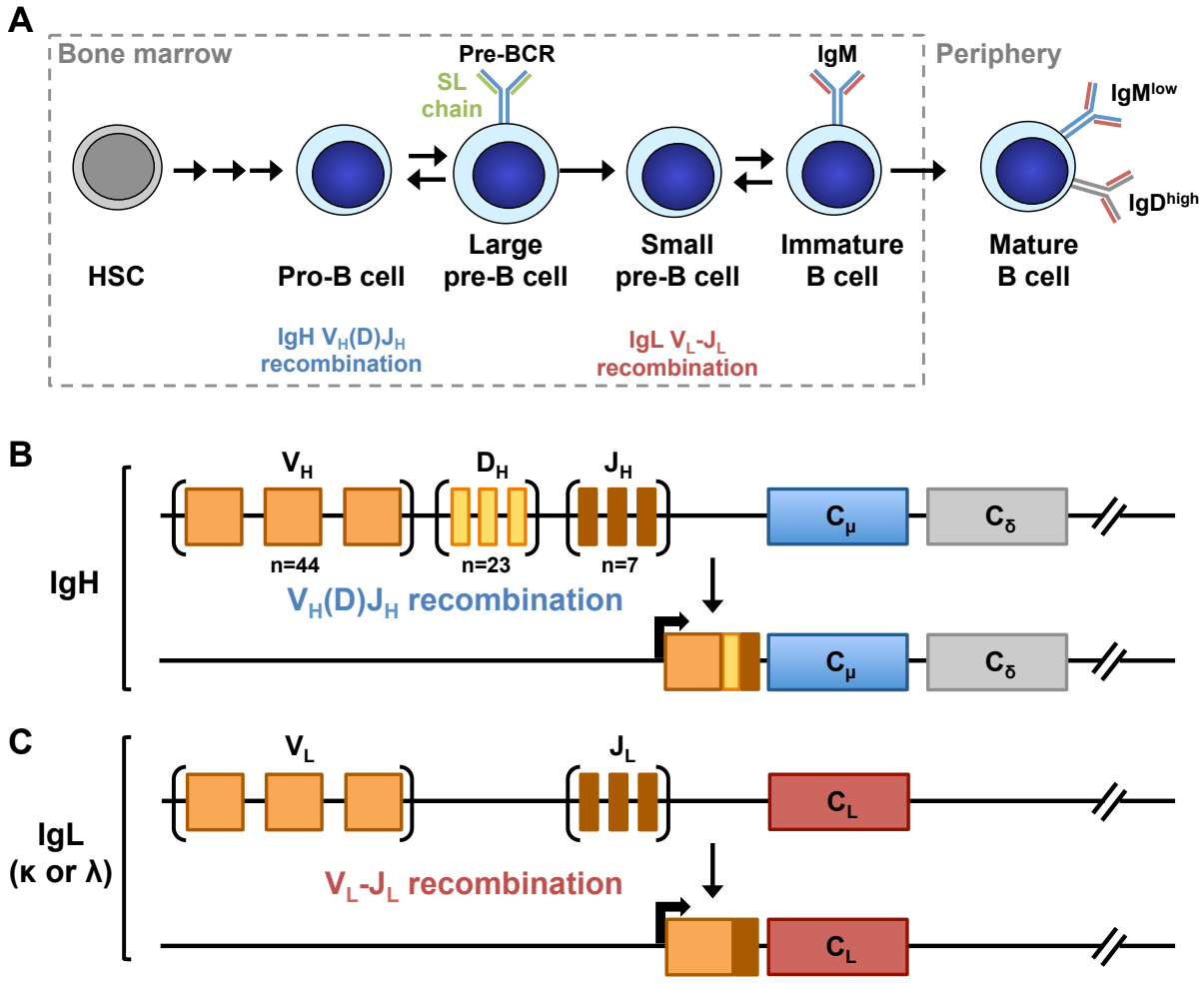


Figure 1.2: B cell development and overview of $V_H(D)J_H$ and V_L-J_L recombination.
(**A**) Overview of different B cell developmental stages within the bone marrow. Haematopoietic stem cells (HSCs) give rise to pro-B cells that initiate immunoglobulin heavy chain (IgH) $V_H(D)J_H$

recombination. Productive IgH rearrangement results in the expression of heavy chains of μ isotype that together with a surrogate light (SL) chain can form pre-B cell receptors (pre-BCR) on the surface of large pre-B cells. Pre-BCR signalling induces proliferation of large pre-B cells and differentiation into resting, small pre-B cells, which undergo C_L-J_L recombination of the light chain (IgL). Only compatible IgH and IgL are expressed as functional IgM on the surface of immature B cells that then exit the bone marrow and become mature B cells expressing predominantly IgD and less IgM. (**B**) Schematic overview of $V_H(D)J_H$ recombination at the IgH locus that joins one of 44 functional V_H segments, one of 23 functional D segments and one of 6 functional J_H segments. Productive recombination products can be expressed together with the constant region of IgM (C_μ) or IgD (C_δ). (**C**) Schematic overview of V_L-J_L recombination at one of the two IgL loci (Igκ or Igλ) that joins one V_L segment and one J_L segment. Productive recombination products can be expressed together with the constant region of the light chain (C_L).

1.1.3 T cell dependent B cell activation

Mature, but naïve B cells continuously circulate throughout the body and migrate from peripheral circulation into secondary lymphoid organs (*e.g.* the spleen or lymph nodes) that contain follicles rich in naïve B cells separated by an interfollicular region from the T cell zone. Naïve B cells scan the lymphoid tissue and the humoral immune response begins when a naïve B cell encounters a T cell dependent antigen that can be bound by its cell surface BCR. Antigen binding by the BCR results in BCR crosslinking, which induces signalling through the ITAM motifs of Ig-α and Ig-β leading to the internalisation of the BRC:antigen complex and fragmentation of the antigen into shorter peptides. These short peptides are then loaded on major histocompatibility complex (MHC) II and are presented on the B cell surface (Blum et al., 2013). B cell activation requires additional signals, which can be provided by a $CD4^+$ T helper (Th) cell that was activated by the same antigen and whose T cell receptor (TCR) recognises the MHCII:antigen complex on the B cell surface (Crotty, 2015). Recognition of MHCII:antigen complexes by TCRs of the activated Th cells induces the expression of the cell surface protein CD40 ligand [CD40L, (Armitage et al., 1992)] and secretion of cytokines, *e.g.* interleukin-4 [IL-4, (Howard et al., 1982)]. CD40L binds to CD40 on the surface of the B cell (Armitage et al., 1992) leading to B cell activation that is further modulated by the different secreted cytokines. The activated B blasts aggregate in primary follicles and start to express the transcriptional repressor B cell lymphoma 6 [BCL6, (Flenghi

et al., 1995; Cattoretti et al., 1995; Onizuka et al., 1995)], the master regulator essential for the formation of germinal centres (GC, see Section 1.1.4) within the centre of these follicles (Dent et al., 1997; Ye et al., 1997). Through repression of members of the DNA damage signalling pathway, *e.g.* tumour protein 53 [TP53, (Phan and Dalla-Favera, 2004)] and ataxia telangiectasia and Rad3 related [ATR, (Ranuncolo et al., 2007)], BCL6 ensures that cells are more resistant to DNA damage to which they are exposed during the GC reaction (see below). On the other side, BCL6 establishes a pro-apoptotic state by repressing the anti-apoptotic protein B cell lymphoma 2 [BCL2, (Saito et al., 2009)], promoting apoptosis (see below) unless positive signals are provided (Liu et al., 1991). Furthermore, BCL6 controls the end of the GC reaction through repression of transcription factors that control downstream differentiation programmes (Shaffer et al., 2000). Additional functions of BCL6 have been recently reviewed (Basso and Dalla-Favera, 2012) together with other factors involved in the formation of GC (De Silva and Klein, 2015; Basso and Dalla-Favera, 2015).

Not all antigen-activated B cells enter the GC reaction, a subset of them remain outside GCs and differentiate into short-lived plasmablasts that secrete low-affinity antibody (Jacob and Kelsoe, 1992).

1.1.4 Germinal centre reaction

GC are composed of a dark zone (DZ) that is densely populated by B cells, and a light zone (LZ) that is less densely populated by B cells and contains a network of follicular dendritic cells (FDC) and follicular Th (T_{FH}) cells (Figure 1.3). In the classic GC reaction model (MacLennan, 1994), larger and highly proliferative B cells within the DZ, named centroblasts, undergo clonal expansion and somatic hypermutation (SHM, see Section 1.1.5.2) of their Ig variable region. Centroblasts then differentiate into smaller, resting centrocytes and migrate into the LZ, where those mutations that increase the binding affinity to the antigen are selected through interactions with antigen-presenting follicular dendritic cells (FDC) and antigen-specific follicular T helper (T_{FH}). FDC provide long-term presentation of intact antigen on their cell surface (Mandel et al., 1980; 1981) enabling the centrocytes to

bind and process the antigen. Originally it was thought that positive selection is predominantly based on the competition between B cells for available antigen with high affinity B cells sequestering all available antigen (MacLennan, 1994). However, more recently it has been suggested that competition for T_{FH} help is the major factor for selection with high affinity B cells being able to present more processed antigen to a limiting amount of T_{FH} cells (Allen et al., 2007a). It was demonstrated that high-affinity B cells present more antigen:MHCII on their surface, which allowed prolonged contacts with a limiting number of T_{FH} cells (Victora et al., 2010) resulting in a refinement of the classic GC reaction model (Victora and Nussenzweig, 2012). Further to this, low affinity or self-reactive centrocytes (Melchers, 2015) who do not receive enough survival signals either undergo apoptosis, due to their default pro-apoptotic state (see above), or re-cycle back to the DZ to undergo additional rounds of SHM and selection for mutations that increase the antigen binding affinity. Upon emergence of Ig with high antigen binding affinity, B cells can undergo class switch recombination (CSR, see Section 1.1.5.2) to IgG, IgA or IgE depending on the type of antigen that triggered the immune response. However, CSR can also occur outside the GC (Stavnezer and Schrader, 2014). Repression of BCL6 expression leads to GC exit and differentiation to either antibody producing plasma cells [reviewed in (Nutt et al., 2015)] or long-lived CD27+ memory B cells [reviewed in (Kurosaki et al., 2015)]. Interestingly, recent studies also report memory B cells derived independently from a GC reaction, but the mechanisms behind this remain unclear (Kurosaki et al., 2015).

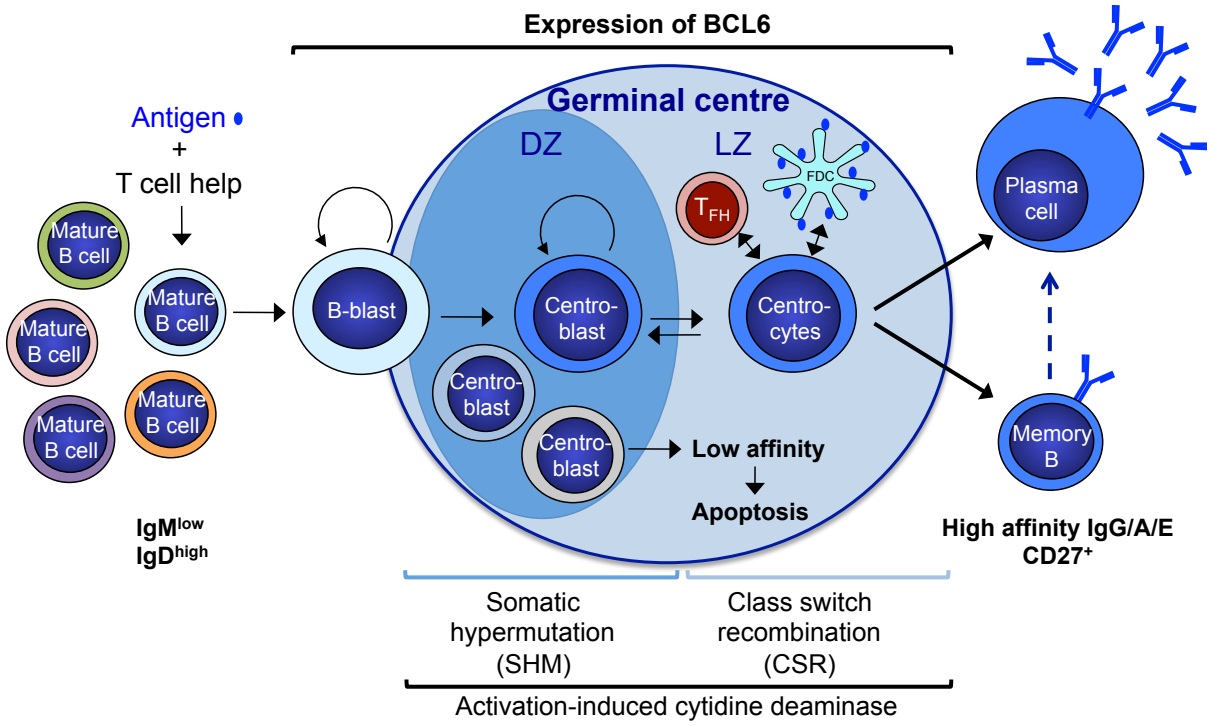


Figure 1.3: Germinal centre reaction after antigen-specific B cell activation.
Mature IgM^{low}/IgD^{high} B cells with distinct antigen binding specificity (represented by the different colours of the mature B cells) are screening secondary lymphoid tissue for their cognate antigen. Encounter between a mature B cell, its cognate antigen and an antigen-specific T cell results in B cell activation. The activated B blast proliferates and expresses BCL6 that initiates the formation of germinal centres (GC). GCs form in the centre of primary follicle and are divided into a dark zone (DZ) and a light zone (LZ). GC B cells express activation-induced cytidine deaminase (AID). Within the DZ, proliferating centroblasts undergo somatic hypermtation (SHM) of their immunoglobulins. Mutations that increase the binding affinity to the antigen are selected for in the LZ through interactions with follicular T helper (T_{FH}) and follicular dendritic cells (FDC). Centrocytes re-enter the DZ where they undergo additional rounds of SHM. Upon emergence of high affinity antigen binding, B cells undergo class switch recombination (CSR) to IgG, IgA or IgE and exit the GC to become either antibody producing plasma cells or long-lived memory B cells. Figure adapted from Prof Martin Allday.

Based on morphological criteria [reviewed in (Nieuwenhuis and Opstelten, 1984)], the larger, proliferating B cells in the DZ were described as centroblasts and the smaller, resting B cells in the LZ as centrocytes. However, more recently studies employing two-photon intravital microscopy allowing direct *in vivo* GC imaging revealed that B cells in the DZ and LZ had similar size and morphology (Schwickert et al., 2007; Allen et al., 2007b; Hauser et al., 2007). The discovery that chemokine receptor CXCR4 (Allen et al., 2004) and the activation molecules CD83 and CD68 (Victora et al., 2010) can be used as markers to

distinguish between centroblasts ($CXCR4^{high}CD83^{low}CD86^{low}$) and centrocytes ($CXCR4^{low}CD83^{high}CD86^{high}$), both in mice and humans (Caron et al., 2009; Victora et al., 2012), allowed further investigation of the differences between DZ and LZ GC B cells. Again, no difference in size could be detected and differences in gene expression between the two populations were limited (Victora et al., 2010; 2012). However, higher expression of genes involved in the regulation of the cell cycle, particularly mitosis, and DNA damage was detected in centroblasts, whereas centrocytes expressed higher levels of genes involved in cell activation (*e.g.* CD40, CD83, CD86), transcription factors (*e.g.* c-MYC) and different regulators of apoptosis (Victora et al., 2010; 2012).

C-MYC can be found in almost all proliferating cells as it supports robust proliferation through transcriptional and non-transcriptional regulation of cellular processes, such as cell cycle progression and metabolism. However, C-MYC was considered to be absent in proliferating GC B cells (Klein et al., 2003). Recently, it has been shown that c-MYC is essential in the early formation of the GC, after which it is repressed by BCL6 in the highly proliferative DZ B cells, before being re-expressed in a subset of LZ B cells that are selected for re-entry into the DZ to undergo further rounds of proliferation and SHM (Dominguez-Sola et al., 2012; Calado et al., 2012; Victora et al., 2010; 2012).

1.1.5 Activation-induced cytidine deaminase

The search for factors involved in SHM and CSR identified activation-induced cytidine deaminase (AID), which is encoded by the gene *AICDA*. AID is a member of the apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like (APOBEC) family of cytidine deaminases. Expression of AID is low in naïve B cells, but increases highly in GC B cells and B cells *in vitro* activated with IL-4, CD40L and tumour growth factor β (TGF-β) (Muramatsu et al., 1999). AID has been shown to be essential for both SHM and CSR. *Aicda* knock-out mice failed to undergo SHM and CSR (Muramatsu et al., 2000) and mutations in *AICDA* were detected in humans patients with hyper-IgM syndrome, characterised by high serum levels of low affinity IgM due to defective SHM and CSR (Revy et al., 2000). In order

to understand the role of AID in SHM and CSR, first the mechanisms that underlie AID-induced mutations have to be explained.

1.1.5.1 Mechanisms underlying AID-induced mutations

AID requires access to single-stranded DNA (ssDNA), *e.g.* at regions transcribed by RNA polymerase II (Chaudhuri et al., 2003), to catalyse the deamination of cytidines (C) to uridines (U) producing uracil/guanine (U/G) mismatches (Figure 1.4A). Normally, U/G mismatches are detected and repaired by error-free DNA repair mechanisms, but when DNA replication over the U/G site precedes its repair, it results in a C>T and G>A transition [purine > purine (A and G) or pyrimidine > pyrimidine (C and T)] mutation (Figure 1.4B).

U/G mismatches can be detected by the base excision repair [BER, (Krokan and Bjørås, 2013)]. This leads to the excision of uracil from the DNA backbone by uracil-DNA-glycosylase (UNG) creating an abasic site (Figure 1.4C). Replication over this abasic site results in both transition and transversion (purine > pyrimidine or pyrimidine > purine) mutations, which are however restricted to the original U/G lesion. UNG deficient mice exhibit a near complete shift towards transition mutations at the U/G site (Rada et al., 2002a). During the normal BER, apurinic/apyridimic endonuclease 1 (APE1) cleaves the DNA backbone at the excised uracil creating a single strand break (SSB), which can lead to a double strand break (DSB) if another SSB is created on the opposite strand (Figure 1.4C).

Beside the BER, the U/G mismatch can be detected by the MSH2-MSH6 heterodimer (Rada et al., 1998; Wiesendanger et al., 2000) of the mismatch repair [MMR, (Fishel, 2015)]. MSH2-MSH6 dimers then recruits exonuclease 1 [Exo1, (Bardwell et al., 2004)], which excises a larger piece of DNA of the mismatch-containing strand producing a SSB (Figure 1.4D). Error-prone polymerase η can fill in the missing nucleotides resulting in a spread of mutations from the original U/G lesion (Zeng et al., 2001; Delbos et al., 2007). Again, the presence of another SSB on the opposite strand results in a DSB (Figure 1.4D). Deficiencies in the MMR pathway, *e.g.* in MSH2 knock-out mice (Rada et al., 1998), results in a reduction in the overall number of mutations and restriction of mutations to the G/C site. Deficiency in

both BER and MMR pathways in UNG and MSH2 double knock-out mice, produces exclusively C/G transition mutations via replication over the U/G lesion and eliminates C/G transversions and mutational spreading (Rada et al., 2004).

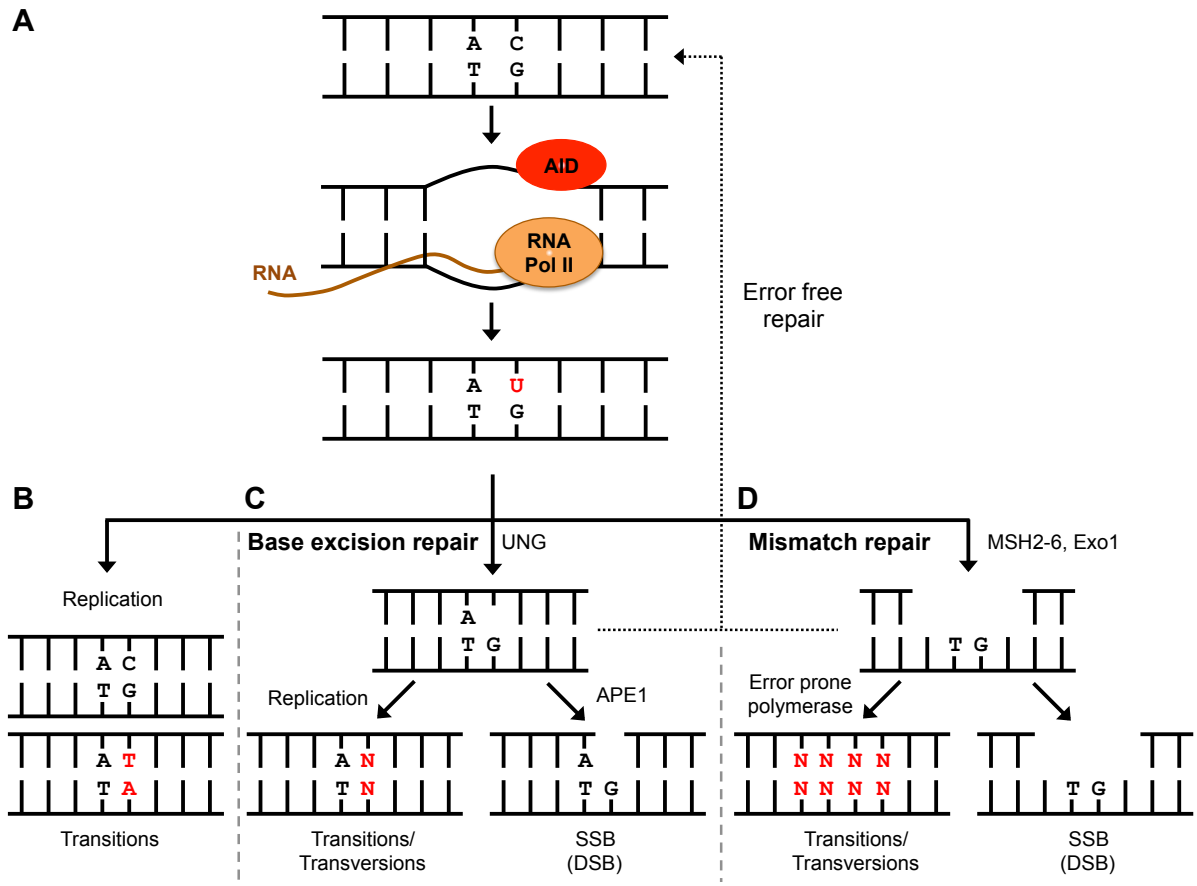


Figure 1.4: Mutations caused by activation-induced cytidine deaminase (AID).
(**A**) AID requires access to single stranded DNA (*e.g.* at sites of active transcription by RNA polymerase II) to deaminate cytidine (C) residues resulting in an uracil/guanine (U/G) mismatch. (**B**) DNA replication of the U/G mismatch leads to one strand with the original C/G and one strand with a T/A mutation (C->T and G->A transitions). (**C**) The U/G mismatch is detected by the base excision repair pathway. Uracil-DNA-glycosylase (UNG) creates an abasic site through excision of uracil. Replication over this abasic site results in transition or transversion mutations limited to the U/G lesion. Recruitment of apurinic/apyridimic endonuclease 1 (APE1) creates a DNA single strand break (SSB). The presence of another SSB on the opposite strand can result in a DNA double strand break (DSB). (**D**) The heterodimer MSH2-MSH6 of the mismatch repair recognises the U/G mismatch and recruits Exonuclease 1 (Exo1) that excises a larger piece of DNA on the mismatch-containing strand creating a SSB. Error-prone polymerases lead to a spread of mutations from the U/G mismatch to neighbouring sites (*e.g.* A/T). Again, two SSB on opposite strands can create a DSB. Figure adapted from (Hwang et al., 2015).

1.1.5.2 Somatic hypermutation (SHM) and class switch recombination (CSR)

AID-induced mutations are essential for both SHM and CSR (Muramatsu et al., 2000; Revy et al., 2000) and occur in GC B cells that express high levels of AID (Muramatsu et al., 1999), with centroblasts expressing higher levels of AID compared to centrocytes (Victora et al., 2012). Replication over AID induced U/G lesions is an important contributing factor for AID-induced mutations (see above) and centroblasts are rapidly proliferating, so mutation rates of up to 10^{-3} per base per generation have been estimated (McKean et al., 1984; Radic, 2008).

For the process of SHM (Figure 1.5A), AID induces predominantly point mutations in the transcribed variable regions [$V_H(D)J_H$ and V_LJ_L] of the recombined IgH and IgL chains. Transcription is required for SHM (Fukita et al., 1998) and mutations occur on both the productive and non-productive IgH and IgL loci, but only mutations on the productive loci affect the positive selection. Mutations accumulate sequentially and particularly within the three CDRs, which mostly affect the antigen binding affinity [(McKean et al., 1984) and reviewed in (Di Noia and Neuberger, 2007; Hwang et al., 2015)]. AID preferentially targets cytidines (highlighted in bold) within WR**C**H or related sequence motifs [(Rogozin and Diaz, 2004) with W = A or T, R = A or G, H = T, C or A]. These WR**C**H sequences are slightly enriched within the variable region of IgH and IgL (Hwang et al., 2015). The mechanisms that target AID, particularly to CDRs, still remain unknown. A recent study employing a $V_H(D)J_H$ passenger allele system, in which SHM can be followed *in vivo* on a passenger IgH locus that is not subjected to antigen binding selection, has shown that not all WR**C**H motifs within the variable region are targeted at the same rate (Yeap et al., 2015). SHM is not a unique feature of Ig sequences and similar frequencies of SHM were observed upon replacement by non-Ig genes, in which WR**C**H motifs were also differentially targeted depending on their position within the sequence. Frequent DSB that caused deletions were detected on the passenger allele, most of which would have probably been eliminated during

the selection, as predominantly point mutations are observed in SHM on the productive IgH alleles (Di Noia and Neuberger, 2007).

For the process of CSR (Figure 1.5B), AID deaminates cytosine residues within switch (S) regions that are long (1 to 10 kb), highly repetitive, noncoding regions upstream of the C_H exons of each Ig isotype class (except IgD which is positioned immediately downstream of IgM). Importantly, each Ig isotype class (except IgD) is part of a separate transcription unit that is under the control of specific cytokine responsive promoters controlling transcription through the corresponding S region and C_H exons (Stavnezer and Schrader, 2014; Hwang et al., 2015). Promoter-specific transcription is induced depending on the cytokines that are secreted by T_{FH} cells or FDCs with the aim to direct CSR to the most effective Ig class to eliminate the antigen. S regions are highly enriched in AGCT, a canonical WR**C**H motifs that provides cytidines on both opposite strands, which is thought to promote a higher frequency of DSB at the two S regions that are selected to undergo CSR (Hwang et al., 2015). The formation of DSB is essential for CSR with deficiencies in either BER [*e.g.* UNG knock-out (Rada et al., 2002a)] or MMR [*e.g.* MSH2 knock-out (Rada et al., 1998)] reducing the rate of CSR and deficiency in both pathways [*e.g.* UNG and MSH2 knock-out (Rada et al., 2004)] completely abrogating it. Deletional DNA recombination between the DSB at the upstream and the downstream S regions causes the excision of the intermediate region (known as switch circle) and positioning of the new Ig isotype class C_H exons immediately downstream of the $V_H(J)D_H$ region (Stavnezer and Schrader, 2014). This results in the expression of another Ig class with a more appropriate effector function, while retaining the high-affinity antigen binding specificity of the $V_H(D)J_H$ region.

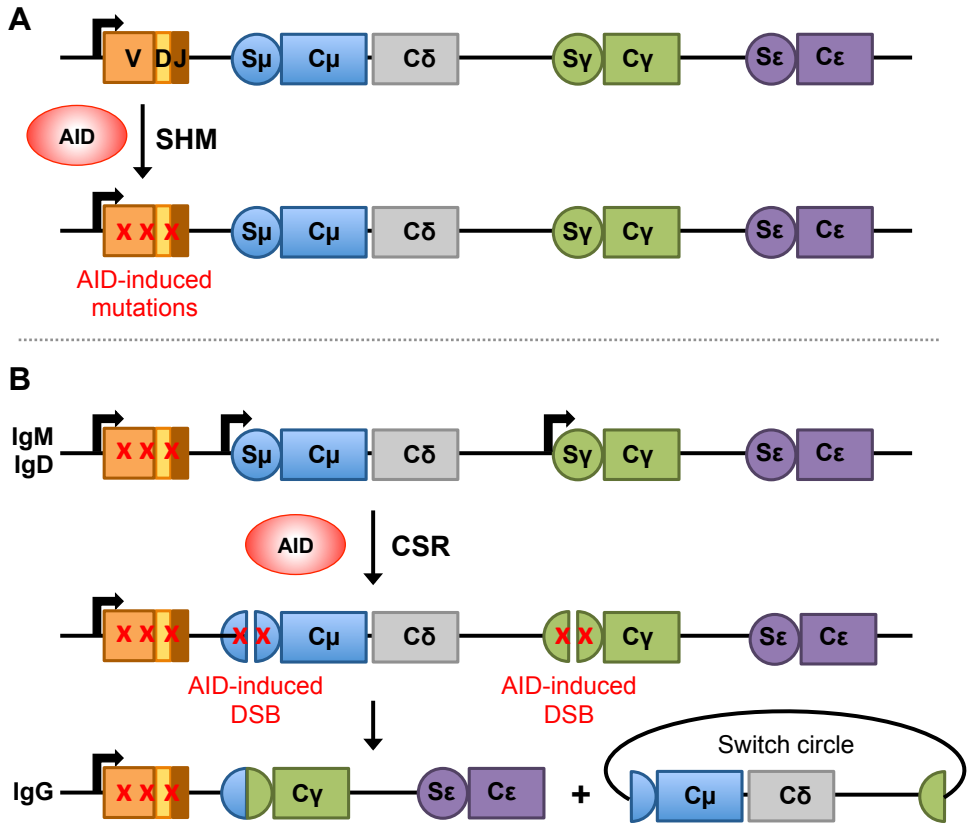


Figure 1.5: Somatic hypermutation (SHM) and class switch recombinantion (CSR) induced by activation-induced cytidine deaminase (AID).

(**A**) AID requires transcription of the immunoglobulin locus (black arrow) to induces mutations within the variable region [schematically shown for IgH $V_H(D)J_H$, but similarly for V_LJ_L] in a process known as SHM. Some of these mutations increase the antigen binding affinity of the immunoglobulin and are positively selected for during the germinal centre (GC) reaction. (**B**) CSR is induced when exogenous signals received by the GC B cell lead to the transcription (black arrows) of two different immunoglobulin classes [shown for IgM (Cμ) and IgG (Cγ)]. AID-induced mutations cause DNA double strand breaks (DSB) in the switch regions (Sμ and Sγ) of the different antibody classes resulting in IgM to IgG class switching through recombination and excision of the intermediate region known as switch circle.

1.1.5.3 Regulation of AID activity

AID has the ability to induce mutations and DSB so it is not surprising that there are several layers of regulation that control the activity of AID (Figure 1.6). High level of AID expression is normally limited to GC B cells (Muramatsu et al., 1999).

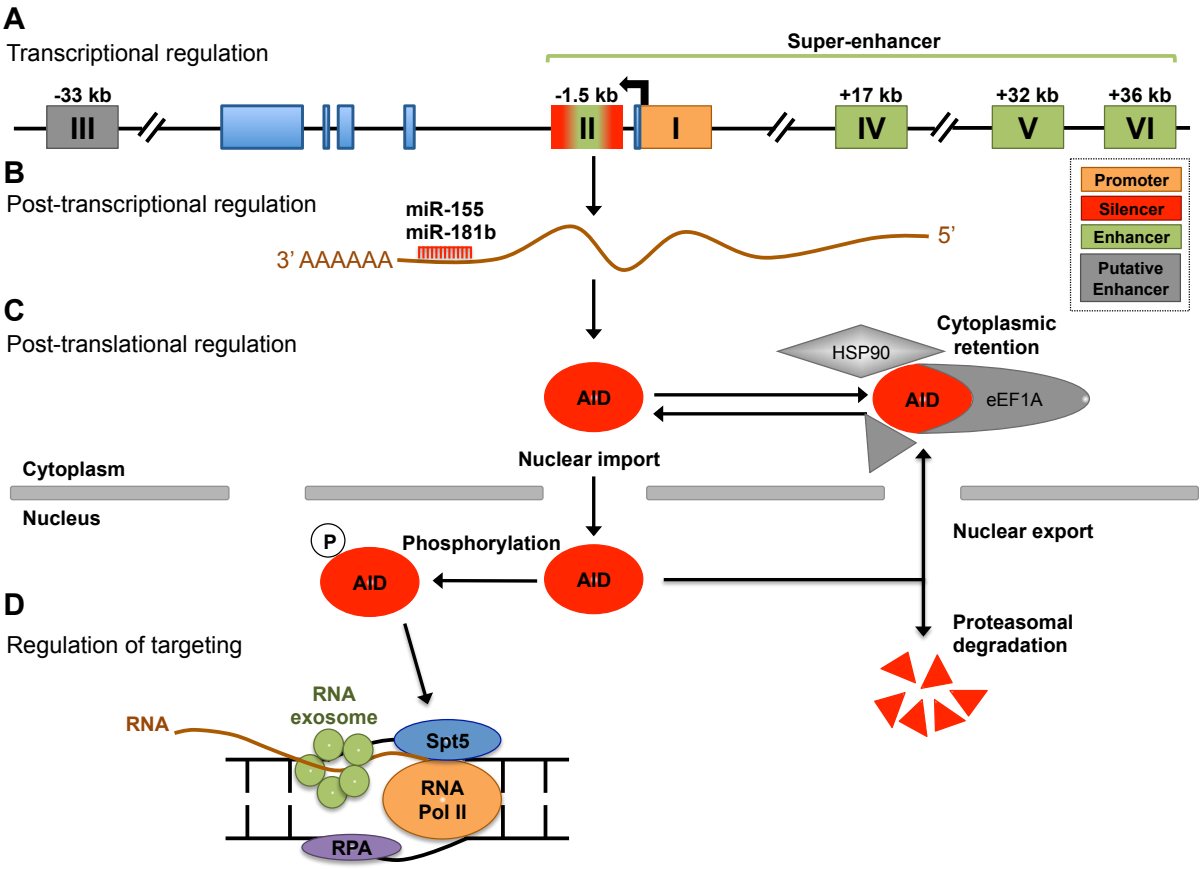


Figure 1.6: Activity of AID is controlled on multiple layers.

(**A**) Regulation of AID transcription (black arrow) is controlled by at least six regulatory regions shown schematically for the human *AICDA* locus (blue boxes represent exons). Region I is a ubiquitously active basal promoter, region II contains both enhancer and silencer elements, region III is a putative enhancer and region IV, V and VI are upstream enhancer elements. (**B**) Regulation of AID at post-transcriptional level through binding of microRNA (*e.g.* miR-155 and miR-181b) to the 3' untranslated region of *AICDA* mRNA leading to mRNA destabilisation and degradation. (**C**) Regulation of AID at post-translational level. Most AID protein is being retained in the cytoplasm by elongation factor eEF1A and heat-shock protein 90 kDa (HSP90) and requires active nuclear import to translocate into the nucleus. Nuclear AID has a short half-life time due to proteosomal degradation and active nuclear export. Phosphorylation on various sites controls the activity of AID. (**D**) AID requires access to single stranded DNA (ssDNA) and often targets transcribed regions that are associated with stalled (Spt5) RNA polymerase II. The ssDNA binding protein (RPA) stabilises ssDNA and the RNA exosome was shown to enable access of AID to both strands.

1.1.5.3.1 Transcriptional regulation of AID

The first layer of control regulates the transcription rate of the *AICDA* gene encoding AID (Figure 1.6A). Comparison between the mouse *Aicda* and the human *AICDA* locus located on chromosomes 6 and 12, respectively, revealed several regions with high DNA sequence conservation in and around the *Aicda/AICDA* loci (Yadav et al., 2006; Crouch et al., 2007; Tran et al., 2010b).

Region I is located immediately upstream of the transcription start site (TSS) with transcription (indicated by the black arrow in Figure 1.6A) from the negative DNA strand running from right to left for human *AICDA*. It contains a minimal promoter and conserved binding sites for various transcription factors (TF) (Yadav et al., 2006). Binding of signal transducers and activators of transcription (STAT) STAT6 and nuclear factor κB (NF-κB) at region I was observed after IL-4 and CD40L-mediated induction of AID in B cell lines *ex vivo* and deficiency of either STAT6 or NF-κB impaired AID expression upon stimulation (Dedeoglu et al., 2004). Homeobox C4 (HoxC4), another TF expressed in GC B cells and up regulated by NF-κB, binds to region I and induces AID expression (Park et al., 2009; 2013). More recent analysis of region I confirmed the basal promoter activity in transient reporter assays, but no increase in luciferase activity was observed upon stimulation with IL-4, CD40L and TGF-β (Tran et al., 2010b).

Region II is located within the first intron just downstream of the TSS and has been shown to contain both stimulatory (enhancer) and repressive (silencer) elements. B cell specific E-proteins, which form homo- or heterodimers to become active TFs, bind to E-box DNA binding sites at region II and enhance the expression of *Aicda* (Sayegh et al., 2003). Pax5, a B cell specific TF, has also been shown to bind to region I and region II in activated B cells (Gonda et al., 2003), however in later studies no binding at region I was observed (Yadav et al., 2006). Overexpression of proteins of the inhibitor of differentiation (Id) family, which can disrupt DNA binding of E-proteins and Pax5, reduced *Aicda* expression (Sayegh et al., 2003; Gonda et al., 2003). The enhancer function of the Pax5 and E-protein DNA

binding sites was confirmed in more extensive analysis using transient reporter assays, but no increase in luciferase activity was observed upon stimulation with IL-4, CD40L and TGF-β (Tran et al., 2010b). In contrast, this study also showed that full-length region II acts as an overall repressor of luciferase activity in transient reporter assays, primarily caused by c-Myb and E2F DNA binding sites (Tran et al., 2010b). Because c-Myb and E2F are expressed in naïve and non-B cells, they are thought to prevent expression of AID in those cells, which is then overcome by Pax5 and E-proteins in activated B cells. Deletion of the two c-Myb and E2F DNA binding sites in transgenic mice, using bacterial artificial chromosomes (BAC) of the *Aicda* locus, increased the number of GC B cells with active *Aicda* expression, but did not result in ectopic expression of *Aicda* (Huong et al., 2013). Importantly, deletion of the entire region II completely abrogated *Aicda* transcription *in vivo*.

Region III is located ~33 kb downstream of the TSS of *AICDA*. BAC transgenic *Aicda*-GFP mice that lack region III did not express *Aicda* (Crouch et al., 2007), which indicates that region III acts as an important enhancer element. In transient reporter assays, however, region III did not alter luciferase expression in transient reporter assays *in vitro* (Tran et al., 2010b), making it a putative enhancer element.

Region IV is located ~17 kb upstream of the TSS of *AICDA* and contains binding sites for STAT6, NF-κB, Smad3/4 and enhancer-binding protein C/EBP. This region increased the basal luciferase activity of region I in transient reporter assays and further enhanced luciferase activity upon stimulation with IL-4, CD40L or TGF-β either alone or in combination resulting in an additive increase (Tran et al., 2010b). Furthermore, Tran and colleagues suggested that IL-4 promotes signalling through STAT6, CD40L through NF-κB, TGF-β through Smad3/4 and C/EBP. Deletion of the C/EBP sites completely abrogated the stimulation by IL-4, CD40L and TGF-β, but presence of the C/EBP sites alone was not sufficient for stimulation. BAC transgenic mice carrying deletions of either the entire region IV or just the C/EBP binding sites within region IV failed to express the *Aicda* reporter (Huong et al., 2013).

Region V and VI are located ~32 and 36 kb, respectively, upstream of the TSS of *AICDA* and have been identified as enhancers of *Aicda* expression through targeted deletion using genome editing techniques (Kieffer-Kwon et al., 2013). Deletion of either region V or region VI completely abrogated *Aicda* expression and IgM -> IgA CSR in activated murine CH12F3 B cells.

Regions II, IV, V and VI form what is known as the AID super-enhancer (Qian et al., 2014). Super-enhancers are formed by clusters of multiple, topologically associated enhancers that can potentiate gene expression and are thought to be co-ordinately regulated during cell differentiation (Figure 1.6A).

1.1.5.3.2 Post-transcriptional regulation of AID

Protein expression can also be modulated at a post-transcriptional level through the binding of microRNA (miR) to a target mRNA sequence resulting in blockage of translation and/or degradation of the target mRNA. Several miRs have been identified that bind in the 3' untranslated region (3' UTR) of *AICDA* mRNA (Figure 1.6B). miR-155 is B lymphocyte specific and is induced in murine B cells undergoing CSR (Teng et al., 2008). Binding of miR-155 in the 3'UTR destabilises *Aicda* mRNA and deletion of the miR-155 binding site not only increased AID expression (Dorsett et al., 2008), but also resulted in expression of AID in post-GC B cells in blood (Teng et al., 2008). In GC B cells, BCL6 inhibits the expression of miR-155 and therefore positively regulates AID expression (Basso et al., 2012). This study also identified miR-361, another BCL6-repressed miR, that is able to down regulate AID protein levels. However, the effect was weaker compared to miR-155 and the role of miR-361 on AID mRNA stability and CSR remains to be determined. Another miR, miR-181b, also binds in the 3'UTR of *Aicda* resulting in the down regulation of AID at mRNA and protein levels and to impair CSR (de Yebenes et al., 2008).

1.1.5.3.3 Compartmentalisation and post-translational regulation of AID

Once AID protein has been translated, there are different ways to control its activity (Figure 1.6C). Most of AID protein localises to the cytoplasm (Rada et al., 2002b; Pasqualucci et al., 2004), probably to restrict its access to DNA. The C-terminus of AID is important for its cytoplasmic retention (Patenaude et al., 2009). Recently the eukaryotic elongation factor 1 (eEF1A) and heat-shock protein 90 kDa (HSP90) were proposed to be responsible for cytoplasmic retention of AID by forming distinct cytoplasmic complexes with AID (Methot et al., 2015). AID localisation is further controlled by nuclear import and nuclear export. N-terminal nuclear localisation and C-terminal nuclear export signals have been identified (Ito et al., 2004) and despite its small molecular weight of 24 kDa active nuclear import is required for AID to enter the nucleus (Patenaude et al., 2009). A short life-span for nuclear AID has been reported involving both ubiquitination-dependent (Aoufouchi et al., 2008) and ubiquitination-independent (Uchimura et al., 2011) proteosomal mechanisms. Furthermore, AID can be phosphorylated at various sites, which can have an inhibitory [*e.g.* phosphorylation on serine 3 (Gazumyan et al., 2011)] or stimulatory [*e.g.* phosphorylation on serine 38, threonine 27 or 140 (Basu et al., 2005; Pasqualucci et al., 2006; McBride et al., 2008)] effect on its activity. For more information about this layer of control please see reviews on compartmentalisation (Patenaude and Di Noia, 2010; Orthwein and Di Noia, 2012) and phosphorylation (Basu et al., 2009).

1.1.5.3.4 Targeting of AID

Inside the nucleus, access of AID to ssDNA involves multiple factors (Figure 1.6D). Interaction between AID and RNA polymerase II was suggested (Nambu et al., 2003), which was shown to be mediated by Suppressor of Ty5 homolog (Spt5), a transcription elongation factor associated with stalled RNA polymerase II (Pavri et al., 2010). Spt5 interacts with AID and depletion of Spt5 reduced the association of AID with RNA polymerase II and the occupancy of AID at the IgH locus reducing the rate of SHM and CSR. Moreover, sites with high Spt5 occupancy displayed increased number of somatic mutations, suggesting that

Spt5 occupancy is predictive for AID-dependent mutations. In addition, AID can interact with the ssDNA binding protein replication protein A (RPA), which promoted deamination of *in vitro* transcribed SHM substrates, but only on the non-template strand, suggesting that binding of RPA to the single stranded non-template strand promotes AID activity (Chaudhuri et al., 2004). In a more recent study, occupancy of AID was associated with promoter-proximal sequences with stalled RNA polymerase II genome-wide, but high levels of RPA were limited to Ig genes (Yamane et al., 2011). This might promote a higher frequency of AID-induced mutations for SHM and CSR, but might be solely a result of introduced DSB, whose single strand ends are bound by RPA (Yamane et al., 2013). *In vitro* DNA deamination was always limited to the non-transcribed strand, probably because the transcribed strand is blocked by RNA:DNA hybridisation of the nascent RNA. The RNA exosome has been shown to enable AID to deaminate the transcribed template strand *in vitro* and to be required for CSR *in vivo* (Basu et al., 2011; Pefanis et al., 2014). The RNA exosome is an exonuclease complex that processes nuclear RNA precursors and can degrade RNA, giving AID access to the single-stranded template strand. Furthermore, AID has been reported to predominantly target highly transcribed super-enhancers where sense and antisense transcription converge, enabling access to both DNA strands (Qian et al., 2014; Meng et al., 2014).

All these mechanisms help in targeting AID to the Ig loci where highest AID activity is observed in activated B cells, however, AID activity is promiscuous and can induce mutations at non-Ig loci (off-target activity). These AID off-targets, *e.g. BCL6* and *cMYC*, usually share key characteristics with the Ig loci, *e.g.* being highly transcribed and controlled by super-enhancers, thus enabling AID to access DNA and to introduce mutations [recently reviewed in (Casellas et al., 2016)].

1.2 Epstein-Barr virus

1.2.1 Overview

Epstein-Barr virus (EBV) is a gamma-herpes virus belonging to the *herpesviridae* family of large dsDNA viruses that establish latent infections in their host. EBV persistently infects over 90% of the human population and usually infections occur early in childhood when they remain mostly asymptomatic. Infections that are delayed until early adolescence or later can cause infectious mononucleosis. Once an individual is infected with EBV, the virus establishes a persistent infection for the lifetime of the host. EBV is B lymphotropic and establishes latent infection within the memory B cell population of the host by closely following the normal biology of mature B lymphocytes (see Section 1.1.4) in what is known as the germinal centre model of persistence (see *in vivo* infection below). At this stage, the virus remains quiescent in resting memory B cells, invisible to the immune system. Latent infection is mostly benign, especially as >90% of the human population is infected, however EBV was the first identified human tumour virus with links to various B cell and non-B cell malignancies (see Section 1.2.5).

1.2.2 Establishment of latent infection

1.2.2.1 In vivo infection

The germinal centre model of persistence (Figure 1.7), proposed by David Thorley-Lawson and colleagues (Thorley-Lawson and Babcock, 1999; Thorley-Lawson, 2015), explains how EBV establishes a latent infection *in vivo* and is now generally accepted as it describes very accurately most aspects of *in vivo* EBV studies and can even be used to mathematically model persistence (Hawkins et al., 2013). EBV is shed in the saliva of an infected person and can infect epithelial cells in a newly infected individual probably to amplify the amount of infectious viral particles.

However, in order to achieve a latent infection, EBV infects mature, resting B cells in the lymphoepithelium of the tonsils. EBV binds to CD21 (Fingeroth et al., 1984; Nemerow et

al., 1985) and MHC II molecules (Li et al., 1997) to enter the cell and cross-linking of CD21 induces signalling events that disrupt the quiescent state of the cell (Sinclair and Farrell, 1995). Newly infected mature B cells express all latent viral gene products (Joseph et al., 2000b; Babcock et al., 2000), six EBV nuclear antigens (EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C and EBNA-LP), three latent membrane proteins (LMP1, LMP2A and LMP2B), two small, non-polyadenylated RNAs (EBER1 and EBER2) and several viral miRs. This viral gene expression programme is known as latency III or the growth programme, as it results in the activation and rapid proliferation of the newly infected B cell. The EBV-activated B cells resemble antigen-activated B blasts (Thorley-Lawson et al., 1982; 1985), but this is solely induced by the viral gene expression (Kempkes et al., 1995) and independent of antigen or T cell help required for normal B cell activation (see Section 1.1.3).

Expression of EBNA2 is not compatible with the GC B cell phenotype (see Section 1.2.4.1.2 on EBNA2). In order to access GCs and to express GC markers (*e.g.* BCL6 and AID), EBV-infected B cells switch their viral gene expression to latency II (Babcock et al., 2000), a more restricted expression programme encoding only four of the nine latent proteins (EBNA1, LMP1, LMP2A/B) together with the EBV-encoded RNAs. Expression of LMP1 and LMP2A/B is thought to provide survival signals in the otherwise pro-apoptotic GC B cell environment (see Sections 1.2.4.2.1 and 1.2.4.2.2 below).

It is thought that the EBV-infected GC B cell then follows the normal differentiation pathway to become a memory B cell, which again requires a switch in the viral gene expression programme to latency 0, which is a complete shut-down of viral protein expression, except expression of EBNA1 (latency I) upon the occasional proliferation of the memory B cell (Hochberg et al., 2004). EBV-infected memory B cells have undergone CSR (Babcock et al., 1998; Joseph et al., 2000a) and also SHM with a pattern similar to antigen-selected memory B cells (Souza et al., 2005). Furthermore, it seems that EBV positive memory B cells are more likely to have undergone CSR and to carry more somatic

mutations compared to EBV negative memory B cells (Souza et al., 2007). It is within the memory B cell compartment that EBV can persist for the lifetime of the host due to the long-life of memory B cells and also because EBV does not need to express any viral proteins and therefore remains undetected by the immune system.

In order to maintain a stable population of infected memory B cells and to spread to a new host, periodic reactivation of the lytic cycle occurs when the memory B cell differentiates into a plasma cell (Laichalk and Thorley-Lawson, 2005). The plasma cell produces new EBV viral particles that can infect mature B cells within the same host or infect a new individual through shedding in the saliva.

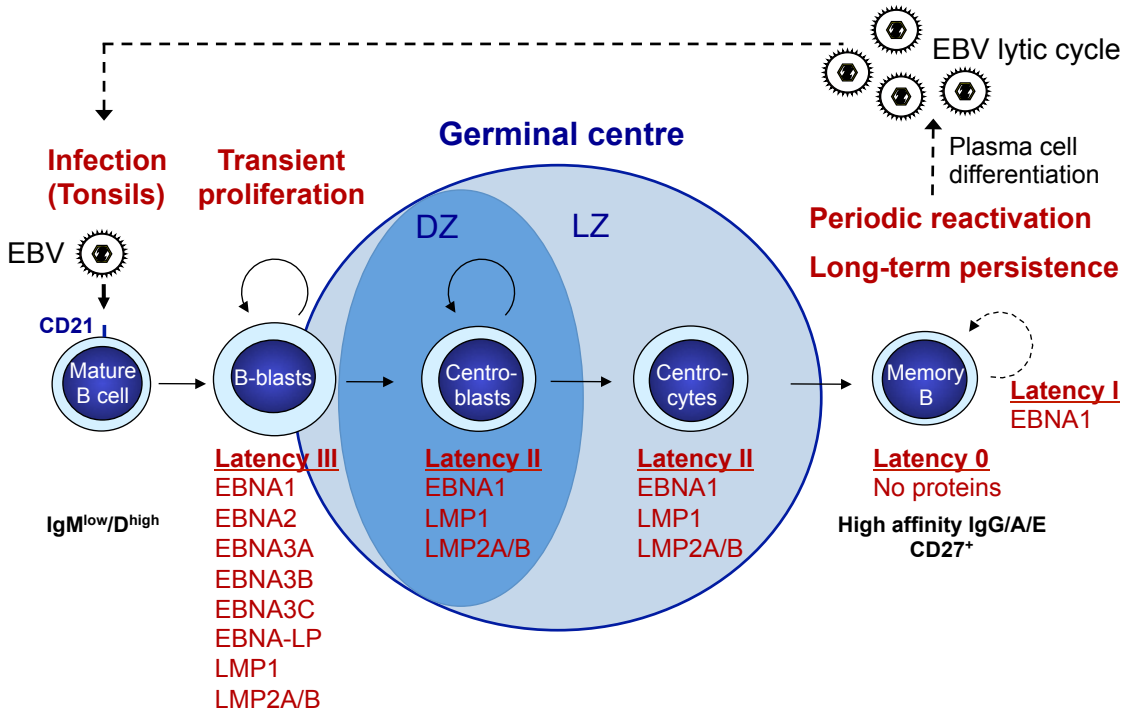


Figure 1.7: *In vivo* infection with EBV in the germinal centre model of persistence.
EBV infects mature, resting B cells in the tonsils by binding to CD21. Through the expression of all nine latent viral proteins comprising six EBV nuclear antigens (EBNA-1, -2, -3A, -3B, -3C and -LP) and three latent membrane proteins (LMP-1, -2A and -2B) EBV-infected mature B cells become activated and undergo rapid proliferation, which resembles antigen-activated B blasts. The viral gene expression programme responsible for this is known as latency III. Activated EBV infected B cells enter germinal centres and restrict viral gene expression to EBNA1, LMP1, LMP2A and known as latency II, resembling centroblasts and centrocytes. These then differentiate and exit the germinal centre to become resting memory B cells, in which no viral proteins are expressed (latency 0), except upon proliferation, in which case EBNA1 is expressed. Figure adapted from Prof Martin Allday.

1.2.2.2 In vitro infection

Infection of mature B cells with EBV *in vitro* is also able to establish a stable infection (Pope et al., 1968; Henle et al., 1967). However, in contrast to *in vivo* EBV infection, *in vitro* infection with EBV results in the creation of continuously proliferating lymphoblastoid cell lines (LCL) that express all nine latent viral protein of the latency III transcriptional programme (Figure 1.8). It is believed that due to the absence of GCs and an immune response to expressed latent proteins, differentiation into memory B cells is neither possible nor necessary. This transformation process of resting B cells into continuously proliferating LCL is sometimes also referred to as immortalisation.

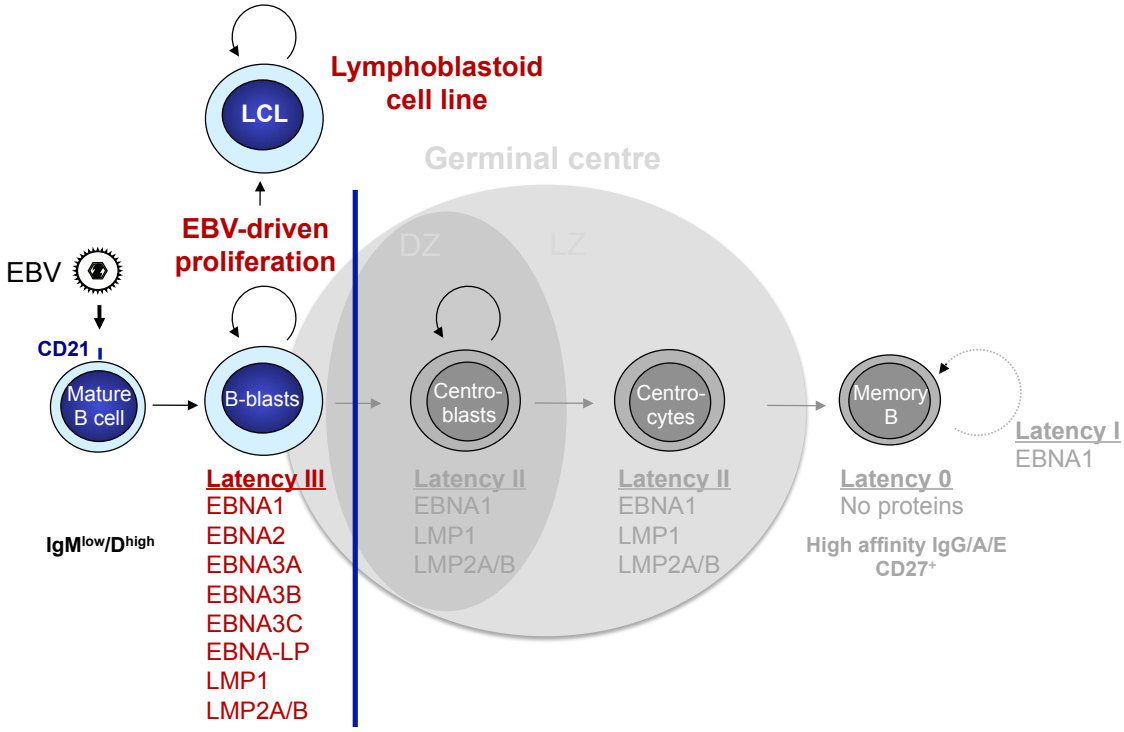


Figure 1.8: *In vitro* infection with EBV leads to the establishment of continuously proliferating lymphoblastoid cell lines (LCL).

EBV infects mature, resting B cells by binding to CD21. Continuous proliferation is driven by the expression of all nine latent viral proteins, six EBV nuclear antigens (EBNAs) and three latent membrane proteins (LMPs). Shut-down of viral gene expression as seen after *in vivo* infection (see Figure 1.7) is either not possible or not required due to the absence of a germinal centre and immune system. Therefore, EBV-driven proliferation continues and leads to the establishment of continuously proliferating LCL.

1.2.3 Latent gene transcription from the circular EBV genome

After infection, the linear EBV genome from the virion circularises and forms a viral episome (Figure 1.9) upon activation of the infected B cell (Hurley and Thorley-Lawson, 1988). This involves recombination of the terminal repeats (TR) located at the ends of the linear viral genome. Furthermore, the EBV genome contains internal repeats (also known as W repeats), which contain the W promoter (Wp) and encode the W exons of EBNA-LP. Early latent viral gene expression is initiated from Wp present in multiple copies in the viral genome (one copy per W repeat), but changes to the C promoter (Cp) in proliferating EBV-driven B blasts (Woisetschlaeger et al., 1990). This transcription leads to very long primary mRNAs coding for all six EBNAs (B cell transcription unit), which are then alternatively spliced to produce functional EBNA proteins. For the latency II and latency I programmes, transcription is initiated from the Q promoter (Qp) resulting only in the expression of EBNA1 (Schaefer et al., 1995; Tsai et al., 1995). LMP1 and LMP2A/B are transcribed independently from the EBNAs.

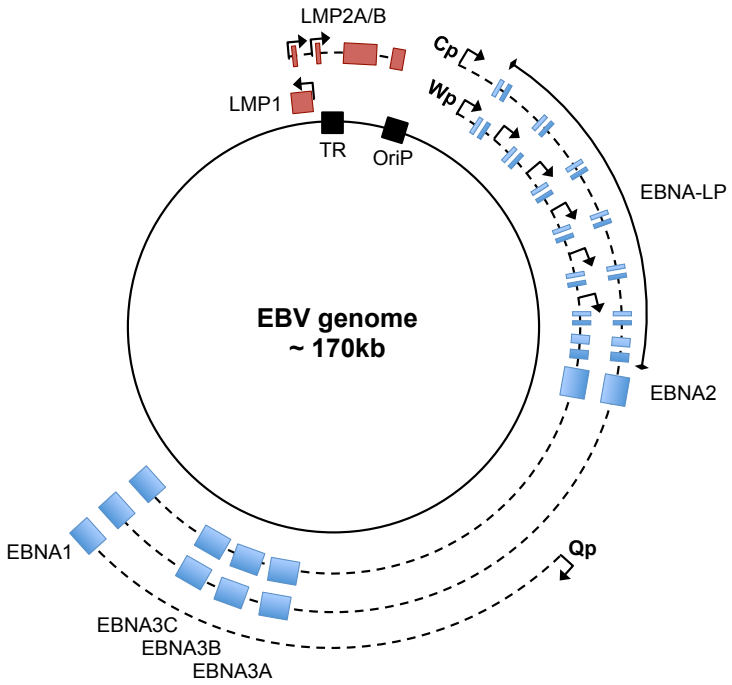


Figure 1.9: Latent protein-coding gene transcription from circular EBV genome.
Latently EBV-infected B cells contain the EBV genome in form of a circular episome. Early after infection, transcription of EBV nuclear antigens (EBNAs, coloured in blue) is initiated from the W promoter (Wp) present in multiple copies, but switches to the C promoter (Cp) in activated EBV-

infected B blasts. All six EBNAs are transcribed as very long primary mRNA and are alternatively spliced and translated into the individual EBNA proteins. During latency II and I, when only EBNA 1 is expressed, transcription is initiated from the Q promoter (Qp). Latent membrane proteins (LMPs, coloured in red) are transcribed independently from the EBNAs. Terminal repeats (TR) and the latent origin of replication (OriP) are shown as black boxes. Figure adapted from (Kalla and Hammerschmidt, 2012).

1.2.4 Latent viral proteins

1.2.4.1 EBV nuclear antigens

1.2.4.1.1 EBNA1

EBNA1 is expressed in all replicating EBV infected B cells, as one of its essential roles is to prevent the loss of the viral genome in the dividing cell. The latent origin of replication (OriP) on the EBV genome (Figure 1.9) enables replication and retention of viral episomes in latently EBV infected cells (Yates et al., 1984) and EBNA1 is essential for both replication and retention (Yates et al., 1985). Binding of EBNA1 to OriP enables the replication of the viral episome (Wysokenski and Yates, 1989), but also tethers the viral episome to cellular chromosomes during metaphase ensuring equal segregation of viral episomes between the two daughter cells (Grogan et al., 1983; Kanda et al., 2001). In addition to this, binding of EBNA1 to OriP transactivates transcription from the Cp (Sugden and Warren, 1989) and LMP promoters (Gahn and Sugden, 1995). Binding of EBNA1 to sequences downstream of its own Qp promoter had an inhibitory effect possibly to auto-regulate its own expression levels (Sample et al., 1992).

1.2.4.1.2 EBNA2

EBNA2 is the first expressed viral gene (Allday et al., 1989) and absolutely essential for B cell transformation (Cohen et al., 1989; Hammerschmidt and Sugden, 1989), as it is required for both initiation (Sinclair et al., 1994) and maintenance (Kempkes et al., 1995) of proliferation of EBV infected B cells. According to sequence differences in the EBNA2 gene, two major EBV clades (type I and type II EBV) are distinguished sharing only 56% of EBNA2 amino acid identity between each other. Type I EBV strains are more efficient in

transforming B cells into LCLs compared to type II strains, which is due to a single amino acid difference between the two types (Tzellos et al., 2014). It should be noted, however, that there are no obvious differences in the pathologies associated with type I or type II EBV.

EBNA2 contains N-terminal (Peng et al., 2004) and C-terminal transactivation domains (Cohen and Kieff, 1991) enabling EBNA2 to interact with transcriptional activators (Tong et al., 1995a; b) – *e.g.* transcription factor IIB (TFIIB), TFIIH , TATA-binding protein associated factor 40 (TAF40) and replication protein A 70 (RPA70) – and co-activators like histone acetyltransferases p300, CBP and PCAF (Wang et al., 2000). This transactivation ability of EBNA2 strongly activates the Cp (Woisetschlaeger et al., 1991; Sung et al., 1991) and LMP promoters (Wang et al., 1990; Ghosh and Kieff, 1990; Fåhraeus et al., 1990; Abbot et al., 1990), probably resulting in the Wp to Cp switch and LMP expression. EBNA2 does not exhibit direct DNA-binding capacity, but is tethered to its responsive promoters through binding to the cellular DNA-binding factor recombination binding protein-J (RBPJ, RBP or RBPJκ) (Grossman et al., 1994; Henkel et al., 1994; Waltzer et al., 1994; Zimber-Strobl et al., 1994). RBPJ is the DNA-binding component of the Notch signalling pathway (see Section 1.4.1) and is also known as CSL protein [C-promoter binding factor 1 (CBF1), Suppressor of hairless Su(H) in *Drosophila melanogaster* and Lag1 in *Caenorhabditis elegans*]. In addition to RBPJ, analysis of the bidirectional LMP1/LMP2B promoter identified a binding site for transcription factor PU.1/SPI1 that was important for EBNA2-mediated induction of the LMP1/LMP2B promoter (Laux et al., 1994; Johannsen et al., 1995) and an interaction between EBNA2 and PU.1/SPI1 was shown in GST-pulldowns (Johannsen et al., 1995; Zhao and Sample, 2000) and co-IP (Yue et al., 2004).

Beside the regulation of viral genes, EBNA2 also differentially regulates a large number of cellular genes. Some of the genes directly activated by EBNA2 are the B cell activation markers CD21 and CD23 (Wang et al., 1987; Cordier et al., 1990; Jochner et al., 1996), HES-1 (Sakai et al., 1998) and c-MYC (Kempkes et al., 1995; Jochner et al., 1996; Kaiser et al., 1999). In contrast to its role as transactivator, EBNA2 was shown to repress

expression of IgM (Jochner et al., 1996), Ig-β (Maier et al., 2006) and the GC-associated genes BCL6 (Boccellato et al., 2007) and AID (Tobollik et al., 2006). Induction of CD21, CD23 and other EBNA2-induced genes has been found to be strictly dependent on the expression of RBPJ, whereas repression of IgM and BCL6 was suggested to be only partially dependent on RBPJ, but no mechanisms have been proposed (Maier et al., 2005; 2006; Boccellato et al., 2007). The induction of c-MYC and repression of BCL6 by EBNA2 is not compatible with GC B cells (see Section 1.1.4). Expression of BCL6 is absolutely required for GC B cells and c-MYC is only expressed before B blasts enter GCs, but not within the GC except for the small subset of centrocytes in the LZ that recycle back into the DZ (Dominguez-Sola et al., 2012; Calado et al., 2012; Victora et al., 2010; 2012). By switching from latency III to latency II, EBV-infected B cells stop expressing EBNA2 and are able to express BCL6 to enter GCs.

See Section 1.4.2 for a model of EBNA2-mediated gene regulation.

1.2.4.1.3 EBNA3A, EBNA3B, EBNA3C

The EBNA3 family of genes is comprised of *EBNA3A*, *EBNA3B* and *EBNA3C*, which are arranged in tandem repeat on the EBV genome and share a similar gene structure with a short 5' exon and a long 3' exon. Alternative splicing of very long primary mRNA transcripts that originate from either the Wp or Cp promoter produces mRNAs for each of the EBNA3s (Figure 1.9). Transcription of EBNA3 is tightly regulated as there are only a few copies of EBNA3 mRNAs in LCLs, *e.g.* less than 3 mRNA copies of EBNA3C per cell (Davies et al., 2010). However, the half-life of EBNA3 proteins is at least 24 hours (Touitou et al., 2005). The two major EBV clades (type I and type II) also differ in their EBNA3 proteins with a amino acid identity between type I and type II of 84%, 80% and 72% for EBNA3A, EBNA3B and EBNA3C, respectively (Sample et al., 1990; Palser et al., 2015). EBNA3s are very large proteins (>900 aa, Figure 1.10A) that do not resemble any known cellular or viral proteins (Allday et al., 2015). However, they do resemble each other although they share only limited amount of sequence similarity. All three contain a N-

terminal homology domain with an amino acid sequence identity of 23-28% between all three (Le Roux et al., 1994). Protein structures have not been solved for any of the EBNA3s, but global protein structure prediction analysis reported a similarity in their secondary structure elements of 71-88% (Yenamandra et al., 2009). The homology domains, in particular, are predicted to fold into alpha helices and beta strands, while no obvious secondary structure was predicted for most of the remaining regions of the proteins [Figure 1.10A, (Allday et al., 2015)]. Although EBNA3s resemble each other there is no indication that they are redundant or can substitute for each other. Early studies indicated that EBNA3A and EBNA3C were essential for B cell transformation (Tomkinson et al., 1993), whereas EBNA3B was dispensable (Tomkinson and Kieff, 1992). More recent studies have shown, however, that under certain conditions (see below), EBNA3A and EBNA3C are dispensable for B cell transformation and therefore not absolutely essential (Hertle et al., 2009; Skalska et al., 2013).

All EBNA3s exhibit strong repressor activity in reporter assays when targeted directly to DNA by fusion with the DNA-binding domain of Gal4 (Bain et al., 1996; Waltzer et al., 1996; Bourillot et al., 1998; Allday et al., 2015), probably through interaction with cellular factors that are involved in transcriptional repression (Figure 1.10B), *e.g.* EBNA3C can interact with histone deacetylase 1 (HDAC1), HDAC2, C-terminal binding protein (CtBP), transcriptional co-repressors Sin3A and silencing mediator of retinoid and thyroid hormone receptors (SMRT/NcoR) (Radkov et al., 1999; Knight et al., 2003; Touitou et al., 2001). Similar to EBNA2, EBNA3 proteins do not contain intrinsic DNA-binding ability, but like EBNA2, all EBNA3s can bind to the DNA sequence binding factor RBPJ (Robertson et al., 1995; 1996; Krauer et al., 1996; Zhao et al., 1996; Marshall and Sample, 1995) and EBNA3C possibly also to PU.1/SPI1 (Zhao and Sample, 2000) and interferon regulatory factor 4 (IRF4) and IRF8 [(Banerjee et al., 2013), (Figure 1.10B)].

Each EBNA3 forms a unique complex with RBPJ (Ohashi et al., 2015), but for all three EBNA3s the RBPJ binding site is situated within their HDs (Robertson et al., 1996;

Zhao et al., 1996). Initially, four core residues ($_{209}$TFGC) within the HD of EBNA3C were identified to be important [(Zhao et al., 1996; Radkov et al., 1997), (Figure 1.10C)]. More recently and only for EBNA3C an additional site ($_{227}$WTP, Figure 1.10C) was identified and mutation of both sites ($_{209}$AAAA and W227S) was required for an effective loss of RBPJ binding as determined by co-immunoprecipitation (Calderwood et al., 2011). The interaction between EBNA3A or EBNA3C and RBPJ was shown to be important for sustaining the proliferation of LCLs (Maruo et al., 2005; 2009; Lee et al., 2009). However, recently it was possible to establish stable LCL from an EBNA3C virus carrying both $_{209}$AAAA and W227S mutations, so an interaction between EBNA3C and RBPJ is not absolutely required for B cell transformation (Kalchschmidt et al., 2016).

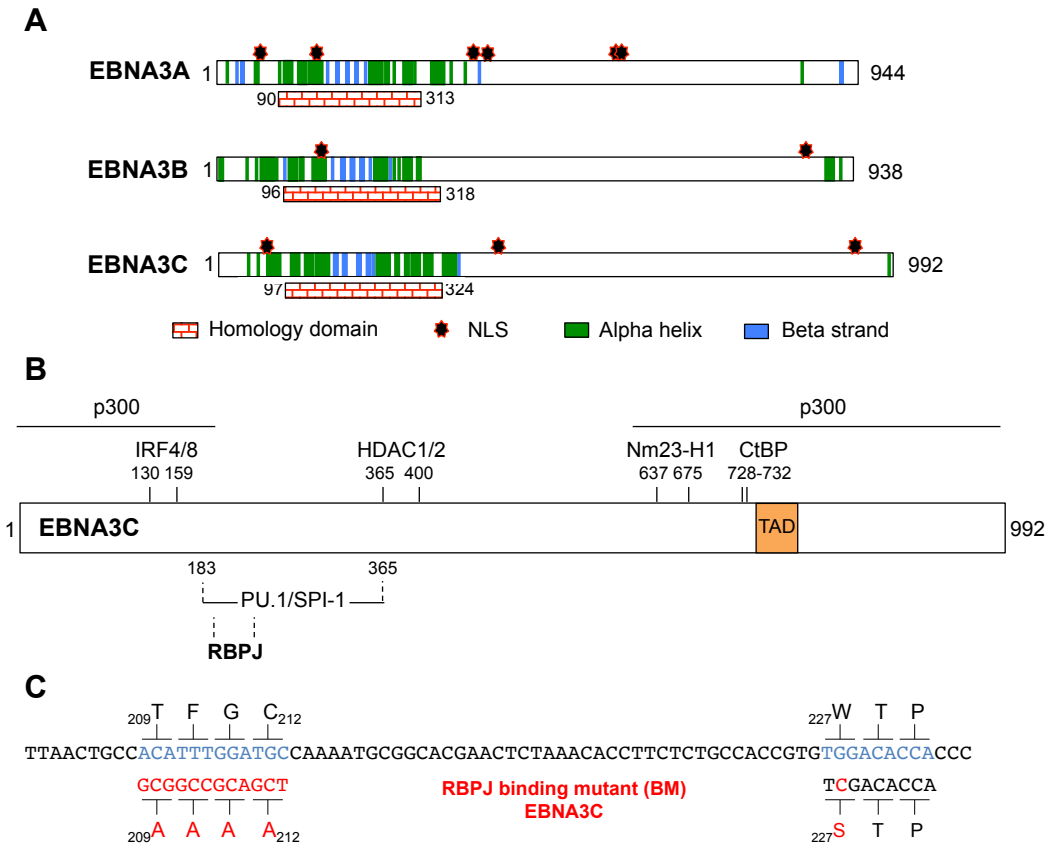


Figure 1.10: Overview of the EBNA3 family of proteins and selected interactions between EBNA3C and cellular factors.

(**A**) Overview of EBNA3 proteins indicating the homology domain, nuclear localisation signal (NLS), predicted alpha helices and beta strands by phyre[2] protein fold recognition server. Numbers indicate amino acids. (**B**) Schematic of full-length EBNA3C showing a transactivation domain (TAD) in orange and important residues for the interaction with selected cellular factors [histone acetyltransferase

p300, histone deacetylases 1 and 2 (HDAC1/2), Nm23-H1, C terminal binding protein (CtBP), and the transcription factors IRF4/8, PU.1/SPI1 and RBPJ]. (**C**) The amino acid residues $_{209}$TFGC and $_{227}$WTP are important for the interaction between EBNA3C and RBPJ. The nucleotide sequence of wild type EBNA3C is shown in black and mutations encoding $_{209}$AAAA and $_{227}$STP in the RBPJ binding mutant (BM) EBNA3C are shown in red. Panel A is adapted from (Allday et al., 2015), panel B from a drawing by Prof Martin Allday and panel C from (Kalchschmidt et al., 2016).

Beside their potent repressive activity on gene expression, EBNA3s have been also reported in gene activation of viral and cellular genes. EBNA3C has been shown to induce expression of LMP1 (Allday et al., 1993; Jiménez-Ramírez et al., 2006) and further studies showed that EBNA3C potentiates EBNA2-mediated activation of the LMP1 promoter (Marshall and Sample, 1995), which was independent of RBPJ, but required PU.1/SPI1 binding sites (Zhao and Sample, 2000; Lin et al., 2002). In contrast, repression of the LMP1 promoter by EBNA3C was also observed, which might depend on the feeding condition of the cells during the transient reporter assays (Robertson et al., 1995; Marshall and Sample, 1995). Furthermore, EBNA3C contains a C-terminal transactivation domain that increases transcription in reporter assays when targeted directly to DNA by fusion with the DNA-binding domain of Gal4 (Marshall and Sample, 1995) and both N- and C-termini have been reported to interact with histone acetyltransferase p300 [(Cotter and Robertson, 2000; Subramanian et al., 2002), (Figure 1.10A)]. Moreover, EBNA3C interacts with the nucleoside diphosphate kinase Nm23-H1 (Subramanian et al., 2001; Subramanian and Robertson, 2002), which has been reported to co-operate with EBNA3C in induction of matrix metalloproteinase 9 (Kuppers et al., 2005) and cyclo-oxygenase-2 gene expression (Kaul et al., 2006).

From gene expression analyses in various EBNA3 knock-out (KO) or conditional cell lines, it is now well established that all EBNA3s are transcriptional co-regulators of over 1000 host genes and often more than one EBNA3 cooperate in the regulation of one gene (Hertle et al., 2009; White et al., 2010; Zhao et al., 2011b; Skalska et al., 2010; McClellan et al., 2012). Two of the genes more extensively studied are the two tumour suppressors

BCL2L11 (BIM) and *CDKN2A* ($p16^{INK4A}$) that are both repressed by EBNA3A together with EBNA3C.

BIM (BCL2-interacting mediator) is a pro-apoptotic BCL-2 family member and is up regulated upon unscheduled cellular proliferation induced by oncogenes such as Myc (Egle et al., 2004). BIM is a potent initiator of apoptosis as it can inactivate BCL-2 and other pro-survival BCL-2 family members through high affinity binding. Both EBNA3A and EBNA3C cooperate in the repression of BIM (Anderton et al., 2007), which involved recruitment of polycomb repressive complex 2 (PRC2) and the deposition of repressive epigenetic histone modification H3K27me3 [(Paschos et al., 2009; 2012), see Section 1.3 for details on epigenetic marks and PRC2]. EBNA3A and EBNA3C were shown to bind proximal to the TSS of BIM (Paschos et al., 2012; McClellan et al., 2013).

$p16^{INK4A}$ is a cyclin-dependent kinase (CDK) inhibitor that binds to cyclin D dependent kinases (CDK4 and CDK6), which prevents their binding to D-type cyclins and inhibits the phosphorylation of retinoblastoma protein (Rb). Hypo-phosphorylation of Rb produced by high levels of $p16^{INK4A}$ results in G1 cell cycle arrest [reviewed in (Gil and Peters, 2006)]. Inactivation of EBNA3C in EBNA3C-conditional LCL, which carry an EBNA3C-estrogen receptor fusion protein that renders EBNA3C conditional to the presence of 4-hydroxytamoxifen (HT), through removal of HT resulted in the increase in $p16^{INK4A}$, dephosphorylation of Rb and cell cycle arrest (Maruo et al., 2006). EBNA3A is also involved in the repression of $p16^{INK4A}$ and together with EBNA3C is necessary for the establishment of repressive state by PRC2 and H3K27me3 (Skalska et al., 2010; Maruo et al., 2011). Repression also depends, in part, on the ability of EBNA3A and EBNA3C to interact with CtBP (Skalska et al., 2010). Furthermore, it was shown that EBNA3C binds proximal to the TSS of $p14^{ARF}$, located at the same locus as $p16^{INK4A}$ (Jiang et al., 2013), and that proliferation of LCL does not decrease after inactivation of EBNA3C in B cells with a homozygous deletion in $p16^{INK4A}$ ($p16^{INK4A}$-null) that ablates expression of functional $p16^{INK4A}$

(Skalska et al., 2013). Therefore, conditional EBNA3C LCL could be established from $p16^{INK4A}$-null primary B cells without expressing functional EBNA3C.

Least is known about EBNA3B, which is completely dispensable for *in vitro* B cell transformation. However, it appears that in a mouse model, EBNA3B acts as a tumour suppressor. Humanised mice infected with EBNA3B KO virus displayed a faster expansion of EBV infected B cells and lower secretion of the T cell attractant chemokine CXCL10 leading to more aggressive lymphoma development (White et al., 2012). Furthermore, mutations and deletions in EBNA3B have been identified in human EBV-associated diffuse large B cell lymphoma (White et al., 2012).

See Section 1.4.2 for a model of EBNA3-mediated gene regulation.

1.2.4.1.4 EBNA-LP

EBNA leader protein (EBNA-LP), which is alternatively spliced from transcripts initiated first from Wp and later from Cp (Figure 1.9), is highly expressed early after latent infection (Allday et al., 1989). EBNA-LP is composed of a various number of repeats, encoded by the W1 and W2 exons, and a unique C-terminal region encoded by the Y1 and Y2 exons (Figure 1.9), resulting in the expression of EBNA-LP with different molecular weights depending on the number of W repeats. It is believed that the main function of EBNA-LP is to augment EBNA2-mediated gene activation. EBNA-LP co-operates with EBNA2 in the induction of cyclin D2 to transition resting B cells from G0 into G1 of the cell cycle (Sinclair et al., 1994) and in the activation of the Cp and LMP promoters (Harada and Kieff, 1997; Nitsche et al., 1997). The mechanisms by which EBNA-LP co-operates with EBNA2 remain mainly unknown. Recently, it was proposed that EBNA-LP displaces the co-repressor NCoR from RBPJ and thereby promotes transactivation by EBNA2 (Portal et al., 2011; 2013), but this remains controversial [reviewed in (Kempkes and Ling, 2015)].

1.2.4.2 Latent membrane proteins

1.2.4.2.1 LMP1

LMP1 expression is induced by EBNA2 and can be potentiated by EBNA3C and EBNA-LP. Originally LMP1 was considered to be absolutely required for B cell transformation (Kaye et al., 1993), but in the presence of feeder cells it is possible to establish LMP1 KO LCL although at dramatically reduced rate compared to LMP1 competent wild type virus (Dirmeier et al., 2003). LMP1 consists of a short N-terminal cytoplasmic domain followed by six transmembrane domains and a longer C-terminal cytoplasmic domain that comprises two C-terminal activation regions (CART1 and CART2), which are essential for its signalling function. LMP1 forms a constitutively active receptor, because the transmembrane domains mediate homo-oligomerisation, which mimics ligand-induced cross-linking and results in constitutive signalling (Gires et al., 1997). Therefore, LMP1 is often considered to be equivalent to a constitutively active CD40 receptor (Hatzivassiliou et al., 1998; Kilger et al., 1998). LMP1 signals via various pathways, *e.g.* NF-κB, mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3-K), IRF7 and STAT pathways and has multiple function in EBV biology [reviewed in (Kieser and Sterz, 2015)]. LMP1 expression *in vivo* is restricted to latency III expressing EBV-induced B blast and latency II expressing GC B cells. By mimicking CD40L signals of T_{FH} cells, LMP1 was considered to help EBV-infected B cells to survive in the GC environment, but more recent data suggest that EBV-infected B cells are still undergoing antigen-selection (Souza et al., 2005; 2007). However, there is no doubt that LMP1 provides survival signals to EBV-infected B cells, *e.g.* through induction of BCL2 (Henderson et al., 1991), which is repressed by BCL6 in GC B cells (Saito et al., 2009). Furthermore, when ectopically over expressed LMP1 can induce the expression of AID probably through NF-κB signalling (He et al., 2003).

1.2.4.2.2 LMP2A and LMP2B

LMP2A and LMP2B exons span the two ends of the linear EBV genome and can only be expressed upon circularisation of the EBV genome (Figure 1.9). They are transcribed from two different promoters, but only differ in the first exon and share the remaining eight exons (exons 2-9). LMP2A and LMP2B are induced by EBNA2 (Zimber-Strobl et al., 1991) and repressed by EBNA3s (Le Roux et al., 1994). The first exon of LMP2A encodes for an N-terminal cytoplasmic domain with signalling ability, whereas the first exon of LMP2B is non-coding and therefore lacking signalling ability. Both proteins contain 12 transmembrane domains that anchor LMP2A and LMP2B in the cell membrane and a short C-terminal cytoplasmic tail that mediates homo- or hetero-dimerisation between LMP2A and LMP2B. The N-terminal cytoplasmic domain of LMP2A contains a consensus immunoreceptor tyrosine-based activation motif (ITAM) that interacts with components of the BCR signalling pathway. This allows LMP2A to mimic constitutively active BCR signalling (Engels et al., 2012) and to activate multiple downstream pathways, *e.g.* NFκB (Swanson-Mungerson et al., 2005), MAPK (Engels et al., 2012) and PI3K (Swart et al., 2000; Portis and Longnecker, 2004), resulting for example in increased cell survival and proliferation. Furthermore, LMP2A is able to induce Notch signalling, which can stimulate LMP2A expression in an auto-regulatory loop, possibly to maintain LMP2A expression in the absence of EBNA2 expression (Anderson and Longnecker, 2008). LMP2A and LMP2B expression *in vivo* is restricted to latency III expressing EBV-induced B blast and latency II expressing GC B cells. The ability of LMP2A to mimic constitutively active BCR signalling is thought to help EBV-infected B cells to survive in the absence of strong antigen-binding signals from the endogenous BCR (Cen and Longnecker, 2015).

1.2.5 EBV-associated lymphomas

EBV infection has been linked with different cancers, including B cell lymphomas, T cell lymphomas and different carcinomas. Only some of the more frequent B cell lymphomas

will be discussed here, which are derived from different stages of latently EBV infected B cells and express different latency programmes.

1.2.5.1 B cell lymphomagenesis

B cell lymphomas are far more frequent (up to 95% of lymphomas in the Western world) than T cell lymphomas although the number of B and T cells in humans are roughly equal. Most B cell lymphomas originate from GC B cells or post-GC B cells [reviewed in (Küppers, 2005)]. One hallmark – and possibly an early event in the development of B cell lymphomas – is the development of reciprocal chromosomal translocations. Translocations require DSB on two heterologous chromosomes that are in proximity of each other in order to be ligated together. Translocations found in B cell lymphomas often occur between one of the Ig loci and a proto-oncogene (*e.g. cMYC*), which places the proto-oncogene under the transcriptional control of the active Ig locus and therefore causes constitutive oncogene expression (Küppers, 2005). Multiple studies have shown that promiscuous AID activity can introduce DSB not only at the Ig loci, but also at non-Ig loci that are frequently translocated [reviewed in (Robbiani and Nussenzweig, 2013)]. This indicates an important role of AID in the development of B cell lymphomas and provides a logical explanation for the frequent GC or post-GC origin.

1.2.5.2 Burkitt's lymphomas

Burkitt's lymphomas (BL) are derived from GC B cells in the DZ that carry a characteristic translocation between *cMYC* and the Ig loci, predominantly IgH and less frequently IgL [reviewed in (Schmitz et al., 2014)]. This disrupts BCL6-mediated repression of *cMYC* and leads to constitutive activation of cMYC, but additional mutations of tumour suppressors, *e.g. TP53, CDKN2A* or *BCL2L11* (BIM), are required for lymphomagenesis (Schmitz et al., 2014). In addition, ~20% of BL cases carry mutations in *cMYC* that were possibly acquired during the SHM-induced translocation process (Dang et al., 2005) and some of the cMYC mutants failed to induce the tumour suppressor BIM (Hemann et al., 2005). The extent of AID expression is directly linked to the frequency of cMYC translocation

(Takizawa et al., 2008) and AID is required for these translocations in a mouse model of BL (Robbiani et al., 2008). BL can be divided into endemic (eBL), sporadic (sBL) and HIV-associated BL. Almost 100% of eBL and around 10-20% of sBL and HIV-associated BL cases are EBV positive.

eBL is aetiologically linked to EBV and malaria infection and occurs mostly in children in sub-Saharan Africa and Papua New Guinea, where *Plasmodium falciparum* is present resulting in chronic malaria infections. Two recent studies showed that malaria parasite extracts induced AID expression in human tonsilar B cells and that chronic malaria infection increased the transition of B cells through the GC (Torgbor et al., 2014), which was also observed in a mouse model of chronic malaria infection that created a favorable GC environment for the development of AID-dependent mature B cell lymphomas (Robbiani et al., 2015). It is thought that the role of EBV in the development of eBL is mostly to repress tumour suppressors and apoptosis-related factors in order to compensate for the c-MYC induced proliferative stress [reviewed in (Allday, 2009; Rowe et al., 2009)].

Many BL cell lines express only EBNA1 from the Qp latency I promoter, but showed a certain degree of dependence on the presence of EBV as expression of a dominant-negative EBNA1 that is unable to retain the EBV episome resulted in reduced growth and increased apoptosis (Vereide and Sugden, 2011). This effect was stronger in Wp-restricted BL cell lines (Kelly et al., 2002) that only express EBNA1, EBNA3A, EBNA3B, EBNA3C and EBNA-LP, but not EBNA2, LMP1 and LMP2A/B, because of a genomic deletion of EBNA2 [reviewed in (Speck, 2002)]. Deletion of EBNA2 in Wp-restricted BL is probably required in order to circumvent the strong repressive effect of EBNA2 on the translocated cMYC on the IgH locus (Jochner et al., 1996). This Wp-restriction was observed in 3 out of 15 (20%) of BL cell lines (Kelly et al., 2002) and more recently in 2 out of 20 (10%) and 4 out of 26 (15%) cases of eBL (Abate et al., 2015). The Wp-restricted gene expression programme might be positively selected because it allows EBV-infected BL cells to express EBNA3 proteins and other viral factors that help to suppress cMYC-induced apoptosis (Allday, 2009). However, it

is probably balanced, because EBNA3 proteins are highly immunogenic (see Section 1.2.5.4), which makes Wp-restricted BL cells good targets for cytotoxic T cells (CTLs).

1.2.5.3 Hodgkin lymphomas

Classical Hodgkin lymphomas (HL) are comprised of mononucleated Hodgkin and multinucleated Reed-Sternberg cells (HRS) that are derived from GC or post-GC B cells (Küppers et al., 1994). Malignant HRS cells represent only a minority of the tumour mass with the majority coming from non-malignant T and B lymphocytes that infiltrate tumour. The malignant HRS cells often carry crippling mutations that result in expression of non-functional BCRs (Kanzler et al., 1996). This can occur as part of the GC reaction, but normally B cells without a functional BCR undergo apoptosis due to the lack of survival signals from absent BCR signalling. It is believed that HSR cells originate from pre-apoptotic GC B cells that are rescued from apoptosis. Classical HLs are EBV positive in around 40% of cases and express the latency II programme (EBNA1, LMP1, LMP2A and LMP2B) – consistent with them being GC-derived. The expression of LMP1 and LMP2A is thought to help rescue the pre-apoptotic HRS cells from apoptosis through providing CD40 and BCR signalling, respectively [reviewed in (Küppers, 2009)].

1.2.5.4 Immunodeficiency-associated lymphoproliferative disease

The full oncogenic potential of EBV can be observed in immunocompromised patients that develop EBV-positive B cell lymphomas. These can be observed in patients on immunosuppression after bone marrow or solid organ transplantation (post-transplant lymphoproliferative disease, PTLD), in late-stage HIV infected patients or in patients with congenital immunodeficiencies. It is believed that the absence of an effective immune response against the EBV-infected B cells allows expression of the full set of latent viral proteins (latency III), which drives uncontrolled B cell proliferation [reviewed in (Hislop et al., 2007)]. Particularly the EBNA3 proteins are highly immunogenic (Hislop et al., 2007) and PTLDs often respond well when reactive CTLs are used in T cell therapy [reviewed in (Gottschalk et al., 2005)].

1.3 Gene regulation

1.3.1 Overview

Eukaryotic gene regulation is a highly complex process that involves many layers whose combinatorial effect determines the gene expression profile of an individual cell. Although all cells in the human body contain the same DNA blueprint (genotype), there are a multitude of different cell types each with distinct cellular functions (phenotype). Gene expression between two cells can differ through genetic and epigenetic mechanisms. Genetic mechanisms involve nucleotide changes in the DNA sequence that directly or indirectly affect gene expression (different genotype – different phenotype). Beside genetic mechanisms, there are epigenetic mechanisms that lead to a heritable phenotype without changes in the DNA sequence [same genotype – different phenotype, (Berger et al., 2009)]. Here some of the mechanisms that lead to differential gene expression will be described in further detail.

1.3.2 Positioning and higher order chromatin structure of chromatin

The interphase eukaryotic nucleus is not a uniform compartment, but is comprised of discrete compartments, *e.g.* euchromatin and heterochromatin distinguishable by their degree of chromatin compaction (Passarge, 1979). Historically, the organisation within the nucleus and the positioning of chromosomes, sub-chromosome structures or individual genes were visualised and defined by fluorescence *in situ* hybridisation (FISH). More recently chromosome conformation capture was developed allowing analysis of the spatial proximity of two genomic loci (Dekker et al., 2002). This involves crosslinking to preserve chromatin contacts, restriction enzyme digestion, ligation of digested fragments and PCR quantification of ligation products. PCR products are interpreted as chromatin contacts between the two interrogated loci that had to be in close proximity before cross-linking to allow subsequent ligation and amplification of the digested fragments. Derivatives of this technique such as Hi-C (Lieberman-Aiden et al., 2009) were developed, which allowed the analysis of the three-dimensional (3D) conformation of whole genomes by measuring

genome-wide chromatin interactions. Both FISH and chromosome conformation capture studies have dramatically advanced the understanding of the organisation of the nucleus in recent years.

1.3.2.1 Chromosome territories

Interphase chromosomes are not randomly distributed, but rather localise in chromosome territories, which are distinct spaces particular chromosomes preferentially occupy within the nucleus (Figure 1.11A). Generally, gene-dense chromosomes localise preferentially in the centre of the nucleus and gene-poor chromosomes are found towards the nuclear periphery (Boyle et al., 2001; Cremer et al., 2001). The localisation of chromosomes within the interphase nucleus can have a direct effect on their gene expression. For instance, artificial tethering of a chromosome to the nuclear periphery can prevent the expression of some of the genes located on the tethered chromosome (Kumaran and Spector, 2008; Finlan et al., 2008; Reddy et al., 2008). Importantly, chromosomes are not isolated within their chromosome territory, as high-resolution FISH studies have suggested that up to 40% of each individual chromosome is intermingled with other chromosomes (Branco and Pombo, 2006). These intermingled regions are mostly at the boundaries of two adjacent chromosome territories and the larger the intermingling volume between a pair of chromosomes, the higher the translocation frequency between these two. Furthermore, the position of an individual chromosome is not fixed within the nucleus. Rapid changes in the positioning of entire chromosomes have been observed after serum starvation in human primary fibroblasts (Mehta et al., 2010) or the position of individual gene loci during development. The Ig loci for example move from the nuclear periphery in non-Ig expressing HSCs, to a more central position within the nucleus in Ig-expressing pro-B cells (Kosak et al., 2002).

1.3.2.2 Topologically associating domains

On the sub-chromosomal level, the initial Hi-C study that analysed genome-wide chromatin interactions observed a compartmentalisation into two classes termed A and B and observed more frequent interaction within the same class compared to in between the two classes (Lieberman-Aiden et al., 2009). Compartment A exhibited more open chromatin compared to compartment B and also higher gene density and gene expression. With higher resolution Hi-C, highly enriched local chromatin interactions were detected in domains with an average size of ~900 kb known as topologically associating domains (TADs) that spanned ~90% of the murine genome (Dixon et al., 2012). Within TADs a high frequency of chromatin interactions between loci is observed and these end abruptly at the boundaries of the TAD (Figure 1.11B). The boundary regions adjacent to TADs are enriched in housekeeping genes and various genetic elements, such as binding sites for the chromatin factors cohesin and CCCTC-binding factor (CTCF). Cohesin links sister chromatids throughout S-, G2- and M-phase of the cell cycle and has been found together with CTCF to control higher order chromatin structures [reviewed in (Merkenschlager and Odom, 2013)]. The exact role of cohesin and CTCF in establishing or maintaining TADs remains to be elucidated. Other factors are probably also involved, because mice with conditional knock-out of CTCF in the limb mesenchyme displayed a similar expression pattern of the *Hoxd* gene cluster, which was suggested to be isolated by CTCF from transcriptional interference to enable coordinated regulation during development (Soshnikova et al., 2010). Moreover, deletion of CTCF in murine macrophages produced a relatively mild phenotype and weakly affected global gene expression compared to wild type macrophages (Nikolic et al., 2014).

1.3.2.3 Long-range chromatin interactions

Within a TAD, long-range interactions between genetic elements can directly influence gene expression (Figure 1.11C). Such elements are often referred to as enhancers or silencers depending on their stimulating or repressive effect, respectively, on gene expression. Enhancers and silencers are often distal from the regulated promoter, but

through looping out of the intervening chromatin, long-range interactions with the promoter can be established and detected by FISH and chromosome conformation capture techniques. This can lead to the formation of an active chromatin hub, which might stabilise the TAD [reviewed in (Pombo and Dillon, 2015)]. Cohesin and CTCF are two of the factors that are thought to mediate these long-range interactions between genetic elements within a TAD (Merkenschlager and Odom, 2013). However, in addition to their role in controlling long-range interactions, CTCF binding sites that are positioned between an enhancer and its promoter can block their interaction, which is why CTCF binding sites are sometimes referred to as insulators (Merkenschlager and Odom, 2013).

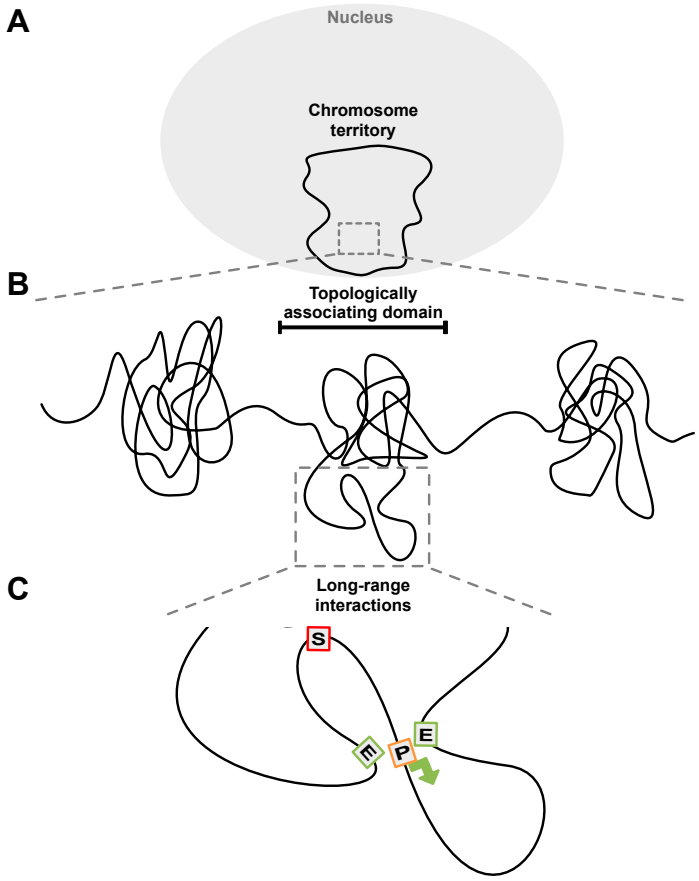


Figure 1.11: Positioning and higher order chromatin structures in the nucleus.
(**A**) Chromosomes have a preferred organisation within the nucleus and localise in distinct chromosome territories. (**B**) The chromatin of each chromosome is arrange in topologically associating domains (TADs) with high intra-TAD and low inter-TAD interaction frequency. (**C**) Within a TAD, long-range interactions between regulatory elements control gene expression, *e.g.* looping between the promoter (P) and upstream and downstream enhancers (E) leads to looping-out of a silencer (S) and active gene expression (green arrow).

1.3.3 Chromatin and chromatin modifications

Within the nucleus, the DNA double helix is wrapped around a histone octamer formed by two H3-H4 dimers and two H2A-H2B dimers (Campos and Reinberg, 2009). The complex of DNA wrapped around a histone octamer is known as nucleosomes, which form the basic building unit of chromatin (Figure 1.12). Various classes of proteins regulate the state and composition of chromatin. For example, histone remodelling complexes (*e.g.* SWI/SNF) can assemble, disassemble, exchange or reposition nucleosomes in an energy dependent manner, which can make chromatin more or less accessible to chromatin interacting proteins [reviewed in (Becker and Workman, 2013)]. Chromatin can be further modified either directly at the level of DNA or at the level of histones.

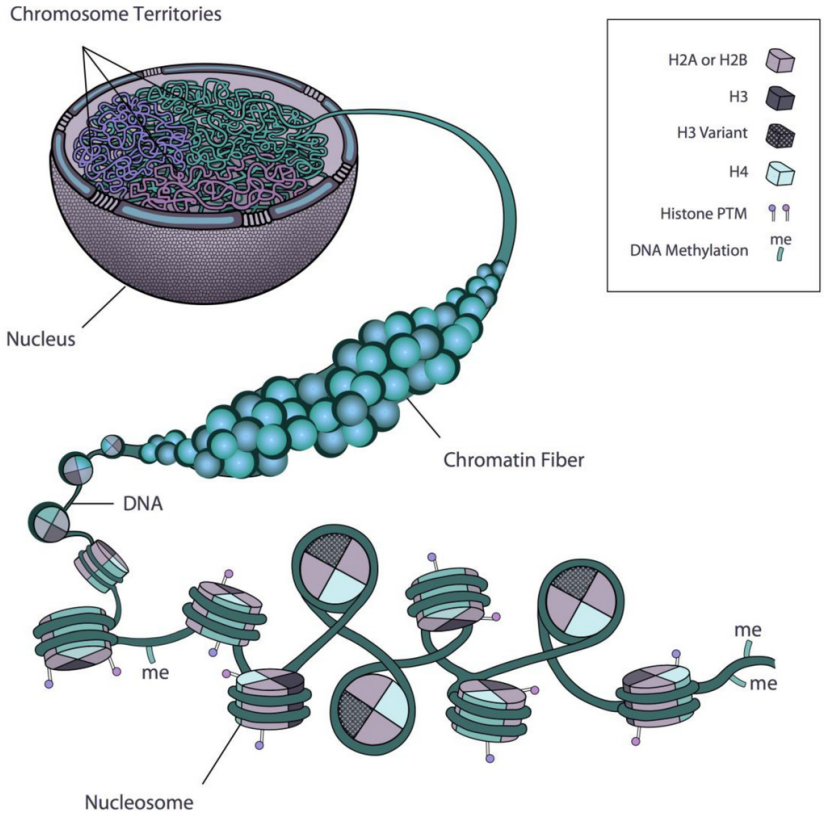


Figure 1.12: Chromatin structure and modifications.
DNA is tightly wrapped around a histone octamer formed by two H3-H4 dimers and two H2A-H2B dimers forming a nucleosome. Nucleosomes are the building unit of chromatin, which further organises into chromatin fibers that are localised in chromosome territories within the nucleus. Chromatin can be modified by DNA methylation and posttranscriptional modifications (PTM) of histones. Figure taken from (Rosa and Shaw, 2013), licensed under a Creative Commons Attribution 3.0 Unported License, http://creativecommons.org/licenses/by/3.0/.

1.3.3.1 DNA methylation

DNA methyltransferases (DNMTs) catalyse the methylation of cytosine residues resulting in 5-methylcytosines predominantly within cytosine-phosphate-guanine (CpG) dinucleotides. Most CpG dinucleotides (70-80%) are methylated in mammalian somatic tissue, but CpG islands, genomic regions rich in CpG dinucleotides that often span the promoters regions of genes, are often non-methylated. The pattern of DNA methylation is maintained after DNA replication by DNMT1 or altered by DNMT3A/B that catalyse *de novo* DNA methylation [reviewed in (Li and Zhang, 2014)].

DNA methylation changes during development and is generally associated with gene silencing (Li and Zhang, 2014). Methylated CpGs are bound by proteins containing methyl-CpG binding domains (Meehan et al., 1989), which can recruit corepressor complexes such as Sin3A and HDACs (Jones et al., 1998; Nan et al., 1998). Methylated CpGs can also prevent binding of activating TFs and thus leading to repression (Watt and Molloy, 1988). In contrast, DNA methylation of CTCF binding sites distal to the insulin-like growth factor 2 (IGF2) gene prevented binding of CTCF to these sites, allowing looping between the promoter and a distal enhancer and therefore expression from the methylated paternal allele, whereas the non-methylated maternal allele was bound by CTCF, allowing CTCF to act as an insulator by preventing enhancer-promoter looping and expression of IGF2 (Bell and Felsenfeld, 2000).

DNA methylation can be removed passively through lack of maintenance of existing DNA methylation by DNMT1 after DNA replication. However, active DNA demethylation was observed in the absence of DNA replication (Li and Zhang, 2014) and, only recently, ten eleven translocation (TET) proteins were discovered that are able to remove DNA methylation. TET proteins catalyse the oxidation of 5-methylcytosine to 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine, which are thought to be either actively removed by the BER pathway or passively lost after replication, because they do not induce methylation by DNMT1 [reviewed in (Kriaucionis and Tahiliani, 2014; Li and

Zhang, 2014)]. In addition, AID has also been proposed to erase CpG DNA methylation by either deaminating 5-methylcytosine into thymine or 5-methylcytosine into 5-methyluridine, resulting in base pair mismatches that activate the BER. However, the role of AID in DNA demethylation is controversial and it appears unlikely that AID is involved in global DNA demethylation in all cell types [critically reviewed in (Ramiro and Barreto, 2015)].

1.3.3.2 Histone modifications and modifiers

Histones contain protruding N-terminal and C-terminal tails that can be modified by post-translational modifications (PTM), *e.g.* acetylation, methylation, phosphorylation and ubiquitination (Figure 1.13). Because both the type of PTM, but also the position of the residue that has been modified can affect the local chromatin state either supporting or preventing active transcription, a histone code has been proposed (Strahl and Allis, 2000; Turner, 2000). Some histone modifications are mutually exclusive, for example different modifications at the same residue, whereas some modifications can enhance or inhibit other modifications on neighbouring residues [reviewed in (Bannister and Kouzarides, 2011)].

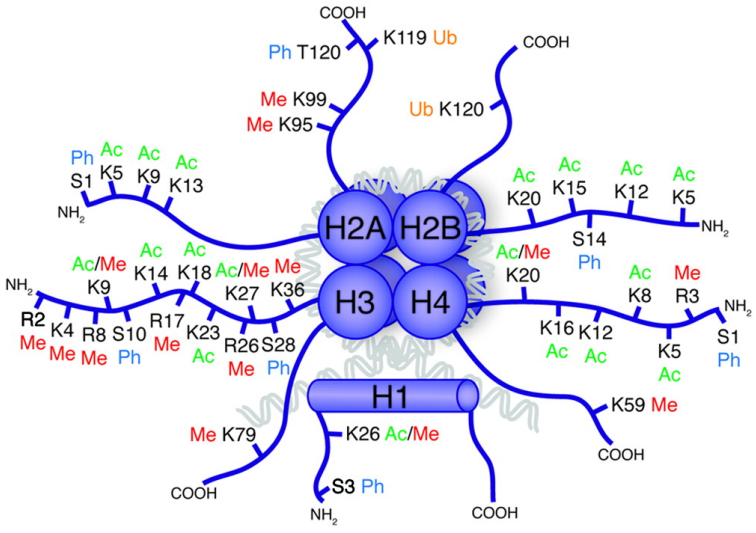


Figure 1.13: Overview of histone modifications.
Schematic nucleosome representation showing posttranscriptional modifications on the N-terminal (NH_2) or C-terminal (COOH) tails of histone 2A (H2A), H2B, H3, H4 and the linker H1. Arginine (R), lysine (K), serine (S) and threonine (T) residues can be modified by acetylation (Ac), methylation (Me), phosphorylation (Ph) and ubiquitination (Ub). Figure taken from (Tollervey and Lunyak, 2012), licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License, https://creativecommons.org/licenses/by-nc/3.0/.

Histone modifying proteins catalyse the addition or removal of these histone PTMs. For example, histone acetyltransferases (HATs) catalyse the acetylation of lysine residues, which neutralises the positive charge of lysine and might loosen the wrapping of DNA around the acetylated histone, whereas HDACs remove acetyl groups and therefore restore the positive charge of lysine residues and possibly strengthen the association between DNA and histones (Bannister and Kouzarides, 2011). Therefore, histone PTMs are thought to directly alter the local chromatin state, which can allow or prevent direct access of DNA binding proteins. In addition, histone PTMs can facilitate binding of proteins containing specific domains or prevent binding of others (Patel and Wang, 2013). For example, bromodomains recognise acetylated lysine residues and are often found in HATs and histone remodelling complexes, which might facilitate the recruitment of SWI/SNF complexes to acetylated promoter regions (Hassan et al., 2002). In contrast, DNMT3L, an inactive DNMT that interacts with DNMT3A/B, can only bind to non-methylated, but not to methylated histone 3 lysine 4 (H3K4) residues and is therefore thought to establish DNA methylation at non-methylated H3K4 nucleosomes (Ooi et al., 2007).

1.3.3.2.1 Polycomb-group proteins

Polycomb-group proteins (PcG) form histone modifying complexes, whose founding member was first discovered in *Drosophila melanogaster*, in which mutation of the Polycomb gene Pc resulted in additional sex combs, not only on the first, but also on the second and third pair of legs of male flies. Subsequently, mutations in other genes causing a similar phenotype were identified and grouped into the Polycomb family. PcGs control body segmentation during the embryonic development of *Drosophila* by repressing the homeobox (*HOX*) gene cluster in a coordinated way to regulate cell specific activation of the *HOX* genes (Lewis, 1978). PcGs are evolutionary conserved across species and form complexes that are generally considered as repressors of gene expression.

The two most studied polycomb complexes are polycomb-repressive complex 1 (PRC1) and PRC2 (Figure 1.14). PRC1 and PRC2 complexes have different enzymatic activities. The PRC1 subunits RING1A and RING1B contain E3 ligase activity and can catalyse ubiquitination of histone 2A lysine 119 [H2AK119ub, (de Napoles et al., 2004; Wang et al., 2004)], whereas EZH1/2 of PRC2 contain methyltransferase activity and catalyse histone 3 lysine 27 mono/di/trimethylation [H3K27me1/me2/me3, (Kuzmichev et al., 2002; Müller et al., 2002; Czermin et al., 2002; Cao et al., 2002; Ferrari et al., 2013)]. H2AK119ub and H3K27me3 are thought to directly block gene transcription [reviewed in (Aranda et al., 2015)].

At least six distinct PRC1 complexes (Gao et al., 2012) can be distinguished depending on the identity of the polycomb group factor (PCGF) subunit (PRC1.1. contains PCGF1, PRC1.2 contains PCGF2, etc.). All PRC1 complexes (Figure 1.14A) contain one RING1 protein (RING1A or RING1B) and one PCGF member, but can be further divided into canonical (cPRC1) and non-canonical (ncPRC1) complexes. There are two cPRC1 complexes containing either PCGF2 (MEL18) or PCGF4 (BMI1) and proteins of the chromobox-domain (CBX) family (CBX2/4/6/7/8) and the human polyhomeotic homolog (HPH1-3) family. ncPRC1 complexes lack CBX and HPH proteins, but contain RING1/YY1-binding protein [RYBP or its homolog YY1-asociated factor 2 (YAF2)] together with additional auxiliary proteins [reviewed in (Aranda et al., 2015; Kondo et al., 2016)].

The core PRC2 complex (Figure 1.14B) is composed of suppressor of zeste 12 (SUZ12) and one family member of enhancer of zeste 1 or 2 (EZH1/2) and embryonic ectoderm development 1-4 (EED1/2/3/4). Additional proteins associate with PRC2 [reviewed in (Vizán et al., 2015)].

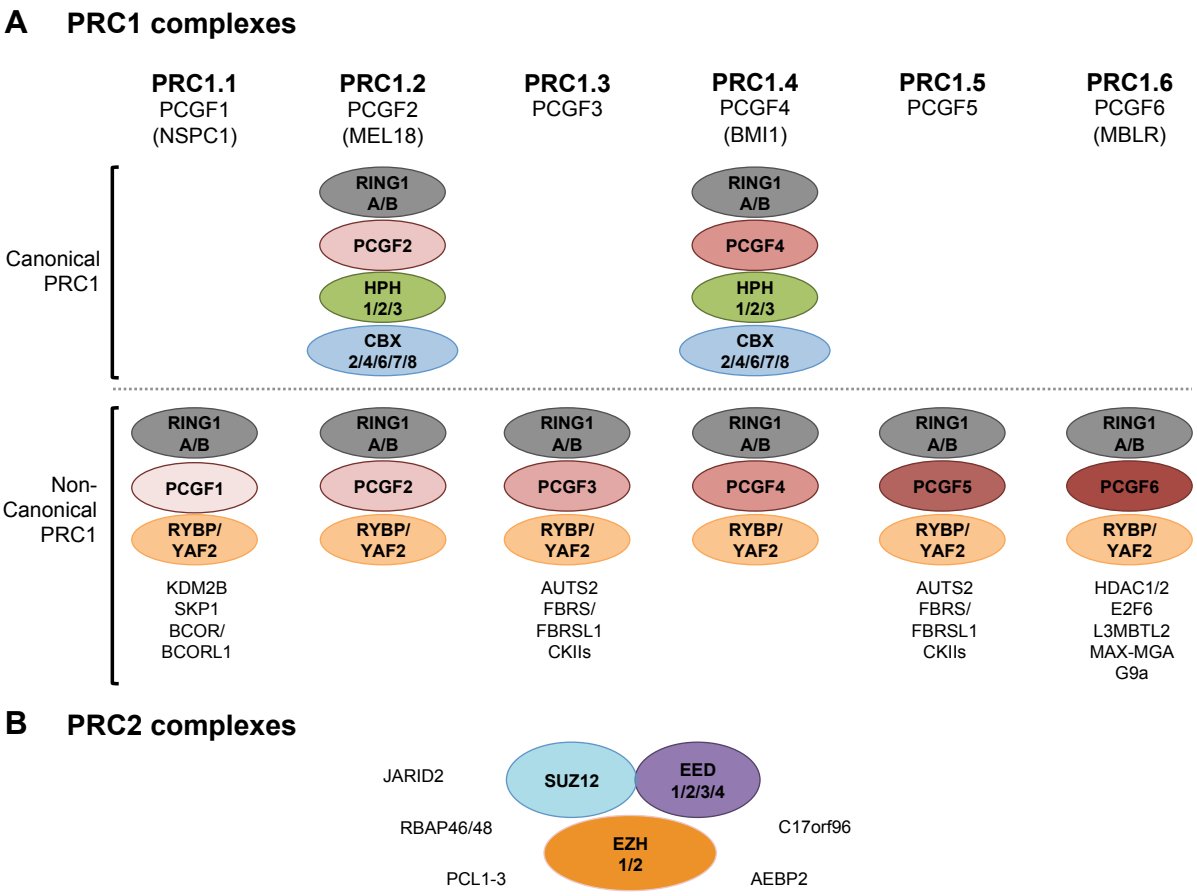


Figure 1.14: Composition of mammalian PCR1 and PRC2 complexes.
(**A**) There are at least six distinct PRC1 complexes [PRC1.1-PRC1.6 (Gao et al., 2012)] distinguishable by the incorporated PCGF protein (PCGF1-6). The two canonical PCR1 complexes (PRC1.2 and PRC1.4) contain either PCGF2 (MEL18) or PCGF4 (BMI1), respectively, and additionally one member of RING1 (RING1A or RING1B), CBX (CBX2/4/6/7/8) and HPH (HPH1/2/3). Non-canonical PRC1 lack CBX and HPH proteins, but contain one RING1 protein (RING1A or RING1B), one PCGF protein (PCGF1-6) and RYBP (or its homolog YAF2). Non-canonical PRC1 also associate with additional auxiliary proteins. (**B**) The core PRC2 is composed of SUZ12, EZH (EZH1 or EZH2) and EED (EED1/2/3/4). Additional auxiliary proteins interact with PRC2.

In the sequential model of polycomb recruitment, PRC2 was believed to recruit PRC1 complexes, because H3K27me3 modified chromatin can be bound by the CBX subunit of cPRC1 complexes (Kuzmichev et al., 2002; Czermin et al., 2002; Cao et al., 2002; Min et al., 2003). The PCGF subunits of cPRC1 (BMI1 and MEL18) have been shown to stimulate the ubiquitination activity of RING1 (Cao et al., 2005; Elderkin et al., 2007). In addition to inducing H3K27me3 and H2AK119ub, PRC1 and PRC2 can cause chromatin compaction and the local formation of heterochromatin during specific gene or genomic locus silencing

(Shao et al., 1999; Levine et al., 2002; Francis et al., 2004; Margueron et al., 2008; Eskeland et al., 2010). However, this model can only apply to cPCR1 complexes that contain CBX subunits, but not the more recently characterised ncPRC1 complexes, which do not contain any CBX subunits (Gao et al., 2012; Tavares et al., 2012).

Furthermore, results of recent studies have challenged the classical, sequential recruitment model in which PRC2-induced modifications recruit PRC1. It has been shown that PRC1 can be recruited independently from PRC2 and the H3K27me3 modification (Schoeftner et al., 2006; Pasini et al., 2007; Tavares et al., 2012; Yu et al., 2012). Furthermore, PRC1 can actually recruit PRC2 through the deposition of H2AK119Ub (Blackledge et al., 2014; Cooper et al., 2014; Kalb et al., 2014). In addition to PRC1 and PRC2 recruiting each other, DNA sequence-specific factors have been proposed to recruit PcG complexes. For example murine Pcgf4 (Bmi1) was shown to directly interact with the DNA sequence-specific binding factor RUNT-related transcription factor 1 (Runx1) and depletion of Runx1 in primary murine thymocytes reduced the occupancy of Ring1b at Runx1 binding sites (Yu et al., 2012). Interaction between PRC2 proteins with long non-coding RNA was also observed, such as HOX transcript antisense RNA (*HOTAIR*) expressed from the *HOXC* gene locus that target PRC2 to the *HOXD* gene (Rinn et al., 2007), and in a RNA sequence-independent manner also with short RNA that are transcribed from the 5' end of polycomb-repressed genes (Kanhere et al., 2010). In addition, non-sequence specific DNA elements have been identified, with PRC1 and PRC2 binding sites frequently found in or adjacent to CpG islands (Lee et al., 2006; Tanay et al., 2007; Ku et al., 2008; Mendenhall et al., 2010). The auxiliary proteins of PRC1.1 (KDM2B) and PRC2 (AEBP2 and JARID2) could be responsible for this, because they specifically recognise non-methylated CpG dinucleotides (Farcas et al., 2012; He et al., 2013; Wu et al., 2013) or generally GC-rich DNA (Kim et al., 2009; Li et al., 2010), respectively. However, the exact mechanism(s) by which different PRC1 and PRC2 complexes are recruited remain poorly understood.

Promoters of polycomb-repressed genes marked by H3K27me3 in ESCs are likely to acquired DNA methylation during differentiation, possibly to further reinforce silencing of these genes in somatic cells [reviewed in (Margueron and Reinberg, 2011)]. This promoted the idea that polycomb-mediated repression can lead to CpG DNA methylation, but much remains unknown about the potential crosstalk between PcG silencing and DNA methylation [reviewed in (Jin et al., 2011; Rose and Klose, 2014)]. Consistent with this idea, EZH2 was shown to directly interact with DNMTs (Viré et al., 2006). However, PRC2 was also reported to interact with TET1 in murine ESCs, but not in murine embryonic fibroblasts (Neri et al., 2013), which might ensure non-methylated CpG dinucleotides at polycomb-regulated gene promoters in ESCs.

Around 70% of polycomb-repressed genes in murine embryonic stem cells (ESCs) are contained in polycomb domains enclosed between two CTCF and cohesin binding sites (polycomb-repressed TADs) with an average size of ~110 kb (Dowen et al., 2014). These polycomb-repressed TADs exhibit very high levels of EZH2, SUZ12 and H3K27me3 and disruption of the TAD through deletion of either of the two CTCF binding sites at the boundary de-repressed the enclosed gene to some degree. Polycomb-repressed domains are subject to changes, for example during the sequential activation of the *HOX* genes, which are initially in a polycomb-repressed TAD, but progressively cluster in an active TAD during embryonic development (Noordermeer et al., 2011). Again, deletion of specific CTCF binding sites within the boundaries result in invasion of activation-associated histone marks into the polycomb-repressed TAD and expression of some of the *HOX* genes (Xu et al., 2014; Narendra et al., 2015). Therefore, CTCF and cohesin binding seem to play a crucial role in maintaining polycomb-repressive TADs and possibly TADs in general. Furthermore, deletion of both catalytic PRC1 subunits RING1A and RING1B in murine ESCs completely disrupted the *Hox* gene cluster, by abolishing promoter-promoter contacts, but maintaining promoter-enhancer contacts, leading to de-repression of *Hox* genes (Schoenfelder et al., 2015). The effect of depletion of the PRC2 subunit EED on the organisation of the *Hox* locus

was much more subtle and disrupted only a few interactions, but maintained the overall network (Denholtz et al., 2013; Schoenfelder et al., 2015). An interaction between the *Drosophila* PRC1 homolog of CBX and cohesin subunits has been reported (Strübbe et al., 2011), which might help PRC1 complexes to establish repressive interactions.

Historically, PcGs have always been associated with repression, however, several recent studies implicate PRC1 complexes in gene activation (Kondo et al., 2014; Gao et al., 2014; Creppe et al., 2014; Frangini et al., 2013; Pemberton et al., 2014; Morey et al., 2015; Schaaf et al., 2013). For example, in resting murine B and T cells cPRC1.4 (Ring1B, Bmi1 and Cbx7) occupy many actively transcribed promoters. cPRC1.4 co-immuniprecipates with Aurora kinase B and both co-occupy the same active promoters, at which Aurora kinase B phosphorylates H3S28 resulting in the inhibition of PRC1-catalysed ubiquitination of H2AK119 and the recruitment of deubiquitinase USP16. Conditional KO of either Ring1B or Aurora kinase B resulted in reduced transcription of these genes and decreased cell viability (Frangini et al., 2013).

A role of PRC2 complexes in gene activation has been also reported (Xu et al., 2012b; Mousavi et al., 2012; Ferrari et al., 2013; Xu et al., 2015). For example, EZH1 containing PRC2 complexes are recruited to newly activated genes and depletion of EZH1 impaired the expression of these genes possibly through ablating a stimulatory interaction between EZH1 and elongating RNA polymerase II (Mousavi et al., 2012).

1.3.4 Control of gene expression by chromatin interactions and modifications

Figure 1.15 provides an overview of chromatin interactions and modifications that control gene expression, showing selected histone modifications and histone modifiers relevant for this project.

Transcriptional activation of RNA-polymerase II-transcribed genes (Figure 1.15A) is often initiated by TF that bind to regulatory DNA sequences (*e.g.* distal enhancer elements or proximal promoter elements) and recruit co-activators such as HATs and histone

remodelling complexes. These catalyse activation-associated histone modifications such as histone 3 lysine 4 trimethylation (H3K4me3), histone 3 lysine 9 acetylation (H3K9ac) and histone 3 lysine 27 acetylation (H3K27ac) and reorganise nucleosomes making chromatin more accessible to chromatin-binding proteins. This also recruits the mediator complex that is comprised of 26 core subunits and is considered as bridging factor between TF and RNA polymerase II [reviewed in (Allen and Taatjes, 2015)]. Mediator and other factors (*e.g.* cohesin and CTCF) are thought to establish and/or stabilise long-range interactions (see Section 1.3.2.3) between distal enhancer elements and the TSS to promote initiation of transcription. For example, recently it was shown that mediator was recruited to the IgH locus during CSR and depletion of two mediator subunits impaired CSR possibly as a result of reduced frequency of long-range chromatin interactions between the IgH enhancer and the acceptor S region and reduced rate of transcription through the acceptor S region (Thomas-Claudepierre et al., 2016). The Mediator complex recruits and stabilises the RNA polymerase II pre-initiation complex. In the classical model of RNA polymerase II pre-initiation complex assembly [reviewed in (Sainsbury et al., 2015)], the initiation factor TFIID, a multiprotein complex containing the TATA box-binding protein (TBP) and 13 distinct TBP-associated factors (TAFs), binds to the TATA box ~30bp upstream of the TSS of TATA-containing promoters, but probably also at TATA-less promoters. TFIIA and TFIIB binding on either side of TFIID stabilises the complex and facilitates binding of RNA polymerase II and TFIIF, which together form the core pre-initiation complex. Binding of TFIIE to RNA polymerase II promotes the recruitment of TFIIH, a complex of 10 subunits with energy-dependent DNA helicase and kinase activity. TFIIH opens the DNA double helix and initiates transcription through phosphorylation of the C-terminal domain serine 5 residue of RNA polymerase II [reviewed in (Compe and Egly, 2012; Eick and Geyer, 2013)].

Gene expression can be repressed (Figure 1.15B) through the recruitment of HDACs, PRC1 and PRC2 resulting in histone deacetylation and deposition of the repressive histone modifications H2AK119ub and H3K27me3 (see Section 1.3.3.2.1). This can lead to a higher

nucleosome density, which makes the distal enhancer element inaccessible to DNA binding factors, disturbs the long-range enhancer-TSS interaction and prevents the assembly of the RNA polymerase II pre-initiation complex.

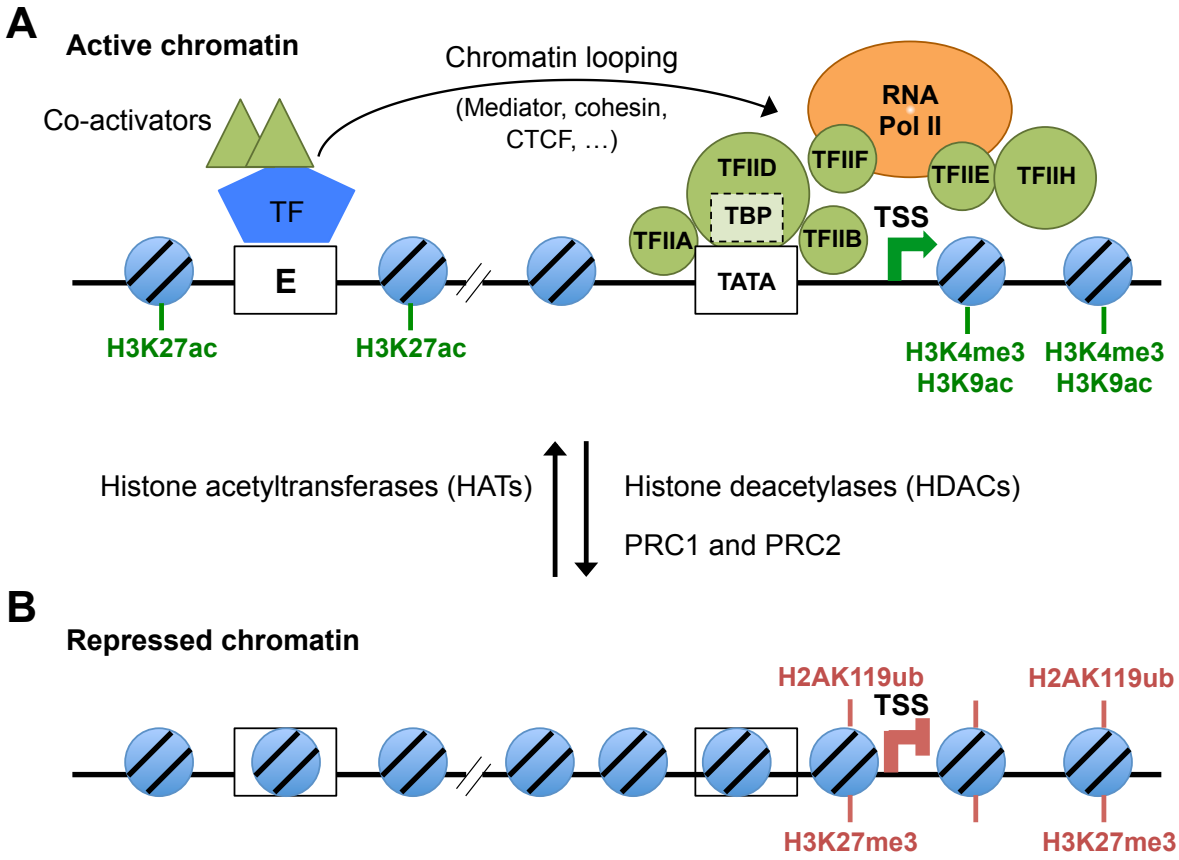


Figure 1.15: Regulation of gene expression.

(**A+B**) Chromatin is composed of nucleosomes, which consist of DNA (black line) wrapped around a histone octamer (blue circle) of two H3-H4 dimers and two H2A-H2B dimers. Histone modifiers catalyse posttranscriptional modifications (PTMs) on histone tails that can alter the chromatin state (active or repressed). (**A**) Active chromatin is characterised by a lower nucleosome density, which allows DNA-sequence-binding transcription factors (TF) to bind to distal enhancer elements (E) and to recruit co-activators that catalyse activation-associated histone marks, such as histone 3 lysine 4 trimethylation (H3K4me3) and acetylation of H3K9 and H3K27 (H3K9ac and H3K27ac). Mediator and cohesin/CTCF promote long-range interactions (arrow) with the transcription start site (TSS). This allows the assembly of the RNA polymerase II pre-initiation complex at the TATA-box (see text for details) and initiation of transcription. (**B**) Histone deacetylases remove acetyl groups from histone residues and polycomb repressive complexes 1 and 2 (PRC1 and PRC2) catalyse histone H2A lysine 119 ubiquitination (H2AK119ub) and histone 3 lysine 27 trimethylation (H3K27me3). This results in repressed chromatin with a higher nucleosome density, which makes the enhancer element inaccessible to TF binding and prevents the assembly of RNA polymerase II pre-initiation complex.

1.4 Background to this project

1.4.1 Notch signalling pathway

The Notch signalling pathway was first discovered in *Drosophila*, but is highly conserved across species and has an important role in developmental processes in embryonic and adult tissue, *e.g.* cell lineage decisions [reviewed in (Bray, 2006; Hori et al., 2013)]. RBPJ is the nuclear DNA binding protein of the Notch pathway thought to stably bind to Notch target genes (Bray, 2006).

In the absence of active Notch signalling, RBPJ represses these genes through interaction with TFIIA and TFIID (Olave et al., 1998), which prevents transcription, and also direct recruitment of repressor complexes containing HDAC1 and 2, SMRT/NcoR, SMRT/HDAC1-associated repressor protein (SHARP/MINT/SPEN), CBF1-interacting co-repressor (CIR), CtBP, CtBP-interacting protein (CtIP) and KyoT2 (Kao et al., 1998; Hsieh et al., 1999; Oswald et al., 2002; 2005; Taniguchi et al., 1998).

Activation of Notch signalling through ligand binding to the Notch receptors in the cell membrane induces a series of proteolytic cleavages of the Notch receptor, which result in the release of Notch intracellular domain (NICD) from the cell membrane (Mumm and Kopan, 2000; Fortini, 2002; Selkoe and Kopan, 2003). Cleaved NICD translocates into the nucleus where it binds to RBPJ via its RBPJ-associated molecule (RAM) domain WΦP (Φ = hydrophobic residue) motif (Tamura et al., 1995). Binding of NICD to RBPJ leads to the displacement of repressor complexes (Kao et al., 1998) and recruitment of the strong co-activators Mastermind (Petcherski and Kimble, 2000a; b; Wu et al., 2000) and HATs, *e.g.* p300 (Kurooka and Honjo, 2000; Oswald et al., 2001; Wallberg et al., 2002). This forms a stable activating complex and induces the expression of the repressed Notch target genes.

1.4.2 Gene regulation by EBNA2 and EBNA3 family

Neither EBNA2 nor EBNA3 proteins bind directly to DNA, but all bind to RBPJ. Therefore the following models of EBNA2-mediated gene activation and EBNA3-mediated gene repression were proposed (Figure 1.16). RBPJ alone stably binds to DNA and recruits co-repressors that inhibit gene expression (Figure 1.16A). Similar to NICD, EBNA2 binds to RBPJ via a RAM domain WWP motif, which is thought to displace co-repressors from RBPJ and to recruit co-activators through the transactivation domains of EBNA2 resulting in gene activation (Figure 1.16B). Therefore, EBNA2 operationally resembles NICD, but NICD cannot substitute for EBNA2 in maintaining LCL proliferation, so they are not completely equivalent [reviewed in (Zimber-Strobl and Strobl, 2001)].

EBNA3 proteins prevented EBNA2-mediated activation in transient reporter assays (Le Roux et al., 1994; Waltzer et al., 1996; Robertson et al., 1995), but only when they were able to bind to RBPJ (Radkov et al., 1997; Zhao et al., 1996). Concurrently, it was shown that EBNA3s compete with EBNA2 for binding to RBPJ and hence form mutually exclusive complexes with RBPJ (Robertson et al., 1995; Johannsen et al., 1996). Furthermore, binding of EBNA3s to RBPJ disrupted DNA binding of RBPJ *in vitro*, in electrophoretic mobility shift assays (Robertson et al., 1995; 1996; Waltzer et al., 1996). Therefore, a model was proposed in which binding of EBNA3s to RBPJ displaces EBNA2 and disrupts DNA-binding of RBPJ (Figure 1.16C), which therefore counteracts EBNA2-mediated gene activation (Robertson et al., 1995; 1996; Waltzer et al., 1996; Young and Rickinson, 2004; Hayward, 2004; Hayward et al., 2006). Later, a model was proposed in which EBNA2 is also replaced by EBNA3s, but RBPJ remains statically bound to DNA and activators are replaced by repressors [Figure 1.16D, (Harth-Hertle et al., 2013; Wang et al., 2015)]. For the replacement model, no net change in RBPJ occupancy was observed when EBNA2 was replaced by EBNA3s (Wang et al., 2015).

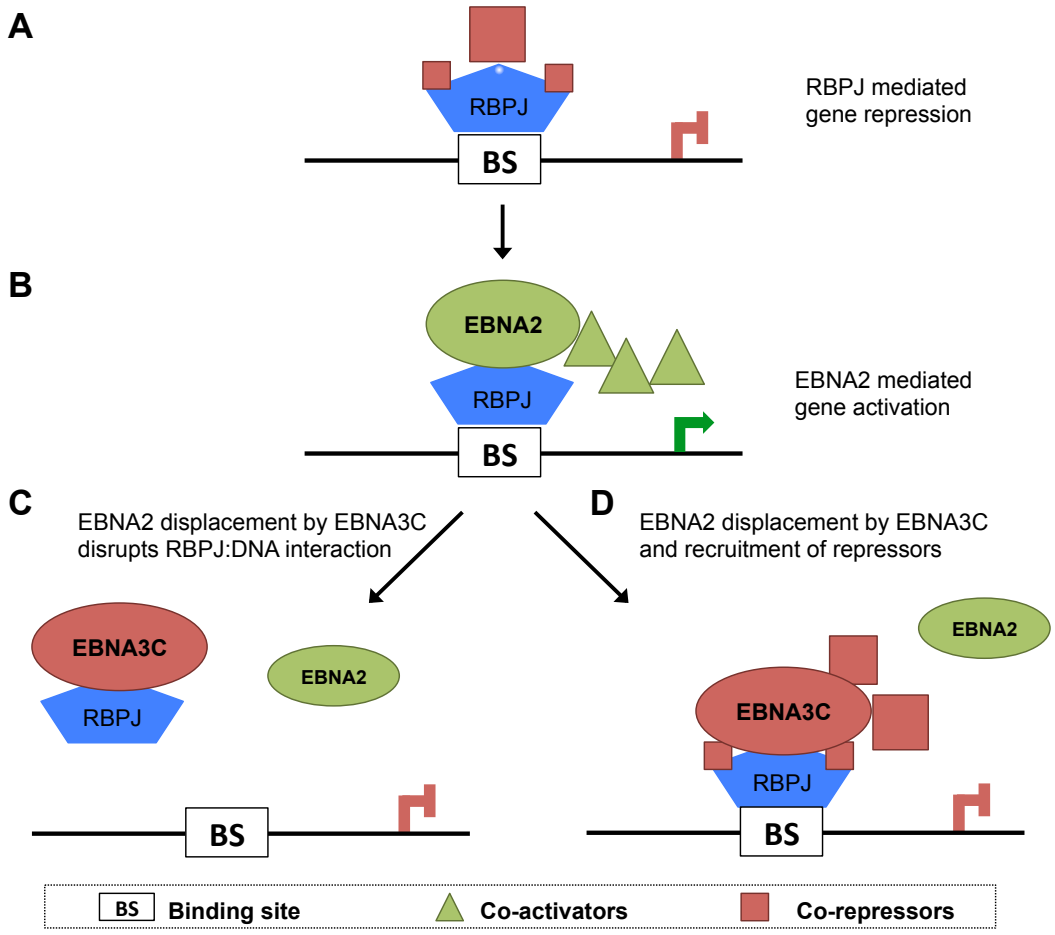


Figure 1.16: Model of gene regulation by EBNA2 and EBNA3C.
(**A**) RBPJ is bound to its binding site on DNA and recruits co-repressors that prevent gene expression. (**B**) EBNA2 binding to RBPJ replaces co-repressors and recruits co-activators that induce gene expression. (**C+D**) EBNA3C replaces EBNA2 as they bind to the same site on RBPJ. This was thought to either disrupt binding of RBPJ to DNA and thereby repressing gene expression through loss of co-activators (**C**) or to remain bound on DNA and to recruit co-repressors that actively inhibit gene expression (**D**).

However, these models are far from complete or thoroughly proven using EBV-recombinants and only describe EBNA2 as a potent transactivator and EBNA3 proteins as potent repressors without taking into account the fact that both can function as either repressor or activator. Therefore, it is fair to say that currently important aspects of the role of RBPJ in gene regulation by EBNA2 and EBNA3 proteins remain largely unknown.

Interestingly, all genes analysed in EBNA2-mediated gene activation (*e.g.* CD21 and CD23) have been found to be strictly dependent on the presence of RBPJ, whereas EBNA2-

mediated repression (IgM and BCL6) was only partially dependent on RBPJ (Maier et al., 2005; 2006; Boccellato et al., 2007). Moreover, EBNA3C-mediated gene activation of the viral LMP1 promoter was independent of RBPJ, but required PU.1/SPI1 binding sites (Zhao and Sample, 2000; Lin et al., 2002). Therefore, it is tempting to speculate that RBPJ is required for EBNA2-mediated gene activation and EBNA3-mediated gene repression (models above) and that other factors (*e.g.* PU.1/SPI1 to which both EBNA2 and EBNA3C can bind) are involved in EBNA2-mediated repression and EBNA3C-mediated activation.

Initial chromatin immunoprecipitation (ChIP) coupled to deep-sequencing (ChIP-Seq) experiments identified >5,000 EBNA2 binding sites and >10,000 RBPJ binding sites in IB4 LCL that carry a deletion of EBNA3B (Zhao et al., 2011a). The majority (72%) of EBNA2 binding sites overlapped with RBPJ binding sites, but only 14% were within 2 kb of a TSS. Furthermore, EBNA2 binding sites were enriched in motifs of other TF binding sites beside RBPJ, such as RUNX, early B cell factor (EBF) and PU.1. EBNA2 transactivation of the LMP1 promoter was shown to be dependent on an EBF binding site, because mutation of the EBF site strongly prevented EBNA2-mediated induction of the LMP1 promoter, whereas mutation of the previously identified PU.1 site completely abrogated EBNA2-induction (Zhao et al., 2011a).

More recently and only after this study began, ChIP-Seq for EBNA2 and EBNA3s in the latency III BL cell line Mutu III identified >20,000 EBNA2 binding sites and >7,000 EBNA3 binding sites, as determined using an antibody that recognizes all EBNA3 proteins (McClellan et al., 2013). There appeared to be a remarkable overlap between EBNA3 and EBNA2 binding sites, suggesting co-regulation by EBNA2 and EBNA3 proteins. Similar to results from Zhao *et al.*, EBNA2 sites were enriched in DNA-recognition motifs for EBF1 and RBPJ, while EBNA3C sites were enriched in motifs for PU.1. Again only a small proportion of EBNA2 and EBNA3 binding sites were found proximal to the TSS, but are frequently 10-50 kb away from the nearest TSS, suggesting the involvement of long-range chromatin interactions. McClellan and colleagues identified a binding site for EBNA2 and all EBNA3s

within the *CTBP2* gene, which encodes one of the two homologues of CtBP, and proposed that EBNA3 proteins repress *CTBP2* expression by preventing looping between the binding site and the *CTBP2* promoter, while the absence of EBNA3A was sufficient to allow looping and correlated with increase expression of *CTBP2*. The role of EBNA2 at the binding site is not fully understand, but was proposed to weaken repression by competing with EBNA3s for access to the binding site, because re-ChIP experiments (ChIP with a one antibody, but instead of reversing the cross-links performing a re-ChIP with either the same or another antibody) did not detect chromatin with co-bound EBNA2 and EBNA3. In contrast, at the *ADAM28-ADAMDEC1* locus, EBNA3A and EBNA3C were proposed to enable repressive chromatin looping between an intergenic EBNA3A and EBNA3C binding site and the promoters of *ADAM28* and *ADAMDEC1* (for more details see Chapter III: Gene repression by EBNA3C). In contrast, activating chromatin looping between enhancer elements and promoter has been shown during the epigenetic activation of the miR-221/miR-222 locus by EBNA3A and EBNA3C (Bazot et al., 2015).

Other ChIP-Seq studies using HA and Flag-tagged EBNA3A (Schmidt et al., 2014) or EBNA3C (Jiang et al., 2013) identified 10,000 EBNA3A and >13,000 EBNA3C binding sites, again predominantly at distal regulatory elements. In contrast to EBNA2, however, only 16% of EBNA3A and EBNA3C binding sites overlapped with RBPJ binding sites (8-9% with joint EBNA2/RBPJ sites and 7-8% with RBPJ only sites), as determined by reanalysis of previous RBPJ ChIP-Seq (Zhao et al., 2011a). Therefore, RBPJ does not appear to be the main TF that tethers EBNA3 proteins to DNA, which could either support the disrupted DNA-binding model (Figure 1.16C) or indicate that EBNA3 proteins are mostly tethered to DNA by other TFs. Importantly, most EBNA3A and EBNA3C sites overlapped with ENCODE TF ChIP-Seq binding sites for RUNX3 (63-64%), BATF (55-65%) and IRF4 (41-44%) in LCL GM12878. Interestingly, highest EBNA3A signals were found at RUNX3 or BATF-IRF4 sites and highest EBNA3C signals at BATF-IRF4 and SPI1-IRF4 sites (Schmidt et al., 2014; Jiang et al., 2013).

Similarly, in another ChIP-Seq analysis of all EBNA3 proteins approximately 1600 EBNA3A, 3000 EBNA3B and 3500 EBNA3C binding sites were identified (Wang et al., 2015). Wang and colleagues also performed ChIP-Seq for RBPJ and reanalysed previous EBNA2 and RBPJ ChIP-Seq data (Zhao et al., 2011a). RBPJ binding was observed at 34% of EBNA3A, 43% of EBNA3B and 44% of EBNA3C binding sites, which is slightly higher than 16% in previous reports for EBNA3A and EBNA3C (Schmidt et al., 2014; Jiang et al., 2013). Analysis of sites that display EBNA2, EBNA3 and RBPJ binding in cells with conditional EBNA3 proteins, showed that in the presence of functional EBNA3 proteins very little EBNA2 was found at these sites, whereas inactivation of EBNA3 proteins resulted in a switch to EBNA2 at these sites without affecting RBPJ occupancy, which supports the replacement model (Figure 1.16D). Consistent with the previous studies from Jiang *et al.* and Schmidt *et al.*, EBNA3 binding sites largely overlapped with ENCODE TF ChIP-Seq binding sites for multiple TF (*e.g.* RUNX3, EBF1, PAX5, BATF, SP1 and IRF4) in LCL GM12878. Overexpression of IRF4 in BJAB cells, which express little IRF4, together with EBNA3C could increase binding of EBNA3C at some selected EBNA3C-bound sites, but IRF4 did not seem to affect binding of EBNA3A or EBNA3B at selected EBNA3A- or EBNA3B-target sites, respectively.

In summary, at the beginning of this project, nothing was known about the genome-wide distribution of EBNA3 proteins. It was only during the course of this project that multiple ChIP-Seq studies emerged on EBNA3 proteins [(McClellan et al., 2013; Schmidt et al., 2014; Jiang et al., 2013; Wang et al., 2015) and unpublished ChIP-Seq from our lab] in different cell lines either expressing un-tagged, Flag-HA-tagged or Strep-Flag-tagged EBNA3 proteins. From these ChIP-Seq studies, it became apparent that EBNA2 and EBNA3 proteins target predominantly distal regulatory elements and probably influence gene expression by modulating higher-order chromatin structures, but some gene promoters are also targeted directly. RBPJ is thought to be the main factor for EBNA2 recruitment as the majority of EBNA2 and RBPJ binding sites overlap (Zhao et al., 2011a), whereas RBPJ was

only seen at a minority of EBNA3 sites (Schmidt et al., 2014; Jiang et al., 2013; Wang et al., 2015). Nevertheless, EBNA3s are still thought to compete with EBNA2 for binding to RBPJ and either disrupt DNA binding of RBPJ (Figure 1.16C) or replace EBNA2 (Figure 1.16D). However, in addition to RBPJ, other TFs such as RUNX3, BATF and IRF4 have been suggested to mediate DNA binding of EBNA3 proteins. Therefore, despite past and recent efforts in obtaining a better understanding of EBNA2 and EBNA3s in gene regulation, much is yet to be determined and our knowledge remains far from complete.

1.4.3 Available recombinant EBV

In order to study EBNA3C-mediated gene regulation, recombinant EBV were used during this project carrying modifications or defined mutations that are introduced in this section. B95-8 is the most commonly used EBV strain and in this study will be referred to as wild type (wt) strain. However, it is important to mention that this is a laboratory adapted strain that was selected for its transformation capacity of B cells in culture and has been passaged in marmoset lymphocytes. Compared to other EBV strains, B95-9 has a genomic deletion (Raab-Traub et al., 1980; Parker et al., 1990) that removed most of the EBV encoded miRs and one lytic origin of replication. B95-8 has been cloned into a BAC, which allowed the creation of recombinant EBV within *E. coli* (Delecluse et al., 1998).

Figure 1.17A shows an overview of recombinant EBV used during this project. Recombinant knock-out (KO) EBV were created that lack each single EBNA3 individually (3AKO, 3BKO or 3CKO) or all three EBNA3s (E3KO) and their corresponding revertant, in which the deleted fragment has been re-introduced to assay for off-target secondary mutations that might have been introduced during the recombination and would differ from the wt. All KO and revertant viruses have been described (Anderton et al., 2007; White et al., 2010). Fusion of tandem affinity purification (TAP) tags to the C-terminus of EBNA3A or EBNA3C introduced Flag and Strep tags that can be used to ChIP for the individual EBNA3 (Paschos et al., 2012). Conditional EBNA3C (3CHT) contains a C-terminal fusion to a modified murine estrogen receptor that renders EBNA3C conditional to the presence of 4-

hydroxytamoxifen (HT) in the culture medium (Skalska et al., 2010). In the absence of HT, EBNA3C is retained in the cytoplasm where it is inactive and degraded, whereas addition of HT to the culture medium allows the translocation of EBNA3C into the nucleus where it is stable and active. Furthermore, the RBPJ BM EBNA3C recombinant EBV has been created containing the $_{209}$AAAA and W227S (Figure 1.10B and C) mutations that almost completely prevent interaction between EBNA3C and RBPJ (Kalchschmidt et al., 2016).

In order to produce infectious EBV virions (Figure 1.17B), these BACs are transfected into HEK293 cells and virus producer cell lines are established under hygromycin selection, which is encoded on the EBV BAC. Induction of the lytic cycle results in the production and secretion of infectious EBV recombinant virions into the supernatant, which can be used to infect B cells.

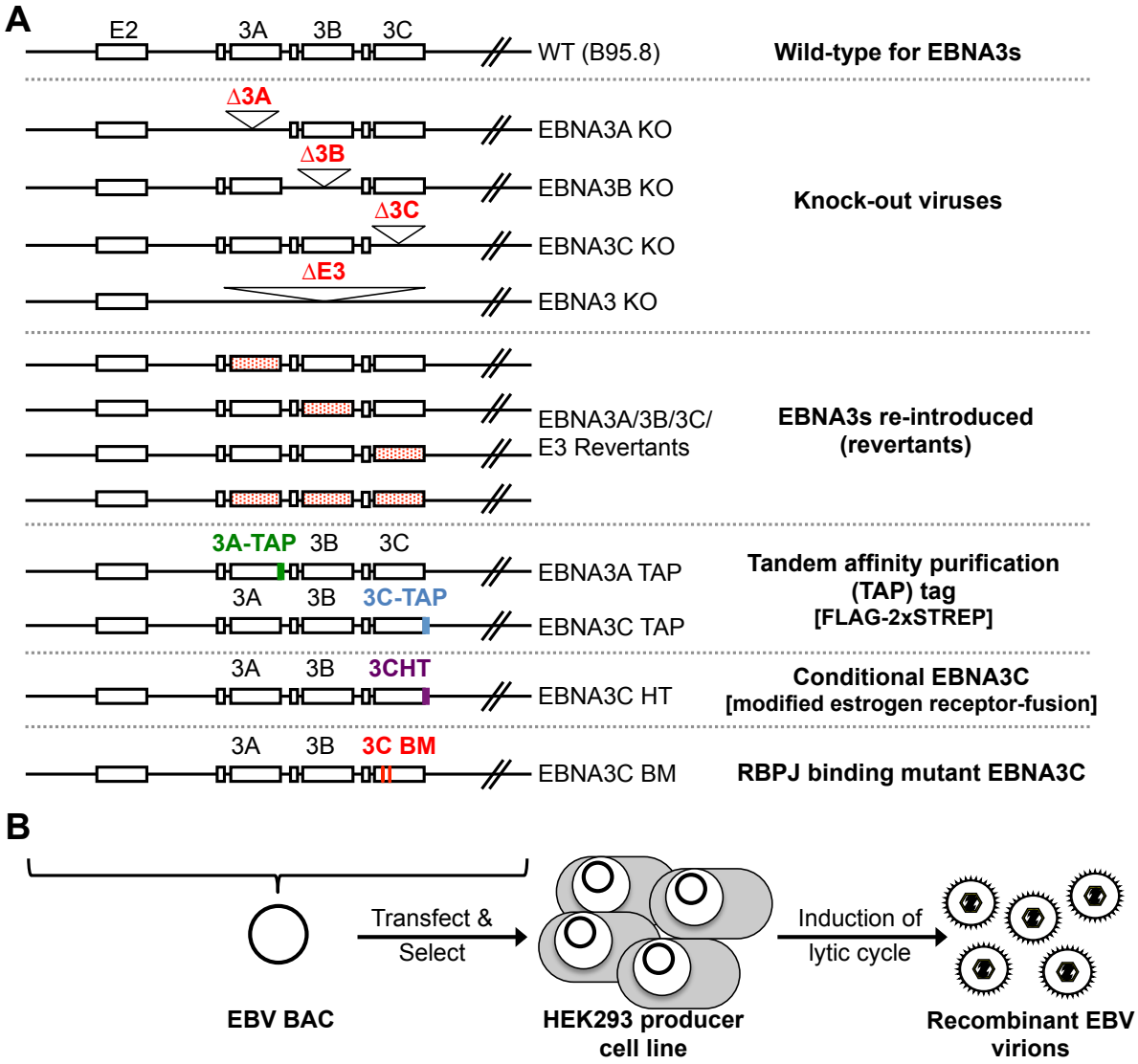


Figure 1.17: Overview of recombinant EBV available and production of infectious EBV virions. (**A**) Overview of recombinant EBV derivations from wild type B95.8 BAC. Knock-out viruses are deleted for either EBNA3A (Δ3A), EBNA3B (Δ3B), EBNA3C (Δ3C) or all three EBNA3s (ΔE3) and revertants contain reintroduction of previously deleted segments [described in (Anderton et al., 2007) and (White et al., 2010)]. Tandem affinitiy purification (TAP) are fused to the C-terminus of EBNA3A or EBNA3C and contain a Flag-2xStrep tags that can be used in ChIP [described in (Paschos et al., 2012)]. Conditional EBNA3C contains a C-terminal fusion to modified estrogen receptor [described in (Skalska et al., 2010)]. RBPJ BM EBNA3C recombinant EBV has been created containing the $_{209}$AAAA and W227S (Figure 1.10B) mutations that almost completely prevent interaction between EBNA3C and RBPJ [described in (Kalchschmidt et al., 2016)].(**B**) Recombinant EBV BACs are transfected in HEK293 cells and selected to establish a HEK293 producer cell line. Induction of the lytic cycle leads to the production and release of recombinant EBV virions into the supernatant.

1.4.4 Identification of EBNA3C target genes

In order to select host genes that are potential direct targets of EBNA3C, a previous exon-microarray using the 3CHT conditional virus in a p16-null background was interrogated (Skalska et al., 2013). Due to the absence of p16 in these cells cellular proliferation does not decrease in the absence of functional EBNA3C, which allows direct comparison between cells with inactive and active EBNA3C.

Table 1.1 and Table 1.2 list the top ten genes that are repressed or activated by EBNA3C. *COBLL1* and *AICDA* are the two genes most repressed or activated by EBNA3C, respectively, and neither was previously reported as regulated by EBNA3C. Therefore, both genes were selected. However, nothing was known about the binding of EBNA3C to either locus and it was possible that both genes were indirect targets of EBNA3C. In contrast, *ADAM28* and *ADAMDEC1*, the 2nd and 3rd most highly repressed genes by EBNA3C, were previously reported as regulated by EBNA3A and EBNA3C (see Section 3.1) and a study published shortly before the start of this project identified an EBNA3 binding site located in between the two adjacent *ADAM28* and *ADAMDEC1* genes (McClellan et al., 2012). For these reasons, the EBNA3C-repressed genes *COBLL1*, *ADAM28* and *ADAMDEC1* and the EBNA3C-activated gene *AICDA* were selected for investigation in this PhD project.

Table 1.1: Top 10 genes repressed by EBNA3C

#	Gene name	Genomic location *	p-value	Fold change
1	COBLL1	chr2: 165,536,695 - 165,698,678	1.300E-12	-63.57
2	ADAM28	chr8: 24,151,553 - 24,216,531	7.950E-08	-20.68
3	ADAMDEC1	chr8: 24,241,798 - 24,263,526	1.399E-06	-17.36
4	RP11-624C23.1	chr8: 24,158,398 - 24,769,586	8.995E-06	-13.69
5	ANKMY2	chr7: 16,639,401 - 16,685,442	4.309E-07	-8.60
6	TREM1	chr6: 41,242,999 - 41,254,457	4.797E-04	-6.40
7	TSPAN13	chr7: 16,793,351 - 16,824,161	4.265E-09	-5.65
8	PPM1L	chr3: 160,473,996 - 160,788,817	1.417E-07	-4.04
9	MUC15	chr11: 26,580,579 - 26,593,815	6.294E-05	-3.98
10	EML5	chr14: 89,081,174 - 89,259,096	2.735E-06	-3.89
* genomic location of RefSeq isoform 1 in case of multiple variants				

Table 1.2: Top 10 genes induced by EBNA3C

#	Gene name	Genomic location *	p-value	Fold change
1	AICDA	chr12: 8,754,762 - 8,765,442	2.852E-07	4.82
2	TOX	chr8: 59,717,977 - 60,031,767	2.573E-06	3.83
3	RP4-678E16.4	chr1: 45,190,179 - 45,191,263	2.366E-04	3.41
4	MOXD1	chr6: 132,617,194 - 132,722,673	1.553E-06	3.31
5	GLYATL2	chr11: 58,601,540 - 58,611,997	1.003E-05	3.21
6	ASCL1	chr12: 103,351,452 - 103,354,294	3.206E-04	3.21
7	SLCO4C1	chr5: 101,569,692 - 101,632,253	2.223E-05	3.17
8	HNRPLL	chr2: 38,790,328 - 38,830,178	6.349E-05	3.15
9	SPRED2	chr2: 65,537,985 - 65,659,656	2.810E-06	3.12
10	CD1C	chr1: 158,259,563 - 158,264,564	2.432E-04	3.09
* genomic location of RefSeq isoform 1 in case of multiple variants				

1.5 Aims of this PhD project

Our current understanding of the mechanisms underlying EBNA3C-mediated gene regulation is very limited. Therefore, the aims of this PhD project were to analyse:

1. Mechanisms involved in EBNA3C-mediated repression of *COBLL1*, *ADAM28* and *ADAMDEC1*
2. Mechanisms involved in EBNA3C-mediated activation of *AICDA*

2 Chapter II: Materials and Methods

2.1 Solutions and buffers

All solutions were prepared in Milli-Q-H_2O, unless stated otherwise.

Alkaline SDS

1 % (w/v)	SDS
0.2 M	NaOH

Blocking solution for western blot

5 % (w/v) Skim milk powder (Sigma) in PBS-Tween

ChIP SDS Lysis buffer (Millipore)

50 mM	Tris, pH8.1
10 mM	EDTA
1 % (w/v)	SDS

ChIP Dilution buffer (Millipore)

16.7 mM	Tris-HCl, pH 8.1
167 mM	NaCl
1.2 mM	EDTA
1.1 % (v/v)	Triton-X-100
0.01 % (w/v)	SDS

ChIP swelling buffer

25 mM	HEPES, pH 7.8	
1.5 mM	$MgCl_2$	
10 mM	KCl	
0.1% (v/v)	NP-40	
1 mM	DTT	(add immediately prior to use)

ChIP sonication buffer

50 mM	HEPES, pH 7.8
140 mM	NaCl
1 mM	EDTA, pH 8
1 % (v/v)	Triton X-100
0.1 % (w/v)	Sodium deoxycholate
0.1 % (w/v)	SDS

ChIP low salt buffer (Millipore)

20 mM	Tris-HCl, pH 8
150 mM	NaCl
2 mM	EDTA, pH 8
1 % (v/v)	Triton X-100
0.1 % (w/v)	SDS

ChIP high salt buffer (Millipore)

20 mM	Tris-HCl, pH 8
500 mM	NaCl
2 mM	EDTA, pH 8
1 % (v/v)	Triton X-100
0.1 % (w/v)	SDS

ChIP LiCl buffer (Millipore)

10 mM	Tris-HCl, pH 8
250 mM	LiCl
1 mM	EDTA, pH 8
1 % (v/v)	NP40
1 % (w/v)	Sodium deoxycholate

DNA sample loading buffer (6x)

30% (v/v)	Glycerol
0.15% (w/v)	Bromophenol blue

β-estradiol

1 mM β-estradiol (1,000x) resuspended in 100% ethanol was a gift from Claudio Elgueta Karstegl (Paul Farrell laboratory) and was kept at 4°C in the dark.

4-Hydroxytamoxifen

4mM (10,000x) stock was prepared by resuspending 5mg of 4-hydroxytamoxifen (Sigma) in 3.22 ml of 100% ethanol. Aliquots were stored at -20°C in the dark.

KPB (20x)

4.6 % (w/v)	KH_2PO_4
24.3 % (w/v)	K_2HPO_4

LB

1% (w/v)	Bacto-tryptone
0.5% (w/v)	Yeast extract
1% (w/v)	NaCl

Autoclave and store at room temperature.

LB-Agar

1% (w/v)	Bacto-tryptone
0.5% (w/v)	Yeast extract
1% (w/v)	NaCl
1.5% (w/v)	Agar

Autoclave and store at room temperature.

PBS

10 mM	Phosphate buffer, pH 7.4
2.7 mM	KCl
137 mM	NaCl

PBS-Tween

0.1 % Tween-20 (Sigma, UK) in 1xPBS

Peptide 6

Re-suspend 50 mg peptide 6 in 50ml OptiMEM I (Life technologies) in autoclaved flask under sterile conditions and leave covered with aluminium foil overnight in fridge, before aliquoting in sterile cryovials and storage at -80°C.

PMFS

100 mM of PMSF resuspended in isopropanol and stored at -20°C

RIPA (5x)

250 mM	Tris-HCl, pH 8.0
750 mM	NaCl
5% (v/v)	NP40
2.5% (w/v)	Sodium deoxycholate
0.5% (v/v)	SDS

Autoclave and store at room temperature.

SDS-PAGE running buffer (10x)

250 mM	Tris
1.92 M	Glycine
1% (w/v)	SDS

SDS-PAGE sample loading buffer (5x)

250mM	Tris-HCl, pH6.8
50% (v/v)	Glycerol
10% (w/v)	SDS
5% (v/v)	β-mercaptoethanol
0.1% (w/v)	Bromophenol blue

SDS-PAGE transfer buffer

10%	10x SDS-PAGE running buffer
23.3% (v/v)	Ethanol

STET

8 % (w/v)	Sucrose
0.1 % (v/v)	Triton-X-100
50 mM	EDTA, pH 8
50 mM	Tris, pH 8

Superbroth

1.2 % (w/v)	Bactotryptone
2.4 % (w/v)	Yeast extract
0.4 % (v/v)	Glycerol

Autoclave and store at room temperature. Add KPB before usage.

5 % (v/v)	20x KPB (sterile filtered)

10x TBE buffer

0.89 M	Tris
0.89 M	Boric acid
20 mM	EDTA

TE buffer

10 mM	Tris-HCl, pH 8
1 mM	EDTA, pH 8

2.2 Cell culture

2.2.1 Cell lines and culture conditions

All cell lines were cultured in RPMI 1640 medium (Life Technologies) supplemented with 10% FBS (Gibco), penicillin and streptomycin except for the following adjustments. Adherent HEK293 virus producers were grown in the presence of additional 2 mM L-Glutamine (Life Technologies) and 100 µg/ml hygromycin B (Roche). Newly infected primary B cells were cultured in the presence of 15% FBS. LCLs with conditional EBNA3C were grown either in absence or presence of 400nM HT. EREB 2-5 cells with conditional EBNA2 were cultured in the presence of 1 µM β-estradiol. All cells were incubated at 37°C and 10 % CO_2 in a humidified incubator. Suspension cells were passaged two to three times per week by diluting appropriately in fresh warmed medium. Adherent cell lines (HEK293) were split twice weekly. For this, cells were washed once in warm PBS and incubated with 0.05% trypsin/EDTA (Invitrogen) at 37°C until cells detached. Fresh medium was added and cells were transferred into a new flask depending on the density of the parental flask. Table 2.1 lists all cell lines used in this project.

Table 2.1: Cell lines used in this project

Cell line name	Cell type	Description	Reference
HEK293 wtI6	HEK293	Wild type EBV producer	
Del3A-M11 30-3	HEK293	3AKO EBV producer	(Anderton et al., 2007)
HB9-r3A-M4	HEK293	3ARev EBV producer	
Del3B-C14	HEK293	3BKO EBV producer	
HB9-r3B-M8	HEK293	3BRev EBV producer	
Del3C-M12 30-3	HEK293	3CKO EBV producer	
HB9-r3C-M15	HEK293	3CRev EBV producer	
E3KO-C34	HEK293	E3KO EBV producer	(White et al., 2010)
E3Rev-C2	HEK293	E3Rev EBV producer	
LCL JK1 wtI6	LCL	Wild type I6 from 1°B infection	This study
LCL JK1 3AKO M11 30-3	LCL	3AKO from 1°B infection	
LCL JK1 3ARev M4	LCL	3ARev from 1°B infection	
LCL JK1 3BKO C14	LCL	3BKO from 1°B infection	
LCL JK1 3BRev M8	LCL	3BRev from 1°B infection	
LCL JK1 3CKO M12 30-3	LCL	3CKO from 1°B infection	
LCL JK1 3CRev M15	LCL	3CRev from 1°B infection	
LCL D11 wt	LCL	Wild type in D11	
LCL D11 04 AC1	LCL	TAP-tagged EBNA3C in D11	(Paschos et al., 2012)
LCL D11 05 AC21	LCL	TAP-tagged EBNA3A in D11	
LCLΔp16 3CHT A13 never	LCL	$p16^{INK4A}$-null, EBNA3C-conditional	(Skalska et al., 2013)
LCLΔp16 3CHT C19C2 washed	LCL	$p16^{INK4A}$-null, EBNA3C-conditional	
EREB 2-5	LCL	Conditional EBNA2	(Kempkes et al., 1995)
GM12878	LCL	Tier 1 ENCODE	
DG75	BL	EBV-negative	(Ben-Bassat et al., 1977)
DG75 RBPJ KO	BL	EBV-negative, lacking RBPJ	(Maier et al., 2005)
Ramos	BL	EBV-negative	(Klein et al., 1975)
Raji	BL	EBV-positive, EBNA3C-null	(Epstein et al., 1966; Allday et al., 1993)
Akata	BL	EBV-positive, latency I	(Takada et al., 1991)
Oku	BL	EBV-positive, Wp-restricted	(Kelly et al., 2002)
Namalwa	BL	EBV-positive, latency III	(Klein et al., 1972)
BJAB	B cell lymphoma	EBV-negative, initially described as BL, but lacking the cMYC:Ig translocation (Klein et al., 1976; Glazer and Summers, 1985)	(Menezes et al., 1975)
K562	CML	EBV-negative	(Lozzio and Lozzio, 1975)

2.2.2 Freezing and thawing cells

For long-term storage of cells in liquid nitrogen, suspension cells were split the day before freezing and adherent cell lines were frozen at the time of passaging. For this, cells were pelleted at 1300 rpm for 4 min and 4°C (Megafuge 40R, Thermo Scientific) and re-suspended in RPMI 1640 medium (Life Technologies) supplemented with 20 % FBS (Gibco) and 10 % DMSO (VWR Chemicals). Cells were transferred into sterile cryovials, placed into a Nalgene Mr. Frosty™ Freezing Container filled with isopropanol, stored at -80°C overnight before being transferred to liquid nitrogen.

To recover cells from liquid nitrogen, cryovials were rapidly thawed in a water bath at 37°C and slowly added into 10 ml of culture medium. Cells were pelleted, the medium aspirated to remove the DMSO and the cells resuspended in fresh medium. Recovered cells from liquid nitrogen were cultured for at least one week before the start of any experiment.

2.2.3 Harvesting cells

For all experiments, unless stated otherwise, on the day prior to harvesting, viable cell counts were performed under a light microscope using a Neubauer improved haemocytometer. For this, cells were diluted 1:1 with 0.4% trypan blue (Sigma) and live cells that do not incorporate the blue dye were counted. Suspension cells were seeded at a density of $3x10^5$ cells/ml. The next day, for RNA or protein extraction, cells were pelleted at 1300 rpm for 4 min and 4°C (Megafuge 40R, Thermo Scientific) and washed once with ice cold PBS. The supernatant was aspirated, cell pellets resuspended in 1 ml ice cold PBS and transferred into 1.5 ml tubes. Cells were centrifuged at 2000 rpm for 5 min (5424R, Eppendorf) and the supernatant removed to obtain dry cell pellets. Samples for protein extraction were immediately snap frozen on dry ice. For RNA extraction, dry cell pellets were lysed by vortexing in RLT buffer containing 1% (v/v) β-mercaptoethanol and snap frozen on dry ice. Samples were stored at -80°C until further extraction.

2.2.4 Time-course experiments

2.2.4.1 EBNA3C-conditional LCL

EBNA3C conditional LCLs (3CHT) on the p16-null background were used for multiple time-course experiments (Skalska et al., 2013).

LCLΔp16 3CHT A13 never (referred to as 3CHT A13) was established in the absence of HT and used in an initial time-course experiment over 60 days with samples harvested for RNA, protein and ChIP every three days over the first 30 days and every ten days until day 60. For this, cells were counted and diluted to $3x10^5$ cells/ml at every time point until day 30 and three times a week subsequently, but seeded at $3x10^5$ cells/ml the day prior to harvesting. After harvesting cells on day 30, half of the +HT culture was centrifuged and the medium replaced by fresh medium without HT and cultured without HT until day 60 (washed).

Replicate time-courses using again 3CHT A13 but also LCLΔp16 3CHT C19C2 washed (referred to as 3CHT C19) were used in time-course experiments over 30 days. 3CHT C19 were established in the presence of HT, washed and subsequently grown without HT in the medium for more than three months before the experiment was started. Samples for RNA, protein and ChIP were harvested on day 0, 3, 6, 9, 15 and 30, but this time, cells were counted and split to $3x10^5$ cells/ml one day before harvesting.

2.2.4.2 EBNA2-conditional LCL

EREB 2-5 cells are LCLs that were established by co-infection of umbilical cord blood B cells with the EBNA2-deleted P3HR1 EBV strain and a recombinant mini-EBV genome carrying an estrogen receptor-EBNA2 fusion protein, which renders EBNA2 conditional to the presence of β-estradiol in the medium (Kempkes et al., 1995). Cells were cultured in the presence of 1 μM β-estradiol and samples for RNA, protein and ChIP were taken. EBNA2 was inactivated by removing β-estradiol from the media by washing the cells twice with β-estradiol-free medium before re-suspending the cells in β-estradiol-free medium. RNA,

protein and ChIP samples were taken on day 1, day 2 and day 3 after removal of β-estradiol. EBNA2 was re-activated through re-addition of β-estradiol to the cells and additional samples were taken 2, 4, 8, 24 and 48 hours post re-addition.

2.3 Validation of HEK293 EBV producing cell lines and virus production

2.3.1 Validation of HEK293 EBV producing cell lines by episomal rescue

Before producing recombinant EBV from HEK293 virus producing cell lines (Anderton et al., 2007), the identity and integrity of the recombinant EBV BACs was validated for each cell line. For this, recombinant EBV BACs were extracted from HEK293 virus producing cell lines, transformed into *E. coli*, extracted from single *E. coli* colonies, restriction enzyme digested and analysed by pulsed-field gel electrophoresis.

2.3.1.1 Extraction of EBV BACs from HEK293 virus producers

During the passaging of HEK293 virus producing cell lines, cells were washed in PBS and pelleted. The cell pellet was re-suspended thoroughly in 60 µl STET and cells were lysed by addition of 130 µl alkaline SDS while vortexing. Lysis was stopped by neutralisation with 110 µl 7.5 M ammonium acetate, again added while vortexing. The mixture was incubated on ice for 5 min and centrifuged at 13,000 rpm and 4°C for 30 min. The supernatant was transferred into a 1.5 ml MaXtract High Density Gel tube (Qiagen), 200 µl of Phenol:Chloroform:Isoamyl alcohol (25:24:1, Sigma) was added and the tube rapidly inverted for 30 sec. The mixture was centrifuged at 10,000 rpm for 6 min and the upper aqueous phase was transferred into a new MaXtract High Density Gel tube using cut-off tips to avoid shearing of the episomes. DNA was washed with 200 µl chloroform as described above and the aqueous phase was transferred into a fresh 1.5 ml tube after centrifugation. DNA was precipitated by addition of 670 µl of 100% ethanol and 20 min centrifugation at maximum speed. Precipitated DNA was washed with 200 µl of 70% ethanol and 5 min centrifugation at maximum speed. The DNA pellet was air-dried for 15 min, re-suspended in 50 µl TE buffer containing 5 µg/ml RNase A and kept at 4°C.

2.3.1.2 Transformation of recombinant EBV BACs into E. coli

Cut-off tips were used to add 1 µl of the extracted EBV BACs to 20 µl of Electromax electrocompetent DH10B *E. coli* (Invitrogen). Bacteria were kept on ice and electroporated at 1.8 kV, 25 µFD and 200 Ω using Gene pulser (BioRad). Electroporated bacteria were recovered in 400 µl of SOC (Invitrogen) and incubated at 37°C for one hour on an orbital shaker. Bacteria were plated on LB agar containing chloramphenicol and incubated overnight at 37°C.

2.3.1.3 Extraction of EBV BACs from E. coli

Single *E. coli* colonies were used to inoculate 1.5 ml superbroth (SB) containing chloramphenicol and grown at 37°C overnight on an orbital incubator. Bacteria were pelleted at 2,500 rpm for 5 min and room temperature and were re-suspended in 70 µl STET by vortexing. Bacteria were lysed with 200 µl of alkaline SDS (added while vortexing), which was neutralised by addition of 150 µl of 7.5 M ammonium acetate (again added while vortexing). Samples were incubated for 5 min on ice and centrifuged at 13,000 rpm and 4°C for 20 min. The supernatant was transferred into a new 1.5ml tube and DNA was precipitated by addition of 240 µl isopropanol and inversion of the tube. DNA was pelleted at 10,000 rpm for 6 min and room temperature and washed in 200 µl of 70 % ethanol. The pelleted DNA was air-dried for 10 min and re-suspended in 50 µl TE containing 5 µg/ml RNase A. Re-suspended DNA was stored at 4°C until analysis by restriction enzyme digestion.

2.3.1.4 Restriction enzyme digestion

In a 15 µl reaction, 5 µl of recovered episomal DNA was digested with *Eco*RI (assessment of episome integrity) and *Cla*I (assessment of identity of virus producer) using standard New England BioLabs restriction enzyme digestion protocols.

2.3.1.5 Pulsed-field gel electrophoresis

CHEF-DR II system (BioRad) was used to run pulsed-field 1 % agarose gel electrophoresis. Agarose was dissolved in 0.5x TBE and gels were run in 2 litres of 0.5x TBE. λ DNA-BstEII digest (NEB) and λ DNA-Mono cut mix (NEB) were used in a 1:1 mixture as DNA size ladder and heated to 50°C for 10 min before loading onto the gel. Digested episomes were mixed with 6x agarose gel loading buffer and run alongside the DNA size ladder for 13 hours and 14°C with 6V/cm and 1 sec initial sweep time and 10 sec final sweep time. The gels were stained with 0.5 µg/ml ethidium bromide in 0.5x TBE for 30 min with gentle agitation and visualised on a Gel Doc XR+ system (BioRad).

2.3.2 Virus production

Once the integrity of the BACs and the identity of the HEK293 virus producing cell lines were confirmed, infectious viral particle were produced.

2.3.2.1 Transfection and harvest

Virus production was induced by transfecting pCMV-EB1 and pBALF4 expression plasmids (from Dr Rob White) using the LID (lipofectin/integrin-targeting peptide/DNA) method (Hart et al., 1998). HEK293 virus producer cell lines were seeded in a 10 cm dish. Two days later, 9 µl Lipofectin Reagent (Invitrogen) and 480 µl Peptide 6 were mixed and pre-incubated for at least 5 min. pCMV-EB1 and pBALF4 plasmids (6 µg each) were dissolved in 1.2 ml OptiMEM I (Invitrogen) and added to the Lipofectin/Peptide 6 solution. This LID mixture was incubated for at least 20 min before adjusting the total volume to 5 ml with OptiMEM I. The medium of the seeded HEK293 virus producer cell lines was replaced by the 5 ml LID mix and cells were incubated with the LID mix for at least 6 hours at 37°C, after which the transfection mix was replaced by 10 ml of fresh medium without hygromycin selection and cells incubated for four days. Supernatants containing viral particles were then harvested through MF-Millipore 0.45 µm filters (Millex-HA, Millipore).

2.3.2.2 Virus titration

The virus titre of the supernatants was determined by infection of Raji cells and counting GFP expressing cells [green Raji units (GRU), (Dirmeier et al., 2003)] on an inverted fluorescent microscope 3 days after infection and 24 hours after treatment with 20 nM 12-O-tetradecanoylphorbol-13-acetate and 5 mM sodium butyrate.

2.4 Isolation and infection of primary B cells

2.4.1 Isolation of primary B cells

Peripheral blood mononuclear cells (PBMC) were isolated from buffy coat residues (UK national blood transfusion service) by centrifugation over Ficoll-Paque. To isolate primary B cells, 10^9 PBMC were incubated with 500 µl CD19 microbeads (Miltenyi Biotec) and separated using AutoMACS Separator (Miltenyi Biotec). Purified B cells were re-suspended at $2x10^6$ cells/ml in RPMI/15% FBS medium and incubated at 37°C and 5% CO_2. The next day, B cell purity was assessed using anti-CD20-APC (eBioscience) staining and flow cytometric analysis.

2.4.2 Infection of primary B cells with recombinant EBV

Three million primary B cells in 1.5 ml were infected with 0.5 ml of supernatant containing the following EBV viruses: wild type I6 ($1.1x10^5$ GRU), EBNA3A KO ($2.7x10^5$ GRU), EBNA3A Rev ($3.5x10^5$ GRU), EBNA3B KO ($2.0x10^5$ GRU), EBNA3B Rev ($1.4x10^5$ GRU), EBNA3C KO ($1.1x10^5$ GRU), EBNA3C Rev ($1.9x10^5$ GRU), EBNA3A/B/C KO ($3.4x10^5$ GRU), and EBNA3A/B/C Rev ($3.9x10^5$ GRU). RNA from three million uninfected primary B cells was taken on the day of infection and infected cells were incubated at 37°C and 5% CO_2. For a period of 30 days, every three days 0.5 ml of cells were harvested for RNA extraction and replaced by fresh medium. After this, cells were, where possible, grown into LCL and analysed by western blotting.

2.5 Reverse transcription quantitative PCR (RT-qPCR)

Total RNA was isolated either from freshly harvested cells, from frozen cell lysates in RLT buffer stored at -80°C or cell pellets stored at -80°C. For cells harvested during the course of the primary B cell infections, total RNA including miRNA was extracted using miRNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. Otherwise, RNA was extracted using RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. Both extractions included the optional on column digestion of genomic DNA using RNase-free DNase set (Qiagen). RNA was eluted in RNase-free water (Qiagen) and extracted RNA was quantified using a Nanodrop Lite spectrophotometer (Thermo Scientific).

Reverse transcription was performed on up to one microgram of RNA using SuperScript III first strand synthesis super mix (Invitrogen) according to the manufacturer's instructions and at a cDNA synthesis temperature of 59°C. cDNA was RNase H treated and diluted with RNase-free water (Qiagen) to a final concentration of 5 ng/µl, assuming complete conversion of RNA into cDNA.

For the quantitative PCR (qPCR) 10 ng of cDNA was used in a final reaction volume of 10 µl containing 5 µl Platinum SYBR Green qPCR SuperMix-UDG kit (Invitrogen) and 200nM of RT-qPCR primers (Table 2.2). cDNA was assayed by qPCR on QuantStudio 7 Flex (Life technologies). The qPCR programme included an initial denaturation step for 2 min at 95°C, followed by 40 qPCR cycles with denaturation for 3 sec at 95°C and primer annealing and elongation for 30 sec at 60°C. Dissociation curves were used to assess the specificity of PCR products. Efficiency of the primer was initially assed by serial dilution of pooled cDNA samples. As negative qPCR controls, reverse-transcribed RNase-free water (Qiagen) was used with the same qPCR Master mix as the samples. Each sample was run in triplicates and *GAPDH* or *GNB2L1* were used as endogenous controls. Relative gene expression was calculated using the $2^{-\Delta\Delta Ct}$ method (Schmittgen and Livak, 2008) with $\Delta\Delta Ct = (Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})_{\text{Sample A}} - (Ct_{\text{gene of interest}} - Ct_{\text{endogenous control}})_{\text{Sample B}}$.

Table 2.2: Sequences of RT-qPCR primers

Gene	Forward primer (5'-3')	Reverse primer (5'-3')	Reference
ADAM28	GTTGCAGGGACAATGGCACA	TGAGACGGCTGCAGGAACTG	This study
ADAMDEC1	CCTTGGTATGCCTGATGTTCCA	CAGCAGGCACTTTGGTTTCTGA	
AICDA	AGCCGTTCTTATTGCGAAGA	TGATGAACCGGAGGAAGTTT	
ALAS1	TCCACTGCAGCAGTACACTACCA	ACGGAAGCTGTGTGCCATCT	(Bazot et al., 2015)
BCL6	CTGCAGATGGAGCATGTTGT	TCTTCACGAGGAGGCTTGAT	(Tran et al., 2010a)
COBLL1	CTGTTCAGCTGACAACAGATCG	ACGTTGAACTCTCAGTGGTCCT	This study
GAPDH	Hs_GAPDH_2_SG QuantiTect Primer Assay, QT01192646		Qiagen
GNB2L1	GCTTGCAGTTAGCCAGGTTC	GAGTGTGGCCTTCTCCTCTG	(Zhang et al., 2005)
miR155-HG	GAGACGCTCCTGGCACTGC	TAAACCACAGATTTCCCCTTCCTGG	(van den Berg et al., 2003)

2.6 Western blotting

Protein was extracted from freshly harvested or frozen cell pellets stored at -80°C using 1x RIPA buffer containing 1 mM PMSF and complete protease inhibitor cocktail (Roche). Lysates were left on ice for 5 min and cleared by centrifugation at 14,000 rpm for 15 min at 4°C. The supernatant was transferred into a fresh tube. The protein concentration was determined with the DC protein assay (Bio-Rad) and measured on a Helios spectrophotometer (Thermo Scientific) against bovine serum albumin as standard. Generally, 30 µg of total protein of each sample were mixed with SDS-PAGE sample loading buffer (1x final concentration) and heated to 100°C for 5 min. Gels with an appropriate percentage for the molecular weight of the protein to be analysed were prepared according to Table 2.3. Spectra™ Multicolor Broad Range Protein Ladder (Thermo Scientific) was used on each gel as reference for the molecular weight and protein samples were separated for one hour at 150V using Mini-PROTEAN® II Electrophoresis Cell (Bio-Rad) containing 1xSDS-PAGE running buffer. Samples were transferred on Whatman nitrocellulose membranes (GE Healthcare) at 100V for one hour in Mini Trans-Blot Module (Bio-Rad)

containing SDS-PAGE transfer buffer. Membranes were blocked in SDS-PAGE blocking solution for at least 15 min. Primary antibodies were diluted in SDS-PAGE blocking solution as shown in Table 2.4 and incubated over night at 4°C with shaking. Membranes were washed three times for 5-10 min with PBS-Tween and incubated at room temperature for one hour with appropriated secondary antibodies diluted in SDS-PAGE blocking solution (Table 2.5). Again, membranes were washed three times for 5-10 min with PBS-Tween. Membranes were briefly incubated with Amersham ECL western blotting detection reagents (RPN2106V1 and RPN2106V2, GE Healthcare), exposed on Amersham Hyperfilm MP (GE Healthcare) and developed on automatic X-ray film processor JP-33 (JPI).

Table 2.3: Composition of resolving and stacking gels for SDS-PAGE

Resolving gel	**7.5 %**	**10 %**	**12 %**	**Stacking gel**	
1 M Tris, pH 8.8	2.96 ml	2.96 ml	2.96 ml	1 M Tris, pH 6.8	620 µl
30 % Acrylamide/Bis solution 37.5:1	1.98 ml	2.64 ml	3.3 ml	30 % Acrylamide/Bis solution 37.5:1	828 µl
ddH_2O	2.96 ml	2.3 ml	1.64 ml	ddH_2O	3.47 µl
10 % SDS	79.2 µl	79.2 µl	79.2 µl	10 % SDS	49.6 µl
15 % APS	26.4 µl	26.4 µl	26.4 µl	15% APS	24.8 µl
TEMED	5.3 µl	5.3 µl	5.3 µl	TEMED	5 µl

Table 2.4: Primary antibodies for western blotting

Protein recognised	Species and type	Source	Catalogue number	Dilution	Reference
AID	Rat monoclonal	eBioscience	mAID-2	1:1000	
BMI1	Rabbit polyclonal	Bethyl	A301-694A	1:1000	
COBLL1	Rabbit polyclonal	Abcam	ab64465	1:2000	
EBNA2	Mouse monoclonal	Abcam	ab90543	1:500	
EBNA3A	Sheep polyclonal	Abcam	ab16126	1:1000	
EBNA3B	Rat monoclonal	Dr. Elizabeth Kremmer	6C9	1:10	(White et al., 2010)
EBNA3C	Mouse monoclonal	Prof. Martin Rowe	A10	1:10	(Maunders et al., 1994)
γ-tubulin	Mouse monoclonal	Sigma	T6557	1:8000	
HA tag	Mouse monoclonal	Covance	MMS-101R	1:1000	
LMP1	Mouse monoclonal	Dako	M0897	1:500	
p16	Mouse monoclonal	Prof. Gordon Peters	JC8	1:10	
Rb	Mouse monoclonal	BD Pharmingen	554136	1:500	
Rb (phospho S807/811)	Rabbit polyclonal	Cell Signaling	9308S	1:500	
RBPJ	Rat monoclonal	Prof. Bettina Kempkes	J7A11-161	1:100	(Maier et al., 2005)
SUZ12	Goat polyclonal	Santa Cruz	SC46264	1:1000	

Table 2.5: Secondary antibodies for western blotting

Antibody	Source	Catalogue number	Dilution
ECL sheep anti-mouse IgG HRP	GE Healthcare	NA931V	1:2000
Polyclonal rabbit anti-sheep Ig/HRP	Dako	P0163	1:2000
Polyclonal goat anti-rabbit Ig/HRP	Dako	P0448	1:2000
Polyclonal rabbit anti-goat Ig/HRP	Dako	P0449	1:2000
Polyclonal rabbit anti-rat Ig/HRP	Dako	P0450	1:2000

2.7 Chromatin immunoprecipitation (ChIP)

For ChIPs, cells were split to $3x10^5$ cells/ml the day prior to crosslinking multiples of 15 ml (~$4.5x10^6$ cells) with 1% formaldehyde (Sigma, UK) for 10 min at either 37°C or room temperature. Cells were washed twice in ice-cold PBS containing 1 mM PMSF, 1 µg/ml aprotinin (Roche) and 1 µg/ml pepstatin A (Roche). The equivalent of $4.5x10^6$ cells was pelleted in one 1.5 ml tube and snap-frozen on dry ice. Samples were stored at -80°C. Per ChIP chromatin of $1.5x10^6$ cells was used, so each tube provided enough material for three ChIPs.

All ChIP buffers used in this paragraph were supplemented with 1 mM PMSF, 1 µg/ml aprotinin (Roche) and 1 µg/ml pepstatin A (Roche) immediately prior to use and sample were kept on ice or 4°C until the elution from the beads. For ChIPs assaying histone modifications and SUZ12, chromatin was extracted in 300 µl ChIP SDS Lysis buffer and was sheared using a Bioruptor UCD-200 sonicator (Diagenode) at the highest setting for 12.5 min in pulses of 0.5 min (0.5 min on and 0.5 min off). Samples were then diluted 10-fold in ChIP Dilution buffer and 950 µl were used for each ChIP with 50 µl kept aside as input. For ChIPs against anti-Flag, BMI1, p300 and RBPJ, thawed cells were incubated in 1 ml swelling buffer under rotation for 20 min at 4°C. Nuclei were resuspended in 1 ml sonication buffer, left on ice for a minimum of 30 min and sonicated for one hour using a M220 focused sonicator (Covaris) with 75 W peak power, 26 duty cycle, 200 cycles/burst and 6°C set temperature. Samples were then adjusted to 3 ml using ChIP sonication buffer and 950 µl were used for each ChIP with 50 µl kept aside as input. Sheared chromatin from Bioruptor UCD-200 or M220 sonication was initially verified for size by agarose gel electrophoresis.

Thereafter, the ChIP assay kit from Millipore (17-295) was used, essentially according to the manufacturer's protocol. Protein A Agarose/Salmon Sperm DNA beads (16-157, Millipore) were used to IP rabbit-derived antibodies, Protein G Agarose/Salmon Sperm DNA beads (16-201, Millipore) were used to IP mouse-derived antibodies and protein G

magnetic beads (9006BC, Cell signaling) were used to IP the anti-Flag antibody. QIAquick PCR purification kit (28106, Qiagen) was used to purify DNA fragments instead of phenol/chloroform extraction in the manual. DNA was assayed by qPCR on QuantStudio 7 Flex (Life technologies) using Platinum SYBR Green qPCR SuperMix-UDG kit (11733-046, Life Technologies) and 200nM of ChIP primer. Input DNA was 5% of DNA used in ChIPs and was diluted to 2.5% prior to PCR quantification. The qPCR programme included an initial denaturation step for 2 min at 95°C, followed by 40 qPCR cycles with denaturation for 3 sec at 95°C and primer annealing and elongation for 30 sec at 60°C. Dissociation curves were used to assess specificity of PCR products. For every experiment, standard curves of four 5-fold serial dilutions of a mixture of different inputs were used to determine the efficiency of each primer pair. For this, the undiluted first standard was set to one and the log of each dilution [log(1), log(1/5), log(1/25), log(1/125)] was plotted against the Ct value of that dilution producing a standard curve where the slope represents the efficiency of the primer pair (efficiency = $10^{(-1/\text{slope})}$-1). The efficiency was always close to 100% and the range of the standard curve covered the Ct values obtained from input and ChIP samples. The DNA amount in each sample was determined by using the slope and the y-intercept of the standard curve. Enrichment relative to input was then calculated by dividing the DNA amount in the ChIP sample by the DNA amount in the corresponding input sample. Error bars were calculated as standard deviations from triplicate PCR reactions for both input and ChIP. Antibodies used are listed in Table 2.6 and sequences of ChIP primers in Table 2.7.

Table 2.6: Antibodies used for ChIP

Protein recognised	Species and type	Source	Catalogue number	Quantity
Anti-Flag (DYKDDDDK)	Rabbit polyclonal	Cell Signaling	2368S	4 µg
BMI1	Rabbit polyclonal	Bethyl	A301-694A	4 µg
EBNA2	Mouse monoclonal	Abcam	ab90543	4 µg
H3K4me3	Rabbit monoclonal	Millipore	17-614	4 µg
H3K9ac	Rabbit polyclonal	Millipore	17-658	4 µg
H3K27ac	Mouse monoclonal	Millipore	05-1334	2 µg

Protein recognised	Species and type	Source	Catalogue number	Quantity
H3K27me3	Rabbit polyclonal	Millipore	07-449	4 µg
p300-C20	Rabbit polyclonal	Santa Cruz	SC585X	4 µg
RBPJ	Rabbit polyclonal	Abcam	ab25949	4 µg
SUZ12	Rabbit polyclonal	Abcam	ab12073	4 µg

Table 2.7: Sequences of ChIP primers

Primer name	Sense	Primer sequence (5'-3')	Reference
ADAM28 TSS	Forward	CCATTGTTGCAGGACCACAG	This study
	Reverse	GCCTCCTCTCCAGTGAGACA	
ADAM28 control	Forward	ACAGGAGCATGCACTCTTCA	(McClellan et al., 2012)
	Reverse	GGCAATGTTCTGCTGCAA	
ADAM peak	Forward	CTTCATGGCTACAGACTCTTGG	
	Reverse	CCTATGTCTCGCTTCCTGCT	
ADAMDEC1 TSS	Forward	TAAGGATACTGAACAACAATCG	This study
	Reverse	GGAATGTGTGAACTGAGACT	
AICDA region I	Forward	GAGGAAGGCCAGTGCAATCA	
	Reverse	CAGGGAGGCAAGAAGACACT	
AICDA region II	Forward	GCTAGTTAACTTTGTTGATC	
	Reverse	CTACTCAGGACAGAAATGAC	
AICDA region III	Forward	GCTAGCCAGTCTTTGTGCCT	
	Reverse	ACCAGGATCAGGAGGTGTTG	
AICDA region IV	Forward	TGGACACCAGCTAGATTGTTCA	
	Reverse	TCACACTTTCACCCACACAGA	
AICDA region V	Forward	CCTGTTCCTCTCCTTACCGC	
	Reverse	ACGGAAGCCCTTGTATCTTTGA	
AICDA region VI	Forward	CAGCAAGTTTCCTTCTGCGA	
	Reverse	GCCATTTCTGACTCAGCAGC	
AICDA control	Forward	GTCCTGTACAGTAACTAGAGAAAA	
	Reverse	GCAAAGCAAGACGACAAAGGA	
COBLL1 peak	Forward	CTGAGTAACAAGAGCGAAAGAG	
	Reverse	ATCAGATGTGTTATGACTAACAGC	
COBLL1 control	Forward	CCCTCCAGTATACCCCAGCT	
	Reverse	ACCCCTTCTCTTTACTTGGCC	
COBLL1 TSS	Forward	GCCGCCGTCTCTACAAGGTCTA	
	Reverse	CTACCCAGTAAACCCCACGG	
GAPDH	Forward	CGCTCTCTGCTCCTCC	Millipore
	Reverse	TTTCTCTCCGCCCGTCCAC	
Myoglobin	Forward	GGAGAAAGAAGGGGAATCACA	(Delbarre et al., 2010)
	Reverse	GATAAATATAGCCAACGCCACA	
HES1	Forward	AAGTTTCACACGAGCCGTTC	(Ntziachristos et al., 2012)
	Reverse	GCTGTTATCAGCACCAGCTC	

2.8 Transient reporter assays

2.8.1 Cloning of transient reporter vectors

Genomic DNA was extracted from GM12878 cells according to the instructions of the Blood & Cell Culture DNA Midi Kit (Qiagen) and served as template in all PCR reactions. The promoter regions 1 kb upstream of the RefSeq gene annotation of *ADAM28*, *ADAMDEC1*, *AICDA* or *COBLL1* (short transcripts) were PCR amplified with primers that add *Mlu*I restriction sites (Table 2.8) using Q5 High Fidelity polymerase (NEB). The amplified promoter regions were then cloned upstream of the luciferase gene in pGL3-basic vector (Promega) using the *Mlu*I restriction site (Figure 2.1A). Restriction digestion and DNA sequencing was used to confirm the direction of the insert.

PCR was also used to amplify the *ADAM* peak (1 kb around the *ADAM* peak adding either *Sal*I or *Not*I restriction sites), the *COBLL1* peak (1.5 kb around the *COBLL1* peak adding *Sal*I restriction sites), *AICDA* region II (1 kb around region II adding *Sal*I restriction sites), *AICDA* region III (1 kb around region III adding *Sal*I restriction sites) and *AICDA* region IV (2 kb around region IV adding *Not*I restriction sites). *ADAM* peak *Sal*I and *COBLL1* peak were cloned downstream of the luciferase gene driven by the *ADAM28* or *COBLL1* promoter, respectively, using the *Sal*I restriction site (Figure 2.1B and C). *ADAM* peak *Not*I was cloned upstream of the *ADAMDEC1* promoter in the pGL3-*ADAMDEC1* vector using the *Not*I restriction sites (Figure 2.1D). *AICDA* regions II and III were inserted downstream of the luciferase gene driven by the *AICDA* promoter using the *Sal*I restriction sites (Figure 2.1B) and *AICDA* region IV was cloned upstream of the AICDA promoter using the *Not*I restriction sites (Figure 2.1E).

All vectors were screened for the correct orientation and were sequence verified by Sanger sequencing. Primers used to PCR amplify promoter regions and EBNA3 binding peaks are listed in Table 2.8.

Table 2.8: Sequences of primers used to clone transient reporter vectors

Primer name	Sense	Primer sequence (5'-3')
ADAM28 promoter	Forward	GAATTCACGCGTGGCGCGCCGGAGGCCTCACTCTCATGGTAG
	Reverse	AAGCTTACGCGTGACAGAATGAAACTCCCTGG
ADAMDEC1 promoter	Forward	GAATTCACGCGTGGCGCGCCGAAGGTATGGGTTTTTAGATTGGG
	Reverse	AAGCTTACGCGTGGACACTGTTTGCCCCTGGTGC
ADAM peak NotI	Forward	GAATTCGCGGCCGCTAAAGCCATGGGTCAGATGAGA
	Reverse	GAATTCGCGGCCGCCTGGATTGATTACAAACAAAAT
ADAM peak SalI	Forward	GAATTCGTCGACTAAAGCCATGGGTCAGATGAGACT
	Reverse	AAGCTTGTCGACCTGGATTGATTACAAACAAA
AICDA Reg I (promoter)	Forward	GAATTCACGCGTGGCGCGCCTTGATTTGAAAATCATCAAGG
	Reverse	AAGAGCACGCGTGTCTTAGTCTGACAGTGCATTGGC
AICDA Reg II	Forward	GAATTCGTCGACTTTAAAAGGATAATAATTTTTATGG
	Reverse	AAGCTTGTCGACGTCAGACAATGTCAAATGTC
AICDA Reg III	Forward	GAATTCGTCGACCAACTATGGAACAACTGAGG
	Reverse	AAGCTTGTCGACAGTCTATTTAGGTATTCATG
AICDA Reg IV	Forward	GAATTCGCGGCCGCAGATGTAAGGAAGTGGTCCAG
	Reverse	AAGAGCGCGGCCGCGGCATTGGGAAAGAATTCCTTC
COBLL1 promoter	Forward	GAATTCACGCGTGGCGCGCCTTAAAGGTCAAGTTAAAAAC
	Reverse	AAGCTTACGCGTTCTCTACAAGGTCTAGCGGG
COBLL1 peak	Forward	GAATTCGTCGACGGGCACAGAGAAACTTTCTGC
	Reverse	AAGCTTGTCGACGGATTCACTTGAGGCTAGTATT

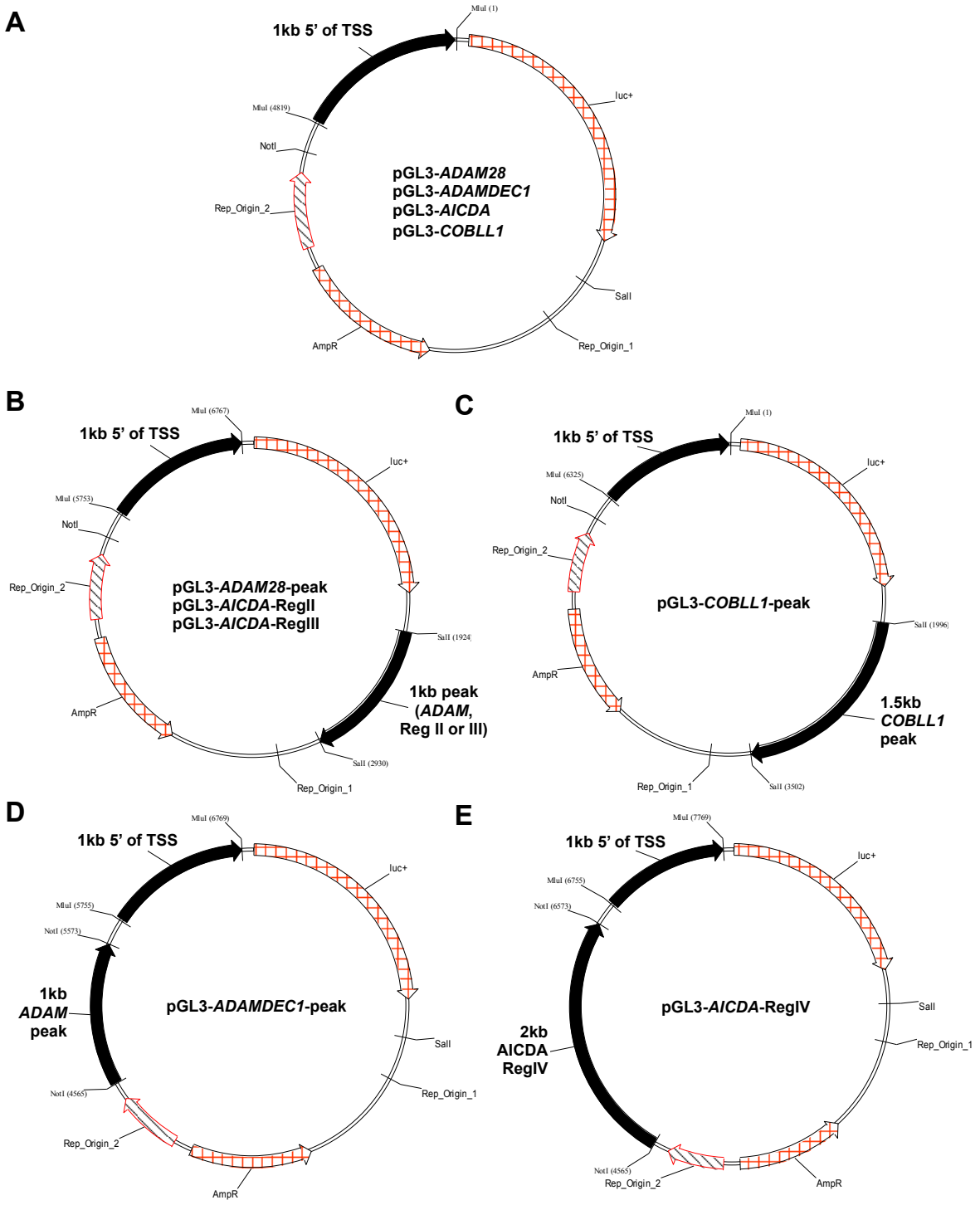


Figure 2.1: Luciferase vectors cloned during this study.

(**A**) The 1kb region 5' of the TSS of *ADAM28*, *ADAMDEC1*, *AICDA* or *COBLL1* were cloned upstream of the luciferase gene (luc+) into pGL3-basic vector using *Mlu*I restriction sites. (**B**) The 1 kb sequence around the *ADAM* peak, *AICDA* region II or III was cloned downstream of luc+ into pGL3-ADAM28 or pGL3-AICDA, as indicated. (**C**) The 1.5 kb sequence around the *COBLL1* peak was cloned downstream of luc+ into pGL3-*COBLL1*. (**D**) The 1 kb sequence around the *ADAM* peak was cloned upstream of the 1kb 5' region of *ADAMDEC1*. (**E**) The 2 kb sequence around *AICDA* region IV was cloned upstream of the 1kb 5' region of *AICDA*.

2.8.2 Cloning of RBPJ binding mutant (BM) EBNA3C into expression vector

The shuttle vector used to create the RBPJ BM EBNA3C recombinant EBV (Kalchschmidt et al., 2016) created by Adam Gillman was digested with *Aar*I and *Spe*I to extract the 716 bp fragment that carries the $_{209}$AAAA and W227S mutations. This fragment was then used to replace the *Aar*I-*Spe*I fragment of pCDNA3-EBNA3C. For this, three-way ligation between the mutated 716 bp fragment together with the 2,056 bp *Spe*I/*Not*I fragment and 5,815 bp *Not*I/*Aar*I fragment of pCDNA3-EBNA3C was performed. Transformed *E. coli* colonies were screened by restriction enzyme digestion for the presence of an additional *Not*I and *Sal*I site introduced with the $_{209}$AAAA and W227S mutations compared to the parental pCDNA3-EBNA3C vector.

2.8.3 Electroporation and transient reporter assays

Vectors used in electroporations are listed in Table 2.9. Luciferase vectors (1 µg) were electroporated into DG75, DG75ΔRBPJ or Raji cells together with 1 µg pSV-β-galactosidase and varying amounts of pCDNA3-EBNA3 expression plasmids. Total amounts of DNA were balanced using an empty pCDNA3 expression plasmid. Electroporations were performed as described previously for DG75 (Hickabottom et al., 2002). The same protocol was used to electroporate Raji cells with the *COBLL1* luciferase constructs, but a voltage of 240V was used. All electroporations were harvested after 48h. Luciferase and β-galactosidase assays were performed as described previously (Hickabottom et al., 2002) and measured on FLUOstar Omega (BMG Labtech). β-galactosidase activity was used to normalise luciferase activities for transfection efficiencies.

Table 2.9: Vectors used in transient reporter assays

Vector name	Description	Reference
pCDNA3-empty	Empty expression vector	Invitrogen
pCDNA3-EBNA3A	Expression vector for full-length EBNA3A	
pCDNA3-EBNA3B	Expression vector for full-length EBNA3B	
pCDNA3-EBNA3C	Expression vector for full-length EBNA3C	
pSG5-HA-EBNA3C 11-549	Expression vector for HA-tagged N-terminal part of EBNA3C (amino acids 11-549)	Gillian Parker
pSG5-HA-EBNA3C 550-992	Expression vector for HA-tagged C-terminal part of EBNA3C (amino acids 550-992)	Gillian Parker
pCDNA3-RBPJ BM EBNA3C	Expression vector for EBNA3C carrying the $_{209}$AAAA and W227S mutations	This study
pSG5-EBNA2	Expression vector for full-length EBNA2	
pSV-β-galactosidase	β-galactosidase expression vector	
pGL3-basic vector	pGL3 luciferase vector without promoter or enhancer	Promega
pGL3-*ADAM28*	pGL3-basic vector + 1kb promoter sequence of *ADAM28*	This study (Figure 2.1)
pGL3-*ADAM28*-peak	pGL3-basic vector + 1kb promoter sequence of *ADAM28* + 1kb sequence around *ADAM* peak	
pGL3-*ADAMDEC1*	pGL3-basic vector + 1kb promoter sequence of *ADAMDEC1*	
pGL3-*ADAMDEC1*-peak	pGL3-basic vector + 1kb promoter sequence of *ADAMDEC1* + 1kb sequence around *ADAM* peak	
pGL3-*COBLL1*	pGL3-basic vector + 1kb promoter sequence of *COBLL1*	
pGL3-*COBLL1*-peak	pGL3-basic vector + 1kb promoter sequence of *COBLL1* + 1.5kb sequence around *COBLL1* peak	
pGL3-*AICDA* Reg I	pGL3-basic vector + 1kb promoter sequence of *AICDA* (Reg I)	
pGL3-*AICDA* Reg II	pGL3-basic vector + 1kb promoter sequence of *AICDA* (Reg I) + 1kb sequence around *AICDA* Reg II	
pGL3-*AICDA* Reg III	pGL3-basic vector + 1kb promoter sequence of *AICDA* (Reg I) + 1kb sequence around *AICDA* Reg III	
pGL3-*AICDA* Reg IV	pGL3-basic vector + 1kb promoter sequence of *AICDA* (Reg I) + 2kb sequence around *AICDA* Reg IV	

3 Chapter III: Gene repression by EBNA3C

3.1 Introduction

This chapter will focus on the role of EBNA3C in the repression of host genes. Based on the exon microarray in EBNA3C-conditional LCL (Skalska et al., 2013), the three host genes most repressed by EBNA3C (Table 1.1) – *COBLL1*, *ADAM28* and *ADAMDEC1* – were selected as model genes in order to explore in more detail the molecular mechanisms and factors that are involved in EBNA3C-mediated gene repression. These three genes have been selected for what appeared to be a very robust repression by EBNA3C and not because of their biological function (see Discussion 3.4.5 for why EBV might target them).

COBLL1 is a novel EBNA3C target gene and has not been previously described as regulated by EBNA3C or other latent EBV proteins. The function of the COBLL1 (Cordon-bleu (COBL) like 1) protein is unknown.

ADAM28 and *ADAMDEC1* are known targets of both EBNA3A and EBNA3C. In addition to the microarray on 3CHT LCL EBNA3C (Skalska et al., 2013), they have been identified in microarray analyses on EBNA3A KO LCL (Hertle et al., 2009), BL31 infected with EBNA3A or EBNA3C KO viruses (White et al., 2010) and BJAB stably expressing EBNA3C (McClellan et al., 2012). They are both members of the ADAM (a disintegrin and metalloprotease) family of cell surface proteins (or in some cases secreted proteins) with proteolytic and adhesive properties. *ADAM28* and *ADAMDEC1* are encoded in adjacent genomic loci. Shortly before the beginning of this study an intergenic EBNA3 binding site was identified between *ADAM28* and *ADAMDEC1* (McClellan et al., 2012). When EBNA3C was expressed reduced levels of the activation-associated histone modification H3K9/14ac and increased levels of the repressive H3K27me3 mark were observed within *ADAM28* and at the TSS of *ADAMDEC1* (McClellan et al., 2012). During the course of this study repressive looping in the presence of EBNA3C was reported between the EBNA3 binding site and the TSSs of *ADAM28* and *ADAMDEC1* (McClellan et al., 2013).

Most of these observations were however obtained from stable transfectants of EBNA3C in the EBV-negative BL cell line BJAB and in the absence of other latent EBV gene products. Before this study, nothing was known about the changes to gene expression of these three genes in primary B cells early after infection with EBV. Furthermore, nothing is known about the molecular mechanisms, including the temporal sequence of events or the factors that are involved in the EBNA3C-mediated gene repression of these genes. Most of the results presented in this Chapter have now been published (Kalchschmidt et al., 2016).

3.2 Results

3.2.1 Validation of EBNA3C repressed genes in EBV-infected primary B cells

Previously established virus producing cell lines (Anderton et al., 2007) were re-validated before producing recombinant EBV viral particles to infect primary B cells in order to establish that *COBLL1*, *ADAM28* and *ADAMDEC1* are genuine targets of EBNA3C and to determine whether EBNA3A or EBNA3B are involved in their regulation.

Viral episomes were rescued from EBV B95.8 wild type (wtI6), EBNA3A (3A), EBNA3B (3B), EBNA3C (3C), triple EBNA3A/3B/3C (E3) knock-out (KO) or revertant (Rev) HEK293 virus producer cell lines (overview in Figure 1.17). Episomes were digested with *Cla*I or *Eco*RI and analysed by pulsed-field gel electrophoresis (Figure 3.1A). All episomes showed the expected restriction enzyme digestion pattern (for list of all fragments see Table 7.1), which validated the integrity of the EBV BACs and the identity of all cell lines. In addition, western blot analysis for the expression of EBNA3 proteins in RIPA extracts of the HEK293 virus producer cell lines further confirmed the identity of all cell lines (Figure 3.1B).

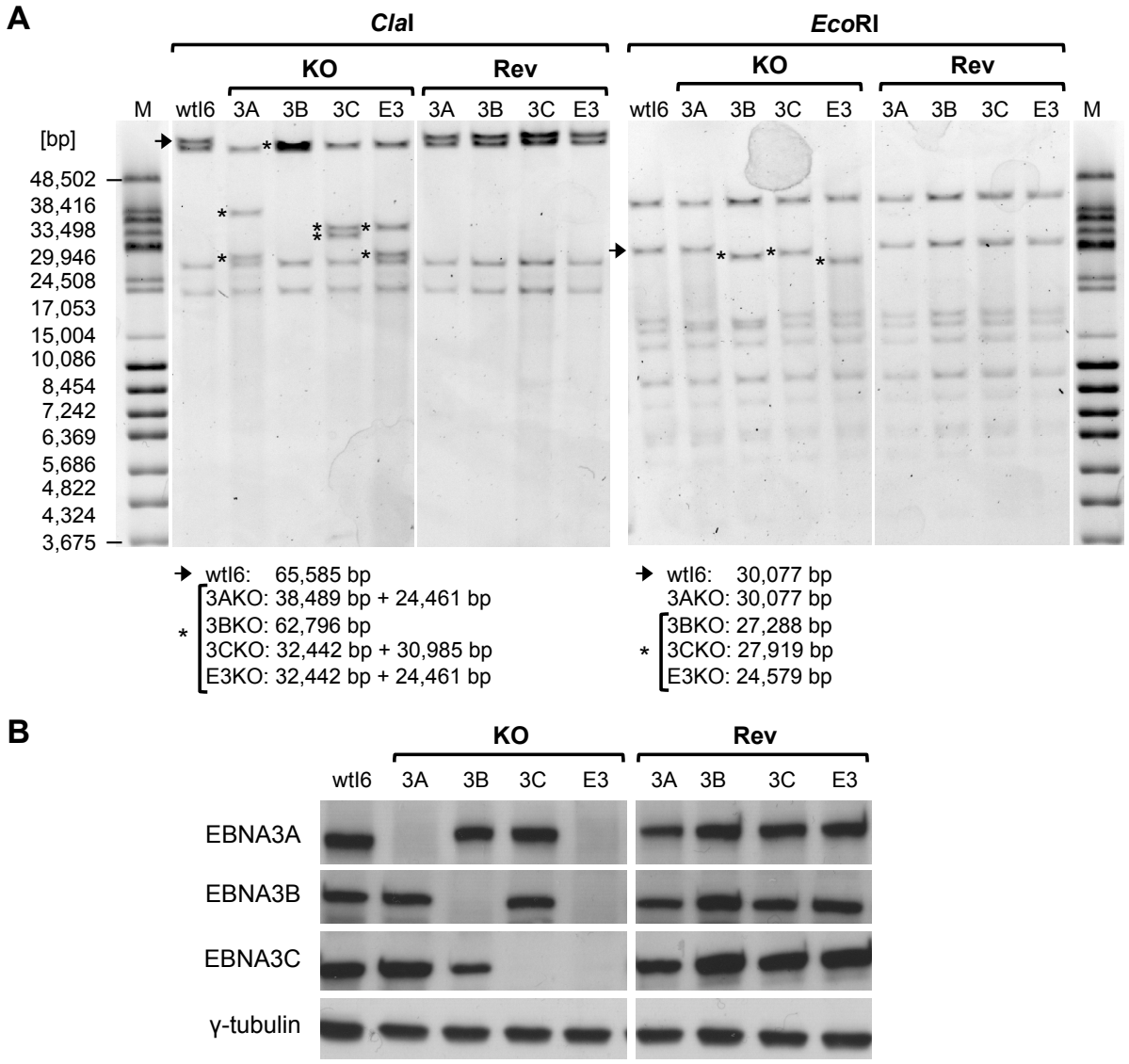


Figure 3.1: Verification of HEK293 virus producer cell lines.

(**A**) Restriction enzyme digestion and pulsed-field agarose gel electrophoresis of viral episomes rescued from HEK293 virus producer cell lines as indicated. Episomes were rescued from EBV B95.8 wild type I6 (wtI6), EBNA3A (3A), EBNA3B (3B), EBNA3C (3C), EBNA3A/3B/3C (E3) knock-out (KO) or revertant (Rev) cell lines and digested with *Cla*I or *Eco*RI. All cell lines show the expected restriction enzyme digestion pattern (see Table 7.1 for full list). Arrows highlight bands in the wild type that are specifically changed in each mutant indicated by asterisks with expected fragment sizes shown below each digestion. (**B**) Western blot analysis of the same HEK293 virus producer cell lines as in A. Cell lines were analysed for expression of EBNA3A, EBNA3B, EBNA3C and γ-tubulin as loading control.

Next, having validated all HEK293 virus producer cell lines, infectious viral particles were produced and used to infect purified primary B cells in a single experiment to establish that *COBLL1*, *ADAM28* and *ADAMDEC1* are genuine targets of EBNA3C and to determine whether EBNA3A or EBNA3B are involved in their regulation. Consistent with the fact that EBNA3A and EBNA3C are both important for B cell transformation, B cells infected with triple E3KO virus could not be cultured past day 21 as cells stopped proliferating and/or died. 3AKO- or 3CKO-infected cultures also struggled around this time, but cells still proliferated. All other infections (wt, Rev and 3BKO) did not exhibit this crisis and proliferated as expected to establish LCLs.

Infection with all EBNA3A- and EBNA3C-competent viruses resulted in a reduction of gene expression of *ADAM28* (2-3 log fold, Figure 3.2A) and *ADAMDEC1* (1-2 log fold, Figure 3.2B) over a period of 30 days after infection. In the absence of all EBNA3s, EBNA3C alone and to a lesser extent EBNA3A alone, there was a failure to repress *ADAM28* and *ADAMDEC1*. This makes *ADAM28* and *ADAMDEC1* targets of EBNA3A and EBNA3C, which is consistent with previous reports (Hertle et al., 2009; McClellan et al., 2012).

For *COBLL1*, infection with all EBNA3C-competent viruses resulted in a robust 3-4 log fold repression over the same period of time (Figure 3.2C). Only the triple E3KO, lacking all EBNA3 proteins, and the 3CKO failed to repress *COBLL1* expression and maintained levels similar to uninfected primary B cells. This is consistent with *COBLL1* being regulated only by EBNA3C.

Analysis of *ALAS1* (Figure 3.2D) and *GNB2L1* (Figure 3.2E), two control genes that are not targeted by EBV proteins, did not exhibit these differences in gene expression between the various virus infections.

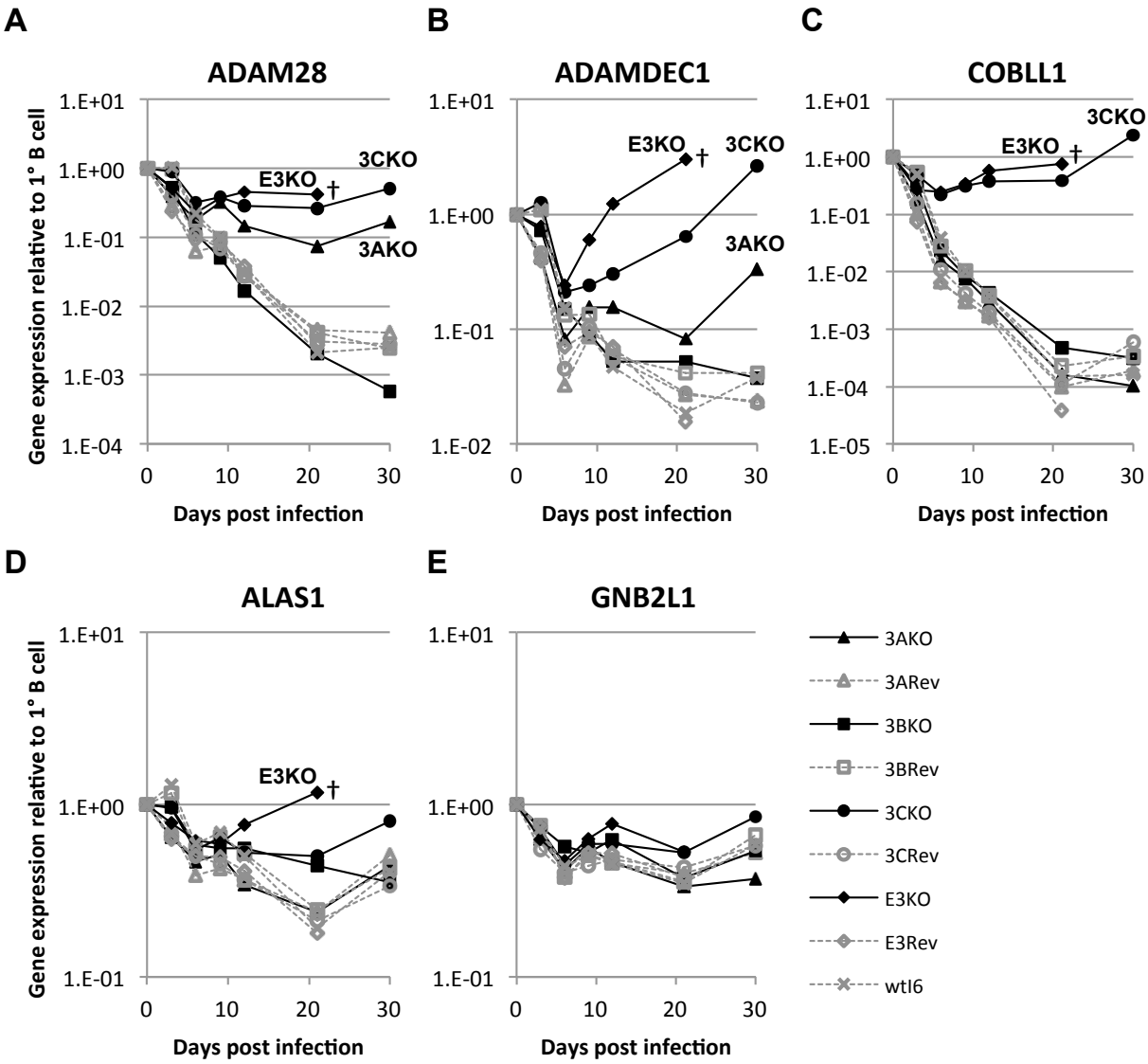


Figure 3.2: Repression of *ADAM28*, *ADAMDEC1* and *COBLL1* in primary B cells infected with recombinant EBV.

Infection of purified primary B cells with EBV B95.8 wild type (wtI6), recombinant EBNA3A (3A), EBNA3B (3B), EBNA3C (3C) or triple EBNA3A/B/C (E3) knock-out (KO) or revertant (Rev) viruses. Over a period of 30 days post infection, RT-qPCR gene expression for *ADAM28* (**A**), *ADAMDEC1* (**B**), *COBLL1* (**C**), *ALAS1* (**D**) or *GNB2L1* (**E**) was normalised to *GAPDH* and is shown relative to uninfected primary B cells. Primary B cells infected with E3KO were rather unhappy on day 21 post infection and could not be cultured past that time-point (†). These results are from a single experiment.

I was able to establish stable LCLs from all primary B cell infections, except for E3KO-infected cells. Western blot analysis of these LCLs revealed that 3AKO and 3CKO LCLs had low levels of retinoblastoma protein (Rb) and phosphorylated Rb (P-Rb) and in addition 3CKO LCL had lost expression of $p16^{INK4a}$ protein (Figure 3.3). This type of clonal selection has been observed before, more frequently however, with cells infected with 3AKO virus (Hertle et al., 2009; Skalska et al., 2010). More importantly and consistent with the gene expression results, western blot analysis of these LCL revealed that COBLL1 protein is only detectable in 3CKO LCL and completely undetectable in all EBNA3C-competent cell lines (Figure 3.3). Unfortunately, protein expression of ADAM28 and ADAMDEC1 could not be assessed because commercial antibodies tested did not produce convincing or reproducible results.

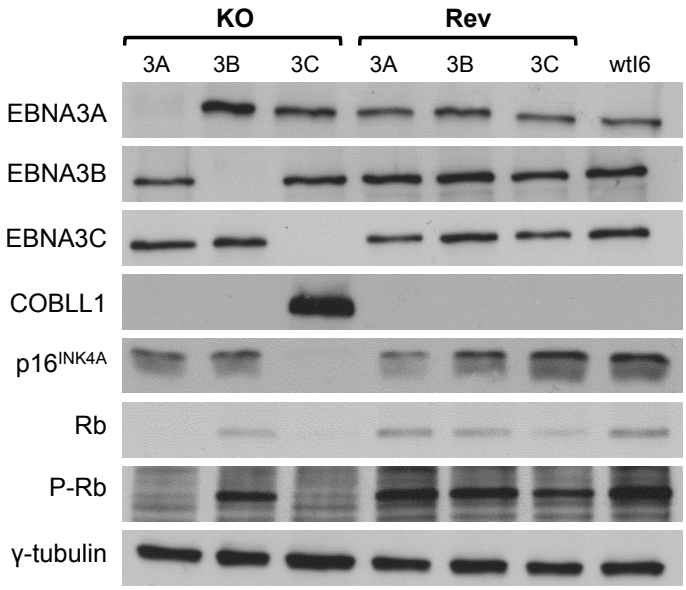


Figure 3.3: Western blot analysis of LCL established from primary B cell infections.
Western blot verification of EBNA3A (3A), EBNA3B (3B), EBNA3C (3C) knock-out (KO), revertant (Rev) or wild type (wtl6) LCL grown out from the primary B cell infection. LCL were analysed on day 85 post infection for expression of EBNA3A, EBNA3B, EBNA3C, COBLL1, $p16^{INK4A}$, Retinoblastoma protein (Rb), Phosphorylated Rb (P-Rb) and γ-tubulin as loading control.

3.2.2 Recapitulation of EBNA3C-mediated gene repression in EBNA3C-conditional LCL

Having established that EBNA3C is the main repressor of both *ADAM*s and *COBLL1* in EBV-infected primary B cells, the next step was to determine whether it is possible to recapitulate this repression in EBNA3C-conditional (3CHT) LCLs. These cells carry recombinant EBV that express endogenous levels of an EBNA3C/estrogen receptor fusion protein, which renders EBNA3C conditional to the presence of 4-hydroxytamoxifen (HT) in the medium (see Section 1.4.3). 3CHT A13 is a p16-null EBNA3C-conditional LCL that has been established in the absence of HT and therefore inactive EBNA3C. Having seen in the infection of primary B cells that the full repression of both *ADAM*s and *COBLL1* required a period of 30 days and to see whether the repression was reversible, the 3CHT time-course experiment was designed to run over a period of 60 days (Figure 3.4A). For this, cells were cultured for 60 days either in the absence of HT (-HT, inactive EBNA3C), presence of HT (+HT, active EBNA3C) or when HT was washed from the medium after 30 days and cells cultured for another 30 days in the absence of HT (washed, inactivated EBNA3C).

In the absence of HT and therefore inactive EBNA3C, these cells express *ADAM28*, *ADAMDEC1* and *COBLL1* at similar levels to uninfected primary B cells and gene expression levels of all three genes were relatively stable over this 60-day period. Activation of EBNA3C by the addition of HT to the medium resulted in a 2-log fold repression of *ADAM28* (Figure 3.4B) and *ADAMDEC1* (Figure 3.4C) and a 3-4 log fold repression of *COBLL1* (Figure 3.4D). However, repression was fully reversible. Inactivation of EBNA3C, by removing HT from the medium from day 30 onwards, resulted in full de-repression over a similar 30-day period. Again, these changes of gene expression could not be detected for the two control genes *ALAS1* (Figure 3.4E) and *GNB2L1* (Figure 3.4F). At the protein level, most of COBLL1 protein was lost three days after activation of EBNA3C and was completely undetectable by day 6 (Figure 3.4G). However, this was also reversible as COBLL1 protein reappeared 20 days after inactivation of EBNA3C.

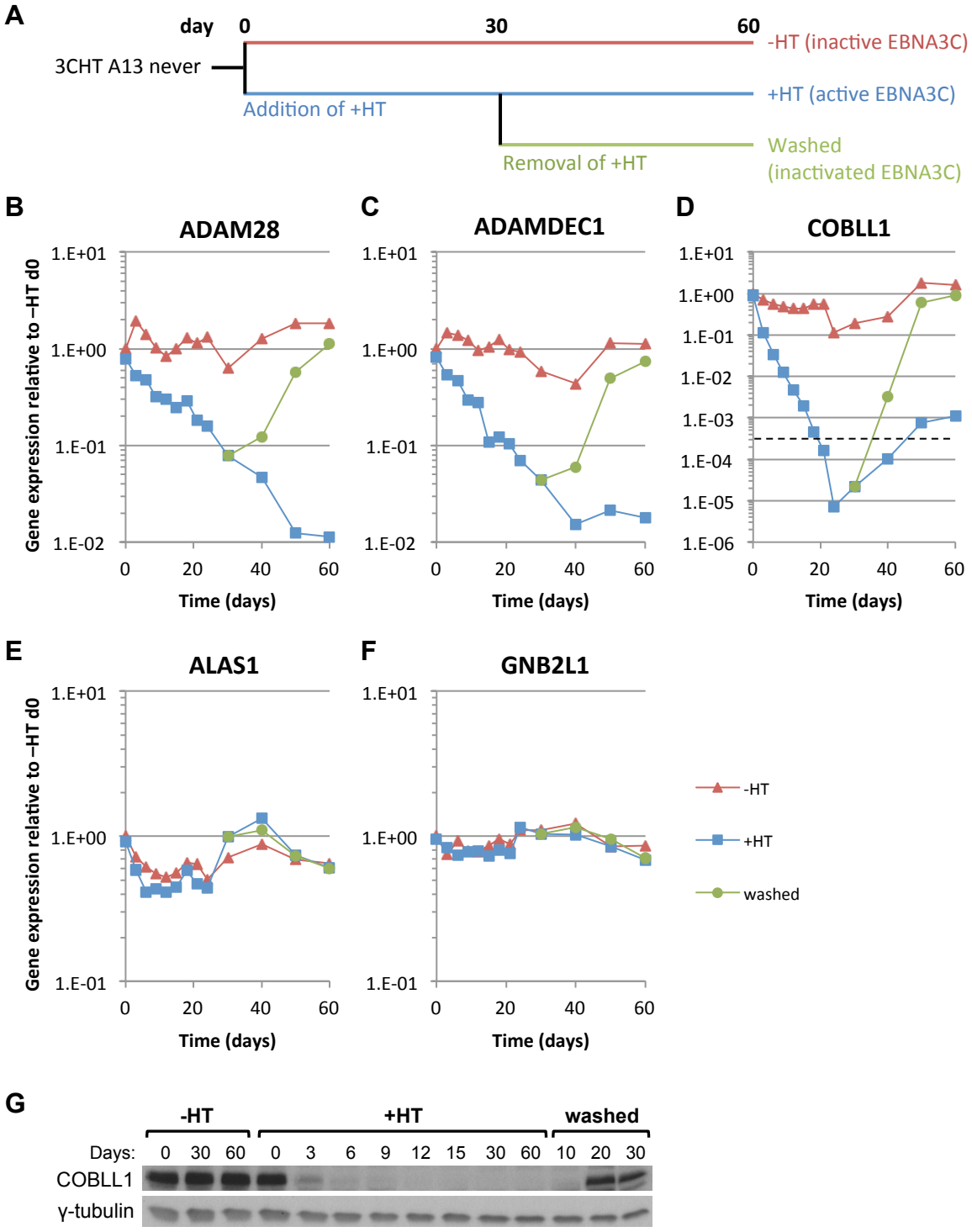


Figure 3.4: Repression of *ADAM28*, *ADAMDEC1* and *COBLL1* in EBNA3C-conditional LCL.
(**A**) Time-course using 3CHT A13 LCL. Cells were grown over 60 days either in absence of HT (-HT), presence of HT (+HT) or with HT removed after 30 days +HT (washed). (**B-F**) RT-qPCR gene expression of *ADAM28* (**B**), *ADAMDEC1* (**C**), *COBLL1* (**D**), *ALAS1* (**E**) and *GNB2L1* (**F**) was normalised to *GAPDH* and is shown relative to day 0 –HT. The dashed line indicates the limit of detection for *COBLL1* expression and values below this line become extremely unreliable. (**G**) Western blot analysis of COBLL1 protein levels during the time-course as indicated with γ-tubulin as loading control. This is a representative result of three biological replicate time-courses.

These changes in gene expression were a direct consequence of the HT-induced activation of EBNA3C, because addition of HT over a period of 30 days to the non-conditional 3CKO LCL, established from the primary B cell infection, did not change the expression levels of any of the three genes (Figure 3.5). HT was however functional and caused the expected gene expression changes when added to 3CHT cells during the same experiment.

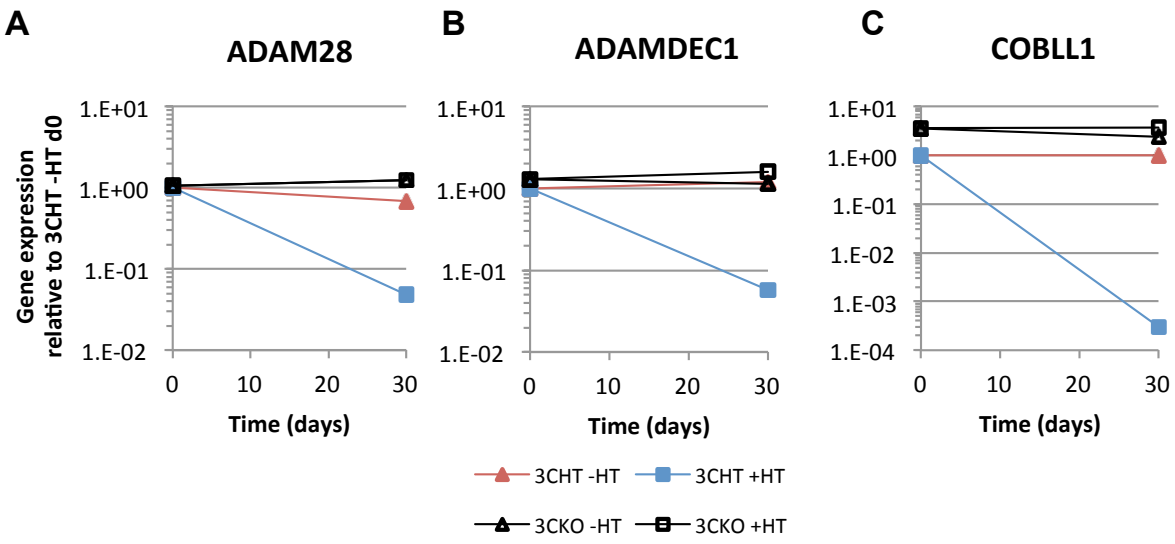


Figure 3.5: HT does not repress EBNA3C-target genes in non-conditional 3CKO LCL.
Time-courses using non-conditional 3CKO or conditional 3CHT A13 LCLs. Cells were grown over 30 days either in absence (-HT) or presence of HT (+HT). RT-qPCR gene expression of *ADAM28* (**A**), *ADAMDEC1* (**B**) and *COBLL1* (**C**) was normalised to *GAPDH* and is shown relative to 3CHT on day 0. These results are from a single experiment.

Furthermore, the results obtained from the 3CHT A13 time-course were highly reproducible, not only in the same cell line (see Section 3.2.3), but also in 3CHT C19, another 3CHT cell line created by an independent 3CHT recombinant virus clone on the p16-null background. This line was grown out in the presence of HT and washed for at least three months before starting the experiment (Appendix, Figure 7.1 and Figure 7.2).

3.2.3 Comparison between primary B and EBNA3C-conditional LCL

In order to determine the extent to which the EBNA3C-conditional 3CHT LCL represent changes to gene expression observed after infection of primary B cell with EBV, gene expression of all three genes were directly compared between the two systems (Figure 3.6). Repression of *ADAM28*, *ADAMDEC1* and *COBLL1* expression in newly EBV-infected primary B cells was very similar to repression observed in the 3CHT LCL after activation of EBNA3C. Repression of all three genes followed highly exponential repression profiles (R^2 > 0.9) in primary B cell infections (wtI6 as representative EBV infection of primary B cells), but also in three independent 3CHT time courses (two replicate A13 time-courses and one C19 time-course). The dynamic range of *COBLL1* repression (more than three logs) was remarkable, following a highly similar exponential repression profile in newly infected primary B cells, but also in all three 3CHT LCL time-courses (Figure 3.6A). Repression of *ADAM28* and *ADAMDEC1* was rather more variable between cell lines, it was slowest in the two replicate 3CHT A13 time-courses and fastest in newly infected primary B cells (Figure 3.6B and C). Together however, these results showed for the first time, that the EBNA3C conditional cell lines can be used to recapitulate efficiently EBNA3C-mediated gene repression observed in B cells early after infection with EBV.

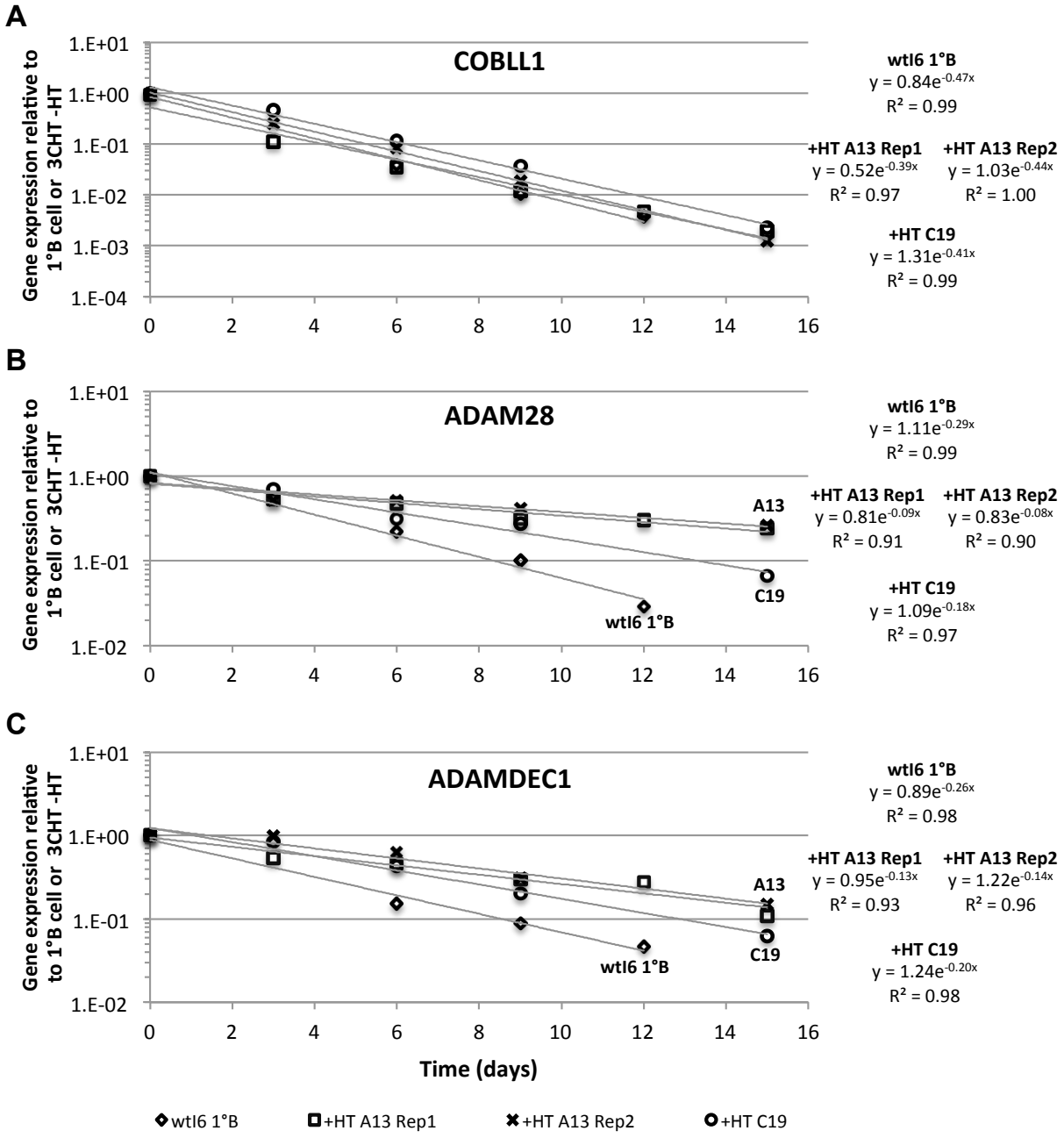


Figure 3.6: Comparative gene expression analysis between newly EBV-infected primary B cells and 3CHT LCL after activation of EBNA3C.

RT-qPCR gene expression comparison for *COBLL1* (**A**), *ADAM28* (**B**) and *ADAMDEC1* (**C**) from the primary B cell infection with wild type (wtI6) EBV and three replicate time-courses using 3CHT A13 and C19 after addition of HT. Gene expression was normalised to *GAPDH* and expressed relative to uninfected primary B cells for wtI6 or 3CHT –HT on day 0 for 3CHT time-courses. Exponential trendlines were fitted with equations and R^2 values shown.

3.2.4 Molecular mechanism of EBNA3C-mediated gene repression

Having confirmed that EBNA3C was the main repressor of *COBLL1*, *ADAM28* and *ADAMDEC1* in newly EBV-infected primary B cells and that this repression could be efficiently replicated in EBNA3C-conditional LCL, my aim was to use the EBNA3C-conditional LCL system to explore some of the molecular details and to identify factors that were involved in this very efficient repression of gene expression by EBNA3C.

3.2.4.1 EBNA3 binding at the COBLL1 and the ADAM28/ADAMDEC1 loci

At the beginning of this project, it was unknown whether *COBLL1* was a direct target of EBNA3C, because nothing was known about EBNA3C binding at the *COBLL1* locus. In contrast, an intergenic EBNA3 binding site in between *ADAM28* and *ADAMDEC1* had been detected by ChIP-Seq using a commercial polyclonal antibody against EBNA3C that also precipitates EBNA3A and EBNA3B (McClellan et al., 2012).

Anti-Flag ChIP-Seq using LCLs expressing either tandem-affinity purification tagged EBNA3A, EBNA3C or wild type (untagged) EBNA3s (Figure 3.7A) was performed by Dr Kostas Paschos in our lab (Paschos *et al.*, manuscript in preparation). This identified a single intragenic EBNA3A and EBNA3C binding site, hereafter called *COBLL1* peak, approximately 80kb downstream of the TSS of *COBLL1* (Figure 3.7B). The anti-Flag ChIP-Seq also detected binding of EBNA3A and EBNA3C at the previously identified intergenic EBNA3 binding site between *ADAM28* and *ADAMDEC1* (McClellan et al., 2012), hereafter called *ADAM* peak (Figure 3.7C). Neither of these peaks was observed in Anti-Flag ChIP-Seq performed on wild type (wt) LCL encoding untagged EBNA3A and EBNA3C (Figure 3.7B and C).

The enrichment levels of EBNA3A and EBNA3C at the *ADAM* and *COBLL1* peaks as determined by ChIP-Seq are shown in Figure 3.7D, which was confirmed by anti-Flag ChIP-qPCR performed by Dr Quentin Bazot (Figure 3.7E). In both ChIP-Seq and ChIP-qPCR, enrichment levels of EBNA3C were higher at the *COBLL1* peak than at the *ADAM* peak.

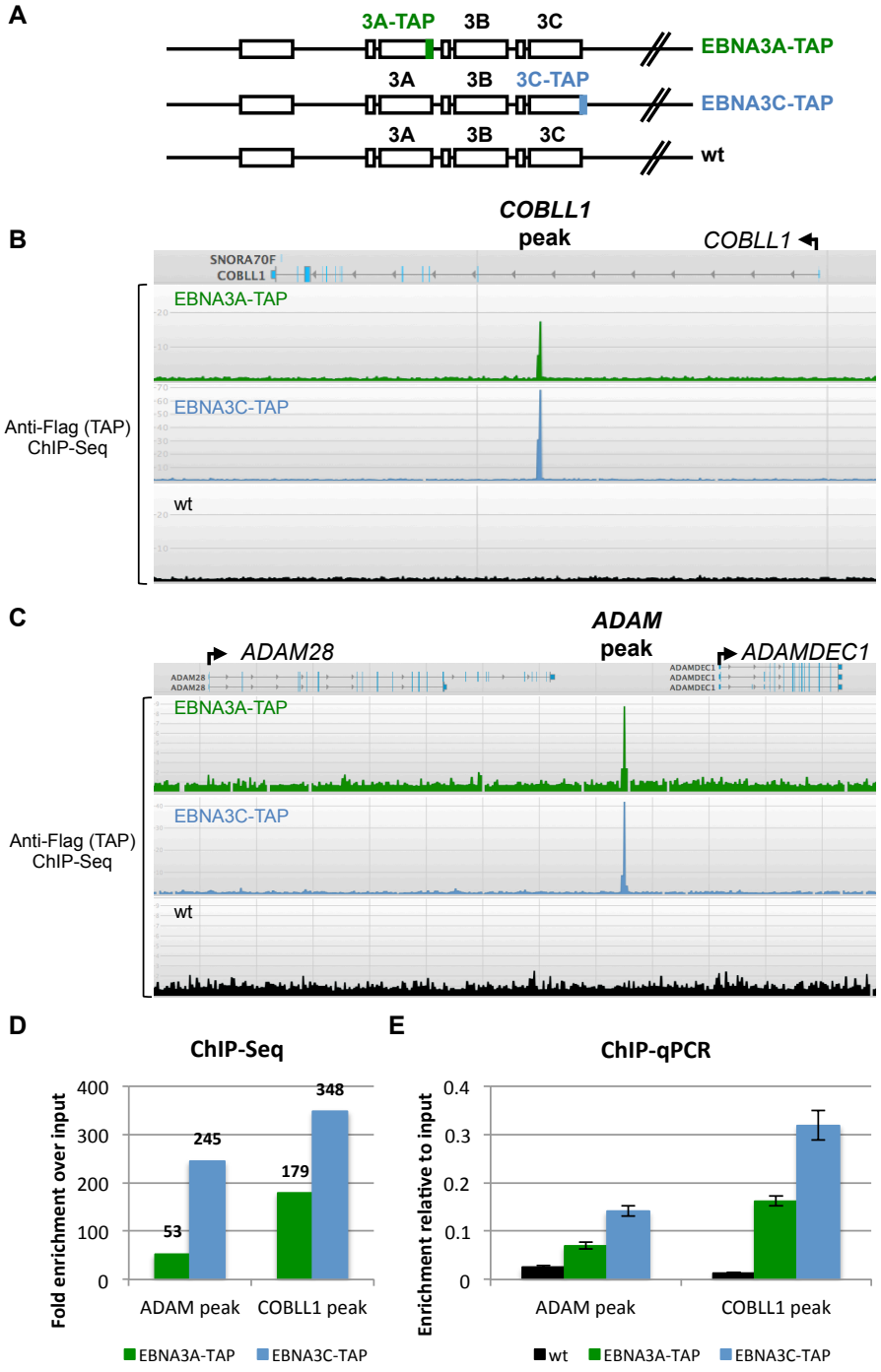


Figure 3.7: EBNA3A and EBNA3C binding at *COBLL1* and *ADAM28/ADAMDEC1*.

(**A**) Overview of recombinant EBV encoding either TAP-tagged (two Strep and one Flag tag) EBNA3A or EBNA3C or wild type (wt) untagged EBNA3s. Cell lines created from these individual viruses were used in anti-Flag ChIP-Seq or ChIP-qPCR experiments. (**B**) Anti-Flag ChIP-Seq on cell lines as indicated and shown for the *COBLL1* locus using Savant genome browser. (**C**) As in (**B**) but shown for the *ADAM28/ADAMDEC1* locus. (**D**) Fold enrichment of EBNA3A-TAP and EBNA3C-TAP as determined by ChIP-Seq at the *ADAM* and *COBLL1* peaks. (**E**) Verification EBNA3 occupancy at the *ADAM* and *COBLL1* peaks by anti-Flag ChIP-qPCR. Dr Kostas Paschos performed the ChIP-Seq and Dr Quentin Bazot the ChIP-qPCR shown here.

3.2.4.2 Changes of epigenetic histone marks correlate with gene expression levels

In order to understand better the sequence of events that led to such robust repression of *COBLL1* and the *ADAM28*/*ADAMDEC1* locus, my aim was to employ ChIP to identify changes in epigenetic marks, factors involved in the regulation and potentially the sequence of events. For this, I designed primers across the *COBLL1* and *ADAM28/ADAMDEC1* locus. The regions around the TSS of all three genes, the EBNA3 binding sites and some additional control sites within each locus were selected. *COBLL1* is located on chromosome 2, has multiple transcript variants and a CpG island around the TSS (Figure 3.8A). *ADAM28* and *ADAMDEC1* are encoded on chromosome 8 and also have multiple transcript variants, but no distinct CpG island (Figure 3.8F). All designed ChIP-qPCR primer pairs were assessed for amplification efficiency and specificity, determined by standard curves derived from serial dilutions and dissociation curves, respectively. Next, various antibodies recognising different epigenetic histone modifications or factors were assayed in ChIP experiments on 3CHT LCL grown for a prolonged time either in the presence or absence of HT. From these preliminary experiments, ChIP antibodies were selected and ChIP was performed on samples harvested throughout the 3CHT A13 time-course.

ChIP for the activation-associated histone marks H3K4me3, H3K9ac and H3K27ac showed a reduction predominantly at the TSS of all three genes after activation of EBNA3C (Figure 3.8B, C, D, G, H and I). Analysis of the repressive H3K27me3 mark revealed an increase mainly at the TSS of *COBLL1* (Figure 3.8E), but at lower levels also across the *ADAM28*/*ADAMDEC1* locus (Figure 3.8J). These consistent changes in epigenetic histone marks could not be observed at *GAPDH* or *Myoglobin*, two controls for expressed and repressed genes, respectively.

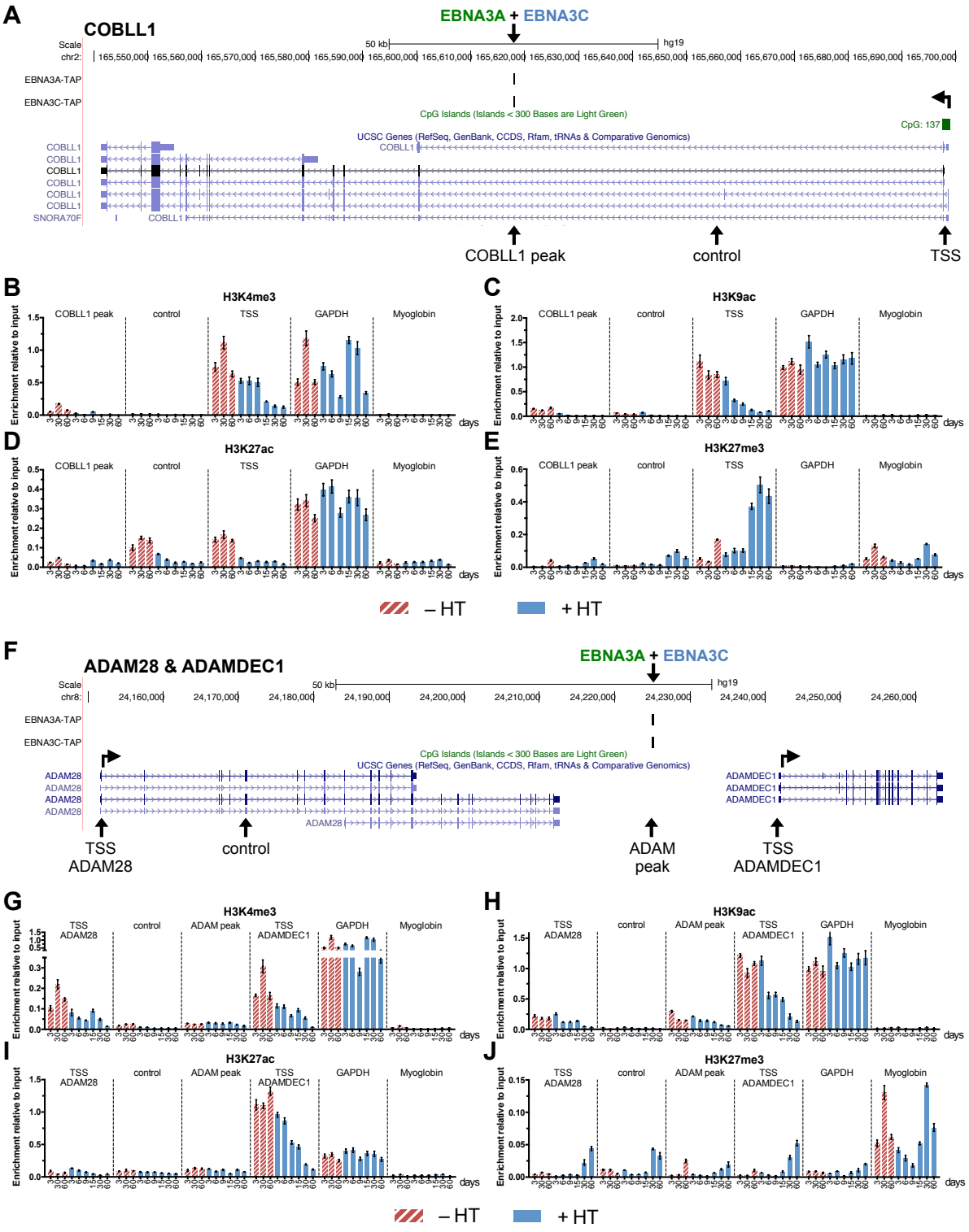


Figure 3.8: Changes in epigenetic marks at the *COBLL1* and *ADAM28/ADAMDEC1* loci after activation of EBNA3C during the 3CHT A13 time-course.

(**A**) UCSC genome browser overview of *COBLL1* genomic locus showing ChIP-Seq tracks for EBNA3A-TAP and EBNA3C-TAP, the TSS (horizontal arrow) and CpG islands. The locations of ChIP-qPCR primer pairs at the TSS, the EBNA3-binding site and a control site (located approximately halfway between the TSS and the EBNA3-binding site) are indicated below the UCSC gene track. (**B-**

E) ChIP for H3K4me3 (**B**), H3K9ac (**C**), H3K27ac (**D**), H3K27me3 (**E**) on samples from 3CHT A13 time-course at locations within the *COBLL1* locus, at *GAPDH* or *myoglobin* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample. (**F-J**) As A-E but for the *ADAM28/ADAMDEC1* locus. This is a representative of two independent biological replicates (see Appendix, Figure 7.3 and Figure 7.4).

Very similar results were obtained from the biological replicate 3CHT C19 time-course (Appendix, Figure 7.3 and Figure 7.4), which allowed a more detailed analysis between the changes in epigenetic histone marks and their effect on gene expression. For this, ChIP enrichment levels of epigenetic histone marks at the TSS of all three genes were normalised for each of the two 3CHT time-courses. Activation-associated histone marks were expressed relative to the first time point –HT (day 3) and repressive histone marks relative to the last time point +HT (day 30). Both time-courses were used to calculate mean values and standard deviations (SD), which revealed that at the TSS of all three genes loss of the activation-associated histone marks H3K9ac, H3K27ac and H3K4me3 preceded any increase in the repressive histone mark H3K27me3 (Figure 3.9A-C top).

Comparison between the changes in histone marks and corresponding mRNA expression levels (Figure 3.9A-C bottom) of each gene over time revealed that the initial repression of mRNA expression correlated well with the loss of all three activation-associated histone modifications (H3K9ac, H3K27ac and H3K4me3) and appeared to be independent of the appearance of the repressive H3K27me3 modification, which was mainly deposited after most gene expression was already repressed.

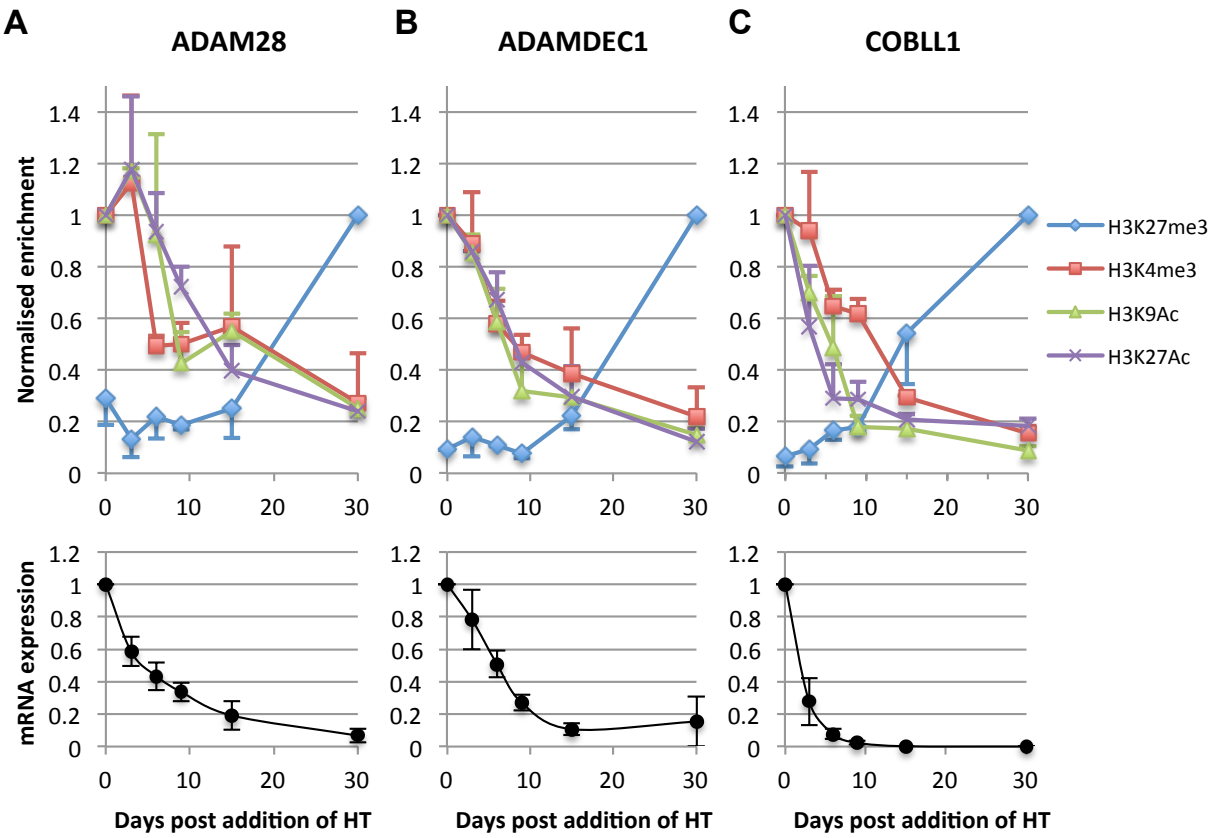


Figure 3.9: Correlation of changes of epigenetic marks at the TSS of *ADAM28*, *ADAMDEC1* and *COBLL1* with gene expression.

ChIP-qPCR values (top) at the TSS of *ADAM28* (**A**), *ADAMDEC1* (**B**) and *COBLL1* (**C**) from 3CHT A13 and 3CHT C19 time-courses were normalised and displayed as mean values ± SD. Activation-associated histone marks H3K4me3, H3K9ac and H3K27ac from the first time-point –HT (day 3) were set to 1 and were used as equivalent for day 0. The repressive histone mark H3K27me3 levels of the last time point +HT (day 30) were set to 1. Mean values ± SD from both replicate time-courses are shown. For better visual clarity, error bars are colour-matching the corresponding histone mark and only upper bars are shown for activation-associated marks and lower bars for H3K27me3. mRNA expression data for each gene (bottom) are shown as mean values ± SD from three replicate time courses (3CHT A13 rep 1+2 and 3CHT C19).

3.2.4.3 Factor recruitment to the COBLL1 or ADAM28/ADAMDEC1 loci

Having seen that changes to epigenetic histone modifications were concurrent for the EBNA3C-mediated gene repression of both *ADAM*s and *COBLL1*, my aim was to determine some of the factors involved in the repression. Since H3K27me3 is catalysed by PRC2 and polycomb complexes were previously found to be involved in the repression of *BCL2L11* (Paschos et al., 2012) and *CDKN2A* (Skalska et al., 2010), ChIP for PRC2 family member SUZ12 was performed. Consistent with the increase in H3K27me3, SUZ12 could be detected at the TSS of *COBLL1* after activation of EBNA3C (Figure 3.10B). However, compared to SUZ12 levels at the TSS of *COBLL1*, no direct enrichment of SUZ12 could be detected across the *ADAM* locus (Figure 3.10F), despite the observed increase in H3K27me3 levels albeit to low levels (Figure 3.8J). In contrast and rather unexpected, ChIP for the PRC1 family member BMI1 (PRC1.4) revealed recruitment of BMI1 to both *COBLL1* peak and *ADAM* peak, but not the TSS of *COBLL1*, with highest BMI1 levels nine days after EBNA3C activation (Figure 3.10C and G).

Finally, since EBNA3C cannot bind directly to DNA and RBPJ is the best-characterised DNA binding factor to which all EBNA3s bind, ChIP for RBPJ was performed. This showed, again rather surprisingly, that RBPJ only accumulated on both the *COBLL1* and *ADAM* peaks, when EBNA3C was functional with highest RBPJ levels six days after activation of EBNA3C (Figure 3.10D and H).

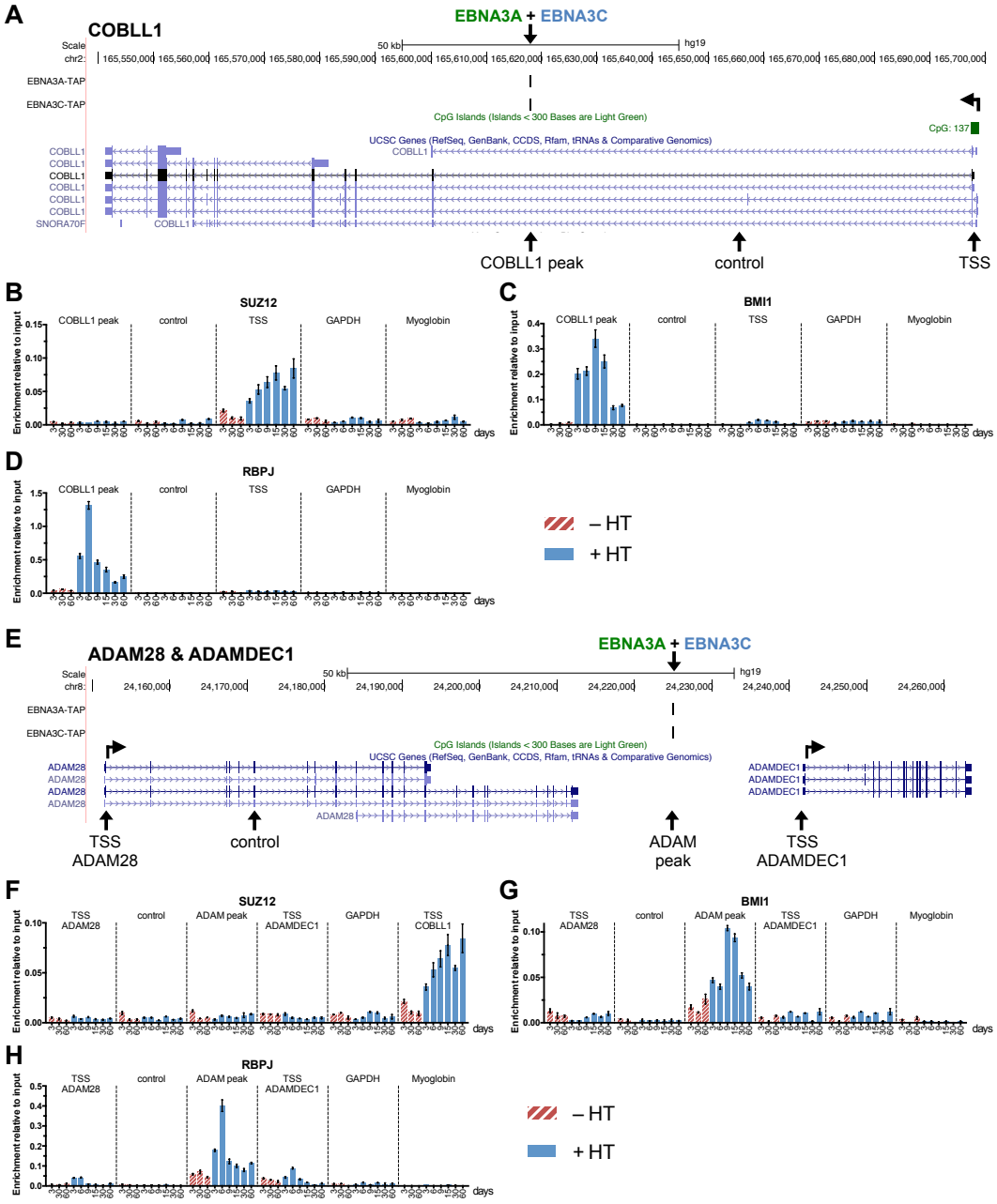


Figure 3.10: Factor binding to the *COBLL1* and *ADAM28/ADAMDEC1* loci during the 3CHT A13 time-course.

(**A**) UCSC genome browser overview of *COBLL1* locus showing ChIP-Seq tracks for EBNA3A-TAP and EBNA3C-TAP, the TSS (horizontal arrow) and CpG islands. The locations of ChIP-qPCR primer pairs at the TSS, the EBNA3-binding site and a control site (located approximately halfway between the TSS and the EBNA3-binding site) are indicated below the UCSC gene track. (**B-D**) ChIP for SUZ12 (**B**), BMI1 (**C**) and RBPJ (**D**) on samples from 3CHT A13 time-course at locations within the *COBLL1* locus, at *GAPDH* or *myoglobin* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample. (**E-H**) As A-D, but for the *ADAM28/ADAMDEC1* locus. This is a representative of two independent biological replicates (see Appendix, Figure 7.3 and Figure 7.4).

The similarity to the ChIP results from the biological replicate time-course 3CHT C19 (Appendix, Figure 7.3 and Figure 7.4) allowed a more detailed analysis of the factor binding using the combined results of the replicate time-courses. For this, enrichment levels for each factor were expressed relative to the time-point with the highest enrichment level during each time-course.

This revealed that maximal SUZ12 levels were reached at the TSS of *COBLL1* by 6-9 days (Figure 3.11A), which preceded the substantial increase in H3K27me3 at this site occurring about 15 days after activation of EBNA3C (Figure 3.9C). High levels of SUZ12 were maintained at the TSS of *COBLL1* over a period of at least 60 days after activation of EBNA3C (Figure 3.10B). This was different for BMI1, in what appeared to be a more transient recruitment to *ADAM* peak and *COBLL1* peak (Figure 3.11B). Maximal BMI1 levels were achieved about nine days after activation of EBNA3C and subsequently dropped, but stayed significantly higher than in cells with inactive EBNA3C. Comparing the recruitment profiles of SUZ12 to the TSS of *COBLL1* and BMI1 to the *COBLL1* peak, it appeared that both were recruited at similar times, but to distinct sites.

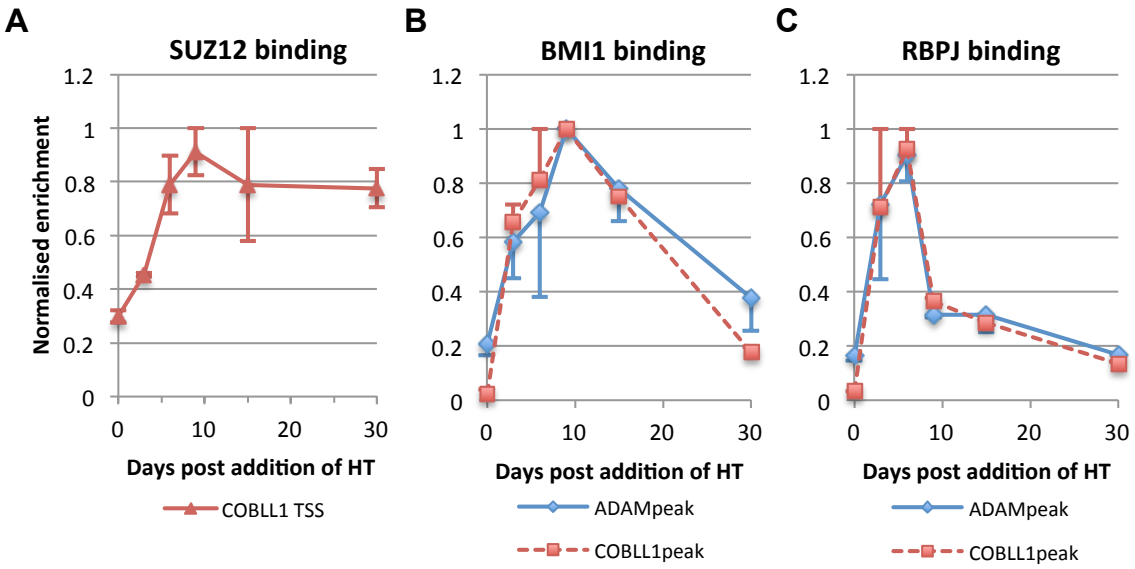


Figure 3.11: Binding of SUZ12 to the TSS of *COBLL1* and BMI1 and RBPJ to *ADAM* peak and *COBLL1* peak after activation of EBNA3C.

ChIP-qPCR values for SUZ12 at the TSS of *COBLL1* (**A**), BMI1 (**B**) and RBPJ (**C**) at the *ADAM* peak and *COBLL1* peak from the two biological time-courses (3CHT A13 and C19) were normalised by

setting the maximal enrichment level during each time-course to one and by using the first time-point –HT (day 3) as equivalent for day 0. Mean values ± SD from both replicate time-courses are shown. For better visual clarity, error bars are colour-matched to either *ADAM* peak (only lower bars displayed) or *COBLL1* peak (only upper bars displayed). For BMI1, day nine of the C19 time-course was treated as an outlier and was not taken into account for the analysis.

Recruitment of RBPJ again seemed to be transient to both *ADAM* and *COBLL1* peaks with highest enrichment levels found on day 3-6 after activation of EBNA3C and subsequently decreasing, but again remaining above the levels seen in cells with inactive EBNA3C (Figure 3.11C). Therefore, the recruitment of RBPJ seemed to precede the recruitment of BMI at both *ADAM* peak and *COBLL1* peak.

Next, because of the observation that both BMI1 and RBPJ appeared to be transiently recruited to both *ADAM* peak and *COBLL1* peak, global BMI1 and RBPJ protein levels were assessed by western blot analysis in samples from the two biological replicate time-courses (Figure 3.12 and Figure 7.2). This showed that there were no consistent changes to the global protein levels of BMI1, SUZ12 or RBPJ that would explain the transient increases of BMI1 and RBPJ at the *ADAM* peak and the *COBLL1* peak.

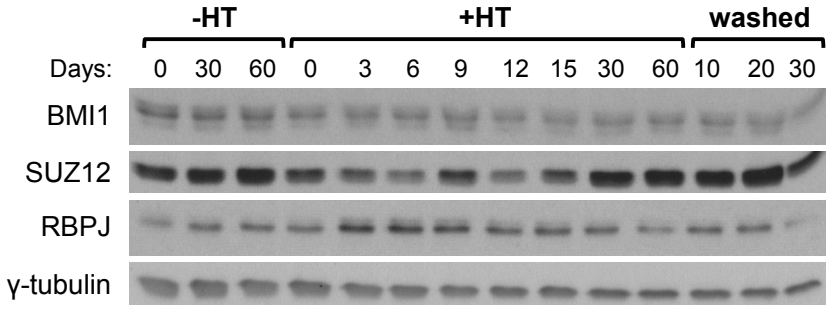


Figure 3.12: Western blot analysis of 3CHT A13 time-course.
Time-course using 3CHT A13 LCL. Cells were grown over 60 days either in absence of HT (-HT), presence of HT (+HT) or with HT removed after 30 days +HT (washed). 30 µg of RIPA extracts from selected samples were analysed for the expression of BMI1, SUZ12 and RBPJ by western blot along γ-tubulin as loading control.

3.2.5 Recapitulation of repression in transient reporter assays

In order to assess whether it is possible to recapitulate the repression of all three genes in transient reporter assays, I cloned the promoter regions of *ADAM28*, *ADAMDEC1* and *COBLL1* upstream of the luciferase gene in the pGL3-basic vector either in the presence or absence of the *ADAM* peak (for *ADAM28* and *ADAMDEC1*, see Figure 3.13A for *ADAM28*) or the *COBLL1* peak (Figure 3.14A). For this, the individual promoter regions or the EBNA3 binding peaks were PCR amplified from genomic DNA extracted from LCL GM12878, a tier 1 LCL used in the ENCODE project, introducing restriction enzyme recognition sequences for subsequent cloning into the pGL3-basic vector (see Material and Methods). The orientation and sequence of the inserts were verified for all luciferase vectors by DNA sequencing and restriction enzyme digests (see Material and Methods).

Preliminary electroporations were performed to assay the luciferase activity of each vector in various cell lines. Luciferase vectors based on the promoter region of ADAM28 resulted in very robust luciferase activity upon electroporation into the EBV-negative BL cell line DG75 (Ben-Bassat et al., 1977) compared to the promoter-less pGL3-basic vector. Luciferase activity was further enhanced by the presence of the *ADAM* peak DNA sequence (Figure 3.13B). Luciferase vectors containing the *ADAMDEC1* promoter region alone did not result in luciferase expression and the presence of the *ADAM* peak resulted only in ~5-fold higher luciferase activities compared to the promoter-less pGL3-basic vector (data not shown). The reasons for this discrepancy between *ADAM28* and *ADAMDEC1* constructs are unclear. It is possible that this was caused by the cloning strategy for the promoter region that was based on the 1kb sequence upstream of the RefSeq gene annotation, which might be too far away from the actual TSS of *ADAMDEC1*. The luciferase expression levels were not high enough to study repression, which is why the *ADAMDEC1* constructs were not used in further assays. Luciferase vectors based on the *COBLL1* promoter resulted in very low luciferase activity in DG75 that do not express endogenous COBLL1 (data not shown). However, robust luciferase activity was observed in the EBV-positive, but EBNA3C-null, cell

line Raji (Allday et al., 1993) that express endogenous COBLL1. Again, the presence of the *COBLL1* peak DNA sequence further enhanced luciferase activities (Figure 3.14B).

Next, in order to determine whether the repression of *ADAM28* by EBNA3A and EBNA3C and the repression of *COBLL1* by EBNA3C could be recapitulated in the transient reporter assays, the dynamic range of these assays was determined by co-electroporating decreasing amounts of EBNA3 expression plasmids along with the luciferase vectors that contain the EBNA3 peak DNA sequences. This revealed that the workable range was between 1 ng and 1 µg of co-electroporated EBNA3 expression plasmid representing approximately one plasmid and 1,000 plasmids per cell, respectively (data not shown).

For the pGL3-*ADAM28*-peak construct, increasing amounts of EBNA3A or EBNA3C, but not EBNA3B, resulted in a reduction of luciferase activity (Figure 3.13C), which is consistent with the results of the primary B cell infection (Figure 3.2A). This repression of luciferase activity was not seen for the pGL3-*ADAM28* construct where only the *ADAM28* promoter is present, but not the *ADAM* peak (data not shown). For the pGL3-*COBLL1*-peak construct, only increasing amounts of EBNA3C, but not EBNA3A or EBNA3B, resulted in a reduction of luciferase activity (Figure 3.14C), which is again consistent with the primary B cell infections (Figure 3.2C). As for the *ADAM28* constructs, without the *COBLL1* peak DNA sequence present in the luciferase vector repression was not observed (data not shown).

An advantage of transient reporter assays is that truncation mutants can be employed in order to try to determine which part of EBNA3C is necessary for the repression of the *ADAM28* and *COBLL1* constructs. Therefore, expression plasmids encoding either an N-terminal part of EBNA3C (amino acids 11-549) or a C-terminal part (amino acids 550-992) were used instead of full-length EBNA3C. These were constructed and used previously. This revealed that the N-terminal part of EBNA3C was able to repress both the pGL3-*ADAM28*-peak (Figure 3.13D) and pGL3-*COBLL1*-peak (Figure 3.14D) constructs, whereas the C-terminal part of EBNA3C failed to repress either construct (Figure 3.13E and Figure 3.14E). The repression of the *ADAM*28 construct by the N-terminal part of EBNA3C seemed to be

impaired compared to full-length EBNA3C, however, luciferase activities of transfections using lower amounts of the N-terminal expression plasmid displayed higher levels compared to the balanced empty vector control (Figure 3.13D). The reasons for this are unknown and for the *COBLL1* construct no such discrepancy could be observed (Figure 3.14D). Therefore, it appeared that the N-terminal part of EBNA3C (amino acids 11-549) was sufficient to repress *ADAM28* and *COBLL1* luciferase constructs.

One of the factors that bind to the N-terminal part of EBNA3C is RBPJ (see Figure 1.10). The observation that this part of EBNA3C was required for the successful repression of the *ADAM28* and *COBLL1* reporters and the fact that RBPJ was recruited or stabilised at the EBNA3-binding site at both endogenous loci (Figure 3.11C), prompted a closer examination of the importance of RBPJ in EBNA3C-mediated repression. For this reason, RBPJ-null DG75 cells [SM224.9, Figure 3.13F, (Maier et al., 2005)] were used for the transient luciferase assays instead of RBPJ-competent DG75 used before. In the absence of RBPJ, both full-length EBNA3A and EBNA3C are unable to repress luciferase activities of the pGL3-*ADAM28*-peak construct (Figure 3.13G). Unfortunately, repression of *COBLL1* construct could not be analysed in these cells due to the low luciferase activity of this construct in DG75.

However, during the course of this project Adam Gillman created the RBPJ binding mutant (BM) EBNA3C (see Introduction and Figure 1.10B and C), which I cloned into a pCDNA3 expression vector. This mutant EBNA3C failed to repress luciferase activity when co-transfected along the pGL3-*ADAM28*-peak construct in RBPJ-competent DG75 (Figure 3.13H) or along the pGL3-*COBLL1*-peak construct in Raji cells (Figure 3.14F). This showed that repression of both *ADAM28* and *COBLL1* in transient assays is dependent on the ability of EBNA3C to interact with RBPJ.

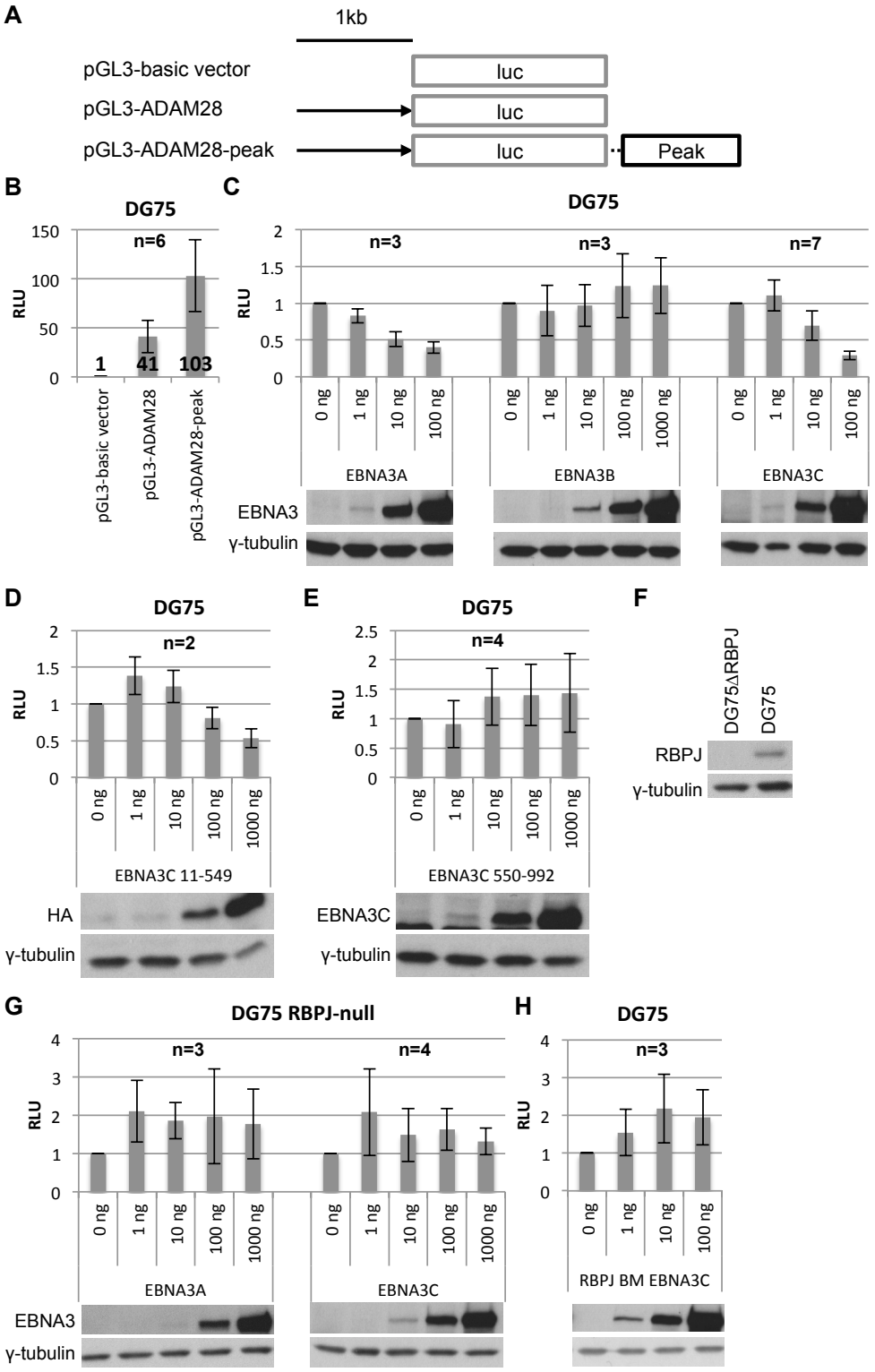


Figure 3.13: Transient reporter assays of *ADAM28* constructs – *ADAM* peak acts as an enhancer and repression by EBNA3A and EBNA3C is dependent on *ADAM* peak and RBPJ.
(**A**) Schematic overview of luciferase vectors used in the transient reporter assays. The 1 kb promoter region upstream of the TSS of *ADAM28* (black arrow) and 1 kb around the *ADAM* peak (black box) was cloned upstream and downstream, respectively, of the luciferase gene (grey box) in the pGL3-basic vector. (**B**) Luciferase reporter assays in DG75 after electroporation of 1 µg of reporter vector as

indicated and expressed relative to pGL3-basic vector. (**C**) Luciferase reporter assays in DG75 after electroporation of 1 µg pGL3-*ADAM28*-peak and increasing amounts of EBNA3 expression plasmids as indicated. (**D**) As in C, but co-transfecting increasing amounts of expression plasmid encoding the N-terminal half of EBNA3C (amino acids 11-549). (**E**) As in C, but co-transfecting increasing amounts of expression plasmid encoding the C-terminal half of EBNA3C (amino acids 550-992). (**F**) Western blot analysis of RBPJ in RBPJ-null DG75 (DG75ΔRBPJ) and RBPJ-competent DG75. (**G**) As in C, but in RBPJ-null DG75. (**H**) As in C, but co-transfecting increasing amounts of expression plasmid encoding the RBPJ binding mutant (BM) of EBNA3C. (**C-H**) Western blots for corresponding EBNA3 proteins or γ-tubulin as loading control. The N-terminal half of EBNA3C expressed as HA-fusion protein was detected by anti-HA antibody. Luciferase units were normalised to β-galactosidase units of the same transfection and are shown as mean relative luciferase units (RLU) ± SD.

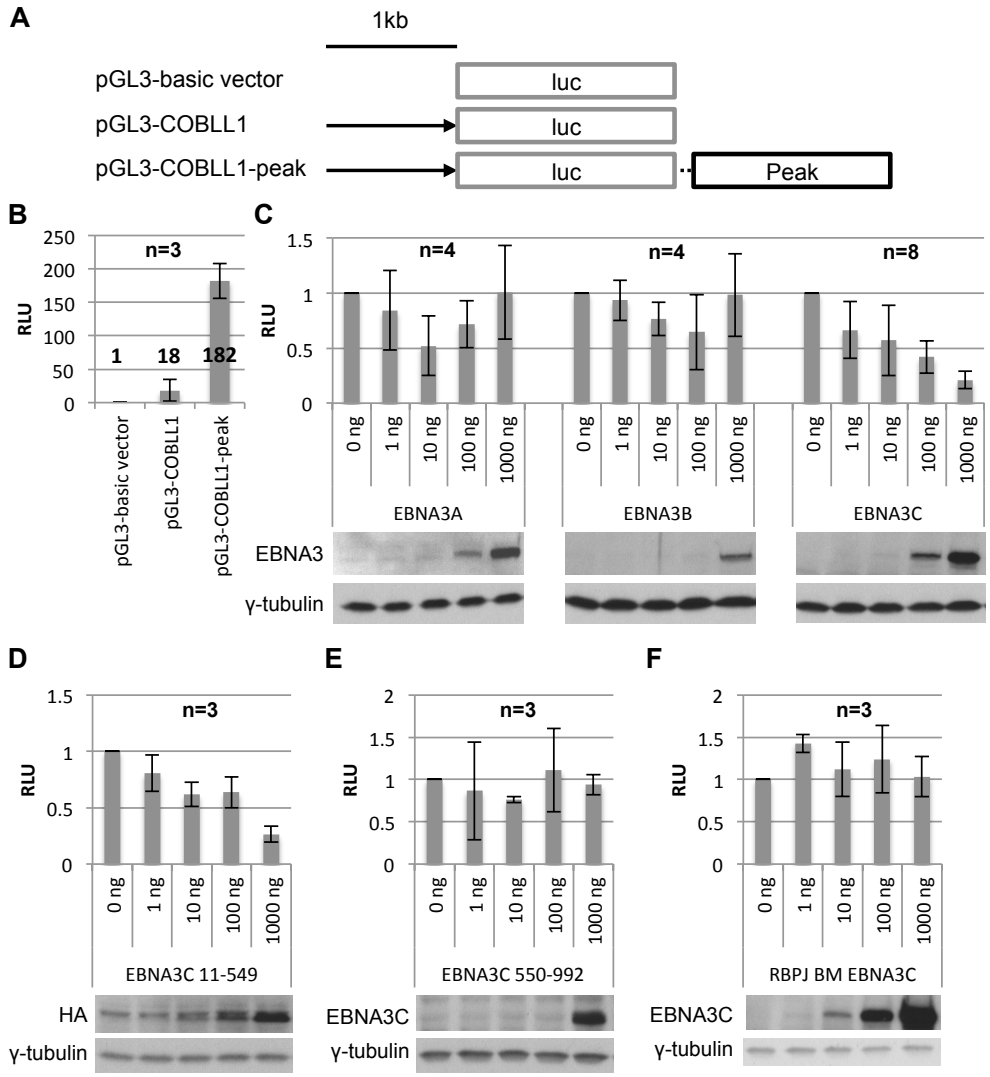


Figure 3.14: Transient reporter assays of *COBLL1* constructs – *COBLL1* peak acts as an enhancer and repression by EBNA3C is dependent on *COBLL1* peak and RBPJ.

(**A**) Schematic overview of luciferase vectors used in the transient reporter assays. The 1 kb promoter region upstream of the TSS of *COBLL1* (black arrow) and 1.5 kb around the *COBLL1* peak (black

box) were cloned upstream and downstream, respectively, of the luciferase gene (grey box) in the pGL3-basic vector. (**B**) Transient luciferase reporter assays in Raji after electroporation of 1 µg of reporter vector as indicated and expressed relative to pGL3-basic vector. (**C**) Transient luciferase reporter assays in Raji after electroporation of 1 µg pGL3-*COBLL1*-peak and increasing amounts of EBNA3 expression plasmids as indicated. (**D**) As in C, but co-transfecting increasing amounts of expression plasmid encoding the N-terminal half of EBNA3C (amino acids 11-549). (**E**) As in C, but co-transfecting increasing amounts of expression plasmid encoding the C-terminal half of EBNA3C (amino acids 550-992). (**F**) As in C, but co-electroporating increasing amounts of RBPJ binding mutant (BM) of EBNA3C. (**C**-**F**) Western blots show corresponding EBNA3 protein and γ-tubulin as loading control. The N-terminal half of EBNA3C expressed as HA-fusion protein was detected by anti-HA antibody. Luciferase units were normalised to β-galactosidase units of the same transfection and are shown as mean relative luciferase units (RLU) ± SD.

3.2.6 Repression is dependent on interaction between RBPJ and EBNA3C

Finally, in order to determine whether binding of EBNA3C to RBPJ was necessary for the repression of endogenous *ADAM28* and *COBLL1* – as suggested by the transient reporter assays – but also *ADAMDEC1* in the context of infection, Adam Gillman created the recombinant EBV carrying the RBPJ BM EBNA3C (Figure 1.10B and Figure 1.17) and used this virus to infect primary B cells. RNA samples were harvested every 5 days from the time of infection. Consistent with the results of the transient luciferase reporter assays and the ChIP studies, infection with the RBPJ BM EBNA3C virus revealed that when EBNA3C cannot bind to RBPJ it was unable to repress *ADAM28*, *ADAMDEC1* or *COBLL1* (Figure 3.15A-C). Regarding the expression levels of all three genes, the RBPJ BM EBNA3C virus behaved similar compared to EBNA3C KO virus, whereas EBNA3C revertant and wild-type viruses resulted in robust repression of all three genes as seen in the previous primary B cell infections (Figure 3.2). As before, the control gene *ALAS1* did not show these differences in gene expression between the various virus infections (Figure 3.15D).

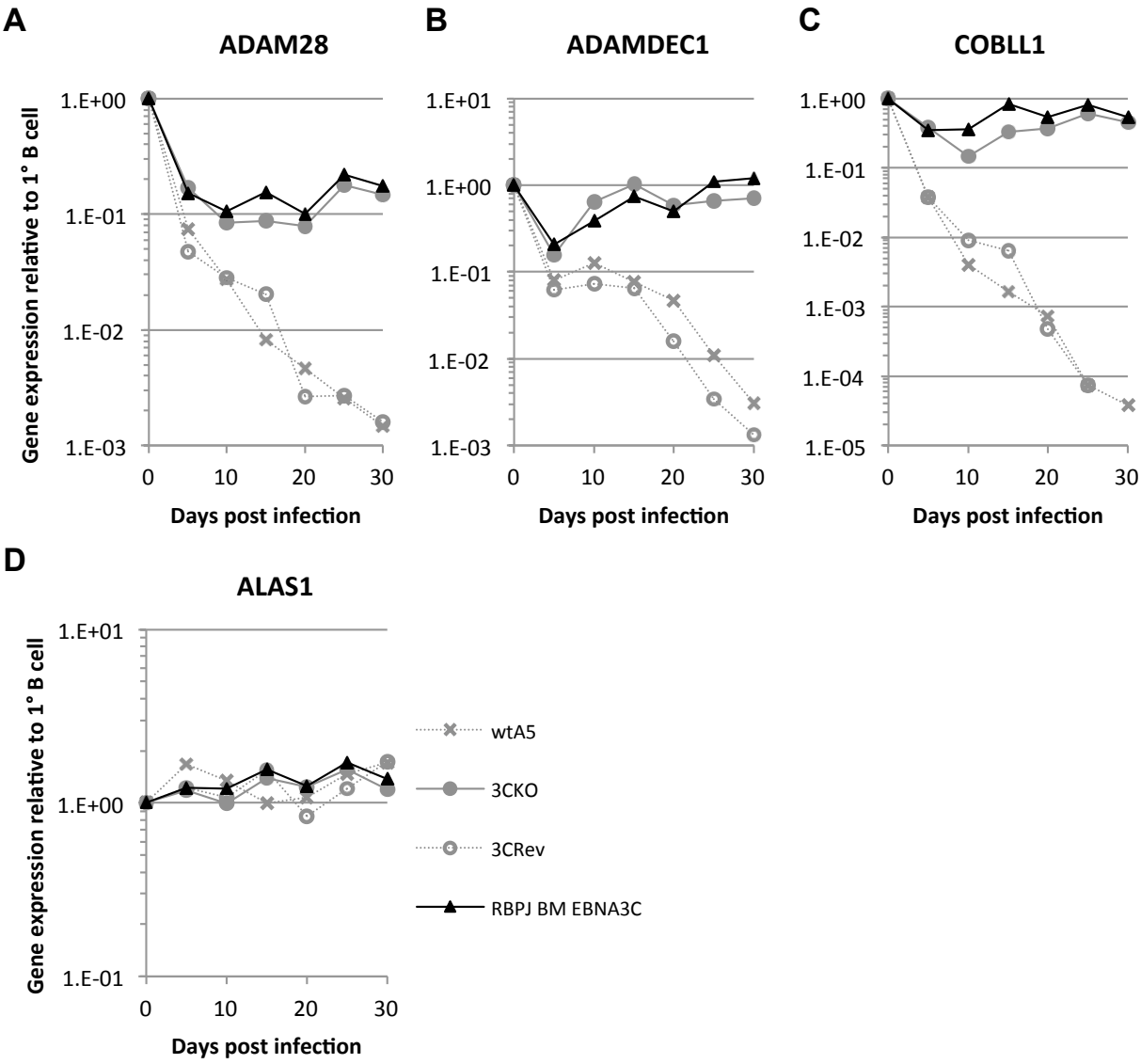


Figure 3.15: Ability of EBNA3C to bind to RBPJ is essential for the repression of *ADAM28*, *ADAMDEC1* and *COBLL1*.

Infection of purified primary B cells with wild type (wtA5), EBNA3C knock-out (3CKO), EBNA3C revertant (3CRev) or RBPJ binding mutant (BM) EBNA3C recombinant EBV. Over a period of 30 days post infection, RT-qPCR gene expression for *ADAM28* (**A**), *ADAMDEC1* (**B**), *COBLL1* (**C**) and *ALAS1* (**D**) was normalised to *GNB2L1* and shown relative to uninfected primary B cells. This is a single experiment that was performed and analysed by Adam Gillman.

3.3 Summary

- In newly EBV-infected primary B cells *ADAM28* & *ADAMDEC1* are repressed by EBNA3A & EBNA3C and *COBLL1* is repressed only by EBNA3C.
- EBNA3C is the main repressor of all three genes.
- EBNA3C-conditional LCL can be used to reliably recapitulate EBNA3C-mediated gene repression seen in newly EBV-infected primary B cells.
- EBNA3C-mediated gene repression follows a highly exponential repression profile.
- Gene repression by EBNA3C seems to take place in two steps:
    1. Initial repression is associated with loss of activation-associated histone marks resulting in the loss of most mRNA transcription.
    2. Subsequent deposition of repressive histone marks by recruited polycomb repressive complexes maintains or further enhances repressed state. PRC1 and PRC2 complexes appeared to be recruited at the same time, but to two distinct sites at the *COBLL1* locus.
- Dynamic recruitment of RBPJ was observed to EBNA3-binding sites only when EBNA3C was functional.
- Repression can be recapitulated in transient reporter assays. This also revealed that the DNA sequences of both EBNA3-binding peaks act as enhancers in the absence of EBNA3 proteins.
- Repression, both in transient reporter assays but more importantly in newly EBV-infected primary B cells, is absolutely dependent on the ability of EBNA3C to bind to RBPJ.
- Most of the results presented in this chapter are now published (Kalchschmidt et al., 2016).

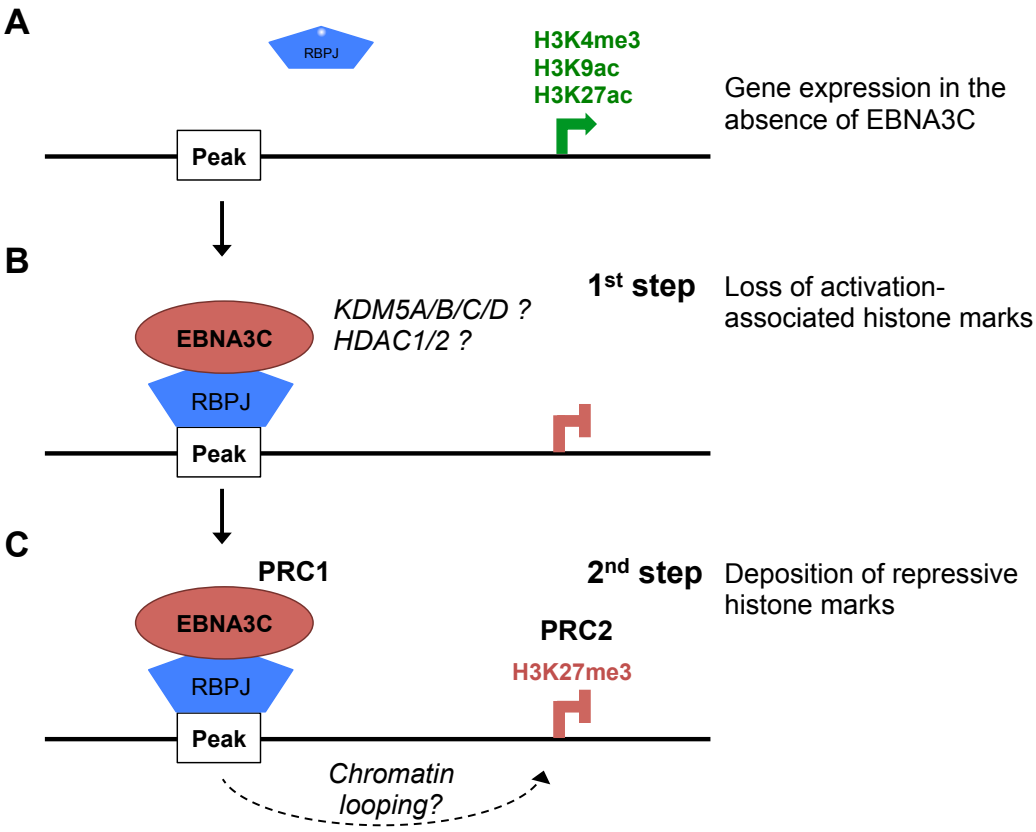


Figure 3.16: Two-step model of EBNA3C-mediated gene repression.

(**A**) Gene expression in the absence of functional EBNA3C through activation-associated histone marks H3K4me3, H3K9ac and H3K27ac. (**B**) Activation of EBNA3C results in the recruitment and/or stabilisation of RBPJ:EBNA3C complexes at the EBNA3C binding peak. EBNA3C potentially recruits H3K4me3 demethylases (KDM5A/B/C/D) and histone deacetylases (HDAC1/2) that catalyse the loss of activation-associated histone marks leading to rapid repression of transcription. (**C**) After most of gene expression has been already repressed, recruitment of polycomb repressive complexes 1 and 2 (PRC1 and PRC2) and deposition of the repressive mark H3K27me3 was observed. Surprisingly, at least at the *COBLL1* locus, recruitment of PRC1 was observed at the EBNA3C binding site, whereas PRC2 was recruited to the gene promoter. It is possible that both sites are involved in chromatin looping that would bring PRC1 and PRC2 together.

3.4 Discussion

The discussion in this Chapter will only focus on issues specific for EBNA3C-mediated gene repression. A broader discussion of the role of EBNA3C in gene regulation in general can be found in Chapter V: General discussion.

Although EBNA3s were known to be transcriptional regulators and EBNA3C is undoubtedly the most studied of the three EBNA3s, very little was known about the precise molecular mechanisms by which EBNA3C regulates gene expression. Here, focusing on two genomic loci that are very robustly repressed by EBNA3C, some of these mechanisms could be identified, which might apply more broadly to EBNA3C-mediated gene repression of host, but possibly also viral gene expression.

3.4.1 Recruitment of RBPJ to chromatin requires functional EBNA3C

Most unexpected, was the dynamic recruitment and/or stabilisation of what are probably RBPJ/EBNA3C complexes at the EBNA3-binding sites at the *ADAM28/ADAMDEC1* and *COBLL1* loci very early after activation of EBNA3C (Figure 3.11C). Unfortunately, ChIP directly for EBNA3C cannot be reliably performed on these samples, because EBNA3C is not TAP-tagged (see Chapter V: General discussion).

This is in contrast to the paradigm that RBPJ is always stably bound at its DNA binding sites to repress gene expression in the absence of Notch signalling (Figure 1.16A). Furthermore, it is not consistent with the previously proposed models that EBNA3s prevent EBNA2-mediated transactivation through disruption of RBPJ binding to DNA (Figure 1.16C) or by replacing EBNA2, but keeping RBPJ stably bound to DNA (Figure 1.16D). No EBNA2 occupancy was observed at the *ADAM* and *COBLL1* peak in the absence of EBNA3C (Figure 7.5).

Recent studies have revealed a more dynamic behaviour of RBPJ binding to DNA. Ectopic activation of Notch signalling can induce *de novo* binding or increased binding of RBPJ in *Drosophila* (Krejcí and Bray, 2007) and mammalian cells (Castel et al., 2013; Wang

et al., 2014) to activate Notch target genes. Increased binding of RBPJ was predominantly observed at distal enhancer sites co-occupied by NICD and resulted in increased p300 occupancy and higher H3K27ac levels (Castel et al., 2013; Wang et al., 2014). Similarly, EBNA2 expression appeared to increase the occupancy of RBPJ at some EBNA2 target genes during EBNA2-mediated gene activation (Portal et al., 2011), which was more recently shown on a large number of genes (Lu et al., 2016). However, to my knowledge no dynamic RBPJ recruitment has been reported during gene repression.

An extended discussion on the role of RBPJ in EBNA3C-mediated gene regulation can be found in Chapter 5.

3.4.2 Two step mechanism of EBNA3C-mediated gene repression

3.4.2.1 First step: Loss of activation-associated histone marks

Rapid loss of activation-associated histone modifications (H3K4me3, H3K9ac and H3K27ac) predominantly at the TSS seemed to coincide with most of the loss of mRNA expression (Figure 3.16A and B). Demethylases are responsible for the removal of methyl groups and KDM5A/B/C/D, which belong to the Jumonji C family of demethylases that can remove mono-, di- or trimethylation on lysine residues, have been shown to be able to catalyse the removal of H3K4me3 (Iwase et al., 2007; Christensen et al., 2007; Klose et al., 2007; Lee et al., 2007; Tahiliani et al., 2007; Kooistra and Helin, 2012). ChIP for KDM5A/B/C/D was not attempted due to limited ChIP material of the 3CHT time-courses and time constrains, but antibodies for all four proteins are commercially available and would need to be tested for specificity before being used in ChIP. The expectation would be that one or several of them are recruited to *ADAM28*/*ADAMDEC1* and *COBLL1* locus very early after activation of EBNA3C.

The removal of acetyl groups on histones is catalysed by HDACs. There are 18 human HDACs grouped into four families according to their homology (HDAC family 1-4) [reviewed in (Haery et al., 2015)]. EBNA3C has been shown to bind to and recruit the class

one HDAC family members 1 and 2 (Radkov et al., 1999; Knight et al., 2003). This makes both of them likely candidates for involvement in the initial repression. ChIP for HDAC1 has been attempted, but was not followed-up due to limited ChIP material of the 3CHT time-courses and time constraints. Again, the expectation would be to observe recruitment very early after activation of EBNA3C.

In addition, both HDAC1 and HDAC2 have been found together with KDM5A/Sin3 or KDM5C/NcoR/REST in multiprotein complexes (Tahiliani et al., 2007; Hayakawa et al., 2007; Hayakawa and Nakayama, 2011). Besides HDAC1 and HDAC2, EBNA3C can bind to both Sin3 and NcoR (Knight et al., 2003). Therefore, it seems likely that EBNA3C recruits multifunctional complexes that are able to remove H3K4me3 and histone acetylation in the initial phase of repression.

ChIPs for these factors might also reveal why the repression of *COBLL1* is faster and more extensive compared to the repression of the *ADAM28*/*ADAMDEC1* locus. A tempting speculation would be that higher levels of EBNA3C on the *COBLL1* peak, recruit more demethylases and deacetylases that are responsible for a faster repression of the *COBLL1* locus compared to the *ADAM28*/*ADAMDEC1* locus where less EBNA3C can be found. Another possibility is that changes in the 3D chromatin organisation (see below) required for gene repression are more quickly established at the *COBLL1* locus compared to the *ADAM28*/*ADAMDEC1* locus.

3.4.2.2 Second step: Recruitment of polycomb proteins and deposition of repressive histone marks

In a second step and what appears to be independent from the initial repression, the PRC1 protein BMI1 was recruited to *ADAM* peak and *COBLL1* peak and the PRC2 protein SUZ12 to the TSS of *COBLL1* (Figure 3.16C). Comparison to gene expression levels revealed that PRC2-catalysed H3K27me3 levels was unlikely to be the direct cause for repression, but rather involved in maintaining or extending the repressive state. It has been a long-standing question in polycomb-mediated repression whether H3K27me3 was a cause

or consequence of gene expression – the results observed here would suggest the latter. This is consistent with the findings of two recent studies. Hosogane and colleagues found that H3K27me3 levels were a consequence of changes in gene expression induced by H-Ras expression in the mouse embryonic fibroblast cell line NIH 3T3 (Hosogane et al., 2013). Unfortunately, changes in activation-associated histone marks were not assessed in this study. Moreover, Riising *et al.* found that PRC2 was dispensable for the initiation of repression, but required for maintenance of the repressed state and that small-molecule inhibition of transcription could trigger recruitment of PRC2 in mouse embryonic stem cells (Riising et al., 2014).

Repeated attempts were made to investigate whether PRC1-catalysed H2AK119ub was deposited at the *ADAM28*/*ADAMDEC1* and *COBLL1* loci, but due to concerns about antibody reliability and specificity this ChIP has not been performed on the A13 and C19 time-courses.

3.4.3 Simultaneous recruitment of PRC1 and PRC2 proteins to two distinct loci

Another surprising finding from the ChIP analysis was that BMI1, probably acting as part of PRC1, was found at the *ADAM* peak and *COBLL1* peak, whereas SUZ12, as part of PRC2, was found at the TSS of *COBLL1*. At the *ADAM28*/*ADAMDEC1* locus, no direct SUZ12 recruitment could be detected although H3K27me3 levels increased at this locus, albeit at much lower levels compared to the *COBLL1* locus. It is possible that SUZ12 binding was therefore below the detection limit or that recruitment of SUZ12 to a discrete site was missed by the choice of primers. This is however unlikely, because previous studies identified that 95% of SUZ12 binding sites localised within 1kb of TSS and that 40% were within 1kb of CpG islands, which by themselves could recruit PRC2, leading to the deposition of H3K27me3 (Lee et al., 2006; Tanay et al., 2007; Ku et al., 2008; Mendenhall et al., 2010). All TSSs were included in the ChIP analysis and only *COBLL1* has a CpG island around its TSS, which might explain the direct recruitment of SUZ12 and the much higher levels of H3K27me3 at *COBLL1* relative to the *ADAM28*/*ADAMDEC1* locus.

Unfortunately, this meant that *COBLL1* was the only locus, where both recruitment of PRC1 proteins and PRC2 proteins could be observed. In contrast to the classical and more recent models of polycomb recruitment (see Introduction), PRC1 and PRC2 complexes appeared to be recruited simultaneously, but to different genomic loci. Higher order chromosome structures can be mediated by polycomb complexes (Lanzuolo et al., 2007; Tiwari et al., 2008a; b) [reviewed in (Cheutin and Cavalli, 2014)], so perhaps repression of *COBLL1* involves chromatin looping between the *COBLL1* peak and the TSS of *COBLL1*, which would bring BMI1 (PRC1) and SUZ12 (PRC2) together (Figure 3.16C). Previously chromosome conformation capture has been successfully used to show repressive loop formation between the *ADAM* peak and the TSS of *ADAM28* and *ADAMDEC1* when EBNA3C was expressed (McClellan et al., 2013). Attempts were made to employ the same technique to analyse the higher-order chromosome structure of the *COBLL1* locus together with the *ADAM* locus as a positive control. Unfortunately, this did not lead to reproducible results. The reasons for this are not known and time constrains did not allow further investigation. However, one possible indication of repressive loop formation between the *COBLL1* peak and the TSS of *COBLL1* or the *ADAM* peak and the TSSs of *ADAM28* and *ADAMDEC1* might be that low levels of BMI1 and RBPJ enrichment could be detected on the TSS of all three genes (Figure 3.10C, D, G and H). It is possible that this resulted from a close spatial proximity between the EBNA3C binding peak and the TSSs when the cells were cross-linked.

Furthermore, it is currently unclear whether polycomb proteins are part of a default repressive mechanism that is activated at these two genomic loci once activation-associated histone marks are removed, or whether they are directly recruited. This and the role of RBPJ in gene regulation mediated by EBNA3C will be discussed in Chapter 5.

3.4.4 Comparisons with EBNA3A-mediated repression

Recently, a similar two-step mechanism of gene repression has been reported for the EBNA3A-mediated repression of *CXCL9* and *CXCL10* (Harth-Hertle et al., 2013). In the

absence of EBNA3A, EBNA2 binds to an intergenic regulatory element and *CXCL9* and *CXCL10* are expressed. Doxycycline-induced overexpression of EBNA3A in EBNA3A KO LCL displaced EBNA2 from the regulatory element resulting in initial de-activation marked by loss of activation-induced histone marks and only subsequently an increase in H3K27me3. This is very similar to the EBNA3C-mediated repression of *ADAM28*, *ADAMDEC1* and *COBLL1*, however, no EBNA2 occupancy and no change in EBNA2 levels could be detected on the *ADAM* peak or *COBLL1* peak during the 3CHT A13 time course or in 3CHT cells grown either in absence or presence of HT for prolonged period of time (Figure 7.5). The results of the luciferase assays show clearly that the DNA sequences of the *ADAM* peak and *COBLL1* peak act as enhancer elements, so in the absence of EBNA3C this might be enough to activate these two loci. Interestingly, EBNA3C seems to target these enhancer elements and somehow converts them into repressive elements. From the luciferase assays it became clear that the *ADAM* peak or the *COBLL1* peak have to be present in the luciferase vectors for repression to take place.

3.4.5 Biological role of *ADAM28*, *ADAMDEC1* and *COBLL1* – why are they repressed by EBV?

Besides the molecular mechanisms of EBNA3C-mediated repression, another question is the biological relevance of the repression of *ADAM28*, *ADAMDEC1* and *COBLL1* for EBV. A closer look at their biological functions might help to explain this.

ADAM28 and ADAMDEC1 are members of the ADAM family of membrane-bound or secreted proteases. Members of this family share a conserved multidomain structure comprised of pro, metalloproteinase, disintegrin, cysteine-rich, epidermal growth factor (EGF)-like, transmembrane and cytoplasmic domains [reviewed in (Edwards et al., 2008; Klein and Bischoff, 2011)]. The pro-domain inhibits the activity of the metalloproteinase domain, ensures correct protein folding and is generally cleaved during the transit through the Golgi secretory system. The metalloprotease domain is responsible for the proteolytic activity, but not all ADAM family members contain active metalloprotease domains. The

disintegrin domain can mediate protein-protein interaction or cell-cell adhesion through interaction with, for example, integrins. The cysteine-rich domain is rather variable between the different members of the ADAM family and faces towards the catalytic site of the metalloproteinase domain, where it is thought to mediate binding to substrates, but also influences desintegrin-integrin interactions. The biological function of the EGF domain is unclear. The cytoplasmic domain of membrane-bound ADAMs is highly variable and is involved in further regulation of ADAM activity and localisation through interaction with intracellular proteins and/or posttranscriptional modifications. With these multiple domains – through regulation of proteolysis, signalling, migration and cell adhesion – ADAM proteins have various roles in human physiology, such as egg-sperm interaction [reviewed in (Cho, 2012)], ectodomain shedding of cytokines or growth factors [reviewed in (Reiss and Saftig, 2009; Taylor et al., 2014)] or regulated intramembrane proteolysis, *e.g.* during Notch signalling where extracellular S2 proteolysis of Notch by ADAM10 or ADAM17 induces the intracellular S3 cleavage [reviewed in (Christian, 2012)]. Deregulation of ADAM expression can result in various diseases, including cancer [reviewed in (Murphy, 2008; Rocks et al., 2008; Mochizuki and Okada, 2007)].

ADAM28 (also known as MDC-L, eMDCII and TECADAM) is highly expressed on B lymphocytes (Roberts et al., 1999; McGinn et al., 2011) and is able to bind to $\alpha_4\beta_1$, $\alpha_4\beta_7$, $\alpha_9\beta_1$ integrins (Bridges, 2001; Bridges et al., 2003; 2005; McGinn et al., 2011) and P-selectin glycoprotein ligand 1 (Shimoda et al., 2007). Furthermore, ADAM28 has a membrane-bound and a secreted isoform (Roberts et al., 1999), both of which are catalytically active (Howard et al., 2001) and able to cleave CD23 (Fourie, 2003), insulin-like growth factor binding protein 3 (Mochizuki et al., 2004), connective tissue growth factor (Mochizuki et al., 2010), von Willebrand factor (Mochizuki et al., 2012), tumour necrosis factor α (Jowett et al., 2012) and CD200 (Twito et al., 2013). Probably as a result of its above listed functions, ADAM28 over expression has been reported in multiple human cancers, *e.g.* in breast carcinoma

(Mitsui et al., 2006), non-small cell lung carcinoma (Ohtsuka et al., 2006; Kuroda et al., 2010; Lv et al., 2012) or acute B cell lymphoblastic leukemia (Zhang et al., 2015).

Compared to the other ADAM family proteins, ADAMDEC1 (ADAM-like, decysin 1) has a non-canonical metalloprotease domain, only contains half of the disintegrin domain and lacks the cysteine-rich region, the EGF-motif, the transmembrane domain and the cytoplasmic tail (Bates et al., 2002). Due to the lack of the transmembrane domain, ADAMDEC1 is strictly secreted. The non-canonical metalloprotease domain is catalytically active and able to cleave α_2-macroglobulin, casein and carboxymethylated transferrin (Lund et al., 2013; 2015). The physiological role of ADAMDEC1 is unknown. Reduced expression of ADAMDEC1 was detected during tumour progression of colorectal cancer (Macartney-Coxson et al., 2008) and in a poor prognosis subgroup of gastric adenocarcinoma (Pasini et al., 2014). Increased expression has been reported within unstable regions of atherosclerotic plaques (Papaspyridonos et al., 2006), in sarcoidosis lung tissues (Crouser et al., 2009) and in craniopharyngioma cells (Xu et al., 2012a). A recent study found high levels of ADAMDEC1 expression in kidney transplants that were rejected by T cell-mediated mechanisms compared to other mechanisms, *e.g.* antibody-mediated rejection (Venner et al., 2014).

The repression of ADAM28 and ADAMDEC1 by EBNA3A and EBNA3C is likely to be part of an immune modification strategy, involving reduced shedding of cytokines and chemokines and reduced cell-to-cell interaction between the EBV infected B cells and immune cells. This would be important during *in vivo* infections. Currently, nothing is known about the *in vivo* role of EBNA3A or EBNA3C. As a part of this PhD project, infectious EBNA3A, EBNA3C and triple EBNA3 KO and Rev viruses were produced at a large scale, to determine their role in a humanised mouse model of EBV infection [reviewed in (Münz, 2015)] in collaboration with Donal McHugh and Dr Christian Münz. Preliminary results, revealed substantially reduced viral loads in animals infected with EBNA3A and EBNA3C

KO viruses compared to wild type or revertant viruses. However, further work is ongoing to determine the activation of immune cells during these infections.

COBLL1 (COBL-like 1, formerly also called KIAA0977 protein) is conserved in vertebrates and its N-terminus shares 37% sequence homology with the actin-nucleation factor Cordon-bleu (COBL) (Mori et al., 2001; Carroll et al., 2003; Ahuja et al., 2007) [reviewed in (Kessels et al., 2011)]. The gene function of *COBLL1* remains unknown.

However, more recently abberant *COBLL1* expression has been repeatedly reported in various cancers and loss of *COBLL1* expression was often associated with disease progression and a poor prognosis. Frame-shift mutations in *COBLL1* have been identified in 24% of colorectal cancers (Mori et al., 2001). Multiple studies also identified *COBLL1* as prognostic a biomarker that was down-regulated in CLL with a non-mutated IgH variable region, which is the subgroup that is linked with poor prognosis. In contrast higher COBLL1 expression was found in CLL with mutated IgH variable region, associated with a more favourable outcome (Haslinger et al., 2004; Abruzzo et al., 2007; Mansouri et al., 2012; Ronchetti et al., 2013). Interestingly, mutations in the functional domain of ADAM28 were also reported in some of these CLL patients (Mansouri et al., 2012). In DLBCL (Tompkins et al., 2013) and malignant pleural mesothelioma (Gordon et al., 2003; 2005; 2009) lower COBLL1 expression was again associated with a less favourable outcome. In prostate cancer, however, *COBLL1* expression was increased in tumours (Singh et al., 2002; Wang and Gotoh, 2009) and in proximal gastric cancers *COBLL1* expressing tumours were associated with disease progression and poor prognosis (Wang et al., 2013).

Beside cancer studies, the *COBLL1* locus has been associated with alcohol addiction (Dick et al., 2010), type 2 diabetes (Manning et al., 2012; Albrechtsen et al., 2013; Randall et al., 2013; Mancina et al., 2013; Desmarchelier et al., 2014; Kraja et al., 2014), thickness of the left Heschl's gyrus (Cai et al., 2014) and periodontitis (Sima et al., 2015). In a biochemical study, COBLL1 was identified by mass spectrometry to interact with nardilysin that binds to H3K4me2 and plays a role in transcriptional regulation through association with

NCoR/SMRT complex and HDAC3 (Li et al., 2012). However, this interaction has still to be confirmed. In conclusion, the biological function of COBLL1 protein is unknown and it remains unclear why EBV via EBNA3C so robustly represses *COBLL1* expression.

3.4.6 Remarks

It is worth noting that the 4th most highly repressed gene by EBNA3C (*RP11-624C23.1,* see Table 1.1) is a long non-coding RNA with four isoforms of different lengths that run on the negative strand across the *ADAM28/ADAMDEC1* locus (Figure 7.9). Analysis of microarray data (http://www.epstein-barrvirus.org.uk) is consistent with the whole locus being co-ordinately regulated by EBNA3A and EBNA3C.

Furthermore, EBNA2 up regulates *ADAMDEC1* (Lucchesi et al., 2008), but not *ADAM28*, probably through binding to an enhancer element proximal to the TSS of *ADAMDEC1* (Figure 7.9). However, EBNA3A- and EBNA3C-mediated repression of the entire *ADAM28*/*ADAMDEC1* locus dominates over induction of *ADAMDEC1* by EBNA2.

3.4.7 Conclusion

The findings presented here are a first step into a better understanding of how EBNA3C represses cellular and perhaps also viral genes. However, further studies including additional EBNA3C-repressed genes are needed in order to determine whether the findings presented here apply to all EBNA3C repressed genes and to fully understand the molecular mechanisms.

The role of EBNA3C in gene regulation in general including the recruitment of RBPJ and the kinetics of the regulation will be further discussed in Chapter V: General discussion.

4 Chapter IV: Gene activation by EBNA3C

4.1 Introduction

Having analysed EBNA3C-mediated gene repression in the previous chapter, this chapter will now focus on EBNA3C-mediated gene activation. Therefore, *AICDA*, the gene encoding AID, was selected based on the previous Affymetrix exon microarray study using EBNA3C-conditional LCL [Table 1.2 and (Skalska et al., 2013)]. This gene will serve as model gene to better understand molecular mechanisms involved in EBNA3C-mediated induction of gene expression. As discussed in the introduction AID expression is controlled on multiple levels (Figure 1.6). Various studies have reported that EBV-infected B cells express AID [e.g. (Gil et al., 2007; Epeldegui et al., 2007)], but currently only LMP1 and EBNA2 have been identified to deregulate AID expression. Over-expression of LMP1 alone induced AID expression in LMP1-conditional BJAB cells resulting in CSR, which could be prevented by inhibiting NFκB signalling (He et al., 2003). More recently, it was suggested that LMP1 activates AID through NFκB signalling and binding of the TF early growth response 1 (EGR-1) to the AID promoter (Kim et al., 2013). EBNA2 on the other hand has been reported to repress expression of AID in three different EBNA2-conditional cell lines, EREB 2-5 cells, BL41 and BJAB cells (Tobollik et al., 2006). The repression of AID by EBNA2 had very similar kinetics compared to the repression of IgM by EBNA2, but no mechanism for the repression has been reported or proposed. EBNA3C has not been previously linked to the regulation of AID expression.

4.2 Results

4.2.1 Validation of *AICDA* as an EBNA3C-activated gene in primary B cells infected with EBV

First of all, in order to validate that *AICDA* is up regulated by EBNA3C and to determine whether EBNA3A or EBNA3B are involved in the regulation, *AICDA* mRNA expression was analysed in primary B cells infected with wild type, knock-out (3AKO, 3BKO,

3CKO or E3KO) or revertant (3ARev, 3BRev, 3CRev, E3Rev) viruses (Figure 4.1A). This revealed that all EBNA3C-competent viruses induced *AICDA* expression over a 30-day period. Only infections with viruses lacking EBNA3C (3CKO or E3KO) failed to induce *AICDA* expression. In GC B cells, BCL6 positively regulates AID expression through repression of miR-155 (Basso et al., 2012). The EBNA3C-mediated induction of *AICDA* was however independent of the expression of *BCL6,* which was highly repressed by all recombinant viruses (Figure 4.1B). Potentially due to the lack of repression by BCL6, expression of *miR-155HG*, a long-precursor transcript that contains miR-155, increased after infection with all EBV recombinant viruses (Figure 4.1C). Although 3CKO-infected cells displayed the highest levels of *miR155-HG*, the level was only marginally higher compared with the other recombinant virus infections and therefore it is unlikely that miR-155 contributes to the low expression levels of *AICDA* in 3CKO-infected cells.

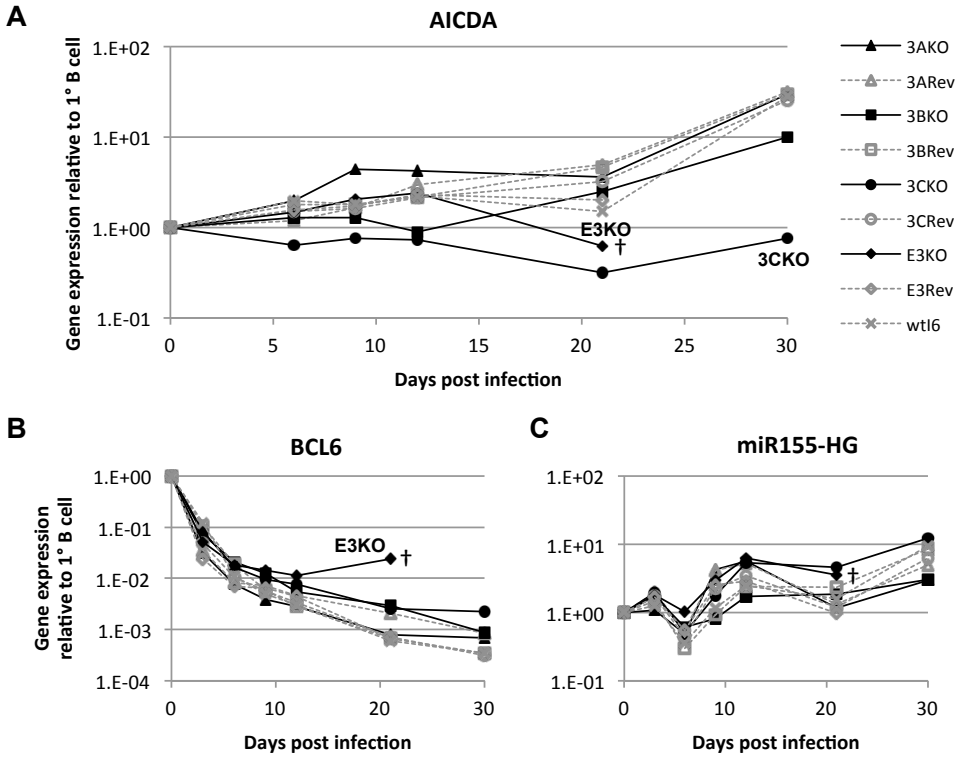


Figure 4.1: Induction of *AICDA* by EBNA3C in primary B cells infected with recombinant EBV. Infection of purified primary B cells with wild type (wtI6), recombinant EBNA3A (3A), EBNA3B (3B), EBNA3C (3C) or triple EBNA3A/B/C (E3) knock-out (KO) or revertant (Rev) viruses. Over a period of 30 days post infection, RT-qPCR gene expression for *AICDA* (**A**), *BCL6* (**B**) and *miR-155-HG* (**C**) was normalised to *GAPDH* and is shown relative to uninfected primary B cells. Primary B cells infected with E3KO could not be cultured past that time-point (†). These results are from a single experiment.

4.2.2 Recapitulation of EBNA3C-mediated gene activation of *AICDA* in EBNA3C-conditional LCL

Having seen that *AICDA* is a genuine target of EBNA3C in newly EBV-infected B cells, the p16-null 3CHT LCL were again used in order to determine whether it is possible to recapitulate the induction of *AICDA* by EBNA3C in these cells. The 3CHT A13 time-course showed increased *AICDA* mRNA levels after activation of EBNA3C with maximal expression levels being attained about 30 days after the addition of HT (Figure 4.2A). This induction was reversible, removal of HT from the culture medium resulted in a decrease of *AICDA* expression back to levels similar to those seen in cells with inactive EBNA3C. The regulation of AID could also be observed at the protein level as revealed by western blotting with two independent antibodies (*e.g.* Figure 4.2B).

When compared to other cells, 3CHT LCL with inactive EBNA3C expressed AID at levels very similar to purified primary B cells, whereas 3CHT LCL with active EBNA3C expressed AID at levels similar to several other B95.8 EBV transformed LCLs (Figure 4.2C and D). Some LCL appeared to express very high levels of AID and it is worth mentioning that LCL GM12878, a LCL used as tier 1 cell line in ENCODE, was one of these. BL derived cell lines (Ramos, Akata, Oku and Namalwa), irrespective of EBV status expressed high levels of AID consistent with them having a GC origin. The two EBNA3C expressing BL lines Oku (Wp restricted) and Namalwa (latency III) express higher levels than Ramos (EBV negative) and Akata (latency I), that do not express EBNA3C. BJAB, a EBV-negative B cell lymphoma derived cell line that is not thought to be GC-derived (Klein et al., 1976; Glazer and Summers, 1985), and K562 derived from a patient with chronic myelogenous leukemia (Lozzio and Lozzio, 1975), expressed low or undetectable levels of AID, respectively.

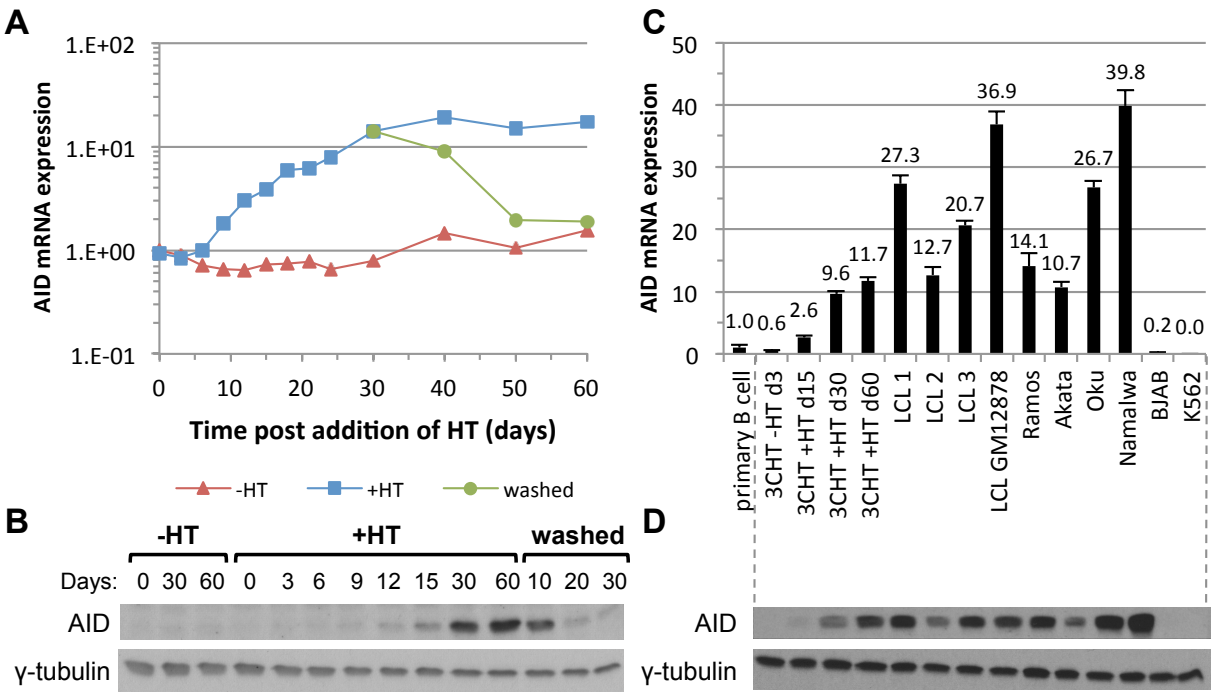


Figure 4.2: Induction of *AICDA* during the 3CHT A13 time-course and comparison across various cell lines.

(**A**) Time-course using EBNA3C-conditional LCL 3CHT A13. Cells were grown over 60 days either in absence of HT (-HT), presence of HT (+HT) or with HT removed after 30 days +HT (washed). RT-qPCR gene expression for *AICDA* was normalised to *GAPDH* and is shown relative to -HT on day 0. (**B**) AID and γ-tubulin protein expression in selected samples of the 3CHT A13 time-course shown in A. (**C**) RT-qPCR AID mRNA expression in 3CHT LCLs, three in house B95.8 EBV transformed LCL, LCL GM12878, Ramos (EBV-negative BL), Akata (EBV-positive BL, latency 1), Oku (EBV-positive, latency 3), Namalwa (EBV-positive BL, latency 3), BJAB (EBV-negative B cell lymphoma) and K562 (EBV-negative erythromyeloblastoid leukemia) was normalised to *GAPDH* and shown relative to primary B cells. Results are shown as mean ± SD of three technical replicates. (**D**) AID and γ-tubulin protein expression in samples shown in C.

The induction of AID was a direct consequence of the HT-induced activation of EBNA3C, because addition of HT to non-conditional 3CKO LCL established from the primary B cell infection did not result in activation of AID (Figure 4.3). HT was functional and caused induction of AID in EBNA3C-conditional 3CHT LCL over the same experiment.

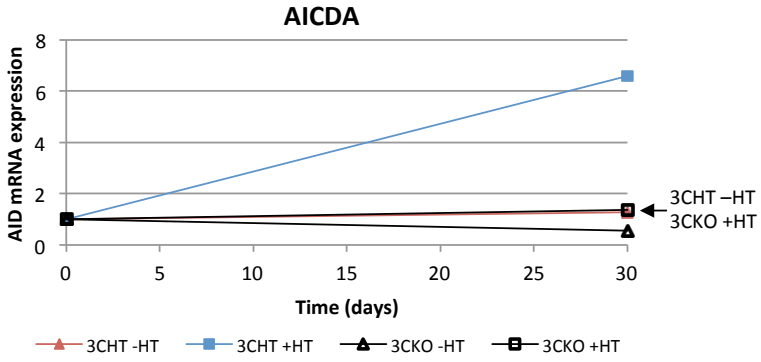


Figure 4.3: HT does not induce AID in non-conditional 3CKO LCL.
Time-courses using non-conditional 3CKO LCL or conditional 3CHT A13 LCL. Cells were grown over 30 days either in the presence or absence of HT. Gene expression of AID was normalised to *GAPDH* and is shown relative to day 0 for 3CHT or 3CKO. This result is from a single experiment.

4.2.3 Comparison between primary B and EBNA3C-conditional LCL

Direct comparisons of *AICDA* gene expression in newly EBV-infected primary B cells with 3CHT LCLs after activation of EBNA3C revealed very similar kinetics and extent (Figure 4.4). Similar to the EBNA3C-repressed genes (see Section 3.2.3), EBNA3C-mediated induction of *AICDA* appeared to follow a highly exponential regulation profile (R^2 > 0.95) over a time period of 30 days when functional EBNA3C was present. Therefore, 3CHT LCL can be also used to recapitulate efficiently EBNA3C-mediated gene activation.

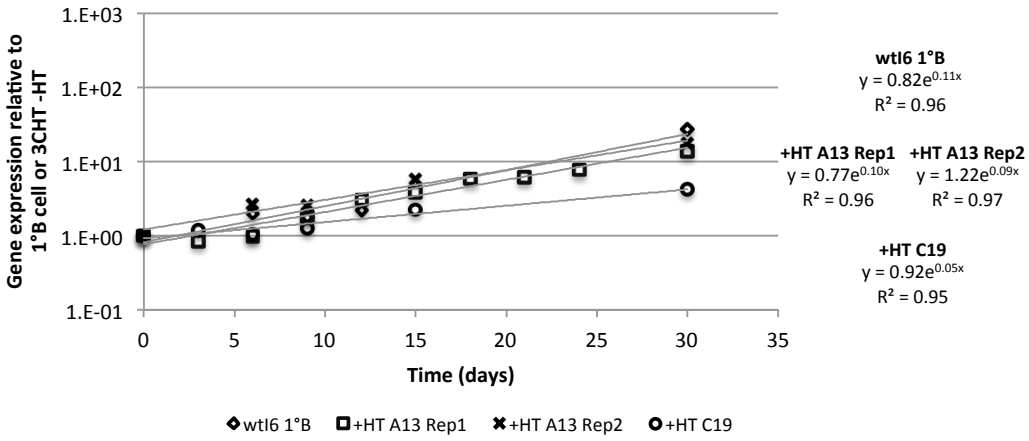


Figure 4.4: Comparison of *AICDA* induction between newly infected primary B cells and 3CHT LCL after activation of EBNA3C.
Comparison of *AICDA* expression between wild type (wtl6) EBV infected primary B cells and three replicate time-courses in 3CHT A13 and C19 after addition of HT. Gene expression was normalised to *GAPDH* and expressed relative to uninfected primary B cells for wtl6 or 3CHT –HT on day 0 for 3CHT time-courses. Exponential trendlines were fitted and equations and R^2 values are shown.

4.2.4 Molecular mechanism of EBNA3C-mediated gene activation

4.2.4.1 EBNA3C-binding at regulatory regions of AICDA

The regulation of *AICDA* in B cells is tightly controlled and regulated on multiple layers to prevent expression outside a GC environment (see Figure 1.6). From the results of the primary B cell infection, it appeared that EBNA3C did not regulate *AICDA* expression on a post-transcriptional level via miR155 (Figure 4.1C) or indirectly via the major GC transcription factor BCL6 – so it is not just inducing the cells to become GC-like (Figure 4.1B). In order to determine whether EBNA3C regulates the *AICDA* locus directly (Figure 4.5A), the ChIP-Seq on EBNA3C-TAP or wild type LCL performed by Dr Kostas Paschos was interrogated (Figure 4.5B). EBNA3C occupancy could be detected in four of the previously identified regulatory regions that control expression of *AICDA*. Importantly, all regulatory regions indisputably identified as enhancers, namely region II, region IV, region V and region VI, showed binding of EBNA3C in the order of magnitude region VI>IV>II>V (Figure 4.5B and C). The result of the ChIP-Seq could be confirmed by ChIP-qPCR using EBNA3C-TAP LCL and untagged wt LCL, showing a similar order of EBNA3C occupancy at the individual regulatory regions (Figure 4.5D). In addition to this, there might be some occupancy of EBNA3C at the putative enhancer region III, although for reasons that are unclear, the background signal appeared to be remarkably higher at this site compared to the other regulatory regions (Figure 4.5D). However, no signal has been detected at this site in the ChIP-Seq (Figure 4.5B and C).

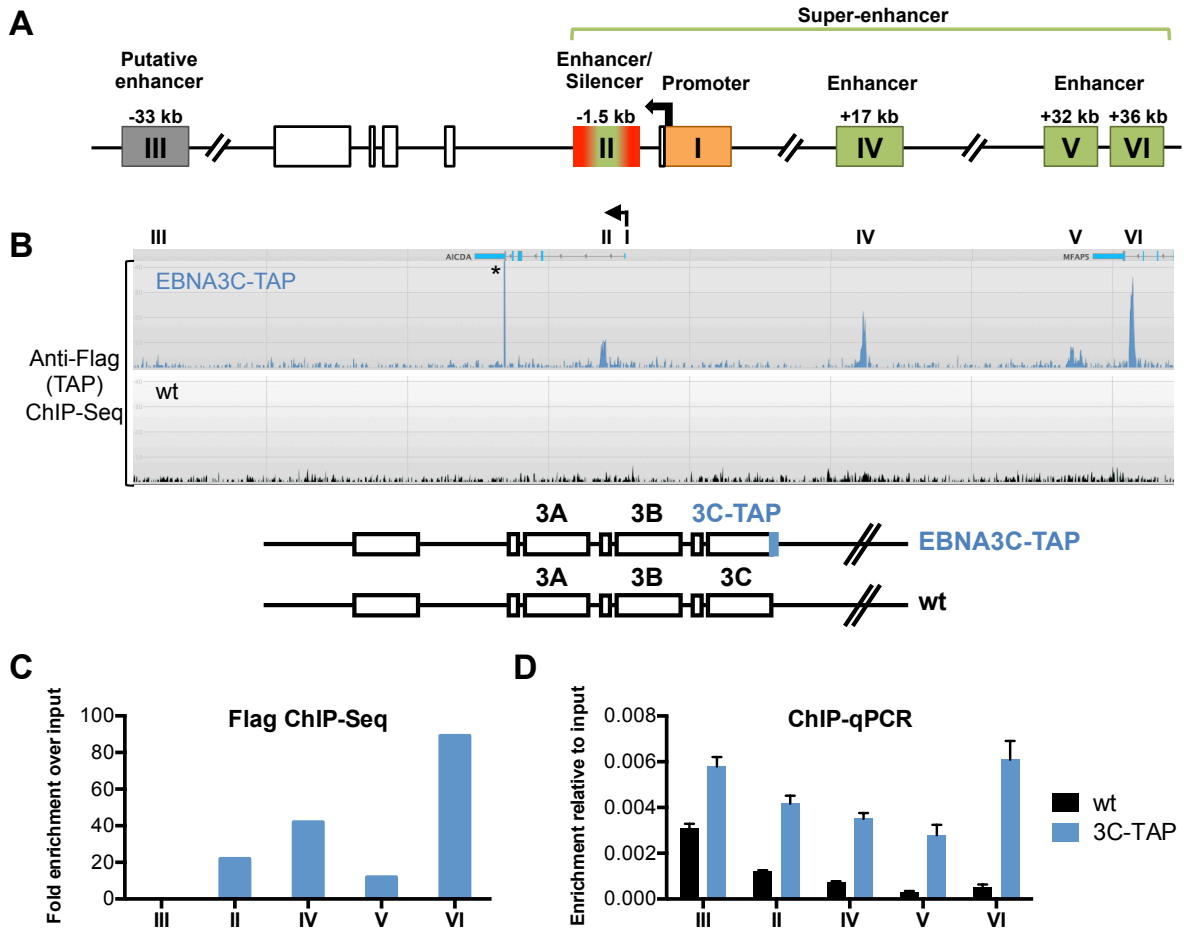


Figure 4.5: EBNA3C binding at the *AICDA* locus.

(**A**) Schematic overview of previously characterised regulatory regions that control transcriptional levels of AID expression (see Introduction). (**B**) Anti-Flag ChIP-Seq on EBNA3C-TAP (Flag and Strep) or untagged (wt) LCL displayed at the *AICDA* locus using Savant genome browser. The asterisk marks a site of identical reads also observed in the input sample (not shown), which is not identified as a genuine ChIP-Seq peak by the Model-based analysis of ChIP-Seq (MACS) algorithm. Dr Kostas Paschos performed the ChIP-Seqs. (**C**) Calculated fold enrichment of EBNA3C-TAP at indicated regulatory regions of *AICDA* as determined by ChIP-Seq. (**D**) Verification of EBNA3C occupancy by anti-Flag ChIP-qPCR in EBNA3C-TAP and wt LCL. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample. This is a representative of two independent biological replicates.

4.2.4.2 Changes of epigenetic histone marks control expression levels

In order to understand better how *AICDA* is up regulated by EBNA3C after having detected EBNA3C binding at most regulatory elements of *AICDA* and having established that 3CHT LCL can be used to recapitulate efficiently gene expression changes of *AICDA*, ChIPs for activation-associated epigenetic histone marks H3K4me3, H3K9ac and H3K27ac were performed on the 3CHT A13 time course (Figure 4.6B-D). This revealed that when expression of *AICDA* was induced by EBNA3C there were higher levels of H3K4me3 around the promoter region of *AICDA* (region I), but also at region II ~1.5 kb downstream of the TSS (Figure 4.6B). In addition, H3K9ac and H3K27ac increased at these two regions, but also at region V, one of the upstream enhancers ~32 kb upstream of the TSS (Figure 4.6C and D).

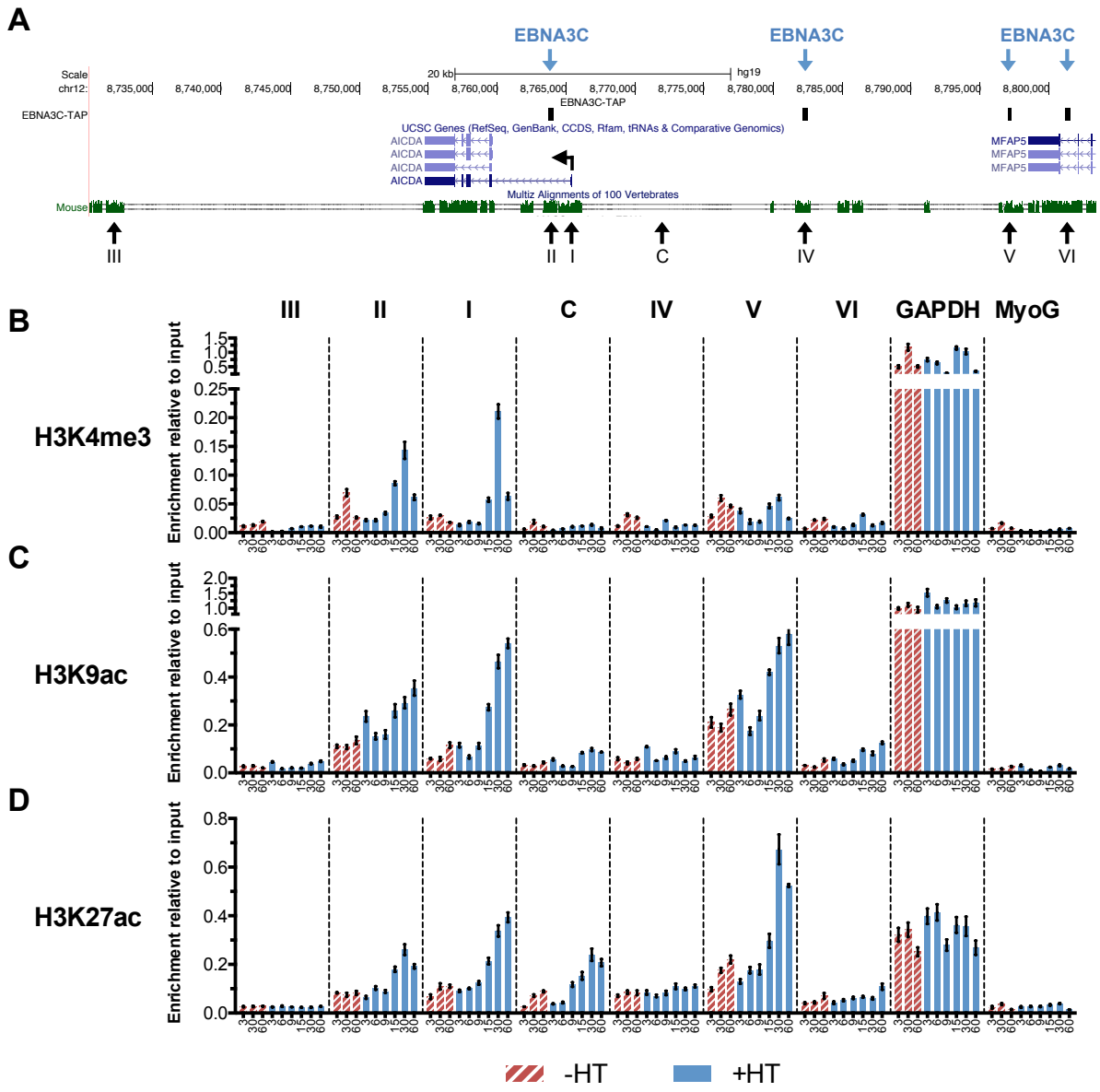


Figure 4.6: Changes in epigenetic marks across the *AICDA* locus during the 3CHT A13 time-course.

(**A**) UCSC genome browser overview of the *AICDA* genomic locus showing ChIP-Seq tracks for EBNA3C-TAP with occupancy highlighted in blue. DNA sequence conservation to murine AICDA locus is shown in green and previously identified regulatory regions I, II, III, IV, V and VI are highlighted by black vertical arrows, which also indicate locations for primers used in ChIP-qPCR (C indicates a control region). The horizontal arrow indicates the TSS of *AICDA*. (**B-D**) ChIP for H3K4me3 (**B**), H3K9ac (**C**) and H3K27ac (**D**) on samples from the 3CHT A13 time-course at locations across the *AICDA* locus, at *GAPDH* or *myoglobin* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample. This is a representative of two independent biological replicates (see Appendix Figure 7.6).

Similar results were observed in the biological replicate time-course 3CHT C19 (Appendix, Figure 7.6A-C). Therefore, normalised histone marks H3K4me3 and H3K9ac at region I from both replicate time-courses could be averaged and compared with *AICDA* mRNA expression. H3K4me3 and H3K9ac increased at the TSS from day 15 onwards (Figure 4.7A), which correlates well with *AICDA* gene expression (Figure 4.7B).

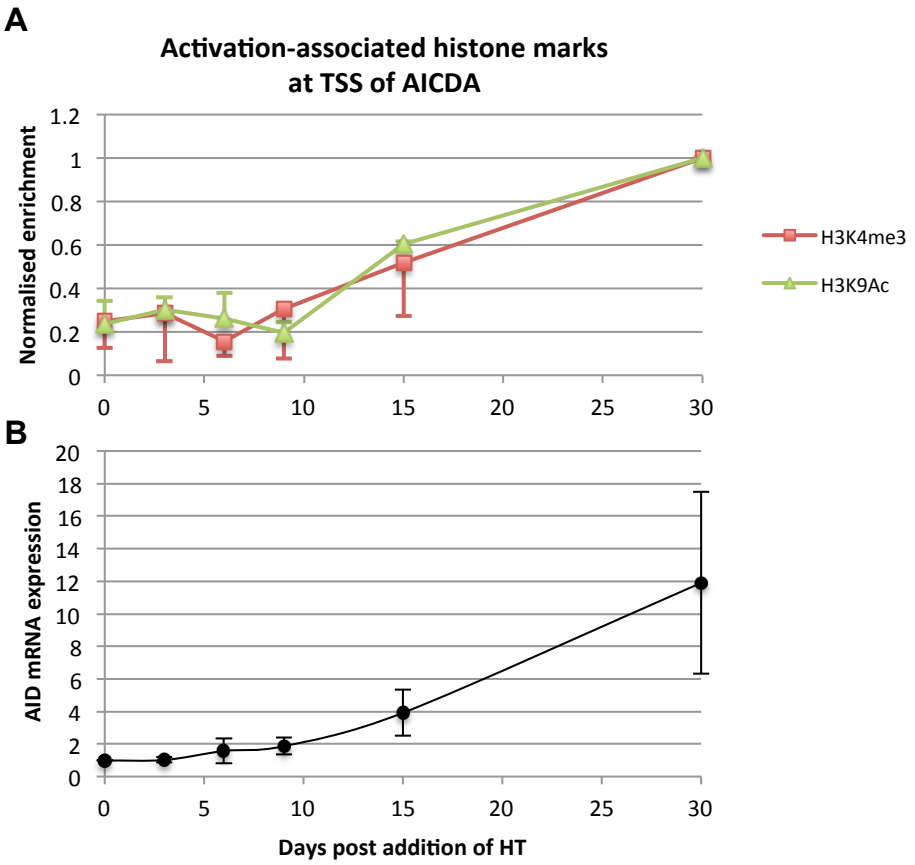


Figure 4.7: Correlation between changes of histone marks H3K4me3 and H3K9ac at the TSS of *AICDA* and gene expression.

(**A**) ChIP-qPCR values at the TSS of *AICDA* (region I) from the 3CHT A13 and 3CHT C19 time-courses were normalised by setting H3K4me3 and H3K9ac levels from the last time-point +HT (day 30) to 1 and levels from the first time-point –HT (day3) were used as equivalent for day 0. Mean values ± SD from both replicate time-courses are shown. For better visual clarity, error bars are colour-matching the corresponding histone mark and only upper bars are shown for H3K9ac and lower bars for H3K4me3. (**B**) RT-qPCR gene expression for *AICDA* from three replicate time courses (3CHT A13 rep 1+2 and 3CHT C19) was normalised to *GAPDH* and is shown as mean values ± SD.

4.2.4.3 Factor recruitment to the AICDA locus

Having seen an increase in activation-associated histone marks when *AICDA* was activated by EBNA3C, I focused on identifying recruited factors that catalysed these changes. Historically, PRC1 was linked to gene repression, but more recently PRC1 complexes have been found at actively expressed genes (see Introduction) and EBNA3C appears to induce its recruitment (see previous Chapter and General Discussion). Therefore, ChIP for BMI1 was performed on the 3CHT A13 time-course, which showed an increase in BMI1 occupancy when *AICDA* was up regulated by functional EBNA3C (Figure 4.8B). Increased BMI1 levels were seen mainly at region IV and VI, but, to a lesser extent, also at region II and region V – all of which are regions where EBNA3C was detected.

Furthermore, ChIP for histone acetyltransferease p300 – one of the few transcriptional co-activators with which EBNA3C has been reported to interact in GST-pull down experiments (Subramanian et al., 2002) – revealed that there was a transient recruitment of p300 mainly to regions IV and VI, again only when EBNA3C was functional (Figure 4.8C).

Since EBNA3C cannot bind directly to DNA, again ChIP for the DNA binding factor RBPJ was performed (Figure 4.8D). This showed that RBPJ occupancy transiently increased at all regions where EBNA3C was found and the levels of recruited RBPJ correlated well with the levels of EBNA3C detected by ChIP-Seq (region VI>IV>II>V).

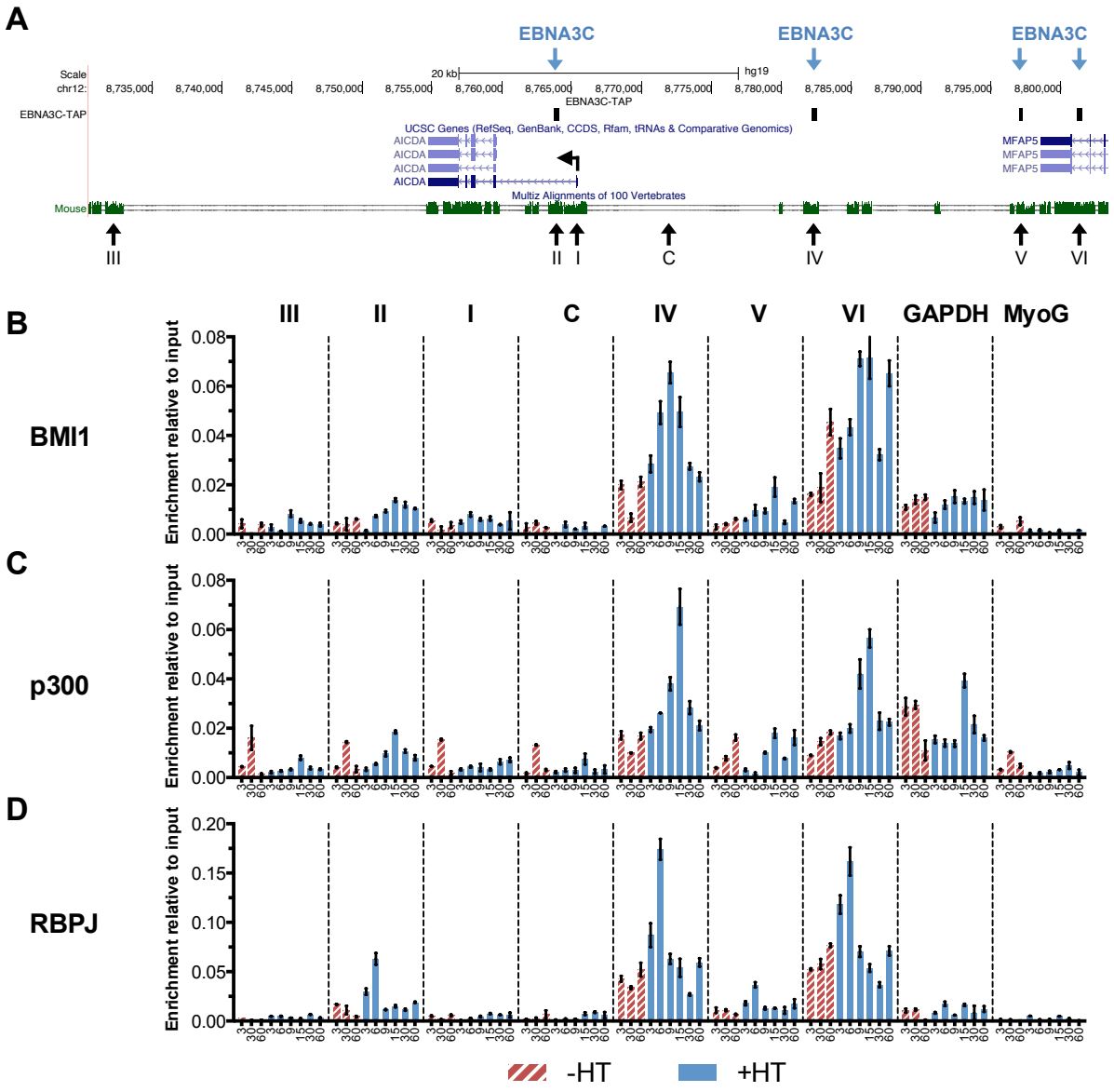


Figure 4.8: Factor binding to the *AICDA* locus during the 3CHT A13 time-course.

(**A**) UCSC genome browser overview of the *AICDA* genomic locus showing ChIP-Seq tracks for EBNA3C-TAP with occupancy highlighted in blue. DNA sequence conservation to murine AICDA locus is shown in green and previously identified regulatory regions I, II, III, IV, V and VI are highlighted by black vertical arrows, which also indicate locations for primers used in ChIP-qPCR (C indicates a control region). The horizontal arrow indicates the TSS of *AICDA*. (**B-D**) ChIP for BMI1 (**B**), p300 (**C**) and RBPJ (**D**) on samples from the 3CHT A13 time-course at locations across the *AICDA* locus, at *GAPDH* or *myoglobin* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample.

Similar ChIP results were observed in the 3CHT C19 time-course (Appendix Figure 7.6D-F), which again allowed a more detailed analysis of the recruited factors to *AICDA* regions IV and VI using the combined results of the 3CHT A13 and C19 time-courses (Figure 4.9). This analysis revealed that maximal BMI1 levels could be detected about 9 days after activation of EBNA3C, however, again these levels were not maintained and subsequently decreased to levels similar or marginally higher than in cells with inactive EBNA3C (Figure 4.9A). Acetyltransferase p300 showed maximal levels around day 15 after activation of EBNA3C, but here levels seem to remain higher compared to cells with inactive EBNA3C (Figure 4.9B). RBPJ occupancy sharply peaked 3-6 days after activation and quickly decreased to levels similar to those in cells with inactive EBNA3C (Figure 4.9C).

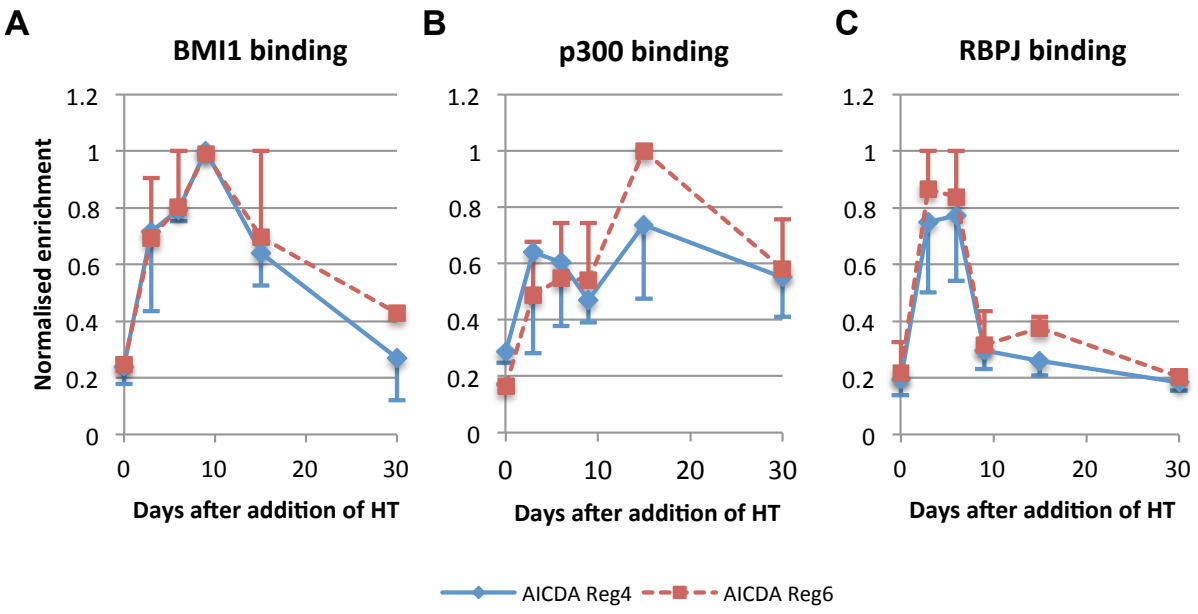


Figure 4.9: Binding of BMI1, p300 and RBPJ to *AICDA* region IV and VI after activation of EBNA3C.

ChIP-qPCR values for BMI1 (**A**), p300 (**B**) and RBPJ (**C**) at *AICDA* region IV and VI from the two biological time-courses (3CHT A13 and C19) were normalised by setting the maximal enrichment level during each time-course to one and by using the first time-point –HT (day 3) as equivalent for day 0. Mean values ± SD from both replicate time-courses are shown. For better visual clarity, error bars are colour-matched and only lower bars are displayed for region IV and upper bars for region IV. For BMI1, day nine of the C19 time-course was treated as an outlier and was not taken into account for the analysis.

4.2.5 Induction of *AICDA* is dependent on interaction between RBPJ and EBNA3C

In order to determine whether the interaction between EBNA3C and RBPJ is essential for the EBNA3C-mediated up regulation of *AICDA*, Adam Gillman analysed primary B cells infected with the RBPJ BM EBNA3C virus (see Figure 1.10B, Figure 1.17 and previous Chapter) for expression of *AICDA* (Figure 4.10). This revealed that EBNA3C failed to induce expression of *AICDA* when it cannot bind to RBPJ, which was similar to infection with EBNA3C KO virus. However, all recombinant viruses where EBNA3C could interact with RBPJ (wtA5 and 3CRev) showed induction of *AICDA* comparable to the previous primary B cell infections (Figure 4.1).

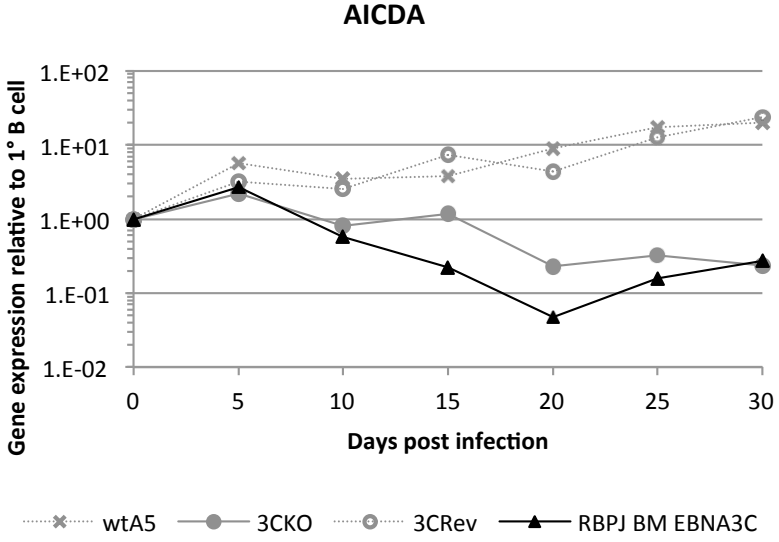


Figure 4.10: Ability of EBNA3C to bind to RBPJ is essential for the activation of *AICDA*.
Infection of purified primary B cells with wild type (wtA5), EBNA3C knock-out (3CKO), EBNA3C revertant (3CRev) or RBPJ binding mutant (BM) EBNA3C recombinant EBV. Over a period of 30 days post infection, RT-qPCR gene expression for *AICDA* was normalised to *GNB2L1* and shown relative to uninfected primary B cells. This is a single experiment that was performed and analysed by Adam Gillman.

4.2.6 EBNA2 represses and LMP1 induces expression of AID

The regulation of *AICDA* by EBV is very complex and appears to involve factors other than EBNA3C. Previously, it has been reported that LMP1 and to a lesser extent LMP2A can also induce expression of *AICDA* (He et al., 2003), whereas EBNA2 has been reported to repress expression of *AICDA* (Tobollik et al., 2006).

In order to confirm that AID is regulated by EBNA2, a time-course experiment in EBNA2-conditional EREB 2-5 cells was performed based on the previous study from Tobollik and colleagues. EREB 2-5 cells are infected with EBV strain P3HR1 that is truncated for EBNALP and deleted for EBNA2, but contain a stable expression plasmid encoding an EBNA2-estrogen receptor fusion protein that renders EBNA2-conditional to the presence of β-estradiol in the culture medium (Kempkes et al., 1995). Withdrawal of β-estradiol from the medium inactivates EBNA2 and cell proliferation ceases, because EBNA2 is essential to maintain B cell proliferation. A time-course experiment in these cells confirmed that inactivation of EBNA2 led to an increase in AID mRNA and protein expression (Figure 4.11A and B). This increase in AID is not caused by LMP1, because LMP1 protein levels quickly decrease in the absence of EBNA2. Re-activation of EBNA2 resulted in a very rapid repression of AID at the mRNA level and remarkably slower at the protein level. A minor re-induction was observed at 8 and 24 hours, which was potentially caused by the EBNA2-induced transiently high levels of LMP1 (Figure 4.11A and B). These results are very similar to reports from Tobollik and colleagues with the exception that the LMP1-mediated induction of AID, probably through increased NFκB signalling, was less pronounced. This confirms that EBNA2 is a potent repressor of AID expression, whereas LMP1 seems to be only a weak activator of AID expression.

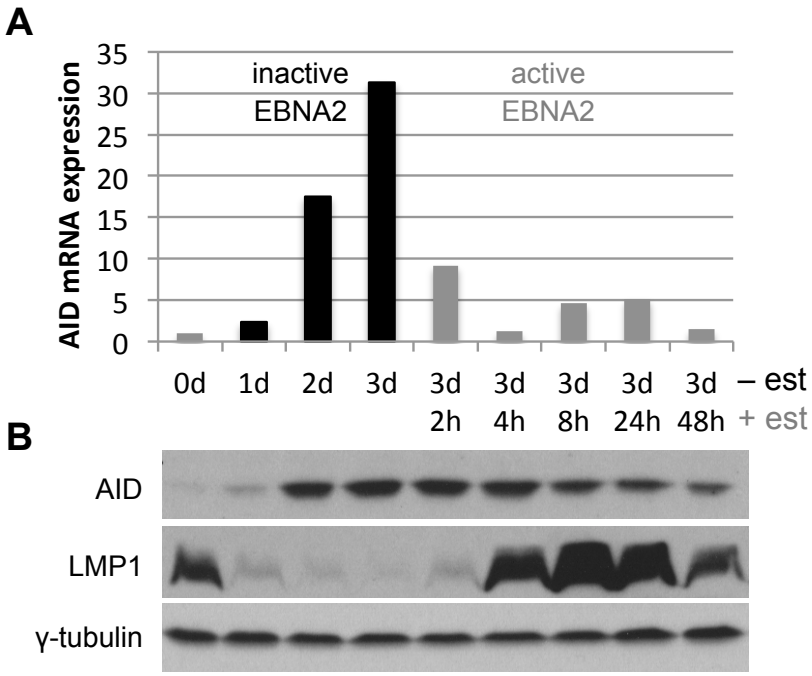


Figure 4.11: Repression of AID by EBNA2 in EREB 2-5 cells.

(**A**) EBNA2-conditional time-course using EREB 2-5 cells. Before the start of the time-course cells were always cultured in the presence of β-estradiol (est), which was removed on day 0 and cells were cultured for 3 days in the absence of β-estradiol (-est) before re-addition of β-estradiol (+est). RT-qPCR gene expression of *AICDA* was normalised to *GAPDH* and is shown relative to day 0. (**B**) Western blot analysis for AID, LMP1 and γ-tubulin in the same samples as in A.

This prompted me to analyse protein levels of EBNA2 and LMP1 in samples of the two 3CHT time-courses. Levels of EBNA3 proteins did not consistently change during either time-course, however a decrease in EBNA2 and LMP1 protein levels was observed in samples with functional EBNA3C (+HT) during the A13 (Figure 4.12A) and the C19 (Figure 7.2) time-courses, which was more pronounced in the A13 than in the C19 time-course.

This raised the question whether the induction of AID was directly caused by EBNA3C or merely a consequence reduced EBNA2-mediated repression of AID. The very rapid repression of AID by EBNA2 within 4 hours after re-activation of EBNA2 suggested a direct regulation and analysis of ChIP-Seq data for EBNA2 in Mutu III cells (McClellan et al., 2013) revealed EBNA2 occupancy at *AICDA* regulatory regions III, IV, V and VI (Figure 4.12B). EBNA3C is found at the same upstream enhancers (IV, V and VI) and it is possible that both compete for access to these sites probably through competitive binding to RBPJ,

which was detected during the 3CHT time-courses (Figure 4.8D and Figure 7.6F) and also in ChIP-Seq analysis in IB4 LCL (Zhao et al., 2011a).

In order to determine whether the induction of *AICDA* by EBNA3C was merely a consequence of lower global EBNA2 levels leading to reduced local EBNA2 occupancy at the regulatory sites of AID, EBNA2 ChIP was performed on samples of the 3CHT A13 and C19 time-courses. This revealed very marginal reduction or no changes to EBNA2 levels at the regulatory regions of *AICDA* in samples with functional EBNA3C compared to samples with inactive EBNA3C (Figure 4.12D and Figure 7.6G). EBNA2 binding at the promoter region of *HES1* (Sakai et al., 1998), a well-characterised EBNA2-regulated gene, was used as positive control, which appeared to be quite variable between samples, but no consistent change in samples with functional EBNA3C compared to samples with inactive EBNA3C could be detected. Therefore, it does not appear that high levels of EBNA2 are actively repressing expression of AID in the absence of functional EBNA3C, which is consistent with a basal rate of *AICDA* transcription comparable to primary B cells.

The fact that EBNA2 represses and EBNA3C activates the expression of AID is surprising, because usually EBNA2 recruits co-activators, whereas EBNA3C recruits co-repressors (see Introduction). This will be discussed in the Discussion section of this and the next Chapter.

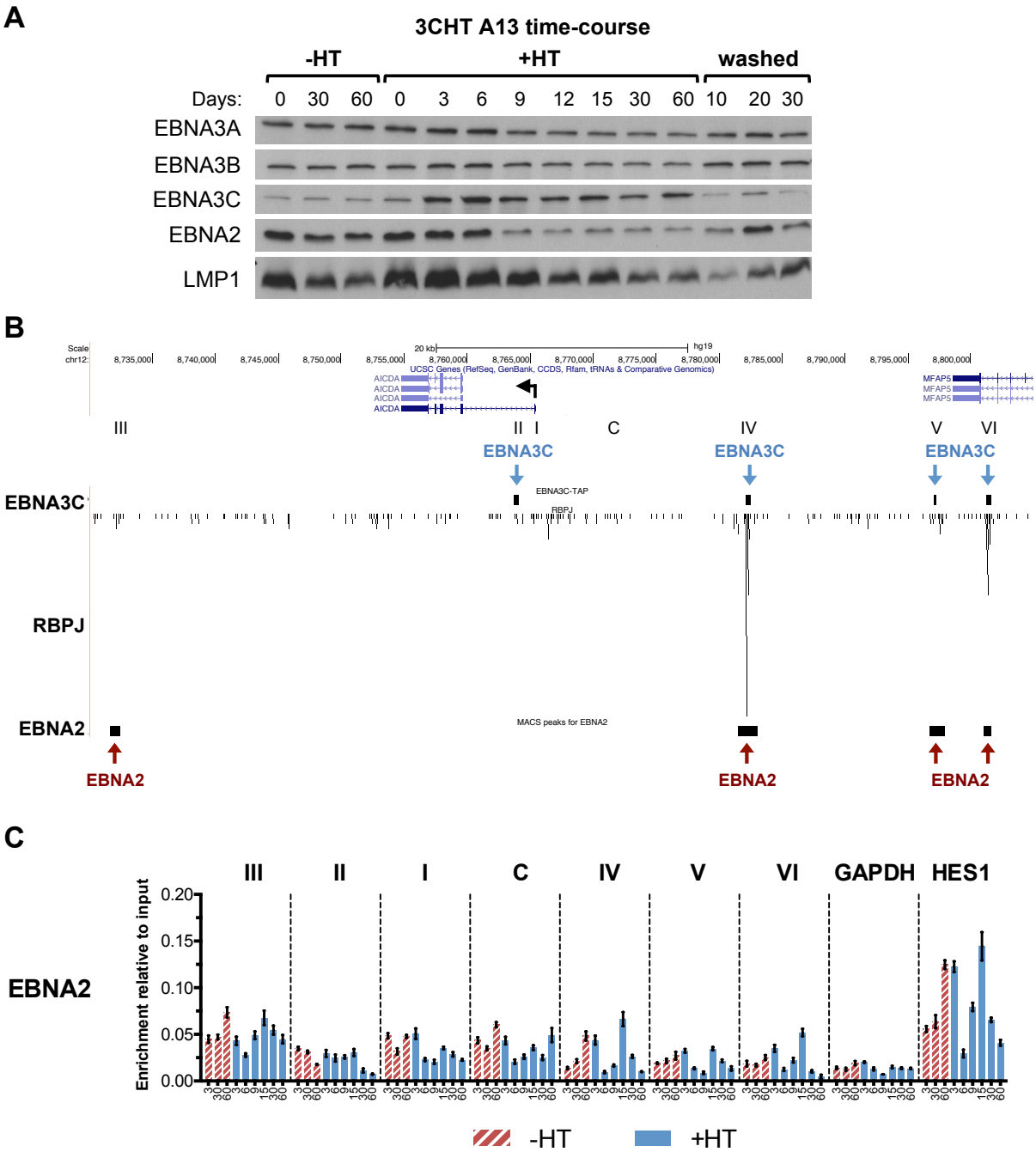


Figure 4.12: Global EBNA2, EBNA3 and LMP1 protein levels and occupancy of EBNA2 at the *AICDA* locus during the 3CHT A13 time-course.

(**A**) Western blot analysis on samples of the 3CHT A13 time-course for EBNA3A, EBNA3B, EBNA3C, EBNA2, LMP1 and γ-tubulin. (**B**) UCSC genome browser overview of *AICDA* locus as in Figure 4.6A, but also showing ChIP-Seq data for RBPJ in IB4 LCL (Zhao et al., 2011a) and EBNA2 in Mutu III (McClellan et al., 2013). EBNA2 binding sites at *AICDA* regulatory regions are highlighted in red. (**C**) ChIP for EBNA2 on samples from 3CHT A13 time-course at locations across the *AICDA* locus, at *GAPDH* or *HES1* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample.

4.2.7 Transient reporter assays in an attempt to recapitulate regulation of *AICDA* by EBNA3C and EBNA2

In order to determine whether the regulation of *AICDA* by EBNA3C and EBNA2 can be recapitulated in transient reporter assays, all the regulatory elements of *AICDA* that were identified at that time, namely region I, II, III and IV [region V and VI were only reported later in (Kieffer-Kwon et al., 2013)], were PCR amplified from genomic DNA extracted from LCL GM12878 introducing restriction enzyme recognition sequences for subsequent cloning into the pGL3-basic vector (see Material and Methods and Figure 4.13A).

Electroporation of these luciferase vectors into DG75 cells revealed a basal promoter activity of region I, a subtle repressive capacity of region II and enhancer activities of region III and IV (Figure 4.13B). To my knowledge, this is the first time that human regulatory regions of *AICDA* were used in transient reporter assays. However, these results are similar to reports from transient luciferase assays using the corresponding murine regulatory sequences of mouse *Aicda*. Mouse region I was reported to increase luciferase activity between 5-15 fold above pGL3-basic vector (Yadav et al., 2006; Park et al., 2009; Tran et al., 2010b), which was slightly repressed by the presence of region II (Park et al., 2009; Tran et al., 2010b) or increased by the presence of region IV (Tran et al., 2010b). In contrast, mouse region III did not affect luciferase activity in these transient assays (Tran et al., 2010b), but was shown to be necessary for *Aicda* expression in a BAC transgenic mouse model (Crouch et al., 2007). The results from human region III suggest a positive effect on luciferase activity (Figure 4.13B).

Having established that human *AICDA* regulatory regions have a similar effect on luciferase activity as those from mouse *Aicda*, my next aim was to determine their role in the regulation by EBNA3C and EBNA2. Therefore, increasing amounts of EBNA3C or EBNA2 expression plasmids were transfected together with luciferase vectors containing region I and in addition either region II (Figure 4.13C), region III (Figure 4.13D) or region IV (Figure 4.13E). Unfortunately, no consistent effect on luciferase activity could be observed for region

II and IV. For region III, large amounts of EBNA3C seemed to have a repressive effect, which was unexpected and very likely due to levels of expression outside the physiologic range (5,000-10,000 expression plasmids per cell in the transfection). From these results it also seems that EBNA2 might activate region III, which again contradicts the expectation and the technical variation between individual experiments was rather large (Figure 4.13D). For reasons that remain unclear, it was not possible to recapitulate regulation of *AICDA* by EBNA3C or EBNA2 in transient reporter assays using sub-genomic fragments. This strategy was therefore not pursued further.

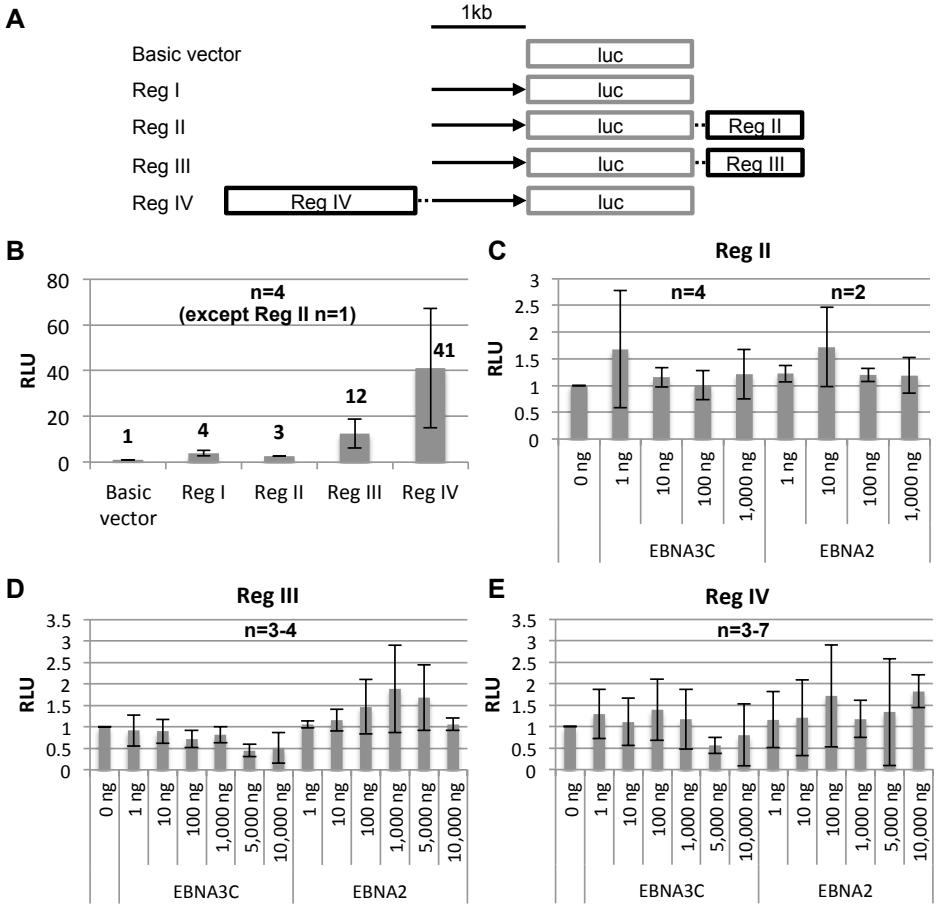


Figure 4.13: Attempt to recapitulate regulation of *AICDA* in transient reporter assays.
(**A**) Overview of luciferase vectors. The 1 kb promoter region of *AICDA* (region I, black arrow) and 1 kb around the *AICDA* regulatory regions II and III or 2 kb for region IV (black boxes) were cloned upstream or downstream of the luciferase gene (grey box). (**B**) Transient luciferase reporter assays in DG75 cells after electroporation of 1 µg of reporter vector as indicated and expressed relative to the promoter-less basic vector. (**C-E**) Transient luciferase reporter assays in DG75 after electroporation of 1 µg luciferase vector – Reg II (**C**), Reg III (**D**) or Reg IV (**E**) – and increasing amounts of EBNA3C or EBNA2 expression plasmids as indicated. Luciferase units were normalised to β-galactosidase units of the same transfection and are shown as mean relative luciferase units (RLU) ± SD.

4.3 Summary

- EBNA3C is essential for the up regulation of *AICDA* in newly EBV-infected primary B cells.
- EBNA3C-conditional LCL can be used to reliably recapitulate EBNA3C-mediated induction of *AICDA* seen in newly EBV-infected primary B cells.
- EBNA3C directly targets all regulatory regions of *AICDA* that have been previously identified as enhancers.
- Activation-associated epigenetic marks are deposited across the *AICDA* locus when EBNA3C is functional probably through recruitment of acetyltransferase p300. The role of recruited BMI1 is unclear, but raises interesting questions concerning PRC1 in gene activation.
- Ability of EBNA3C to bind to RBPJ is absolutely essential for the induction of *AICDA*.
- Other EBV proteins are involved in the regulation of *AICDA*:
  - EBNA2 represses AID, possibly by targeting most of the regulatory regions of *AICDA* that are also targeted by EBNA3C.
  - LMP1 up regulates AID, possibly through NFκB signalling.
- Regulation by EBNA3C or EBNA2 could not be recapitulated in transient reporter assays.
- Most of the results presented in this Chapter have been submitted to JEM for publication and are currently under peer review.

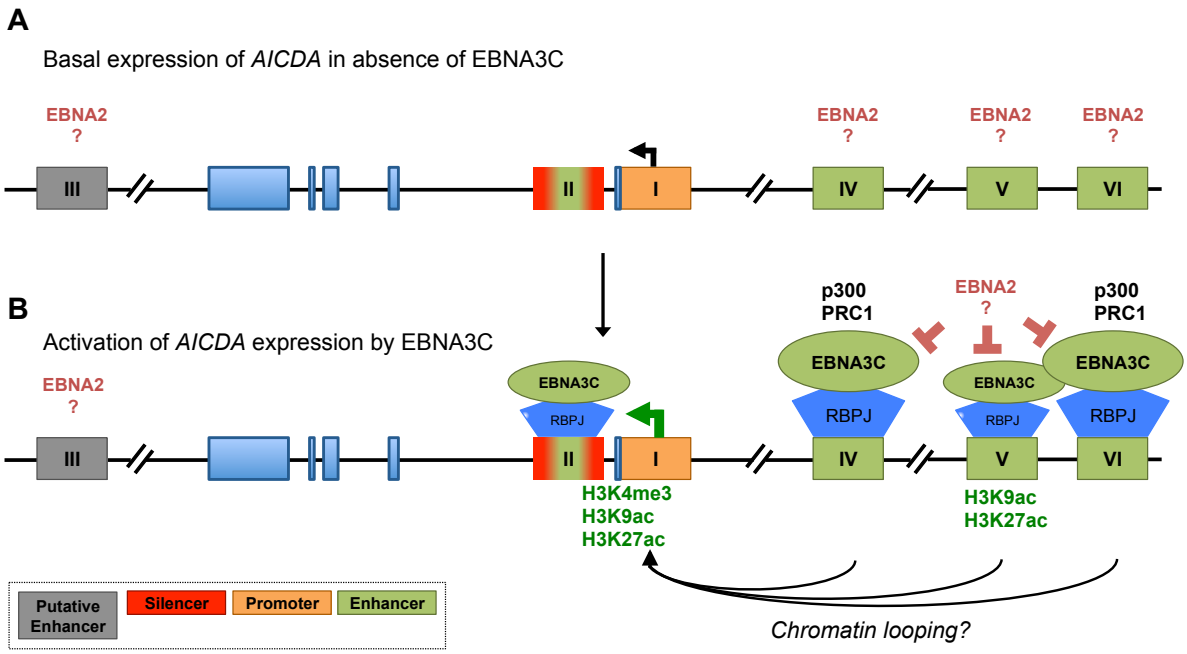


Figure 4.14: Model of gene activation by EBNA3C at AID genomic locus.

(**A**) In the absence of functional EBNA3C, *AICDA* is expressed at basal levels comparable to primary B cells. EBNA2 has been detected to bind to regulatory regions III, IV, V and VI and has the potential to negatively affect *AICDA* transcription. (**B**) Presence of functional EBNA3C, leads to the recruitment/stabilisation of what are probably RBPJ:EBNA3C complexes at regulatory regions II, IV, V and VI (occupancy levels in the following order: VI>IV>II>V). Acetyltransferase p300 and the PRC1 subunit BMI1 are recruited to region IV and VI and an increase in activation-associated histone marks can be detected at region I and region V. Chromatin looping between the upstream enhancers and the promoter probably supports high transcription levels of AID.

4.4 Discussion

The discussion of this Chapter will only focus on issues specific for EBNA3C-mediated gene activation. A broader discussion of issues regarding the role of EBNA3C in gene regulation can be found in Chapter 5.

At the onset of this study, compared with EBNA3C-mediated gene repression, near to nothing was known about the role of EBNA3C in gene activation. Here, focusing on the gene most robustly induced by EBNA3C as revealed in an exon microarray screen, some insights could be obtained. Infection of primary B cells with recombinant EBV showed clearly that EBNA3C is essential for the induction of AID expression. AID activity is controlled on multiple levels (see Figure 1.6), but it appeared that EBNA3C only acts at the transcriptional level. It specifically targets most of the DNA regulatory elements that have been previously identified to control expression of AID, most importantly all previously identified enhancer elements that together form the AID super-enhancer are targeted by EBNA3C. Activation of EBNA3C function resulted in the deposition of activation-associated histone modifications mainly around the TSS of *AICDA*. The levels of these histone marks correlated well with AID mRNA expression.

4.4.1 Dynamic recruitment of RBPJ during EBNA3C-mediated induction of *AICDA*

Again, most surprising was the dynamic recruitment and/or stabilisation of what are probably RBPJ/EBNA3C complexes at the EBNA3-binding sites in the enhancer regions of AID after activation of conditional EBNA3C by the activating ligand HT. The observation that EBNA3C activates and EBNA2 represses the AID locus, possibly through competitive binding to RBPJ, is also very surprising (Figure 4.14). This turns the traditional models of EBNA2- and EBNA3C-mediated gene regulation on their head (Figure 1.16C and Figure 1.16D), because typically EBNA2 is considered as potent transactivator and EBNA3C as strong repressor when anchored to DNA. Although EBNA2 and EBNA3C have been linked to gene repression and activation, respectively (see Introduction), to my knowledge

repression by EBNA2 and activation by EBNA3C have not been previously reported for the same gene. Dynamic recruitment of RBPJ in gene activation has been reported during Notch signalling (in both *Drosophila* and mammals) and also in response to EBNA2 (see Section 3.4.1), but to my knowledge this is the first time it has been observed for EBNA3C-mediated gene activation.

Previous studies found that the EBNA3C-mediated gene activation of the viral LMP1 promoter (in transient reporter assays) was independent of RBPJ, but required PU.1/SPI1 binding sites (Zhao and Sample, 2000; Lin et al., 2002). Here, however, binding of EBNA3C to RBPJ appears to be essential for the activation of *AICDA*, as determined by infection of primary B cells with the RBPJ BM EBNA3C virus (Figure 4.10). It is possible that EBNA2 inhibits activation of AID by competing with EBNA3C for binding to RBPJ, however, mechanisms independent of RBPJ might also be involved (see Section 4.4.3).

A more extended discussion on the role of RBPJ in EBNA3C-mediated gene regulation can be found in Chapter 5.

4.4.2 Role of BMI1 (PRC1) and p300 in EBNA3C-mediated gene activation

Another surprise was the transient recruitment of BMI1 to the regulatory sites of *AICDA* with levels correlating with EBNA3C-occupancy (VI>IV>II>V), because BMI1 as part of PRC1 was traditionally associated with gene repression and silencing. However, both PRC1 and PRC2 complexes have been recently associated with gene activation (see Section 1.3.3.2.1). No recruitment of the PRC2-family member SUZ12 could be detected during the 3CHT A13 and C19 time-courses at sites across the *AICDA* locus (Figure 7.7). It is unclear what role BMI1 plays at the *AICDA* locus, but it has been reported at actively expressed genes (Frangini et al., 2013).

Frangini and colleagues found that PRC1 (Ring1B, Bmi1 and Cbx7) and Aurora B could be co-immunoprecipitated and co-occupied many active promoters in resting murine B and T cells. Knock-out of either Ring1B or Aurora kinase B resulted in reduced transcription

of these genes and decreased cell viability. Aurora kinase B was found to phosphorylate H3S28, inhibit PRC1-catalysed ubiquitination of H2AK119 and recruit deubiquitinase USP16 at these promoters (Frangini et al., 2013). Interestingly, EBNA3C has been shown to directly interact with Aurora kinase B by co-immunoprecipitation and GST pull downs, which mapped the interaction to the N-terminal residues 90-160 of EBNA3C (Jha et al., 2013). As PRC1 proteins and EBNA3C can independently bind to Aurora kinase B, it is therefore possible that an activating EBNA3C-PRC1-Aurora kinase B complex is recruited to the *AICDA* locus. ChIP for Aurora kinase B has been attempted, but not pursued, because Aurora kinase B is globally recruited to condensed chromatin during mitosis, which would make data interpretation very difficult in LCL that continuously proliferate. ChIP to assess phosphorylation levels of H3S28 across the 3CHT A13 and C19 time course was not performed because activated B cells display this mark only on mitotic chromosomes (Sabbattini et al., 2014).

Some PRC1 complexes have been found to interact with autism susceptibility gene 2 (AUTS2) protein that activated gene expression through recruitment of p300 to promoter regions controlling neural gene expression during the development of the mouse central nervous system (Gao et al., 2014). However, it is important to mention that the AUTS2 containing PRC1 complexes were limited to PRC1.3 and PRC1.5, but have not been described for canonical or non-canonical PRC1.4 that contain BMI1 (see Figure 1.14). Interrogation of the 3CHT exon-microarray indicated that AUTS2 is expressed in LCL and it is possible that other PRC1 complexes, which can interact with AUTS2, helped with the recruitment of p300, but this has not been investigated.

Amounts of the acetyltransferase p300 predominantly increased at *AICDA* regulatory regions IV and VI, the two sites with highest levels of EBNA3C. However, no direct increase in H3K9 and H3K27 acetylation was detected at these two sites, but at regulatory regions I, II and V, which are sites up to 20kb away from where p300 was detected. Analysis of publically available HiC data from LCL GM12878 cells [Figure 4.15A, (Rao et al., 2014)],

which highly express AID mRNA and protein and constitutively express EBNA3C (Figure 4.2 C and D), suggested that all upstream *AICDA* regulatory regions (IV, V and VI) are in contact with the *AICDA* promoter region (Figure 4.15B). Similarly, in cytokine activated murine B cells that express high levels of AID, all upstream *Aicda* regulatory regions (IV, V and VI) were found to be associated in long-range interactions with the promoter region determined by chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) of polymerase II (Kieffer-Kwon et al., 2013). Furthermore, it has been suggested that PRC1 complexes together with CTCF are involved in the formation of higher order chromatin structures that are essential for gene activation (Schaaf et al., 2013; Kondo et al., 2014).

Therefore, it is likely that PRC1 proteins are recruited in order to either directly help in the recruitment of transactivators like p300 and Aurora kinase B and/or to provide support in the formation of the higher order chromatin structure that facilitates EBNA3C-mediated recruited p300 to acetylate regulatory regions that are located up to 20kb away from its main binding sites (Figure 4.14).

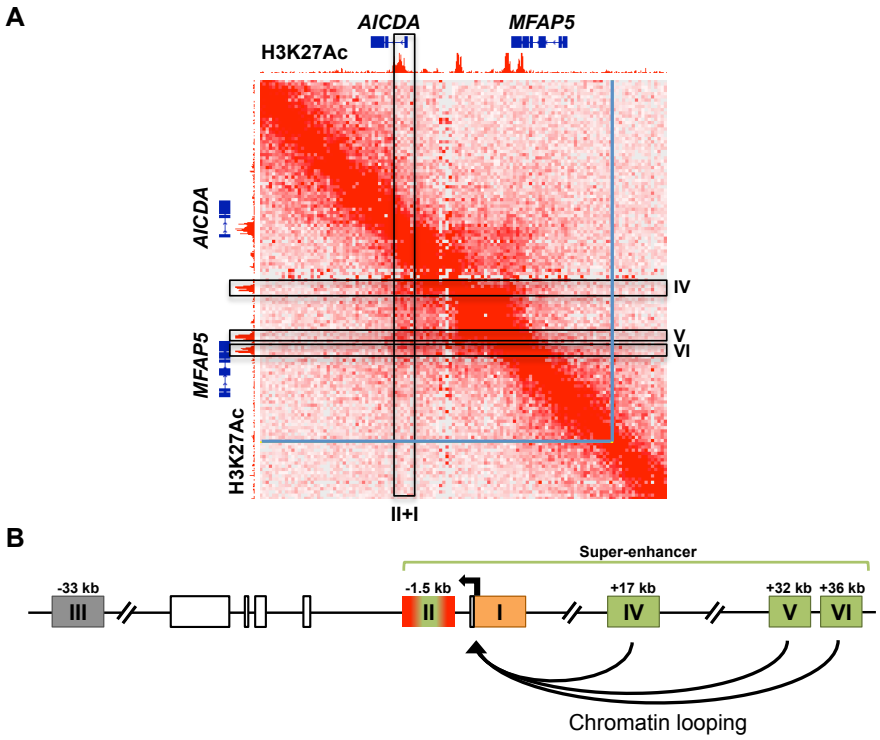


Figure 4.15: Higher order chromatin structure of the *AICDA* locus in LCL GM12878.
(**A**) Hi-C map GM12878 in situ MboI combined (4.9B) at 1kb resolution was displayed using Juicebox [http://www.aidenlab.org/juicebox/, (Rao et al., 2014)]. The observed number (balance normalised) of

contacts between a pair of loci is represented by the intensity of each pixel. *AICDA* and *MFAP5* gene locations, H3K27ac in LCL GM12878, a topological associating domain (blue box) and locations of AID regulatory regions I, II, IV, V and VI are indicated. (**B**) Schematic overview of higher order chromatin structure of the *AICDA* locus in LCL GM12878 with chromatin looping indicated by black arrows between the promoter region of *AICDA* and upstream enhancer elements IV, V and VI.

The importance of the correct higher order chromatin structure is also one possible explanation why EBNA3C-mediated activation and EBNA2-mediated repression of *AICDA* could not be recapitulated in transient luciferase reporter assays, when regulatory regions II, III and IV were taken out of their context and cloned individually into pGL3-basic vector containing the *AICDA* basal promoter (region I). To circumvent this problem, luciferase reporter BACs could be constructed that drive the expression of a luciferase gene cloned in place of the first intron of *AICDA* and otherwise contain either the entire *AICDA* locus or are deleted for individual regulatory regions.

4.4.3 Regulation of *AICDA* by EBNA2 and LMP1

The results of the EBNA2-conditional time-course using EREB 2-5 cells clearly confirmed EBNA2 as a negative regulator of AID expression (Tobollik et al., 2006) and LMP1 as a moderate inducer of AID expression (He et al., 2003). It is surprising, that in the regulation of AID, EBNA2, that is traditionally seen as a strong transactivator, is acting as a very potent repressor. In contrast, EBNA3C, that is more readily associated with repression, is essential for gene activation. It is important to mention that AID expression can be induced by β-estradiol via binding of estrogen receptors to the AID promoter (Pauklin et al., 2009). However this is not a concern for this study, because addition of β-estradiol to EBNA2-conditional EREB2-5 cells results in a potent repression of AID expression and was reported not to affect AID expression in non-conditional EBNA2 cell lines (Tobollik et al., 2006). Furthermore, addition of HT, an estrogen antagonist, to 3CHT LCL results in the induction of AID and addition to non-conditional 3CKO LCL did not activate AID expression (Figure 4.3).

The induction of AID expression in EREB 2-5 cells after inactivation of EBNA2 is probably a cumulative effect of lost repression by inactivated EBNA2 and maintained

activation by EBNA3C. Expression of EBNA3C is dependent on active EBNA2, but due to the relative long half-life of EBNA3C protein (Touitou et al., 2005), active EBNA3C is probably still present after EBNA2 has been inactivated. Unfortunately, EREB 2-5 cells express type II EBNA3C that cannot be detected by the anti-EBNA3C A10 antibody in western blots (Maunders et al., 1994). Interrogation of EBNA2 ChIP-Seq indicated that EBNA2 and EBNA3C seem to target mostly the same regulatory elements. Analysis of ChIP-Seq for RBPJ in IB4 LCL (Zhao et al., 2011a) revealed that RBPJ occupancy is highest at regulatory regions IV, VI and V, which is similar to the RBPJ ChIP results from the 3CHT A13 and C19 time-course. Therefore, it seemed likely that EBNA2 and EBNA3C compete for binding to RBPJ to obtain access to the upstream regulatory regions of *AICDA* (Figure 4.14). However, ChIP for EBNA2 on samples of the 3CHT A13 and C19 time-courses did not reveal changes in EBNA2 occupancy at these sites that would support a competitive binding model of EBNA2 and EBNA3C to these sites. Furthermore, repression of IgM and BCL6 by EBNA2 has been reported to be only partially dependent on RBPJ, as repression still occurs after activation of EBNA2 in RBPJ-null DG75 (Maier et al., 2005; Boccellato et al., 2007). Therefore, beside competing with EBNA3C for binding to RBPJ, other mechanisms may be involved in EBNA2-mediated gene repression of AID. However, the up regulation of AID in newly EBV-infected B cells indicates that EBNA3C-mediated induction dominates over EBNA2-mediated repression of AID.

The effect of LMP1 on the expression of AID was rather moderate and it did not seem to play a role in the initial up regulation of AID after inactivation of EBNA2, because LMP1 protein levels decreased rapidly and were barely detectable while AID levels still increased. Even the high levels of LMP1 after re-activation of EBNA2 only resulted in a relative moderate transient increase in AID mRNA levels, which was however in contrast to the more pronounced effect of LMP1 at these time-points observed by Tobollik and colleagues. LMP1 was shown to up regulate AID expression via NFκB and EGR-1 signalling (Kim et al., 2013). NFκB binding sites have been reported within *AICDA* regulatory regions I (Dedeoglu et al.,

2004; Gourzi et al., 2007; Park et al., 2009; Tran et al., 2010b), region II (Yadav et al., 2006; Tran et al., 2010b) and region IV (Tran et al., 2010b). Interrogation of ChIP-Seq for NFκB subunits RelA, RelB, cRel, p50 and p52 in GM12878 (Zhao et al., 2014) revealed occupancy of all five NFκB subunits predominantly at region IV, but also at region II, V and VI (Figure 4.16). In order to investigate whether changes in NFκB signalling during the 3CHT A13 and C19 time-courses were involved in the up regulation of AID expression, ChIP for RelA was attempted. This produced very low enrichment levels at all *AICDA* regulatory sites, however the ChIP did not seem to work, because no enrichment of RelA could be detected at the promoter region of *EBI3*, a LMP1/NFκB-regulated gene chosen as a positive control (data not shown). Furthermore, no increase, but rather a decrease in LMP1 expression was observed during the 3CHT A13 and C19 time-courses (Figure 3.12 and Figure 7.2). Assuming unaltered LMP1/NFκB signalling in cells with active and inactive EBNA3C, this suggests that LMP1 does not play a major role in the induction of AID. However, in order to determine whether LMP1 has to be present for the EBNA3C-mediated induction of AID and to assess the extent of LMP1 in the regulation, primary B cells would have to be infected with LMP1 KO virus. This however is complicated by the fact that LMP1 is important for B cell transformation by EBV. LMP1 KO LCL can be established, but only in the presence of feeder cells and with a low efficiency (Dirmeier et al., 2003).

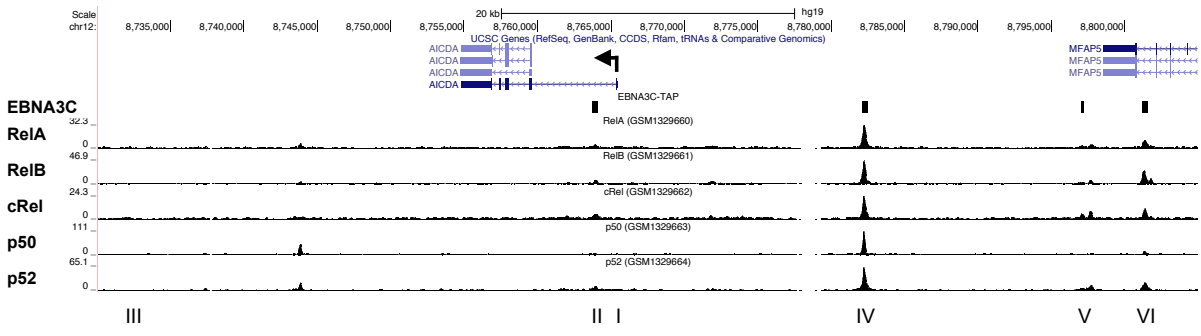


Figure 4.16: NFκB binding at *AICDA* locus in GM12878.

UCSC genome browser overview of AICDA genomic locus showing ChIP-Seq data for NFκB subunits RelA, RelB, cRel, p50 and p52 in GM12878 (Zhao et al., 2014). Positions of the AICDA regulatory regions (I-VI) are indicated below the ChIP-Seq tracks.

In summary, the precise interplay between EBNA3C and EBNA2 in the regulation of *AICDA* remains to be determined. The results presented here suggest that both nuclear proteins are directly involved in the regulation and might compete with each other for binding to RBPJ. Infection of primary B cells with EBV and the 3CHT conditional system showed that the presence of functional EBNA3C is absolutely essential for the activation of AID and further depends on the ability of EBNA3C to bind to RBPJ, indicating that EBNA3C-mediated induction dominates over EBNA2-mediated repression of AID. LMP1, probably through NFκB signalling, seems to play only a minor role in the induction of AID.

4.4.4 Why does EBV activate AID expression and what at are the consequences for the EBV-infected B cells?

In addition to providing insight into the molecular mechanisms by which EBV deregulates AID expression, this study raises very interesting questions about why EBV activates AID expression in newly infected B cells, what consequences this has for EBV-infected B cells and why it might favour EBV.

Expression of AID at high levels is generally limited to GC B cells (see Section 1.1.5.3) and expression of EBNA3C *in vivo* is thought to be limited to EBV-infected B blasts prior to their entry into GCs (Figure 4.17A). Therefore, EBNA3C expression has the potential to induce AID in pre-GC EBV-infected B blasts resulting in AID expression for a longer time and possibly at higher levels. Although LMP1 and LMP2A were believed to provide the necessary pro-survival signals for EBV-infected B cells to survive the GC reaction, it has been found that EBV-infected B cells still undergo CSR (Babcock et al., 1998; Joseph et al., 2000a) and SHM with a pattern that is similar to antigen-selected memory B cells (Souza et al., 2005). Therefore, EBV-infected B cells appear to participate in the normal B cell selection process in GCs and it is possible that increased levels of AID might provide them with a higher probability of surviving the GC reaction. This would be consistent with the observation that EBV positive memory B cells are more likely to have undergone CSR and to contain a higher number of SHM compared to EBV negative memory B cells (Souza et al.,

2007). Although the induction of AID by EBNA3C *in vitro* was relatively slow (see Section 5.1.1 for an extended discussion on the kinetics of EBNA3C-mediated gene regulation) this might be sufficient to favour EBV-infected B cells by expressing AID before they enter GCs and therefore over a longer time-frame and maybe at higher levels. It is also possible that the induction is more rapid *in vivo*.

This does however not explain why EBNA2 has the ability to counter-act EBNA3C-mediated induction of AID. It is possible that EBNA2 is required to fine-tune AID expression. EBNA2 expression has to cease in order to allow EBV-infected B blasts to express BCL6, to enter GCs and to differentiate. This will also result in the loss of EBNA3C expression, but AID expression probably remains high due to signals from T_{fh} cells in the GC environment. Repression of AID by EBNA2 might be a viral fail-safe mechanism that prevents overly high levels of EBNA3C-induced AID, which could potentially be disadvantageous for the host and virus (see below), when EBV-infected cells fail to switch off EBNA2 and therefore cannot transit through the GC to become memory B cells (Figure 4.17B).

Interestingly, recently EBV infected B cells were discovered in the marginal zone surrounding GCs and these had the phenotype of non-class switched memory B cells (IgD+ and CD27+). In addition, they expressed the EBV latency III programme and showed signs of extensive proliferation [Figure 4.17B, (Thorley-Lawson, 2015)]. These cells seem to have failed to switch off EBNA2 and could therefore not enter a GC, but it remains be confirmed that they are BCL6-negative and AID-positive. If they are AID-positive, EBNA2 might ensure that AID expression levels are restricted by counter-acting the EBNA3C-mediated activation of AID. Latency III expressing B blasts are effectively targeted by CTLs (see Section 1.2.5.4) so it remains to be determined how common these marginal zone EBV-infected B blasts are, how long they survive and whether they are important in the EBV life cycle.

These cells could be the *in vivo* counterparts of LCLs established by infection of naïve B cells with EBV in culture, which remain BCL6-, become CD27+, express AID and undergo SHM [(Siemer et al., 2008) and this study], but not CSR without external T cell

stimuli and therefore remain IgD+ (Heath et al., 2012). It has been suggested that EBV infection *per se* could be sufficient to reach the memory B cell compartment without the need for cells to transit GCs [Figure 4.17B, (Heath et al., 2012)]. However, this remains to be confirmed *in vivo* and currently an adequate explanation of how EBV would switch directly from latency III in proliferating B blasts to latency 0 in resting memory B cells is lacking (Thorley-Lawson, 2015).

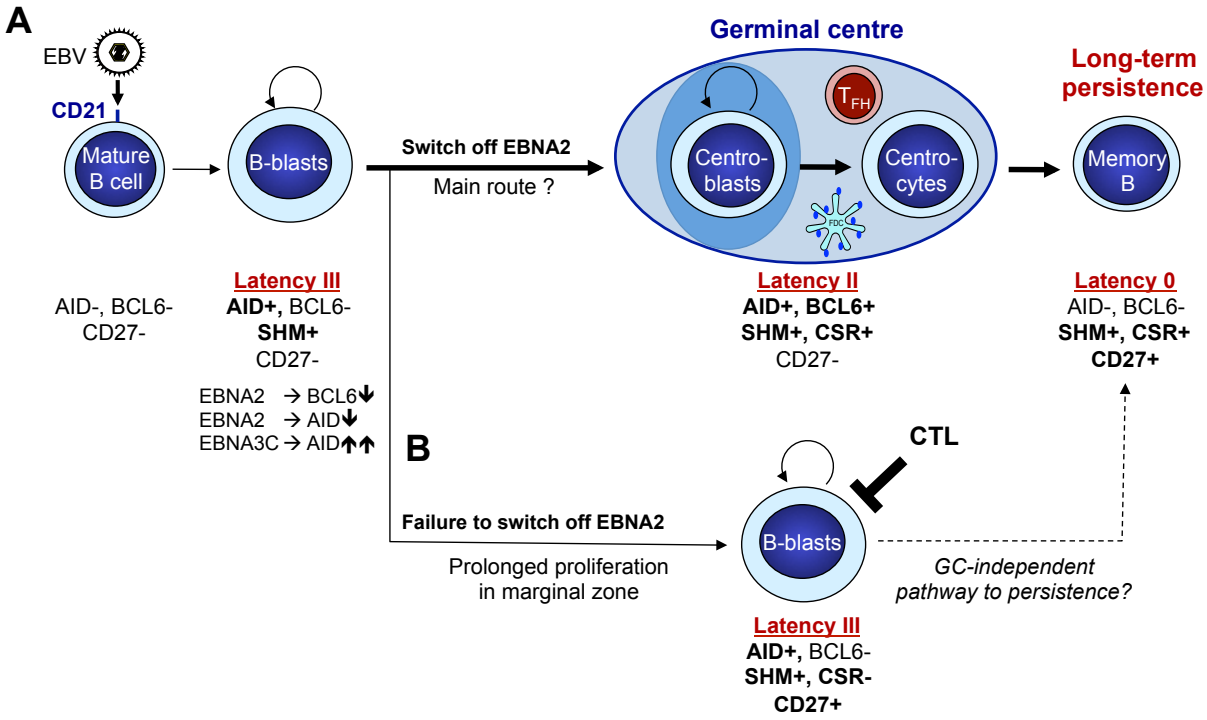


Figure 4.17: Are there different routes to EBV latency and what is the role of AID?

(**A**) Classical germinal centre (GC) model of persistence (see Figure 1.7 for details). Latency III expressing EBV-infected B blasts potentially start expressing AID and to undergo somatic hypermution (SHM) as a result of EBNA3C expression, but are unable to enter GCs due to the repression of BCL6 by EBNA2. Switch from latency III to latency II allows B blasts to enter GC and to undergo class switch recombination (CSR). CD27+ memory B cells emerge from GC reaction and enable persistent infection. (**B**) Prolonged proliferation of latency III expressing EBV-infected B blasts could be possible in the marginal zone surrounding GCs. These cells would originate from cells that fail to repress EBNA2 and might represent the *in vivo* counterpart of LCL (AID+, BCL6-, SHM+, CSR- and CD27+). They represent targets for cytotoxic T lymphocytes (CTL) as they express latency III. It might be possible that the memory B cell compartment is reached via GC-independent pathways.

A major unanswered question is what effect does AID have on EBV-infected B cells. Previous *in vitro* studies have detected ongoing SHM in the IgH V(D)J region of an established LCL derived from a rheumatoid arthritis patient (Gil et al., 2007; Chezar et al., 2008; Laskov et al., 2011), but in newly EBV-infected naïve B cells only around 30-35% of the developing clones exhibited IgH V(D)J SHM (Siemer et al., 2008; Heath et al., 2012). Heath and colleagues reported that in 5 of 6 limiting dilution experiments, EBV-infected naïve B cells that contained signs of SHM became either the dominant or significant sub-dominant clone and hypothesised that these might have obtained a growth advantage through off-target SHM at non-Ig (Heath et al., 2012). Epeldegui *et al.* reported that infection of B cells with EBV resulted in mutations at *BCL6* and *TP53* that accumulated over time in the culture suggestive of ongoing SHM caused by AID (Epeldegui et al., 2007). However, by the time that AID reaches full-activity in LCLs, expression of *BCL6* is highly repressed [see Figure 4.1 and (Gil et al., 2007; Siemer et al., 2008)] and it is therefore very unlikely that BCL6 is an off-target locus for AID activity in LCL. Consistent with this, Heath and colleagues failed to detect significant levels of mutations at the *BCL6* locus in EBV-infected naïve B cells compared to CD27+ memory B cells (Heath et al., 2012). Therefore the observations from Epeldegui and colleagues are most likely explained by pre-existing mutations at the *BCL6* locus in some of the EBV-infected B cells that potentially became more dominant over time in the culture. B cell populations infected with EBV-infection remain polyclonal for around four weeks post infection, but thereafter they quickly become more clonal and usually approach monoclonality by 8-12 weeks post infection (Ryan et al., 2006; Heath et al., 2012). It is currently unclear what drives this clonal development of a LCL culture, but it has to be considered when performing mutational analysis.

In order to determine whether the high levels of EBNA3C-induced AID was functional in the 3CHT LCL, SHM at the IgH V(D)J region was analysed in samples of the 3CHT A13 time-course in collaboration with Dr Rachael Bashford-Rogers and Prof Paul Kellam. For this, RNA from samples from the 3CHT A13 time-course was converted into cDNA and used

as template to PCR amplify the IgH V(D)J region. The PCR products were subjected to DNA deep-sequencing and Dr Bashford-Rogers performed the bioinformatics analysis. The clonal distribution within each sample was determined using the fact that each B cell clone has a uniquely recombined IgH V(D)J region, which act as molecular barcode. This revealed that at the start of the 3CHT time-course four major clones composed >90% of the cell population, which thereafter developed similarly in samples with active or active EBNA3C over the 60 day time-course (Figure 7.8A). All four major clones were subjected to V(D)J sequence analysis, which revealed that from day 30 onwards there were significantly more mutated V(D)J sequences in samples with active EBNA3C compared to samples with inactive EBNA3C (Figure 7.8B). This correlated with the time when maximal expression levels of AID were observed (Figure 4.2).

Having found that EBNA3C-activated AID is functionally active in 3CHT LCL, we initiated a collaboration with Dr Peter Campbell aimed at determining AID-induced mutations at non-Ig loci across the whole human genome. Selected samples of the 3CHT A13 time-course were subjected to HiSeq X Ten whole-genome sequencing (Illumina) with 30x coverage. Unfortunately, the analysis of this is still ongoing, because due to changes in the clonal composition of the LCL culture over the course of the time-course (Figure 7.8A) it is challenging to differentiate between genuine *de novo* mutations (AID-dependent and –independent) and pre-existing mutations that start appearing because of an expansion in clone size. However, preliminary results indicate that EBNA3C-activated AID also introduced mutations at non-Ig loci that appear as kataegis, localised clusters of hypermutations (Nik-Zainal et al., 2012; Alexandrov et al., 2013). Single-cell sorting of 3CHT cells to establish a clonal 3CHT culture that could be sequenced before and after AID is activated by EBNA3C could circumvent this problem, however LCLs do not generally grow at low cell densities. BL cell lines do not exhibit this limitation allowing this approach to be employed to determine AID off-targets in the human BL cell line Ramos (Qian et al., 2014). Additional overexpression of AID and inhibition of UNG was also used in order to increase the

detection rate of AID-induced mutations. AID off-target mutations were observed in kataegic clusters at loci that showed characteristic marks of super-enhancers, *e.g.* IgH, *cMYC* and *BCL6* (Qian et al., 2014). Therefore, it is likely that EBNA3C-induced AID can also introduce mutations at non-Ig loci, which might directly affect the growth and/or survival rate of the EBV-infected B cell and hence – *in vivo* – potentiate the risk of B cell lymphomagenesis (see Section 1.2.5).

Further analysis that remains to be performed on the existing whole genome sequencing data is to determine the mutational impact of AID on EBV episomes. Interestingly, members of the APOBEC family of cytidine deaminases, to which AID belongs, have been frequently described as cellular restriction factors that confer innate immunity against viruses, such as HIV, human T lymphotropic virus 1 (HTLV-1), hepatitis B and C viruses (HBV and HCV), human papilloma virus (HPV) and Kaposi's sarcoma associated herpes virus (KSHV) [reviewed in (Moris et al., 2014; Harris and Dudley, 2015)]. For example it has been reported that infection of B cells with the gamma-herpesvirus KSHV resulted in increase AID levels and the overexpression of AID reduced lytic reactivation and therefore production of infectious viral particles (Bekerman et al., 2013). Interestingly, two KSHV-encoded miR were identified capable of binding in the 3'UTR of AID and reducing AID protein levels. In this light it appears surprising that EBV specifically upregulates AID expression unless the benefits (see above) outweigh the associated risks for EBV.

4.4.5 Conclusion

The findings presented in this Chapter provide novel insights into how EBNA3C activates gene expression at a particularly complex locus. EBNA3C targets most of the regulatory elements of AID and induces AID expression and this depends on the ability of EBNA3C to bind to RBPJ. It could be confirmed that EBNA2 represses AID expression and so is tempting to speculate that EBV via EBNA3C might activate AID expression in order to increase its chances of accessing the memory B cell compartment to establish a persistent infection. The risk this poses to both the cell and viral genomes is possibly controlled by the

repressive activity of EBNA2 limiting AID expression levels. However, much work is still required to determine the reasons and consequences of AID induction by EBV *in vitro*, but more importantly *in vivo*.

AICDA is a very interesting gene to study because of the complexity of its regulation. However, further studies, preferentially on genes that are regulated by EBNA3C alone are needed in order to determine whether the findings presented here apply more generally to EBNA3C activated genes and to better understand the precise molecular mechanisms underlying EBNA3C-mediated gene activation.

5 Chapter V: General discussion of gene repression and activation by EBNA3C

The analysis of *ADAM28*, *ADAMDEC1* and *COBLL1* as EBNA3C repressed genes and *AICDA* as an induced gene by EBNA3C revealed similarities and differences in EBNA3C-mediated gene regulation mechanisms that will be further discussed here.

It was remarkable to observe such dramatic changes in gene expression, *e.g.* at least one log fold up regulation of *AICDA* and 3-4 log fold repression of *COBLL1*, occurring simultaneously within the same virus infected cell population and probably within the same cell. Unfortunately, techniques employed here only describe changes at a population level and do not allow one to comment on gene expression changes in single cells. Attempts to analyse protein expression at a single cell level, *e.g.* by flow cytometry or immunofluorescence did not produce reliable results probably due to lack of antibody sensitivity or specificity in these techniques. As an alternative, RNA abundance could have been assessed directly by RNA flow cytometry, which is likely to exhibit a greater specificity and sensitivity, in order to allow conclusions on whether gene expression changes uniformly in all cells or only in a small subset of cells within the culture. Currently, I see no reasons why there should be a significant difference in the response to EBNA3C between individual cells within the culture, but this remains to be determined. Nevertheless, interesting observations were made, but it is important to remember that these are derived from a population of cells over many population doubling times.

Comparative gene expression analysis of EBNA3C repressed (Figure 3.6) and activated (Figure 4.4) genes between EBV-infected primary B cell and the EBNA3C-conditional time-courses after activation of EBNA3C revealed that the 3CHT system could be efficiently used to recapitulate both repressive and activating changes in gene expression seen early after infection of primary B cell with EBV. This is the first time that this has been shown, which is important, because it validates the full functionality of the 3CHT fusion

proteins and their utility as a model system to study EBNA3C-mediated gene regulation. No changes in gene expression of EBNA3C target genes were observed after addition of HT to non-conditional 3CKO LCL proving that changes in gene expression in 3CHT LCL are a direct consequence of EBNA3C activation (Figure 3.5 and Figure 4.3).

5.1 Similarities between EBNA3C-mediated repression and activation

5.1.1 Kinetics of EBNA3C-mediated gene regulation

It is striking that a relatively long period of 30-60 days post infection or activation of EBNA3C by HT was required for full repression/activation of EBNA3C target genes and equally long for full de-repression/de-activation after inactivation of EBNA3C (Figure 3.4 and Figure 4.2). It is however not the first time that this has been observed. Previous studies on EBNA3-mediated gene repression, *e.g. CDKN2A* (Skalska et al., 2010; 2013), *BCL2L11* (Paschos et al., 2012) and *CXCL9/10* (Harth-Hertle et al., 2013), or gene activation, *e.g. AICDA* (Gil et al., 2007; Heath et al., 2012), also reported time-scales of up to four weeks or sometimes even longer for full activation or repression.

It is possible that it takes that long for EBNA3C associated complexes, to fully establish or remove epigenetic histone marks through recruitment of auxiliary factors – particularly as it was found that changes in activation-associated histone marks correlated very well with the levels of gene expression. However, taking into account the proliferation rate of LCL with a doubling time of approximately 24-48 hours, this probably means that most cells undergo up to 30 cell divisions by the time complete repression or activation is achieved. DNA replication and mitosis are both processes that disrupt the chromatin structure locally and globally, respectively. Most TFs dissociate from mitotic chromatin that is up to 50-fold more condensed compared to interphase chromatin [reviewed in (Belmont, 2006)] and the disintegration of the nuclear envelope results in a major change to the nuclear architecture. Interphase nuclear organisation has to be re-established in the two post-mitotic daughter cells. Epigenetic bookmarking has been proposed to help re-

establishing gene expression profiles in post-mitotic cells [reviewed in (Zaidi et al., 2010)]. For this, certain epigenetic histone modifications such as H3K4me2/me3 and H3K27me3 are maintained during mitosis [reviewed in (Wang and Higgins, 2013)] and while most TF dissociate from mitotic chromatin, some (*e.g.* RUNX factors) are retained during mitosis [reviewed in (Kadauke and Blobel, 2013)]. Intriguingly, RBPJ has recently been found to remain stably associated with chromatin during mitosis and to interact with CTCF (Lake et al., 2014). Since on chromatin EBNA3C appears to be associated with RBPJ, this might ensure coordination of EBNA3C-mediated gene regulation over many cell divisions. Although 30 generations appear remarkably long, it is possible that EBNA3C-mediated gene regulation is somehow linked to the proliferation rate of the cells (see below). Currently, however, the reasons why EBNA3C-induced changes in gene expression occur over such long periods of time remain unknown, especially when compared for example to EBNA2-mediated repression of AID, which only required four hours for a 30-fold change in *AICDA* expression (Figure 4.12A). ChIP analysis on samples of the EREB 2-5 time-course are required to better understand these differences in kinetics between EBNA2- and EBNA3C-mediated gene regulation.

As surprising as the long time-frame required to achieve maximal regulation is the fact that all EBNA3C-regulated genes displayed very similar and highly exponential regulatory profiles (R^2>0.98) in newly EBV-infected primary B cells and also after activation of conditional EBNA3C in the 3CHT system (Figure 5.1A). The only difference between the individual genes beside the direction of the regulation (repressed vs activated) is how fast the changes occur for each gene. Consistent with the idea that the *ADAM* locus is co-ordinately repressed by EBNA3C both *ADAM28* and *ADAMDEC1* display almost identical regulation profiles. Intriguingly, an almost identical exponential behaviour in the opposite direction was observed for de-repression/de-activation after inactivation of EBNA3C (Figure 5.1B-E). The reasons for this are currently unknown.

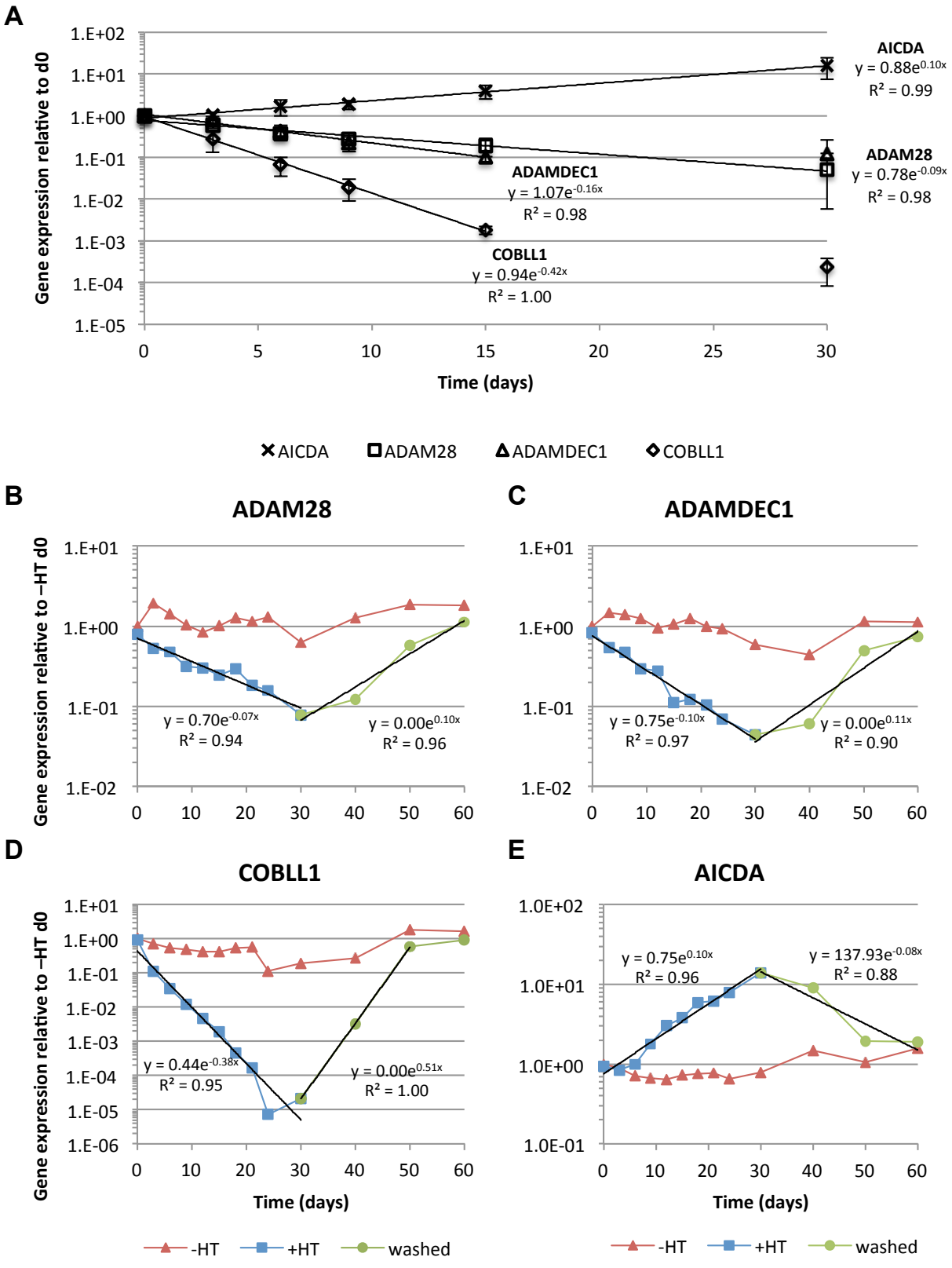


Figure 5.1: Gene expression analysis of EBNA3C-regulated host genes.

(**A**) Gene expression data for *AICDA*, *ADAM28*, *ADAMDEC1* and *COBLL1* from primary B cell infection with wtI6 EBV and three replicate 3CHT time-courses after activation of EBNA3C (A13 Rep1+2 and C19) are displayed as mean values ± SD. Exponential trendlines were fitted up to day 30

for *AICDA* and *ADAM28* and up to day 15 for *ADAMDEC1* and *COBLL1*, because repression is either achieved by this time (*ADAMDEC1*) or slowed down after this time-point (*COBLL1*). (**B-E**) Gene expression data for *ADAM28* (**A**), *ADAMDEC1* (**B**), *COBLL1* (**C**) and *AICDA* (**D**) over the 3CHT A13 time-course of cells grown in the absence (-HT), presence (+HT) of HT or in cells with HT removed at day 30 (washed). Exponential trendlines were fitted over the period of 30 days for +HT and washed conditions (trendline for washed COBLL1 expression is only fitted for the first 20 days after removal of HT). Equations and R^2 values are shown for each trendline.

Because both primary B cells and LCLs were kept mostly in the exponential growth phase during these time-courses together with the observed exponential regulation profile, it would be interesting to analyse whether there is a link between gene regulation by EBNA3C and cellular proliferation and whether an altered proliferation rate has an effect on its kinetics. For example, it would be possible to alter the proliferation rate of 3CHT LCL either through changes in cell culture conditions (*e.g.* different temperature) or through a fine-tuned ectopic regulation of cMYC expression. Increased expression levels of cMYC in LCL were shown to increase the percentage of cells in S-phase from 33% to 47% and hence resulted in a higher overall growth rate (Cutrona et al., 1995) without affecting the expression levels of latent EBV gene products (Hotchin et al., 1990). In contrast, inhibition of cMYC expression, either by chemical inhibition or by expression of a dominant-negative cMYC mutant, resulted in growth arrest of LCL (Faumont et al., 2009). Fine-regulation of the cMYC inhibition might only decreased the proliferation rate without inducing growth arrest. Unfortunately, cMYC regulates many genes and its deregulation has widespread affects on gene transcription.

5.1.2 Recruitment of RBPJ is dependent on presence of EBNA3C

Originally, RBPJ was thought to be stably associated with sequence-specific binding sites on DNA (Figure 1.16A) to silence gene expression by recruiting corepressors, which can be replaced by activators upon binding of NICD or EBNA2 (Figure 1.16B). Dynamic changes to RBPJ binding have been previously described for both NICD and EBNA2 (see Section 3.4.1), but this is the first time this has been observed for EBNA3C in gene repression and activation.

Low level RBPJ occupancy was observed in the absence of functional EBNA3C at all analysed EBNA3C-binding sites, which increased dramatically upon activation of EBNA3C suggesting that functional EBNA3C can promote the accumulation of high levels of RBPJ to specific sites on chromatin (Figure 3.10D+H and Figure 4.8D). It is important to mention that, although lower levels of RBPJ were detected at the EBNA3C-binding sites in the absence of functional EBNA3C, these levels were higher than at control sites on each of the three loci. This might be important for the initial recruitment of EBNA3C. It is then possible that the interaction between EBNA3C and RBPJ increases the binding to and/or stabilises EBNA3C/RBPJ complexes on regulatory elements that control expression of EBNA3C target genes, which might then serve as assembly platform for multi-protein complexes, comprised of additional repressors or transactivators (see also 5.3.1).

Consistent with this, the ChIP analysis on the 3CHT A13 and C19 time-courses for all three genomic loci revealed that RBPJ was the first factor to increase at the EBNA3C binding sites (Figure 5.2A). This was subsequently followed by an increase in BMI1 occupancy at these sites (Figure 5.2B). The role of BMI1 in repression (Chapter 3.4) and activation (Chapter 4.4) has been discussed previously. Interestingly, no difference in the recruitment profile of RBPJ and BMI1 could be detected between EBNA3C-repressed or activated genes. Surprisingly, it appeared that the occupancy of RBPJ and BMI1 at the EBNA3C-binding sites decreased over time. Either high levels of these factors are only needed to initiate EBNA3C-mediated gene regulation or it is conceivable that the assembly of a multi-protein complexes can mask the epitope detected by the antibodies used in the ChIP experiments, which would prevent detection at later time-points and look like reduced binding.

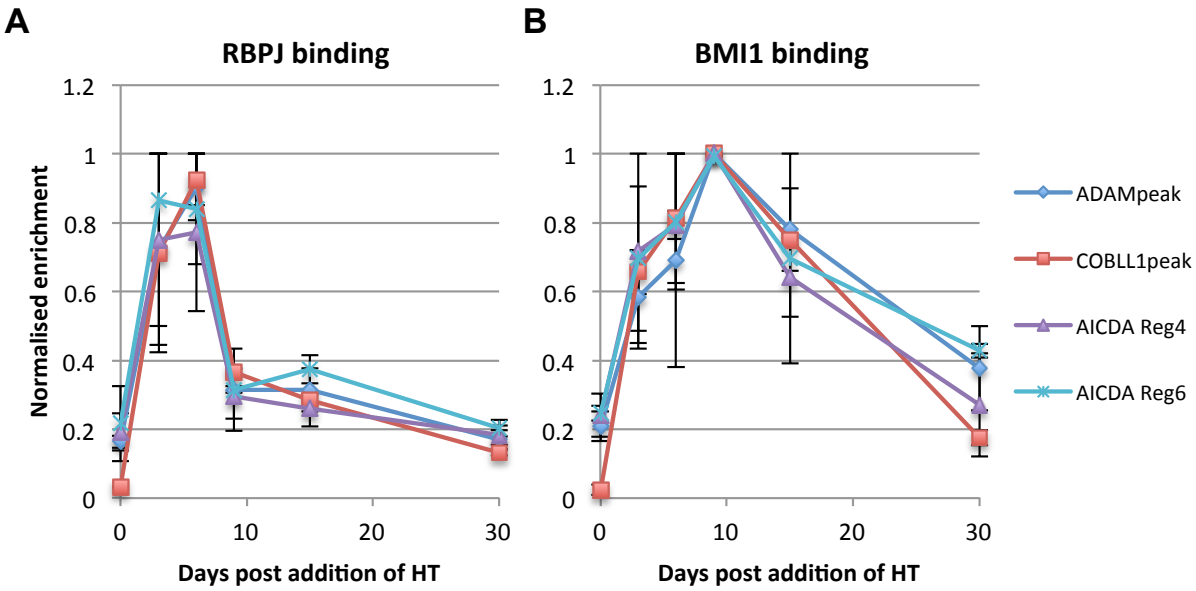


Figure 5.2: Binding of RBPJ and BMI1 to EBNA3C repressed and activated genes.
ChIP-qPCR values for RBPJ (**A**) and BMI1 (**B**) at *ADAM* peak, *COBLL1* peak, *AICDA* regions IV and VI from the two biological time-courses (3CHT A13 and C19) were normalised by setting the maximal enrichment level during each time-course to one and by using the first time-point –HT (day 3) as representative value for day 0. Mean values ± SD from both replicate time-courses are shown. For BMI1, day nine of the C19 time-course was treated as an outlier and was not taken into account for the analysis.

Interestingly, DNA sequence analysis using publically available transcription factor binding prediction software [PROMO (Messeguer et al., 2002; Farré et al., 2003) and Patch 1.0 (BIOBASE)] did not reveal strictly canonical RBPJ binding sites at the EBNA3C peaks at *ADAM28*/*ADAMDEC1*, *COBLL1* or *AICDA*. Using a less stringent interpretation several possible sites at each EBNA3C peak were identified, but it was not possible to determine whether any of these was responsible for recruiting RBPJ. Recent ChIP-Seq analysis of EBNA3C also reported that DNA sites occupied by EBNA3C were not specifically enriched in RBPJ recognition motifs (Wang et al., 2015). It is also possible that binding of EBNA3C to RBPJ changed the DNA binding preference of RBPJ or that other TF are involved in the recruitment (see Section 5.3.1).

The phenotypic importance of the interaction between EBNA3C and RBPJ first became apparent in transient reporter assays using RBPJ BM EBNA3C and RBPJ knock-out cells to recapitulate the repression of *ADAM28* and *COBLL1*, but more importantly also

in the context of viral infection – the RBPJ BM EBNA3C virus failed to repress *ADAM28*, *ADAMDEC1* and *COBLL1* or to up regulate *AICDA*. It is possible that the mutations introduced in the RBPJ BM EBNA3C alter the interaction with other yet to be described factor(s) required for EBNA3C-mediated gene regulation, but even wild type EBNA3C did not successfully repress *ADAM28* luciferase constructs in the RBPJ-null DG75 cells, clearly indicating that the presence of RBPJ is essential for EBNA3C-mediated gene regulation.

In contrast to previous studies (Lee et al., 2009; Maruo et al., 2009; Calderwood et al., 2011), the RBPJ BM EBNA3C virus was able to establish a stable LCL indicating that the interaction between EBNA3C and RBPJ is not absolutely required for B cell transformation. Further studies are required to investigate the importance of the interaction between EBNA3C and RBPJ.

5.2 Differences between EBNA3C-mediated repression and activation

Surprisingly few differences were detected between EBNA3C-mediated gene repression and activation. The most prominent were observed changes to epigenetic histone marks that differed between EBNA3C induced repression or activation. For the repressed genes *ADAM28*, *ADAMDEC1* and *COBLL1*, activation-associated histone marks quickly disappeared – predominantly at the TSS of each gene – while for induced *AICDA* these marks appeared, again predominantly around the TSS, but also at a distal enhancer element. There was a good correlation between activation-associated marks and levels of gene expression (Figure 3.9 and Figure 4.7). Increased levels of the repressive histone modification H3K27me3 were observed at all EBNA3C-repressed loci and PRC2 at the *COBLL1* locus, neither of which were observed at the activated *AICDA* locus. In contrast, acetlytransferase p300 was detected at some of the regulatory regions of *AICDA*, but not consistently at the repressed *ADAM28*/*ADAMDEC1* and *COBLL1* loci. More work is needed to understand how the mechanisms underlying EBNA3C-mediated gene repression and activation differ and why EBNA3C reproducibly acts as a very potent repressor on some

genes, but as a transactivator on others in the same timeframe and in the same population of cells (see Section 5.3.4). This leads to some important questions that remain unanswered.

5.3 Open questions

5.3.1 How is EBNA3C recruited to chromatin and which factors are important?

The observations that RBPJ levels increased shortly after activation of EBNA3C at all EBNA3C binding sites at the *ADAM*, *COBLL1* and *AICDA* loci and the fact that EBNA3C failed to regulate all three loci when it was unable to bind to RBPJ, suggest that RBPJ is an essential factor for the recruitment of EBNA3C. Unfortunately, in my hands, direct ChIP for EBNA3C during the 3CHT A13 and C19 time-courses did not yield reproducible results using a commercial polyclonal antibody against EBNA3C that also precipitates EBNA3A and EBNA3B (McClellan et al., 2013). The reasons for this are unknown, but it is worth mentioning that when the EBNA3 ChIP seemed to work, it essentially produced the same result as ChIP for RBPJ, again supporting the idea that both are required to form stable complexes at the EBNA3 peaks. In order to bypass this, ChIP was attempted using an antibody against the modified estrogen-receptor that is fused to EBNA3C in the 3CHT LCL, but this resulted in very poor enrichment relative to input (data not shown) preventing its use on samples of the 3CHT time-courses. In order to reliably ChIP for EBNA3C during these time-courses, the conditional EBNA3C on the p16-null background would also need to be TAP-tagged, but this virus and cell line are not yet available and it might not be possible to construct them because it is not known how the combined tags would affect the function and stability of EBNA3C.

In addition to RBPJ, it is likely that other transcription factors are involved in targeting EBNA3C to DNA. Previous ChIP-Seq studies reported that EBNA3 binding sites coincided with ENCODE ChIP-Seq binding sites of multiple transcription factors in LCL GM12878, *e.g.* RUNX3, BATF, IRF4, BCL11A and PAX5 (McClellan et al., 2013; Jiang et al., 2013; Schmidt

et al., 2014; Wang et al., 2015), while only 16% (Jiang et al., 2013) or 44% (Wang et al., 2015) of EBNA3C binding sites were reported to be co-bound by RBPJ (see Section 1.4.2). All of these transcription factors appear to occupy the EBNA3C binding sites at *ADAM28*/*ADAMDEC1*, *COBLL1* and *AICDA* in LCL GM12878, in which EBNA3C is present and all genes are robustly regulated by EBNA3C (Figure 7.9B, Figure 7.10B and Figure 7.11B). IRF4, in particular, is thought to be involved in the recruitment of EBNA3C (Jiang et al., 2013) and overexpression of IRF4 resulted in increased EBNA3C occupancy at EBNA3C-binding sites (Wang et al., 2015). Preliminary results from Dr Kostas Paschos in our lab indicate that shRNA-mediated knock down of IRF4 in EBNA3C-TAP LCL reduces EBNA3C occupancy at the *COBLL1* peak and *AICDA* regulatory regions II and IV compared to the non-targeting control shRNA. Therefore, it is likely that IRF4 could act as a co-factor to help directing EBNA3C/RBPJ to these target sites.

5.3.2 Why does EBNA3A bind to the COBLL1 peak and the regulatory regions of AICDA although it is not involved in their regulation?

The ChIP-Seq on EBNA3A-TAP LCL detected binding of EBNA3A to the *COBLL1* peak (Figure 3.8A) and most of the regulatory regions of *AICDA* (Figure 7.11B), which has been verified by ChIP-qPCR (see Figure 3.7 for EBNA3A at the *COBLL1* peak). It is currently unclear why EBNA3A can be detected at these two genes although the infection of primary B cell with 3AKO virus showed that EBNA3A was not required for in the repression of *COBLL1* (Figure 3.2C) or the up regulation of *AICDA* (Figure 4.1A). Furthermore, in LCL with conditional EBNA3A, no changes to the expression levels of *COBLL1* or *AICDA* could be detected upon activation or inactivation of EBNA3A (Dr Quentin Bazot, personal communication).

It is unlikely that the occupancy of EBNA3A at these two genes is caused by an artefact of the TAP-tagged LCL cell lines, because a recently published ChIP-seq experiment using anti-HA antibody in EBNA3A-Flag-HA LCL (Schmidt et al., 2014) also detected EBNA3A on the *COBLL1* peak (Figure 7.10A) and *AICDA* regulatory regions II, IV,

VI and possibly V (Figure 7.11A). One possible explanation of why EBNA3A is found at these two genes, although it does not have an apparent biological role, could be due to its ability to bind to EBNA3C (Calderwood et al., 2007; Paschos et al., 2012; McClellan et al., 2013). This could also explain why EBNA3C binding can be detected together with EBNA3A at the regulatory regions of genes that are only regulated by EBNA3A (Dr Quentin Bazot, personal communication). Consistent with this, ChIP-Seq analysis of our TAP-tagged EBNA3 cell lines revealed that 83% of EBNA3A binding sites are co-occupied by EBNA3C and 37% of EBNA3C sites also display binding of EBNA3A (Dr Kostas Pachos, personal communication). This is even higher compared to previous studies reporting that 26% (Wang et al., 2015) or 50% (Schmidt et al., 2014) of EBNA3A binding sites are co-occupied by EBNA3C. Although the numbers vary between studies, all suggest extensive overlap between EBNA3A and EBNA3C targeting.

However, it remains to be determined whether EBNA3A and EBNA3C bind to these regulatory sites at the same time and in the same cell. Because ChIP results are averaged binding events observed in many cells, it is possible that a site is only bound by EBNA3A in some cells, but by EBNA3C in others. In order to determine whether both bind at the same time, re-ChIP has to be performed on EBNA3A and EBNA3C carrying different epitope-tags. Unfortunately, no recombinant EBV is currently available that carries differentially epitope tagged EBNA3A and EBNA3C. To circumvent this, it would be possible to overexpress HA-tagged EBNA3A in EBNA3C-TAP LCL or vice versa, but this might disrupt normal targeting of EBNA3 proteins to their natural binding sites. Novel insights into the targeting of EBNA3A and EBNA3C is likely to result from newly created (by Dr Quentin Bazot) EBV recombinants that carry TAP-tagged EBNA3A and conditional EBNA3C allowing one to analyse EBNA3A targeting to the *COBLL1* peak and the regulatory regions of *AICDA* in the absence or presence of EBNA3C.

Taking into account the low transcript number (Davies et al., 2010) and low protein abundance of EBNA3A and EBNA3C per individual cell, it is very surprising and hard to

believe that these proteins are accidentally targeted to sites where they do not have any biological role. Additionally, although both EBNA3A and EBNA3C are involved in the repression of the *ADAM28*/*ADAMDEC1* locus, it remains unclear why EBNA3A plays only a subsidiary role relative to EBNA3C (Figure 3.2A and B). Further studies are needed to understand better the interplay between the different EBNA3s.

5.3.3 How are PRC1 and PRC2 proteins recruited?

The role of polycomb proteins in EBNA3C-mediated gene repression (3.4.2.2) and activation (4.4.2) was discussed previously. Beside their mechanistic role in gene regulation, it remains unclear how they are targeted to EBNA3C-regulated genes.

5.3.3.1 Default mechanism

The recruitment of polycomb proteins might be a default mechanism of these genes when they are repressed or activated. It is conceivable that removal of activation-associated histone modifications on *ADAM28*, *ADAMDEC1* and *COBLL1* is enough to attract polycomb proteins. If this is the case, preventing the removal of histone acetylation by addition of HDAC inhibitors during a 3CHT time-course might be enough to ensure that *ADAM28*, *ADAMDEC1* and *COBLL1* are not repressed by EBNA3C and polycomb proteins are not recruited even in the presence of EBNA3C. Unfortunately, LCLs tolerate HDAC inhibitors very poorly, even over relatively short time-courses of 1-2 days (Paschos et al., 2009). Alternatively, small molecule inhibition of transcription, similar to Riising *et al.* 2014, could be used in 3CHT LCL with inactive EBNA3C to determine whether EBNA3C-independent inhibition of *ADAM28*, *ADAMDEC1* and *COBLL1* transcription is sufficient to recruit polycomb proteins. On the other hand, IL4 and CD40L activation of primary B cells combined with ChIP analysis could be used to determine whether EBNA3C-independent up regulation of *AICDA* involves the recruitment of BMI1 to the regulatory regions of *AICDA*.

5.3.3.2 Direct recruitment

Another possibility is that polycomb proteins were directly recruited either by EBNA3C or another factor in an EBNA3C-associated complex. One initial aim of this PhD project was to characterise and map the interaction between SUZ12 and EBNA3C. Unpublished results from our lab indicated that EBNA3C could be co-immunoprecipitated with SUZ12 in EBV-infected BL31 cells and LCL, and also in GST-SUZ12 pull-down experiments (Sara Temelso, 2012, Master's thesis, Imperial College London). However, the GST-SUZ12 fusion protein purified from bacteria was not very pure and was badly degraded. Therefore, I established different production and purification strategies, optimised with help from Dr Goedele Maertens. Unfortunately, it turned out that the production and purification of full length SUZ12 was technically very challenging. SUZ12 was found to be either highly unstable or subjected to protein degradation in bacteria especially at increased temperatures. The best result was obtained when SUZ12 was produced from an IPTG-inducible pET28a-His-SUMO-SUZ12-TAP vector in Rosetta II at room temperature and harvested two hours after induction. The purification strategy involved a four-step process (His-tag purification using HIS-Select, Strep-tag purification using StrepTactin, ionic exchange chromatography using HiTrap SP FF and size exclusion chromatography using a SD200 column). Unfortunately, Strep-tag pull-down experiments using the purified SUZ12-TAP fusion protein and *in vitro* translated EBNA3C could not reproduce a strong interaction with EBNA3C relative to the beads alone. This avenue of investigation was therefore dropped.

Moreover, subsequently the ChIP results on the 3CHT A13 and C19 time-courses showed that EBNA3C and SUZ12 did not bind at the same site, with EBNA3C at the *COBLL1* peak and SUZ12 at the TSS of *COBLL1*, and that SUZ12 was not directly detected at the *ADAM28/ADAMDEC1* locus. Therefore it is unlikely that SUZ12 is directly recruited by EBNA3C, even so it is possible that both sites loop in the presence of EBNA3C (see 3.4.3).

It now seems more likely that EBNA3C can recruit BMI1. The ChIP results on the 3CHT A13 and C19 time-courses showed that BMI1 was recruited wherever EBNA3C bound (the *ADAM* peak, the *COBLL1* peak and the *AICDA* regulatory regions II, VI, V and VI). Recruitment levels of BMI1 correlated well with the expected occupancy of EBNA3C at each site. This was further supported by recent ChIP-Seq for BMI1 and SUZ12 in LCLs used for the EBNA3C-TAP ChIP-Seq, which revealed that 47% of EBNA3C sites overlap with BMI1 sites, but only 1% with SUZ12 sites (Dr Kostas Paschos, personal communication). Furthermore, using proximity-ligation techniques coupled to immunofluorescence it could be shown that EBNA3C localises in close proximity to BMI1 (Gillian Parker, personal communication).

In addition, it has been shown that murine Bmi1 can be recruited to chromatin by direct interaction with the TF Runx1 (Yu et al., 2012). RUNX1 together with RUNX2 and RUNX3 are family members of tissue-specific core DNA binding transcription factors, which share a high level of protein homology. In LCLs EBNA2 upregulates RUNX3 (Spender et al., 2002; Gunnell et al., 2016), which in turn represses RUNX1 (Brady et al., 2009). Interestingly, 64-80% of EBNA3C binding sites have been found to be co-occupied by RUNX3 [(Jiang et al., 2013; Wang et al., 2015) and Dr Kostas Paschos, personal communication]. Furthermore, a strong overlap between binding sites for RUNX and RBPJ was reported in T-lymphoblastic leukemia cells, in which 90% dynamic RBPJ binding sites were found at distal enhancers with strong RUNX binding and characteristics of super-enhancers (Wang et al., 2011; 2014). Although concerns about the specificity of the RBPJ antibody used in these studies were raised as it also detects REST beside RBPJ (Castel et al., 2013), it is nevertheless possible that BMI1 is recruited by a large multiprotein complex composed of RUNX:EBNA3C:RBPJ, in which RUNX and RBPJ could mediate DNA binding and potentially also ensure mitotic bookmarking, while EBNA3C is perhaps required to stabilise the whole complex.

5.3.4 What determines whether a locus is repressed or activated by EBNA3C?

5.3.4.1 Distinct forms of repressive or activating EBNA3C complexes

It is conceivable that different forms of EBNA3C proteins exist, *e.g.* through different post-transcriptional modifications (PTM) on EBNA3C that determine the interaction with associated co-factors. Thus one form could be responsible for recruiting co-repressors whereas another for recruiting transactivators. Biochemical and mass spectrometric analysis could be employed to detect different PTMs. Furthermore, in order to determine whether activating forms or repressive forms of EBNA3C are preferentially targeted to induced or repressed genes, respectively, end-targeting proteomics of isolated chromatin fragments (ePICh) could be attempted (Déjardin and Kingston, 2009; Ide and Déjardin, 2015). This technique allows the identification of proteins by mass spectrometry that are bound at a distinct chromatin site. Currently it is limited to abundant loci (>100 copies per cell), therefore large BAC constructs containing loci targeted by EBNA3C could be used to provide up to several hundred BAC copies per cell (Bian and Belmont, 2010). By using BAC constructs containing EBNA3C repressed or induced genes, respectively, it might be possible to detect different forms of EBNA3C together with potentially novel interacting factors. The question that would remain, however, is how activating or repressive forms of EBNA3C are specifically recruited to induced or repressed genes, respectively.

5.3.4.2 Sequence of the EBNA3-binding site

It is possible, that all necessary information required for EBNA3C to activate or repress a specific locus is encoded by the DNA sequence at the EBNA3C binding site. To test this, the regulatory elements, to which EBNA3C is tethered, could be exchanged between repressed and activated genes using genome-editing techniques, *e.g.* CRISPR/CAS9 [reviewed in (Hsu et al., 2014; Doudna and Charpentier, 2014; Wiles et al., 2015)]. This could be used for example to exchange the DNA region around the *COBLL1* peak with that around *AICDA* region IV. If the DNA sequence alone is important, this should result in *COBLL1* being activated and *AICDA* being repressed (or not induced) after

activation of EBNA3C, in which case the DNA sequences of the EBNA3C binding sites will have to be analysed more precisely in order to detect what confers this information.

5.3.4.3 Locus specific features or factors

If the DNA sequence of the EBNA3C binding site does not contain regulatory information, there might be additional locus-specific features or factors in that particular genomic context, which are required to provide the regulatory information. Analysis of the higher order chromatin structure(s) or its localization within the nucleus when EBNA3C is active or inactive could provide important information regarding the effect of the locus. Again, genome editing techniques could be used to alter sites at the locus that are likely to change the 3D conformation (*e.g.* CTCF sites) and consequently maybe its location in nuclear space in order to assess the effect of the changed locus on EBNA3C-mediated gene regulation.

It is likely that not a single, but rather a combination of all three above listed possibilities act together in order to control EBNA3C-mediated gene regulation.

5.4 Final conclusions

The findings made during the course of this project provide important information on the EBNA3C-mediated regulation of three repressed genes and one activated gene. First of all, it was important to verify that all four genes were genuine targets of EBNA3C by infecting primary B cells with recombinant EBV and to identify that EBNA3C-conditional LCLs can be used to recapitulate reliably EBNA3C-mediated changes in gene expression observed early after infection of primary B cells with EBV. This was shown for the first time and revealed the full functionality of the 3CHT system, which will be also important for future studies on EBNA3C-mediated gene regulation.

ChIP analysis has revealed similarities and differences between EBNA3C-mediated gene repression and activation occurring simultaneously in the same cell population and has provided novel insights into the temporal sequence of events and the dynamics of factor

recruitment. For EBNA3C-mediated gene repression a two-step mechanism was observed, in which activation-associated histone marks were removed, resulting in the extensive loss of gene expression, before repressive marks were deposited by simultaneous, but spatially distinct recruitment of PRC1 (BMI1) and PRC2 (SUZ12). For EBNA3C-mediated gene activation, increases in activation-associated histone marks were detected that correlated well with recruited histone acetyltransferase p300 and PRC1 (BMI1) and increased levels of gene expression. Most surprisingly and common to both EBNA3C-mediated gene repression and activation was the dynamic recruitment and/or stabilisation of RBPJ/EBNA3C complexes on chromatin only in the presence of active EBNA3C. The functional importance of the interaction between EBNA3C and RBPJ became apparent in transient reporter assays and was confirmed in primary B cell infection using an EBV recombinant encoding a RBPJ binding mutant EBNA3C that failed to regulate all four genes.

Further studies are needed in order to elucidate the exact molecular mechanisms underlying EBNA3C-mediated gene regulation, to determine whether the findings presented here apply to all genes regulated by EBNA3C and to discover factors that determine whether a locus is repressed or activated by EBNA3C. Extensive analysis is also required to identify the role that EBNA3C-regulated genes play for EBV biology (*e.g. in vivo* role of *ADAM28* and *ADAMDEC1* repression and AID activation) and to determine the biological consequences for the EBV-infected B cells (*e.g.* activation of AID).

6 References

Abate, F., M.R. Ambrosio, L. Mundo, M.A. Laginestra, F. Fuligni, M. Rossi, S. Zairis, S. Gazaneo, G. De Falco, S. Lazzi, C. Bellan, B.J. Rocca, T. Amato, E. Marasco, M. Etebari, M. Ogwang, V. Calbi, I. Ndede, K. Patel, D. Chumba, P.P. Piccaluga, S. Pileri, L. Leoncini, and R. Rabadan. 2015. Distinct Viral and Mutational Spectrum of Endemic Burkitt Lymphoma. *PLoS Pathog*. 11:e1005158. doi:10.1371/journal.ppat.1005158.

Abbot, S.D., M. Rowe, K. Cadwallader, A. Ricksten, J. Gordon, F. Wang, L. Rymo, and A.B. Rickinson. 1990. Epstein-Barr virus nuclear antigen 2 induces expression of the virus-encoded latent membrane protein. *Journal of Virology*. 64:2126–2134.

Abruzzo, L.V., L.L. Barron, K. Anderson, R.J. Newman, W.G. Wierda, S. O'brien, A. Ferrajoli, M. Luthra, S. Talwalkar, R. Luthra, D. Jones, M.J. Keating, and K.R. Coombes. 2007. Identification and validation of biomarkers of IgV(H) mutation status in chronic lymphocytic leukemia using microfluidics quantitative real-time polymerase chain reaction technology. *J Mol Diagn*. 9:546–555. doi:10.2353/jmoldx.2007.070001.

Ahuja, R., R. Pinyol, N. Reichenbach, L. Custer, J. Klingensmith, M.M. Kessels, and B. Qualmann. 2007. Cordon-bleu is an actin nucleation factor and controls neuronal morphology. *Cell*. 131:337–350. doi:10.1016/j.cell.2007.08.030.

Albrechtsen, A., N. Grarup, Y. Li, T. Sparsø, G. Tian, H. Cao, T. Jiang, S.Y. Kim, T. Korneliussen, Q. Li, C. Nie, R. Wu, L. Skotte, A.P. Morris, C. Ladenvall, S. Cauchi, A. Stančáková, G. Andersen, A. Astrup, K. Banasik, A.J. Bennett, L. Bolund, G. Charpentier, Y. Chen, J.M. Dekker, A.S.F. Doney, M. Dorkhan, T. Forsen, T.M. Frayling, C.J. Groves, Y. Gui, G. Hallmans, A.T. Hattersley, K. He, G.A. Hitman, J. Holmkvist, S. Huang, H. Jiang, X. Jin, J.M. Justesen, K. Kristiansen, J. Kuusisto, M. Lajer, O. Lantieri, W. Li, H. Liang, Q. Liao, X. Liu, T. Ma, X. Ma, M.P. Manijak, M. Marre, J. Mokrosiński, A.D. Morris, B. Mu, A.A. Nielsen, G. Nijpels, P. Nilsson, C.N.A. Palmer, N.W. Rayner, F. Renström, R. Ribel-Madsen, N. Robertson, O. Rolandsson, P. Rossing, T.W. Schwartz, D.E.S.I.R. Study Group, P.E. Slagboom, M. Sterner, DIAGRAM Consortium, M. Tang, L. Tarnow, T. Tuomi, E. van't Riet, N. van Leeuwen, T.V. Varga, M.A. Vestmar, M. Walker, B. Wang, Y. Wang, H. Wu, F. Xi, L. Yengo, C. Yu, X. Zhang, J. Zhang, Q. Zhang, W. Zhang, H. Zheng, Y. Zhou, D. Altshuler, L.M. 't Hart, P.W. Franks, B. Balkau, P. Froguel, M.I. McCarthy, M. Laakso, L. Groop, et al. 2013. Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. *Diabetologia*. 56:298–310. doi:10.1007/s00125-012-2756-1.

Alexandrov, L.B., S. Nik-Zainal, D.C. Wedge, S.A.J.R. Aparicio, S. Behjati, A.V. Biankin, G.R. Bignell, N. Bolli, A. Borg, A.-L. Børresen-Dale, S. Boyault, B. Burkhardt, A.P. Butler, C. Caldas, H.R. Davies, C. Desmedt, R. Eils, J.E. Eyfjörd, J.A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinski, M. Imielinsk, N. Jäger, D.T.W. Jones, D. Jones, S. Knappskog, M. Kool, S.R. Lakhani, C. López-Otín, S. Martin, N.C. Munshi, H. Nakamura, P.A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J.V. Pearson, X.S. Puente, K. Raine, M. Ramakrishna, A.L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T.N. Schumacher, P.N. Span, J.W. Teague, Y. Totoki, A.N.J. Tutt, R. Valdés-Mas, M.M. van Buuren, L. van 't Veer, A. Vincent-Salomon, N. Waddell, L.R. Yates, Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, ICGC PedBrain, J. Zucman-Rossi, P.A. Futreal, U. McDermott, P. Lichter, M. Meyerson, S.M. Grimmond, R. Siebert, E. Campo, T. Shibata, S.M. Pfister, P.J. Campbell, and M.R. Stratton. 2013. Signatures of mutational processes in human cancer. *Nature*. 500:415–421. doi:10.1038/nature12477.

Allday, M.J. 2009. How does Epstein-Barr virus (EBV) complement the activation of Myc in the pathogenesis of Burkitt's lymphoma? *Semin. Cancer Biol.* 19:366–376. doi:10.1016/j.semcancer.2009.07.007.

Allday, M.J., D.H. Crawford, and B.E. Griffin. 1989. Epstein-Barr virus latent gene expression during the initiation of B cell immortalization. *J. Gen. Virol.* 70 (Pt 7):1755–1764.

Allday, M.J., D.H. Crawford, and J.A. Thomas. 1993. Epstein-Barr virus (EBV) nuclear antigen 6 induces expression of the EBV latent membrane protein and an activated phenotype in Raji cells. *J. Gen. Virol.* 74 (Pt 3):361–369.

Allday, M.J., Q. Bazot, and R.E. White. 2015. The EBNA3 Family: Two Oncoproteins and a Tumour Suppressor that Are Central to the Biology of EBV in B Cells. *Curr. Top. Microbiol. Immunol.* 391:61–117. doi:10.1007/978-3-319-22834-1_3.

Allen, B.L., and D.J. Taatjes. 2015. The Mediator complex: a central integrator of transcription. *Nat. Rev. Mol. Cell Biol.* 16:155–166. doi:10.1038/nrm3951.

Allen, C.D.C., K.M. Ansel, C. Low, R. Lesley, H. Tamamura, N. Fujii, and J.G. Cyster. 2004. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat. Immunol.* 5:943–952. doi:10.1038/ni1100.

Allen, C.D.C., T. Okada, and J.G. Cyster. 2007a. Germinal-center organization and cellular dynamics. *Immunity*. 27:190–202. doi:10.1016/j.immuni.2007.07.009.

Allen, C.D.C., T. Okada, H.L. Tang, and J.G. Cyster. 2007b. Imaging of germinal center selection events during affinity maturation. *Science*. 315:528–531. doi:10.1126/science.1136736.

Alt, F.W., and D. Baltimore. 1982. Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D-JH fusions. *Proc. Natl. Acad. Sci. U.S.A.* 79:4118–4122.

Anderson, L.J., and R. Longnecker. 2008. An auto-regulatory loop for EBV LMP2A involves activation of Notch. *Virology*. 371:257–266. doi:10.1016/j.virol.2007.10.009.

Anderton, E., J. Yee, P. Smith, T. Crook, R.E. White, and M.J. Allday. 2007. Two Epstein–Barr virus (EBV) oncoproteins cooperate to repress expression of the proapoptotic tumour-suppressor Bim: clues to the pathogenesis of Burkitt's lymphoma. *Oncogene*. 27:421–433. doi:10.1038/sj.onc.1210668.

Aoufouchi, S., A. Faili, C. Zober, O. D'Orlando, S. Weller, J.-C. Weill, and C.-A. Reynaud. 2008. Proteasomal degradation restricts the nuclear lifespan of AID. *J. Exp. Med.* 205:1357–1368. doi:10.1084/jem.20070950.

Aranda, S., G. Mas, and L. Di Croce. 2015. Regulation of gene transcription by Polycomb proteins. *Sci Adv*. 1:e1500737–e1500737. doi:10.1126/sciadv.1500737.

Armitage, R.J., W.C. Fanslow, L. Strockbine, T.A. Sato, K.N. Clifford, B.M. Macduff, D.M. Anderson, S.D. Gimpel, T. Davis-Smith, and C.R. Maliszewski. 1992. Molecular and biological characterization of a murine ligand for CD40. *Nature*. 357:80–82. doi:10.1038/357080a0.

Babcock, G.J., D. Hochberg, and A.D. Thorley-Lawson. 2000. The expression pattern of

Epstein-Barr virus latent genes in vivo is dependent upon the differentiation stage of the infected B cell. *Immunity*. 13:497–506.

Babcock, G.J., L.L. Decker, M. Volk, and D.A. Thorley-Lawson. 1998. EBV persistence in memory B cells in vivo. *Immunity*. 9:395–404.

Bain, M., R.J. Watson, P.J. Farrell, and M.J. Allday. 1996. Epstein-Barr virus nuclear antigen 3C is a powerful repressor of transcription when tethered to DNA. *Journal of Virology*. 70:2481–2489.

Banerjee, S., J. Lu, Q. Cai, A. Saha, H.C. Jha, R.K. Dzeng, and E.S. Robertson. 2013. The EBV Latent Antigen 3C Inhibits Apoptosis through Targeted Regulation of Interferon Regulatory Factors 4 and 8. *PLoS Pathog*. 9:e1003314. doi:10.1371/journal.ppat.1003314.

Bannister, A.J., and T. Kouzarides. 2011. Regulation of chromatin by histone modifications. *Cell Res.* 21:381–395. doi:10.1038/cr.2011.22.

Bardwell, P.D., C.J. Woo, K. Wei, Z. Li, A. Martin, S.Z. Sack, T. Parris, W. Edelmann, and M.D. Scharff. 2004. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. *Nat. Immunol.* 5:224–229. doi:10.1038/ni1031.

Basso, K., and R. Dalla-Favera. 2012. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol. Rev.* 247:172–183. doi:10.1111/j.1600-065X.2012.01112.x.

Basso, K., and R. Dalla-Favera. 2015. Germinal centres and B cell lymphomagenesis. *Nat. Rev. Immunol.* 15:172–184. doi:10.1038/nri3814.

Basso, K., C. Schneider, Q. Shen, A.B. Holmes, M. Setty, C. Leslie, and R. Dalla-Favera. 2012. BCL6 positively regulates AID and germinal center gene expression via repression of miR-155. *J. Exp. Med.* 209:2455–2465. doi:10.1084/jem.20121387.

Basu, U., A. Franklin, and F.W. Alt. 2009. Post-translational regulation of activation-induced cytidine deaminase. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 364:667–673. doi:10.1098/rstb.2008.0194.

Basu, U., F.-L. Meng, C. Keim, V. Grinstein, E. Pefanis, J. Eccleston, T. Zhang, D. Myers, C.R. Wasserman, D.R. Wesemann, K. Januszyk, R.I. Gregory, H. Deng, C.D. Lima, and F.W. Alt. 2011. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell*. 144:353–363. doi:10.1016/j.cell.2011.01.001.

Basu, U., J. Chaudhuri, C. Alpert, S. Dutt, S. Ranganath, G. Li, J.P. Schrum, J.P. Manis, and F.W. Alt. 2005. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature*. 438:508–511. doi:10.1038/nature04255.

Bates, E., W. Fridman, and C. Mueller. 2002. The ADAMDEC1 (decysin) gene structure: evolution by duplication in a metalloprotease gene cluster on Chromosome 8p12. *Immunogenetics*. 54:96–105. doi:10.1007/s00251-002-0430-3.

Bazot, Q., K. Paschos, L. Skalska, J.S. Kalchschmidt, G.A. Parker, and M.J. Allday. 2015. Epstein-Barr Virus Proteins EBNA3A and EBNA3C Together Induce Expression of the Oncogenic MicroRNA Cluster miR-221/miR-222 and Ablate Expression of Its Target p57KIP2. *PLoS Pathog*. 11:e1005031. doi:10.1371/journal.ppat.1005031.

Becker, P.B., and J.L. Workman. 2013. Nucleosome remodeling and epigenetics. *Cold Spring Harb Perspect Biol*. 5:a017905–a017905. doi:10.1101/cshperspect.a017905.

Bekerman, E., D. Jeon, M. Ardolino, and L. Coscoy. 2013. A Role for Host Activation-Induced Cytidine Deaminase in Innate Immune Defense against KSHV. *PLoS Pathog*. 9:e1003748. doi:10.1371/journal.ppat.1003748.

Bell, A.C., and G. Felsenfeld. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*. 405:482–485. doi:10.1038/35013100.

Belmont, A.S. 2006. Mitotic chromosome structure and condensation. *Curr. Opin. Cell Biol.* 18:632–638. doi:10.1016/j.ceb.2006.09.007.

Ben-Bassat, H., N. Goldblum, S. Mitrani, T. Goldblum, J.M. Yoffey, M.M. Cohen, Z. Bentwich, B. Ramot, E. Klein, and G. Klein. 1977. Establishment in continuous culture of a new type of lymphocyte from a “Burkitt like” malignant lymphoma (line D.G.-75). *Int. J. Cancer*. 19:27–33.

Berger, S.L., T. Kouzarides, R. Shiekhattar, and A. Shilatifard. 2009. An operational definition of epigenetics. *Genes Dev.* 23:781–783. doi:10.1101/gad.1787609.

Bian, Q., and A.S. Belmont. 2010. BAC TG-EMBED: one-step method for high-level, copy-number-dependent, position-independent transgene expression. *Nucleic Acids Research*. 38:e127–e127. doi:10.1093/nar/gkq178.

Blackledge, N.P., A.M. Farcas, T. Kondo, H.W. King, J.F. McGouran, L.L.P. Hanssen, S. Ito, S. Cooper, K. Kondo, Y. Koseki, T. Ishikura, H.K. Long, T.W. Sheahan, N. Brockdorff, B.M. Kessler, H. Koseki, and R.J. Klose. 2014. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell*. 157:1445–1459. doi:10.1016/j.cell.2014.05.004.

Blum, J.S., P.A. Wearsch, and P. Cresswell. 2013. Pathways of antigen processing. *Annu. Rev. Immunol.* 31:443–473. doi:10.1146/annurev-immunol-032712-095910.

Boccellato, F., E. Anastasiadou, P. Rosato, B. Kempkes, L. Frati, A. Faggioni, and P. Trivedi. 2007. EBNA2 interferes with the germinal center phenotype by downregulating BCL6 and TCL1 in non-Hodgkin's lymphoma cells. *Journal of Virology*. 81:2274–2282. doi:10.1128/JVI.01822-06.

Bourillot, P.Y., L. Waltzer, A. Sergeant, and E. Manet. 1998. Transcriptional repression by the Epstein-Barr virus EBNA3A protein tethered to DNA does not require RBP-Jkappa. *J. Gen. Virol.* 79 (Pt 2):363–370.

Boyle, S., S. Gilchrist, J.M. Bridger, N.L. Mahy, J.A. Ellis, and W.A. Bickmore. 2001. The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Hum. Mol. Genet.* 10:211–219.

Brady, G., H.J. Whiteman, L.C. Spender, and P.J. Farrell. 2009. Downregulation of RUNX1 by RUNX3 requires the RUNX3 VWRPY sequence and is essential for Epstein-Barr virus-driven B-cell proliferation. *Journal of Virology*. 83:6909–6916. doi:10.1128/JVI.00216-09.

Branco, M.R., and A. Pombo. 2006. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol.* 4:e138. doi:10.1371/journal.pbio.0040138.

Bray, S.J. 2006. Notch signalling: a simple pathway becomes complex. *Nat. Rev. Mol. Cell Biol.* 7:678–689. doi:10.1038/nrm2009.

Bridges, L.C. 2001. The Lymphocyte Metalloprotease MDC-L (ADAM 28) Is a Ligand for the Integrin alpha 4beta 1. *Journal of Biological Chemistry*. 277:3784–3792. doi:10.1074/jbc.M109538200.

Bridges, L.C., D. Sheppard, and R.D. Bowditch. 2005. ADAM disintegrin-like domain recognition by the lymphocyte integrins alpha4beta1 and alpha4beta7. *Biochem. J.* 387:101–108. doi:10.1042/BJ20041444.

Bridges, L.C., K.R. Hanson, P.H. Tani, T. Mather, and R.D. Bowditch. 2003. Integrin alpha4beta1-dependent adhesion to ADAM 28 (MDC-L) requires an extended surface of the disintegrin domain. *Biochemistry*. 42:3734–3741. doi:10.1021/bi026871y.

Cai, D.-C., H. Fonteijn, T. Guadalupe, M. Zwiers, K. Wittfeld, A. Teumer, M. Hoogman, A. Arias-Vásquez, Y. Yang, J. Buitelaar, G. Fernández, H.G. Brunner, H. van Bokhoven, B. Franke, K. Hegenscheid, G. Homuth, S.E. Fisher, H.J. Grabe, C. Francks, and P. Hagoort. 2014. A genome-wide search for quantitative trait loci affecting the cortical surface area and thickness of Heschl's gyrus. *Genes Brain Behav.* 13:675–685. doi:10.1111/gbb.12157.

Calado, D.P., Y. Sasaki, S.A. Godinho, A. Pellerin, K. Köchert, B.P. Sleckman, I.M. de Alborán, M. Janz, S. Rodig, and K. Rajewsky. 2012. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat. Immunol.* 13:1092–1100. doi:10.1038/ni.2418.

Calderwood, M.A., K. Venkatesan, L. Xing, M.R. Chase, A. Vazquez, A.M. Holthaus, A.E. Ewence, N. Li, T. Hirozane-Kishikawa, D.E. Hill, M. Vidal, E. Kieff, and E. Johannsen. 2007. Epstein-Barr virus and virus human protein interaction maps. *Proc. Natl. Acad. Sci. U.S.A.* 104:7606–7611. doi:10.1073/pnas.0702332104.

Calderwood, M.A., S. Lee, A.M. Holthaus, S.C. Blacklow, E. Kieff, and E. Johannsen. 2011. Epstein-Barr virus nuclear protein 3C binds to the N-terminal (NTD) and beta trefoil domains (BTD) of RBP/CSL; only the NTD interaction is essential for lymphoblastoid cell growth. *Virology*. 414:19–25. doi:10.1016/j.virol.2011.02.018.

Campos, E.I., and D. Reinberg. 2009. Histones: annotating chromatin. *Annu. Rev. Genet.* 43:559–599. doi:10.1146/annurev.genet.032608.103928.

Cao, R., L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R.S. Jones, and Y. Zhang. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*. 298:1039–1043. doi:10.1126/science.1076997.

Cao, R., Y.-l. Tsukada, and Y. Zhang. 2005. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell*. 20:845–854. doi:10.1016/j.molcel.2005.12.002.

Caron, G., S. Le Gallou, T. Lamy, K. Tarte, and T. Fest. 2009. CXCR4 expression functionally discriminates centroblasts versus centrocytes within human germinal center B cells. *J. Immunol.* 182:7595–7602. doi:10.4049/jimmunol.0804272.

Carroll, E.A., D. Gerrelli, S. Gasca, E. Berg, D.R. Beier, A.J. Copp, and J. Klingensmith. 2003. Cordon-bleu is a conserved gene involved in neural tube formation. *Dev. Biol.* 262:16–31. doi:10.1038/nrn1261.

Casellas, R., U. Basu, W.T. Yewdell, J. Chaudhuri, D.F. Robbiani, and J.M. Di Noia. 2016. Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. *Nat. Rev. Immunol.* 16:164–176. doi:10.1038/nri.2016.2.

Castel, D., P. Mourikis, S.J.J. Bartels, A.B. Brinkman, S. Tajbakhsh, and H.G. Stunnenberg. 2013. Dynamic binding of RBPJ is determined by Notch signaling status. *Genes Dev.* 27:1059–1071. doi:10.1101/gad.211912.112.

Cattoretti, G., C.C. Chang, K. Cechova, J. Zhang, B.H. Ye, B. Falini, D.C. Louie, K. Offit, R.S. Chaganti, and R. Dalla-Favera. 1995. BCL-6 protein is expressed in germinal-center B cells. *Blood*. 86:45–53.

Cen, O., and R. Longnecker. 2015. Latent Membrane Protein 2 (LMP2). *Curr. Top. Microbiol. Immunol.* 391:151–180. doi:10.1007/978-3-319-22834-1_5.

Chaudhuri, J., C. Khuong, and F.W. Alt. 2004. Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. *Nature*. 430:992–998. doi:10.1038/nature02821.

Chaudhuri, J., M. Tian, C. Khuong, K. Chua, E. Pinaud, and F.W. Alt. 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature*. 422:726–730. doi:10.1038/nature01574.

Cheutin, T., and G. Cavalli. 2014. Polycomb silencing: from linear chromatin domains to 3D chromosome folding. *Curr. Opin. Genet. Dev.* 25:30–37. doi:10.1016/j.gde.2013.11.016.

Chezar, I., L. Lobel-Lavi, M. Steinitz, and R. Laskov. 2008. Ongoing somatic hypermutation of the rearranged VH but not of the V-lambda gene in EBV-transformed rheumatoid factor-producing lymphoblastoid cell line. *Mol. Immunol.* 46:80–90. doi:10.1016/j.molimm.2008.07.002.

Cho, C. 2012. Testicular and epididymal ADAMs: expression and function during fertilization. *Nat Rev Urol*. 9:550–560. doi:10.1038/nrurol.2012.167.

Christensen, J., K. Agger, P.A.C. Cloos, D. Pasini, S. Rose, L. Sennels, J. Rappsilber, K.H. Hansen, A.E. Salcini, and K. Helin. 2007. RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. *Cell*. 128:1063–1076. doi:10.1016/j.cell.2007.02.003.

Christian, L.M. 2012. The ADAM family: Insights into Notch proteolysis. *Fly (Austin)*. 6:30–34. doi:10.4161/fly.18823.

Cohen, J.I., and E. Kieff. 1991. An Epstein-Barr virus nuclear protein 2 domain essential for transformation is a direct transcriptional activator. *Journal of Virology*. 65:5880–5885.

Cohen, J.I., F. Wang, J. Mannick, and E. Kieff. 1989. Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation. *Proc. Natl. Acad. Sci. U.S.A.* 86:9558–9562.

Compe, E., and J.-M. Egly. 2012. TFIIH: when transcription met DNA repair. *Nat. Rev. Mol. Cell Biol.* 13:343–354. doi:10.1038/nrm3350.

Constantinescu, A., and M.S. Schlissel. 1997. Changes in locus-specific V(D)J recombinase activity induced by immunoglobulin gene products during B cell development. *J. Exp. Med.* 185:609–620.

Cooper, S., M. Dienstbier, R. Hassan, L. Schermelleh, J. Sharif, N.P. Blackledge, V. De Marco, S. Elderkin, H. Koseki, R. Klose, A. Heger, and N. Brockdorff. 2014. Targeting polycomb to pericentric heterochromatin in embryonic stem cells reveals a role for H2AK119u1 in PRC2 recruitment. *Cell Rep*. 7:1456–1470. doi:10.1016/j.celrep.2014.04.012.

Cordier, M., A. Calender, M. Billaud, U. Zimber, G. Rousselet, O. Pavlish, J. Banchereau, T. Tursz, G. Bornkamm, and G.M. Lenoir. 1990. Stable transfection of Epstein-Barr virus (EBV) nuclear antigen 2 in lymphoma cells containing the EBV P3HR1 genome induces expression of B-cell activation molecules CD21 and CD23. *Journal of Virology*. 64:1002–1013.

Cotter, M.A., and E.S. Robertson. 2000. Modulation of histone acetyltransferase activity through interaction of epstein-barr nuclear antigen 3C with prothymosin alpha. *Mol. Cell. Biol.* 20:5722–5735.

Cremer, M., J. von Hase, T. Volm, A. Brero, G. Kreth, J. Walter, C. Fischer, I. Solovei, C. Cremer, and T. Cremer. 2001. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res.* 9:541–567.

Creppe, C., A. Palau, R. Malinverni, V. Valero, and M. Buschbeck. 2014. A Cbx8-containing polycomb complex facilitates the transition to gene activation during ES cell differentiation. *PLoS Genet.* 10:e1004851. doi:10.1371/journal.pgen.1004851.

Crotty, S. 2015. A brief history of T cell help to B cells. *Nat. Rev. Immunol.* 15:185–189. doi:10.1038/nri3803.

Crouch, E.E., Z. Li, M. Takizawa, S. Fichtner-Feigl, P. Gourzi, C. Montaño, L. Feigenbaum, P. Wilson, S. Janz, F.N. Papavasiliou, and R. Casellas. 2007. Regulation of AID expression in the immune response. *J. Exp. Med.* 204:1145–1156. doi:10.1084/jem.20061952.

Crouser, E.D., D.A. Culver, K.S. Knox, M.W. Julian, G. Shao, S. Abraham, S. Liyanarachchi, J.E. Macre, M.D. Wewers, M.A. Gavrilin, P. Ross, A. Abbas, and C. Eng. 2009. Gene Expression Profiling Identifies MMP-12 and ADAMDEC1 as Potential Pathogenic Mediators of Pulmonary Sarcoidosis. *American Journal of Respiratory and Critical Care Medicine*. 179:929–938. doi:10.1164/rccm.200803-490OC.

Cutrona, G., M. Ulivi, F. Fais, S. Roncella, and M. Ferrarini. 1995. Transfection of the c-myc oncogene into normal Epstein-Barr virus-harboring B cells results in new phenotypic and functional features resembling those of Burkitt lymphoma cells and normal centroblasts. *J. Exp. Med.* 181:699–711.

Czermin, B., R. Melfi, D. McCabe, V. Seitz, A. Imhof, and V. Pirrotta. 2002. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*. 111:185–196.

Dal Porto, J.M., S.B. Gauld, K.T. Merrell, D. Mills, A.E. Pugh-Bernard, and J. Cambier. 2004. B cell antigen receptor signaling 101. *Mol. Immunol.* 41:599–613. doi:10.1016/j.molimm.2004.04.008.

Dang, C.V., K.A. O'donnell, and T. Juopperi. 2005. The great MYC escape in tumorigenesis. *Cancer Cell*. 8:177–178. doi:10.1016/j.ccr.2005.08.005.

Davies, M.L., S. Xu, J. Lyons-Weiler, A. Rosendorff, S.A. Webber, L.R. Wasil, D. Metes, and

D.T. Rowe. 2010. Cellular factors associated with latency and spontaneous Epstein-Barr virus reactivation in B-lymphoblastoid cell lines. *Virology*. 400:53–67. doi:10.1016/j.virol.2010.01.002.

de Napoles, M., J.E. Mermoud, R. Wakao, Y.A. Tang, M. Endoh, R. Appanah, T.B. Nesterova, J. Silva, A.P. Otte, M. Vidal, H. Koseki, and N. Brockdorff. 2004. Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev. Cell*. 7:663–676. doi:10.1016/j.devcel.2004.10.005.

De Silva, N.S., and U. Klein. 2015. Dynamics of B cells in germinal centres. *Nat. Rev. Immunol.* 15:137–148. doi:10.1038/nri3804.

de Yebenes, V.G., L. Belver, D.G. Pisano, S. González, A. Villasante, C. Croce, L. He, and A.R. Ramiro. 2008. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J. Exp. Med.* 205:2199–2206. doi:10.1084/jem.20080579.

Dedeoglu, F., B. Horwitz, J. Chaudhuri, F.W. Alt, and R.S. Geha. 2004. Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFkappaB. *Int. Immunol.* 16:395–404.

Dekker, J., K. Rippe, M. Dekker, and N. Kleckner. 2002. Capturing chromosome conformation. *Science*. 295:1306–1311. doi:10.1126/science.1067799.

Delbarre, E., B.M. Jacobsen, A.H. Reiner, A.L. Sørensen, T. Küntziger, and P. Collas. 2010. Chromatin environment of histone variant H3.3 revealed by quantitative imaging and genome-scale chromatin and DNA immunoprecipitation. *Mol. Biol. Cell*. 21:1872–1884. doi:10.1091/mbc.E09-09-0839.

Delbos, F., S. Aoufouchi, A. Faili, J.-C. Weill, and C.-A. Reynaud. 2007. DNA polymerase eta is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. *J. Exp. Med.* 204:17–23. doi:10.1084/jem.20062131.

Delecluse, H.J., T. Hilsendegen, D. Pich, R. Zeidler, and W. Hammerschmidt. 1998. Propagation and recovery of intact, infectious Epstein-Barr virus from prokaryotic to human cells. *Proc. Natl. Acad. Sci. U.S.A.* 95:8245–8250.

Denholtz, M., G. Bonora, C. Chronis, E. Splinter, W. De Laat, J. Ernst, M. Pellegrini, and K. Plath. 2013. Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell*. 13:602–616. doi:10.1016/j.stem.2013.08.013.

Dent, A.L., A.L. Shaffer, X. Yu, D. Allman, and L.M. Staudt. 1997. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science*. 276:589–592.

Desmarchelier, C., J.-C. Martin, R. Planells, M. Gastaldi, M. Nowicki, A. Goncalves, R. Valéro, D. Lairon, and P. Borel. 2014. The postprandial chylomicron triacylglycerol response to dietary fat in healthy male adults is significantly explained by a combination of single nucleotide polymorphisms in genes involved in triacylglycerol metabolism. *J. Clin. Endocrinol. Metab.* 99:E484–8. doi:10.1210/jc.2013-3962.

Déjardin, J., and R.E. Kingston. 2009. Purification of proteins associated with specific genomic Loci. *Cell*. 136:175–186. doi:10.1016/j.cell.2008.11.045.

Di Noia, J.M., and M.S. Neuberger. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* 76:1–22.

doi:10.1146/annurev.biochem.76.061705.090740.

Dick, D.M., J. Meyers, F. Aliev, J. Nurnberger Jr., J. Kramer, S. Kuperman, B. Porjesz, J. Tischfield, H.J. Edenberg, T. Foroud, M. Schuckit, A. Goate, V. Hesselbrock, and L. Bierut. 2010. Evidence for genes on chromosome 2 contributing to alcohol dependence with conduct disorder and suicide attempts. *Am. J. Med. Genet.* 9999B:n/a–n/a. doi:10.1002/ajmg.b.31089.

Dirmeier, U., B. Neuhierl, E. Kilger, G. Reisbach, M.L. Sandberg, and W. Hammerschmidt. 2003. Latent membrane protein 1 is critical for efficient growth transformation of human B cells by epstein-barr virus. *Cancer Research*. 63:2982–2989.

Dixon, J.R., S. Selvaraj, F. Yue, A. Kim, Y. Li, Y. Shen, M. Hu, J.S. Liu, and B. Ren. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature*. 485:376–380. doi:10.1038/nature11082.

Dominguez-Sola, D., G.D. Victora, C.Y. Ying, R.T. Phan, M. Saito, M.C. Nussenzweig, and R. Dalla-Favera. 2012. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat. Immunol.* 13:1083–1091. doi:10.1038/ni.2428.

Dorsett, Y., K.M. McBride, M. Jankovic, A. Gazumyan, T.-H. Thai, D.F. Robbiani, M. Di Virgilio, B. Reina-San-Martin, G. Heidkamp, T.A. Schwickert, T. Eisenreich, K. Rajewsky, and M.C. Nussenzweig. 2008. MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity*. 28:630–638. doi:10.1016/j.immuni.2008.04.002.

Doudna, J.A., and E. Charpentier. 2014. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science*. 346:1258096–1258096. doi:10.1126/science.1258096.

Dowen, J.M., Z.P. Fan, D. Hnisz, G. Ren, B.J. Abraham, L.N. Zhang, A.S. Weintraub, J. Schuijers, T.I. Lee, K. Zhao, and R.A. Young. 2014. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell*. 159:374–387. doi:10.1016/j.cell.2014.09.030.

Edwards, D.R., M.M. Handsley, and C.J. Pennington. 2008. The ADAM metalloproteinases. *Mol. Aspects Med.* 29:258–289. doi:10.1016/j.mam.2008.08.001.

Egle, A., A.W. Harris, P. Bouillet, and S. Cory. 2004. Bim is a suppressor of Myc-induced mouse B cell leukemia. *Proc. Natl. Acad. Sci. U.S.A.* 101:6164–6169. doi:10.1073/pnas.0401471101.

Eick, D., and M. Geyer. 2013. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem. Rev.* 113:8456–8490. doi:10.1021/cr400071f.

Elderkin, S., G.N. Maertens, M. Endoh, D.L. Mallery, N. Morrice, H. Koseki, G. Peters, N. Brockdorff, and K. Hiom. 2007. A phosphorylated form of Mel-18 targets the Ring1B histone H2A ubiquitin ligase to chromatin. *Mol. Cell*. 28:107–120. doi:10.1016/j.molcel.2007.08.009.

Engels, N., G. Yigit, C.H. Emmerich, D. Czesnik, D. Schild, and J. Wienands. 2012. Epstein-Barr virus LMP2A signaling in statu nascendi mimics a B cell antigen receptor-like activation signal. *Cell Commun. Signal*. 10:9. doi:10.1186/1478-811X-10-9.

Epeldegui, M., Y.P. Hung, A. McQuay, R.F. Ambinder, and O. Martínez-Maza. 2007.

Infection of human B cells with Epstein-Barr virus results in the expression of somatic hypermutation-inducing molecules and in the accrual of oncogene mutations. *Mol. Immunol.* 44:934–942. doi:10.1016/j.molimm.2006.03.018.

Epstein, M.A., B.G. Achong, Y.M. Barr, B. Zajac, G. Henle, and W. Henle. 1966. Morphological and virological investigations on cultured Burkitt tumor lymphoblasts (strain Raji). *J. Natl. Cancer Inst.* 37:547–559.

Eskeland, R., M. Leeb, G.R. Grimes, C. Kress, S. Boyle, D. Sproul, N. Gilbert, Y. Fan, A.I. Skoultchi, A. Wutz, and W.A. Bickmore. 2010. Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Mol. Cell*. 38:452–464. doi:10.1016/j.molcel.2010.02.032.

Farcas, A.M., N.P. Blackledge, I. Sudbery, H.K. Long, J.F. McGouran, N.R. Rose, S. Lee, D. Sims, A. Cerase, T.W. Sheahan, H. Koseki, N. Brockdorff, C.P. Ponting, B.M. Kessler, and R.J. Klose. 2012. KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *Elife*. 1:e00205. doi:10.7554/eLife.00205.

Farré, D., R. Roset, M. Huerta, J.E. Adsuara, L. Roselló, M.M. Albà, and X. Messeguer. 2003. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. *Nucleic Acids Research*. 31:3651–3653.

Faumont, N., S. Durand-Panteix, M. Schlee, S. Grömminger, M. Schuhmacher, M. Hölzel, G. Laux, R. Mailhammer, A. Rosenwald, L.M. Staudt, G.W. Bornkamm, and J. Feuillard. 2009. c-Myc and Rel/NF-kappaB are the two master transcriptional systems activated in the latency III program of Epstein-Barr virus-immortalized B cells. *Journal of Virology*. 83:5014–5027. doi:10.1128/JVI.02264-08.

Fåhraeus, R., A. Jansson, A. Ricksten, A. Sjöblom, and L. Rymo. 1990. Epstein-Barr virus-encoded nuclear antigen 2 activates the viral latent membrane protein promoter by modulating the activity of a negative regulatory element. *Proc. Natl. Acad. Sci. U.S.A.* 87:7390–7394.

Ferrari, K.J., A. Scelfo, S. Jammula, A. Cuomo, I. Barozzi, A. Stützer, W. Fischle, T. Bonaldi, and D. Pasini. 2013. Polycomb-Dependent H3K27me1 and H3K27me2 Regulate Active Transcription and Enhancer Fidelity. *Mol. Cell*. doi:10.1016/j.molcel.2013.10.030.

Fingeroth, J.D., J.J. Weis, T.F. Tedder, J.L. Strominger, P.A. Biro, and D.T. Fearon. 1984. Epstein-Barr virus receptor of human B lymphocytes is the C3d receptor CR2. *Proc. Natl. Acad. Sci. U.S.A.* 81:4510–4514.

Finlan, L.E., D. Sproul, I. Thomson, S. Boyle, E. Kerr, P. Perry, B. Ylstra, J.R. Chubb, and W.A. Bickmore. 2008. Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet.* 4:e1000039. doi:10.1371/journal.pgen.1000039.

Fishel, R. 2015. Mismatch repair. *J. Biol. Chem.* 290:26395–26403. doi:10.1074/jbc.R115.660142.

Flenghi, L., B.H. Ye, M. Fizzotti, B. Bigerna, G. Cattoretti, S. Venturi, R. Pacini, S. Pileri, F. Lo Coco, and E. Pescarmona. 1995. A specific monoclonal antibody (PG-B6) detects expression of the BCL-6 protein in germinal center B cells. *Am. J. Pathol.* 147:405–411.

Fortini, M.E. 2002. Gamma-secretase-mediated proteolysis in cell-surface-receptor signalling. *Nat. Rev. Mol. Cell Biol.* 3:673–684. doi:10.1038/nrm910.

Fourie, A.M. 2003. Catalytic Activity of ADAM8, ADAM15, and MDC-L (ADAM28) on Synthetic Peptide Substrates and in Ectodomain Cleavage of CD23. *Journal of Biological Chemistry*. 278:30469–30477. doi:10.1074/jbc.M213157200.

Francis, N.J., R.E. Kingston, and C.L. Woodcock. 2004. Chromatin compaction by a polycomb group protein complex. *Science*. 306:1574–1577. doi:10.1126/science.1100576.

Frangini, A., M. Sjöberg, M. Roman-Trufero, G. Dharmalingam, V. Haberle, T. Bartke, B. Lenhard, M. Malumbres, M. Vidal, and N. Dillon. 2013. The aurora B kinase and the polycomb protein ring1B combine to regulate active promoters in quiescent lymphocytes. *Mol. Cell*. 51:647–661. doi:10.1016/j.molcel.2013.08.022.

Fukita, Y., H. Jacobs, and K. Rajewsky. 1998. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity*. 9:105–114.

Gahn, T.A., and B. Sugden. 1995. An EBNA-1-dependent enhancer acts from a distance of 10 kilobase pairs to increase expression of the Epstein-Barr virus LMP gene. *Journal of Virology*. 69:2633–2636.

Galler, G.R., C. Mundt, M. Parker, R. Pelanda, I.-L. Mårtensson, and T.H. Winkler. 2004. Surface mu heavy chain signals down-regulation of the V(D)J-recombinase machinery in the absence of surrogate light chain components. *J. Exp. Med.* 199:1523–1532. doi:10.1084/jem.20031523.

Gao, Z., J. Zhang, R. Bonasio, F. Strino, A. Sawai, F. Parisi, Y. Kluger, and D. Reinberg. 2012. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell*. 45:344–356. doi:10.1016/j.molcel.2012.01.002.

Gao, Z., P. Lee, J.M. Stafford, M. von Schimmelmann, A. Schaefer, and D. Reinberg. 2014. An AUTS2-Polycomb complex activates gene expression in the CNS. *Nature*. 516:349–354. doi:10.1038/nature13921.

Gazumyan, A., K. Timachova, G. Yuen, E. Siden, M. Di Virgilio, E.M. Woo, B.T. Chait, B. Reina-San-Martin, M.C. Nussenzweig, and K.M. McBride. 2011. Amino-terminal phosphorylation of activation-induced cytidine deaminase suppresses c-myc/IgH translocation. *Mol. Cell. Biol.* 31:442–449. doi:10.1128/MCB.00349-10.

Ghosh, D., and E. Kieff. 1990. cis-acting regulatory elements near the Epstein-Barr virus latent-infection membrane protein transcriptional start site. *Journal of Virology*. 64:1855–1858.

Gil, J., and G. Peters. 2006. Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. *Nat. Rev. Mol. Cell Biol.* 7:667–677. doi:10.1038/nrm1987.

Gil, Y., S. Levy-Nabot, M. Steinitz, and R. Laskov. 2007. Somatic mutations and activation-induced cytidine deaminase (AID) expression in established rheumatoid factor-producing lymphoblastoid cell line. *Mol. Immunol.* 44:494–505. doi:10.1016/j.molimm.2006.02.012.

Gires, O., U. Zimber-Strobl, R. Gonnella, M. Ueffing, G. Marschall, R. Zeidler, D. Pich, and W. Hammerschmidt. 1997. Latent membrane protein 1 of Epstein-Barr virus mimics a constitutively active receptor molecule. *EMBO J.* 16:6131–6140. doi:10.1093/emboj/16.20.6131.

Glazer, P.M., and W.C. Summers. 1985. Oncogene expression in isogenic, EBV-positive and -negative Burkitt lymphoma cell lines. *Intervirology*. 23:82–89.

Gonda, H., M. Sugai, Y. Nambu, T. Katakai, Y. Agata, K.J. Mori, Y. Yokota, and A. Shimizu. 2003. The balance between Pax5 and Id2 activities is the key to AID gene expression. *J. Exp. Med.* 198:1427–1437. doi:10.1084/jem.20030802.

Gordon, G.J., G.N. Rockwell, P.A. Godfrey, R.V. Jensen, J.N. Glickman, B.Y. Yeap, W.G. Richards, D.J. Sugarbaker, and R. Bueno. 2005. Validation of genomics-based prognostic tests in malignant pleural mesothelioma. *Clin. Cancer Res.* 11:4406–4414. doi:10.1158/1078-0432.CCR-04-2181.

Gordon, G.J., L. Dong, B.Y. Yeap, W.G. Richards, J.N. Glickman, H. Edenfield, M. Mani, R. Colquitt, G. Maulik, B. Van Oss, D.J. Sugarbaker, and R. Bueno. 2009. Four-Gene Expression Ratio Test for Survival in Patients Undergoing Surgery for Mesothelioma. *JNCI Journal of the National Cancer Institute*. 101:678–686. doi:10.1093/jnci/djp061.

Gordon, G.J., R.V. Jensen, L.-L. Hsiao, S.R. Gullans, J.E. Blumenstock, W.G. Richards, M.T. Jaklitsch, D.J. Sugarbaker, and R. Bueno. 2003. Using gene expression ratios to predict outcome among patients with mesothelioma. *J. Natl. Cancer Inst.* 95:598–605.

Gottschalk, S., C.M. Rooney, and H.E. Heslop. 2005. Post-transplant lymphoproliferative disorders. *Annu. Rev. Med.* 56:29–44. doi:10.1146/annurev.med.56.082103.104727.

Gourzi, P., T. Leonova, and F.N. Papavasiliou. 2007. Viral induction of AID is independent of the interferon and the Toll-like receptor signaling pathways but requires NF-kappaB. *J. Exp. Med.* 204:259–265. doi:10.1084/jem.20061801.

Grawunder, U., T.M. Leu, D.G. Schatz, A. Werner, A.G. Rolink, F. Melchers, and T.H. Winkler. 1995. Down-regulation of RAG1 and RAG2 gene expression in preB cells after functional immunoglobulin heavy chain rearrangement. *Immunity*. 3:601–608.

Grogan, E.A., W.P. Summers, S. Dowling, D. Shedd, L. Gradoville, and G. Miller. 1983. Two Epstein-Barr viral nuclear neoantigens distinguished by gene transfer, serology, and chromosome binding. *Proc. Natl. Acad. Sci. U.S.A.* 80:7650–7653.

Grossman, S.R., E. Johannsen, X. Tong, R. Yalamanchili, and E. Kieff. 1994. The Epstein-Barr virus nuclear antigen 2 transactivator is directed to response elements by the J kappa recombination signal binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 91:7568–7572.

Gunnell, A., H.M. Webb, C.D. Wood, M.J. McClellan, B. Wichaidit, B. Kempkes, R.G. Jenner, C. Osborne, P.J. Farrell, and M.J. West. 2016. RUNX super-enhancer control through the Notch pathway by Epstein-Barr virus transcription factors regulates B cell growth. *Nucleic Acids Research*. gkw085. doi:10.1093/nar/gkw085.

Haery, L., R.C. Thompson, and T.D. Gilmore. 2015. Histone acetyltransferases and histone deacetylases in B- and T-cell development, physiology and malignancy. *Genes Cancer*. 6:184–213.

Hammerschmidt, W., and B. Sugden. 1989. Genetic analysis of immortalizing functions of Epstein-Barr virus in human B lymphocytes. *Nature*. 340:393–397. doi:10.1038/340393a0.

Harada, S., and E. Kieff. 1997. Epstein-Barr virus nuclear protein LP stimulates EBNA-2 acidic domain-mediated transcriptional activation. *Journal of Virology*. 71:6611–6618.

Harris, R.S., and J.P. Dudley. 2015. APOBECs and virus restriction. *Virology*. 479-480:131–145. doi:10.1016/j.virol.2015.03.012.

Hart, S.L., C.V. Arancibia-Cárcamo, M.A. Wolfert, C. Mailhos, N.J. O'Reilly, R.R. Ali, C. Coutelle, A.J. George, R.P. Harbottle, A.M. Knight, D.F. Larkin, R.J. Levinsky, L.W. Seymour, A.J. Thrasher, and C. Kinnon. 1998. Lipid-mediated enhancement of transfection by a nonviral integrin-targeting vector. *Hum. Gene Ther.* 9:575–585. doi:10.1089/hum.1998.9.4-575.

Harth-Hertle, M.L., B.A. Scholz, F. Erhard, L.V. Glaser, L. Dölken, R. Zimmer, and B. Kempkes. 2013. Inactivation of Intergenic Enhancers by EBNA3A Initiates and Maintains Polycomb Signatures across a Chromatin Domain Encoding CXCL10 and CXCL9. *PLoS Pathog*. 9:e1003638. doi:10.1371/journal.ppat.1003638.

Haslinger, C., N. Schweifer, S. Stilgenbauer, H. Döhner, P. Lichter, N. Kraut, C. Stratowa, and R. Abseher. 2004. Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. *J. Clin. Oncol.* 22:3937–3949. doi:10.1200/JCO.2004.12.133.

Hassan, A.H., P. Prochasson, K.E. Neely, S.C. Galasinski, M. Chandy, M.J. Carrozza, and J.L. Workman. 2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell*. 111:369–379.

Hatzivassiliou, E., W.E. Miller, N. Raab-Traub, E. Kieff, and G. Mosialos. 1998. A fusion of the EBV latent membrane protein-1 (LMP1) transmembrane domains to the CD40 cytoplasmic domain is similar to LMP1 in constitutive activation of epidermal growth factor receptor expression, nuclear factor-kappa B, and stress-activated protein kinase. *J. Immunol.* 160:1116–1121.

Hauser, A.E., T. Junt, T.R. Mempel, M.W. Sneddon, S.H. Kleinstein, S.E. Henrickson, U.H. von Andrian, M.J. Shlomchik, and A.M. Haberman. 2007. Definition of germinal-center B cell migration in vivo reveals predominant intrazonal circulation patterns. *Immunity*. 26:655–667. doi:10.1016/j.immuni.2007.04.008.

Hawkins, J.B., E. Delgado-Eckert, D.A. Thorley-Lawson, and M. Shapiro. 2013. The cycle of EBV infection explains persistence, the sizes of the infected cell populations and which come under CTL regulation. *PLoS Pathog*. 9:e1003685. doi:10.1371/journal.ppat.1003685.

Hayakawa, T., and J.-I. Nakayama. 2011. Physiological roles of class I HDAC complex and histone demethylase. *J. Biomed. Biotechnol.* 2011:129383–10. doi:10.1155/2011/129383.

Hayakawa, T., Y. Ohtani, N. Hayakawa, K. Shinmyozu, M. Saito, F. Ishikawa, and J.-I. Nakayama. 2007. RBP2 is an MRG15 complex component and down-regulates intragenic histone H3 lysine 4 methylation. *Genes Cells*. 12:811–826. doi:10.1111/j.1365-2443.2007.01089.x.

Hayward, S.D. 2004. Viral interactions with the Notch pathway. *Semin. Cancer Biol.* 14:387–396. doi:10.1016/j.semcancer.2004.04.018.

Hayward, S.D., J. Liu, and M. Fujimuro. 2006. Notch and Wnt signaling: mimicry and manipulation by gamma herpesviruses. *Sci. STKE*. 2006:re4–re4. doi:10.1126/stke.3352006re4.

He, B., N. Raab-Traub, P. Casali, and A. Cerutti. 2003. EBV-encoded latent membrane protein 1 cooperates with BAFF/BLyS and APRIL to induce T cell-independent Ig heavy chain class switching. *J. Immunol.* 171:5215–5224.

He, J., L. Shen, M. Wan, O. Taranova, H. Wu, and Y. Zhang. 2013. Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nat. Cell Biol.* 15:373–384. doi:10.1038/ncb2702.

Heath, E., N. Begue-Pastor, S. Chaganti, D. Croom-Carter, C. Shannon-Lowe, D. Kube, R. Feederle, H.-J. Delecluse, A.B. Rickinson, and A.I. Bell. 2012. Epstein-Barr virus infection of naïve B cells in vitro frequently selects clones with mutated immunoglobulin genotypes: implications for virus biology. *PLoS Pathog*. 8:e1002697. doi:10.1371/journal.ppat.1002697.

Hemann, M.T., A. Bric, J. Teruya-Feldstein, A. Herbst, J.A. Nilsson, C. Cordon-Cardo, J.L. Cleveland, W.P. Tansey, and S.W. Lowe. 2005. Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants. *Nature*. 436:807–811. doi:10.1038/nature03845.

Henderson, S., M. Rowe, C. Gregory, D. Croom-Carter, F. Wang, R. Longnecker, E. Kieff, and A. Rickinson. 1991. Induction of bcl-2 expression by Epstein-Barr virus latent membrane protein 1 protects infected B cells from programmed cell death. *Cell*. 65:1107–1115.

Henkel, T., P.D. Ling, S.D. Hayward, and M.G. Peterson. 1994. Mediation of Epstein-Barr virus EBNA2 transactivation by recombination signal-binding protein J kappa. *Science*. 265:92–95.

Henle, W., V. Diehl, G. Kohn, H. Zur Hausen, and G. Henle. 1967. Herpes-type virus and chromosome marker in normal leukocytes after growth with irradiated Burkitt cells. *Science*. 157:1064–1065.

Hertle, M.L., C. Popp, S. Petermann, S. Maier, E. Kremmer, R. Lang, J. Mages, and B. Kempkes. 2009. Differential gene expression patterns of EBV infected EBNA-3A positive and negative human B lymphocytes. *PLoS Pathog*. 5:e1000506. doi:10.1371/journal.ppat.1000506.

Hess, J., A. Werner, T. Wirth, F. Melchers, H.M. Jäck, and T.H. Winkler. 2001. Induction of pre-B cell proliferation after de novo synthesis of the pre-B cell receptor. *Proc. Natl. Acad. Sci. U.S.A.* 98:1745–1750. doi:10.1073/pnas.041492098.

Hickabottom, M., G.A. Parker, P. Freemont, T. Crook, and M.J. Allday. 2002. Two nonconsensus sites in the Epstein-Barr virus oncoprotein EBNA3A cooperate to bind the co-repressor carboxyl-terminal-binding protein (CtBP). *J. Biol. Chem.* 277:47197–47204. doi:10.1074/jbc.M208116200.

Hislop, A.D., G.S. Taylor, D. Sauce, and A.B. Rickinson. 2007. Cellular responses to viral infection in humans: lessons from Epstein-Barr virus. *Annu. Rev. Immunol.* 25:587–617. doi:10.1146/annurev.immunol.25.022106.141553.

Hochberg, D., J.M. Middeldorp, M. Catalina, J.L. Sullivan, K. Luzuriaga, and D.A. Thorley-Lawson. 2004. Demonstration of the Burkitt's lymphoma Epstein-Barr virus phenotype in dividing latently infected memory cells in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 101:239–244. doi:10.1073/pnas.2237267100.

Hori, K., A. Sen, and S. Artavanis-Tsakonas. 2013. Notch signaling at a glance. *J. Cell. Sci.* 126:2135–2140. doi:10.1242/jcs.127308.

Hosogane, M., R. Funayama, Y. Nishida, T. Nagashima, and K. Nakayama. 2013. Ras-induced changes in H3K27me3 occur after those in transcriptional activity. *PLoS Genet.* 9:e1003698. doi:10.1371/journal.pgen.1003698.

Hotchin, N.A., M.J. Allday, and D.H. Crawford. 1990. Deregulated c-myc expression in Epstein-Barr-virus-immortalized B-cells induces altered growth properties and surface phenotype but not tumorigenicity. *Int. J. Cancer*. 45:566–571.

Howard, L., Y. Zheng, M. Horrocks, R.A. Maciewicz, and C. Blobel. 2001. Catalytic activity of ADAM28. *FEBS Lett.* 498:82–86.

Howard, M., J. Farrar, M. Hilfiker, B. Johnson, K. Takatsu, T. Hamaoka, and W.E. Paul. 1982. Identification of a T cell-derived b cell growth factor distinct from interleukin 2. *J. Exp. Med.* 155:914–923.

Hsieh, J.J., S. Zhou, L. Chen, D.B. Young, and S.D. Hayward. 1999. CIR, a corepressor linking the DNA binding factor CBF1 to the histone deacetylase complex. *Proc. Natl. Acad. Sci. U.S.A.* 96:23–28.

Hsu, P.D., E.S. Lander, and F. Zhang. 2014. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*. 157:1262–1278. doi:10.1016/j.cell.2014.05.010.

Huong, L.T., M. Kobayashi, M. Nakata, G. Shioi, H. Miyachi, T. Honjo, and H. Nagaoka. 2013. In vivo analysis of Aicda gene regulation: a critical balance between upstream enhancers and intronic silencers governs appropriate expression. *PLoS ONE*. 8:e61433. doi:10.1371/journal.pone.0061433.

Hurley, E.A., and D.A. Thorley-Lawson. 1988. B cell activation and the establishment of Epstein-Barr virus latency. *J. Exp. Med.* 168:2059–2075.

Hwang, J.K., F.W. Alt, and L.-S. Yeap. 2015. Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. *Microbiol Spectr*. 3:MDNA3–0037–2014. doi:10.1128/microbiolspec.MDNA3-0037-2014.

Ide, S., and J. Déjardin. 2015. End-targeting proteomics of isolated chromatin segments of a mammalian ribosomal RNA gene promoter. *Nat Commun*. 6:6674. doi:10.1038/ncomms7674.

Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata, and T. Honjo. 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl. Acad. Sci. U.S.A.* 101:1975–1980. doi:10.1073/pnas.0307335101.

Iwase, S., F. Lan, P. Bayliss, L. de la Torre-Ubieta, M. Huarte, H.H. Qi, J.R. Whetstine, A. Bonni, T.M. Roberts, and Y. Shi. 2007. The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell*. 128:1077–1088. doi:10.1016/j.cell.2007.02.017.

Jacob, J., and G. Kelsoe. 1992. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for periarteriolar lymphoid sheath-associated foci and germinal centers. *J. Exp. Med.* 176:679–687.

Jha, H.C., J. Lu, A. Saha, Q. Cai, S. Banerjee, M.A.J. Prasad, and E.S. Robertson. 2013. EBNA3C-mediated regulation of Aurora Kinase B contributes to EBV-induced B-cell proliferation through modulation of the activities of Rb and apoptotic Caspases. *Journal of Virology*. 87:12121–12138. doi:10.1128/JVI.02379-13.

Jiang, S., B. Willox, H. Zhou, A.M. Holthaus, A. Wang, T.T. Shi, S. Maruo, P.V. Kharchenko, E.C. Johannsen, E. Kieff, and B. Zhao. 2013. Epstein-Barr Virus Nuclear Antigen 3C binds to BATF/IRF4 or SPI1/IRF4 composite sites and recruits Sin3A to repress CDKN2A. *Proc. Natl. Acad. Sci. U.S.A.* doi:10.1073/pnas.1321704111.

Jiménez-Ramírez, C., A.J. Brooks, L.P. Forshell, K. Yakimchuk, B. Zhao, T.Z. Fulgham, and C.E. Sample. 2006. Epstein-Barr virus EBNA-3C is targeted to and regulates expression from the bidirectional LMP-1/2B promoter. *Journal of Virology*. 80:11200–11208. doi:10.1128/JVI.00897-06.

Jin, B., Y. Li, and K.D. Robertson. 2011. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer*. 2:607–617. doi:10.1177/1947601910393957.

Jochner, N., D. Eick, U. Zimber-Strobl, M. Pawlita, G.W. Bornkamm, and B. Kempkes. 1996. Epstein-Barr virus nuclear antigen 2 is a transcriptional suppressor of the immunoglobulin mu gene: implications for the expression of the translocated c-myc gene in Burkitt's lymphoma cells. *EMBO J.* 15:375–382.

Johannsen, E., C.L. Miller, S.R. Grossman, and E. Kieff. 1996. EBNA-2 and EBNA-3C extensively and mutually exclusively associate with RBPJkappa in Epstein-Barr virus-transformed B lymphocytes. *Journal of Virology*. 70:4179–4183.

Johannsen, E., E. Koh, G. Mosialos, X. Tong, E. Kieff, and S.R. Grossman. 1995. Epstein-Barr virus nuclear protein 2 transactivation of the latent membrane protein 1 promoter is mediated by J kappa and PU.1. *Journal of Virology*. 69:253–262.

Jones, P.L., G.J. Veenstra, P.A. Wade, D. Vermaak, S.U. Kass, N. Landsberger, J. Strouboulis, and A.P. Wolffe. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet*. 19:187–191. doi:10.1038/561.

Joseph, A.M., G.J. Babcock, and D.A. Thorley-Lawson. 2000a. EBV persistence involves strict selection of latently infected B cells. *J. Immunol.* 165:2975–2981.

Joseph, A.M., G.J. Babcock, and D.A. Thorley-Lawson. 2000b. Cells expressing the Epstein-Barr virus growth program are present in and restricted to the naive B-cell subset of healthy tonsils. *Journal of Virology*. 74:9964–9971.

Jowett, J.B.M., Y. Okada, P.J. Leedman, J.E. Curran, M.P. Johnson, E.K. Moses, H.H.H. Goring, S. Mochizuki, J. Blangero, L. Stone, H. Allen, C. Mitchell, and V.B. Matthews. 2012. ADAM28 is elevated in humans with the metabolic syndrome and is a novel sheddase of human tumour necrosis factor-α. *Immunol. Cell Biol.* 90:966–973. doi:10.1038/icb.2012.44.

Kadauke, S., and G.A. Blobel. 2013. Mitotic bookmarking by transcription factors. *Epigenetics Chromatin*. 6:6. doi:10.1186/1756-8935-6-6.

Kaiser, C., G. Laux, D. Eick, N. Jochner, G.W. Bornkamm, and B. Kempkes. 1999. The proto-oncogene c-myc is a direct target gene of Epstein-Barr virus nuclear antigen 2. *Journal of Virology*. 73:4481–4484.

Kalb, R., S. Latwiel, H.I. Baymaz, P.W.T.C. Jansen, C.W. Müller, M. Vermeulen, and J. Müller. 2014. Histone H2A monoubiquitination promotes histone H3 methylation in Polycomb repression. *Nat. Struct. Mol. Biol.* 21:569–571. doi:10.1038/nsmb.2833.

Kalchschmidt, J.S., A.C.T. Gillman, K. Paschos, Q. Bazot, B. Kempkes, and M.J. Allday. 2016. EBNA3C Directs Recruitment of RBPJ (CBF1) to Chromatin during the Process of Gene Repression in EBV Infected B Cells. *PLoS Pathog*. 12:e1005383. doi:10.1371/journal.ppat.1005383.

Kalla, M., and W. Hammerschmidt. 2012. Human B cells on their route to latent infection--early but transient expression of lytic genes of Epstein-Barr virus. *European Journal of Cell Biology*. 91:65–69. doi:10.1016/j.ejcb.2011.01.014.

Kanda, T., M. Otter, and G.M. Wahl. 2001. Coupling of mitotic chromosome tethering and replication competence in epstein-barr virus-based plasmids. *Mol. Cell. Biol.* 21:3576–3588. doi:10.1128/MCB.21.10.3576-3588.2001.

Kanhere, A., K. Viiri, C.C. Araújo, J. Rasaiyaah, R.D. Bouwman, W.A. Whyte, C.F. Pereira, E. Brookes, K. Walker, G.W. Bell, A. Pombo, A.G. Fisher, R.A. Young, and R.G. Jenner. 2010. Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol. Cell*. 38:675–688. doi:10.1016/j.molcel.2010.03.019.

Kanzler, H., R. Küppers, M.L. Hansmann, and K. Rajewsky. 1996. Hodgkin and Reed-Sternberg cells in Hodgkin's disease represent the outgrowth of a dominant tumor clone derived from (crippled) germinal center B cells. *J. Exp. Med.* 184:1495–1505.

Kao, H.Y., P. Ordentlich, N. Koyano-Nakagawa, Z. Tang, M. Downes, C.R. Kintner, R.M. Evans, and T. Kadesch. 1998. A histone deacetylase corepressor complex regulates the Notch signal transduction pathway. *Genes Dev.* 12:2269–2277.

Karasuyama, H., A. Kudo, and F. Melchers. 1990. The proteins encoded by the VpreB and lambda 5 pre-B cell-specific genes can associate with each other and with mu heavy chain. *J. Exp. Med.* 172:969–972.

Karasuyama, H., A. Rolink, Y. Shinkai, F. Young, F.W. Alt, and F. Melchers. 1994. The expression of Vpre-B/lambda 5 surrogate light chain in early bone marrow precursor B cells of normal and B cell-deficient mutant mice. *Cell*. 77:133–143.

Kaul, R., S.C. Verma, M. Murakami, K. Lan, T. Choudhuri, and E.S. Robertson. 2006. Epstein-Barr virus protein can upregulate cyclo-oxygenase-2 expression through association with the suppressor of metastasis Nm23-H1. *Journal of Virology*. 80:1321–1331. doi:10.1128/JVI.80.3.1321-1331.2006.

Kaye, K.M., K.M. Izumi, and E. Kieff. 1993. Epstein-Barr virus latent membrane protein 1 is essential for B-lymphocyte growth transformation. *Proc. Natl. Acad. Sci. U.S.A.* 90:9150–9154.

Kelly, G., A. Bell, and A. Rickinson. 2002. Epstein-Barr virus-associated Burkitt lymphomagenesis selects for downregulation of the nuclear antigen EBNA2. *Nat. Med.* 8:1098–1104. doi:10.1038/nm758.

Kempkes, B., and P.D. Ling. 2015. EBNA2 and Its Coactivator EBNA-LP. *Curr. Top. Microbiol. Immunol.* 391:35–59. doi:10.1007/978-3-319-22834-1_2.

Kempkes, B., D. Spitkovsky, P. Jansen-Dürr, J.W. Ellwart, E. Kremmer, H.J. Delecluse, C. Rottenberger, G.W. Bornkamm, and W. Hammerschmidt. 1995. B-cell proliferation and induction of early G1-regulating proteins by Epstein-Barr virus mutants conditional for EBNA2. *EMBO J.* 14:88–96.

Kessels, M.M., L. Schwintzer, D. Schlobinski, and B. Qualmann. 2011. Controlling actin cytoskeletal organization and dynamics during neuronal morphogenesis. *European Journal of Cell Biology*. 90:926–933. doi:10.1016/j.ejcb.2010.08.011.

Kieffer-Kwon, K.-R., Z. Tang, E. Mathe, J. Qian, M.-H. Sung, G. Li, W. Resch, S. Baek, N. Pruett, L. Grøntved, L. Vian, S. Nelson, H. Zare, O. Hakim, D. Reyon, A. Yamane, H. Nakahashi, A.L. Kovalchuk, J. Zou, J.K. Joung, V. Sartorelli, C.-L. Wei, X. Ruan, G.L. Hager, Y. Ruan, and R. Casellas. 2013. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell*. 155:1507–1520. doi:10.1016/j.cell.2013.11.039.

Kieser, A., and K.R. Sterz. 2015. The Latent Membrane Protein 1 (LMP1). *Curr. Top. Microbiol. Immunol.* 391:119–149. doi:10.1007/978-3-319-22834-1_4.

Kilger, E., A. Kieser, M. Baumann, and W. Hammerschmidt. 1998. Epstein-Barr virus-mediated B-cell proliferation is dependent upon latent membrane protein 1, which simulates an activated CD40 receptor. *EMBO J.* 17:1700–1709. doi:10.1093/emboj/17.6.1700.

Kim, H., K. Kang, and J. Kim. 2009. AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2. *Nucleic Acids Research*. 37:2940–2950. doi:10.1093/nar/gkp149.

Kim, J.H., W.S. Kim, and C. Park. 2013. Epstein-Barr virus latent membrane protein 1 increases genomic instability through Egr-1-mediated up-regulation of activation-induced cytidine deaminase in B-cell lymphoma. *Leuk. Lymphoma*. 54:2035–2040. doi:10.3109/10428194.2013.769218.

Kitamura, D., and K. Rajewsky. 1992. Targeted disruption of mu chain membrane exon causes loss of heavy-chain allelic exclusion. *Nature*. 356:154–156. doi:10.1038/356154a0.

Klein, G., B. Giovanella, A. Westman, J.S. Stehlin, and D. Mumford. 1975. An EBV-genome-negative cell line established from an American Burkitt lymphoma; receptor characteristics. EBV infectibility and permanent conversion into EBV-positive sublines by in vitro infection. *Intervirology*. 5:319–334.

Klein, G., J. Zeuthen, P. Terasaki, R. Billing, R. Honig, M. Jondal, A. Westman, and G. Clements. 1976. Inducibility of the Epstein-Barr virus (EBV) cycle and surface marker properties of EBV-negative lymphoma lines and their in vitro EBV-converted sublines. *Int. J. Cancer*. 18:639–652.

Klein, G., L. Dombos, and B. Gothoskar. 1972. Sensitivity of Epstein-Barr virus (EBV) producer and non-producer human lymphoblastoid cell lines to superinfection with EB-virus. *Int. J. Cancer*. 10:44–57.

Klein, T., and R. Bischoff. 2011. Active metalloproteases of the A Disintegrin and Metalloprotease (ADAM) family: biological function and structure. *J. Proteome Res.* 10:17–33. doi:10.1021/pr100556z.

Klein, U., Y. Tu, G.A. Stolovitzky, J.L. Keller, J. Haddad, V. Miljkovic, G. Cattoretti, A. Califano, and R. Dalla-Favera. 2003. Transcriptional analysis of the B cell germinal center reaction. *Proc. Natl. Acad. Sci. U.S.A.* 100:2639–2644. doi:10.1073/pnas.0437996100.

Klose, R.J., Q. Yan, Z. Tothova, K. Yamane, H. Erdjument-Bromage, P. Tempst, D.G. Gilliland, Y. Zhang, and W.G. Kaelin. 2007. The retinoblastoma binding protein RBP2 is an H3K4 demethylase. *Cell*. 128:889–900. doi:10.1016/j.cell.2007.02.013.

Knight, J.S., K. Lan, C. Subramanian, and E.S. Robertson. 2003. Epstein-Barr virus nuclear antigen 3C recruits histone deacetylase activity and associates with the corepressors mSin3A and NCoR in human B-cell lines. *Journal of Virology*. 77:4261–4272.

Kondo, T., K. Isono, K. Kondo, T.A. Endo, S. Itohara, M. Vidal, and H. Koseki. 2014. Polycomb potentiates meis2 activation in midbrain by mediating interaction of the promoter with a tissue-specific enhancer. *Dev. Cell*. 28:94–101. doi:10.1016/j.devcel.2013.11.021.

Kondo, T., S. Ito, and H. Koseki. 2016. Polycomb in Transcriptional Phase Transition of Developmental Genes. *Trends Biochem. Sci.* 41:9–19. doi:10.1016/j.tibs.2015.11.005.

Kooistra, S.M., and K. Helin. 2012. Molecular mechanisms and potential functions of histone demethylases. *Nat. Rev. Mol. Cell Biol.* 13:297–311. doi:10.1038/nrm3327.

Kosak, S.T., J.A. Skok, K.L. Medina, R. Riblet, M.M. Le Beau, A.G. Fisher, and H. Singh. 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science*. 296:158–162. doi:10.1126/science.1068768.

Kraja, A.T., D.I. Chasman, K.E. North, A.P. Reiner, L.R. Yanek, T.O. Kilpeläinen, J.A. Smith, A. Dehghan, J. Dupuis, A.D. Johnson, M.F. Feitosa, F. Tekola-Ayele, A.Y. Chu, I.M. Nolte, Z. Dastani, A. Morris, S.A. Pendergrass, Y.V. Sun, M.D. Ritchie, A. Vaez, H. Lin, S. Ligthart, L. Marullo, R. Rohde, Y. Shao, M.A. Ziegler, H.K. Im, Cross Consortia Pleiotropy Group, Cohorts for Heart and, Aging Research in Genetic Epidemiology, Genetic Investigation of Anthropometric Traits Consortium, Global Lipids Genetics Consortium, Meta-Analyses of Glucose, Insulin-related traits Consortium, Global BPgen Consortium, ADIPOGen Consortium, Women's Genome Health Study, Howard University Family Study, R.B. Schnabel, T. Jørgensen, M.E. Jørgensen, T. Hansen, O. Pedersen, R.P. Stolk, H. Snieder, A. Hofman, A.G. Uitterlinden, O.H. Franco, M.A. Ikram, J.B. Richards, C. Rotimi, J.G. Wilson, L. Lange, S.K. Ganesh, M. Nalls, L.J. Rasmussen-Torvik, J.S. Pankow, J. Coresh, W. Tang, W.H. Linda Kao, E. Boerwinkle, A.C. Morrison, P.M. Ridker, D.M. Becker, J.I. Rotter, S.L.R. Kardia, R.J.F. Loos, M.G. Larson, Y.-H. Hsu, M.A. Province, R. Tracy, B.F. Voight, D. Vaidya, C.J. O'Donnell, E.J. Benjamin, B.Z. Alizadeh, I. Prokopenko, J.B. Meigs, and I.B. Borecki. 2014. Pleiotropic genes for metabolic syndrome and inflammation. *Mol. Genet. Metab.* 112:317–338. doi:10.1016/j.ymgme.2014.04.007.

Krauer, K.G., N. Kienzle, D.B. Young, and T.B. Sculley. 1996. Epstein-Barr nuclear antigen-3 and -4 interact with RBP-2N, a major isoform of RBP-J kappa in B lymphocytes. *Virology*. 226:346–353. doi:10.1006/viro.1996.0662.

Krejcí, A., and S. Bray. 2007. Notch activation stimulates transient and selective binding of Su(H)/CSL to target enhancers. *Genes Dev.* 21:1322–1327. doi:10.1101/gad.424607.

Kriaucionis, S., and M. Tahiliani. 2014. Expanding the epigenetic landscape: novel modifications of cytosine in genomic DNA. *Cold Spring Harb Perspect Biol*. 6:a018630–

a018630. doi:10.1101/cshperspect.a018630.

Krokan, H.E., and M. Bjørås. 2013. Base excision repair. *Cold Spring Harb Perspect Biol*. 5:a012583–a012583. doi:10.1101/cshperspect.a012583.

Ku, M., R.P. Koche, E. Rheinbay, E.M. Mendenhall, M. Endoh, T.S. Mikkelsen, A. Presser, C. Nusbaum, X. Xie, A.S. Chi, M. Adli, S. Kasif, L.M. Ptaszek, C.A. Cowan, E.S. Lander, H. Koseki, and B.E. Bernstein. 2008. Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. *PLoS Genet.* 4:e1000242. doi:10.1371/journal.pgen.1000242.

Kudo, A., and F. Melchers. 1987. A second gene, VpreB in the lambda 5 locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes. *EMBO J.* 6:2267–2272.

Kumaran, R.I., and D.L. Spector. 2008. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J. Cell Biol.* 180:51–65. doi:10.1083/jcb.200706060.

Kuppers, D.A., K. Lan, J.S. Knight, and E.S. Robertson. 2005. Regulation of matrix metalloproteinase 9 expression by Epstein-Barr virus nuclear antigen 3C and the suppressor of metastasis Nm23-H1. *Journal of Virology*. 79:9714–9724. doi:10.1128/JVI.79.15.9714-9724.2005.

Kuroda, H., S. Mochizuki, M. Shimoda, M. Chijiiwa, K. Kamiya, Y. Izumi, M. Watanabe, H. Horinouchi, M. Kawamura, K. Kobayashi, and Y. Okada. 2010. ADAM28 is a serological and histochemical marker for non-small-cell lung cancers. *Int. J. Cancer*. 127:1844–1856. doi:10.1002/ijc.25212.

Kurooka, H., and T. Honjo. 2000. Functional interaction between the mouse notch1 intracellular region and histone acetyltransferases PCAF and GCN5. *J. Biol. Chem.* 275:17211–17220. doi:10.1074/jbc.M000909200.

Kurosaki, T., K. Kometani, and W. Ise. 2015. Memory B cells. *Nat. Rev. Immunol.* 15:149–159. doi:10.1038/nri3802.

Kuzmichev, A., K. Nishioka, H. Erdjument-Bromage, P. Tempst, and D. Reinberg. 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* 16:2893–2905. doi:10.1101/gad.1035902.

Küppers, R. 2005. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer*. 5:251–262. doi:10.1038/nrc1589.

Küppers, R. 2009. The biology of Hodgkin's lymphoma. *Nat Rev Cancer*. 9:15–27. doi:10.1038/nrc2542.

Küppers, R., K. Rajewsky, M. Zhao, G. Simons, R. Laumann, R. Fischer, and M.L. Hansmann. 1994. Hodgkin disease: Hodgkin and Reed-Sternberg cells picked from histological sections show clonal immunoglobulin gene rearrangements and appear to be derived from B cells at various stages of development. *Proc. Natl. Acad. Sci. U.S.A.* 91:10962–10966.

Laichalk, L.L., and D.A. Thorley-Lawson. 2005. Terminal differentiation into plasma cells initiates the replicative cycle of Epstein-Barr virus in vivo. *Journal of Virology*. 79:1296–

1307. doi:10.1128/JVI.79.2.1296-1307.2005.

Lake, R.J., P.-F. Tsai, I. Choi, K.-J. Won, and H.-Y. Fan. 2014. RBPJ, the major transcriptional effector of Notch signaling, remains associated with chromatin throughout mitosis, suggesting a role in mitotic bookmarking. *PLoS Genet.* 10:e1004204. doi:10.1371/journal.pgen.1004204.

Lanzuolo, C., V. Roure, J. Dekker, F. Bantignies, and V. Orlando. 2007. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat. Cell Biol.* 9:1167–1174. doi:10.1038/ncb1637.

Laskov, R., V. Yahud, R. Hamo, and M. Steinitz. 2011. Preferential targeting of somatic hypermutation to hotspot motifs and hypermutable sites and generation of mutational clusters in the IgVH alleles of a rheumatoid factor producing lymphoblastoid cell line. *Mol. Immunol.* 48:733–745. doi:10.1016/j.molimm.2010.10.009.

Laux, G., B. Adam, L.J. Strobl, and F. Moreau-Gachelin. 1994. The Spi-1/PU.1 and Spi-B ets family transcription factors and the recombination signal binding protein RBP-J kappa interact with an Epstein-Barr virus nuclear antigen 2 responsive cis-element. *EMBO J.* 13:5624–5632.

Le Roux, A., B. Kerdiles, D. Walls, J.F. Dedieu, and M. Perricaudet. 1994. The Epstein-Barr virus determined nuclear antigens EBNA-3A, -3B, and -3C repress EBNA-2-mediated transactivation of the viral terminal protein 1 gene promoter. *Virology*. 205:596–602. doi:10.1006/viro.1994.1687.

Lee, M.G., J. Norman, A. Shilatifard, and R. Shiekhattar. 2007. Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. *Cell*. 128:877–887. doi:10.1016/j.cell.2007.02.004.

Lee, S., S. Sakakibara, S. Maruo, B. Zhao, M.A. Calderwood, A.M. Holthaus, C.-Y. Lai, K. Takada, E. Kieff, and E. Johannsen. 2009. Epstein-Barr virus nuclear protein 3C domains necessary for lymphoblastoid cell growth: interaction with RBP-Jkappa regulates TCL1. *Journal of Virology*. 83:12368–12377. doi:10.1128/JVI.01403-09.

Lee, T.I., R.G. Jenner, L.A. Boyer, M.G. Guenther, S.S. Levine, R.M. Kumar, B. Chevalier, S.E. Johnstone, M.F. Cole, K.-I. Isono, H. Koseki, T. Fuchikami, K. Abe, H.L. Murray, J.P. Zucker, B. Yuan, G.W. Bell, E. Herbolsheimer, N.M. Hannett, K. Sun, D.T. Odom, A.P. Otte, T.L. Volkert, D.P. Bartel, D.A. Melton, D.K. Gifford, R. Jaenisch, and R.A. Young. 2006. Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*. 125:301–313. doi:10.1016/j.cell.2006.02.043.

Levine, S.S., A. Weiss, H. Erdjument-Bromage, Z. Shao, P. Tempst, and R.E. Kingston. 2002. The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. *Mol. Cell. Biol.* 22:6070–6078. doi:10.1128/MCB.22.17.6070-6078.2002.

Lewis, E.B. 1978. A gene complex controlling segmentation in Drosophila. *Nature*. 276:565–570.

Li, E., and Y. Zhang. 2014. DNA methylation in mammals. *Cold Spring Harb Perspect Biol*. 6:a019133–a019133. doi:10.1101/cshperspect.a019133.

Li, G., R. Margueron, M. Ku, P. Chambon, B.E. Bernstein, and D. Reinberg. 2010. Jarid2 and PRC2, partners in regulating gene expression. *Genes Dev.* 24:368–380.

doi:10.1101/gad.1886410.

Li, J., M. Chu, S. Wang, D. Chan, S. Qi, M. Wu, Z. Zhou, J. Li, E. Nishi, J. Qin, and J. Wong. 2012. Identification and characterization of nardilysin as a novel dimethyl H3K4-binding protein involved in transcriptional regulation. *J. Biol. Chem.* 287:10089–10098. doi:10.1074/jbc.M111.313965.

Li, Q., M.K. Spriggs, S. Kovats, S.M. Turk, M.R. Comeau, B. Nepom, and L.M. Hutt-Fletcher. 1997. Epstein-Barr virus uses HLA class II as a cofactor for infection of B lymphocytes. *Journal of Virology*. 71:4657–4662.

Lieberman-Aiden, E., N.L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B.R. Lajoie, P.J. Sabo, M.O. Dorschner, R. Sandstrom, B. Bernstein, M.A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L.A. Mirny, E.S. Lander, and J. Dekker. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 326:289–293. doi:10.1126/science.1181369.

Lin, J., E. Johannsen, E. Robertson, and E. Kieff. 2002. Epstein-Barr virus nuclear antigen 3C putative repression domain mediates coactivation of the LMP1 promoter with EBNA-2. *Journal of Virology*. 76:232–242.

Liu, Y.J., D.Y. Mason, G.D. Johnson, S. Abbot, C.D. Gregory, D.L. Hardie, J. Gordon, and I.C. MacLennan. 1991. Germinal center cells express bcl-2 protein after activation by signals which prevent their entry into apoptosis. *Eur. J. Immunol.* 21:1905–1910. doi:10.1002/eji.1830210819.

Lozzio, C.B., and B.B. Lozzio. 1975. Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood*. 45:321–334.

Lu, F., H.-S. Chen, A.V. Kossenkov, K. DeWispeleare, K.-J. Won, and P.M. Lieberman. 2016. EBNA2 Drives Formation of New Chromosome Binding Sites and Target Genes for B-Cell Master Regulatory Transcription Factors RBP-jκ and EBF1. *PLoS Pathog*. 12:e1005339. doi:10.1371/journal.ppat.1005339.

Lucchesi, W., G. Brady, O. Dittrich-Breiholz, M. Kracht, R. Russ, and P.J. Farrell. 2008. Differential gene regulation by Epstein-Barr virus type 1 and type 2 EBNA2. *Journal of Virology*. 82:7456–7466. doi:10.1128/JVI.00223-08.

Lund, J., L. Troeberg, H. Kjeldal, O.H. Olsen, H. Nagase, E.S. Sorensen, H.R. Stennicke, H.H. Petersen, and M.T. Overgaard. 2015. Evidence for Restricted Reactivity of ADAMDEC1 with Protein Substrates and Endogenous Inhibitors. *J. Biol. Chem.* jbc.M114.601724. doi:10.1074/jbc.M114.601724.

Lund, J., O.H. Olsen, E.S. Sorensen, H.R. Stennicke, H.H. Petersen, and M.T. Overgaard. 2013. ADAMDEC1 - a Metzincin metalloprotease with dampened proteolytic activity. *J. Biol. Chem.* doi:10.1074/jbc.M113.474536.

Lv, Y.-L., D.-M. Yuan, Q.-B. Wang, P. Zhan, L. Luo, T.-F. Lv, H.-B. Liu, Y.-F. Li, J. Wen, and Y. Song. 2012. Baseline and decline of serum ADAM28 during chemotherapy of advanced non-small cell lung cancer: a probable predictive and prognostic factor. *Med. Oncol.* 29:2633–2639. doi:10.1007/s12032-012-0180-5.

Macartney-Coxson, D.P., K.A. Hood, H.-J. Shi, T. Ward, A. Wiles, R. O'Connor, D.A. Hall, R.A. Lea, J.A. Royds, R.S. Stubbs, and S. Rooker. 2008. Metastatic susceptibility locus, an 8p hot-spot for tumour progression disrupted in colorectal liver metastases: 13

candidate genes examined at the DNA, mRNA and protein level. *BMC Cancer*. 8:187. doi:10.1186/1471-2407-8-187.

MacLennan, I.C. 1994. Germinal centers. *Annu. Rev. Immunol.* 12:117–139. doi:10.1146/annurev.iy.12.040194.001001.

Maier, S., G. Staffler, A. Hartmann, J. Höck, K. Henning, K. Grabusic, R. Mailhammer, R. Hoffmann, M. Wilmanns, R. Lang, J. Mages, and B. Kempkes. 2006. Cellular target genes of Epstein-Barr virus nuclear antigen 2. *Journal of Virology*. 80:9761–9771. doi:10.1128/JVI.00665-06.

Maier, S., M. Santak, A. Mantik, K. Grabusic, E. Kremmer, W. Hammerschmidt, and B. Kempkes. 2005. A somatic knockout of CBF1 in a human B-cell line reveals that induction of CD21 and CCR7 by EBNA-2 is strictly CBF1 dependent and that downregulation of immunoglobulin M is partially CBF1 independent. *Journal of Virology*. 79:8784–8792. doi:10.1128/JVI.79.14.8784-8792.2005.

Mancina, R.M., M.A. Burza, C. Maglio, C. Pirazzi, F. Sentinelli, M. Incani, T. Montalcini, A. Pujia, T. Congiu, S. Loche, S. Pilia, O. Wiklund, J. Borén, S. Romeo, and M.G. Baroni. 2013. The COBLL1 C allele is associated with lower serum insulin levels and lower insulin resistance in overweight and obese children. *Diabetes Metab. Res. Rev.* 29:413–416. doi:10.1002/dmrr.2408.

Mandel, T.E., R.P. Phipps, A. Abbot, and J.G. Tew. 1980. The follicular dendritic cell: long term antigen retention during immunity. *Immunol. Rev.* 53:29–59.

Mandel, T.E., R.P. Phipps, A.P. Abbot, and J.G. Tew. 1981. Long-term antigen retention by dendritic cells in the popliteal lymph node of immunized mice. *Immunology*. 43:353–362.

Manning, A.K., M.-F. Hivert, R.A. Scott, J.L. Grimsby, N. Bouatia-Naji, H. Chen, D. Rybin, C.-T. Liu, L.F. Bielak, I. Prokopenko, N. Amin, D. Barnes, G. Cadby, J.-J. Hottenga, E. Ingelsson, A.U. Jackson, T. Johnson, S. Kanoni, C. Ladenvall, V. Lagou, J. Lahti, C. Lecoeur, Y. Liu, M.T. Martinez-Larrad, M.E. Montasser, P. Navarro, J.R.B. Perry, L.J. Rasmussen-Torvik, P. Salo, N. Sattar, D. Shungin, R.J. Strawbridge, T. Tanaka, C.M. van Duijn, P. An, M. de Andrade, J.S. Andrews, T. Aspelund, M. Atalay, Y. Aulchenko, B. Balkau, S. Bandinelli, J.S. Beckmann, J.P. Beilby, C. Bellis, R.N. Bergman, J. Blangero, M. Boban, M. Boehnke, E. Boerwinkle, L.L. Bonnycastle, D.I. Boomsma, I.B. Borecki, Y. Böttcher, C. Bouchard, E. Brunner, D. Budimir, H. Campbell, O. Carlson, P.S. Chines, R. Clarke, F.S. Collins, A. Corbatón-Anchuelo, D. Couper, U. de Faire, G.V. Dedoussis, P. Deloukas, M. Dimitriou, J.M. Egan, G. Eiriksdottir, M.R. Erdos, J.G. Eriksson, E. Eury, L. Ferrucci, I. Ford, N.G. Forouhi, C.S. Fox, M.G. Franzosi, P.W. Franks, T.M. Frayling, P. Froguel, P. Galan, E. de Geus, B. Gigante, N.L. Glazer, A. Goel, L. Groop, V. Gudnason, G. Hallmans, A. Hamsten, O. Hansson, T.B. Harris, C. Hayward, S. Heath, S. Hercberg, A.A. Hicks, A. Hingorani, A. Hofman, et al. 2012. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet*. 44:659–669. doi:10.1038/ng.2274.

Mansouri, L., R. Gunnarsson, L.-A. Sutton, A. Ameur, S.D. Hooper, M. Mayrhofer, G. Juliusson, A. Isaksson, U. Gyllensten, and R. Rosenquist. 2012. Next generation RNA-sequencing in prognostic subsets of chronic lymphocytic leukemia. *Am. J. Hematol.* 87:737–740. doi:10.1002/ajh.23227.

Margueron, R., and D. Reinberg. 2011. The Polycomb complex PRC2 and its mark in life. *Nature*. 469:343–349. doi:10.1038/nature09784.

Margueron, R., G. Li, K. Sarma, A. Blais, J. Zavadil, C.L. Woodcock, B.D. Dynlacht, and D. Reinberg. 2008. Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol. Cell*. 32:503–518. doi:10.1016/j.molcel.2008.11.004.

Marshall, D., and C. Sample. 1995. Epstein-Barr virus nuclear antigen 3C is a transcriptional regulator. *Journal of Virology*. 69:3624–3630.

Maruo, S., B. Zhao, E. Johannsen, E. Kieff, J. Zou, and K. Takada. 2011. Epstein-Barr virus nuclear antigens 3C and 3A maintain lymphoblastoid cell growth by repressing p16INK4A and p14ARF expression. *Proc. Natl. Acad. Sci. U.S.A.* 108:1919–1924. doi:10.1073/pnas.1019599108.

Maruo, S., E. Johannsen, D. Illanes, A. Cooper, B. Zhao, and E. Kieff. 2005. Epstein-Barr virus nuclear protein 3A domains essential for growth of lymphoblasts: transcriptional regulation through RBP-Jkappa/CBF1 is critical. *Journal of Virology*. 79:10171–10179. doi:10.1128/JVI.79.16.10171-10179.2005.

Maruo, S., Y. Wu, S. Ishikawa, T. Kanda, D. Iwakiri, and K. Takada. 2006. Epstein-Barr virus nuclear protein EBNA3C is required for cell cycle progression and growth maintenance of lymphoblastoid cells. *Proc. Natl. Acad. Sci. U.S.A.* 103:19500–19505. doi:10.1073/pnas.0604919104.

Maruo, S., Y. Wu, T. Ito, T. Kanda, E.D. Kieff, and K. Takada. 2009. Epstein-Barr virus nuclear protein EBNA3C residues critical for maintaining lymphoblastoid cell growth. *Proc. Natl. Acad. Sci. U.S.A.* 106:4419–4424. doi:10.1073/pnas.0813134106.

Maunders, M.J., L. Petti, and M. Rowe. 1994. Precipitation of the Epstein-Barr virus protein EBNA 2 by an EBNA 3c-specific monoclonal antibody. *J. Gen. Virol.* 75 (Pt 4):769–778.

McBride, K.M., A. Gazumyan, E.M. Woo, T.A. Schwickert, B.T. Chait, and M.C. Nussenzweig. 2008. Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J. Exp. Med.* 205:2585–2594. doi:10.1084/jem.20081319.

McClellan, M.J., C.D. Wood, O. Ojeniyi, T.J. Cooper, A. Kanhere, A. Arvey, H.M. Webb, R.D. Palermo, M.L. Harth-Hertle, B. Kempkes, R.G. Jenner, and M.J. West. 2013. Modulation of enhancer looping and differential gene targeting by epstein-barr virus transcription factors directs cellular reprogramming. *PLoS Pathog*. 9:e1003636. doi:10.1371/journal.ppat.1003636.

McClellan, M.J., S. Khasnis, C.D. Wood, R.D. Palermo, S.N. Schlick, A.S. Kanhere, R.G. Jenner, and M.J. West. 2012. Downregulation of Integrin Receptor-Signaling Genes by Epstein-Barr Virus EBNA 3C via Promoter-Proximal and -Distal Binding Elements. *Journal of Virology*. 86:5165–5178. doi:10.1128/JVI.07161-11.

McGinn, O.J., W.R. English, S. Roberts, A. Ager, P. Newham, and G. Murphy. 2011. Modulation of integrin α4β1 by ADAM28 promotes lymphocyte adhesion and transendothelial migration. *Cell Biol. Int.* 35:1043–1053. doi:10.1042/CBI20100885.

McKean, D., K. Huppi, M. Bell, L. Staudt, W. Gerhard, and M. Weigert. 1984. Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *Proc. Natl. Acad. Sci. U.S.A.* 81:3180–3184.

Meehan, R.R., J.D. Lewis, S. McKay, E.L. Kleiner, and A.P. Bird. 1989. Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*. 58:499–507.

Mehta, I.S., M. Amira, A.J. Harvey, and J.M. Bridger. 2010. Rapid chromosome territory relocation by nuclear motor activity in response to serum removal in primary human fibroblasts. *Genome Biol.* 11:R5. doi:10.1186/gb-2010-11-1-r5.

Melchers, F. 2015. Checkpoints that control B cell development. *J. Clin. Invest.* 125:2203–2210. doi:10.1172/JCI78083.

Mendenhall, E.M., R.P. Koche, T. Truong, V.W. Zhou, B. Issac, A.S. Chi, M. Ku, and B.E. Bernstein. 2010. GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet.* 6:e1001244. doi:10.1371/journal.pgen.1001244.

Menezes, J., W. Leibold, G. Klein, and G. Clements. 1975. Establishment and characterization of an Epstein-Barr virus (EBC)-negative lymphoblastoid B cell line (BJA-B) from an exceptional, EBV-genome-negative African Burkitt's lymphoma. *Biomedicine*. 22:276–284.

Meng, F.-L., Z. Du, A. Federation, J. Hu, Q. Wang, K.-R. Kieffer-Kwon, R.M. Meyers, C. Amor, C.R. Wasserman, D. Neuberg, R. Casellas, M.C. Nussenzweig, J.E. Bradner, X.S. Liu, and F.W. Alt. 2014. Convergent Transcription at Intragenic Super-Enhancers Targets AID-Initiated Genomic Instability. *Cell*. 159:1538–1548. doi:10.1016/j.cell.2014.11.014.

Merkenschlager, M., and D.T. Odom. 2013. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell*. 152:1285–1297. doi:10.1016/j.cell.2013.02.029.

Messeguer, X., R. Escudero, D. Farré, O. Núñez, J. Martínez, and M.M. Albà. 2002. PROMO: detection of known transcription regulatory elements using species-tailored searches. *Bioinformatics*. 18:333–334.

Methot, S.P., L.C. Litzler, F. Trajtenberg, A. Zahn, F. Robert, J. Pelletier, A. Buschiazzo, B.G. Magor, and J.M. Di Noia. 2015. Consecutive interactions with HSP90 and eEF1A underlie a functional maturation and storage pathway of AID in the cytoplasm. *J. Exp. Med.* 212:581–596. doi:10.1084/jem.20141157.

Min, J., Y. Zhang, and R.-M. Xu. 2003. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes Dev.* 17:1823–1828. doi:10.1101/gad.269603.

Mitsui, Y., S. Mochizuki, T. Kodama, M. Shimoda, T. Ohtsuka, T. Shiomi, M. Chijiiwa, T. Ikeda, M. Kitajima, and Y. Okada. 2006. ADAM28 Is Overexpressed in Human Breast Carcinomas: Implications for Carcinoma Cell Proliferation through Cleavage of Insulin-like Growth Factor Binding Protein-3. *Cancer Research*. 66:9913–9920. doi:10.1158/0008-5472.CAN-06-0377.

Mochizuki, S., and Y. Okada. 2007. ADAMs in cancer cell proliferation and progression. *Cancer Sci.* 98:621–628. doi:10.1111/j.1349-7006.2007.00434.x.

Mochizuki, S., K. Soejima, M. Shimoda, H. Abe, A. Sasaki, H.J. Okano, H. Okano, and Y. Okada. 2012. Effect of ADAM28 on carcinoma cell metastasis by cleavage of von Willebrand factor. *JNCI Journal of the National Cancer Institute*. 104:906–922. doi:10.1093/jnci/djs232.

Mochizuki, S., M. Shimoda, T. Shiomi, Y. Fujii, and Y. Okada. 2004. ADAM28 is activated by MMP-7 (matrilysin-1) and cleaves insulin-like growth factor binding protein-3. *Biochemical and Biophysical Research Communications*. 315:79–84.

doi:10.1016/j.bbrc.2004.01.022.

Mochizuki, S., R. Tanaka, M. Shimoda, J. Onuma, Y. Fujii, H. Jinno, and Y. Okada. 2010. Connective tissue growth factor is a substrate of ADAM28. *Biochemical and Biophysical Research Communications*. 402:651–657. doi:10.1016/j.bbrc.2010.10.077.

Morey, L., A. Santanach, E. Blanco, L. Aloia, E.P. Nora, B.G. Bruneau, and L. Di Croce. 2015. Polycomb Regulates Mesoderm Cell Fate-Specification in Embryonic Stem Cells through Activation and Repression Mechanisms. *Cell Stem Cell*. 17:300–315. doi:10.1016/j.stem.2015.08.009.

Mori, Y., J. Yin, A. Rashid, B.A. Leggett, J. Young, L. Simms, P.M. Kuehl, P. Langenberg, S.J. Meltzer, and O.C. Stine. 2001. Instabilotyping: comprehensive identification of frameshift mutations caused by coding region microsatellite instability. *Cancer Research*. 61:6046–6049.

Moris, A., S. Murray, and S. Cardinaud. 2014. AID and APOBECs span the gap between innate and adaptive immunity. *Front. Microbio.* 5:534. doi:10.3389/fmicb.2014.00534.

Mousavi, K., H. Zare, A.H. Wang, and V. Sartorelli. 2012. Polycomb protein Ezh1 promotes RNA polymerase II elongation. *Mol. Cell*. 45:255–262. doi:10.1016/j.molcel.2011.11.019.

Mumm, J.S., and R. Kopan. 2000. Notch signaling: from the outside in. *Dev. Biol.* 228:151–165. doi:10.1006/dbio.2000.9960.

Muramatsu, M., K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai, and T. Honjo. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*. 102:553–563.

Muramatsu, M., V.S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N.O. Davidson, and T. Honjo. 1999. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J. Biol. Chem.* 274:18470–18476.

Murphy, G. 2008. The ADAMs: signalling scissors in the tumour microenvironment. *Nat Rev Cancer*. 8:929–941. doi:10.1038/nrc2459.

Müller, J., C.M. Hart, N.J. Francis, M.L. Vargas, A. Sengupta, B. Wild, E.L. Miller, M.B. O'Connor, R.E. Kingston, and J.A. Simon. 2002. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell*. 111:197–208.

Münz, C. 2015. EBV Infection of Mice with Reconstituted Human Immune System Components. *Curr. Top. Microbiol. Immunol.* 391:407–423. doi:10.1007/978-3-319-22834-1_14.

Nambu, Y., M. Sugai, H. Gonda, C.-G. Lee, T. Katakai, Y. Agata, Y. Yokota, and A. Shimizu. 2003. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science*. 302:2137–2140. doi:10.1126/science.1092481.

Nan, X., H.H. Ng, C.A. Johnson, C.D. Laherty, B.M. Turner, R.N. Eisenman, and A. Bird. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*. 393:386–389. doi:10.1038/30764.

Narendra, V., P.P. Rocha, D. An, R. Raviram, J.A. Skok, E.O. Mazzoni, and D. Reinberg. 2015. Transcription. CTCF establishes discrete functional chromatin domains at the Hox

clusters during differentiation. *Science*. 347:1017–1021. doi:10.1126/science.1262088.

Nemerow, G.R., R. Wolfert, M.E. McNaughton, and N.R. Cooper. 1985. Identification and characterization of the Epstein-Barr virus receptor on human B lymphocytes and its relationship to the C3d complement receptor (CR2). *Journal of Virology*. 55:347–351.

Neri, F., D. Incarnato, A. Krepelova, S. Rapelli, A. Pagnani, R. Zecchina, C. Parlato, and S. Oliviero. 2013. Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome Biol.* 14:R91. doi:10.1186/gb-2013-14-8-r91.

Nieuwenhuis, P., and D. Opstelten. 1984. Functional anatomy of germinal centers. *Am. J. Anat.* 170:421–435. doi:10.1002/aja.1001700315.

Nik-Zainal, S., L.B. Alexandrov, D.C. Wedge, P. Van Loo, C.D. Greenman, K. Raine, D. Jones, J. Hinton, J. Marshall, L.A. Stebbings, A. Menzies, S. Martin, K. Leung, L. Chen, C. Leroy, M. Ramakrishna, R. Rance, K.W. Lau, L.J. Mudie, I. Varela, D.J. McBride, G.R. Bignell, S.L. Cooke, A. Shlien, J. Gamble, I. Whitmore, M. Maddison, P.S. Tarpey, H.R. Davies, E. Papaemmanuil, P.J. Stephens, S. McLaren, A.P. Butler, J.W. Teague, G. Jönsson, J.E. Garber, D. Silver, P. Miron, A. Fatima, S. Boyault, A. Langerød, A. Tutt, J.W.M. Martens, S.A.J.R. Aparicio, A. Borg, A.V. Salomon, G. Thomas, A.-L. Børresen-Dale, A.L. Richardson, M.S. Neuberger, P.A. Futreal, P.J. Campbell, M.R. Stratton, Breast Cancer Working Group of the International Cancer Genome Consortium. 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell*. 149:979–993. doi:10.1016/j.cell.2012.04.024.

Nikolic, T., D. Movita, M.E.H. Lambers, C. Ribeiro de Almeida, P. Biesta, K. Kreefft, M.J.W. de Bruijn, I. Bergen, N. Galjart, A. Boonstra, and R. Hendriks. 2014. The DNA-binding factor Ctcf critically controls gene expression in macrophages. *Cell. Mol. Immunol.* 11:58–70. doi:10.1038/cmi.2013.41.

Nitsche, F., A. Bell, and A. Rickinson. 1997. Epstein-Barr virus leader protein enhances EBNA-2-mediated transactivation of latent membrane protein 1 expression: a role for the W1W2 repeat domain. *Journal of Virology*. 71:6619–6628.

Noordermeer, D., M. Leleu, E. Splinter, J. Rougemont, W. De Laat, and D. Duboule. 2011. The dynamic architecture of Hox gene clusters. *Science*. 334:222–225. doi:10.1126/science.1207194.

Ntziachristos, P., A. Tsirigos, P. Van Vlierberghe, J. Nedjic, T. Trimarchi, M.S. Flaherty, D. Ferres-Marco, V. da Ros, Z. Tang, J. Siegle, P. Asp, M. Hadler, I. Rigo, K. De Keersmaecker, J. Patel, T. Huynh, F. Utro, S. Poglio, J.B. Samon, E. Paietta, J. Racevskis, J.M. Rowe, R. Rabadan, R.L. Levine, S. Brown, F. Pflumio, M. Dominguez, A. Ferrando, and I. Aifantis. 2012. Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. *Nat. Med.* 18:298–301. doi:10.1038/nm.2651.

Nutt, S.L., P.D. Hodgkin, D.M. Tarlinton, and L.M. Corcoran. 2015. The generation of antibody-secreting plasma cells. *Nat. Rev. Immunol.* 15:160–171. doi:10.1038/nri3795.

Oettinger, M.A., D.G. Schatz, C. Gorka, and D. Baltimore. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science*. 248:1517–1523.

Ohashi, M., A.M. Holthaus, M.A. Calderwood, C.-Y. Lai, B. Krastins, D. Sarracino, and E.

Johannsen. 2015. The EBNA3 Family of Epstein-Barr Virus Nuclear Proteins Associates with the USP46/USP12 Deubiquitination Complexes to Regulate Lymphoblastoid Cell Line Growth. *PLoS Pathog*. 11:e1004822. doi:10.1371/journal.ppat.1004822.

Ohtsuka, T., T. Shiomi, M. Shimoda, T. Kodama, A. Amour, G. Murphy, E. Ohuchi, K. Kobayashi, and Y. Okada. 2006. ADAM28 is overexpressed in human non-small cell lung carcinomas and correlates with cell proliferation and lymph node metastasis. *Int. J. Cancer*. 118:263–273. doi:10.1002/ijc.21324.

Olave, I., D. Reinberg, and L.D. Vales. 1998. The mammalian transcriptional repressor RBP (CBF1) targets TFIID and TFIIA to prevent activated transcription. *Genes Dev.* 12:1621–1637.

Onizuka, T., M. Moriyama, T. Yamochi, T. Kuroda, A. Kazama, N. Kanazawa, K. Sato, T. Kato, H. Ota, and S. Mori. 1995. BCL-6 gene product, a 92- to 98-kD nuclear phosphoprotein, is highly expressed in germinal center B cells and their neoplastic counterparts. *Blood*. 86:28–37.

Ooi, S.K.T., C. Qiu, E. Bernstein, K. Li, D. Jia, Z. Yang, H. Erdjument-Bromage, P. Tempst, S.-P. Lin, C.D. Allis, X. Cheng, and T.H. Bestor. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*. 448:714–717. doi:10.1038/nature05987.

Orthwein, A., and J.M. Di Noia. 2012. Activation induced deaminase: how much and where? *Semin. Immunol.* 24:246–254. doi:10.1016/j.smim.2012.05.001.

Oswald, F., B. Täuber, T. Dobner, S. Bourteele, U. Kostezka, G. Adler, S. Liptay, and R.M. Schmid. 2001. p300 acts as a transcriptional coactivator for mammalian Notch-1. *Mol. Cell. Biol.* 21:7761–7774. doi:10.1128/MCB.21.22.7761-7774.2001.

Oswald, F., M. Winkler, Y. Cao, K. Astrahantseff, S. Bourteele, W. Knöchel, and T. Borggrefe. 2005. RBP-Jkappa/SHARP recruits CtIP/CtBP corepressors to silence Notch target genes. *Mol. Cell. Biol.* 25:10379–10390. doi:10.1128/MCB.25.23.10379-10390.2005.

Oswald, F., U. Kostezka, K. Astrahantseff, S. Bourteele, K. Dillinger, U. Zechner, L. Ludwig, M. Wilda, H. Hameister, W. Knöchel, S. Liptay, and R.M. Schmid. 2002. SHARP is a novel component of the Notch/RBP-Jkappa signalling pathway. *EMBO J.* 21:5417–5426. doi:10.1093/emboj/cdf549.

Palser, A.L., N.E. Grayson, R.E. White, C. Corton, S. Correia, M.M. Ba Abdullah, S.J. Watson, M. Cotten, J.R. Arrand, P.G. Murray, M.J. Allday, A.B. Rickinson, L.S. Young, P.J. Farrell, and P. Kellam. 2015. Genome Diversity of Epstein-Barr Virus from Multiple Tumour Types and Normal Infection. *Journal of Virology*. JVI.03614–14. doi:10.1128/JVI.03614-14.

Papaspyridonos, M., A. Smith, K.G. Burnand, P. Taylor, S. Padayachee, K.E. Suckling, C.H. James, D.R. Greaves, and L. Patel. 2006. Novel candidate genes in unstable areas of human atherosclerotic plaques. *Arterioscler. Thromb. Vasc. Biol.* 26:1837–1844. doi:10.1161/01.ATV.0000229695.68416.76.

Park, S.-R., H. Zan, Z. Pal, J. Zhang, A. Al-Qahtani, E.J. Pone, Z. Xu, T. Mai, and P. Casali. 2009. HoxC4 binds to the promoter of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. *Nat. Immunol.* 10:540–550. doi:10.1038/ni.1725.

Park, S.-R., P.-H. Kim, K.-S. Lee, S.-H. Lee, G.-Y. Seo, Y.-C. Yoo, J. Lee, and P. Casali. 2013. APRIL stimulates NF-κB-mediated HoxC4 induction for AID expression in mouse B cells. *Cytokine*. 61:608–613. doi:10.1016/j.cyto.2012.10.018.

Parker, B.D., A. Bankier, S. Satchwell, B. Barrell, and P.J. Farrell. 1990. Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B95-8 deletion region. *Virology*. 179:339–346.

Parker, M.J., S. Licence, L. Erlandsson, G.R. Galler, L. Chakalova, C.S. Osborne, G. Morgan, P. Fraser, H. Jumaa, T.H. Winkler, J. Skok, and I.-L. Mårtensson. 2005. The pre-B-cell receptor induces silencing of VpreB and lambda5 transcription. *EMBO J.* 24:3895–3905. doi:10.1038/sj.emboj.7600850.

Paschos, K., G.A. Parker, E. Watanatanasup, R.E. White, and M.J. Allday. 2012. BIM promoter directly targeted by EBNA3C in polycomb-mediated repression by EBV. *Nucleic Acids Research*. 40:7233–7246. doi:10.1093/nar/gks391.

Paschos, K., P. Smith, E. Anderton, J.M. Middeldorp, R.E. White, and M.J. Allday. 2009. Epstein-Barr Virus Latency in B Cells Leads to Epigenetic Repression and CpG Methylation of the Tumour Suppressor Gene Bim. *PLoS Pathog*. 5:e1000492. doi:10.1371/journal.ppat.1000492.g010.

Pasini, D., A.P. Bracken, J.B. Hansen, M. Capillo, and K. Helin. 2007. The polycomb group protein Suz12 is required for embryonic stem cell differentiation. *Mol. Cell. Biol.* 27:3769–3779. doi:10.1128/MCB.01432-06.

Pasini, F.S., B. Zilberstein, I. Snitcovsky, R.A. Roela, F.R.R. Mangone, U. Ribeiro, S. Nonogaki, G.C. Brito, G.D. Callegari, I. Cecconello, V.A.F. Alves, J. Eluf-Neto, R. Chammas, and M.H.H. Federico. 2014. A gene expression profile related to immune dampening in the tumor microenvironment is associated with poor prognosis in gastric adenocarcinoma. *J. Gastroenterol.* 49:1453–1466. doi:10.1007/s00535-013-0904-0.

Pasqualucci, L., R. Guglielmino, J. Houldsworth, J. Mohr, S. Aoufouchi, R. Polakiewicz, R.S.K. Chaganti, and R. Dalla-Favera. 2004. Expression of the AID protein in normal and neoplastic B cells. *Blood*. 104:3318–3325. doi:10.1182/blood-2004-04-1558.

Pasqualucci, L., Y. Kitaura, H. Gu, and R. Dalla-Favera. 2006. PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *Proc. Natl. Acad. Sci. U.S.A.* 103:395–400. doi:10.1073/pnas.0509969103.

Passarge, E. 1979. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. 31. Elsevier. 10 pp.

Patel, D.J., and Z. Wang. 2013. Readout of epigenetic modifications. *Annu. Rev. Biochem.* 82:81–118. doi:10.1146/annurev-biochem-072711-165700.

Patenaude, A.-M., A. Orthwein, Y. Hu, V.A. Campo, B. Kavli, A. Buschiazzo, and J.M. Di Noia. 2009. Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nat. Struct. Mol. Biol.* 16:517–527. doi:10.1038/nsmb.1598.

Patenaude, A.-M., and J.M. Di Noia. 2010. The mechanisms regulating the subcellular localization of AID. *Nucleus*. 1:325–331. doi:10.4161/nucl.1.4.12107.

Pauklin, S., I.V. Sernández, G. Bachmann, A.R. Ramiro, and S.K. Petersen-Mahrt. 2009. Estrogen directly activates AID transcription and function. *J. Exp. Med.* 206:99–111.

doi:10.1084/jem.20080521.

Pavri, R., A. Gazumyan, M. Jankovic, M. Di Virgilio, I. Klein, C. Ansarah-Sobrinho, W. Resch, A. Yamane, B. Reina-San-Martin, V. Barreto, T.J. Nieland, D.E. Root, R. Casellas, and M.C. Nussenzweig. 2010. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell*. 143:122–133. doi:10.1016/j.cell.2010.09.017.

Pefanis, E., J. Wang, G. Rothschild, J. Lim, J. Chao, R. Rabadan, A.N. Economides, and U. Basu. 2014. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature*. doi:10.1038/nature13580.

Pemberton, H., E. Anderton, H. Patel, S. Brookes, H. Chandler, R. Palermo, J. Stock, M. Rodriguez-Niedenführ, T. Racek, L. de Breed, A. Stewart, N. Matthews, and G. Peters. 2014. Genome-wide co-localization of Polycomb orthologs and their effects on gene expression in human fibroblasts. *Genome Biol.* 15:R23. doi:10.1186/gb-2014-15-2-r23.

Peng, C.-W., B. Zhao, and E. Kieff. 2004. Four EBNA2 domains are important for EBNALP coactivation. *Journal of Virology*. 78:11439–11442. doi:10.1128/JVI.78.20.11439-11442.2004.

Petcherski, A.G., and J. Kimble. 2000a. LAG-3 is a putative transcriptional activator in the C. elegans Notch pathway. *Nature*. 405:364–368. doi:10.1038/35012645.

Petcherski, A.G., and J. Kimble. 2000b. Mastermind is a putative activator for Notch. *Curr. Biol.* 10:R471–3.

Phan, R.T., and R. Dalla-Favera. 2004. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature*. 432:635–639. doi:10.1038/nature03147.

Pombo, A., and N. Dillon. 2015. Three-dimensional genome architecture: players and mechanisms. *Nat. Rev. Mol. Cell Biol.* 16:245–257. doi:10.1038/nrm3965.

Pope, J.H., M.K. Horne, and W. Scott. 1968. Transformation of foetal human keukocytes in vitro by filtrates of a human leukaemic cell line containing herpes-like virus. *Int. J. Cancer*. 3:857–866.

Portal, D., B. Zhao, M.A. Calderwood, T. Sommermann, E. Johannsen, and E. Kieff. 2011. EBV nuclear antigen EBNALP dismisses transcription repressors NCoR and RBPJ from enhancers and EBNA2 increases NCoR-deficient RBPJ DNA binding. *Proc. Natl. Acad. Sci. U.S.A.* 108:7808–7813. doi:10.1073/pnas.1104991108.

Portal, D., H. Zhou, B. Zhao, P.V. Kharchenko, E. Lowry, L. Wong, J. Quackenbush, D. Holloway, S. Jiang, Y. Lu, and E. Kieff. 2013. Epstein-Barr virus nuclear antigen leader protein localizes to promoters and enhancers with cell transcription factors and EBNA2. *Proc. Natl. Acad. Sci. U.S.A.* 110:18537–18542. doi:10.1073/pnas.1317608110.

Portis, T., and R. Longnecker. 2004. Epstein-Barr virus (EBV) LMP2A mediates B-lymphocyte survival through constitutive activation of the Ras/PI3K/Akt pathway. *Oncogene*. 23:8619–8628. doi:10.1038/sj.onc.1207905.

Qian, J., Q. Wang, M. Dose, N. Pruett, K.-R. Kieffer-Kwon, W. Resch, G. Liang, Z. Tang, E. Mathe, C. Benner, W. Dubois, S. Nelson, L. Vian, T.Y. Oliveira, M. Jankovic, O. Hakim, A. Gazumyan, R. Pavri, P. Awasthi, B. Song, G. Liu, L. Chen, S. Zhu, L. Feigenbaum, L. Staudt, C. Murre, Y. Ruan, D.F. Robbiani, Q. Pan-Hammarström, M.C. Nussenzweig,

and R. Casellas. 2014. B Cell Super-Enhancers and Regulatory Clusters Recruit AID Tumorigenic Activity. *Cell*. 159:1524–1537. doi:10.1016/j.cell.2014.11.013.

Raab-Traub, N., T. Dambaugh, and E. Kieff. 1980. DNA of Epstein-Barr virus VIII: B95-8, the previous prototype, is an unusual deletion derivative. *Cell*. 22:257–267.

Rada, C., G.T. Williams, H. Nilsen, D.E. Barnes, T. Lindahl, and M.S. Neuberger. 2002a. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr. Biol.* 12:1748–1755.

Rada, C., J.M. Di Noia, and M.S. Neuberger. 2004. Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol. Cell*. 16:163–171. doi:10.1016/j.molcel.2004.10.011.

Rada, C., J.M. Jarvis, and C. Milstein. 2002b. AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc. Natl. Acad. Sci. U.S.A.* 99:7003–7008. doi:10.1073/pnas.092160999.

Rada, C., M.R. Ehrenstein, M.S. Neuberger, and C. Milstein. 1998. Hot spot focusing of somatic hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. *Immunity*. 9:135–141.

Radic, M. 2008. Tracking and trapping somatic mutations in Ig genes. *J. Immunol.* 180:5763–5764.

Radkov, S.A., M. Bain, P.J. Farrell, M. West, M. Rowe, and M.J. Allday. 1997. Epstein-Barr virus EBNA3C represses Cp, the major promoter for EBNA expression, but has no effect on the promoter of the cell gene CD21. *Journal of Virology*. 71:8552–8562.

Radkov, S.A., R. Touitou, A. Brehm, M. Rowe, M. West, T. Kouzarides, and M.J. Allday. 1999. Epstein-Barr virus nuclear antigen 3C interacts with histone deacetylase to repress transcription. *Journal of Virology*. 73:5688–5697.

Ramiro, A.R., and V.M. Barreto. 2015. Activation-induced cytidine deaminase and active cytidine demethylation. *Trends Biochem. Sci.* 40:172–181. doi:10.1016/j.tibs.2015.01.006.

Randall, J.C., T.W. Winkler, Z. Kutalik, S.I. Berndt, A.U. Jackson, K.L. Monda, T.O. Kilpeläinen, T. Esko, R. Mägi, S. Li, T. Workalemahu, M.F. Feitosa, D.C. Croteau-Chonka, F.R. Day, T. Fall, T. Ferreira, S. Gustafsson, A.E. Locke, I. Mathieson, A. Scherag, S. Vedantam, A.R. Wood, L. Liang, V. Steinthorsdottir, G. Thorleifsson, E.T. Dermitzakis, A.S. Dimas, F. Karpe, J.L. Min, G. Nicholson, D.J. Clegg, T. Person, J.P. Krohn, S. Bauer, C. Buechler, K. Eisinger, DIAGRAM Consortium, A. Bonnefond, P. Froguel, MAGIC Investigators, J.-J. Hottenga, I. Prokopenko, L.L. Waite, T.B. Harris, A.V. Smith, A.R. Shuldiner, W.L. McArdle, M.J. Caulfield, P.B. Munroe, H. Grönberg, Y.-D.I. Chen, G. Li, J.S. Beckmann, T. Johnson, U. Thorsteinsdottir, M. Teder-Laving, K.-T. Khaw, N.J. Wareham, J.H. Zhao, N. Amin, B.A. Oostra, A.T. Kraja, M.A. Province, L.A. Cupples, N.L. Heard-Costa, J. Kaprio, S. Ripatti, I. Surakka, F.S. Collins, J. Saramies, J. Tuomilehto, A. Jula, V. Salomaa, J. Erdmann, C. Hengstenberg, C. Loley, H. Schunkert, C. Lamina, H.E. Wichmann, E. Albrecht, C. Gieger, A.A. Hicks, A. Johansson, P.P. Pramstaller, S. Kathiresan, E.K. Speliotes, B. Penninx, A.-L. Hartikainen, M.-R. Jarvelin, U. Gyllensten, D.I. Boomsma, H. Campbell, J.F. Wilson, S.J. Chanock, M. Farrall, A. Goel, C. Medina-Gomez, F. Rivadeneira, et al. 2013. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genet.* 9:e1003500. doi:10.1371/journal.pgen.1003500.

Ranuncolo, S.M., J.M. Polo, J. Dierov, M. Singer, T. Kuo, J. Greally, R. Green, M. Carroll, and A. Melnick. 2007. Bcl-6 mediates the germinal center B cell phenotype and lymphomagenesis through transcriptional repression of the DNA-damage sensor ATR. *Nat. Immunol.* 8:705–714. doi:10.1038/ni1478.

Rao, S.S.P., M.H. Huntley, N.C. Durand, E.K. Stamenova, I.D. Bochkov, J.T. Robinson, A.L. Sanborn, I. Machol, A.D. Omer, E.S. Lander, and E.L. Aiden. 2014. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell*. 159:1665–1680. doi:10.1016/j.cell.2014.11.021.

Reddy, K.L., J.M. Zullo, E. Bertolino, and H. Singh. 2008. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature*. 452:243–247. doi:10.1038/nature06727.

Reiss, K., and P. Saftig. 2009. The “a disintegrin and metalloprotease” (ADAM) family of sheddases: physiological and cellular functions. *Semin. Cell Dev. Biol.* 20:126–137. doi:10.1016/j.semcdb.2008.11.002.

Revy, P., T. Muto, Y. Levy, F. Geissmann, A. Plebani, O. Sanal, N. Catalan, M. Forveille, R. Dufourcq-Labelouse, A. Gennery, I. Tezcan, F. Ersoy, H. Kayserili, A.G. Ugazio, N. Brousse, M. Muramatsu, L.D. Notarangelo, K. Kinoshita, T. Honjo, A. Fischer, and A. Durandy. 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*. 102:565–575.

Riising, E.M., I. Comet, B. Leblanc, X. Wu, J.V. Johansen, and K. Helin. 2014. Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol. Cell*. 55:347–360. doi:10.1016/j.molcel.2014.06.005.

Rinn, J.L., M. Kertesz, J.K. Wang, S.L. Squazzo, X. Xu, S.A. Brugmann, L.H. Goodnough, J.A. Helms, P.J. Farnham, E. Segal, and H.Y. Chang. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*. 129:1311–1323. doi:10.1016/j.cell.2007.05.022.

Robbiani, D.F., A. Bothmer, E. Callen, B. Reina-San-Martin, Y. Dorsett, S. Difilippantonio, D.J. Bolland, H.T. Chen, A.E. Corcoran, A. Nussenzweig, and M.C. Nussenzweig. 2008. AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. *Cell*. 135:1028–1038. doi:10.1016/j.cell.2008.09.062.

Robbiani, D.F., and M.C. Nussenzweig. 2013. Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. *Annu Rev Pathol*. 8:79–103. doi:10.1146/annurev-pathol-020712-164004.

Robbiani, D.F., S. Deroubaix, N. Feldhahn, T.Y. Oliveira, E. Callen, Q. Wang, M. Jankovic, I.T. Silva, P.C. Rommel, D. Bosque, T. Eisenreich, A. Nussenzweig, and M.C. Nussenzweig. 2015. Plasmodium Infection Promotes Genomic Instability and AID-Dependent B Cell Lymphoma. *Cell*. 162:727–737. doi:10.1016/j.cell.2015.07.019.

Roberts, C.M., P.H. Tani, L.C. Bridges, Z. Laszik, and R.D. Bowditch. 1999. MDC-L, a novel metalloprotease disintegrin cysteine-rich protein family member expressed by human lymphocytes. *J. Biol. Chem.* 274:29251–29259.

Robertson, E.S., J. Lin, and E. Kieff. 1996. The amino-terminal domains of Epstein-Barr virus nuclear proteins 3A, 3B, and 3C interact with RBPJ(kappa). *Journal of Virology*. 70:3068–3074.

Robertson, E.S., S. Grossman, E. Johannsen, C. Miller, J. Lin, B. Tomkinson, and E. Kieff. 1995. Epstein-Barr virus nuclear protein 3C modulates transcription through interaction with the sequence-specific DNA-binding protein J kappa. *Journal of Virology*. 69:3108–3116.

Rocks, N., G. Paulissen, M. El Hour, F. Quesada, C. Crahay, M. Gueders, J.M. Foidart, A. Noel, and D. Cataldo. 2008. Emerging roles of ADAM and ADAMTS metalloproteinases in cancer. *Biochimie*. 90:369–379. doi:10.1016/j.biochi.2007.08.008.

Rogozin, I.B., and M. Diaz. 2004. Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. *J. Immunol.* 172:3382–3384.

Ronchetti, D., L. Mosca, G. Cutrona, G. Tuana, M. Gentile, S. Fabris, L. Agnelli, G. Ciceri, S. Matis, C. Massucco, M. Colombo, D. Reverberi, A.G. Recchia, S. Bossio, M. Negrini, P. Tassone, F. Morabito, M. Ferrarini, and A. Neri. 2013. Small nucleolar RNAs as new biomarkers in chronic lymphocytic leukemia. *BMC Med Genomics*. 6:27. doi:10.1186/1755-8794-6-27.

Rosa, S., and P. Shaw. 2013. Insights into chromatin structure and dynamics in plants. *Biology (Basel)*. 2:1378–1410. doi:10.3390/biology2041378.

Rose, N.R., and R.J. Klose. 2014. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim. Biophys. Acta*. 1839:1362–1372. doi:10.1016/j.bbagrm.2014.02.007.

Rowe, M., G.L. Kelly, A.I. Bell, and A.B. Rickinson. 2009. Burkitt's lymphoma: the Rosetta Stone deciphering Epstein-Barr virus biology. *Semin. Cancer Biol.* 19:377–388. doi:10.1016/j.semcancer.2009.07.004.

Ryan, J.L., W.K. Kaufmann, N. Raab-Traub, S.E. Oglesbee, L.A. Carey, and M.L. Gulley. 2006. Clonal evolution of lymphoblastoid cell lines. *Lab. Invest.* 86:1193–1200. doi:10.1038/labinvest.3700472.

Sabbattini, P., M. Sjöberg, S. Nikic, A. Frangini, P.-H. Holmqvist, N. Kunowska, T. Carroll, E. Brookes, S.J. Arthur, A. Pombo, and N. Dillon. 2014. An H3K9/S10 methyl-phospho switch modulates Polycomb and Pol II binding at repressed genes during differentiation. *Mol. Biol. Cell*. 25:904–915. doi:10.1091/mbc.E13-10-0628.

Sainsbury, S., C. Bernecky, and P. Cramer. 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 16:129–143. doi:10.1038/nrm3952.

Saito, M., U. Novak, E. Piovan, K. Basso, P. Sumazin, C. Schneider, M. Crespo, Q. Shen, G. Bhagat, A. Califano, A. Chadburn, L. Pasqualucci, and R. Dalla-Favera. 2009. BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc. Natl. Acad. Sci. U.S.A.* 106:11294–11299. doi:10.1073/pnas.0903854106.

Sakaguchi, N., and F. Melchers. 1986. Lambda 5, a new light-chain-related locus selectively expressed in pre-B lymphocytes. *Nature*. 324:579–582. doi:10.1038/324579a0.

Sakai, T., Y. Taniguchi, K. Tamura, S. Minoguchi, T. Fukuhara, L.J. Strobl, U. Zimber-Strobl, G.W. Bornkamm, and T. Honjo. 1998. Functional replacement of the intracellular region of the Notch1 receptor by Epstein-Barr virus nuclear antigen 2. *Journal of Virology*. 72:6034–6039.

Sample, J., E.B. Henson, and C. Sample. 1992. The Epstein-Barr virus nuclear protein 1 promoter active in type I latency is autoregulated. *Journal of Virology*. 66:4654–4661.

Sample, J., L. Young, B. Martin, T. Chatman, A. Rickinson, and E. Kieff. 1990. Epstein-Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B, and EBNA-3C genes. *Journal of Virology*. 64:4084–4092.

Sayegh, C.E., M.W. Quong, Y. Agata, and C. Murre. 2003. E-proteins directly regulate expression of activation-induced deaminase in mature B cells. *Nat. Immunol.* 4:586–593. doi:10.1038/ni923.

Schaaf, C.A., Z. Misulovin, M. Gause, A. Koenig, D.W. Gohara, A. Watson, and D. Dorsett. 2013. Cohesin and polycomb proteins functionally interact to control transcription at silenced and active genes. *PLoS Genet.* 9:e1003560. doi:10.1371/journal.pgen.1003560.

Schaefer, B.C., J.L. Strominger, and S.H. Speck. 1995. Redefining the Epstein-Barr virus-encoded nuclear antigen EBNA-1 gene promoter and transcription initiation site in group I Burkitt lymphoma cell lines. *Proc. Natl. Acad. Sci. U.S.A.* 92:10565–10569.

Schatz, D.G., M.A. Oettinger, and D. Baltimore. 1989. The V(D)J recombination activating gene, RAG-1. *Cell*. 59:1035–1048.

Schmidt, S.C.S., S. Jiang, H. Zhou, B. Willox, A.M. Holthaus, P.V. Kharchenko, E.C. Johannsen, E. Kieff, and B. Zhao. 2014. Epstein-Barr virus nuclear antigen 3A partially coincides with EBNA3C genome-wide and is tethered to DNA through BATF complexes. *Proc. Natl. Acad. Sci. U.S.A.* 201422580. doi:10.1073/pnas.1422580112.

Schmittgen, T.D., and K.J. Livak. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc*. 3:1101–1108.

Schmitz, R., M. Ceribelli, S. Pittaluga, G. Wright, and L.M. Staudt. 2014. Oncogenic mechanisms in Burkitt lymphoma. *Cold Spring Harb Perspect Med*. 4:a014282–a014282. doi:10.1101/cshperspect.a014282.

Schoeftner, S., A.K. Sengupta, S. Kubicek, K. Mechtler, L. Spahn, H. Koseki, T. Jenuwein, and A. Wutz. 2006. Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. *EMBO J.* 25:3110–3122. doi:10.1038/sj.emboj.7601187.

Schoenfelder, S., R. Sugar, A. Dimond, B.-M. Javierre, H. Armstrong, B. Mifsud, E. Dimitrova, L. Matheson, F. Tavares-Cadete, M. Furlan-Magaril, A. Segonds-Pichon, W. Jurkowski, S.W. Wingett, K. Tabbada, S. Andrews, B. Herman, E. LeProust, C.S. Osborne, H. Koseki, P. Fraser, N.M. Luscombe, and S. Elderkin. 2015. Polycomb repressive complex PRC1 spatially constrains the mouse embryonic stem cell genome. *Nat Genet*. 47:1179–1186. doi:10.1038/ng.3393.

Schwickert, T.A., R.L. Lindquist, G. Shakhar, G. Livshits, D. Skokos, M.H. Kosco-Vilbois, M.L. Dustin, and M.C. Nussenzweig. 2007. In vivo imaging of germinal centres reveals a dynamic open structure. *Nature*. 446:83–87. doi:10.1038/nature05573.

Selkoe, D., and R. Kopan. 2003. Notch and Presenilin: regulated intramembrane proteolysis links development and degeneration. *Annu. Rev. Neurosci.* 26:565–597. doi:10.1146/annurev.neuro.26.041002.131334.

Shaffer, A.L., X. Yu, Y. He, J. Boldrick, E.P. Chan, and L.M. Staudt. 2000. BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity*. 13:199–212.

Shao, Z., F. Raible, R. Mollaaghababa, J.R. Guyon, C.T. Wu, W. Bender, and R.E. Kingston. 1999. Stabilization of chromatin structure by PRC1, a Polycomb complex. *Cell*. 98:37–46. doi:10.1016/S0092-8674(00)80604-2.

Shimoda, M., G. Hashimoto, S. Mochizuki, E. Ikeda, N. Nagai, S. Ishida, and Y. Okada. 2007. Binding of ADAM28 to P-selectin glycoprotein ligand-1 enhances P-selectin-mediated leukocyte adhesion to endothelial cells. *J. Biol. Chem.* 282:25864–25874. doi:10.1074/jbc.M702414200.

Siemer, D., J. Kurth, S. Lang, G. Lehnerdt, J. Stanelle, and R. Küppers. 2008. EBV transformation overrides gene expression patterns of B cell differentiation stages. *Mol. Immunol.* 45:3133–3141. doi:10.1016/j.molimm.2008.03.002.

Sima, C., Q. Cheng, J. Rautava, C. Levesque, P. Sherman, and M. Glogauer. 2015. Identification of quantitative trait loci influencing inflammation-mediated alveolar bone loss: insights into polygenic inheritance of host-biofilm disequilibria in periodontitis. *J. Periodont. Res.* n/a–n/a. doi:10.1111/jre.12303.

Sinclair, A.J., and P.J. Farrell. 1995. Host cell requirements for efficient infection of quiescent primary B lymphocytes by Epstein-Barr virus. *Journal of Virology*. 69:5461–5468.

Sinclair, A.J., I. Palmero, G. Peters, and P.J. Farrell. 1994. EBNA-2 and EBNA-LP cooperate to cause G0 to G1 transition during immortalization of resting human B lymphocytes by Epstein-Barr virus. *EMBO J.* 13:3321–3328.

Singh, D., P.G. Febbo, K. Ross, D.G. Jackson, J. Manola, C. Ladd, P. Tamayo, A.A. Renshaw, A.V. D'Amico, J.P. Richie, E.S. Lander, M. Loda, P.W. Kantoff, T.R. Golub, and W.R. Sellers. 2002. Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell*. 1:203–209.

Skalska, L., R.E. White, and M.J. Allday. 2010. Epigenetic Repression of p16INK4A by Latent Epstein-Barr Virus Requires the Interaction of EBNA3A and EBNA3C with CtBP. *PLoS Pathog*. 6:e1000951. doi:10.1371/journal.ppat.1000951.g010.

Skalska, L., R.E. White, G.A. Parker, A.J. Sinclair, K. Paschos, and M.J. Allday. 2013. Induction of p16(INK4a) is the major barrier to proliferation when Epstein-Barr virus (EBV) transforms primary B cells into lymphoblastoid cell lines. *PLoS Pathog*. 9:e1003187. doi:10.1371/journal.ppat.1003187.

Soshnikova, N., T. Montavon, M. Leleu, N. Galjart, and D. Duboule. 2010. Functional analysis of CTCF during mammalian limb development. *Dev. Cell*. 19:819–830. doi:10.1016/j.devcel.2010.11.009.

Souza, T.A., B.D. Stollar, J.L. Sullivan, K. Luzuriaga, and D.A. Thorley-Lawson. 2005. Peripheral B cells latently infected with Epstein-Barr virus display molecular hallmarks of classical antigen-selected memory B cells. *Proc. Natl. Acad. Sci. U.S.A.* 102:18093–18098. doi:10.1073/pnas.0509311102.

Souza, T.A., B.D. Stollar, J.L. Sullivan, K. Luzuriaga, and D.A. Thorley-Lawson. 2007. Influence of EBV on the peripheral blood memory B cell compartment. *J. Immunol.*

179:3153–3160.

Speck, S.H. 2002. EBV framed in Burkitt lymphoma. *Nat. Med.* 8:1086–1087. doi:10.1038/nm1002-1086.

Spender, L.C., G.H. Cornish, A. Sullivan, and P.J. Farrell. 2002. Expression of transcription factor AML-2 (RUNX3, CBF(alpha)-3) is induced by Epstein-Barr virus EBNA-2 and correlates with the B-cell activation phenotype. *Journal of Virology*. 76:4919–4927. doi:10.1128/JVI.76.10.4919-4927.2002.

Stavnezer, J., and C.E. Schrader. 2014. IgH chain class switch recombination: mechanism and regulation. *J. Immunol.* 193:5370–5378. doi:10.4049/jimmunol.1401849.

Strahl, B.D., and C.D. Allis. 2000. The language of covalent histone modifications. *Nature*. 403:41–45. doi:10.1038/47412.

Strübbe, G., C. Popp, A. Schmidt, A. Pauli, L. Ringrose, C. Beisel, and R. Paro. 2011. Polycomb purification by in vivo biotinylation tagging reveals cohesin and Trithorax group proteins as interaction partners. *Proc. Natl. Acad. Sci. U.S.A.* 108:5572–5577. doi:10.1073/pnas.1007916108.

Subramanian, C., and E.S. Robertson. 2002. The metastatic suppressor Nm23-H1 interacts with EBNA3C at sequences located between the glutamine- and proline-rich domains and can cooperate in activation of transcription. *Journal of Virology*. 76:8702–8709. doi:10.1128/JVI.76.17.8702-8709.2002.

Subramanian, C., M.A. Cotter, and E.S. Robertson. 2001. Epstein-Barr virus nuclear protein EBNA-3C interacts with the human metastatic suppressor Nm23-H1: a molecular link to cancer metastasis. *Nat. Med.* 7:350–355. doi:10.1038/85499.

Subramanian, C., S. Hasan, M. Rowe, M. Hottiger, R. Orre, and E.S. Robertson. 2002. Epstein-Barr virus nuclear antigen 3C and prothymosin alpha interact with the p300 transcriptional coactivator at the CH1 and CH3/HAT domains and cooperate in regulation of transcription and histone acetylation. *Journal of Virology*. 76:4699–4708. doi:10.1128/JVI.76.10.4699-4708.2002.

Sugden, B., and N. Warren. 1989. A promoter of Epstein-Barr virus that can function during latent infection can be transactivated by EBNA-1, a viral protein required for viral DNA replication during latent infection. *Journal of Virology*. 63:2644–2649.

Sung, N.S., S. Kenney, D. Gutsch, and J.S. Pagano. 1991. EBNA-2 transactivates a lymphoid-specific enhancer in the BamHI C promoter of Epstein-Barr virus. *Journal of Virology*. 65:2164–2169.

Swanson-Mungerson, M.A., R.G. Caldwell, R. Bultema, and R. Longnecker. 2005. Epstein-Barr virus LMP2A alters in vivo and in vitro models of B-cell anergy, but not deletion, in response to autoantigen. *Journal of Virology*. 79:7355–7362. doi:10.1128/JVI.79.12.7355-7362.2005.

Swart, R., I.K. Ruf, J. Sample, and R. Longnecker. 2000. Latent membrane protein 2A-mediated effects on the phosphatidylinositol 3-Kinase/Akt pathway. *Journal of Virology*. 74:10838–10845.

Tahiliani, M., P. Mei, R. Fang, T. Leonor, M. Rutenberg, F. Shimizu, J. Li, A. Rao, and Y. Shi. 2007. The histone H3K4 demethylase SMCX links REST target genes to X-linked

mental retardation. *Nature*. 447:601–605. doi:10.1038/nature05823.

Takada, K., K. Horinouchi, Y. Ono, T. Aya, T. Osato, M. Takahashi, and S. Hayasaka. 1991. An Epstein-Barr virus-producer line Akata: establishment of the cell line and analysis of viral DNA. *Virus Genes*. 5:147–156.

Takizawa, M., H. Tolarová, Z. Li, W. Dubois, S. Lim, E. Callen, S. Franco, M. Mosaico, L. Feigenbaum, F.W. Alt, A. Nussenzweig, M. Potter, and R. Casellas. 2008. AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *J. Exp. Med.* 205:1949–1957. doi:10.1084/jem.20081007.

Tamura, K., Y. Taniguchi, S. Minoguchi, T. Sakai, T. Tun, T. Furukawa, and T. Honjo. 1995. Physical interaction between a novel domain of the receptor Notch and the transcription factor RBP-J kappa/Su(H). *Curr. Biol.* 5:1416–1423.

Tanay, A., A.H. O'Donnell, M. Damelin, and T.H. Bestor. 2007. Hyperconserved CpG domains underlie Polycomb-binding sites. *Proc. Natl. Acad. Sci. U.S.A.* 104:5521–5526. doi:10.1073/pnas.0609746104.

Taniguchi, Y., T. Furukawa, T. Tun, H. Han, and T. Honjo. 1998. LIM protein KyoT2 negatively regulates transcription by association with the RBP-J DNA-binding protein. *Mol. Cell. Biol.* 18:644–654.

Tavares, L., E. Dimitrova, D. Oxley, J. Webster, R. Poot, J. Demmers, K. Bezstarosti, S. Taylor, H. Ura, H. Koide, A. Wutz, M. Vidal, S. Elderkin, and N. Brockdorff. 2012. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell*. 148:664–678. doi:10.1016/j.cell.2011.12.029.

Taylor, S.R., M.G. Markesbery, and P.A. Harding. 2014. Heparin-binding epidermal growth factor-like growth factor (HB-EGF) and proteolytic processing by a disintegrin and metalloproteinases (ADAM): a regulator of several pathways. *Semin. Cell Dev. Biol.* 28:22–30. doi:10.1016/j.semcdb.2014.03.004.

Teng, G., P. Hakimpour, P. Landgraf, A. Rice, T. Tuschl, R. Casellas, and F.N. Papavasiliou. 2008. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity*. 28:621–629. doi:10.1016/j.immuni.2008.03.015.

Thomas-Claudepierre, A.-S., I. Robert, P.P. Rocha, R. Raviram, E. Schiavo, V. Heyer, R. Bonneau, V.M. Luo, J.K. Reddy, T. Borggrefe, J.A. Skok, and B. Reina-San-Martin. 2016. Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination. *J. Exp. Med.* 213:303–312. doi:10.1084/jem.20141967.

Thorley-Lawson, D.A. 2015. EBV Persistence-Introducing the Virus. *Curr. Top. Microbiol. Immunol.* 390:151–209. doi:10.1007/978-3-319-22822-8_8.

Thorley-Lawson, D.A., and G.J. Babcock. 1999. A model for persistent infection with Epstein-Barr virus: the stealth virus of human B cells. *Life Sci.* 65:1433–1453.

Thorley-Lawson, D.A., L.M. Nadler, A.K. Bhan, and R.T. Schooley. 1985. BLAST-2 [EBVCS], an early cell surface marker of human B cell activation, is superinduced by Epstein Barr virus. *J. Immunol.* 134:3007–3012.

Thorley-Lawson, D.A., R.T. Schooley, A.K. Bhan, and L.M. Nadler. 1982. Epstein-Barr virus

superinduces a new human B cell differentiation antigen (B-LAST 1) expressed on transformed lymphoblasts. *Cell*. 30:415–425.

Tiwari, V.K., K.M. McGarvey, J.D.F. Licchesi, J.E. Ohm, J.G. Herman, D. Schübeler, and S.B. Baylin. 2008a. PcG proteins, DNA methylation, and gene repression by chromatin looping. *PLoS Biol.* 6:2911–2927. doi:10.1371/journal.pbio.0060306.

Tiwari, V.K., L. Cope, K.M. McGarvey, J.E. Ohm, and S.B. Baylin. 2008b. A novel 6C assay uncovers Polycomb-mediated higher order chromatin conformations. *Genome Res.* 18:1171–1179. doi:10.1101/gr.073452.107.

Tobollik, S., L. Meyer, M. Buettner, S. Klemmer, B. Kempkes, E. Kremmer, G. Niedobitek, and B. Jungnickel. 2006. Epstein-Barr virus nuclear antigen 2 inhibits AID expression during EBV-driven B-cell growth. *Blood*. 108:3859–3864. doi:10.1182/blood-2006-05-021303.

Tollervey, J.R., and V.V. Lunyak. 2012. Epigenetics: judge, jury and executioner of stem cell fate. *Epigenetics*. 7:823–840. doi:10.4161/epi.21141.

Tomkinson, B., and E. Kieff. 1992. Use of second-site homologous recombination to demonstrate that Epstein-Barr virus nuclear protein 3B is not important for lymphocyte infection or growth transformation in vitro. *Journal of Virology*. 66:2893–2903.

Tomkinson, B., E. Robertson, and E. Kieff. 1993. Epstein-Barr virus nuclear proteins EBNA-3A and EBNA-3C are essential for B-lymphocyte growth transformation. *Journal of Virology*. 67:2014–2025.

Tompkins, V.S., S.-S. Han, A. Olivier, S. Syrbu, T. Bair, A. Button, L. Jacobus, Z. Wang, S. Lifton, P. Raychaudhuri, H.C. Morse, G. Weiner, B. Link, B.J. Smith, and S. Janz. 2013. Identification of candidate B-lymphoma genes by cross-species gene expression profiling. *PLoS ONE*. 8:e76889. doi:10.1371/journal.pone.0076889.

Tong, X., F. Wang, C.J. Thut, and E. Kieff. 1995a. The Epstein-Barr virus nuclear protein 2 acidic domain can interact with TFIIB, TAF40, and RPA70 but not with TATA-binding protein. *Journal of Virology*. 69:585–588.

Tong, X., R. Drapkin, D. Reinberg, and E. Kieff. 1995b. The 62- and 80-kDa subunits of transcription factor IIH mediate the interaction with Epstein-Barr virus nuclear protein 2. *Proc. Natl. Acad. Sci. U.S.A.* 92:3259–3263.

Torgbor, C., P. Awuah, K. Deitsch, P. Kalantari, K.A. Duca, and D.A. Thorley-Lawson. 2014. A multifactorial role for P. falciparum malaria in endemic Burkitt's lymphoma pathogenesis. *PLoS Pathog*. 10:e1004170. doi:10.1371/journal.ppat.1004170.

Touitou, R., J. O'Nions, J. Heaney, and M.J. Allday. 2005. Epstein-Barr virus EBNA3 proteins bind to the C8/alpha7 subunit of the 20S proteasome and are degraded by 20S proteasomes in vitro, but are very stable in latently infected B cells. *J. Gen. Virol.* 86:1269–1277. doi:10.1099/vir.0.80763-0.

Touitou, R., M. Hickabottom, G. Parker, T. Crook, and M.J. Allday. 2001. Physical and functional interactions between the corepressor CtBP and the Epstein-Barr virus nuclear antigen EBNA3C. *Journal of Virology*. 75:7749–7755. doi:10.1128/JVI.75.16.7749-7755.2001.

Tran, T.H., F.E. Utama, J. Lin, N. Yang, A.B. Sjolund, A. Ryder, K.J. Johnson, L.M. Neilson,

C. Liu, K.L. Brill, A.L. Rosenberg, A.K. Witkiewicz, and H. Rui. 2010a. Prolactin inhibits BCL6 expression in breast cancer through a Stat5a-dependent mechanism. *Cancer Research*. 70:1711–1721. doi:10.1158/0008-5472.CAN-09-2314.

Tran, T.H., M. Nakata, K. Suzuki, N.A. Begum, R. Shinkura, S. Fagarasan, T. Honjo, and H. Nagaoka. 2010b. B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. *Nat. Immunol.* 11:148–154. doi:10.1038/ni.1829.

Tsai, C.N., S.T. Liu, and Y.S. Chang. 1995. Identification of a novel promoter located within the Bam HI Q region of the Epstein-Barr virus genome for the EBNA 1 gene. *DNA Cell Biol.* 14:767–776.

Tsubata, T., and M. Reth. 1990. The products of pre-B cell-specific genes (lambda 5 and VpreB) and the immunoglobulin mu chain form a complex that is transported onto the cell surface. *J. Exp. Med.* 172:973–976.

Turner, B.M. 2000. Histone acetylation and an epigenetic code. *Bioessays*. 22:836–845. doi:10.1002/1521-1878(200009)22:9<836::AID-BIES9>3.0.CO;2-X.

Twito, T., Z. Chen, I. Khatri, K. Wong, D. Spaner, and R. Gorczynski. 2013. Ectodomain shedding of CD200 from the B-CLL cell surface is regulated by ADAM28 expression. *Leuk. Res.* doi:10.1016/j.leukres.2013.04.014.

Tzellos, S., P.B. Correia, C.E. Karstegl, L. Cancian, J. Cano-Flanagan, M.J. McClellan, M.J. West, and P.J. Farrell. 2014. A single amino acid in EBNA-2 determines superior B lymphoblastoid cell line growth maintenance by Epstein-Barr virus type 1 EBNA-2. *Journal of Virology*. 88:8743–8753. doi:10.1128/JVI.01000-14.

Uchimura, Y., L.F. Barton, C. Rada, and M.S. Neuberger. 2011. REG-γ associates with and modulates the abundance of nuclear activation-induced deaminase. *J. Exp. Med.* 208:2385–2391. doi:10.1084/jem.20110856.

van den Berg, A., B.-J. Kroesen, K. Kooistra, D. de Jong, J. Briggs, T. Blokzijl, S. Jacobs, J. Kluiver, A. Diepstra, E. Maggio, and S. Poppema. 2003. High expression of B-cell receptor inducible gene BIC in all subtypes of Hodgkin lymphoma. *Genes Chromosom. Cancer*. 37:20–28. doi:10.1002/gcc.10186.

Venner, J.M., K.S. Famulski, D. Badr, L.G. Hidalgo, J. Chang, and P.F. Halloran. 2014. Molecular landscape of T cell-mediated rejection in human kidney transplants: prominence of CTLA4 and PD ligands. *Am. J. Transplant.* 14:2565–2576. doi:10.1111/ajt.12946.

Vereide, D.T., and B. Sugden. 2011. Lymphomas differ in their dependence on Epstein-Barr virus. *Blood*. 117:1977–1985. doi:10.1182/blood-2010-05-285791.

Victora, G.D., and M.C. Nussenzweig. 2012. Germinal centers. *Annu. Rev. Immunol.* 30:429–457. doi:10.1146/annurev-immunol-020711-075032.

Victora, G.D., D. Dominguez-Sola, A.B. Holmes, S. Deroubaix, R. Dalla-Favera, and M.C. Nussenzweig. 2012. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. *Blood*. 120:2240–2248. doi:10.1182/blood-2012-03-415380.

Victora, G.D., T.A. Schwickert, D.R. Fooksman, A.O. Kamphorst, M. Meyer-Hermann, M.L. Dustin, and M.C. Nussenzweig. 2010. Germinal center dynamics revealed by

multiphoton microscopy with a photoactivatable fluorescent reporter. *Cell*. 143:592–605. doi:10.1016/j.cell.2010.10.032.

Viré, E., C. Brenner, R. Deplus, L. Blanchon, M. Fraga, C. Didelot, L. Morey, A. Van Eynde, D. Bernard, J.-M. Vanderwinden, M. Bollen, M. Esteller, L. Di Croce, Y. de Launoit, and F. Fuks. 2006. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*. 439:871–874. doi:10.1038/nature04431.

Vizán, P., M. Beringer, C. Ballaré, and L. Di Croce. 2015. Role of PRC2-associated factors in stem cells and disease. *FEBS J.* 282:1723–1735. doi:10.1111/febs.13083.

Wallberg, A.E., K. Pedersen, U. Lendahl, and R.G. Roeder. 2002. p300 and PCAF act cooperatively to mediate transcriptional activation from chromatin templates by notch intracellular domains in vitro. *Mol. Cell. Biol.* 22:7812–7819. doi:10.1128/MCB.22.22.7812-7819.2002.

Waltzer, L., F. Logeat, C. Brou, A. Israel, A. Sergeant, and E. Manet. 1994. The human J kappa recombination signal sequence binding protein (RBP-J kappa) targets the Epstein-Barr virus EBNA2 protein to its DNA responsive elements. *EMBO J.* 13:5633–5638.

Waltzer, L., M. Perricaudet, A. Sergeant, and E. Manet. 1996. Epstein-Barr virus EBNA3A and EBNA3C proteins both repress RBP-J kappa-EBNA2-activated transcription by inhibiting the binding of RBP-J kappa to DNA. *Journal of Virology*. 70:5909–5915.

Wang, A., R. Welch, B. Zhao, T. Ta, S. Keleş, and E. Johannsen. 2015. EBNA3 proteins regulate EBNA2 binding to distinct RBPJ genomic sites. *Journal of Virology*. JVI.02737–15. doi:10.1128/JVI.02737-15.

Wang, F., and J.M.G. Higgins. 2013. Histone modifications and mitosis: countermarks, landmarks, and bookmarks. *Trends Cell Biol.* 23:175–184. doi:10.1016/j.tcb.2012.11.005.

Wang, F., C.D. Gregory, M. Rowe, A.B. Rickinson, D. Wang, M. Birkenbach, H. Kikutani, T. Kishimoto, and E. Kieff. 1987. Epstein-Barr virus nuclear antigen 2 specifically induces expression of the B-cell activation antigen CD23. *Proc. Natl. Acad. Sci. U.S.A.* 84:3452–3456.

Wang, F., S.F. Tsang, M.G. Kurilla, J.I. Cohen, and E. Kieff. 1990. Epstein-Barr virus nuclear antigen 2 transactivates latent membrane protein LMP1. *Journal of Virology*. 64:3407–3416.

Wang, H., C. Zang, L. Taing, K.L. Arnett, Y.J. Wong, W.S. Pear, S.C. Blacklow, X.S. Liu, and J.C. Aster. 2014. NOTCH1-RBPJ complexes drive target gene expression through dynamic interactions with superenhancers. *Proc. Natl. Acad. Sci. U.S.A.* 111:705–710. doi:10.1073/pnas.1315023111.

Wang, H., J. Zou, B. Zhao, E. Johannsen, T. Ashworth, H. Wong, W.S. Pear, J. Schug, S.C. Blacklow, K.L. Arnett, B.E. Bernstein, E. Kieff, and J.C. Aster. 2011. Genome-wide analysis reveals conserved and divergent features of Notch1/RBPJ binding in human and murine T-lymphoblastic leukemia cells. *Proc. Natl. Acad. Sci. U.S.A.* 108:14908–14913. doi:10.1073/pnas.1109023108.

Wang, H., L. Wang, H. Erdjument-Bromage, M. Vidal, P. Tempst, R.S. Jones, and Y. Zhang. 2004. Role of histone H2A ubiquitination in Polycomb silencing. *Nature*. 431:873–878.

doi:10.1038/nature02985.

Wang, L., S.R. Grossman, and E. Kieff. 2000. Epstein-Barr virus nuclear protein 2 interacts with p300, CBP, and PCAF histone acetyltransferases in activation of the LMP1 promoter. *Proc. Natl. Acad. Sci. U.S.A.* 97:430–435.

Wang, X., and O. Gotoh. 2009. Accurate molecular classification of cancer using simple rules. *BMC Med Genomics*. 2:64. doi:10.1186/1755-8794-2-64.

Wang, Z., Z. Yan, B. Zhang, Z. Rao, Y. Zhang, J. Liu, L. Yu, Y. Zhao, B. Yang, T. Wu, and J. Gao. 2013. Identification of a 5-gene signature for clinical and prognostic prediction in gastric cancer patients upon microarray data. *Med. Oncol.* 30:678–11. doi:10.1007/s12032-013-0678-5.

Watson, C.T., and F. Breden. 2012. The immunoglobulin heavy chain locus: genetic variation, missing data, and implications for human disease. *Genes Immun.* 13:363–373. doi:10.1038/gene.2012.12.

Watt, F., and P.L. Molloy. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev.* 2:1136–1143.

White, R.E., E. Kremmer, and M.J. Allday. 2010. Extensive Co-Operation between the Epstein-Barr Virus EBNA3 Proteins in the Manipulation of Host Gene Expression and Epigenetic Chromatin Modification. *PLoS ONE*. 5:e13979. doi:10.1371/journal.pone.0013979.t001.

White, R.E., P.C. Rämer, K.N. Naresh, S. Meixlsperger, L. Pinaud, C. Rooney, B. Savoldo, R. Coutinho, C. Bödör, J. Gribben, H.A. Ibrahim, M. Bower, J.P. Nourse, M.K. Gandhi, J. Middeldorp, F.Z. Cader, P. Murray, C. Münz, and M.J. Allday. 2012. EBNA3B-deficient EBV promotes B cell lymphomagenesis in humanized mice and is found in human tumors. *J. Clin. Invest.* 122:1487–1502. doi:10.1172/JCI58092.

Wiesendanger, M., B. Kneitz, W. Edelmann, and M.D. Scharff. 2000. Somatic hypermutation in MutS homologue (MSH)3-, MSH6-, and MSH3/MSH6-deficient mice reveals a role for the MSH2-MSH6 heterodimer in modulating the base substitution pattern. *J. Exp. Med.* 191:579–584.

Wiles, M.V., W. Qin, A.W. Cheng, and H. Wang. 2015. CRISPR-Cas9-mediated genome editing and guide RNA design. *Mamm. Genome*. 26:501–510. doi:10.1007/s00335-015-9565-z.

Woisetschlaeger, M., C.N. Yandava, L.A. Furmanski, J.L. Strominger, and S.H. Speck. 1990. Promoter switching in Epstein-Barr virus during the initial stages of infection of B lymphocytes. *Proc. Natl. Acad. Sci. U.S.A.* 87:1725–1729.

Woisetschlaeger, M., X.W. Jin, C.N. Yandava, L.A. Furmanski, J.L. Strominger, and S.H. Speck. 1991. Role for the Epstein-Barr virus nuclear antigen 2 in viral promoter switching during initial stages of infection. *Proc. Natl. Acad. Sci. U.S.A.* 88:3942–3946.

Wu, L., J.C. Aster, S.C. Blacklow, R. Lake, S. Artavanis-Tsakonas, and J.D. Griffin. 2000. MAML1, a human homologue of Drosophila mastermind, is a transcriptional co-activator for NOTCH receptors. *Nat Genet*. 26:484–489. doi:10.1038/82644.

Wu, X., J.V. Johansen, and K. Helin. 2013. Fbxl10/Kdm2b recruits polycomb repressive

complex 1 to CpG islands and regulates H2A ubiquitylation. *Mol. Cell*. 49:1134–1146. doi:10.1016/j.molcel.2013.01.016.

Wysokenski, D.A., and J.L. Yates. 1989. Multiple EBNA1-binding sites are required to form an EBNA1-dependent enhancer and to activate a minimal replicative origin within oriP of Epstein-Barr virus. *Journal of Virology*. 63:2657–2666.

Xu, J., L. Liu, X. Zheng, C. You, and Q. Li. 2012a. Expression and inhibition of ADAMDEC1 in craniopharyngioma cells. *Neurol. Res.* 34:701–706. doi:10.1179/1743132812Y.0000000067.

Xu, J., Z. Shao, D. Li, H. Xie, W. Kim, J. Huang, J.E. Taylor, L. Pinello, K. Glass, J.D. Jaffe, G.-C. Yuan, and S.H. Orkin. 2015. Developmental control of polycomb subunit composition by GATA factors mediates a switch to non-canonical functions. *Mol. Cell*. 57:304–316. doi:10.1016/j.molcel.2014.12.009.

Xu, K., Z.J. Wu, A.C. Groner, H.H. He, C. Cai, R.T. Lis, X. Wu, E.C. Stack, M. Loda, T. Liu, H. Xu, L. Cato, J.E. Thornton, R.I. Gregory, C. Morrissey, R.L. Vessella, R. Montironi, C. Magi-Galluzzi, P.W. Kantoff, S.P. Balk, X.S. Liu, and M. Brown. 2012b. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. *Science*. 338:1465–1469. doi:10.1126/science.1227604.

Xu, M., G.-N. Zhao, X. Lv, G. Liu, L.Y. Wang, D.-L. Hao, J. Wang, D.-P. Liu, and C.-C. Liang. 2014. CTCF controls HOXA cluster silencing and mediates PRC2-repressive higher-order chromatin structure in NT2/D1 cells. *Mol. Cell. Biol.* 34:3867–3879. doi:10.1128/MCB.00567-14.

Yadav, A., A. Olaru, M. Saltis, A. Setren, J. Cerny, and F. Livák. 2006. Identification of a ubiquitously active promoter of the murine activation-induced cytidine deaminase (AICDA) gene. *Mol. Immunol.* 43:529–541. doi:10.1016/j.molimm.2005.05.007.

Yamane, A., D.F. Robbiani, W. Resch, A. Bothmer, H. Nakahashi, T. Oliveira, P.C. Rommel, E.J. Brown, A. Nussenzweig, M.C. Nussenzweig, and R. Casellas. 2013. RPA accumulation during class switch recombination represents 5“-3” DNA-end resection during the S-G2/M phase of the cell cycle. *Cell Rep*. 3:138–147. doi:10.1016/j.celrep.2012.12.006.

Yamane, A., W. Resch, N. Kuo, S. Kuchen, Z. Li, H.-W. Sun, D.F. Robbiani, K. McBride, M.C. Nussenzweig, and R. Casellas. 2011. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat. Immunol.* 12:62–69. doi:10.1038/ni.1964.

Yates, J., N. Warren, D. Reisman, and B. Sugden. 1984. A cis-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. *Proc. Natl. Acad. Sci. U.S.A.* 81:3806–3810.

Yates, J.L., N. Warren, and B. Sugden. 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. *Nature*. 313:812–815.

Ye, B.H., G. Cattoretti, Q. Shen, J. Zhang, N. Hawe, R. de Waard, C. Leung, M. Nouri-Shirazi, A. Orazi, R.S. Chaganti, P. Rothman, A.M. Stall, P.P. Pandolfi, and R. Dalla-Favera. 1997. The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nat Genet*. 16:161–170. doi:10.1038/ng0697-161.

Yeap, L.-S., J.K. Hwang, Z. Du, R.M. Meyers, F.-L. Meng, A. Jakubauskaitė, M. Liu, V. Mani,

D. Neuberg, T.B. Kepler, J.H. Wang, and F.W. Alt. 2015. Sequence-Intrinsic Mechanisms that Target AID Mutational Outcomes on Antibody Genes. *Cell*. 163:1124–1137. doi:10.1016/j.cell.2015.10.042.

Yenamandra, S.P., R. Sompallae, G. Klein, and E. Kashuba. 2009. Comparative analysis of the Epstein-Barr virus encoded nuclear proteins of EBNA-3 family. *Comput. Biol. Med.* 39:1036–1042. doi:10.1016/j.compbiomed.2009.08.006.

Young, L.S., and A.B. Rickinson. 2004. Epstein–Barr virus: 40 years on. *Nat Rev Cancer*. 4:757–768. doi:10.1038/nrc1452.

Yu, M., T. Mazor, H. Huang, H.-T. Huang, K.L. Kathrein, A.J. Woo, C.R. Chouinard, A. Labadorf, T.E. Akie, T.B. Moran, H. Xie, S. Zacharek, I. Taniuchi, R.G. Roeder, C.F. Kim, L.I. Zon, E. Fraenkel, and A.B. Cantor. 2012. Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors. *Mol. Cell*. 45:330–343. doi:10.1016/j.molcel.2011.11.032.

Yue, W., M.G. Davenport, J. Shackelford, and J.S. Pagano. 2004. Mitosis-specific hyperphosphorylation of Epstein-Barr virus nuclear antigen 2 suppresses its function. *Journal of Virology*. 78:3542–3552. doi:10.1128/JVI.78.7.3542-3552.2004.

Zaidi, S.K., D.W. Young, M.A. Montecino, J.B. Lian, A.J. van Wijnen, J.L. Stein, and G.S. Stein. 2010. Mitotic bookmarking of genes: a novel dimension to epigenetic control. *Nat. Rev. Genet.* 11:583–589. doi:10.1038/nrg2827.

Zeng, X., D.B. Winter, C. Kasmer, K.H. Kraemer, A.R. Lehmann, and P.J. Gearhart. 2001. DNA polymerase eta is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nat. Immunol.* 2:537–541. doi:10.1038/88740.

Zhang, X., L. Ding, and A.J. Sandford. 2005. Selection of reference genes for gene expression studies in human neutrophils by real-time PCR. *BMC Mol. Biol.* 6:4. doi:10.1186/1471-2199-6-4.

Zhang, X.-H., C.-C. Wang, Q. Jiang, S.-M. Yang, H. Jiang, J. Lu, Q.-M. Wang, F.-E. Feng, X.-L. Zhu, T. Zhao, and X.-J. Huang. 2015. ADAM28 overexpression regulated via the PI3K/Akt pathway is associated with relapse in de novo adult B-cell acute lymphoblastic leukemia. *Leuk. Res.* 39:1229–1238. doi:10.1016/j.leukres.2015.08.006.

Zhao, B., and C.E. Sample. 2000. Epstein-barr virus nuclear antigen 3C activates the latent membrane protein 1 promoter in the presence of Epstein-Barr virus nuclear antigen 2 through sequences encompassing an spi-1/Spi-B binding site. *Journal of Virology*. 74:5151–5160.

Zhao, B., D.R. Marshall, and C.E. Sample. 1996. A conserved domain of the Epstein-Barr virus nuclear antigens 3A and 3C binds to a discrete domain of Jkappa. *Journal of Virology*. 70:4228–4236.

Zhao, B., J. Zou, H. Wang, E. Johannsen, C.-W. Peng, J. Quackenbush, J.C. Mar, C.C. Morton, M.L. Freedman, S.C. Blacklow, J.C. Aster, B.E. Bernstein, and E. Kieff. 2011a. Epstein-Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. *Proc. Natl. Acad. Sci. U.S.A.* 108:14902–14907. doi:10.1073/pnas.1108892108.

Zhao, B., J.C. Mar, S. Maruo, S. Lee, B.E. Gewurz, E. Johannsen, K. Holton, R. Rubio, K. Takada, J. Quackenbush, and E. Kieff. 2011b. Epstein-Barr virus nuclear antigen 3C

regulated genes in lymphoblastoid cell lines. *Proc. Natl. Acad. Sci. U.S.A.* 108:337–342. doi:10.1073/pnas.1017419108.

Zhao, B., L.A. Barrera, I. Ersing, B. Willox, S.C.S. Schmidt, H. Greenfeld, H. Zhou, S.B. Mollo, T.T. Shi, K. Takasaki, S. Jiang, E. Cahir-McFarland, M. Kellis, M.L. Bulyk, E. Kieff, and B.E. Gewurz. 2014. The NF-κB genomic landscape in lymphoblastoid B cells. *Cell Rep*. 8:1595–1606. doi:10.1016/j.celrep.2014.07.037.

Zimber-Strobl, U., and L.J. Strobl. 2001. EBNA2 and Notch signalling in Epstein-Barr virus mediated immortalization of B lymphocytes. *Semin. Cancer Biol.* 11:423–434. doi:10.1006/scbi.2001.0409.

Zimber-Strobl, U., K.O. Suentzenich, G. Laux, D. Eick, M. Cordier, A. Calender, M. Billaud, G.M. Lenoir, and G.W. Bornkamm. 1991. Epstein-Barr virus nuclear antigen 2 activates transcription of the terminal protein gene. *Journal of Virology*. 65:415–423.

Zimber-Strobl, U., L.J. Strobl, C. Meitinger, R. Hinrichs, T. Sakai, T. Furukawa, T. Honjo, and G.W. Bornkamm. 1994. Epstein-Barr virus nuclear antigen 2 exerts its transactivating function through interaction with recombination signal binding protein RBP-J kappa, the homologue of Drosophila Suppressor of Hairless. *EMBO J.* 13:4973–4982.

7 Appendix

Table 7.1: Overview of restriction enzyme digestion fragments for recombinant EBV BACs used in this study.

Fragments are sorted by size and differences to wild type EBV BAC (wtI6) are highlighted in bold. For *Cla*I digestions, Dam methylation sites were analysed to determine the size of fragments.

***Cla*I digestion**

wtI6: pHB9, 9 *Cla*I sites

ClaI	Dam methylation	Fragment	Size in bp
1	No	1-5	65,585
2	Yes	9-1	61,533
3	Yes	8-9	21,039
4	Yes	5-6	15,107
5	No	6-8	3,491
6	No		
7	Yes		
8	No		
9	No		

3AKO: pB9-/-E3A(+12.5), 10 *Cla*I sites

ClaI	Dam methylation	Fragment	Size in bp
1	No	10-1	61,533
2	Yes	3-6	**38,489**
3	No	1-3	**24,461**
4	Yes	9-10	21,039
5	Yes	6-7	15,107
6	No	7-9	3,491
7	No		
8	Yes		
9	No		
10	No		

3BKO: pB9-/-E3B(+2.2), 9 *Cla*I sites

ClaI	Dam methylation	Fragment	Size in bp
1	No	1-5	**62,796**
2	Yes	9-1	61,533
3	Yes	8-9	21,039
4	Yes	5-6	15,107
5	No	6-8	3,491
6	No		
7	Yes		
8	No		
9	No		

3CKO: pB9-/-E3C(+3.1), 9 *Cla*I sites

ClaI	Dam methylation	Fragment	Size in bp
1	No	9-1	61,533
2	Yes	4-5	**32,442**
3	Yes	1-4	**30,985**
4	Yes	8-9	21,039
5	No	5-6	15,107
6	No	6-8	3,491
7	Yes		
8	No		
9	No		

E3KO: pHB9JY6+3C.3A, 8 *Cla*I sites

ClaI	Dam methylation	Fragment	Size in bp
1	No	8-1	61,533
2	Yes	3-4	**32,442**
3	No	1-3	**24,461**
4	No	7-8	21,039
5	No	4-5	15,107
6	Yes	5-7	3,491
7	No		
8	No		

***Eco*RI digestion**

wtI6: pHB9, 18 *Eco*RI sites

Fragment	Size in bp	Fragment	Size in bp
1	46,444	10	6,218
2	30,077	11	5,466
3	12,430	12	4,161
4	11,905	13	3,152
5	10,503	14	2,077
6	8,501	15	1,259
7	8,413	16	1,107
8	7,139	17	945
9	6,324	18	634

3AKO: pB9-/-E3A(+12.5), 17 *Eco*RI sites

Fragment	Size in bp	Fragment	Size in bp
1	46,444	10	6,218
2	30,077	11	5,466
3	12,430	12	4,161
4	11,905	13	3,152
5	10,503	14	1,259
6	8,501	15	945
7	8,413	16	634
8	7,139	17	**549**
9	6,324		

3BKO: pB9-/-E3B(+2.2), 18 *Eco*RI sites

Fragment	Size in bp	Fragment	Size in bp
1	46,444	10	6,218
2	**27,288**	11	5,466
3	12,430	12	4,161
4	11,905	13	3,152
5	10,503	14	2,077
6	8,501	15	1,259
7	8,413	16	1,107
8	7,139	17	945
9	6,324	18	634

3CKO: pB9-/-E3C(+3.1), 18 *Eco*RI sites

Fragment	Size in bp	Fragment	Size in bp
1	46,444	10	6,218
2	**27,919**	11	5,466
3	12,430	12	4,161
4	11,905	13	3,152
5	10,503	14	2,077
6	8,501	15	1,259
7	8,413	16	1,107
8	7,139	17	945
9	6,324	18	634

E3KO: pHB9JY6+3C.3A, 16 *Eco*RI sites

Fragment	Size in bp	Fragment	Size in bp
1	46,444	10	6,218
2	**24,579**	11	5,466
3	12,430	12	4,161
4	11,905	13	3,152
5	10,503	14	1,259
6	8,501	15	945
7	8,413	16	634
8	7,139		
9	6,324		

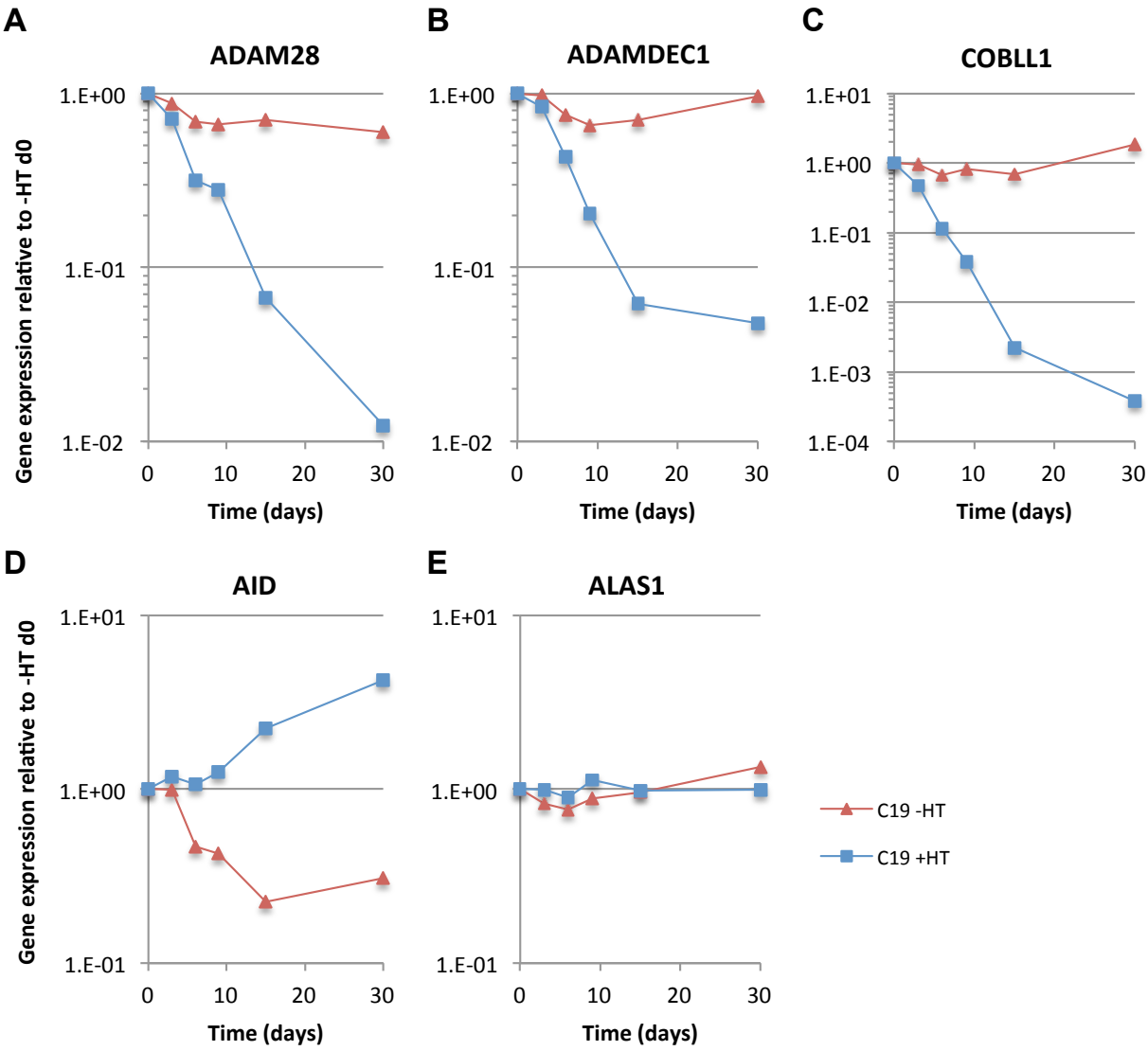


Figure 7.1: Gene expression analysis in the 3CHT C19 time-course.

Time-course using 3CHT C19 LCL. Cells were grown over 30 days either in absence of HT (-HT) or presence of HT (+HT). Gene expression of *ADAM28* (**A**), *ADAMDEC1* (**B**), *COBLL1* (**C**), *AICDA* (**D**) and *ALAS1* (**E**) was normalised to *GAPDH* and is shown relative to day 0 –HT.

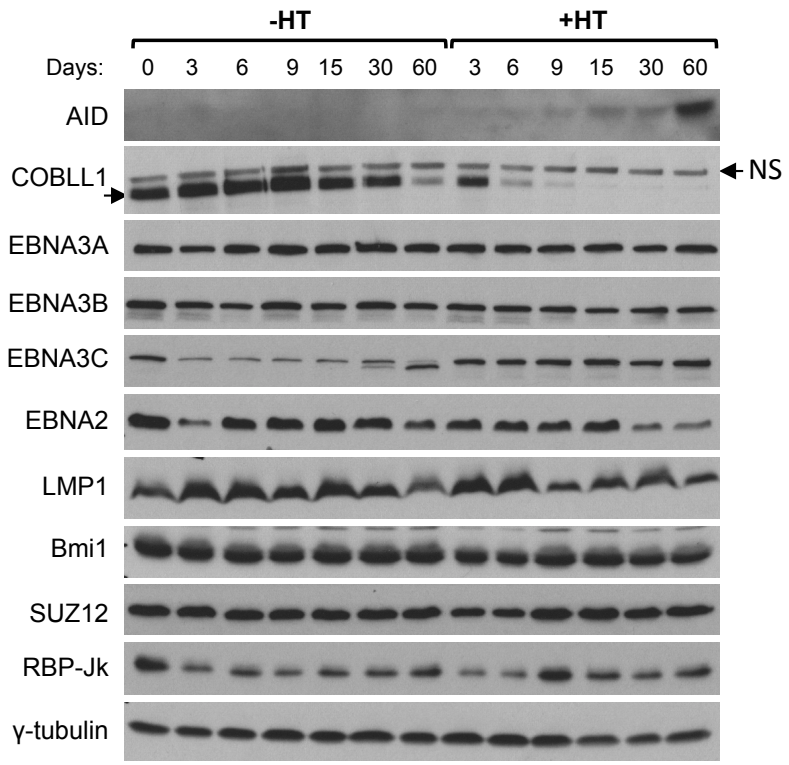


Figure 7.2: Western blot analysis of the 3CHT C19 time-course.

Time-course using 3CHT C19 LCL. Cells were grown over 60 days either in absence of HT (-HT) or presence of HT (+HT). 30 µg of RIPA extracts from selected samples were western blotted for the expression of AID, COBLL1, EBNA3A, EBNA3B, EBNA3C, EBNA2, LMP1, BMI1, SUZ12, RBPJ and γ-tubulin as loading control. NS indicates a non-specific band detected by the anti-COBLL1 antibody.

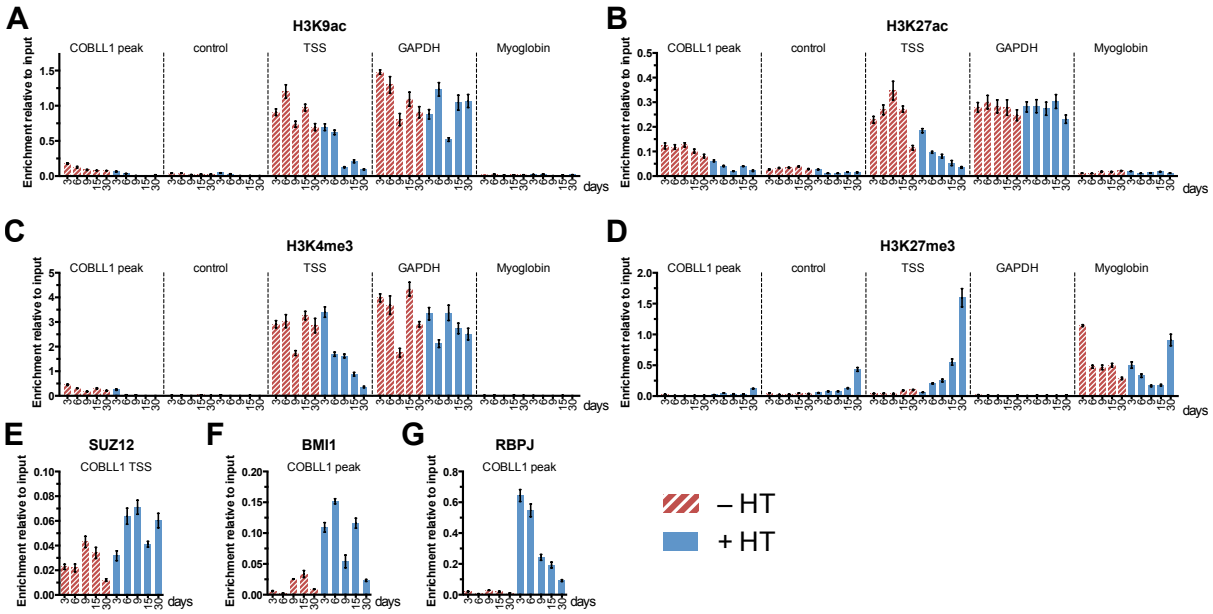


Figure 7.3: ChIP analysis of the *COBLL1* locus during the 3CHT C19 time-course.

(**A-G**) ChIP for H3K9ac (**A**), H3K27ac (**B**), H3K4me3 (**C**), H3K27me3 (**D**), SUZ12 (**E**), BMI1 (**F**) and RBPJ (**G**) on samples from the 3CHT C19 time-course at locations within the *COBLL1* locus (see Figure 3.8), at *GAPDH* or *myoglobin* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample.

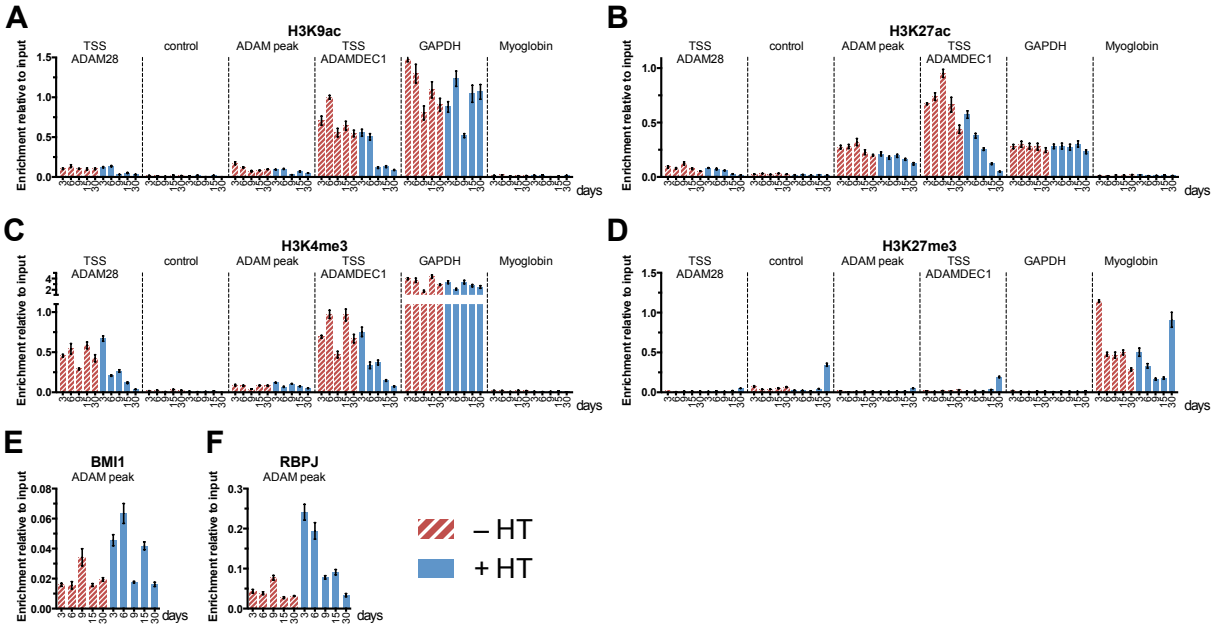


Figure 7.4: ChIP analysis of the *ADAM28-ADAMDEC1* locus during the 3CHT C19 time-course. (**A-F**) ChIP for H3K9ac (**A**), H3K27ac (**B**), H3K4me3 (**C**), H3K27me3 (**D**), BMI1 (**E**) and RBPJ (**F**) on samples from the 3CHT C19 time-course at locations across the *ADAM28-ADAMDEC1* locus (see Figure 3.8), at *GAPDH* or *myoglobin* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample.

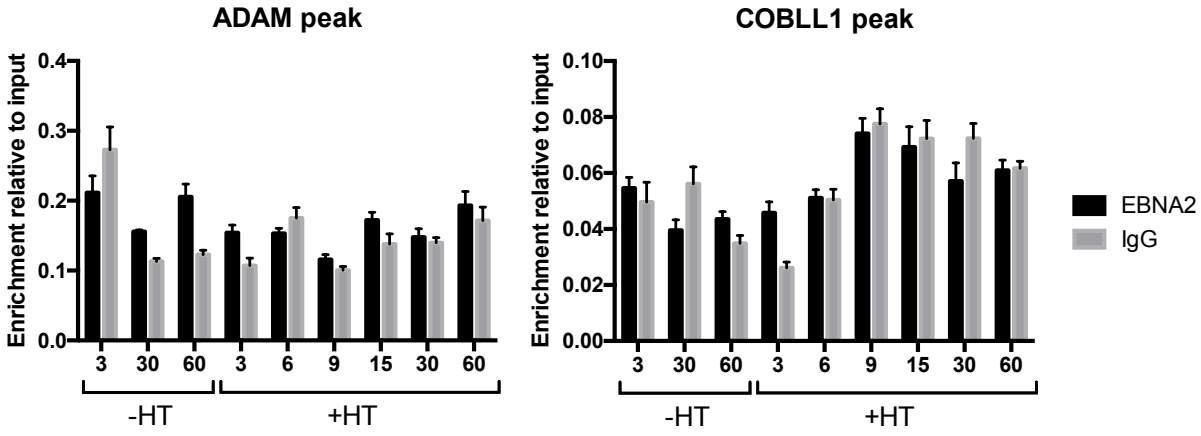


Figure 7.5: No changes to EBNA2 occupancy at *ADAM* peak and *COBLL1* peak during the 3CHT A13 time-course.

ChIP using antibodies against EBNA2 or non-specific IgG control on samples from the 3CHT A13 time-course at *ADAM* peak or *COBLL1* peak. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample.

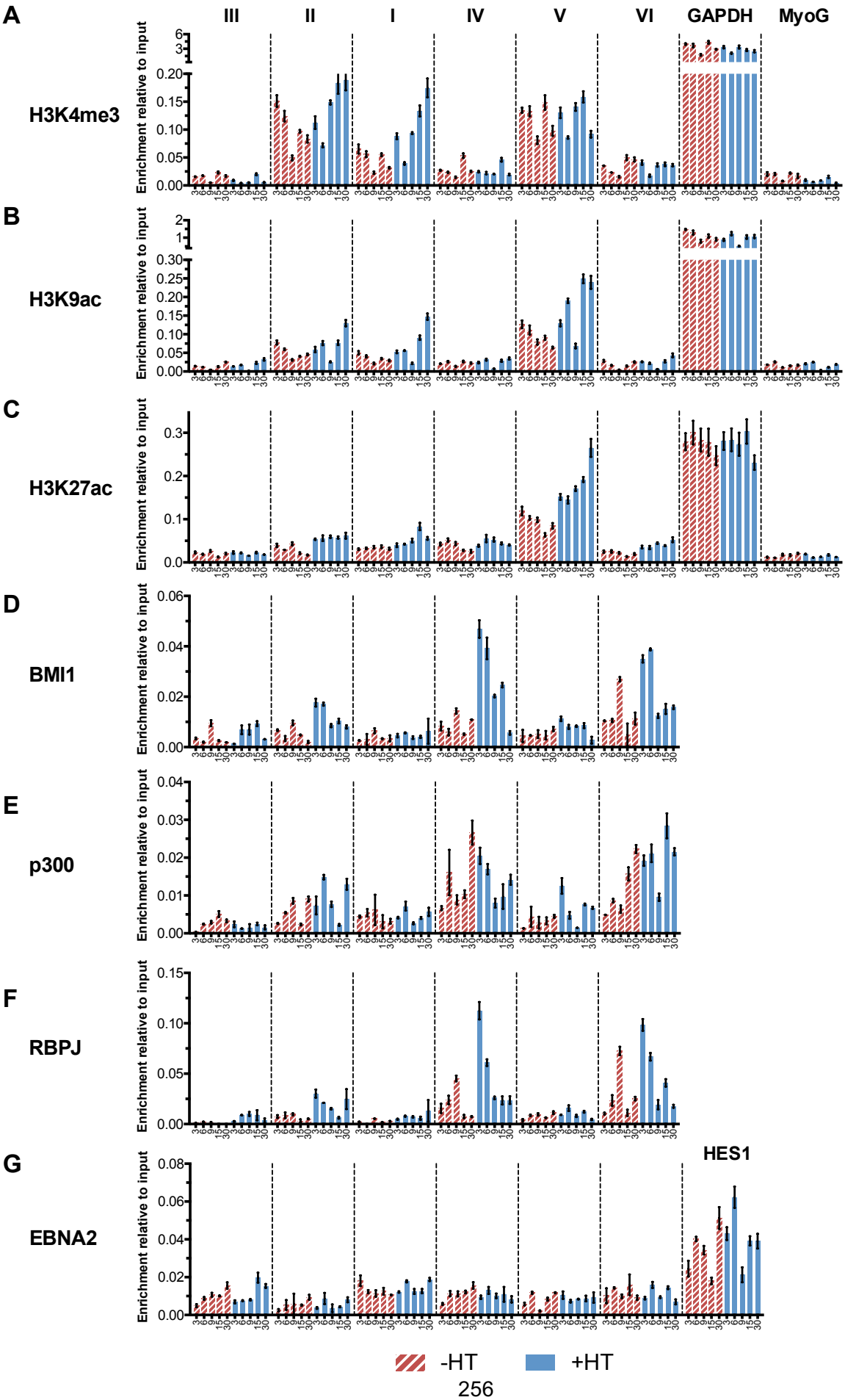

A
B
C
D
E
F
G
H3K4me3
H3K9ac
H3K27ac
BMI1
p300
RBPJ
EBNA2
III
II
I
IV
V
VI
GAPDH
MyoG
HES1
Enrichment relative to input
-HT
+HT

Figure 7.6: ChIP analysis of the *AICDA* locus during the 3CHT C19 time-course.
(**A-G**) ChIP for H3K4me3 (**A**), H3K9ac (**B**), H3K27ac (**C**), BMI1 (**D**), p300 (**E**), RBPJ (**F**) and EBNA2 (**G**) on samples from the 3CHT C19 time-course at locations across the *AICDA* locus (see Figure 4.6), at *GAPDH*, *myoglobin* or *HES1* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample.

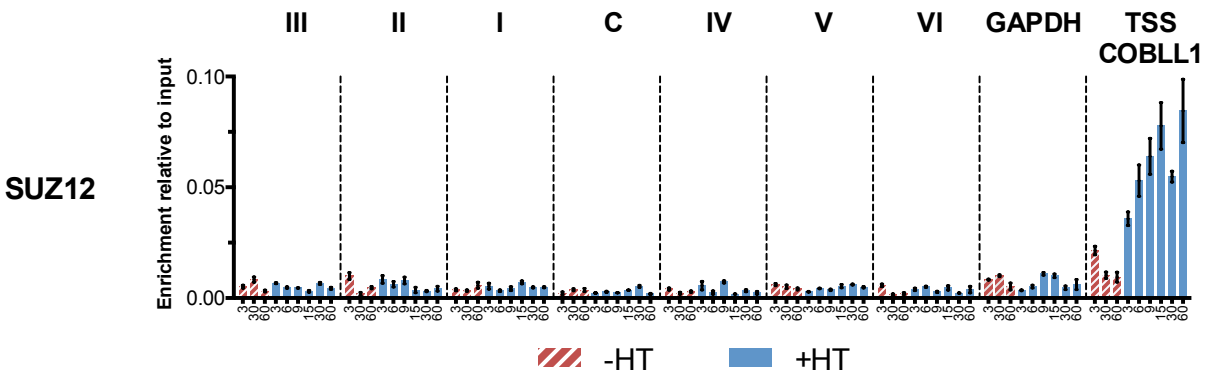


Figure 7.7: Occupancy of SUZ12 across the *AICDA* locus during the 3CHT A13 time-course.
ChIP for SUZ12 on samples from the 3CHT A13 time-course at locations across the *AICDA* locus (see Figure 4.6), at *GAPDH*, or the TSS of *COBLL1* as indicated. Cells were grown in the absence (-HT) or presence of HT (+HT) and numbers indicate the day of harvest. ChIP values represent enrichment relative to input ± SD of triplicate qPCR reactions for ChIP and input of each sample.

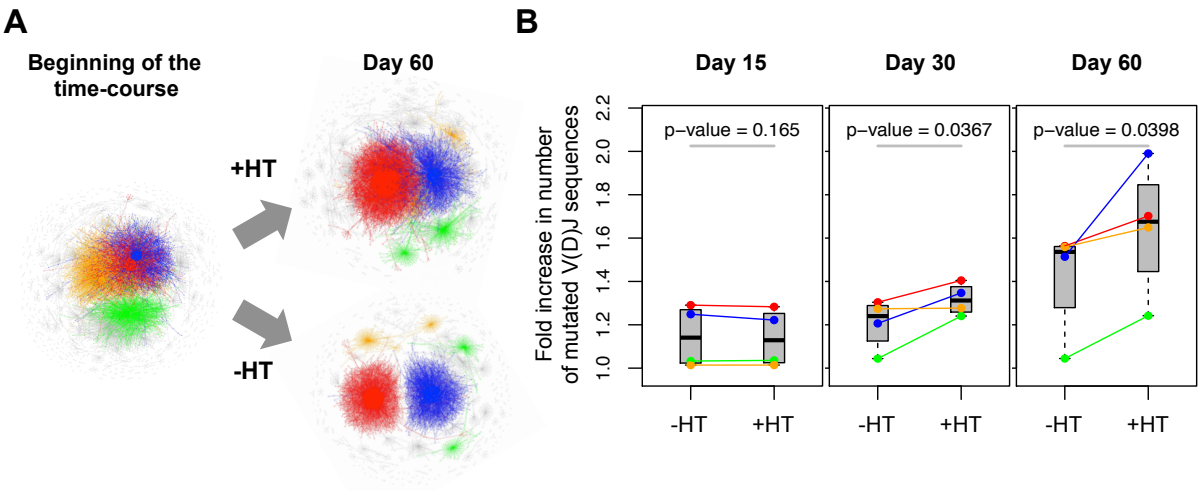


Figure 7.8: SHM and clonal development during 3CHT A13 time-course.

(**A**) Network diagrams of B cell clones according to their rearranged V(D)J IgH sequence in samples of the 3CHT A13 time-course show the clonal development of the four major clones (shown in red, blue, green, orange) along minor clones (shown in grey) from the starting population at day 0 of the time-course and after 60 days of culture with inactive (-HT) or active (+HT) EBNA3C. Each vertex represents a unique sequence, where relative vertex size is proportional to the number of identical reads. Edges join vertices that differ by single nucleotide differences and clusters are collections of related, connected vertices. (**B**) Boxplots showing the fold increase in the number of mutated V(D)J sequences) after 15, 30 and 60 days of culture with inactive (-HT) or active (+HT) EBNA3C for the four largest clones (shown in same colours in A). This is calculated as a cumulative change in the number of mutated V(D)J sequences at each time-point relative to the starting population at day 0. P-values were determined by paired t-test. Dr Rachael Bashford-Rogers performed the analysis and prepared the figures.

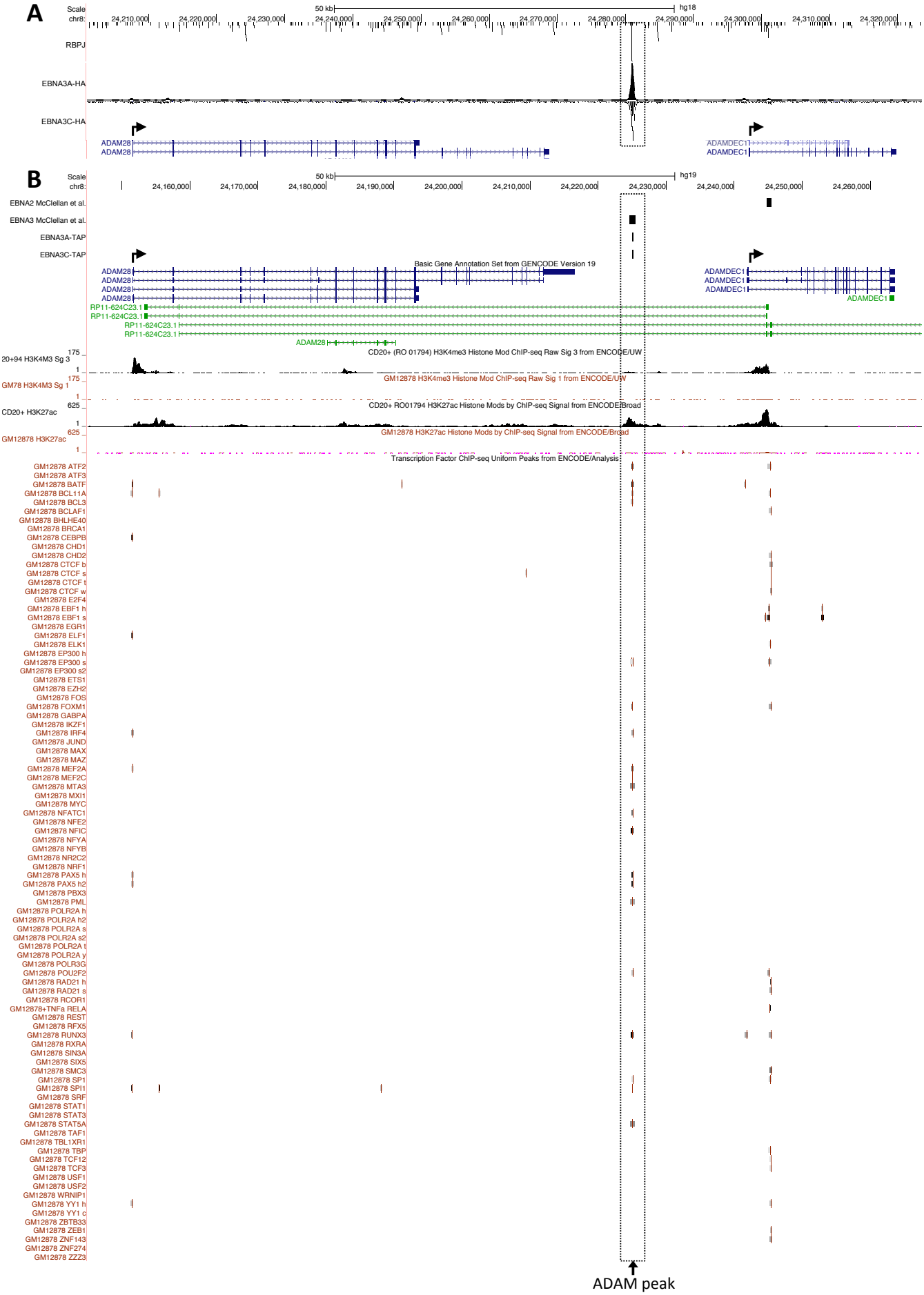

A
B
RBPJ
EBNA3A-HA
EBNA3C-HA
EBNA2 McClellan et al.
EBNA3 McClellan et al.
EBNA3A-TAP
EBNA3C-TAP
Basic Gene Annotation Set from GENCODE Version 19
ADAM28
ADAMDEC1
RP11-624C23.1
CD20+ (RO 01794) H3K4me3 Histone Mod ChIP-seq Raw Sig 3 from ENCODE/UW
GM12878 H3K4me3 Histone Mod ChIP-seq Raw Sig 1 from ENCODE/UW
CD20+ RO01794 H3K27ac Histone Mods by ChIP-seq Signal from ENCODE/Broad
GM12878 H3K27ac Histone Mods by ChIP-seq Signal from ENCODE/Broad
Transcription Factor ChIP-seq Uniform Peaks from ENCODE/Analysis
ADAM peak

Figure 7.9: EBNA3A and EBNA3C binding sites on *ADAM28/ADAMDEC1* locus are co-occupied by multiple transcription factors.

(**A**) UCSC genome browser overview of *ADAM28/ADAMDEC1* genomic locus showing ChIP-seq tracks for RBPJ (Zhao et al., 2011a), EBNA3A-HA (Schmidt et al., 2014) and EBNA3C-HA (Jiang et al., 2013) aligned to hg18. Occupancy levels are merely to indicate the position and are not to scale. (**B**) Hg19 view of same genomic locus as in (**A**) showing ChIP-seq tracks for EBNA2 (McClellan et al., 2013), EBNA3 (McClellan et al., 2013), EBNA3A-TAP, EBNA3C-TAP, the TSSs of *ADAM28* and *ADAMDEC1* (horizontal arrows) on the positive strand (left to right), long non-coding RNA RP11-624C23.1 on the negative strand (right to left), ENCODE ChIP-seq tracks for H3K4me3 and H3K27ac in $CD20^+$ primary B cells (black) or LCL GM12878 (red) and ENCODE ChIP-seq peaks for various transcription factors in GM12878 LCL. The EBNA3A and EBNA3C binding site is highlighted by inclusion in a dashed box.

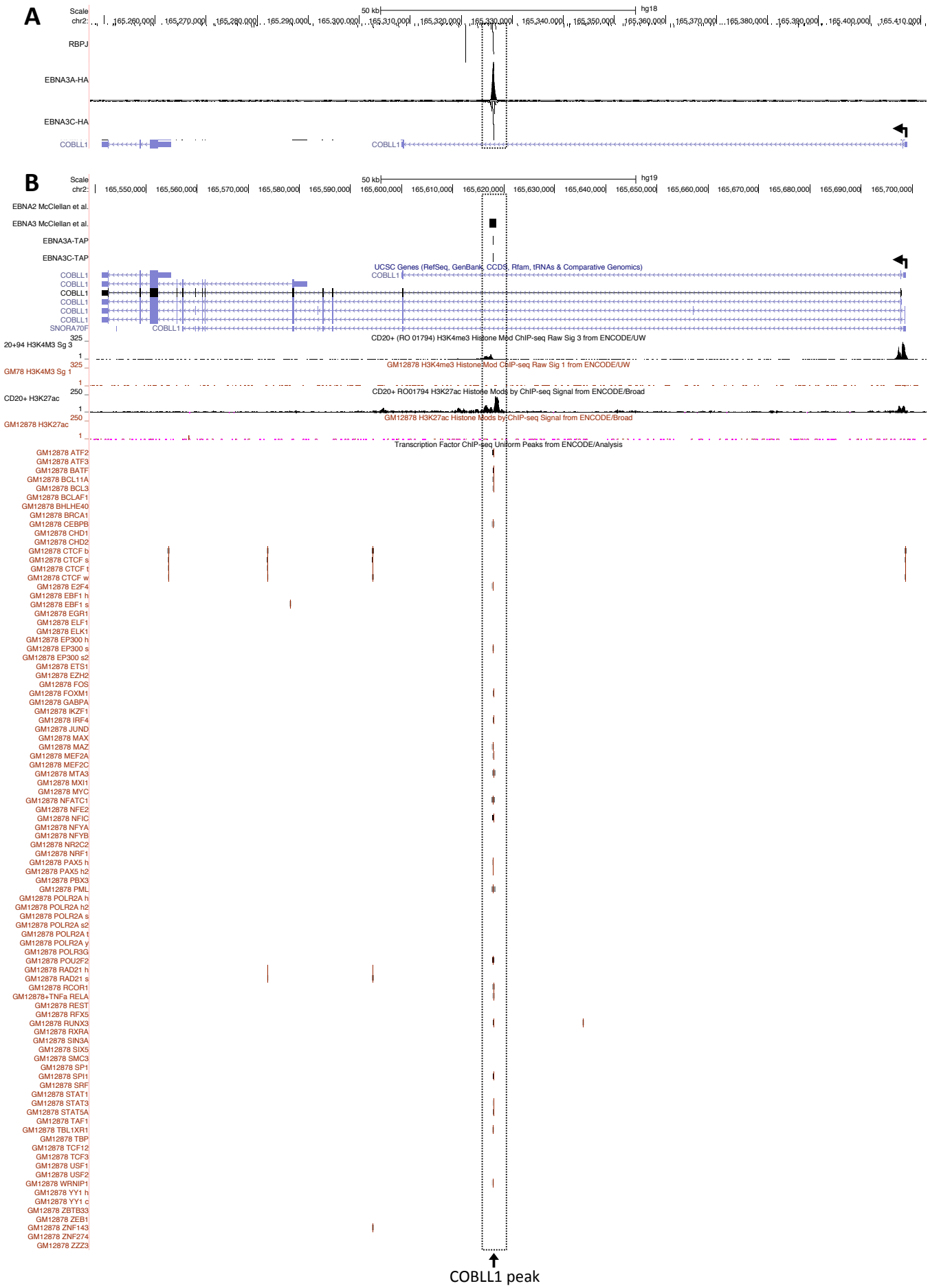

A
B
RBPJ
EBNA3A-HA
EBNA3C-HA
COBLL1
EBNA2 McClellan et al.
EBNA3 McClellan et al.
EBNA3A-TAP
EBNA3C-TAP
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
SNORA70F
20+94 H3K4M3 Sg 3
CD20+ (RO 01794) H3K4me3 Histone Mod ChIP-seq Raw Sig 3 from ENCODE/UW
GM78 H3K4M3 Sg 1
GM12878 H3K4me3 Histone Mod ChIP-seq Raw Sig 1 from ENCODE/UW
CD20+ H3K27ac
CD20+ RO01794 H3K27ac Histone Mods by ChIP-seq Signal from ENCODE/Broad
GM12878 H3K27ac
GM12878 H3K27ac Histone Mods by ChIP-seq Signal from ENCODE/Broad
Transcription Factor ChIP-seq Uniform Peaks from ENCODE/Analysis
COBLL1 peak

Figure 7.10: EBNA3A and EBNA3C binding sites on *COBLL1* locus are co-occupied by multiple transcription factors.

(**A**) UCSC genome browser overview of *COBLL1* genomic locus showing ChIP-seq tracks for RBPJ (Zhao et al., 2011a), EBNA3A-HA (Schmidt et al., 2014) and EBNA3C-HA (Jiang et al., 2013) aligned to hg18. Occupancy levels are merely to indicate the position and are not to scale. (**B**) Hg19 view of same genomic locus as in (**A**) showing ChIP-seq tracks for EBNA2 (McClellan et al., 2013), EBNA3 (McClellan et al., 2013), EBNA3A-TAP, EBNA3C-TAP, the TSS (horizontal arrow), ENCODE ChIP-seq tracks for H3K4me3 and H3K27ac in $CD20^{+}$ primary B cells (black) or LCL GM12878 (red) and ENCODE ChIP-seq peaks for various transcription factors in GM12878 LCL. The EBNA3A and EBNA3C binding site is highlighted by inclusion in a dashed box.

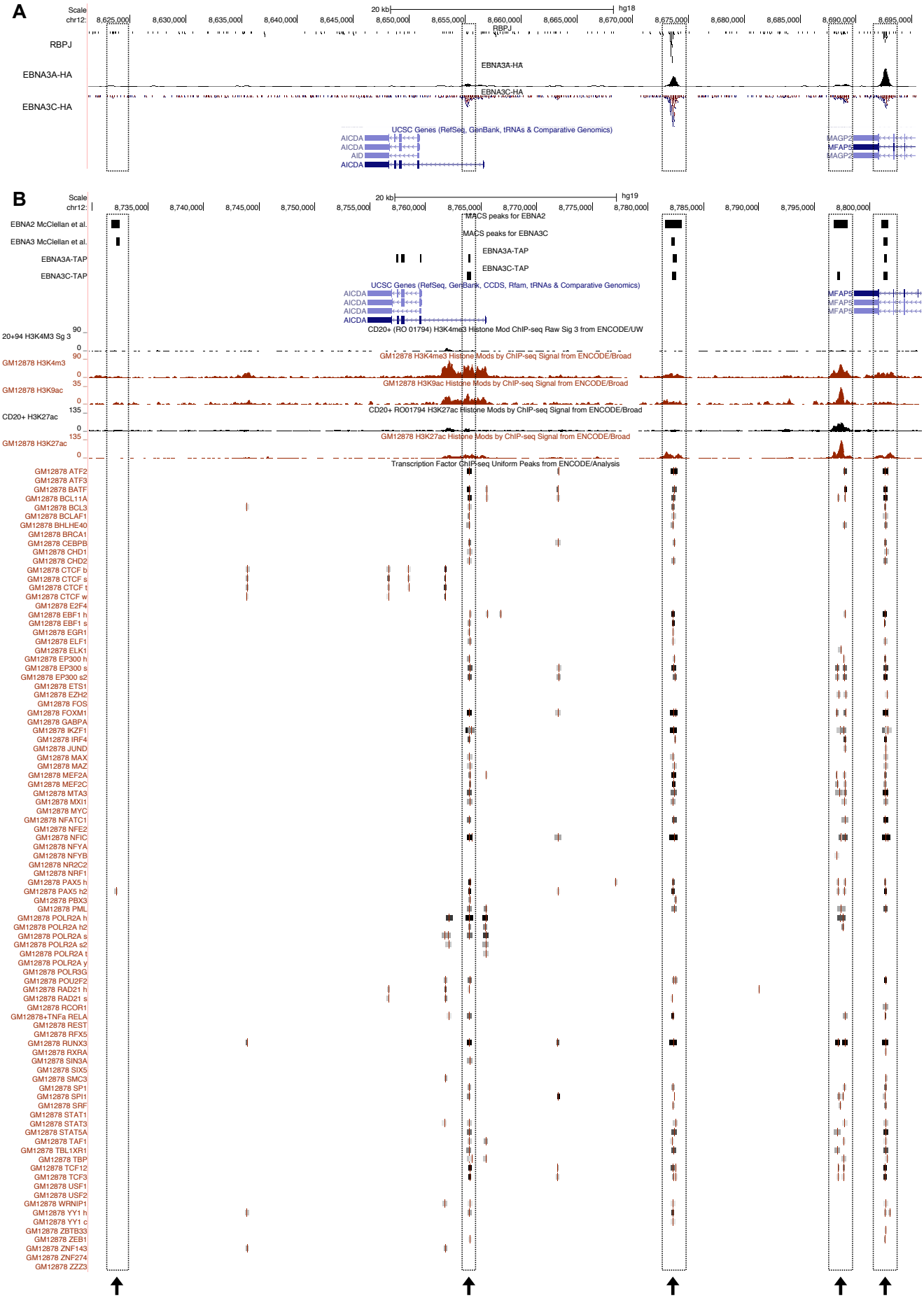

A
Scale
chr12:
20 kb
hg18
RBPJ
EBNA3A-HA
EBNA3C-HA
UCSC Genes (RefSeq, GenBank, tRNAs & Comparative Genomics)
AICDA
AID
B
Scale
chr12:
20 kb
hg19
EBNA2 McClellan et al.
EBNA3 McClellan et al.
EBNA3A-TAP
EBNA3C-TAP
MACS peaks for EBNA2
MACS peaks for EBNA3C
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
MFAP5
20+94 H3K4M3 Sg 3
CD20+ (RO 01794) H3K4me3 Histone Mod ChIP-seq Raw Sig 3 from ENCODE/UW
GM12878 H3K4m3
GM12878 H3K4me3 Histone Mods by ChIP-seq Signal from ENCODE/Broad
GM12878 H3K9ac
GM12878 H3K9ac Histone Mods by ChIP-seq Signal from ENCODE/Broad
CD20+ H3K27ac
CD20+ RO01794 H3K27ac Histone Mods by ChIP-seq Signal from ENCODE/Broad
GM12878 H3K27ac
GM12878 H3K27ac Histone Mods by ChIP-seq Signal from ENCODE/Broad
Transcription Factor ChIP-seq Uniform Peaks from ENCODE/Analysis
III
II
IV
V
VI

Figure 7.11: EBNA3A, EBNA3C and EBNA2 binding sites on *AICDA* locus are co-occupied by multiple transcription factors.

(**A**) UCSC genome browser overview of *AICDA* genomic locus showing ChIP-seq tracks for RBPJ (Zhao et al., 2011a), EBNA3A-HA (Schmidt et al., 2014) and EBNA3C-HA (Jiang et al., 2013) aligned to hg18. Occupancy levels are merely to indicate the position and are not to scale. (**B**) Hg19 view of same genomic locus as in (**A**) showing ChIP-seq tracks for EBNA2 (McClellan et al., 2013), EBNA3 (McClellan et al., 2013), EBNA3A-TAP, EBNA3C-TAP, ENCODE ChIP-seq tracks for H3K4me3 and H3K27ac in $CD20^{+}$ primary B cells (black) or LCL GM12878 (red) and ENCODE ChIP-seq peaks for various transcription factors in GM12878 LCL. The EBNA2, EBNA3A and EBNA3C binding site at the AICDA regulatory regions II, III, IV, V and VI are highlighted by inclusion in dashed boxes.