

Analysis of host gene regulation by Epstein-Barr Virus Nuclear Antigen 3B (EBNA3B)

By

Guiyi Ho

submitted to

Imperial College London

for the degree of

Doctor of Philosophy

2016

**Department of Virology
Imperial College London
St Mary's Campus
Norfolk Place
London W2 1PG**

Declaration of Originality:

I declare that, unless otherwise stated, all work presented in this thesis are my own, and material that are obtained from other sources are referenced appropriately.

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.

Abstract

EBNA3A and EBNA3C are oncoproteins that are essential for EBV-driven transformation of B-cells, whereas *in vivo* studies have demonstrated that EBNA3B has tumour suppressive properties. Previous exon-microarray studies performed by our group have identified a subset of cellular genes that may be regulated by EBNA3B.

In order to investigate the gene regulatory mechanisms of EBNA3B, I established a Doxycycline (Dox)-inducible EBNA3B system by transfecting EBNA3B-null (3BKO) lymphoblastoid cell lines (LCLs) with a pRTS vector expressing EBNA3B with Dox treatment. Using this system I tested 43 predicted EBNA3B-regulated genes using Taqman Low Density Array cards and confirmed that EBNA3B rapidly and robustly represses 7 genes, and induces 1 gene. I have additionally established LCLs expressing epitope (FLAG-Strep-Strep)-tagged EBNA3B and used these for Chromatin Immunoprecipitation linked to high throughput sequencing (ChIP-seq). ChIP-seq analyses uncovered 1931 EBNA3B peaks and revealed that EBNA3B localises distal to its predicted target genes, in enhancer regions that are positive for active histone marks (H3K4me1 $\pm$ H3K27Ac). EBNA3B frequently overlaps with EBNA3C, but EBNA3B-only binding sites are found within currently uncharacterised genomic regions. EBNA3B binds within the same looped chromatin domain as *TERT* and *TNFSF10* genes, and qPCR analyses revealed that induction of EBNA3B using the Dox-inducible system robustly alters the expression of these genes.

Focussed analyses revealed that EBNA3B binds to a putative enhancer 17kb upstream of the *TERT* promoter and strongly represses the expression of *TERT*, a gene encoding the catalytic subunit of the telomere-lengthening telomerase complex. EBNA3B mediates gene repression by facilitating the removal of active histone marks (H3K27Ac, H3K9Ac and H3K4me3) from the *TERT* promoter and EBNA3B-bound region and by recruiting Polycomb Repressive Complex to deposit repressive H3K27me3 at the *TERT* promoter. However, deletion of EBNA3B does not have an obvious effect on stabilising telomere lengths during the first 120 days of B-cell outgrowth.

Acknowledgements

I would like to express my heartfelt gratitude towards my supervisor Prof. Martin Allday and co-supervisor Dr Rob White for their guidance, editorial comments and patience throughout my whole PhD journey. I am also indebted to Dr Kostas Paschos for his help and supervision, especially for the ChIP-seq experiments. Special thanks also go to Dr Quentin Bazot for his continuing work on TERT, for being a great friend, and for spreading his enthusiasm for science (and EBV) within the lab. I would also like to express my deepest appreciation for Ms Gillian Parker, for her effective management of the lab and for her reassurances and genuine interest in the welfare of everyone.

To Jens and Agnieszka, you are the best fellow EBV PhD mates I could ever have hoped for. Thank you for keeping me going with your encouragement and always having my back. I wish you both successful careers. To Adam, thank you for the interesting conversations in the tissue culture room – I will miss our little chats about the most random things in the world (and curly fries). To Christine, thank you for bringing a positive vibe to the lab – especially in stressful times.

To my parents: 老爸老妈，多谢你们那么多年的鼓励，我终于成功了！I would also like to thank my siblings for being so supportive. And of course, I would not have succeeded without dear Jared – thank you for always being there and reminding me that we learn and become stronger after each hurdle. Next, I would like to thank each and every friend, medicine school colleague and house-mate who has in his/her own way contributed to my knowledge and resilience.

Last, but not least, I would also like to express gratitude to A*STAR for funding my MBBS/PhD studies.

Abbreviations

3AKO	EBV BAC or virus with EBNA3A deletion
3ARev	EBNA3A-deleted EBV BAC or virus with EBNA3B reverted back to WT
3B	EBNA3B
3BKO	EBV BAC or virus with EBNA3B deletion
3BRev	EBNA3B-deleted EBV BAC or virus with EBNA3B reverted back to WT
3CKO	EBV BAC or virus with EBNA3C deletion
3CRev	EBNA3C-deleted EBV BAC or virus with EBNA3B reverted back to WT
4HT	4-Hydroxytamoxifen
ALT	Alternate lengthening of telomeres
Amp	Ampicillin
APS	Ammonium Persulphate
BAC	Bacterial artificial chromosome
BART	BamHI-A Rightward Transcript
BIM	Bcl-2 Interacting Mediator of cell death
BL	Burkitt Lymphoma
bp	Base Pair
BSA	Bovine Serum Albumin
CBF-1	C-Promoter Binding Factor-1
CDKI	Cyclin-Dependent Kinase Inhibitor
ChIP	Chromatin Immunoprecipitation
ChIP-seq	Chromatin Immunoprecipitation linked to high throughput sequencing
Cm	Chloramphenicol
Cp	EBV C promoter
CpG	Cytosin-Phosphate-Guanin
CtBP	C-Terminal Binding Protein
CTL	Cytotoxic T Lymphocyte
ddH2O	Double Distilled Water
DLBCL	Diffuse large B-cell lymphoma
DMSO	Dimethyl Sulphoxide
DNA	Deoxyribonucleic Acid
Dox	Doxycycline
E.coli	Escherichia Coli
EBER	EBV-Encoded RNA
EBF1	Early B-cell factor 1
EBNA	EBV Nuclear Antigen
EBNA3C-HT	EBV BAC or virus containing EBNA3C-HT fusion
EBV	Epstein-Barr Virus
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediamine Tetra-Acetic Acid
eGFP	enhanced Green Fluorescent Protein
ER3B	EBV BAC or virus with N-terminal fusion of ERT2 domain at EBNA3B
ERT2	4-hydroxytamoxifen-conditional, selective oestrogen receptor

EtBr	Ethidium Bromide
EtOH	Ethanol
EZH 1/2	Enhancer Of Zeste Homolog 1/2
FACS	Fluorescence-Activated Cell Sorting
FCS	Foetal Calf Serum
FS	FLAG-Strep-Strep tag
FS3B	EBV BAC or virus with N-terminal fusion of FLAG-Strep-Strep tag at EBNA3B
GC	Germinal Centre
h	Hour
H3K27me3	Trimethylation of lysine 27 at histone 3
H3K4me3	Trimethylation of lysine 4 at histone 3
HDAC	Histone Deacetylase
HL	Hodgkin Lymphoma
HPV	Human Papilloma Virus
HSV	Herpes Simplex Virus
hu-NSG	Mice with reconstituted components of the human immune system
IM	Infectious Mononucleosis
IP	Immunoprecipitation
IRES	Internal ribosome entry site
Kan	Kanamycin
kDa	Kilodaltons
KO	Knock-out
KPB	Potassium Phosphate Buffer
LB	Luria-Bertani Broth
LCL	Lymphoblastoid Cell Line(s)
LCV	Lymphocryptovirus
LID	Lipofectin/integrin-targeting peptide/DNA
LMP	Latent Membrane Protein
LP	Leader Protein
LUC	Luciferase
min	Minutes
miR	MicroRNA
mQH ₂ O	milliQ purified water
mRNA	Messenger RNA
NCoR	Nuclear receptor co-repressor
NFkB	Nuclear Factor Kappa-light-chain-enhancer of activated B cells
NGFR	Truncated nerve growth factor receptor
NK	Natural Killer
Notch-IC	Intracellular Notch receptor
NP40	Nonidet P-40
NPC	Nasopharyngeal Carcinoma
ORF	Open reading frame
oriP	Origin Of Replication
PAGE	Poly-Acrylamide Gel Electrophoresis
PBL	Peripheral Blood Lymphocytes
PBS	Phosphate Buffered Saline
PBS-T	Phosphate Buffered Saline Containing 0.5% V/V Tween 20

PcG	Polycomb Group proteins
PCR	Polymerase Chain Reaction
PDL	Population doubling level
PFGE	Pulsed field gel electrophoresis
PMSF	Phenyl Methyl Sulphonyl Fluoride
Pol II	RNA Polymerase II
PRC	Polycomb Repression Complex
PTLD	Post transplant Lymphoproliferative Disorder
Qp	EBV Q promoter
qPCR	Quantitative PCR
Rb	Retinoblastoma Protein
RBPJk/CBF1	Recombinant Binding Protein J Kappa/C-Promoter Binding Factor-1
REF	Rat embryonic fibroblasts
Rev	Revertant
RIPA	Radioimmunoprecipitation
RNA	Ribonucleic Acid
RPMI	Roswell Park Memorial Institute Medium
RT	Real-Time
RTS	Reverse transactivator silencer-1
SB	Superbroth
SD	Standard Deviation
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
seq	Sequencing
Ser 5	Serine 5 (of Pol II)
Sin3A	SIN3 transcription regulator family member A
SOC media	Super optimal broth with catabolite repression
STET	Solution made of sucrose, titron, EDTA and Tris
SV40	Simian Virus 40
TAP	Tandem Affinity Purification
TBE	Tris-Borate-EDTA
TE	Tris-EDTA
TEMED	Tetramethylethylenediamine
TERT	Telomerase reverse transcriptase
Tet	Tetracycline
TF	Transcription Factor
TFBS	Transcription Factor binding site
TLDA	Taqman quantitative real-time PCR Low Density Array
TNF	Tumour Necrosis Factor
TNFRSF10A/TRAIL1	Tumour necrosis factor-related apoptosis-inducing ligand receptor 1
TNFRSF10D/TRAILR4	Tumour necrosis factor-related apoptosis-inducing ligand receptor 4
TNFSF10/TRAIL	Tumour necrosis factor-related apoptosis-inducing ligand
TPA	Tetradecanoylphorbol Acetate
TR	Terminal Repeats
TSS	Transcription Start Site

Tween 20	Polyoxyethylenesorbitanmonolaurate
v/v	Volume by Volume
w/v	Weight by Volume
WB	Western Blotting
Wp	EBV W promoter
WT	Wild-Type

Table of Contents

Abstract	3
Acknowledgements	4
Abbreviations	5
Table of Contents.....	9
List of Figures	15
List of Tables.....	18
 1. Introduction.....	 19
1.1. Discovery of Epstein-Barr Virus	19
1.2. EBV virion and genome structure.....	19
1.3. Transmission and life cycle of EBV	22
1.3.1. Latency-associated gene transcription programmes	23
1.3.2. Persistence of EBV	24
1.3.3. Transcription of EBV latent genes.....	26
1.4. EBV and its associated diseases	29
1.5. EBV latency-associated proteins.....	32
1.6. Epstein-Barr virus Nuclear Antigen 3s (EBNA3s).....	34
1.6.1. Genome organisation and structure	34
1.6.2. Structural features and homology between the EBNA3 family of proteins	36
1.6.3. The EBNA3s are transcriptional regulators that affect host gene expression .	38
1.6.4. EBNA3A & EBNA3C are potent oncogenes.....	39
1.6.5. EBNA3B as a viral tumour suppressor	41
1.7. Mechanisms of gene regulation by the EBNA3s	44
1.7.1. EBNA3s can interact with EBNA2 and RBPJ	44
1.7.2. EBNA3s can bind proximally to genes and direct gene repression by interacting with PcG proteins.....	49
1.7.3. The EBNA3s are able to bind and co-ordinate gene regulation distally from gene TSS.....	50
1.7.4. EBNA3-mediated gene repression follows a complex temporal sequence	53
1.8. Telomerase, TERT and EBV.....	54
1.8.1. Structure and function of telomeres	54
1.8.2. Effects of telomere length and TERT on EBV-driven B-cell immortalisation ...	55
1.9. TRAIL and its receptors	56
1.10. Background to project.....	57
1.11. Aims	59

2	Materials and Methods	61
2.1	Solutions and Buffers	61
2.2	General cloning strategy and protocols for DNA manipulation	64
2.2.1	Extraction of Bacterial plasmid	65
2.2.2	Restriction Digestion	67
2.2.3	Agarose gel electrophoresis.....	68
2.2.4	Pulsed field gel electrophoresis (PFGE).....	68
2.2.5	DNA Gel extraction	69
2.2.6	DNA ligation.....	69
2.2.7	Oligonucleotide linker ligation	69
2.2.8	Transformation of DNA into bacteria	70
2.3	Construction of recombinant EBV BAC	71
2.3.1	Design of cloning constructs	71
2.3.2	Cloning of pKovKan shuttle vector used for BAC recombineering	72
2.3.3	RecA-mediated homologous recombination.....	74
2.4	Construction of ERT2-EBNA3B-citrine construct.....	77
2.5	Cell culture.....	80
2.5.1	Cell lines	80
2.5.2	Cell culture media	83
2.5.3	Maintenance of cells lines in culture.....	84
2.5.4	Freezing and thawing cells.....	84
2.5.5	Harvesting cells	85
2.6	Transfection protocols (mammalian cells)	86
2.6.1	LID vector delivery	86
2.6.2	Electroporation of 3BKO LCLs	87
2.7	Low molecular weight DNA extraction.....	88
2.8	Generation of HEK-293 virus producers and production of recombinant EBV	89
2.8.1	Generation of clonal cell lines	89
2.8.2	Production of recombinant EBV virions and virus titration	90
2.9	EBV infection of B-cells.....	90
2.10	Isolation of PBMCs	91
2.11	Isolation of cells via magnetic sorting	92
2.11.1	Primary B-cells.....	92
2.11.2	Dox-inducible mCherry positive cells	92
2.12	FACs analysis.....	93
2.12.1	Cell Staining protocol	93

2.12.2	Gating Strategy	93
2.13	Western Blot	94
2.14	Reverse transcription quantitative PCR (RT-qPCR)	96
2.14.1	PCR quantitation using SYBR Green	97
2.14.2	PCR quantitation using TaqMan® Gene expression assays	98
2.15	TaqMan® Gene Expression Array Cards	99
2.16	PCR amplification across splice junctions	100
2.17	Confocal microscopy.....	102
2.18	Immunoprecipitation.....	102
2.19	ChIP and ChIP-seq analyses	103
2.19.1	Chromatin Immunoprecipitation – epitope-tagged EBNA3s.....	103
2.19.2	Chromatin precipitation – other factors.....	105
2.19.3	ChIP-qPCR.....	105
2.19.4	High-throughput ChIP sequencing	106
2.20	TRAP assay.....	108
2.21	Telomere length assay.....	109
3	Strategies for validating EBNA3B regulated genes.....	112
3.1	Introduction	112
3.2	Differential expression of CXCL10 & CXCL9 between WT and 3BKO LCLs may be a result of clonal selection	113
3.3	Construction of a recombinant EBV BAC encoding 4HT-conditional EBNA3B (ER3B)	118
3.4	ER3B virus producer cell lines are generated by transfection into HEK-293 cells	122
3.4.1	mRNA transcripts for ER3B are present in HEK-293 ER3B clones.	125
3.4.2	ERT2-EBNA3B-citrine construct can be expressed in isolation.....	126
3.5	Recombinant ER3B viruses are produced and used to generate viable ER3B LCLs	128
3.6	Generation and validation of doxycycline-inducible EBNA3B system in 3BKO LCLs.	133
3.6.1	Doxycycline successfully induced the expression of all RTS-vector components including EBNA3B in 3BKO LCLs.	136
3.6.2	Increased dosage of doxycycline results in a dose-response increase in expression of EBNA3B.....	140
3.6.3	Long-term treatment of doxycycline maintains complementation of EBNA3B in transfected LCLs but shows no effect on selected predicted target genes.	142
3.6.4	Magnetic sorting for NGFR+ cells increases yield of Dox-responsive cells but these cell populations tend to drift back to unsorted phenotypes over time.	143

3.6.5	Four-day dose treatment of Dox (with sorting for NGFR) consistently shows regulation of ITGAL by EBNA3B.	149
3.6.6	Analysis of 43 potential EBNA3B regulated genes using Taqman Low-Density Array card to identify possible rapidly regulated targets of EBNA3B.....	152
3.7	Overview.....	154
4	Determination of Genome-wide EBNA3B binding	159
4.1	Introduction	159
4.2	Construction of recombinant epitope-tagged EBNA3B (FS3B) BAC	160
4.3	Transfection of HEK-293 cells and generation of FS3B virus producer cell lines.	162
4.4	Validation of LCLs established using recombinant FS3B virus.	164
4.4.1	Episome rescued from FS3B LCLs showed correct integrity.....	164
4.4.2	FS3B protein is present at slightly higher levels than wild-type EBNA3B in established FS3B LCLs and BL31s.....	164
4.4.3	Low molecular weight DNA isolated from FS3B LCLs were verified by sequencing	166
4.4.4	Expression levels of several potential EBNA3B targets are closer to wild-type levels than those of 3BKO LCLs.	167
4.4.5	FS3B LCLs are validated for chromatin immunoprecipitation using ITGAL as a positive EBNA3B binding site.....	168
4.5	ChIP coupled with high-throughput sequencing using FS3B LCLs.....	169
4.5.1	Quality control of sequencing data as defined by the ENCODE guidelines. .	171
4.5.2	EBNA3B binds to fewer regions in the genome than EBNA3A, EBNA3C and EBNA2.	176
4.5.3	EBNA3B co-localises mainly with EBNA3C, but also with EBNA3A and EBNA2	179
4.5.4	Majority of the EBNA3B binding peaks do not co-localise with EBNA3A, 3C or EBNA2.	179
4.5.5	EBNA3B binds distally to its putative regulated genes.	181
4.5.6	The distribution of EBNA3B across genome segments alters considerably depending on its co-localization with other EBNA3s and EBNA2	182
4.5.7	EBNA3B co-localise with histone modifications associated with active promoters/enhancers, when also co-localised with other EBNA3s	189
4.5.8	EBNA3B enriched sites are also bound by specific transcription factors	191
4.5.9	EBNA3B regulated genes are found within the same Contact Domains as EBNA3B peaks	194
4.5.10	EBNA3B peaks which localise within the same Contact Domains as their target genes exhibit similar co-localisation patterns with histone modifications and transcription factors, regardless of the nature of the gene regulation.....	197

4.6	List of genes identified as rapidly targeted by EBNA3B and with EBNA3B binding sites within the same contact domain.	199
4.7	Overview.....	201
5	Focused analysis of EBNA3B regulated genes to validate microarrays and ChIP-seq and provide some insight into mechanisms involved.....	203
5.1	Introduction.....	203
5.2	TLDA card study of the DOX system is validated using RT-qPCR – <i>TERT</i> , <i>TRAIL</i> , <i>TRAILR1</i> & <i>TRAIL R4</i>	204
5.3	Detailed study of the regulation of <i>TERT</i> by EBNA3B	204
5.3.1	Primary B-cell infections demonstrated the kinetics of <i>TERT</i> regulation.....	204
5.3.2	Binding of EBNA3s upstream of hTERT confirmed by ChIP-qPCR.....	208
5.3.3	Regulation of hTERT by EBNA3B may involve histone modifications	208
5.3.4	Physiological implications of TERT regulation by EBNA3B	214
5.4	Regulation of TRAIL, TRAILR1 & TRAILR4 by EBNA3B.....	219
5.5	Overview.....	221
6	Discussion.....	223
6.1	The Dox-inducible EBNA3B system can be used to identify bona fide EBNA3B target genes that are robustly and rapidly regulated	223
6.1.1	Magnetically sorted NGFR+ RTS-3B/LUC cells ensures high purity of cells expressing the RTS vector and have to be re-sorted at every time-point	223
6.1.2	The Dox-inducible RTS-3B experiments identifies genes that are robustly and rapidly repressed or induced in the presence of high levels of EBNA3B	225
6.1.3	Variation exists between the response of different RTS-3B cell lines	227
6.1.4	Conclusions and future work with the Dox-inducible RTS-3B system.....	228
6.2	ChIP-seq has identified properties of EBNA3B binding regions that may be important for its function as a transcriptional regulator.....	228
6.2.1	EBNA3B is mainly a transcriptional repressor	229
6.2.2	EBNA3C might be an important mediator in the co-ordinated gene regulation with EBNA3B	230
6.2.3	EBNA3B might function in a different manner to EBNA3A and EBNA3C	232
6.2.4	Data analyses using high stringency produces similar results as this study .	233
6.2.5	EBNA3B co-localises with EBNA2 at strong enhancers and TSS of regulated genes	235
6.2.6	EBNA3B is involved in chromatin looping mechanisms between TSS and distal regulatory elements	235
6.2.7	Overview and future work on EBNA3B ChIP-seq.....	236
6.3	Regulation of <i>TERT</i> by EBNA3B.....	236

6.3.1	EBNA3B regulates <i>TERT</i> expression by mediating epigenetic modifications at the <i>TERT</i> promoter	236
6.3.2	EBNA3B may facilitate TERT regulation by mediating chromatin looping of a distal enhancer to the promoter region	237
6.3.3	Heterogeneity in telomere length in cell cultures could mask subtle changes early in the B-cell outgrowth	238
6.3.4	Telomeres of EBV-infected B-cells may shorten until they arrive at their proliferative crisis despite the <i>TERT</i> status of the cells.....	240
6.3.5	Future work on the repression of <i>TERT</i> by EBNA3B and its physiological implications	241
6.3.6	Overview on repression of <i>TERT</i> by EBNA3B.....	244
6.4	Regulation of <i>TRAIL</i> , <i>TRAILR1</i> and <i>TRAILR4</i> by EBNA3B	244
6.5	Final remarks	245
7	Bibliography	248
Appendix A: Expression of EBNA3C is compromised when EBNA3B is tagged at the C-terminus.....		266
Appendix B: ChIP experiments using monoclonal rat antibodies.....		267
Appendix C: Table of genes that are differentially regulated by EBNA3B in both LCLs and BL31.....		268
Appendix D: Pile-up of ChIP-seq reads at the <i>TERT</i> , <i>TRAIL</i> & <i>TRAILR1/R4 loci</i>		269
Appendix E: α-FLAG ChIP using Donor 13 LCLs expressing epitope-tagged EBNA3B or EBNA3C		270
Appendix F: Comparisons of EBNA3 binding profile.....		271
Appendix G: EBNA3 co-localisation with TFBS.....		272

List of Figures

Figure 1.1. Schematic showing the linear representation of the EBV genome.....	21
Figure 1.2. Schematic showing how EBV may establish persistence in an immunocompetent host by exploiting the normal B-cell differentiation pathway.....	25
Figure 1.3. Schematic of the circularised EBV genome and transcripts initiated by W, C and Q promoters.....	28
Figure 1.4. Schematic showing the splicing of EBNA3 mRNA.....	35
Figure 1.5. Schematic representation of the structural domains of EBNA3s.....	37
Figure 1.6. Schematic showing (A) Chromatin-Immunoprecipitation & (B) Chromatin Conformation Capture (3C) techniques.....	48
Figure 2.1. Targeting vectors used for BAC recombineering.....	73
Figure 2.2. pKovKan Δ Cm-based shuttle vectors used for BAC recombineering.....	74
Figure 2.3. Schematic showing the incorporation of ERT2-tag at the N-terminus of EBNA3B in B95.8 EBV-BAC via RecA-homologous recombination.....	78
Figure 2.4. (A) Oligonucleotide linker and (B) vectors used to create ERT2-EBNA3B-citrine expression construct.....	79
Figure 3.1. Differential expression of (A) CXCL10 and (B) CXCL9 between wild-type and 3BKO LCLs which are established from 3 independent donors (D1, D2 & D3).....	115
Figure 3.2. Relative gene expression of (A) CXCL10 & (B) CXCL9 in LCLs derived from Donor 11 and Donor 13.....	115
Figure 3.3. ChIP for histone modification changes at CXCL10 TSS in WT, 3BKO and 3BRev LCLs.....	116
Figure 3.4. Timecourse following the expression of CXCL10 and CXCL9 over the first 40 days of LCL outgrowth.....	117
Figure 3.5. Schematic showing the insertion of ERT2 tag at the N-terminus of EBNA3B through homologous recombination.....	120
Figure 3.6. Generation of BAC DNA from wild-type (B95.8) to ER3B fusion protein.....	121
Figure 3.7. Verification of maxipreps of ER3B BAC before transfection into HEK-293 cells.....	123
Figure 3.8. Verification of HEK-293 ER3B cell lines.....	124
Figure 3.9. Accurately spliced transcripts for ER3B were detected by agarose gel electrophoresis.....	125
Figure 3.10. ERT2-EBN3B-citrine construct is stabilised with addition of 4HT 48h after transient transfection with pGY15.....	127
Figure 3.11. Verification of ER3B LCLs.....	130
Figure 3.12. PCR amplification of cDNA or low molecular weight DNA recovered from ER3B LCLs and HEK-293 virus producers.....	132
Figure 3.13. Schematic map of Doxycycline (Dox)-inducible RTS-vectors used to transfect 3BKO LCLs.....	135
Figure 3.14. EBNA3B is expressed on the induction of RTS-3B 3BKO LCLs with 50ng/ml of doxycycline after 4 days.....	138
Figure 3.15. Expression of mCherry relative to truncated NGFR in RTS3-3B/LUC LCLs before and after induction with Dox.....	139
Figure 3.16. Dose-response induction of 3BKO RTS3-3B and RTS3-LUC LCLs.....	141
Figure 3.17. Western blot showing EBNA3B protein levels with increasing dosages of Dox.....	142

Figure 3.18. Expression of EBNA3B and its candidate regulated genes in D1.5 3BKO RTS-3B and RTS-LUC LCLs after a 23-day treatment with 20ng/ml of Dox.	144
Figure 3.19. Magnetically-sorted RTS-3B/LUC cell populations from 2 different donors (D1.5 and JK1) drift to unsorted phenotype over time.	146
Figure 3.20. Short-term Dox induction (20ng/ml) of RTS2-3B/LUC cells using ITGAL expression as a positive control gene.	148
Figure 3.21. Flow cytometry histograms of D13.4 RTS-3B cells pre and post-sort for NGFR positivity, four days after Dox treatment.	150
Figure 3.22. Validation of 4-day Dox titration in NGFR-sorted D13.4 RTS-3B/LUC 3BKO LCLs using <i>ITGAL</i> as a positive control.....	151
Figure 3.23. Heatmap showing the expression of candidate EBNA3B-regulated genes in D1.5 and D13.4 3BKO RTS-3B LCLs after induction with increasing dosages of Dox.....	155
Figure 3.24. Heatmap highlighting any differences in gene expression of candidate EBNA3B-regulated genes between parental D1.5 and D13.4 3BKO LCLs and dox-inducible RTS-3B LCLs.	156
Figure 4.1. Generation of FS3B BAC from wild-type B95.8.	161
Figure 4.2. Validation of HEK-293 FS3B cell lines.....	163
Figure 4.3. Validation of FS3B LCLs.	165
Figure 4.4. PCR amplification of the FS3B domain of low molecular weight DNA purified from FS3B LCLs.	166
Figure 4.5. Expression of CD28, SMAD1 & ITGAL in FS3B LCLs.	167
Figure 4.6. Validation of FS3B LCLs for use in ChIP experiments.....	168
Figure 4.7. Agarose gel electrophoresis of pooled input and ChIP samples from D11 C8 FS3B LCLs.	169
Figure 4.8. Bioanalyzer analysis (Agilent 2100 Bioanalyzer) of size-selected ChIP and Input samples from D11 C8 FS3B LCL.	170
Figure 4.9. Cross-correlation of ChIP samples taken from D11 C8 FS3B and WT LCLs. ..	173
Figure 4.10. Calculation of normalised strand co-efficient (NSC) and relative strand co-efficient (RSC).....	174
Figure 4.11. Aligned pile-up of reads from ChIP-seq analysis of FS3B and WT LCLs.....	175
Figure 4.12. EBNA3A, EBNA3B, EBNA3C, pan-specific EBNA3 and EBNA2 enriched peaks in hg19.....	177
Figure 4.13. Overlap of EBNA3A, EBNA3B and EBNA3C ChIP-seq peaks (from our lab) with pan-specific EBNA3 ChIP-seq peaks (McClellan <i>et al</i> , 2012).....	178
Figure 4.14. Distribution of EBNA3B peaks according to their co-localisation with EBNA3A, EBNA3C and/or EBNA2.....	180
Figure 4.15. Overlap of EBNA3B peaks with EBNA3A, EBNA3C, & EBNA2 peaks (McClellan <i>et al.</i> , 2013).....	180
Figure 4.16. Distance of EBNA3B binding sites to transcriptional start sites of (A) all genes and (B) potential EBNA3B target genes (White <i>et al</i> , 2010).	181
Figure 4.17. Distribution of (A) EBNA3B, (B) EBNA3A, (C) EBNA3C, (D) EBNA3 (pan-specific) and (E) EBNA2 ChIP-seq peaks across different genome segments.	186
Figure 4.18. Distribution of EBNA3B binding peaks within different genome states. (part 1)	187
Figure 4.19. Heat map comparing the co-occupancy of EBNA3s and EBNA2 with histone marks.....	190
Figure 4.20. Co-localisation of transcription factor binding sites (TFBS) with EBNA3B peaks.	192

Figure 4.21. Contact Domains co-localising with both EBNA3B enriched peaks and potential EBNA3B-regulated genes.	196
Figure 4.22. Histone modifications associated with EBNA3B peaks in Contact Domains. .	197
Figure 4.23. Co-localisation of TFBS with EBNA3B peaks found in Contact Domains.....	198
Figure 5.1. Verification of TLDA findings using qPCR analyses of (A) <i>TERT</i> , (B) <i>TRAILR1</i> , (C) <i>TRAILR4</i> and (D) <i>TRAIL</i> mRNA transcripts.....	205
Figure 5.3. Outgrowth of recombinant EBV-infected primary B-cells.	207
Figure 5.2. High yield of CD20+ B-cells was obtained using magnetic sorting for cells expressing CD19+ marker.	207
Figure 5.4. EBNA3A, 3B and 3C all bind at a distal regulatory region upstream of <i>TERT</i> transcriptional start site.	209
Figure 5.5. Levels of mRNA transcripts found in WT, EBNA3BKO and 3BRev LCLs derived from Donor 11 and Donor 13.....	210
Figure 5.6. Enrichment of (A) H3K27me3 (B) H3K4me3 (C) H3K9Ac & (D) H3K27Ac at <i>TERT</i> promoter and EBNA3B binding sites.....	211
Figure 5.7. ChIP experiments showing (A) PolII S5 phosphorylation, and (B) SUZ12 at <i>TERT</i> promoter and EBNA3 binding sites.	212
Figure 5.8. Detection of telomerase activity using the TRAPeze® RT (S7710) kit.	215
Figure 5.9. Telomerase activity in 3BKO and 3BRev LCLs of two different donor backgrounds.	216
Figure 5.10. Telomere length during B-cell outgrowth of cells in 4 different genetic backgrounds.	218
Figure 5.11. Outgrowth of recombinant EBV-infected primary B-cells – <i>TRAIL</i> , <i>TRAILR1</i> , <i>TRAILR4</i>	220
Figure 5.12. Figure showing the enrichment of EBNA3A, EBNA3B, EBNA3C and EBNA2 at the gene loci of (A) <i>TRAIL</i> and (B) <i>TRAILR1/R4</i>	222
Figure 6.1. Overlap of genes which are differentially expressed in 3A/3B/3CKO recombinant EBV-infected BL31 cells as compared to WT EBV-infected BL31s.	231
Figure 6.2. View of RBPJ, EBNA2, and EBNA3A/3B/3C binding at the <i>TERT</i> locus using Genome Browser (accessible at genome.ucsc.edu .)	242

List of Tables

Table 1.1. EBV latent gene transcription programme and the associated cell types and B-cell malignancies. Adapted from Price and Luftig (2014).	24
Table 2.1. Solutions used in this study	61
Table 2.2. Vectors and BACs used in this study	64
Table 2.3. Restriction enzyme digests.....	67
Table 2.4. Cell lines used in project.....	81
Table 2.5 Cell culture conditions for each cell line	83
Table 2.6. Titres in Green Raji Units (GRU)/ μ l of viruses used in this study	91
Table 2.7. Recipe for stacking and resolving SDS-PAGE gels	95
Table 2.8. Antibodies used for Western Blots.....	96
Table 2.9. Primer sequences used in this study	98
Table 2.10. RT-qPCR probe assays used in this study	99
Table 2.11. Primers used for PCR amplification across the splice junction between U-exon and EBNA3B.....	100
Table 2.12. TaqMan® Assays (Applied Biosystems) used in this study.....	101
Table 2.13. Antibodies and IgG controls used for ChIP	105
Table 2.14. Primers used for ChIP-qPCR analysis	106
Table 2.15. ChIP-seq and Hi-C data sets used in this study.....	108
Table 2.16. Primers used for quantification of telomere length	110
Table 3.1. Table showing virus titres determined by Raji cell infections.....	129
Table 3.2. Criteria used for determining EBNA3B regulation	153
Table 3.3. EBNA3B-repressed genes as determined by the Dox-inducible EBNA3B system.	157
Table 3.4. EBNA3B-induced genes as determined by Dox-inducible EBNA3B system.	158
Table 4.1. Table showing virus titres of FS3B HEK-293 cell lines as determined by Raji cell infections. Green Raji Units (GRU) are determined by counting the number of Raji cells infected per μ l of virus used.	162
Table 4.2. Distribution of reads obtained from high-throughput sequencing of ChIP samples.	172
Table 4.3. Description of genome segments as described by the ENCODE consortium. ..	183
Table 4.4. Transcription factor motifs associated with EBNA3B binding regions.	194
Table 4.5. EBNA3B-regulated genes which are found within the same contact domain (CD) as EBNA3B peaks.	200

1. Introduction

1.1. Discovery of Epstein-Barr Virus

In 1958, surgeon Denis Burkitt reported a jaw tumour that affected children in equatorial Africa, which came to be known as Burkitt's lymphoma (BL) (Burkitt, 1958). Using biopsies obtained from BL tumour tissue, Epstein and colleagues successfully established a liquid suspension culture of BL cells and subsequently visualised herpesvirus-like particles in this culture via electron microscopy (Epstein *et al.*, 1964a). Despite being morphologically similar to other known herpesviruses, such as herpes simplex virus (HSV), the virus particles identified by Epstein and colleagues were smaller than HSV, and did not exhibit the same characteristics as other known members of the herpes virus family in various isolation and immunofluorescence experiments (Epstein and Achong, 1968; Epstein *et al.*, 1964b; Henle and Henle, 1966). This newly discovered herpesvirus was named Epstein-Barr virus (EBV), after Anthony Epstein and Yvonne Barr, who first cultured the BL cells. Further, antibodies against specific EBV antigens in the cultured BL cells were found in the BL patients, and the EBV deoxyribonucleic acid (DNA) was discovered in BL cell lines and in biopsies from tumours of nasopharyngeal carcinoma (NPC) (Epstein, 1975; Gunven *et al.*, 1970; zur Hausen *et al.*, 1970; Henle *et al.*, 1969). Hence, EBV was recognised as a novel herpesvirus and was proposed to be the first human tumour virus to be discovered.

1.2. EBV virion and genome structure

EBV, which is now also known as Human Herpesvirus 4 (HHV-4), belongs to the Lymphocryptovirus (LCV) genus that is within the Gammaherpesvirinae subfamily of the Herpesviridae family. As the only known LCV to infect humans, EBV has probably co-evolved with *Homo sapiens* and its ancestors over millennia of evolutionary development (Ehlers *et al.*, 2010).

The EBV virion contains a linear, approximately 172kb double-stranded DNA genome that is wrapped around a central protein core. The DNA core is enclosed within a 162-

capsomere icosahedral nucleocapsid, which is surrounded by a layer of tegument that comprises both viral-encoded and host cell proteins. This tegument layer is in turn coated by a lipid envelope studded with glycoproteins such as gp350/220, gB, gH/gL and gp42, that are essential for host cell recognition and entry (Johannsen *et al.*, 2004; Longnecker *et al.*, 2013).

The linear EBV genome contains multiple tandem repeats of the same 538bp sequence (TRs) on either ends. Upon infection, these TRs are involved in the circularization of the linear genome (Kintner and Sugden, 1979; Zimmermann and Hammerschmidt, 1995), hence the EBV genome exists as several extra-chromosomal plasmids (episomes) in the host cell during the latent phase (Hung *et al.*, 2001). The number of TRs in the episomes generally remains unchanged during DNA replication in latently infected host cells. Hence, daughter cells which arose from monoclonal expansion of a single parent cell would contain episomes having the same number of TRs as their progenitor cell. Likewise, variable number of TRs in episomes obtained from a group of cells would suggest polyclonal expansion of multiple latently infected cells (Raab-Traub and Flynn, 1986; Walling *et al.*, 2004).

In addition to the TRs, there are also internal direct repeat regions (IR) of varying sizes (between 100 to over 3000bp units) interspersed within the EBV genome (Cheung and Kieff, 1982). The large IR1 repeat region (3kb units) divides the EBV genome into short (US) and long (UL) unique sequence domains (Figure 1.1), and together these two unique domains contain nearly all of the EBV protein-coding sequences (Longnecker *et al.*, 2013).

The structure of the DNA and mechanisms of DNA replication vary according to the stage of the life cycle of viral infection. Depending on whether the virus is going through latent or lytic reproduction, initiation of replication can take place at different origins of replication within the genome (at OriP for latent phase and two OriLyt for lytic phase replication) (Figure 1.1), which are each associated with distinct viral and cellular replicative proteins and mechanisms (Hammerschmidt and Sugden, 2013). The 1.7kb OriP region

contains an approximately 120bp-long dyad symmetry (DS) element and an array of tandem repeats (known as family of repeats, FR), that is around 650bp in total length, and located 1kb upstream of the DS. The DS element comprises binding sites for the viral latent protein EBNA1, which recruits the replication machinery to initiate the bi-directional synthesis of a daughter DNA episome. The FR element is tethered via EBNA1 to host chromosomes, which ensures that the circularised EBV genome can propagate during host cell replication hence minimising episome loss in proliferating host cells during latent infection (Hammerschmidt and Sugden, 2013).

In contrast, DNA replication in the lytic phase involves a different set of viral proteins (such as BZLF1, BALF5, BALF2, BMRF1, BBLF4, BSLF1 and BBLF2/3) interacting with and recruiting replication complexes at the OriLyt, resulting in several hundred-fold amplification of the EBV genome (Hammerschmidt and Sugden, 2013). Aside from laboratory strains with genomic deletions (such as the lab strain B95.8 - Figure 1.1), two OriLyts are present and spaced approximately 100kb apart within the genome of all EBV strains. The end products of lytic replication are long concatemers that contain multiple EBV genome units arranged in series. These are cleaved, processed and packed as linear genomes into the EBV virions (Hammerschmidt and Sugden, 2013; Longnecker *et al.*, 2013).

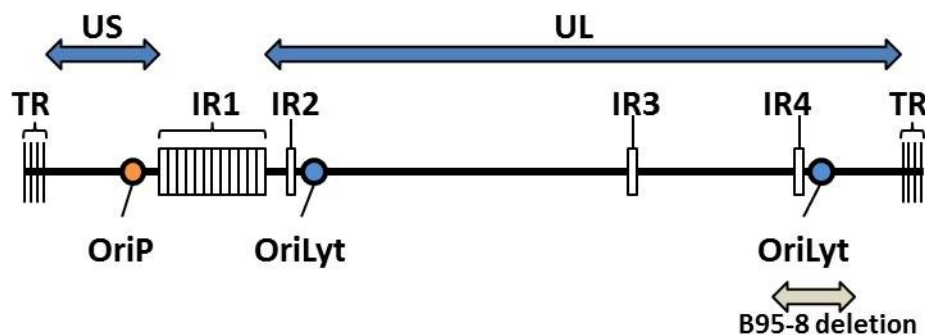


Figure 1.1. Schematic showing the linear representation of the EBV genome.

Unique short (US) and long (UL) genome regions are shown, along with terminal repeats (TR) and internal direct repeats (IR1-4). The positions of the origins of replication in latent infection (OriP) and lytic infection (OriLyt) are indicated with orange and blue circles respectively. The deletion in the B95.8 genome of EBV from a marmoset B-cell line are indicated by the double-headed grey arrow. Adapted from Longnecker *et al.* (2013) with modifications.

1.3. Transmission and life cycle of EBV

EBV is orally transmitted via salivary contact with infected individuals. The site of primary infection is still a contentious issue, since EBV appears to exhibit dual tropism for epithelial cells and B-cells depending on the glycoprotein content – notably levels of gp42 – on its viral envelope (Borza and Hutt-Fletcher, 2002; Shannon-Lowe and Rowe, 2011; Wang *et al.*, 1998b). EBV establishes latency within B-cells, and in contrast, infection of epithelial cells is generally associated with lytic activation and virus production (Temple *et al.*, 2014). It is postulated that EBV enters the tonsillar epithelium of Waldeyer's ring in the oropharynx, where the virus replicates and amplifies via initiation of lytic cycle in the epithelial cells. As a lymphotropic virus, EBV subsequently infects the surrounding naïve B-cells in the lymphoid tissue of the Waldeyer's ring and establishes persistence in the host by exploiting the normal physiology of B-cell differentiation (further details in section 1.3.2; Thorley-Lawson and Gross, 2004). In another model, EBV is thought to directly infect B-cells in the oropharynx and establish persistence in memory B-cells. Some of the latently infected memory B-cells may become re-activated, differentiating into plasma cells and entering the lytic phase (Faulkner *et al.*, 2000; Laichalk and Thorley-Lawson, 2005). These re-activated B-cells either release viruses directly into the saliva, or produce viruses that infect and initiate the lytic reproductive phase in the surrounding epithelial cells of the oropharynx, thus amplifying the viral load in the saliva for transmission to another host (Faulkner *et al.*, 2000; Laichalk and Thorley-Lawson, 2005; Longnecker *et al.*, 2013).

EBV exhibits a tropism for B-cells due to the interaction between the viral envelope glycoproteins gp350/220 and the B-cell CD21 surface receptor (Young *et al.*, 2007a). Endocytosis and fusion of EBV envelope with the host cell membrane is triggered by the binding of another viral envelope glycoprotein gp42 to the human leucocyte antigen class II molecule (HLA) on B-cells and facilitated by three other EBV glycoproteins gB and gH/gL (Hutt-Fletcher, 2007; Li *et al.*, 1997; Young *et al.*, 2007a). EBV can also fuse with the cell membrane and enter CD21-negative lymphocytes and epithelial cells without endocytosis,

although this is an inefficient process which appears less well understood (Hutt-Fletcher, 2007; Shannon-Lowe and Rowe, 2014). Once the virus is internalised, the viral capsid is ejected and releases the linear EBV genome into the cell nucleus, where it circularises and propagates in latently-infected B-cells as episomes.

Sporadically, the differentiation of infected B-cells into plasma cells is thought to re-activate lytic replication (Laichalk and Thorley-Lawson, 2005), expressing over 80 lytic proteins in a temporal sequence, starting with the activation of immediate-early transactivator genes (BZLF1 & BRLF1) which initiate viral DNA replication at the OriLyt, and ending with late gene products such as envelope glycoproteins and structural proteins (Young *et al.*, 2007b). The EBV virions are assembled in the nucleus and acquire their viral envelope by budding through the host cell membrane.

1.3.1. Latency-associated gene transcription programmes

During latent infection, EBV can undergo different highly restricted latency-associated gene transcription programmes which may be initiated by different latency promoters (Qp, Wp & Cp) depending on the location and differentiation state of the B-cells (Table 1.1; Klein *et al.*, 2007; Price and Luftig, 2014). A maximum of 9 proteins, comprising 6 EBV nuclear antigens (EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C and EBNA-LP) and 3 integral latent membrane proteins (LMP1, LMP2A and LMP2B), may be expressed during latency. Some of these proteins, including EBNA2, EBNA3A, EBNA3C and LMP1 are now widely accepted as growth-promoting and anti-apoptotic oncogenic proteins (Young and Rickinson, 2004). In addition, two non-polyadenylated non-coding viral RNAs (EBER1 & EBER2) and untranslated microRNAs from the BamHI A rightward transcripts (BARTs) and BHRF1 transcripts (miR-BHRF1) of the EBV genome are also expressed during the latent phase. In recent years, it has been demonstrated that a number of canonical lytic genes, including BHRF1 and BALF1 which are viral homologues of anti-apoptotic human BCL genes, and a number of EBV immunoevasins such as BCRF1 (homologue of human IL10) and BNLF2a [inhibitor of the transporter associated with antigen processing (TAP)] (Jochum *et al.*, 2012),

are expressed very early after infection at the “pre-latent phase” (reviewed in Price and Luftig, 2014). It is postulated that the latency-associated genes together promote growth and contribute towards the establishment of latency in the B-cells by preventing apoptosis.

Table 1.1. EBV latent gene transcription programme and the associated cell types and B-cell malignancies. Adapted from Price and Luftig (2014).

Latent gene transcription programme	Viral genes expressed	Associated differentiated B-cell type	Associated malignancies
Latency 0	EBER1&2, BARTs	Peripheral non-dividing memory B-cell	-
Latency I	EBNA1, EBER1&2, BARTs	Dividing peripheral memory B-cell	BL
Latency II	EBNA1, LMP1, LMP2A/B, EBER1&2, BARTs	Germinal centre cells	HL
Latency III	EBNA 1, EBNA2, EBNA3A/B/C, EBNA-LP, LMP1, LMP2A/B, EBER1 &2, BARTs, miR-BHRF1	Activated naïve B-cells	AIDS-associated DLBCL, PTLD

1.3.2. Persistence of EBV

The widely accepted model for EBV persistence *in vivo* proposes that the virus is able to persist and evade host immune surveillance by taking advantage of the normal B-cell differentiation pathway. In short, this involves the controlled stepwise restriction of the full panel of viral protein expression from latency III in infected activated B-lymphoblasts to latency 0 in the circulating memory B-cells, where none of the EBV latent proteins are expressed (reviewed by Thorley-Lawson *et al.*, 2013). Briefly, EBV infects naïve B-cells in the tonsils, driving them to differentiate into activated and proliferating B-lymphoblasts which express all known EBV latent genes (latency III). This is mostly a transient event in an immunocompetent host because most of the infected and activated B-lymphoblasts would be recognised and removed by EBV-specific cytotoxic CD8⁺ T-lymphocytes (CTLs), and to a smaller extent CD4⁺ T-cells and Natural Killer (NK) cells (Hislop, 2015; Hislop *et al.*, 2007; Long *et al.*, 2011). The CD8⁺ response is most commonly generated against the immuno-dominant EBNA3 and LMP2 proteins, whereas the CD4⁺ cell response is predominantly triggered by EBNA1 (Hislop *et al.*, 2007; Khanna *et al.*, 1992; Münz *et al.*, 2000; Murray *et*

al., 1992; Rickinson and Moss, 1997). Infected B-lymphoblasts which escape immune surveillance migrate to the germinal centre (GC) in lymph nodes where they switch to the latency II gene expression pattern whereby the expression of EBNA2, EBNA-LP and the EBNA3s are shut down (Roughan and Thorley-Lawson, 2009). In the GC, the B-lymphoblasts differentiate into centroblasts and centrocytes which are protected from apoptosis by the cell survival signals mimicked by LMP1 and LMP2A (Panagopoulos *et al.*, 2004; Rastelli *et al.*, 2008). The surviving EBV-infected centrocytes differentiate into memory B-cells and enter the latency 0 programme, where all viral protein expression is repressed. When these infected memory B-cells divide, they express EBNA1 (latency I), which allows the viral genome to propagate (see section 1.2). Hence, EBV is able to stay undetectable by the host immune system through silent persistence in the latently-infected memory B-cells. Viral re-activation and entry into the lytic phase occurs when these memory B-cells are activated and differentiate into plasma cells (Laichalk and Thorley-Lawson, 2005). Steady levels of latent EBV-infected B-cells (0.5-50 per million B-cells) are maintained in the infected healthy individual via continuous immune surveillance from circulating CTLs and antibodies against EBV proteins (Khan *et al.*, 1996; Thorley-Lawson, 2001).

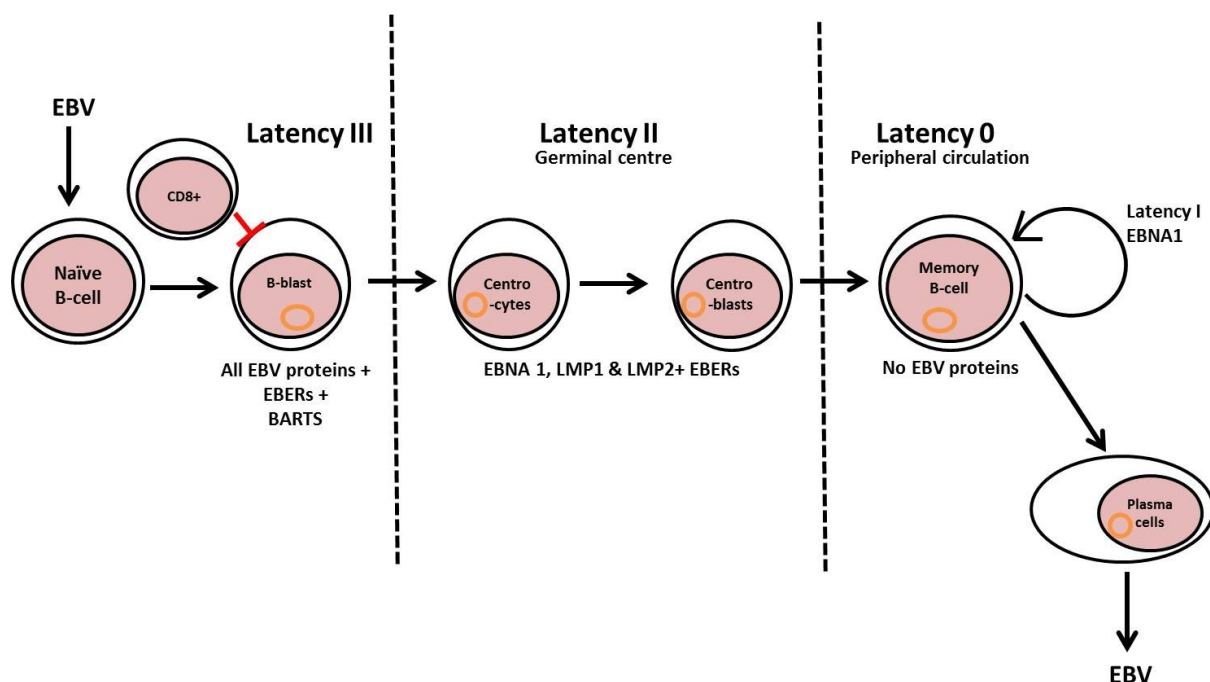


Figure 1.2. Schematic showing how EBV may establish persistence in an immunocompetent host by exploiting the normal B-cell differentiation pathway.

Refer to text in section 1.3.2 for details and references. Adapted from Allday (2009)

In vitro, EBV is able to transform primary B-cells into continuously proliferating lymphoblastoid cell lines (LCLs) (Henle *et al.*, 1967; Pope *et al.*, 1968). Each cell in an LCL contains multiple copies of circular episomes and expresses the full set of latency-associated genes (latency III) (Longnecker *et al.*, 2013; Young *et al.*, 2007b). Hence, LCLs are used frequently as a model system to study the effects of EBV latent proteins on host gene expression and their potential biological effects.

1.3.3. Transcription of EBV latent genes

The transition from early infection into different latency programmes (Table 1.1) relies on the usage of different viral latency promoters to initiate viral gene transcription. Once the EBV genome enters the nucleus, it circularises, fusing both ends of terminal repeats together (Figure 1.3). Within the first 24h post EBV infection, the expression of EBNA2 and EBNA-LP are initiated via the viral W promoter (Wp), which is present within each IR1 repeat fragment (Price and Luftig, 2014; Young *et al.*, 2007b). Importantly, EBNA2 interacts with DNA-bound host transcription factors such as RBPJ (a.k.a. RBP-Jk, CBF-1, and CSL) and strongly activates the viral C promoter (Cp), LMP1, LMP2A and LMP2B promoters (Laux *et al.*, 1994; Rickinson, 1990; Sung *et al.*, 1991; Wang *et al.*, 1990; Zimmer-Strobl and Strobl, 2001). In addition, EBNA-LP supports the activation of these promoters by EBNA2, with the exception of LMP2A promoter (Kempkes and Ling, 2015; Price and Luftig, 2014). EBNA-LP can potentiate EBNA2-mediated activation of LMP1/LMP2B promoter and Cp in transient reporter gene assays (Harada and Kieff, 1997; Nitsche *et al.*, 1997). It was also shown that evolutionarily conserved regions within the W repeat domains of EBNA-LP were essential for the co-operation of EBNA-LP with EBNA2 in transactivating Cp and LMP1 promoter (Nitsche *et al.*, 1997; Peng *et al.*, 2000). The activity of Wp declines as EBNA2 accumulates and transcriptional activity switches from Wp to the upstream Cp (Woisetschlaeger *et al.*, 1990, 1991).

While the primary role of EBNA1 is in episome maintenance (see section 1.2), it can also play a role in transactivating viral Cp. Notably, transcripts of EBNA1 can be detected

prior to the switch from Wp to Cp (Schlager and Speck, 1996), and EBNA1 has been shown to transactivate Cp and LMP1 promoter by binding to an enhancer region within the OriP (Gahn and Sugden, 1995; Sugden and Warren, 1989). Hence it is likely that EBNA1 co-operates with EBNA2 to induce the switch of transcriptional activity from Wp to Cp (Schlager and Speck, 1996).

The Cp initiates the transcription of a set of long (around 100kb) primary polycistronic transcripts that are alternatively spliced and translated to generate EBNA-LP, EBNA2, EBNA3A, EBNA3B, EBNA3C and EBNA1 (Lieberman, 2013; Price and Luftig, 2014). How this type III latency may switch to more restricted expression patterns in type II and type I latency (Table 1.1) remains unclear, although this may be related to DNA and histone methylation – and hence silencing and reduced availability – of important transcription factor binding sites at the Cp and the LMP promoters (Lieberman, 2013). In the highly restricted type I and also type II latency, Cp is repressed and EBNA1 is expressed from the Q promoter (Q promoter; Lieberman, 2013).

1.4. EBV and its associated diseases

Shortly after its discovery, in 1968, Henle and colleagues serendipitously discovered that EBV is the causative agent for infectious mononucleosis (IM, commonly known as glandular fever) (Henle *et al.*, 1968), a benign and generally self-limiting lymphoproliferative disease. In developed or industrialised nations, EBV infection occurs more frequently in adolescents and young adults, and may result in IM in some cases (Henle *et al.*, 1969; Luzuriaga and Sullivan, 2010; Thorley-Lawson and Gross, 2004). IM patients typically present with a febrile illness that is characterised by fatigue, lymphadenopathy and splenomegaly with atypical lymphocytosis in blood films, which are the clinical manifestations of the immune response to the initial EBV infection (Luzuriaga and Sullivan, 2010). In contrast, in developing countries, primary infection with EBV is generally asymptomatic and occurs within the first few months of life or early childhood (Hsu and Glaser, 2000; Luzuriaga and Sullivan, 2010). In recent years, increasing evidence has surfaced to suggest that EBV may also play a role in the aetiology of autoimmune conditions such as multiple sclerosis, systemic lupus, reactive arthritis and Sjögren's syndrome (Draborg *et al.*, 2016).

An intact immune system is vital for bringing EBV infection under control. EBV infection in immunocompromised patients – such as those who have undergone organ transplants, or those co-infected with HIV or malaria – increase the risk of these individuals developing EBV-associated immunoblastic B-cell lymphomas such as diffuse large B-cell lymphoma (DLBCL) and post-transplant lymphoproliferative disease (PTLD) (Saha and Robertson, 2011; Young and Rickinson, 2004). These immunoblastic lymphoproliferative diseases are generally derived from EBV-infected B-lymphoblasts that express the full complement of latency III growth-promoting genes (Table 1.1) (Rowe *et al.*, 1991; Thomas *et al.*, 1990) such as EBNA2, EBNA3A, EBNA3C, EBNA-LP and LMP1 which together predispose the B-cells to malignant transformation and proliferation by upregulating anti-apoptotic proteins (such as BCL2) and cytokines and repression of pro-apoptotic proteins

[such as BCL2-like protein 11 (BIM) and cyclin-dependent kinase inhibitors (CDKIs) such as p16^{INK4A}]. These B-lymphoblasts are less likely to be removed by CTLs in a background of reduced immune surveillance, so encouraging the development of lymphoproliferative disease (Rowe and Zuo, 2010; Straus *et al.*, 1993). The treatments for these conditions include reducing levels of immunosuppressive drugs, administration of rituximab (anti-CD20 monoclonal antibodies) and the adoptive transfer of EBV-specific CTLs (Petrara *et al.*, 2015; Styczynski *et al.*, 2016).

The first experimental evidence that supported the notion of EBV as an oncogenic virus is its ability to efficiently infect and transform foetal leukocytes into proliferating “immortalised” cells *in vitro* (Henle *et al.*, 1967; Pope *et al.*, 1968). As the first human tumour virus that is described in literature, EBV is currently associated with several malignant conditions even in immunocompetent hosts. Over the past five decades, through various epidemiological studies involving comparisons of viral antibodies, viral load and the presence of EBV DNA in tumour samples, EBV is now aetiologically linked to B-cell lymphomas [BL and Hodgkin’s Lymphoma (HL)], EBV-associated T- and NK-cell lymphomas, and epithelial carcinomas such as NPC and gastric adenocarcinoma (Young *et al.*, 2009).

Similar to the situation in immunodeficient individuals, EBV may be contributing indirectly as a predisposing factor that, together with a myriad of other cellular host cell factors, makes the infected B-cell susceptible to transformation even in immunocompetent hosts (Thorley-Lawson and Allday, 2008). This is supported by the observation that only a small number of infected immunocompetent individuals develop EBV-associated malignancies whereas the large majority would not suffer any adverse health consequences. Importantly, several EBV-associated malignancies occur at higher incidences within defined geographical locations (such as BL in equatorial Africa) and/or certain ethnic populations (such as NPC in Southern Chinese and South-East Asians), which suggests other essential co-factors may be involved in EBV-associated carcinogenesis.

In all cases of BL, the translocation of the *MYC* proto-oncogene into an actively transcribed Immunoglobulin locus in B-cells of the germinal centre is deemed to be a causative event of lymphomagenesis (Klein, 1983). The prevalence of EBV-positive BL is geographically restricted; less than 25% of the BL in some regions of the world are associated with EBV, in contrast to the 90% of BL in equatorial Africa which are EBV-positive (Thorley-Lawson and Allday, 2008). There are strong indications that the co-infection of endemic malaria and EBV may be a reason for this geographical restriction of EBV-associated BL (reviewed in Thorley-Lawson and Allday, 2008). On top of its overall immunosuppressive effects, infection with malaria may up-regulate activation-induced cytidine deaminase (AID), the enzyme responsible for introducing double strand breaks and DNA nicks during class switch recombination and somatic hyper-mutation within the GC, which can predispose the B-cells to acquire to *MYC* translocations and generalised genomic instability (Robbiani *et al.*, 2015; Torgbor *et al.*, 2014). It has been proposed that the epigenetic repression of pro-apoptotic proteins (such as BIM) and senescence-inducing proteins such as p16^{INK4A} and p14^{ARF} by EBV oncogenic proteins EBNA3A and EBNA3C expressed during the transient expansion of EBV-infected B-lymphoblasts (Anderton *et al.*, 2008; Maruo *et al.*, 2011; Paschos *et al.*, 2012; Skalska *et al.*, 2010) would later prevent apoptosis of B-cells harbouring *MYC* translocations within the GC, hence allowing these cells to survive and proliferate indefinitely (Allday, 2009; Thorley-Lawson and Allday, 2008). The infected B-cells leave the GC and differentiate into memory B-cells, but do not exit the cell cycle due to continuous *MYC*-driven proliferation, hence establishing BL tumours which will classically express latency I gene transcription programmes (Table 1.1) (Thorley-Lawson and Allday, 2008).

Similarly, only 40-50% of HL tumours are EBV-positive, and EBV-positive HLs are more common in lower socio-economic populations, young children and the elderly (Jarrett *et al.*, 1991; Pallesen *et al.*, 1991; Vockerodt *et al.*, 2014). The Hodgkin/Reed Sternberg (HRS) cells that make up the malignant component of EBV-positive HL tumours lack

functional B-cell receptors (BCR), which, under normal physiological conditions, would have designated these cells for apoptosis within the GC. The EBV-infected HRS cells display characteristics of GC cells, and express the latency II transcription programme (Table 1.1) (Deacon *et al.*, 1993), which makes it likely that these cells arise from EBV-infected B-cells of GC origin and, due to unknown reasons, are blocked from further differentiation (Thorley-Lawson and Gross, 2004). Although the exact role of EBV in the pathogenesis of HL is still unclear, it is speculated that the transformation process of EBV-infected cells involves LMP1-mediated activation of cellular survival signalling pathways such as the nuclear factor-kappa B (NF- κ B), phosphatidylinositol 3-kinase/Protein kinase-B (PI3K/Akt) and Janus kinase/signal transducers and activators of transcription (JAK/STAT) pathways (Tao *et al.*, 2006; Vockerodt *et al.*, 2014). LMP2A, another EBV latent protein, likely contributes to driving the proliferation of BCR-negative HRS cells by mimicking the pro-survival signals of an activated BCR (Caldwell *et al.*, 1998; Vockerodt *et al.*, 2014). Hence, in HL, EBV may contribute to lymphomagenesis by preventing the apoptosis of BCR-negative GC cells through the provision of appropriate pro-survival signals that would normally require antigen signalling and T-cell help. In EBV-negative HL cells, it is likely that other cellular mechanisms are at play in the pathogenesis.

1.5. EBV latency-associated proteins

It is clear that EBV latent proteins play important roles not only in viral persistence within the healthy host, but also in the pathogenesis of several B-cell malignancies. Reverse genetics experiments have revealed that EBNA2, EBNA3A, EBNA3C, LMP1 and EBNA-LP are critical for efficient *in vitro* transformation of B-cells (Ring, 1994; Young and Murray, 2003). In addition, the demonstration of distinct restricted patterns of latent gene expression in each EBV-related malignancy has highlighted the importance of studying the contributions of each EBV latent protein in the tumorigenic process. Understanding their specific functions and interactions with cellular factors may help illuminate strategies for targeted treatment against EBV-positive lymphomas and lymphoproliferative diseases. In particular, our group

is interested in the functions of the EBNA3 family of proteins, and their interactions with EBNA2.

EBNA2 is a powerful viral transactivator of host and viral genes and is functionally similar to an activated intracellular Notch receptor (Notch-IC). The membrane-bound Notch receptor comprises extracellular (Notch-EC), transmembrane (Notch-TM) and intracellular domains (reviewed by Kopan and Ilagan, 2009). Briefly, ligands for the Notch receptor (such as members of the Delta-like and Jagged families) bind to the extracellular domain, which induce the TNF- α ADAM metalloprotease converting enzyme (TACE)-mediated cleavage of the Notch-EC (Brou *et al.*, 2000). Subsequently, removal of Notch-EC triggers the cleavage of Notch-IC from Notch-TM, releasing the activated Notch-IC (Mumm *et al.*, 2000; Strooper *et al.*, 1999). The activated Notch-IC translocates to the nucleus where it binds to DNA binding protein RBPJ (a.k.a. RBP-Jk, CBF-1, and CSL). RBPJ is a transcription factor that is conventionally thought to bind constitutively to many regulatory elements and in the absence of activated Notch-IC, will form complexes with co-repressors such as histone deacetylases (HDACs) to repress expression of cellular genes. In the presence of activated Notch-IC, RBPJ recruits the co-activator mastermind (MAM) which facilitates further recruitment of other co-activators such as histone acetyltransferase p300 (Oswald *et al.*, 2001).

EBNA2 does not bind directly to DNA, but is targeted to EBNA2-responsive promoters and enhancers through its interaction with DNA binding proteins such as RBPJ (Grossman *et al.*, 1994; Waltzer *et al.*, 1994). Like Notch-IC, EBNA2 binds to the repressive domain of RBPJ, and replaces the co-repressor complexes with its transactivation domains (reviewed by Zimmer-Strobl and Strobl, 2001). In this way, EBNA2 mimics the Notch signalling pathway and activates the expression of host genes that are important in controlling growth and B-cell activation (such as *MYC*, *CD23* and *CD21*), (Cordier *et al.*, 1990; Kaiser *et al.*, 1999; Wang *et al.*, 1987; Zimmer-Strobl and Strobl, 2001). EBNA2 can also transactivate the viral Cp and up-regulate LMP1 and LMP2 expression (Longnecker *et al.*, 2013; Sjöblom *et al.*, 1995).

Genome-wide association studies in LCLs and BL cell lines have revealed that EBNA2 is preferentially sited distal to promoter regions of cellular genes: within intragenic or intergenic regions that are most commonly transcriptional enhancers enriched with acetylated lysine 27 residue on histone 3 (H3K27Ac) (McClellan *et al.*, 2013; Zhao *et al.*, 2011a; Zhou *et al.*, 2015). EBNA2 frequently co-localises with RBPJ, as well as other B-cell transcription factors such as Early B-cell Factor 1 (EBF1) that are important in the determination of cell fate (Lu *et al.*, 2016; McClellan *et al.*, 2013; Zhao *et al.*, 2011a). Furthermore, chromatin conformation capture experiments revealed that an enhancer site situated 428kb upstream of *MYC* gene is co-occupied by EBNA2 and RBPJ and this enhancer is brought into proximity of the *MYC* promoter through EBNA2-dependent chromatin looping, hence presenting evidence that EBNA2 transactivates cellular genes by mediating long-range interactions between regulatory regions of the genome (Wood *et al.*, 2016; Zhao *et al.*, 2011a). Importantly, EBNA2 also interacts with other EBV latent proteins. Although the exact mechanisms remain unclear, EBNA-LP is known to co-operate with EBNA2 in the activation of viral Cp and LMP promoters and in regulating some host genes (reviewed in Kempkes and Ling, 2015). Notably, EBNA2 also interacts with EBNA3 proteins (see section 1.7) in the regulation of key viral and cellular genes involved in cell proliferation and differentiation.

1.6. Epstein-Barr virus Nuclear Antigen 3s (EBNA3s)

1.6.1. Genome organisation and structure

The EBNA3 family of proteins (EBNA3A, 3B and 3C) have no known human homologues. It is likely that these three genes likely arose from historic gene duplication events of a single progenitor gene during EBV evolution. The EBNA3 genes have similar structures and are tandemly arranged in the EBV genome. Each EBNA3 gene comprises a short (~360 nucleotides) 5'- exon followed by intron that is shorter than 100bp, and then a long (~2.5kb) 3'-exon (Figure 1.4A) (Allday *et al.*, 2015; Joab *et al.*, 1987; Petti and Kieff, 1988; Petti *et al.*, 1988). The individual EBNA3 mRNAs are alternatively spliced from long

pre-mRNA transcripts initiated primarily from the Cp (Figure 1.4B) (Lieberman, 2013; Price and Luftig, 2014). This mRNA transcript contains an internal ribosome entry site (IRES) within the U exon, where ribosomes are recruited directly to initiate the translation of individual EBNA3s and EBNA1 (Isaksson *et al.*, 2003). Complex alternative splicing events involving EBNA2 and EBNA-LP upstream of the U-exon may inadvertently contribute to less frequent transcription of downstream open reading frames (ORFs), resulting in low levels of EBNA3 mRNA and proteins (Allday *et al.*, 2015). Additionally, there appears to be a tight regulation of EBNA3 protein levels since immunoblotting experiments have shown that the levels of EBNA3 proteins remain relatively constant in different LCLs and the proteins are stable with half-lives of more than 24h (Allday *et al.*, 2015). This could be a result of strict transcriptional and translational regulation involving mRNA splicing and IRES function or protein metabolism, however, the specific mechanisms underlying these processes remain unclear (Allday *et al.*, 2015).

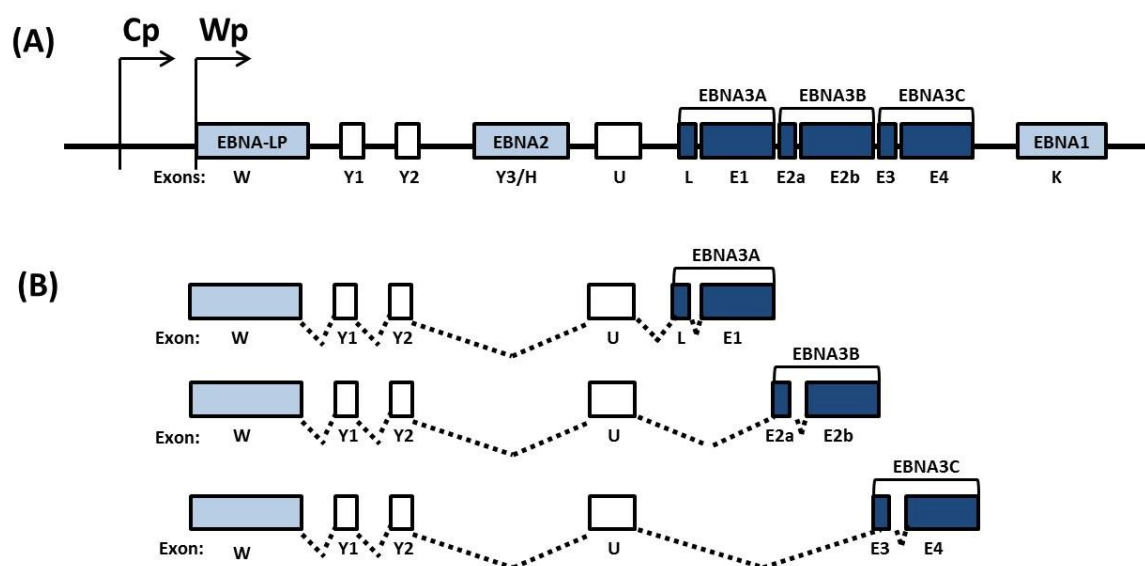


Figure 1.4. Schematic showing the splicing of EBNA3 mRNA.

Exons are depicted using blue boxes with their corresponding exon name below. EBNA-LP is represented as one long exon whereas in reality, it is composed of a variable number of small W exon repeats. **(A)** Immediately after infection with EBV, multiple copies of Wp present within the BamHI W repeats are activated and initiate transcription of long pre-mRNA comprising exons coding for all the EBNA proteins. During latency III initiation of transcription switches to the single upstream Cp. **(B)** Alternative splicing depicted by dotted lines produces the individual EBNA3 mRNA. EBNA3B transcripts additionally contain the open reading frame for EBNA3C but this is omitted for simplicity. An internal ribosome entry site (IRES) is found within the U-exon where ribosomes are recruited for initiation of translation. The IRES allows translation of EBNA3s or EBNA1 depending on which open reading frame is retained, whereas EBNA-LP and EBNA2 are translated in a cap-dependent manner. Figure adapted from Allday *et al.* (2015) with some modifications.

1.6.2. Structural features and homology between the EBNA3 family of proteins

The EBNA3s are large proteins that are each composed of more than 900 amino acids. Immunofluorescence experiments have shown that each of the EBNA3s localise exclusively to the nucleus with a punctate staining pattern and sparing of the nucleolus (Allday *et al.*, 1988; Hennessy *et al.*, 1985; Krauer *et al.*, 2004b; Petti and Kieff, 1988; Petti *et al.*, 1988; Ricksten *et al.*, 1988; Young *et al.*, 2008). Each of these proteins contains more than one nuclear localisation signal (NLS), as shown in Figure 1.5. There are six NLSs within EBNA3A (Buck *et al.*, 2006; Le Roux *et al.*, 1993), two within EBNA3B (Burgess *et al.*, 2006) and three within EBNA3C (Allday *et al.*, 1988; Krauer *et al.*, 2004a). The organisation of each protein is grossly similar, although there is limited amino acid (aa) sequence homology between the EBNA3s (Figure 1.5). Each of the EBNA3s contains a hydrophilic N-terminus (between 90-97 aa), a 220-230 aa homology region, short imperfect repeat regions, and a proline-rich C-terminal region (Le Roux *et al.*, 1994). Within the homology domain, EBNA3A and EBNA3C share 28% aa sequence identity, EBNA3B and EBNA3C share 27% identity and EBNA3A and EBNA3B share 23% identity (Allday *et al.*, 2015; Le Roux *et al.*, 1994). Despite the partial sequence homology between the EBNA3s, their predicted secondary structures are largely similar (Yenamandra *et al.*, 2009). *In silico* comparative analyses performed by Yenamandra and colleagues (2009) predicted the percentage similarity in secondary structure between the EBNA3s as follows: 88.30% between EBNA3A and EBNA3B, 71.81% between EBNA3B and EBNA3C and 71.52% between EBNA3A and EBNA3C. A large majority of the predicted secondary structures are alpha-helices and there appears to be fewer beta-sheets (Figure 1.5; Allday *et al.*, 2015; Yenamandra *et al.*, 2009). EBNA3C appears to have a unique predicted leucine zipper within the homology domain; disruption of the structure at this leucine zipper interferes with EBNA3C's ability to antagonise EBNA2 activation of genes and its ability to interact with RBPJ (West *et al.*, 2004).

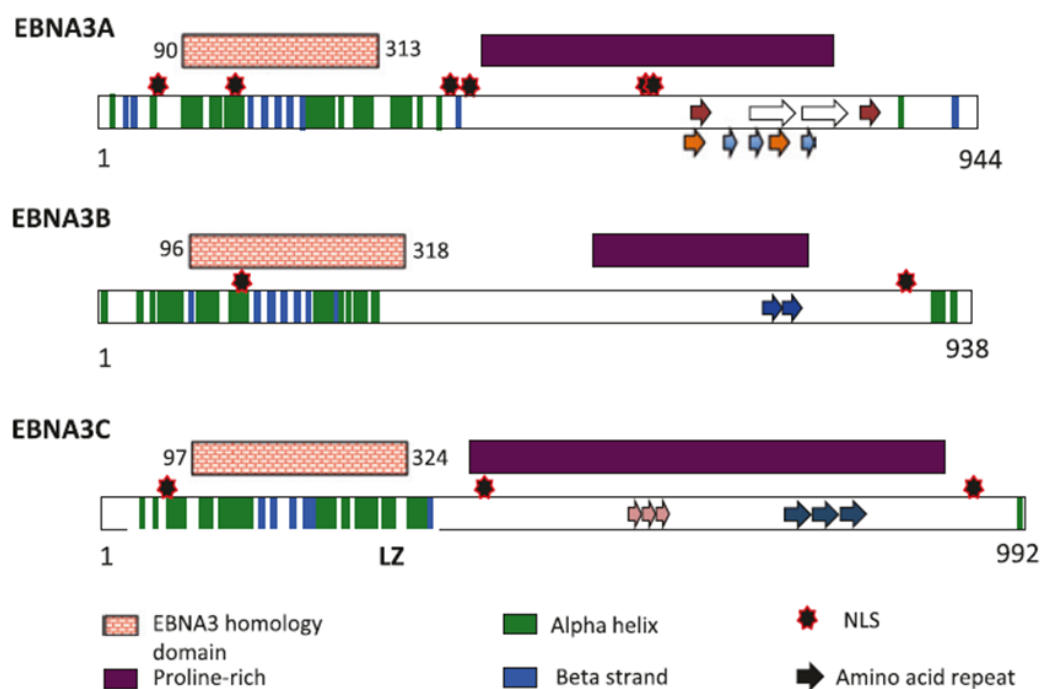


Figure 1.5. Schematic representation of the structural domains of EBNA3s.

Nuclear localisation signals (NLS) located in each protein are shown with black stars. The EBNA3 homology domains are shown as indicated. Structural motifs such as alpha-helices and beta strands are depicted with green and blue rectangles respectively. Repeat regions which are similar are illustrated with arrows of the same colour. Proline-rich regions are indicated using purple rectangles. The location of the unique leucine zipper domain is indicated with "LZ". Figure reproduced from Allday *et al.* (2015). The schematic is not drawn to scale.

1.6.3. The EBNA3s are transcriptional regulators that affect host gene expression

Exon-microarray studies comparing the relative expression of mRNA transcripts in cell lines (either BL or LCLs) established with wild-type or recombinant EBV expressing deletions or truncated versions of EBNA3A, EBNA3B or EBNA3C proteins have been performed by several groups (Chen *et al.*, 2006; Hertle *et al.*, 2009; White *et al.*, 2010; Zhao *et al.*, 2011b) and data from two of these groups are publicly available on <http://www.epstein-barrvirus.org> (Harth-Hertle *et al.*, 2013; Skalska *et al.*, 2010; White *et al.*, 2010). These microarray studies have revealed subsets of genes controlled by each EBNA3 protein, and demonstrated that together the EBNA3s may regulate over a 1000 of the host cell genes (White *et al.*, 2010). A substantial majority of these genes are regulated by cooperating EBNA3s. EBNA3C is involved in the majority of the co-regulation, altering gene expression with either EBNA3A or EBNA3B or both.

The EBNA3 proteins are localised within the nucleus where they can exert their regulatory functions on host cell gene expression (Allday *et al.*, 2015; Burgess *et al.*, 2006; Hennessy *et al.*, 1985; Krauer *et al.*, 2004b). Although – like EBNA2 – they do not bind directly to DNA, the EBNA3s exhibit strong repressor activity in transient transfection assays when they are targeted directly to a reporter gene by fusion with the DNA-binding domain of yeast transactivator Gal4 (Allday *et al.*, 2015; Bain *et al.*, 1996; Bourillot *et al.*, 1998 and our unpublished data). Furthermore, EBNA3A and EBNA3C can interact with host cell transcriptional repressors and factors associated with gene silencing. These include C-terminal binding protein (CtBP), histone deacetylases (HDAC) -1 and -2, SIN3 transcription regulator family member A (Sin3A) and nuclear receptor co-repressor (NCoR) (Hickabottom *et al.*, 2002; Knight *et al.*, 2003; Radkov *et al.*, 1999; Touitou *et al.*, 2001a). Importantly, the EBNA3s have been shown to interact with RBPJ, and in co-transfection assays, each EBNA3 has the ability to antagonise the activation of EBV promoters – such as *LMP2A* and *Cp* – by EBNA2. This suggests that the EBNA3s may be involved in the negative auto-regulation of latency III genes by antagonising EBNA2 transactivation of the *Cp* and *LMP2A*

promoters (Robertson *et al.*, 1996; Le Roux *et al.*, 1994; Waltzer *et al.*, 1994, 1996). However, the EBNA3s do not function solely as transcriptional repressors, because various microarray studies have shown that the EBNA3s play a role in activating cellular genes as well as in their repression (Anderton *et al.*, 2008; Hertle *et al.*, 2009; Skalska *et al.*, 2010, 2013; White *et al.*, 2010). Recently Bazot and colleagues have found that EBNA3A and EBNA3C can directly activate the oncogenic microRNA cluster miR221/222, indirectly repressing the downstream target of miR221/222, the cyclin-dependent kinase inhibitor p57^{KIP2} (Bazot *et al.*, 2015). In addition, EBNA3C recruits DNA-binding protein RBPJ, binds to well-characterised regulatory regions in the *AICDA* locus (coding for the AID protein – see section 1.4), and induces *AICDA* gene transcription by mediating the deposition of activation-associated histone marks such as acetylated lysine residue 27 of histone 3 (H3K27Ac), acetylated lysine residue 9 of histone 3 (H3K9Ac) and tri-methylated lysine residue 4 of histone 3 (H3K4me3) (Kalchschmidt *et al.*, 2016a). The deposition of activating histone marks around the *AICDA* locus is associated with the recruitment of co-activator histone acetyl-transferase p300 by EBNA3C (Kalchschmidt *et al.*, 2016a). EBNA3C can also localise to and directly transactivate the bi-directional *LMP1/2B* promoter (Jiménez-Ramírez *et al.*, 2006; Zhao and Sample, 2000). However, the exact molecular mechanisms of gene repression and activation by the EBNA3s still remain to be elucidated (see section 1.7 for more details).

1.6.4. EBNA3A & EBNA3C are potent oncogenes

Despite their similarities, the EBNA3s each have individual and largely non-redundant roles in regulating cellular gene expression. EBNA3A and EBNA3C are now widely recognised as viral oncogenes which play critical roles in mediating cell-cycle progression and apoptosis. Reverse genetic studies using mutant viruses expressing truncated EBNA3A or EBNA3C have shown that both EBNA3A and EBNA3C are essential for B-cell transformation (Tomkinson *et al.*, 1993). Expression of either EBNA3A or EBNA3C in rat embryonic fibroblasts (REF) can overcome G1 arrest in a state of premature

senescence induced by oncogenic Ha-Ras protein, hence driving the transformation of REFs (Hickabottom *et al.*, 2002; Parker *et al.*, 1996; Touitou *et al.*, 2001b). The ability of EBNA3A and EBNA3C to drive oncogenic transformation in collaboration with Ha-Ras is dependent on their interaction with cellular co-repressor CtBP (Hickabottom *et al.*, 2002; Touitou *et al.*, 2001a). EBNA3C has also been reported to interact with numerous cell cycle regulators such as cyclin A, p53 and its antagonist MDM2 (Knight and Robertson, 2004; Saha *et al.*, 2009; Yi *et al.*, 2009).

In order to study the biological consequences of host gene regulation by EBNA3A and EBNA3C, several groups have chosen to fuse the 4-hydroxytamoxifen (4HT)-conditional, selective oestrogen receptor (ERT2) to the each of EBNA3A and EBNA3C (EBNA3A-HT and EBNA3C-HT respectively). The 4HT-conditional system comprises fusion of a modified oestrogen receptor that is selectively responsive to 4HT, and non-responsive to endogenous oestrogen to the C-terminus of EBNA3A or EBNA3C. By cloning EBNA3s that are individually 4HT-responsive, the tagged EBNA3 proteins will be sequestered to the cytoplasm for degradation in the absence of 4HT, and, in the presence of 4HT, will remain in the nucleus where they can employ their gene regulatory functions. Using these EBNA3A-HT and EBNA3C-HT systems, EBNA3A and EBNA3C are shown to repress the expression of senescence-inducing p16^{INK4A} and p14^{ARF} proteins and the pro-apoptotic BIM (Anderton *et al.*, 2008; Maruo *et al.*, 2003, 2006, 2011; Paschos *et al.*, 2012; Skalska *et al.*, 2010, 2013). Notably, CtBP plays a role in mediating the EBNA3A/3C-driven epigenetic silencing at the p16^{INK4A} locus (Skalska *et al.*, 2010). Inactivation of EBNA3C in EBNA3C-HT LCLs leads to a subsequent increase of p16^{INK4A} and reduced hyper-phosphorylated (activated) retinoblastoma (Rb) protein – a negative regulator of cell cycle – in the cells, which is accompanied by reduced cell cycle progression (Maruo *et al.*, 2006). Likewise, inactivation of EBNA3A in LCLs expressing 4HT-conditional EBNA3A resulted in growth arrest and a decline in proliferation, which can only be salvaged by the re-activation of EBNA3A (Maruo *et al.*, 2003). In addition, EBNA3A can also interact with MYC interacting finger protein 1

(MIZ-1) in the repression of cyclin-dependent kinase inhibitor p15^{INK4b}, although the impact of this on cell cycle progression of B-cells is unknown (Bazot *et al.*, 2014). Hence, together, EBNA3A and EBNA3C exert their oncogenic properties through the transcriptional regulation of and interaction with components of cell cycle checkpoint regulators.

1.6.5. EBNA3B as a viral tumour suppressor

To date, interest in the EBNA3s has mainly focussed on the gene regulatory functions of EBNA3A and EBNA3C, and the mechanisms of actions of these two oncogenic proteins have been rigorously examined. In contrast to the other two EBNA3 proteins, EBNA3B may paradoxically function as a viral-encoded tumour suppressor (White *et al.*, 2012). EBNA3B is not essential for *in vitro* transformation of B-cells (Chen *et al.*, 2005; Tomkinson and Kieff, 1992) and on the contrary, B-cells infected with EBNA3B-null recombinant EBV (3BKO) appear to proliferate earlier and faster in comparison with those infected with wildtype (WT) B95.8 and EBNA3B-revertant (3BRev) recombinant EBV (White *et al.*, 2012). Furthermore, established 3BKO LCLs grow robustly *in vitro* and spontaneous deletions in EBNA3B have been observed both in adult B-cell- (Allday *et al.*, 2015; unpublished observations) and umbilical cord blood-derived LCLs (Allday and Farrell, 1994; Chen *et al.*, 2005) that were originally established using WT B95.8 EBV. Put together, these observations indicate that at least *in vitro*, EBNA3B plays a role in impeding proliferation and growth. Hence, the loss of EBNA3B can confer growth advantages to EBNA3B-null cells.

It is likely that EBNA3B serves an important function *in vivo* since EBNA3B has not been counter-selected despite millennia of virus-host co-evolution. Exon-microarray analyses done by our group revealed that 3BKO LCLs/BLs express a transcriptome that is distinctly different from that of WT cell lines (White *et al.*, 2010): about 200 cellular genes demonstrated altered expression in EBNA3B-null LCLs (more than 2-fold change, $p < 0.001$), and are therefore considered to be EBNA3B-regulated. In addition, of these EBNA3B-regulated genes, numerous encode cell surface proteins or chemokines which play important roles in regulating cell-cell interaction (including CD28, LAIR1, CXCL10, IL10, IL19

and ITGAL) and may have important and significant functions *in vivo* (White *et al.*, 2010). Hence, in order to uncover the *in vivo* function of EBNA3B, White *et al.* (2012) utilised immuno-deficient mice (NOD/LtSz-scid/IL2R γ null) which were engrafted with CD34+ human foetal haematopoietic progenitor cells, therefore generating mice with reconstituted components of the human immune system (hu-NSG mice). These hu-NSG mice were infected using equal titres of WT, 3BKO and 3BRev recombinant EBV, and the effects on lymphomagenesis *in vivo* were observed.

At least half of the 3BKO-infected hu-NSG mice had enlarged spleens with large tumours (White *et al.*, 2012). Morphologically, these tumours are monomorphic and phenotypically similar to non-germinal centre, activated B-cell DLBCL. More than 90% of the cells are Ki67 positive, an indication of rapid cellular proliferation. Hence infection of hu-NSG mice with 3BKO EBV promotes a highly proliferative and aggressive lymphoma. Conversely, hu-NSG mice infected with WT and 3BRev viruses did not exhibit splenomegaly and developed a lymphoma consisting of infiltrating B-cells at varying stages of differentiation that is more consistent with the phenotype of polymorphic human PTLD. In addition, markedly lower percentages of the WT/3BRev-associated B cells are Ki-67 positive, indicating that the WT/3BRev EBV-associated lymphoproliferations are less mitotic and far less aggressive than the tumours associated with EBNA3B-null EBV.

Perhaps the most interesting finding of White *et al.* (2012) was that splenic lesions found in mice infected with 3BKO EBV were significantly deficient in T-cell infiltration as compared to those found within WT and 3BRev-infected hu-NSG mice, despite there being greater T-cell expansion in 3BKO EBV-infected mice. This is partly due to 3BKO EBV-induced tumour cells secreting reduced levels of CXCL10, a T-cell chemo-attractant. The expression of CXCL10 is consistently reduced in the 3BKO hu-NSG mice tumour-derived LCLs, cell lines derived from two patients known to have truncations in the EBNA3B gene (see below) and in 3BKO LCLs, as compared to their WT/3BRev counterparts (White *et al.*, 2010, 2012). Furthermore, absence of EBNA3B in LCLs not only reduces the basal CXCL10

expression, but also diminishes the responsiveness of CXCL10 induction by interferon-gamma (IFN γ) (White *et al.*, 2012). Importantly, impaired T-cell migration is observed in cell supernatants prepared from 3BKO tumour-derived LCLs which is moderately improved by restoring the levels of CXCL10 in the media (White *et al.*, 2012). Finally, White and colleagues (2012) demonstrated using *in vivo* killing assays that the restoration of CXCL10 expression in 3BKO LCLs via lentiviral transduction renders these cells more susceptible to T-cell mediated cytotoxicity. Hence, the evidence strongly suggest that *in vivo*, EBNA3B plays a role in enhancing T-cell recruitment and immune surveillance (White *et al.*, 2012).

EBNA3B is likely to assist the transition and differentiation of infected B-cells from activated B-lymphoblasts to the GC where interactions of activated B-lymphoblasts with helper T-cells facilitate class-switch recombination and the subsequent B-cell selection. This is consistent with the observation from exon-microarray studies that EBNA3B down-regulates some genes that are repressed in germinal centre B-cells (White *et al.*, 2010). Conceivably, defective B-cell differentiation arising from inadequate interactions with T-helper cells in the GC is a plausible explanation for the monomorphic DLBCL phenotype found in 3BKO hu-NSG mice (Allday *et al.*, 2015; White *et al.*, 2012).

Naturally-occurring EBNA3B mutations and truncations can be found in patients with EBV-positive B-cell lymphoma, which may manifest as aggressive disease. In 2001, Gottschalk and colleagues reported a case of aggressive and fatal PTLD which was associated with the selection of EBNA3B-deleted EBV during treatment with donor-derived CTLs that predominantly targeted 2 epitopes on EBNA3B. A second case of PTLD with an EBNA3B truncation from insertion mutation was described by White *et al.* (2012), although this case was detected early and removal of affected lymphoid tissue coupled with adoptive T-cell immunotherapy sufficiently cured the patient. Cell lines established from these two PTLD patients demonstrated similar phenotype and gene expression profiles as the 3BKO LCLs derived from hu-NSG mice and primary B-cell infections (White *et al.*, 2012). Further screening of 11 EBV-positive non-germinal centre DLBCLs revealed one other tumour

associated with EBV expressing truncated EBNA3B. In addition, deletions and potential EBNA3B mutations have been detected in several HL and BL cell lines (White *et al.*, 2012). However, it is not possible to determine the prevalence of genuine EBNA3B mutations associated with lymphomagenesis without comprehensive knowledge of the diversity of EBNA3B sequence and any naturally-occurring, harmless polymorphisms in the general population (Allday *et al.*, 2015; White *et al.*, 2012).

In summary, EBNA3B displays features of a viral-encoded tumour suppressor that has an essential function in enhancing immune surveillance. The aggressive case of PTLD reported by Gottschalk *et al.* (2001), along with our observations from the hu-NSG mouse experiments demonstrate that *in vivo*, EBNA3B exerts an important tumour suppressive function that counteracts the potent oncogenic properties of other EBV latency proteins such as EBNA3A, EBNA3C and LMP1 during the initial and transient proliferative phase of infected B-lymphoblasts. Thus, EBNA3B can prevent severe mortality and morbidity resulting from uncontrolled lymphoproliferation, ultimately ensuring the persistence of the virus and its host.

1.7. Mechanisms of gene regulation by the EBNA3s

1.7.1. EBNA3s can interact with EBNA2 and RBPJ

It has been speculated that the EBNA3s antagonise EBNA2 transactivation of both cellular and viral genes (see also section 1.6.3) by competing for binding with DNA-bound RBPJ (Johannsen *et al.*, 1996; Robertson *et al.*, 1996). The EBNA3s are able to disrupt EBNA2-mediated transactivation of genes in transient reporter assays (Robertson *et al.*, 1995; Le Roux *et al.*, 1994; Waltzer *et al.*, 1996; Zhao *et al.*, 1996). In addition, electrophoretic mobility shift assays using *in vitro* translated EBNA3s showed that the EBNA3s are able to disrupt RBPJ binding to DNA (Robertson *et al.*, 1995, 1996; Waltzer *et al.*, 1996). Although RBPJ is one of the most extensively studied EBNA3-associated transcription factor, the molecular details of how RBPJ interacts with the EBNA3s are only slowly becoming uncovered. Each EBNA3 protein can directly interact with RBPJ through

their homology domains which are located between amino acid 170-221 for EBNA3A, 176-227 for EBNA3B and 180-231 for EBNA3C (see section 1.6.2; reviewed in Allday *et al.*, 2015). Deletions of these domains in EBNA3A and EBNA3C ablate their ability to antagonise EBNA2-mediated transactivation of reporter genes in transient transfection assays and their capacity to transform B-cells and maintain LCL proliferation (Lee *et al.*, 2009; Maruo *et al.*, 2005, 2009).

The EBNA3s and EBNA2 bind to the same site on RBPJ (Robertson *et al.*, 1995, 1996), however, even though EBNA2 and EBNA3 proteins can be co-immunoprecipitated readily with RBPJ, each of these proteins form distinct and mutually exclusive complexes with RBPJ (Johannsen *et al.*, 1996; Ohashi *et al.*, 2015).

Put together, all of the early evidence have contributed to the development of two models for the antagonism of EBNA2 activity. The first model proposes that the EBNA3s prevent the EBNA2-mediated transactivation of genes by binding to RBPJ and disrupting the binding of EBNA2/RBPJ complex to DNA (Allday *et al.*, 2015; Hayward, 2004; Young and Rickinson, 2004). In another model, EBNA3s directly displace EBNA2 from DNA-bound RBPJ, subsequently recruiting transcriptional co-repressors at the EBNA3/RBPJ-bound sites (Allday *et al.*, 2015; Harth-Hertle *et al.*, 2013). It is possible that different mechanisms of action may exist at different target gene loci.

Chromatin Immunoprecipitation (ChIP) experiments are used extensively to detect the localisation of DNA-associated proteins on any given genome. Briefly, DNA-associated protein complexes are cross-linked to fragmented DNA using paraformaldehyde, and then selectively immunoprecipitated using specific antibodies. Associated DNA is eluted and the sequence determined typically by quantitative PCR (ChIP-qPCR) (Figure 1.6A, more details in section 2.19 and section 4.1). Examination of genome-wide association of each of these proteins can also be done using ChIP coupled with high throughput sequencing (ChIP-seq)

which scrutinises the entire human genome for relative enrichment of the DNA-associated protein found at specific gene loci.

Consistent with current understanding that there is an extensive interplay between the EBNA3 proteins and EBNA2 in the regulation of gene expression (Harth-Hertle *et al.*, 2013; McClellan *et al.*, 2013; Le Roux *et al.*, 1994; Waltzer *et al.*, 1996), ChIP-seq studies have shown that a significant proportion of EBNA3 binding regions identified so far co-localise with EBNA2 (McClellan *et al.*, 2013; Wang *et al.*, 2015). Although only a small proportion (30-40%) of EBNA3-bound regions co-localise with RBPJ; those which overlap with RBPJ are also frequently enriched with EBNA2 (Wang *et al.*, 2015). McClellan and colleagues (2013) revealed that EBNA3s and EBNA2 compete for binding at shared binding sites near their regulated genes, and proposed that EBNA2 and EBNA3s exert opposing effects on chromatin remodelling, and consequently gene regulation. Furthermore, inactivation of EBNA3A or EBNA3C in 4HT-permissive LCLs results in no change of RBPJ occupancy, but increased enrichment of EBNA2 at the EBNA3A/RBPJ or EBNA3C/RBPJ-bound regions at the loci of some EBNA3A- (HDAC7 and CDH1) or EBNA3C- (BACH2, JAK1 and CXCR5) repressed genes (Wang *et al.*, 2015). Put together, these findings support the model that EBNA3A and EBNA3C can compete with EBNA2 for binding to DNA-bound RBPJ at specific genomic locations, hence preventing the transactivation of these genes.

Recently, Kalchschmidt *et al.* (2016a, 2016b) made the observation that, in contrast to the classic picture of RBPJ being statically bound to DNA, EBNA3C is able to recruit RBPJ to the loci of three EBNA3C-repressed genes *COBLL1* and *ADAM28/ADAMDEC1* and one EBNA3C-induced gene *AICDA* – all of which are not co-bound by EBNA2. This observation has created a paradigm shift in our understanding of the interaction between EBNA3 and RBPJ, introducing a new model that EBNA3/RBPJ complexes may be dynamically recruited and stabilised at various gene loci, hence facilitating gene regulation. Similarly, Lu and colleagues (2016) revealed that EBNA2 increases the stability/recruitment of RBPJ and EBF1 to numerous genomic loci. Put together, the observations of Lu *et al.* (2016) and

Kalchschmidt *et al.* (2016a, 2016b), indicate that EBNA2 and EBNA3s are able to recruit transcription factors (such as RBPJ and EBF1) to a subset of their regulated genes. Hence it is clear that the interaction of RBPJ with EBNA3s and/or EBNA2 is important in the regulation of a distinct subset of EBNA3-regulated genes, although the mechanisms through which this is done may be varied and would require more investigation at specific gene targets.

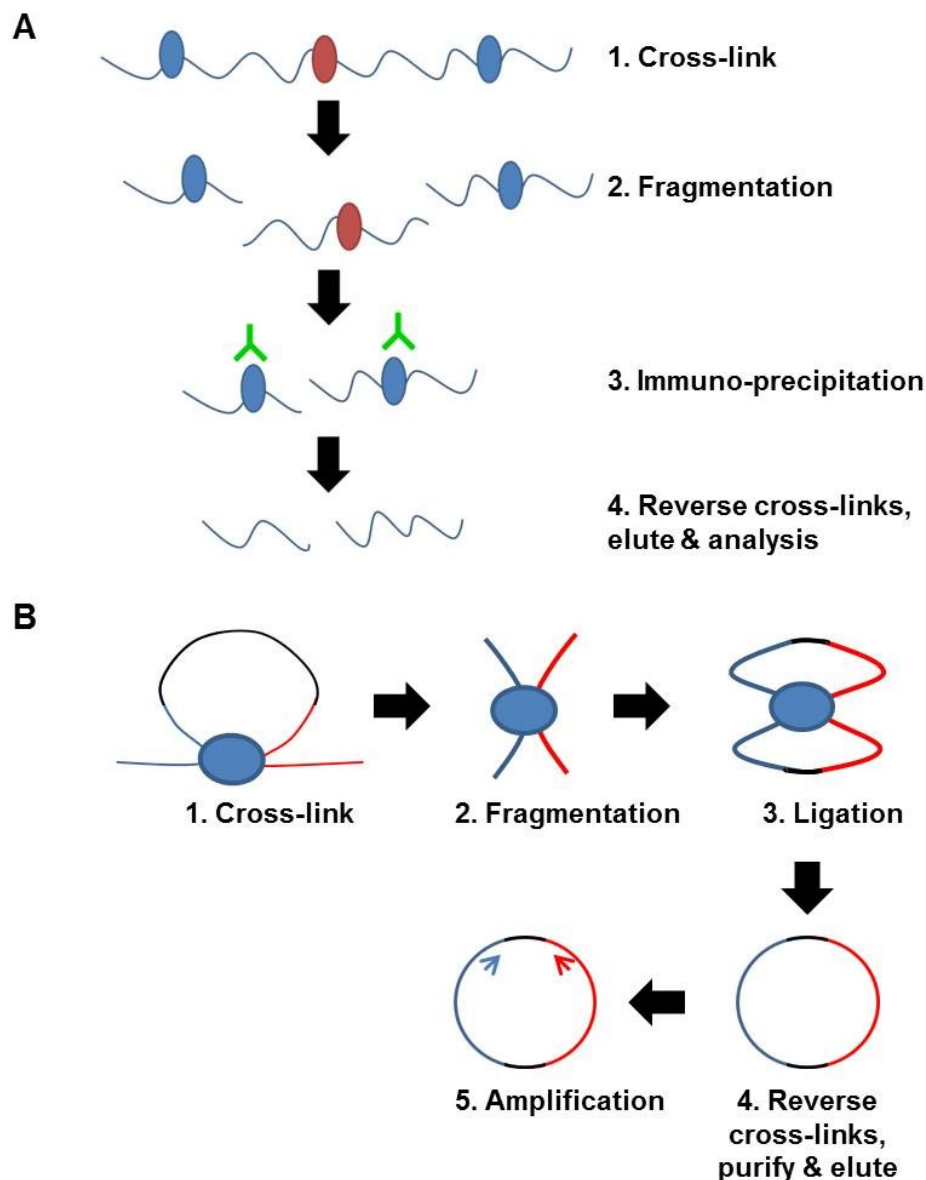


Figure 1.6. Schematic showing (A) Chromatin-Immunoprecipitation & (B) Chromatin Conformation Capture (3C) techniques.

(A) Proteins are first cross-linked to DNA using formaldehyde, after which the cross-linked DNA are fragmented using restriction digestion or sonication. Antibodies specific to the protein of interest are used to immunoprecipitate the DNA-bound proteins, after which the protein-DNA crosslinks are reversed and the DNA eluted for further quantification with PCR. **(B)** The architecture of human chromatin is highly complex and distal regions are often brought together via various looping mechanisms that are orchestrated and maintained by proteins and/or protein complexes. Briefly explained, in 3C, the DNA and proteins are cross-linked using formaldehyde in order to maintain the spatial 3D conformation. Fragmentation by restriction digestion or sonication results in fragments composed of DNA that are brought into close proximity as a result of the 3D structure. This is followed by ligation of the fragmented DNA at dilute concentrations in order to encourage intra-fragmental ligation. Next, the cross-links are reversed, and the DNA purified and eluted. PCR amplification of the DNA using primer pairs specific to the two otherwise distal regions is done to verify that these distal regions are indeed brought into close proximity via chromatin loops.

1.7.2. EBNA3s can bind proximally to genes and direct gene repression by interacting with PcG proteins

EBNA3s are able to mediate epigenetic silencing of their target genes via interactions with polycomb group proteins (PcG) within Polycomb Repressive Complexes (PRC). The PRC2 complex is a multi-protein complex that catalyses and maintains tri-methylation of lysine 27 residue of histone 3 (H3K27me3) and is composed of three core proteins: SUZ12, EED, and EZH2, where EZH2 is responsible for the deposition of H3K27me3 (Di Croce and Helin, 2013; Simon and Kingston, 2009). The presence of H3K27me3 can in turn recruit PRC1 complex which comprises components such as BMI1, RING1, MEL18 and CBX proteins. This PRC1 complex catalyses the mono-ubiquitylation of lysine 119 in histone H2A (H2AK119Ub), consequently resulting in local chromatin compaction and hence gene silencing (Cheutin and Cavalli, 2014; Di Croce and Helin, 2013; Simon and Kingston, 2009). Paschos *et al.* (2012) showed that EBNA3C binds to the *BIM* promoter and represses the expression of *BIM* in part by recruiting PRC2 and directing deposition of repressive H3K27me3 at the *BIM* TSS. In particular, it has been shown that PcG components (SUZ12 and EZH2) can only be recruited to the transcriptional start site of *BIM* in the presence of both EBNA3A and EBNA3C, and the use of conditional EBNA3C-HT cell lines demonstrated that this epigenetic repression was reversible by switching between permissive and non-permissive conditions (Paschos *et al.*, 2012). Since the *BIM* promoter is enriched with trimethylation at the lysine 4 residue of histone 3 (H3K4me3) – a marker of activated chromatin – it is postulated that EBNA3A and EBNA3C are able to maintain a “poised” and bivalent chromatin state at the *BIM* promoter by facilitating the deposition of repressive H3K27me3. The end result should be a promoter that can be readily activated or repressed, poised at the initiation stage of mRNA transcription with the presence RNA polymerase II (RNA PolII) that is phosphorylated at the serine 5 residue (PolII S5) (Bernstein *et al.*, 2006; Brookes and Pombo, 2009; Levine, 2011). However, in the presence of EBNA3A and EBNA3C, PolII S5 phosphorylation is inhibited at the *BIM* promoter, which prevents the initiation of mRNA transcription. Hence, at the *BIM* locus, EBNA3C and EBNA3A keep the

promoter in a non-classical “poised” state by facilitating the deposition and maintenance of repressive H3K27me3 and also preventing the Ser5 phosphorylation of RNA PolII.

Similar to the situation at the BIM locus, EBNA3C can also localise directly to the TSS of *CDKN2A* (p16^{INK4A} & p14^{ARF}) and *CDKN2B* (p15^{INK4B}) loci, and EBNA3A and EBNA3C are both associated with increased H3K27me3 deposition at the TSS (Maruo *et al.*, 2005, 2009; Skalska *et al.*, 2010). This whole *CDKN2A/2B* locus is thought to be co-ordinately regulated by EBNA3A and EBNA3C, and it is possible that the interaction between EBNA3C and IRF4/BATF-containing complexes at these loci contributes to the EBNA3-mediated regulation (Allday *et al.*, 2015; Jiang *et al.*, 2014; Skalska *et al.*, 2010). In addition, EBNA3s have also been found to localise proximal to the TSS at other gene loci – including *ITGB1* and *ITGB4* – and direct H3K27me3-mediated gene repression (McClellan *et al.*, 2012). Put together, the evidence obtained from studying specific gene loci (as described above) support a model whereby the EBNA3s are able to mediate cellular gene repression by recruitment of PcG proteins and, as a consequence, facilitate H3K27me3 deposition and the associated chromatin silencing of cellular genes.

1.7.3. The EBNA3s are able to bind and co-ordinate gene regulation distally from gene TSS

Although the EBNA3s have been shown to bind locally to regulate cellular genes, there is now a common consensus that the EBNA3s are mostly localised in regions that are distal to the promoters or transcriptional start sites of their target genes and they can influence conformational changes of the chromatin to direct gene regulation. Advancements in sequencing and ChIP technologies have spawned numerous publications detailing the genome-wide localisation of each EBNA3 protein in LCLs and BL cell lines (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015; Wang *et al.*, 2015). All groups concur that the EBNA3s predominantly bind distally to TSS of genes, and in the case of Mutu III cells (EBV-positive latency III BL cell line), 84% of the EBNA3-bound sites are localised more

than 4kb away from any gene TSS (McClellan *et al.*, 2013). Each EBNA3 has a distinct set of binding sites, but can also co-localise with one or both of the other EBNA3 proteins, which is consistent with results from our exon-microarray studies (White *et al.*, 2010), and indicates that the EBNA3s co-operate to regulate host genes. Comparisons of EBNA3 ChIP-seq with transcription factor binding data for the LCL GM12878 cell line (publicly available on UCSC Genome browser: <https://genome.ucsc.edu>) revealed that the EBNA3s co-localise with a plethora of B-cell transcription factors, including IRF4, BATF, SP1, PAX5, EBF1, RUNX3 and p300 (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015; Wang *et al.*, 2015). However, how these transcription factors may individually interact with each of the EBNA3s in the regulation of gene expression remains to be determined, although effort has been made to interrogate the role of specific transcription factors. In particular, the transcription factor IRF4 plays a crucial role in recruiting EBNA3C to the chromatin at specific sites, but IRF4 does not seem to significantly affect the binding of either EBNA3A or EBNA3B (Jiang *et al.*, 2014; Wang *et al.*, 2015). Hence, it is likely that many transcription factors are important in tethering the EBNA3s to the genome, and that the binding specificity of each EBNA3 is determined by the distinct combination of transcription factors and complexes that it associates with (Wang *et al.*, 2015).

Even though there are strong indications that the EBNA3s are primarily repressors of gene expression (see section 1.6.1), the EBNA3s localise mainly to regions that are active (H3K4me1+ and H3K27Ac+) or “poised” (H3K4me1+, H3K27Ac-) enhancers and relatively fewer of the EBNA3-bound sites co-localised with repressive marks such as H3K27me3 or heterochromatin-associated H3K9me3 (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015; Wang *et al.*, 2015). Recently, using conditional EBNA3A-HT and EBNA3C-HT LCLs, Bazot *et al.* (2015) reported for the first time that both EBNA3A and EBNA3C bind to distal intragenic enhancer regions and transactivate the oncogenic microRNA miR221/miR222 cluster by promoting enhancer-promoter looping. Importantly, in the presence of EBNA3A and EBNA3C, the levels of histone marks indicative of active

chromatin (H3K4me3, H3K27Ac, H3K9Ac) are raised at the miR221/222 TSS and at the EBNA3-bound enhancer regions, indicating that the EBNA3s additionally mediate the deposition of such modifications. This is not surprising given that EBNA3C has already been shown to interact with histone acetyl transferase p300 (Cotter and Robertson, 2000).

In a landmark study which examined the EBNA3-mediated repression of *CTBP2*, *WEE1*, *ADAM28* and *ADAMDEC1*, McClellan and colleagues (2013) utilised Chromatin conformation capture (3C) technology to uncover different mechanisms by which the EBNA3s can repress their target genes. Similar to the ChIP techniques, 3C comprises covalently cross-linking DNA that are in close spatial proximity, fragmentation of the cross-linked chromatin via restriction digestion or sonication, and ligation of the fragmented and cross-linked chromatin to form hybrid DNA molecules that are each representative of an interaction between two distal loci (Figure 1.6B) (Dekker *et al.*, 2002). The DNA is then purified; further detection and quantification via PCR is done using primer pairs that amplify across two known, normally distal loci. Hence 3C is classically used to interrogate two known non-proximal regions that are suspected to physically interact through chromatin looping mechanisms. Two different EBNA3 gene regulatory mechanisms were proposed by McClellan *et al.* (2013). For the first model described, EBNA3A binds to an intragenic enhancer within the *CTBP2* locus and mediates the repression of *CTBP2* by preventing the formation of enhancer-promoter loops (facilitated by EBNA2) that would otherwise activate the promoter. In contrast, in the case of *WEE1* and the *ADAM28/ADAMDEC1* loci, EBNA3C binds to intergenic enhancers and facilitates and/or stabilises enhancer-promoter looping, which is associated with transcriptional repression of these genes (McClellan *et al.*, 2013). In the case of *WEE1*, EBNA2 competes with EBNA3s and it is proposed that in the absence of EBNA3s, EBNA2 can prevent chromatin looping between distal enhancer and *WEE1* promoter. The results presented by McClellan *et al.* (2013) have shown that the EBNA3s act distally on their regulated genes and can have important roles in either initiating or stabilising

3D chromatin structures – therefore either preventing or encouraging the looping of distal regulatory regions into close proximity with the promoters of host cell genes.

1.7.4. EBNA3-mediated gene repression follows a complex temporal sequence

The deposition of H3K27me3 at TSS of EBNA3-regulated genes such as *CXCL9/10*, *ADAM28*, *ADAMDEC1* and *COBLL1* is an event that happens late after the initial de-activation of these genes. In order to study the kinetics of EBNA3A-mediated gene regulation, Harth-Hertle and colleagues (2013) stably transfected EBNA3A-null LCLs with an episomal vector that expresses EBNA3A with Doxycycline (Dox) treatment (description of this system can be found in section 3.6). Although there appears to be rapid reduction in *CXCL9* and *CXCL10* expression on induction of Dox-inducible EBNA3A, the establishment and maintenance of H3K27me3 marks across the *CXCL9/10* gene loci is relatively delayed in comparison (Harth-Hertle *et al.*, 2013). Here, Harth-Hertle and colleagues (2013) speculated that EBNA3A binds to and inactivates distal intergenic enhancers between *CXCL9* and *CXCL10* genes. It was proposed that the inactivation of enhancer activity is secondary to the displacement of RBPJ-bound EBNA2 by EBNA3A, which interferes with H3K27 acetylation and prevents the recruitment of RNA PolII to the enhancer, thus disrupting any enhancer-promoter looping mechanisms and resulting in rapid reduction in the transcription of *CXCL9* and *CXCL10* genes. Similarly, on re-introducing 4HT to the culture medium of EBNA3C-HT LCLs, Kalchschmidt and colleagues (2016) observed a delay between the initial reduction of gene expression and the deposition of H3K27me3 marks at the TSS of *ADAM28*, *ADAMDEC1* and *COBLL1*. In this case, it was proposed that the de-activation of gene expression by EBNA3C was initiated by the loss of histone markers of activation (such H3K9Ac, H3K27Ac, and H3K4me3) at the TSS of these three genes, and that PcG-mediated H3K27me3 only appears late after the mRNA levels have been substantially reduced (Kalchschmidt *et al.*, 2016b). Similarly, Paschos *et al.* (2012) determined that de-acetylation of histone H3 is concurrently associated with deposition of H3K27me3 at the BIM locus, although the relative kinetics of the deposition of these histone

marks were not examined. These observations are not surprising, given that the EBNA3s have been shown to interact and bind with HDAC1 and HDAC2 (Knight *et al.*, 2003; Radkov *et al.*, 1999), and might facilitate histone deacetylations by recruiting these HDACs to their target gene loci. Hence, the EBNA3s can repress genes in a temporal fashion: first by facilitating the removal of histone marks of activation, and subsequently by recruiting PcG proteins and the consequent deposition and maintenance of the repressive H3K27me3.

1.8. Telomerase, TERT and EBV

1.8.1. Structure and function of telomeres

Telomeres are tandem repeats of TTAGGG found at the ends of linear cellular DNA. Due to the inability of DNA polymerase to completely replicate the 3' termini of each DNA strand, telomeres in somatic cells are progressively shortened with each cycle of DNA replication until they approach a critical length (Hayflick limit), where the cells would undergo replicative senescence (Counter, 1996; Shay *et al.*, 2001). A complex of proteins composed of TRF1, TRF2, RAP1, TIN2, TPP1 and POT1 – collectively named the shelterin complex – protect the exposed telomeres from recognition by DNA damage repair mechanisms which can trigger chromosomal fusions and inadvertently result in cellular apoptosis (reviewed by De Lange, 2005). Stabilisation of telomeres is a feature present in many cancers (Kim *et al.*, 1994), and is associated with elevated expression of TERT and/or increased activity of the telomerase complex (Counter *et al.*, 1992; Kataoka *et al.*, 1998; Kim *et al.*, 1994). *De novo* synthesis of telomeres is catalysed by the telomerase complex, which is a ribonucleoprotein complex composed of an RNA template (TERC) for the elongation of telomeres and telomerase reverse transcriptase (TERT), the rate limiting enzyme subunit (Autexier and Lue, 2006; Hu and Insel, 1999; Nakayama *et al.*, 1998; Weinrich *et al.*, 1997). In addition, telomeres can concomitantly be elongated by an alternate lengthening of telomeres (ALT) pathway involving the homologous recombination of telomeres (Dunham *et al.*, 2000). The mechanisms of action of this pathway remain unclear, but cells undergoing the ALT pathway would exhibit distinct characteristics including an abundance of extra-chromosomal telomeric

DNA, promyelocytic leukaemia nuclear bodies (PBL) containing telomeric chromatin and highly varied chromosomal telomere lengths (reviewed by Cesare and Reddel, 2010).

1.8.2. Effects of telomere length and TERT on EBV-driven B-cell immortalisation

Although it was initially believed that EBV can immortalise infected B-cells *in vitro*, most established cell lines eventually reach a proliferative crisis between 50-160 population doubling levels (PDLs) (Counter *et al.*, 1994; Kataoka *et al.*, 1998; Sugimoto *et al.*, 1999). Only EBV-infected transformed B-cells which consistently express high levels of TERT and maintain elevated telomerase activity are able to overcome cellular senescence and become truly immortalised (Counter *et al.*, 1994; Kataoka *et al.*, 1998; Sugimoto *et al.*, 1999). Even though MYC is a known regulator of TERT expression and telomerase activity (Wang *et al.*, 1998a; Wu *et al.*, 1999), LMP1 has been proposed as the main inducer of TERT activation in EBV positive B-cells (Terrin *et al.*, 2008). LMP1 induces TERT expression by activating the NF- κ B and MAPK/ERK1/2 pathways (Terrin *et al.*, 2008). In contrast, we have found in our microarray assays that EBNA3B is a strong candidate for the repression of TERT in LCLs and EBV-infected BL cell lines (White *et al.*, 2010), although the mechanisms behind this repression are unknown. Hence it appears that LMP1/MYC and EBNA3B exert opposing effects on the activation of *TERT* gene, and might be an interesting area of study. Intriguingly, it was recently reported that the ALT pathway is activated early in EBV-driven immortalisation and it is the main mechanism maintaining telomere length in telomerase-negative LCLs prior to the activation of telomerase, which may only occur after 100 PDLs in some cell lines (Kamranvar *et al.*, 2013; Sugimoto *et al.*, 2004). Here, EBNA1 may play a role in activating the ALT pathway since EBNA1 induces features characteristic of ALT-mediated telomere elongation in B-cells, such as increased heterogeneity in telomere length, accumulation of extrachromosomal telomeres, and assembly of DNA repair proteins at telomeres (Kamranvar and Masucci, 2011). Regardless of the mechanism, it is clear that maintaining telomere length is a crucial factor that is important in the immortalisation of EBV-infected B-cells (Counter *et al.*, 1994; Kamranvar *et al.*, 2013; Kataoka *et al.*, 1998;

Sugimoto *et al.*, 1999, 2004). Hence, the relationship between TERT and EBV proteins is definitely one worth exploring.

1.9. TRAIL and its receptors

Evading apoptosis is a hallmark of oncogenesis that contributes to tumour formation and growth. Tumour necrosis factor-related apoptosis-inducing ligand (TRAIL or TNFSF10) belongs to the tumour necrosis factor (TNF) family of cytokines and is a transmembrane protein with an extracellular domain that can be proteolytically cleaved. TRAIL induces apoptosis in cells upon binding to its transmembrane protein receptors TRAILR1 and TRAILR2, also known as DR4 and DR5 respectively (reviewed by LeBlanc and Ashkenazi, 2003). Activation of the death receptors recruits Fas-associated protein with Death Domain (FADD) and caspases 8 or 10, forming the Death Inducing Stimulating Complex (DISC) which initiates a cascade of intracellular signalling events that ultimately lead to cellular apoptosis (Abdulghani and El-Deiry, 2010; LeBlanc and Ashkenazi, 2003). Inhibition of this pathway is mediated by cellular FLICE-inhibitory protein (c-FLIP) which antagonises the apoptotic cascade by inhibiting the activation of caspase 8 (reviewed in Shirley and Micheau, 2013). TRAIL can also bind to its decoy receptors: TRAILR3 (DcR1) and TRAILR4 (DcR2). In contrast to TRAILR1 and TRAILR2, binding of TRAIL to its decoy receptors does not trigger apoptosis, and overexpression of these receptors can confer a protective effect to TRAIL-induced programmed cell death (Almasan and Ashkenazi, 2003; Degli-Esposti *et al.*, 1997; Sheridan *et al.*, 1997).

Currently, there is limited literature on the relationship between EBV and TRAIL. Although TRAIL has been shown to effectively induce apoptosis in cancer cells, sporadic reports in the literature have revealed that EBV can contribute towards resistance to TRAIL-induced apoptosis in EBV-positive BLs and LCLs derived from PTLN patients (Mouzakiti and Packham, 2003; Snow *et al.*, 2006; Tafuku *et al.*, 2006). This anti-apoptotic effect is in part mediated by LMP1 via the action of NF- κ B, which upregulates c-FLIP expression and inhibits caspase 8 activation (Mouzakiti and Packham, 2003; Snow *et al.*, 2006; Tafuku *et al.*, 2006).

We have additionally found using our microarray studies that EBNA3B induces the expression of TRAIL, but represses the expression of TRAILR1 and TRAILR4 (White *et al.*, 2010), although the biological significance of this is yet to be identified.

1.10. Background to project

The discovery of EBNA3B as a viral-encoded tumour suppressor (see section 1.6.5) has fuelled our interest in uncovering the molecular details of its functions as a transcriptional regulator. However, there have been few further studies focussing exclusively on the mechanisms of EBNA3B-mediated gene regulation. McClellan *et al.* (2013) confirmed our EBNA3B exon microarray results and showed de-repression of *ITGAL* with EBNA3B deletion in LCLs and EBV-infected Mutu III BL cell lines. Transient transfection assays showed that each of the EBNA3 proteins can inhibit EBNA2 transactivation of *ITGAL* promoter (McClellan *et al.*, 2013). Using ChIP and re-ChIP experiments, McClellan and colleagues have also demonstrated that EBNA3B competes with EBNA2 to bind directly at the promoter of *ITGAL* (McClellan *et al.*, 2013). In addition, the same group has also determined that EBNA3B binds at an intragenic enhancer within *CTBP2*, and speculated that EBNA3B facilitates repression of *CTBP2* by preventing enhancer-promoter looping, although experimental results supporting the latter were not actually shown. The studies on EBNA3B thus far have highlighted that EBNA3B plays an essential role in gene regulation, but it is clear that the exact mechanisms of actions of EBNA3B will have to be elucidated through the use of specific and verified EBNA3B target genes, just as we have done for EBNA3A and EBNA3C.

The insertion of F-factor plasmid derived from *Escherichia coli* (*E. coli*) bacteria into various sites of wild-type EBV genome – including the B95.8 and Akata strains – has generated bacterial artificial chromosomes (BACs) which can be successfully propagated and manipulated in *E.coli* (Chen *et al.*, 2005; Delecluse *et al.*, 1998; Feederle *et al.*, 2010; Kanda *et al.*, 2004). These EBV BACs are transfected into human epithelial cell lines which are induced to produce viable infectious virus particles, after which the recombinant viruses

are harvested and used to produce B-cell lines for further study. The EBV BAC system has revolutionised the way we study the effects of EBV genes (Allday *et al.*, 2015; Feederle *et al.*, 2010). Modifications of the viral genome using RecA-mediated homologous recombination techniques in prokaryotes (White, 2003; described in detail in section 2.3.3 and section 3.2) have permitted us to generate mutant EBV BAC harbouring viral gene deletions or insertions of epitope tags and/or 4HT-conditional fusion peptides, in a more physiological cellular milieu that is consistent with normal expression of other EBV latent proteins.

In order to study the function of EBNA3B, it is first necessary to verify and identify *bona fide* EBNA3B gene targets. Exon-microarray studies have already revealed a large set of host cell genes that are likely to be regulated by EBNA3B (Chen *et al.*, 2006; White *et al.*, 2010). However, since the exon-microarray studies examine the steady state expression of genes in established LCLs/EBV-infected BL31s, information is lacking on whether EBNA3B-mediated regulation of these genes is direct or indirect (as a consequence of activation or repression of other cellular genes). In addition, it is unclear whether clonal selection occurring during outgrowth of 3BKO and WT LCLs used for the microarray studies has an effect on the transcriptome data obtained, potentially leading to genes being erroneously identified as EBNA3B targets. One way of overcoming this effect is to establish and utilise a 4HT-conditional EBNA3B system – a tool that can be used to study the gene regulatory function of EBNA3B by comparing differences in the expression of cellular genes between 4HT permissive and non-permissive states within the same cellular background. An advantage of using this system is that the expression of the conditional protein in the permissive state would be similar to physiological levels. Other potential uses of such a system include the ability to track the kinetics of gene regulation of one or several genes over a set period (such as 30 days), and performing unbiased microarray or RNA-seq studies to compare cellular gene expression in cells with functional EBNA3B to those without. These have already been done for conditional EBNA3C-HT and EBNA3A-HT cell

lines (Bazot *et al.*, 2015; Maruo *et al.*, 2003, 2006; Skalska *et al.*, 2010; and our unpublished data).

Another method of examining the function of EBNA3B is to study its localisation on chromatin with respect to its regulated genes. Although ChIP studies using antibodies specific to each EBNA3 have been performed, there have been reports of cross-reactivity between these antibodies (McClellan *et al.*, 2012). To avoid issues of cross-reactivity, our lab and others have engineered recombinant EBV with epitope tags fused to the C-terminus of each EBNA3 protein, and are able to successfully immunoprecipitate each tagged protein using robust antibodies against the epitope tags that do not cross-react with other untagged EBNA3s (Bazot *et al.*, 2015; Ohashi *et al.*, 2015; Paschos *et al.*, 2012; Wang *et al.*, 2015). However, although the reasons remain unknown, addition of C-terminal fusions to EBNA3B has unfortunately affected the splicing or expression of EBNA3C (our unpublished observations). As a result, successfully outgrown epitope-tagged EBNA3B cell lines struggle to proliferate because they express very low or negligible levels of EBNA3C, making it difficult to collect enough material for ChIP experiments (Ohashi *et al.*, 2015 and our unpublished data). In addition, since it is highly likely EBNA3B co-operates and interacts with EBNA3C to regulate a subset of its target genes, low levels of EBNA3C can also skew the overall genomic distribution of EBNA3B binding. Therefore, an attempt at fusing peptides at the N-terminus of EBNA3B is required.

1.11. Aims

The overall objective of the project is to study and understand the function of EBNA3B and its role in gene regulation. Hence, the specific aims of the project and the strategies undertaken are:

1. Determine *bona fide* EBNA3B-regulated genes by establishing conditional EBNA3B cell lines.

- a. This will involve generating a recombinant EBV expressing 4HT-conditional EBNA3B by fusion of ERT2 to the N-terminus of EBNA3B.
 - b. Transfection of episomal vector expressing Dox-inducible EBNA3B into 3BKO LCLs as an alternate approach to 1a.
 - c. Using the 4HT-conditional/Dox-inducible EBNA3B system to examine the expression of multiple genes (using Taqman Gene Expression Array Cards) under permissive and non-permissive conditions.
2. Map out the genome-wide binding regions of EBNA3B via using ChIP and ChIP-seq techniques.
 - a. Generation of recombinant EBV expressing N-terminally epitope-tagged EBNA3B suitable for immunoprecipitation
 - b. Genome-wide association studies using N-terminal tagged EBNA3B cell lines for ChIP-seq studies
 - c. Comparisons of EBNA3B-bound regions with publicly available histone modification and transcriptional factor binding data.
 - d. Identify EBNA3B target genes by comparing current available chromatin conformation data on GM12878 LCL with our EBNA3B ChIP-seq data and the list of EBNA3B-regulated genes verified by the conditional EBNA3B system.
3. Examine the mechanisms of gene regulation of specific, validated target genes (TERT and TRAIL).
 - a. By observing the expression of these EBNA3B target genes during the outgrowth of WT, 3BKO and 3BRev EBV-infected B-cells.
 - b. Focus on the epigenetic mechanisms underlying EBNA3B-mediated regulation of its target genes by examining the chromatin landscape at the gene promoter and at any EBNA3B-bound distal regions.
4. Explore the biological implications of EBNA3B regulation of the TERT gene.
 - a. By examining the telomere length during the outgrowth of WT, 3BKO and 3BRev EBV-infected B-cells.

2 Materials and Methods

2.1 Solutions and Buffers

All solutions and buffers are made up with ultrapure Milli-Q® water (mQH₂O), unless otherwise stated. All chemical grade reagents were supplied by BDH chemicals (VWR) and were of AnalR grade purity unless otherwise stated.

Table 2.1. Solutions used in this study

Solution	Composition
Agarose gel loading buffer (6x)	25% (w/v) Ficoll-400 25mM EDTA 0.1% (w/v) Bromophenol blue 0.1% (w/v) Xylene cyanol
Alkaline SDS	1% (w/v) SDS 0.2M NaOH
Ammonium acetate (7.5M)	Add 800ml of mQH ₂ O to 500g of Ammonium acetate (Sigma)
Ammonium Persulphate (10%)	Dissolve 1g in 10ml of water, aliquot and store at -20°C
Ampicillin (500X)	Dissolve 250mg in 50ml of 70% ethanol, store at -20°C
Blocking solution for Western Blot	5% (w/v) of skim milk powder in PBS-Tween
ChIP Protease inhibitors	Aprotinin 1mg/ml (1000x working concentration): Dissolve 10mg in 10ml of water Pepstatin A 1mg/ml (1000x working concentration): Dissolve 10mg in 10mg of 100% ethanol
ChIP Sonication/IP buffer	50mM Hepes (pH7.9) 140mM NaCl 1mM EDTA 1% (v/v) Triton-X100 0.1 % (w/v) Sodium deoxycholate 0.1% (w/v) SDS Add ChIP protease inhibitors immediately before use
ChIP swelling buffer	25M Hepes (pH7.8) 1.5MM MgCl ₂ 10mM KCl 0.1% (v/v) NP-40 1mM Dithiothreitol (DTT)
ChIP low salt wash buffer	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 wash buffer	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 wash buffer	10 mM Tris-HCl (pH 8) 250 mM LiCl 1 mM EDTA (pH 8) 1 % (v/v) NP40 1 % (w/v) Sodium deoxycholate
Chloramphenicol (1000X)	Dissolve 625mg in 50ml of 70% ethanol, store at -20°C
Doxycycline (1000X)	Dissolve 5mg in 5ml mQH ₂ O, aliquot and store in the dark at -20°C
GTE	40mM glucose 50mM Tris HCl (pH8) 50mM EDTA Autoclave and store at 4°C
4-Hydroxytamoxifen (4mM, 10,000X)	Resuspend 5mg of 4HT (Sigma) in 3.22 ml of 100% ethanol. Aliquots were stored at -20°C in the dark.
KPB (20X)	4.6 % (w/v) KH ₂ PO ₄ 24.3 % (w/v) K ₂ HPO ₄ Sterile filtered and stored at room temperature
LB	1% (w/v) Bacto-tryptone 0.5% (w/v) Yeast extract 1% (w/v) NaCl Made up with mQH ₂ O, autoclaved, and stored at room temperature
LB-Agar	1%(w/v) Tryptone 0.5% (w/v) Yeast Extract 1% (w/v) NaCl 1.5% (w/v) Agar Made up with mQH ₂ O, autoclaved, and cooled to 55°C in a water bath before pouring. Stored at 4°C.
LB Agar + 5% sucrose	As above, but allowed to cool before adding sucrose solution
IP lysis buffer	50 mM Tris-HCl (pH 8.0) 150 mM NaCl 10% Glycerol 0.5% Triton X-100 2 mM Phenylmethylsulfonyl fluoride (PMSF) 2 mM cOmplete protease inhibitor cocktail (Roche)
Kanamycin (1000X)	Dissolve 250mg in 50 ml of mQH ₂ O, aliquot and store at -20°C
Phosphate Buffered Saline (PBS)	10 mM Phosphate buffer (pH 7.4) 2.7 mM KCl 137 mM NaCl
PBS/1%BSA	Dissolve 1% (w/v) BSA in PBS Sterile filtered (NALGENE® sterile filter units), stored at 4°C
PBS-Tween	0.1 % (v/v) Tween-20 (Sigma, UK) in PBS
Peptide 6	Resuspend 50 mg of peptide 6 in 50ml OptiMEM I (Life technologies) in an autoclaved flask under sterile conditions. Cover with aluminium foil and leave overnight in fridge. Aliquot in sterile cryovials and store at -80°C.
Phenylmethylsulfonyl fluoride (PMSF)	Resuspend 100mM of PMSF in isopropanol. Aliquot and store at -20°C
Protease inhibitor cocktail (cOmplete)	1 tablet of cOmplete™ Protease Inhibitors Cocktail (Roche) was dissolved in 2ml ddH ₂ O. Stored at -20°C for up to 2 months.

RIPA lysis buffer (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 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 (SB)	1.2 % (w/v) Bactotryptone 2.4 % (w/v) Yeast extract 0.4 % (v/v) Glycerol 5 % (v/v) 20x KPB (sterile filtered) Autoclave and store at room temperature. Add KPB before use.
TAE buffer (10X)	40mM Tris (pH7.6) 20mM acetic acid 1mM EDTA
TBE buffer (10X)	89 mM Tris 89 mM Boric acid 20 mM EDTA
TE buffer	10 mM Tris-HCl (pH 8) 1 mM EDTA (pH 8)
T ₁₀ E ₂₅	10 mM Tris-HCl (pH 8) 25 mM EDTA (pH 8)
TE saturated butanol	Make up 1:2 solution of TE and butanol in 50ml Falcon tubes, mix by shaking and centrifuge for 5 min at 3000rpm. Only use the top layer.
Tetracycline (500X)	Dissolve 25 mg in 50ml 100% ethanol, store in the dark at -20°C
1.5M Tris (pH8.8)	Dissolve 1.5M Tris base (Sigma) in 700ml mQH ₂ O and adjust pH to 8.8 using concentrated HCl

2.2 General cloning strategy and protocols for DNA manipulation

Designing of constructs and planning of cloning strategy were all done with the aid of Vector NTI® Advance software v11.5.4 (Invitrogen). To construct a new replicon, DNA fragments carrying required features from different sources were identified. Ideally the same restriction enzymes would be used to excise both the fragments; alternatively, restriction enzymes that generate ligation-compatible overhangs were used. Digested DNA fragments were separated on an agarose gel and the desired DNA fragments were cut from the gel, purified and ligated. Ligated DNA were transformed into bacteria via the heat-shock method and successfully transformed bacteria were selected based on their viability and growth under various conditions (i.e. addition of antibiotics and varying temperatures). Unless stated otherwise, DNA purified from viable bacteria were extracted via mini-prep and analysed using diagnostic restriction digests followed by agarose gel electrophoresis. Successfully cloned plasmids were used for further sub-cloning or transfection experiments. A full list of constructs used for this study can be found in Table 2.2.

Table 2.2. Vectors and BACs used in this study

Name	Description	Source
p4487	pKovKan Δ Cm	White (2003)
p8085.2	pcDNA expression vector +EBNA3B-citrine	Dr Rob White
pAA15.2	pcDNA expression vector +N-terminal fusion of eGFP to full EBNA3B sequence (inclusive of introns)	Abdullah Al-Matrouk (Masters Student, 2006)
pDF25-Tet	RecA gene + tetracyclineresistance, temperatue sensitive replication origin	Laloti and Heath (2001)
pUC57-FOXP1-ERT2	pUC57 targeting vector +FOXP1-ERT2	Dr Rob White (synthesised insert)
p4793.1	pBlueScript II KS (+) targetingvector. +eGFP-EBNA3B (incomplete 3B gene)	Dr Rob White
pGY001-pUC57SF3Bsynth1	pUC57 targeting vector. + FS-tag (synthesised)	This study
pGY002-bsSF3B target	pBlueScript II KS (+) targeting vector. +partial FS-EBNA3B	
pGY003-ERT2-3B-target	pBlueScript II KS (+) targeting vector. Carrying partial ERT2-EBNA3B	
pGY004-FS3B-KovKan	pKovKan Δ Cm backbone. + FS-EBNA3B with EBNA3B homology sequences	
pGY005-ERT2-3B-KovKan	pKovKan Δ Cm backbone. + ERT2-EBNA3B with EBNA3B homology sequences	
pGY009-ERT2-3B-linker	pBlueScript II KS (+) targeting vector. +Partial ERT2-EBNA3B with extra linker	

pGY011-ERT2-full3B	pBlueScript II KS (+) targeting vector +full ERT2-EBNA3B	This study
pGY015-pcdna-ERT2-full3B	pcDNA expression vector +ERT2-EBNA3B-citrine	
pHB9	Subclone of the wild type B95.8 BAC containing 6.6 BamW repeat units.	Delecluse, 1998
pHB9-GY04-FS3B	FS3B recombinant BAC	This study
pHB9-GY05-ER3B	ER3B recombinant BAC	

2.2.1 Extraction of Bacterial plasmid

2.2.1.1 Mini-prep

Freshly grown bacterial colonies stored at -80°C were picked in the evening and cultured in 1.5ml of superbroth (SB) containing the appropriate selective antibiotic(s). The SB culture was set up in a 12ml Falcon™ round bottom polypropylene bacterial culture tube (Thermo Scientific) and incubated overnight in a shaking incubator at an appropriate temperature (generally 37°C). The next day, cultures were transferred into Eppendorf tubes which were centrifuged at 4 min at 4000rpm at room temperature. The supernatant was discarded and 70µl of STET was used to resuspend the cells. Next, the cells were lysed by adding 200µl of alkaline SDS while vortexing gently. The mixture was neutralised with 150µl of 7.5M ammonium acetate. The lysis mixture was left on ice for at least 5 min, and then centrifuged for 20 min at 13,000rpm and 4°C. The supernatant was collected in fresh tubes and 240µl of isopropanol was added at room temperature to precipitate the DNA. The sample tubes were inverted several times and then centrifuged for 6 min at 10,000 rpm at room temperature to pellet the DNA. The supernatant was aspirated carefully. The pelleted DNA was washed with 200µl of 70% ethanol at room temperature and centrifuged at 10,000 rpm for 3 min. The ethanol was aspirated and the pellet was left to air dry for 10 min. Finally, the DNA was resuspended in 50µl of TE mixed with 5µg/ml of RNase A (QIAGEN).

2.2.1.2 Column-based Maxi-prep

Expression vectors that were required in large quantity for transfection experiments in eukaryotic cells were purified using the QIAGEN® plasmid maxi kit and QIAGEN-500 tips, according to the manufacturer's protocol. The DNA was eluted in 500µl of sterile autoclaved

TE and the DNA concentration was determined by UV spectrophotometry at 260nm using Nanodrop Lite spectrophotometer (Thermo Scientific).

2.2.1.3 *Maxi-prep – CsCl gradient*

Low copy number plasmids and BACs were purified from bacteria using the large scale caesium chloride (CsCl) bacterial plasmid maxi-prep protocol. A single colony was picked and inoculated in 2ml of SB including appropriate selective antibiotics. The cells were grown in a shaking incubator during the day at 37°C and 500µl of cells were transferred into a conical flask containing either 250ml (plasmids less than 30kb) or 500ml (BACs) of SB and antibiotics. The cells were placed overnight in a shaking incubator at 37°C. The next day, the culture was poured into a 250ml centrifuge bottle and pelleted by centrifugation at 4000rpm for 15 min in a Sorvall GSA rotor. The supernatant was discarded and the pellet was resuspended in 30ml of GTE. 60ml of alkaline SDS was added while swirling the cell suspension, and inverted several times to ensure adequate mixing. Next, 45ml of 7.5M ammonium acetate was added to the lysis mixture while swirling the solution, which was then mixed by inversion. The mixture was cooled on ice for 15 min. Debris from cell lysis was pelleted by centrifugation at 6000 rpm for 30 min at 4°C. The supernatant was filtered using 70µm Corning Falcon™ Cell Strainers (Thermo Scientific). Approximately 0.6 volumes of isopropanol was added to the lysate to precipitate the DNA, and mixed by inversion. The DNA suspension was left at room temperature for 5-10 min and then centrifuged for 10 min at 4000rpm in a Sorvall GSA rotor at room temperature. The cell pellet was gently rinsed with 6.5ml of 70% ethanol, left to dry by leaving the bottle to drain upside down. The pellet was resuspended in 6ml of T₁₀E₂₅ and the solution was made up to 10ml using additional T₁₀E₂₅ when necessary. 10.55g of CsCl was completely dissolved in 10ml of the DNA in T₁₀E₂₅ by swirling and inversion of the tube and 100µl of ethidium bromide (EtBr, 10mg/ml) was added. The solution was transferred to a Quick-Seal® Polyallomer Tube (Beckman) which was heat-sealed and centrifuged for 18h at 60,000rpm and 20°C using the 70.1 Ti rotor in a Beckman or Hitachi Ultracentrifuge.

The next day, supercoiled plasmid/BAC (lower band) was removed from the gradient. 36.9g of CsCl was dissolved in 37.7ml of T₁₀E₂₅ and the resultant solution was added to the DNA retrieved from the column to achieve a final volume of 12ml. 50µl of EtBr was added to the solution which was then transferred to another heat-sealed tube and spun using the same conditions. The following day, supercoiled DNA was collected and EtBr was extracted by repeated (6 times) extraction against T₁₀E₂₅ saturated butanol. DNA was precipitated with half a volume of 7.5M ammonium acetate, 1µl of GenElute™-LPA neutral carrier for ethanol precipitation (Sigma-Aldrich) and 2.5 volumes of 100% ethanol. The DNA mixture was inverted several times and pelleted by spinning at 13,000rpm for 15 min. Supernatant from the centrifugation was discarded and the pellet was washed with 200µl of 70% ethanol. This was centrifuged at 13,000rpm for 10 min. The supernatant was removed and the pellet left to air dry for 5 min. Finally, depending on the size of the dry pellet, between in 200µl to 1ml of sterile TE was used to resuspend the DNA and the final concentration was determined by UV spectrophotometry at 260nm using Nanodrop Lite spectrophotometer (Thermo Scientific).

2.2.2 Restriction Digestion

All restriction digests were performed using enzymes from New England Biolabs (NEB), according to the recommended protocols. Double digests were performed using the most suitable enzyme buffer. In cases where the enzymes do not function in the same buffer, ethanol precipitation (section 2.6.2.1) was done after the first digestion and the DNA was resuspended in an appropriate buffer for the second enzyme. The restriction digests were set up as illustrated in Table 2.3.

Table 2.3. Restriction enzyme digests

Type of digest	Diagnostic digest	Preparatory digest
10x restriction enzyme buffer	1.5 µl	3 µl
20x BSA	0.75 µl	1.5 µl
DNA	1 µl	3 µl
ddH ₂ O	11.75 µl	22.5 µl
Enzyme	0.2 µl	0.2 µl
Total	15 µl	30 µl

2.2.3 Agarose gel electrophoresis

2.2.3.1 *PCR-amplified or restriction digested DNA*

Generally, gel electrophoresis was performed using 0.8-1.0% UltraPure™ Agarose (Thermo Scientific) dissolved in 1x TBE buffer and stained with SYBR® Safe DNA Gel Stain (Thermo Scientific). DNA samples were mixed with agarose gel loading buffer to 1x final concentration. Gels were run at 80V in Mini-Sub® Cell GT Cell (Bio-Rad) for an appropriate length of time. Band sizes were determined by comparison with 100bp DNA ladder, 1kb DNA ladder or PCR marker obtained from NEB. Gels were visualised under blue light and images were captured using a Molecular Imager® Gel Doc™ XR System (Bio-Rad).

2.2.3.2 *For visualisation of pooled ChIP samples*

Samples pooled from 8 individual experiments were mixed with agarose gel loading buffer. Gels were prepared using 2% agarose gel (Bio-Rad Low Range Ultra, #161-3107) dissolved in 1x TAE buffer and stained with SYBR® Gold Nucleic Acid Gel Stain (Thermo Scientific). Gel electrophoresis was performed initially at 80V for 10 min, followed by 100V for an appropriate amount of time. Gels were visualised under blue light and fragments less than 500bp were excised using QIAGEN MinElute Gel Extraction Kit (see section 2.2.5).

2.2.4 Pulsed field gel electrophoresis (PFGE)

Pulsed field gel electrophoresis was performed on the Bio-Rad CHEF DR II chiller system. Gels were prepared using 0.8-1% UltraPure™ Agarose (Thermo Scientific) in 0.5x TBE and were submerged in 2L of 0.5x TBE. DNA band sizes were determined by comparison with a 1:1 mixture of λ DNA-BstEII Digest and λ DNA-Mono Cut Mix size standards from NEB. After the run, the DNA was stained using 0.15µg/ml of EtBr in 0.5x TBE. The gel was visualised under UV illumination and the pictures were taken using Molecular Imager® Gel Doc™ XR System (Bio-Rad). The running conditions were as follows:

Temperature	14°C
Voltage	6V/cm
Initial pulse time	2 sec
Final pulse time	16 sec
Total run time	14-16h

2.2.5 DNA Gel extraction

Desired fragments were visualised under a blue light transilluminator and excised using a scalpel. DNA from the excised gel was purified using QIAGEN Gel Extraction Kit. Enzyme digested DNA or linker-ligated vectors (see section 2.2.7) were resuspended in 30µl of TE buffer.

Samples for ChIP-seq were resuspended in 20µl of ddH₂O after gel purification. The ChIP samples were quantified with Qubit™ quantitation assay using a Qubit™ 2.0 flurometer (Invitrogen), and diluted to 1ng/µl with ddH₂O. The purity of the sample was determined using the Agilent High Sensitivity DNA Assay with an Agilent 2100 Bioanalyzer, according to the manufacturer's protocol.

2.2.6 DNA ligation

DNA ligation between two purified restriction digested fragments was set up on ice with the constituents described below. Approximately 10ng of DNA was used for the vector and 100ng of DNA was used for the insert. The ligation mixture was left at room temperature for 3-4 hours and used immediately for transformation (section 2.2.8).

<u>Components</u>	<u>Volume</u>	<u>Starting concentration</u>
Vector	0.5 µl	20ng/µl (approximate)
Insert	5 µl	20ng/µl (approximate)
10x Ligase buffer (NEB)	1 µl	-
ddH ₂ O	3.5 µl	-
T4 DNA Ligase (NEB)	0.2 µl	400 units/µl

2.2.7 Oligonucleotide linker ligation

New enzyme sites were introduced to a vector by ligating DNA linkers to the ends of restriction digested DNA. Restriction digested vector DNA was resuspended in 22 µl of ddH₂O. 2µl of this was added to 1µl of 10x NEB ligase buffer and made up to 10µl with H₂O.

5µl of this was set aside as negative control, and the other 5µl as used as a positive control with 0.1µl of T4 DNA ligase (NEB) added. 2.5µl of each 100µM oligonucleotide solution and 2.5µl of 10x NEB ligase buffer (NEB) were added to the remaining 20µl of vector DNA. 1µl of this was used as non-ligated control and 1µl of DNA ligase was added to the remaining 19µl. All samples were incubated overnight in a 16°C water bath. The samples were run on a gel and successful linker-ligated DNA were gel-purified as described above (section 2.2.5)

2.2.8 Transformation of DNA into bacteria

2.2.8.1 Heat-shock

Unless stated otherwise, all transformations were done on competent DH5α bacteria using the heat-shock method. The ligated DNA (section 2.2.6 and 2.2.7) was added directly to 50µl of competent DH5α cells and left on ice for 20 min. Next, the cells were incubated in a 42°C water bath for 1 min and then immediately transferred back into ice. 1ml of LB was then added to allow cells to recover, and the cells were left in a shaking incubator at 37°C for 1h or 30°C for 1.5h. Cells were then plated onto agar plates containing the appropriate selective antibiotics (see Table 2.1 for concentrations) and incubated overnight at the appropriate temperature.

2.2.8.2 Electroporation

The integrity of EBV episomes in mammalian cells was determined by performing low molecular weight DNA extraction (section 2.7), rescuing the episomes by electroporating the DNA into competent bacteria, and then running mini-preps (section 2.2.1.1) to obtain sufficient DNA material for diagnostic restriction digests (section 2.2.2). 17µl of electrocompetent ElectroMAX™ DH10B™ *E.Coli* cells (Invitrogen) were added to 1µl of DNA and 2µl of water on ice. The mixture was transferred into pre-cooled Gene Pulser Electroporation Cuvettes (Bio-Rad) and the cells were electroporated using Gene Pulser electroporator (0.1cm gap, Bio-Rad) with the following conditions:

Capacitance: 25 μ F
Resistance: 200 Ω
Voltage: 1800V

The cells were immediately flushed with 400 μ l of SOC media (Invitrogen) and transferred into a 5ml Falcon™ Round-Bottom Polypropylene Tube (Thermo Scientific). The tube was placed in a shaking incubator for 1h at 37°C. The cell suspension was then plated on 15cm agar plates containing chloramphenicol and incubated overnight at 37°C.

2.3 Construction of recombinant EBV BAC

2.3.1 Design of cloning constructs

RecA-mediated homologous recombination between wild-type EBV bacterial artificial chromosome (pHB9 BAC, Delecluse *et al.*, 1998) and a shuttle vector providing the the sequence of interest (pKov-Kan Δ Cm, White, 2003) was used to engineer N-terminal fusion proteins of EBNA3B in the viral genome (method described in section 2.3.3, adapted from Lalioti and Heath, 2001). The modified human oestrogen receptor ligand binding domain that is responsive to 4HT but insensitive to oestrogen (ERT2) was first described and characterised by Feil *et al.* (1997). The FLAG-StrepII-StrepII (FS) tag sequence was acquired from Gloeckner *et al.* (2007). In order to obtain a final shuttle vector carrying the sequences of ERT2 or FS tags inserted in frame and retaining the first codon of EBNA3B, we designed and used several high copy number vectors (pUC57 and pBluescript) that facilitated the cloning and isolation of desired sequences from *E.Coli* bacteria. For convenience, we planned the cloning experiments in a way that allowed concurrent cloning of EBV recombinant BACs carrying N-terminal fusions of either ERT2 or FS tag to EBNA3B.

Before my project commenced, Dr Rob White had ordered the synthesis of a pUC57-based vector carrying the full sequence of ERT2 (pUC57-FOXP1-ERT2, Figure 2.1). Isolation of the ERT2 fragment by restriction digestion of pUC57-FOXP1-ERT2 with BamHI and BtgZI results in a 3' overhang c(AGCT) that is compatible with a HindIII-digested 5' overhang (TCGA)_n but will not be cut by either enzymes once these are ligated. We

designed a pUC57 vector (pGY001-pUC57-SF3Bsynth1) carrying DNA sequence starting 70bp upstream of EBNA3B to the first 80bp of EBNA3B gene, with the FS sequence embedded after the start codon of EBNA3B (Nter-FS-linker-3B, Figure 2.1). The FS epitope tag was separated from the start codon (ATG) by a peptide linker (GS) comprising a BamHI restriction site, and another peptide linker separated the FS tag from the rest of EBNA3B by a peptide linker (ASNGGSGEASKKA) encompassing a single HindIII restriction site (Figure 2.1). This ensured that the FS sequence can be excised and replaced with the ERT2 fragment using BamHI and HindIII digestion.

2.3.2 Cloning of pKovKan shuttle vector used for BAC recombineering

The pGY001-pUC57-SF3Bsynth1 vector was synthesised by GenScript. Previously, Dr Rob White had constructed a pBlueScript II plasmid carrying N-terminal fusion of enhanced green fluorescence protein (eGFP) at EBNA3B (p4793.1, Figure 2.2). The eGFP sequence was embedded between two homology domains required for subsequent BAC recombineering steps (i.e. 300bp upstream of EBNA3B and 500bp downstream of the EBNA3B start codon). Using PfiMI and EcoRI digestions, the eGFP sequence was replaced with the FS tag (pGY002-bsSF3B, Figure 2.2). A similar targeting vector was created by replacing the FS tag with ERT2 domain (pGY003-ERT2-3B, Figure 2.2) via restriction digestion of pUC57-FOXP1-ERT2 (BamHI and BtgZI) and pGY001-pUC57-SF3Bsynth1 (BamHI and HindIII). Finally, the pKovKan shuttle vectors carrying either ERT2 (pGY005-ERT2-3B-KovKan) or FS-tagged EBNA3B (pGY004-FS3B-KovKan) were created by digestion of pKovKan Δ Cm (White, 2003), and either pGY002-bsSF3B or pGY003-ERT2-3B, with Sall and SacI enzymes (Figure 2.1 and Figure 2.2).

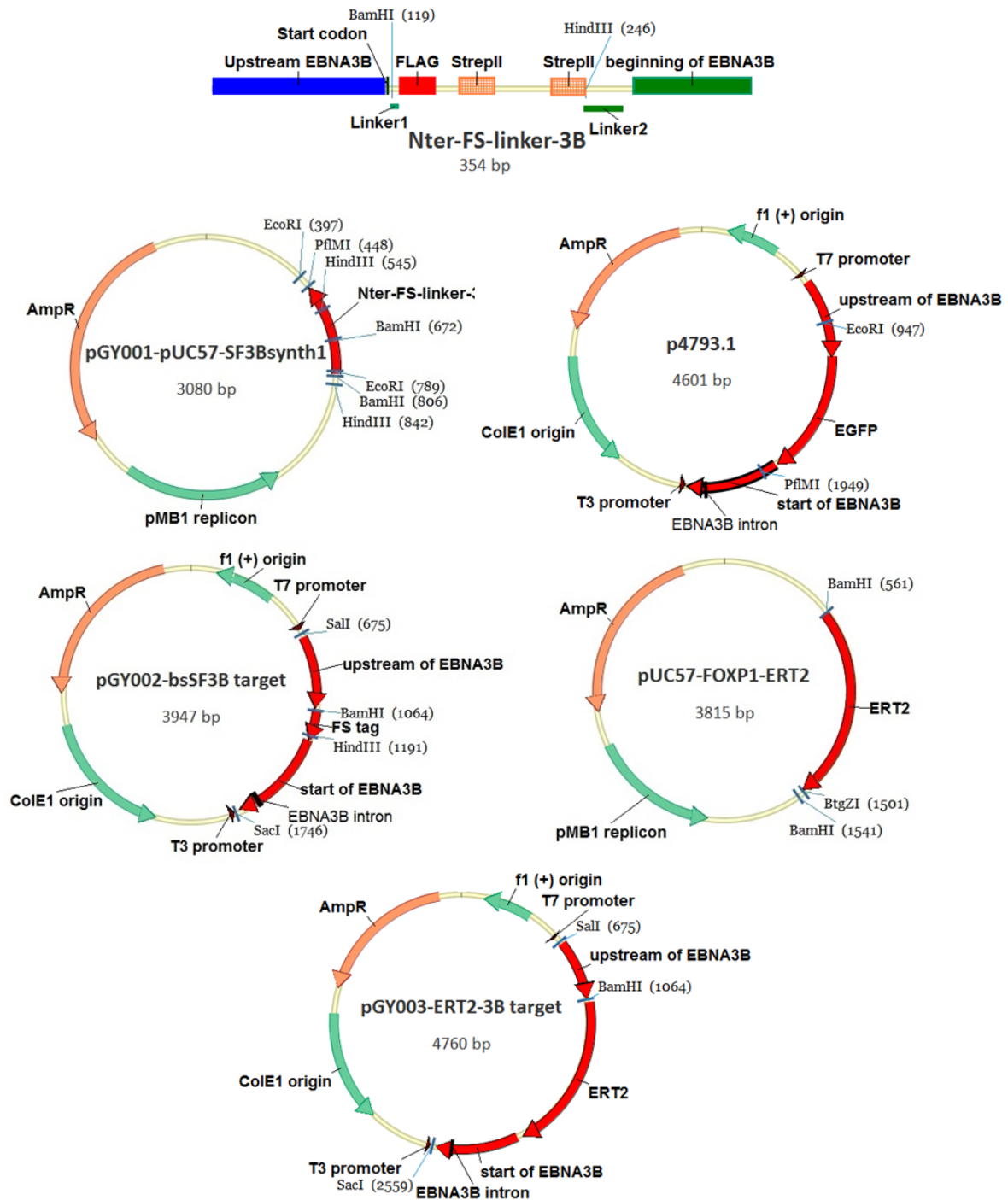


Figure 2.1. Targeting vectors used for BAC recombineering

Refer to text for vector descriptions. For each vector, the insert sequences are labelled as red arrows, origins of replication are labelled green and the ampicillin resistance (AmpR) gene is labelled with a peach arrow. The start of EBNA3B refers to the sequence starting immediately after the ATG of EBNA3B gene to a position 500bp downstream in the B95.8 genome. The upstream sequence refers to 300bp immediately upstream of ATG of EBNA3B in the B96.8 genome. Restriction enzyme digestion sites used for the cloning process are shown.

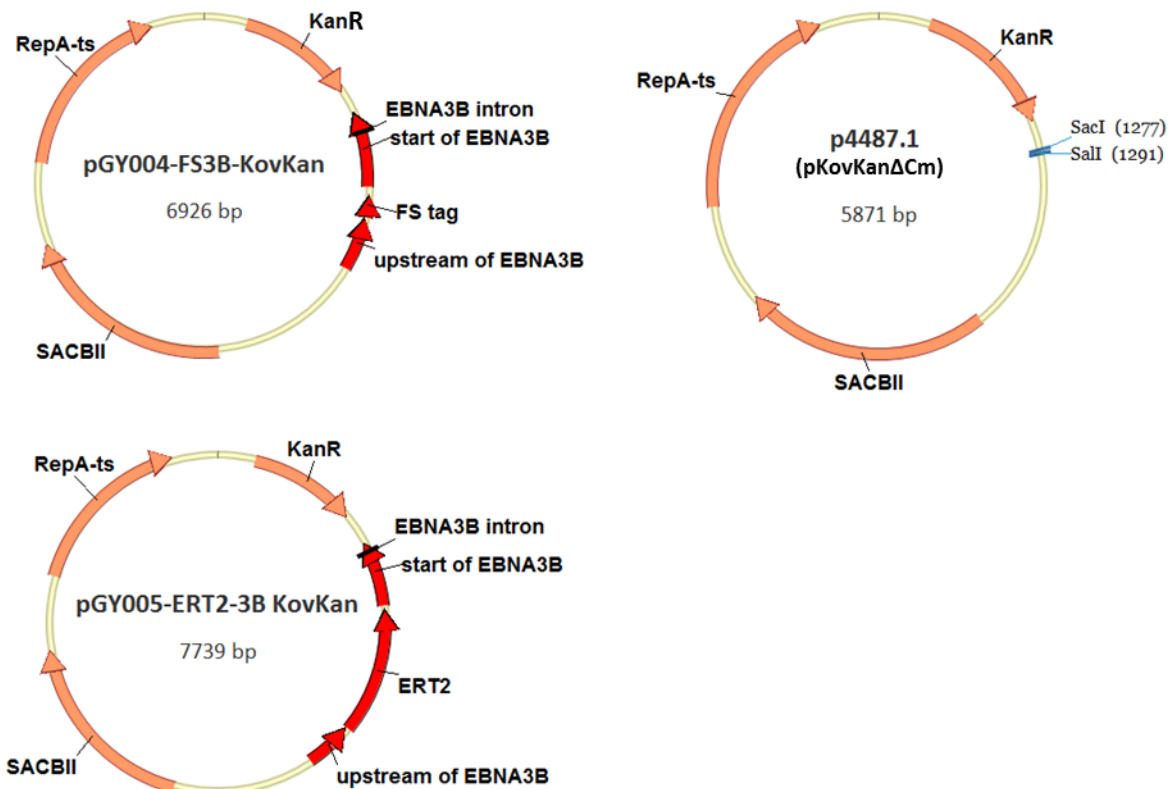


Figure 2.2. pKovKan Δ Cm-based shuttle vectors used for BAC recombineering.

Refer to text for details. Inserts are highlighted with red arrows. Genes coding for temperature-sensitive RepA replicon protein (RepA-ts), kanamycin resistance (KanR), levansucrase (SACBII) are labelled with peach arrows. Restriction sites used for cloning of pGY004 and pGY005 are labelled on the pKovKan Δ Cm backbone.

2.3.3 RecA-mediated homologous recombination

RecA-mediated homologous recombination was performed as described by Lalioti and Heath (2001) and White (2003) and a summary of the method is illustrated in Figure 2.3.

2.3.3.1 Making competent bacteria

To begin, bacteria containing pHB9 (our subclone of p2089 or maxi-EBV) were made competent using CaCl_2 . pHB9 is a B95.8-derived BAC that carries the F factor origin of replication, chloramphenicol resistance, an eGFP gene that is controlled by the CMV immediate-early promoter, and hygromycin resistance gene (controlled by SV40 promoter) used as a selection marker in eukaryotic cells. DH10B *E. Coli* were grown in 1ml SB with the appropriate antibiotic (Chloramphenicol, Cm) and at the appropriate temperature (37°C). On the next day, 100 μ l of the overnight culture was diluted in 10ml of SB with antibiotics and

grown at 37°C for 3-4h. The remaining cells were mini-prepped to check the integrity of the BAC. After 3-4h, the cells were placed on ice for 30 min, and then centrifuged for 15 min at 4000rpm and 4°C. From this point on all work was done in the cold room at 4-6°C. The supernatant from the pelleted cells was discarded and the cells were resuspended in 2ml of 100mM CaCl₂. The cells were then incubated on ice for another 30 min and then centrifuged using the same conditions. The supernatant was again discarded and cells were resuspended in 200µl of 100mM CaCl₂ with 15% glycerol. Aliquots of 50µl were snap frozen using dry ice and ethanol, and stored at -80°C.

2.3.3.2 Preparation of required plasmids

Plasmid DNA from DH5α bacteria transformed with either pKovKan (with inserts) or pDF25-Tet vectors were extracted using the mini-prep protocol (see section 2.2.1.1). The bacteria were incubated overnight at 30°C in a shaking incubator. The pKovKan shuttle vectors carry the kanamycin (Kan) resistance gene and a temperature-sensitive replication origin, hence they can propagate at 30-33°C, but are unable to replicate at 43°C. The pKovKanΔCm vector also carries the levansucrase gene (sacB) which serves as a negative selection marker in the presence of sucrose. pDF25-Tet carries the RecA recombinase and tetracycline (Tet) resistance genes and can only propagate at the permissive temperature of 30-33°C.

2.3.3.3 Transformation of pHB9 with pKov-Kan shuttle vector and pDF-Tet

100µl of competent bacteria harbouring pHB9 were transformed with pDF25-Tet and the pKovKan-based plasmids. Competent bacteria cells were defrosted on ice, and then mixed with 1µl of the pKovKan shuttle plasmid (either pGY004-FS3B-KovKan or pGY005-ERT2-3B-KovKan) and 5µl of pDF25-Tet. The mixture was incubated on ice for 30 min and then heat shocked at 42°C for 1 min. 1ml of LB was added immediately and the mixture shaken for 1.5h at 30°C. Next, 100µl of transformed bacteria were plated on LB agar plates containing Cm, Kan and Tet and then incubated overnight at 30°C to select for bacteria containing the BAC and the two transformed plasmids. The next day, a pool of colonies was

picked into 1ml LB and 200µl of this was immediately plated on plates containing Cm and Kan, which were incubated at 43°C overnight. To test the co-integrants, more than 8 large colonies were picked the next day and grown overnight in 2ml SB at 43°C. 0.5ml was kept as glycerol stock and DNA was extracted from 1.5ml of these cells via mini-prep. The integrity and content of the co-integrants were tested using restriction digestion with appropriate enzymes, followed by analysis of the digested products on a pulsed field gel (see Chapter 3, section 3.3 for details).

2.3.3.4 Resolving co-integrant

DH10B bacteria harbouring successful co-integrant clones (in either directions) were made competent according to the protocol above (section 2.3.3.1) except these cells were grown at 43°C on agar plates containing Cm and Kan. Glycerol stocks of the non-competent bacteria were made by adding equal volume of 50% glycerol to the culture. These were snap frozen and stored at -80°C. 2µl of pDF25-Tet were transformed into 50µl of competent co-integrant cells according to the protocol described above (section 2.3.3.3). 100µl of the transformation mixture was plated on Cm and Tet agar plates and incubated overnight at 30°C. The next morning, using inoculation loops, the colonies were picked and pooled into 1ml SB containing 5% sucrose (Suc), Cm and Tet. These cells were incubated with shaking at 30°C. After 4h, 50-100µl of the culture was transferred in 1ml of fresh SB containing Suc, Cm and Tet and then incubated with shaking at 30°C for another 4h. This procedure was repeated and the bacterial culture left overnight in a shaking incubator at 30°C. The next day, 5µl of the culture was streaked (in a manner that progressively diluted the bacteria) onto 14cm agar plates containing Cm and 5% Sucrose and incubated overnight at 43°C. Around 20 colonies were picked and replica plated onto Cm+Kan and Cm only plates which were incubated overnight at 37°C. Bacteria carrying co-integrant BAC with a SacB mutation would be able to grow both plates, whereas successful recombinants or revertant BACs would grow only on Cm plates. Finally, in order to distinguish the correct recombinants from revertant BACs, DNA was extracted from the bacteria by mini-prep and tested via restriction

digestion and PFGE. Using this method, I successfully constructed pHB9-GY04-FS3B and pHB9-GY05-ER3B recombinant EBV BACs respectively carrying fusions of FS tag and ERT2 domain at the N-terminus of EBNA3B (see Chapter 3, section 3.3 for details).

2.4 Construction of ERT2-EBNA3B-citrine construct

An oligonucleotide linker was designed to introduce a BtgZI digestion site in the pGY003-ERT2-3B targeting construct, with a 5' Sall overhang and a 3' BamHI overhang. This linker (Figure 2.4A) was ligated to the pGY003-ERT2-3B targeting plasmid to create another targeting construct (pGY009-ERT2-3B-linker, Figure 2.4B) with the linker inserted immediately before the ERT2 domain, according to the protocol described in section 2.2.7. The second exon of EBNA3B was excised from pAA15.2 (constructed by Abdullah Al-Matrouk, Figure 2.4B) and inserted into pGY009-ERT2-3B-linker in order to create pGY011-ERT2-full3B targeting vector carrying ERT2 domain within the full EBNA3B gene (Figure 2.4B). This was done by restriction digestion of both plasmids with NotI and SpeI. To create a pcDNA3 expression vector, the ERT2-EBNA3B fragment was excised from pGY011-ERT2-full3B using BtgZI and SfiI and inserted into p8085.2, a pcDNA3 expression vector carrying citrine fused to the C-terminus of EBNA3B (Figure 2.4B). For this, p8085.2 was digested with HindIII, to create a overhang compatible with BtgZI, and SfiI. The final expression vector (pGY015-pcdna3-ERT2-full3B) expressing the full ERT2-EBNA3B-citrine protein was purified and used for transfection experiments in U2OS cells.

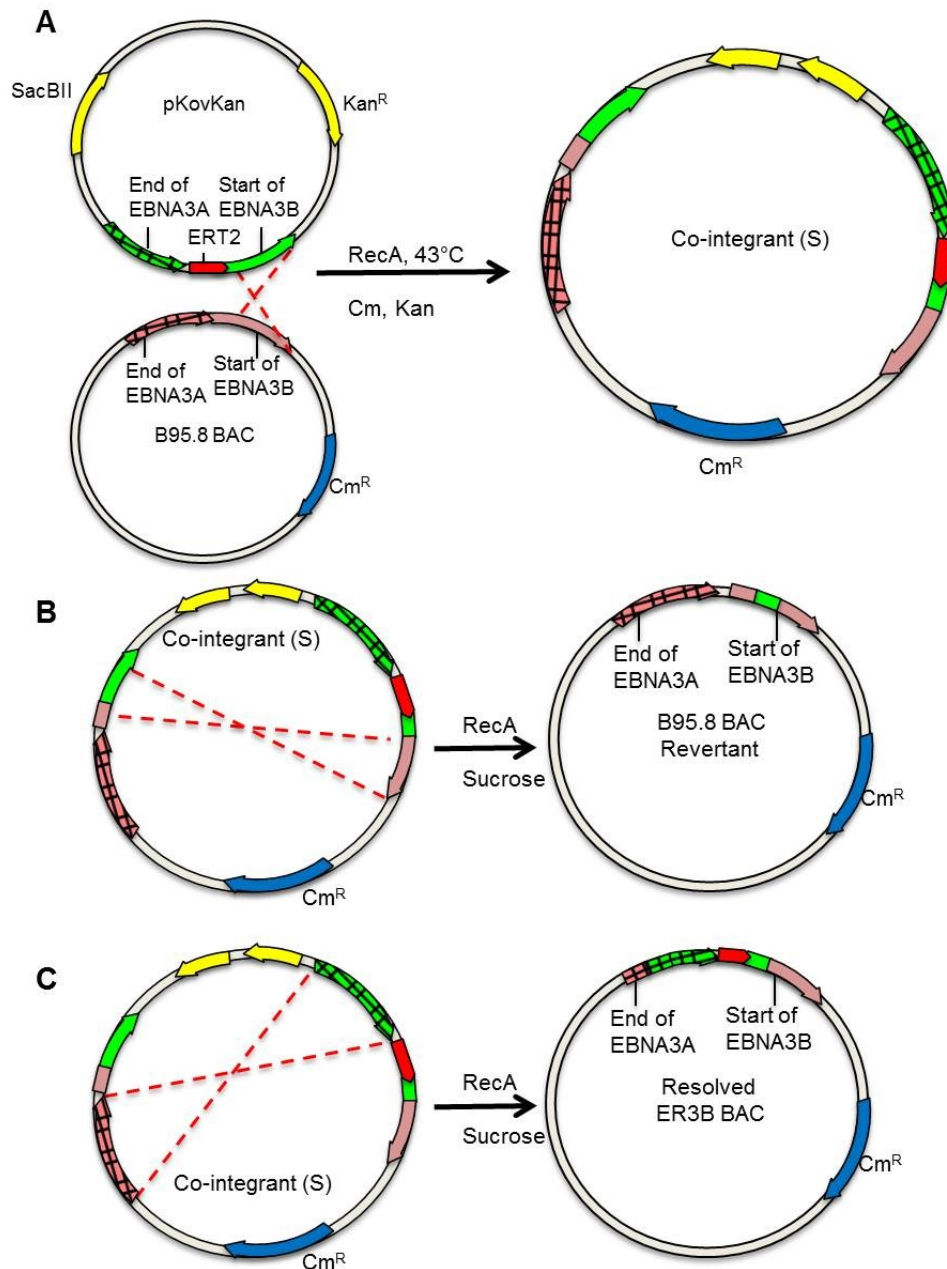


Figure 2.3. Schematic showing the incorporation of ERT2-tag at the N-terminus of EBNA3B in B95.8 EBV-BAC via RecA-homologous recombination.

(A) The ERT2 sequence was cloned into a shuttle plasmid pKovKan (White, 2003) which contains the DNA sequence from the end of wild-type EBNA3A to the start of EBNA3B. The shuttle plasmid and RecA gene are co-transformed into competent E.coli DH10B cells containing the wild-type B95.8 EBV-BAC. Homologous recombination occurs either at the end of EBNA3A or the start of EBNA3B, forming co-integrant BACs in one of two different orientations. Only one configuration (S) is shown in this figure. These are then selected by growth on Chloramphenicol (Cm) and Kanamycin (Kan) at 43°C (non-permissive for the origin of the pKovKan shuttle plasmid). (B) & (C) The co-integrants were resolved to form either the required recombinant, or the original BAC by a second RecA-mediated recombination. These are selected by growth at 43°C on Cm and 5% sucrose. Random mutations in the SacBII gene will produce co-integrants that would survive this selection and would be screened out by replica plating onto kanamycin.

Figure adapted from White (2003) with some modifications

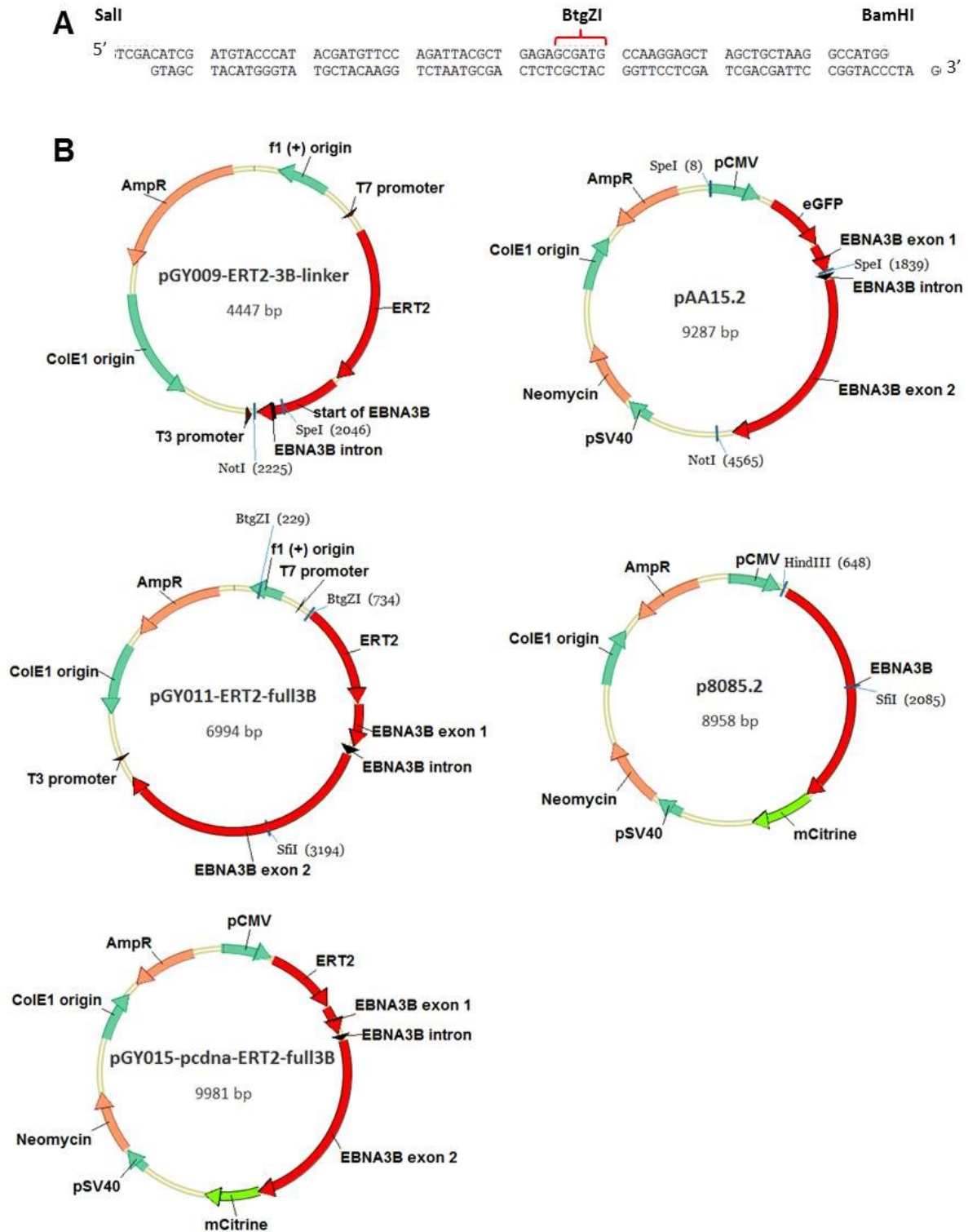


Figure 2.4. (A) Oligonucleotide linker and (B) vectors used to create ERT2-EBNA3B-citrine expression construct

(A) Sequence of linker DNA designed to include BtgZI restriction site into pBlueScript vectors. Description found in section 2.4 of text. (B) Vectors are described in detail in section 2.4. Inserts are illustrated with red arrows and the ampicillin and neomycin resistance genes are labelled with peach arrows. Sequences for citrine are shown with green arrows. Restriction digestion sites used for the cloning process are shown.

2.5 Cell culture

2.5.1 Cell lines

Most experiments were performed on B-cells in suspension cultures. Adherent HEK-293 and U2OS cells were mainly used for virus production and transient transfection experiments respectively.

2.5.1.1 Adherent cells

1. Human embryonic kidney (HEK)-293 cells were derived from human embryonic kidney fibroblasts that are transformed with sheared adenovirus 5 DNA.
2. U2OS cells were derived from a culture of human osteosarcoma tissue taken from a 15-year-old girl.

2.5.1.2 Suspension cells

1. Primary B-cells are CD19+ lymphocytes that are isolated from human peripheral blood (protocol described in 2.10).
2. Lymphoblastoid cell lines (LCLs) are derived from EBV infection of peripheral blood lymphocytes (PBLs) or of primary B-cells.
3. Raji cells were derived from B-cells of a patient with EBV-positive Burkitt's lymphoma.
4. BL31 is an EBV-negative Burkitt's Lymphoma cell line that can be infected with wild-type and recombinant EBV efficiently.

2.5.1.3 Table of cell lines used

Specific descriptions of each cell line used in this project and any associated references are found in Table 2.4

Table 2.4. Cell lines used in project

Name	Cell type	Description	Reference
293 Producers			
HEK-293 ER3B C3	HEK-293	ER3B EBV producer, clone C3	This study, Chapter 3
HEK-293 ER3B C4	HEK-293	ER3B EBV producer, clone C4	
HEK-293 FS3B C8	HEK-293	FS3B EBV producer, clone C8	This study, Chapter 4
HEK-293 FS3B C16	HEK-293	FS3B EBV producer, clone C16	
HEK-293 FS3B E1	HEK-293	FS3B EBV producer, clone E1	
LCLs from infection of PBLs			
D11 WT LCL	LCL	Wild-type B95.8 EBV infection of Donor 11 or Donor 13 PBLs	
D13 WT LCL	LCL		
D11 3BRev LCL	LCL	3BRev EBV infection of Donor 11 or Donor 13 PBLs	
D13 3BRev LCL	LCL		
D11 3A-TAP LCL	LCL	TAP-tagged EBNA3A in LCLs (Donor 11 or Donor 13)	
D13 3A-TAP LCL	LCL		
D11 3C-TAP LCL	LCL	TAP-tagged EBNA3C in LCLs (Donor 11 or Donor 13)	(Paschos <i>et al.</i> , 2012)
D13 3C-TAP LCL	LCL		
D11 FS3B LCL C8	LCL	FS3B EBV (clone 8) infection of Donor 11 PBLs	This study, Chapter 4
D13 FS3B LCL C16	LCL	FS3B EBV (clone 16) infection of Donor 13 PBLs	
D13 FS3B LCL E1	LCL	FS3B EBV (clone E1) infection of Donor 13 PBLs	
D28 ER3B LCL C18	LCL	ER3B EBV infection of Donor 28 PBLs	This study, Chapter 3
D1.4 3BKO LCL	LCL	3BKO EBV (clone C14) infection of Donor 1, 2, 13 and JK1 PBLs	
D2.4 3BKO LCL	LCL		
D13.4 3BKO LCL	LCL		
JK1 3BKO LCL	LCL		
D1.5 3BKO LCL	LCL	3BKO EBV (clone I3) infection of Donor 1 PBLs	
D13.4 RTS2-LUC 3BKO LCL	LCL	Transfection of p8301 (Dox-inducible luciferase gene) into D13.4 3BKO LCLs	From Dr Rob White
D13.4 RTS2-3B 3BKO LCL	LCL	Transfection of p8304 (Dox-inducible EBNA3B) into D13.4 3BKO LCLs	
D1.4 RTS3-LUC 3BKO LCL	LCL	Transfection of p8316 (Dox-inducible luciferase gene) into D1.4, D1.5, D2.4, D13.4 and JK1 3BKO LCLs	This study, Chapter 3
D1.5 RTS3-LUC 3BKO LCL	LCL		
D13.4 RTS3-LUC 3BKO LCL	LCL		
D2.4 RTS3-LUC 3BKO LCL	LCL		
JK1 RTS3-LUC 3BKO LCL	LCL		
D1.4 RTS3-3B 3BKO LCL	LCL	Transfection of p8317 (Dox-inducible EBNA3B) into D1.4, D1.5, D2.4, D13.4	This study, Chapter 3
D1.5 RTS3-3B 3BKO LCL	LCL		

D13.4 RTS3-3B 3BKO LCL	LCL	and JK1 3BKO LCLs	
D2.4 RTS3-3B 3BKO LCL	LCL		
JK1 RTS3-3B 3BKO LCL	LCL		
LCLs from infecting Primary B-cells (1° B-cells)			
GH1+2 3AKO M11 LCL	LCL	3AKO EBV (clone M11) infection of mixed donor (GH1 &GH2) 1° B-cells	This study, Chapter 5
GH1+2 3ARev M4 LCL	LCL	3ARev EBV (clone M4) infection of mixed donor (GH1 &GH2) 1° B-cells	
GH1+2 3BKO C14 LCL	LCL	3BKO EBV (clone C14) infection of mixed donor (GH1 &GH2 or GH3& GH4) or single donor (GH5 or GH6) 1° B-cells	
GH3+4 3BKO C14 LCL	LCL		
GH5 3BKO C14 LCL	LCL		
GH6 3BKO C14 LCL	LCL		
GH5 3BKO I3 LCL	LCL	3BKO EBV (clone I3) infection of single donor (GH5 or GH6) 1° B-cells	
GH6 3BKO I3 LCL	LCL		
GH1+2 3BRev M8 LCL	LCL	3BRev EBV (clone M8) infection of mixed donor (GH1 &GH2 or GH3& GH4) or single donor (GH5 or GH6) 1° B-cells	
GH3+4 3BRev M8 LCL	LCL		
GH5 3BRev M8 LCL	LCL		
GH6 3BRev M8 LCL	LCL		
GH1+2 3CKO M12 LCL	LCL	3CKO EBV (clone M12) infection of mixed donor (GH1 &GH2) 1° B-cells	
GH1+2 3CRev M15 LCL	LCL	3CRev EBV (clone M15) infection of mixed donor (GH1 &GH2) 1° B-cells	
GH1+2 WT A4 LCL	LCL	Wild-type B95.8 EBV (clone A4) infection of mixed donor (GH1 &GH2 or GH3 & GH6) 1° B-cells	
GH3+4 WT A4 LCL	LCL		
GH5 WT A4 LCL	LCL	Wild-type B95.8 EBV (clone A4) infection of single donor (GH5 or GH6) 1° B-cells	
GH6 WT A4 LCL	LCL		
Burkitt's Lymphoma cell lines			
BL31 WT	BL	Wild-type B95.8 EBV infection of BL31 cells	(Anderton <i>et al.</i> , 2008)

2.5.2 Cell culture media

All cell lines were cultured in RPMI 1640 medium (Life Technologies) supplemented with 10% fetal calf serum (FCS; Gibco), penicillin (100units/ml; Sigma-Aldrich) and streptomycin (0.1mg/ml; Sigma-Aldrich) with the following adjustments:

Table 2.5 Cell culture conditions for each cell line

Cell line	Common culture media	Adjustments/Additions
Established LCLs	• RPMI 1640 media • 2 mM L-glutamine • 10% FCS • Penicillin (100units/ml) • Streptomycin (0.1mg/ml)	None
Adherent HEK-293 cells (Obtained from Dr Claire Shannon-Lowe)		Extra 2mM L-glutamine (Life Technologies)
Adherent HEK-293 virus producers		• Extra 2mM L-glutamine • 100 µg/ml hygromycin B (Roche)
Primary B-cells (newly infected)		• 15% FCS (instead of 10%) • Extra 2mM L-glutamine
pRTS-transfected 3BKO LCLs		• Tetracycline-free FCS (Takara Clontech) instead of normal FCS
BL31		• 1 mM sodium pyruvate (Sigma-Aldrich) • 50 µM α-thioglycerol (Sigma-Aldrich)
EBV-infected BL31	• DMEM media (Gibco) • 10% FCS • Penicillin (100units/ml) • Streptomycin (0.1mg/ml)	• 100 µg/ml hygromycin B (Roche) • 1 mM sodium pyruvate (Sigma-Aldrich) • 50 µM α-thioglycerol (Sigma-Aldrich)
U2OS		None

Adherent ER3B HEK-293 virus producer cell lines, ER3B LCLs and U2OS cells that were transfected with pGY15 plasmid (pcDNA3 ERT2-EBNA3B-citrine; see Chapter 3 section 3.4.2) were induced by adding a final concentration of 400nM 4HT (Sigma-Aldrich) to the culture media. 3BKO LCLs that were transfected with doxycycline (Dox)-inducible pRTS3/pRTS2 episomes were induced with between 5ng/ml and 1000ng/ml of doxycycline (Clontech).

2.5.3 Maintenance of cells lines in culture

Cells were maintained in an actively dividing state at 37°C and 10% CO₂ in a humidified incubator. All cells were subcultured in 75cm² flasks.

2.5.3.1 Adherent cells

Adherent cells were split every 3-5 days and diluted 5 to 30-fold depending on the cell density in the parental flask. Old growth medium was removed and the cells washed with 4ml of warm (37°C) phosphate-buffered saline (PBS). After removal of the PBS, 3ml of warm 0.05% Trypsin-EDTA (Invitrogen) was added and the cells were incubated at 37°C until cells have detached. Fresh medium (3ml) was added to quench the action of trypsin and clumps of cells were broken up by vigorous pipetting. Finally, 1-3ml of cell suspension was transferred to a new flask containing 18ml of media.

2.5.3.2 Suspension cells

Suspension cell lines were passaged every 2-5 days and diluted 4 to 10-fold according to the cell density in parental flasks. The flasks were placed upright so as to allow cells to settle to the bottom. Most of the medium was removed from above the cells by vacuum aspiration or pipetting and the remaining concentrated cell suspension was mixed by pipetting. The cells were diluted by transferring the appropriate amount into a new flask containing 18-20ml of fresh medium.

2.5.4 Freezing and thawing cells

Cells were kept in liquid nitrogen for long-term storage. Adherent cells were frozen at the time of passage whereas suspension cells were split one day before freezing. Adherent cells were detached from flask by using Trypsin-EDTA (see section 2.5.3.1), pelleted by centrifugation and washed with normal culture media to remove traces of Trypsin-EDTA. All cells were centrifuged at 1300rpm for 4min at 4°C (MegafugeTM 40R, Thermo Scientific). Approximately 10⁶ cells were resuspended in 1ml of freeze medium containing RPMI 1640 medium (Life Technologies), 20% FCS (Gibco) and 10% dimethyl sulfoxide (DMSO). Cells

were transferred into sterile cryovials (Thermo Scientific) which were placed into a Mr. Frosty™ Freezing Container (Thermo Scientific). These freezing containers were filled with isopropanol and stored at -80°C for at least 24h before the cryovials were transferred into liquid nitrogen tanks.

Cells were recovered from liquid nitrogen by rapidly thawing the cryovials in a 37°C water bath and slowly adding 10ml of normal culture medium to the cells. Cells were pelleted by centrifugation (as above) and the medium removed by vacuum aspiration before addition of fresh culture medium. The resuspended cells were cultured for at least 10 days to allow the cells to recover before being used in any experiments. Unless otherwise stated, cells were discarded after 40 days in culture.

2.5.5 Harvesting cells

2.5.5.1 Suspension cells

All suspension cells were counted and seeded at a density of 3×10^5 cells/ml one day prior to cell harvest, unless stated otherwise. Firstly, cell clumps were broken vigorously by pipetting, and then 50µl of cells were diluted 1:1 with 0.4% trypan blue (Sigma-Aldrich). Viable cells that excluded the blue dye were counted using a haemocytometer under a light microscope. The appropriate amount of fresh medium required to dilute the cells to the correct concentration was calculated and used to seed the cells. The next day, the cells were pelleted at 1300 rpm for 4 min at 4°C (Megafuge 40R, Thermo Scientific) and washed with ice cold PBS. The supernatant was discarded and the cells were resuspended in cold PBS before being transferred to 1.5ml microcentrifuge tubes. Cells were centrifuged at 2000rpm for 5min at 4°C (5424R, Eppendorf) and the supernatant was aspirated. For protein extraction, dry cell pellets were snap frozen in dry ice and then stored at -80°C, or were treated immediately with RIPA buffer. For RNA extraction, RLT buffer (QIAGEN) containing 1% β-mercaptoethanol was added to the cell pellets and the cells were lysed by vigorous

vortexing. Samples were snap frozen in dry ice and stored at -80°C until further extraction using RNeasy Mini Kit from QIAGEN (see section 0).

2.5.5.2 Adherent cells

Adherent HEK-293 virus producer cells were seeded on a 10cm dish at an appropriate density 3-5 days before harvest. To harvest the cells, cell medium was aspirated, and the cell layer was washed carefully with ice cold PBS. The cells were scrapped off using sterile cell scrapers (Greiner Bio-One), resuspended in 1ml of cold PBS and transferred to 1.5ml microcentrifuge tubes. Samples were subsequently treated in the same manner as suspension cells (section 2.5.5.1 above).

2.6 Transfection protocols (mammalian cells)

2.6.1 LID vector delivery

Lipofectin/integrin-targeting peptide/DNA (LID) vector delivery system (Hart *et al.*, 1998) was used to transfect pGY015-pcdna3-ERT2-full3B expression vector into U2OS cells. It was also used to transfect EBV recombinant BACs into HEK-293 cells, and, after outgrowth and selection of successful virus producers, the LID method was used to introduce pCMV-EB1 and pBALF4 expression plasmids (obtained from Dr Rob White) into the HEK-293 virus producer cells in order to induce lytic cycle.

Cells were seeded in 6-well plates 48 hours before transfection. 40µl of 1mg/ml of peptide 6 (synthesised by Genscript; Hart *et al.*, 1998) and 0.75µl of Lipofectin® Reagent (Invitrogen) were mixed and left to incubate at room temperature for 5-10 min. 1µg of the expression vector(s) was made up to a final volume of 100µl of OptiMEM (Invitrogen), and then added to the Lipofectin® and peptide 6 mixture. These were mixed by flicking, and then allowed to stand at room temperature for 30 min. The lipofectin/peptide 6/DNA mixture was made up to a final volume of 1ml with OptiMEM. Culture medium was removed from the cells to be transfected, and 1ml of the LID mixture was added to the cells. The cells were kept in

an incubator at 37°C for at least 6 hours, after which the transfection mix was replaced by 4ml of normal culture media.

The protocol above is described for transfection of 1µg of DNA, but larger scale transfections were performed by increasing the volumes of all reagents proportionally, although final OptiMEM volume was dictated by the size of plate: for transfections in 10cm dishes, the amount used for each reagent is as follows:

<u>Reagent</u>	<u>Small scale (6-well plate)</u>	<u>Large scale (10cm dish)</u>
Lipofectin	0.75 µl	9 µl
Peptide 6	40 µl	480 µl
DNA	1 µg	6 µg
LID incubation volume	100 µl	1.2 ml
Final volume in OptiMEM	1ml	5 ml

The OptiMEM was replaced by 10ml of normal culture media after a minimum of 4h.

For virus production, hygromycin was omitted from the culture medium added after transfection. The cells were incubated for 4 days and virus particles in the media were harvested by passing the media through MF-Millipore 0.45 µm filters (Millex-HA, Millipore).

For experiments involving ERT2-3B-citrine, duplicate wells were transfected with each pcDNA3 vector, and 4HT was added immediately to one of the wells.

2.6.2 Electroporation of 3BKO LCLs

p8304/p8317 (RTS2/3-3B) and p8301/8316 (RTS2/3-LUC, see Chapter 3, section 3.6 for details) plasmids were electroporated into 3BKO LCLs using a NEPA21 electroporator (Sonidel).

2.6.2.1 Ethanol precipitation

DNA required for the electroporation was ethanol precipitated in order to resuspend it at a final concentration of 1µg/µl with sterile TE. First, 40µg of DNA was retrieved from maxi-prepped (CsCl gradient) samples. 0.5 volumes of 7.5M ammonium acetate was added to the DNA, followed by 2.5 volumes of 100% ethanol. The DNA mixture was spun at 15°C for 30

min at 13,000rpm. The supernatant was discarded, and the DNA was washed twice with 70% ethanol and centrifugation at 10,000rpm for 3 mins in room temperature. Finally the DNA was resuspended in sterile TE.

2.6.2.2 Electroporation

40µl of DNA and 1×10^6 cells were mixed in a final reaction volume of 100µl (topped up with OptiMEM) for the electroporation. The cells were placed in EC-002S NEPA Electroporation Cuvettes (2mm gap). Electroporation was performed using the settings below:

Poring Pulse		Transfer Pulse	
Volts	250.0V	Volts	20.0V
Pulse Length	0.3ms	Pulse Length	50.0
Pulse Interval	50.0	Pulse Interval	50.0
Polarity	+	Polarity	+/-
No. of Pulse	2	No. of Pulse	5
Decay	10%	Decay	40%

After electroporation, the cells were transferred into a 12 well plate containing 3ml of warm RPMI that is supplemented with 10%FCS. The cells were mixed well with a pipette, and 150µl of cells were transferred into a 96-well plate, and immediately treated with 1µg/ml of Dox. All cells were left to recover in the incubator at 37°C. Two days later, cells in the 96-well plate were checked for mCherry fluorescence using an inverted microscope. Cells in the 12-well plate were gradually expanded until sufficient volumes were reached for the required experiments.

2.7 Low molecular weight DNA extraction

Low molecular weight DNA was extracted from cells in order to rescue EBV episomes from LCLs and virus producers. Cell pellets were harvested as described in section 2.5.5. The cells were resuspended in 60µl of STET and lysed in 130µl of alkaline SDS. Neutralisation was done by adding 110µl of 7.5M ammonium acetate. The cell lysate was incubated on ice for at least 10 min and the debris was pelleted by centrifugation in 4°C for 30 min at 13,000 rpm. The supernatant was decanted into MaXtract high density tubes

(QIAGEN) and 200µl of phenol:chloroform was added . The tube was inverted vigorously for 1 min to ensure a homogenous suspension, and then centrifuged for 6 min at 10,000rpm. The MaXtract high density gel within the tubes would form a barrier between aqueous phase containing nucleic acids on top, and organic phase containing phenol:chloroform and other contaminants at the bottom of the tube. 200µl of chloroform was added to the top aqueous layer and the tube was inverted and centrifuged again. The top aqueous layer was transferred to a new tube and the DNA was ethanol precipitated (section 2.6.2.1). Finally the DNA was resuspended in 50µl of TE containing 5µg/ml of RNase A.

2.8 Generation of HEK-293 virus producers and production of recombinant EBV

2.8.1 Generation of clonal cell lines

To create virus producers, HEK-293 cells were seeded in a 10 cm dish and transfected with either pHB9-GY04-FS3B or pHB9-GY05-ER3B BACs (refer to sections 2.3 and 2.6.1 for details). In order to select for successfully transfected cells, 150 µg/ml of Hygromycin B (Roche) were added to the growth medium. The cells were left untouched in an incubator at 37°C for 10 days, after which media was changed every 5 days. Transfected cells would form small colonies within the 10cm dishes and up to 20 colonies were picked into 6-well plates for further expansion.

For colony picking, the media was removed and the cells were washed with warm PBS. An autoclaved cloning ring (made by cutting the wide ends off yellow pipette tips) was dipped in autoclaved vacuum grease and placed over the desired colony. 50µl of warm 0.05% Trypsin-EDTA (Invitrogen) was added into the cloning ring and the cells were incubated at 37°C for 3 min. 50µl of normal media was added to the trypsin and the cells were detached by pipetting. Each clone was transferred to a different well containing 4ml of pre-warmed media on a 6-well plate. The colonies were left to grow to confluence, before being expanded to T25 and then T75 flasks. Where possible, early passages of the cells

were frozen and stored at -80°C (see section 2.5.4). Episomes were rescued from these HEK-293 producers and diagnostic digests performed to confirm the integrity and identity of these producers.

2.8.2 Production of recombinant EBV virions and virus titration

Viable and validated HEK-293 producer cell lines were transfected with pBALF4 and pCMV-EB1 to induce virus production, and virus supernatants were harvested as described in section 2.6.1. 111µl of the virus supernatant was added to 1ml of normal media in a 24-well plate. Three serial 10x dilutions (1:10, 1:100 and 1:1000) of the virus media were done by sequentially removing 111µl and adding it to a fresh 1ml of growth medium, with 111µl being discarded from the last well. This generated four wells containing 100, 10, 1 and 0.1 µl of virus respectively. 1×10^5 Raji cells in 0.5ml of media were added to each well, and the cells were left in the 37°C incubator for 2 days. 20 nM of 12-O-tetradecanoylphorbol-13-acetate (TPA) and 5 mM sodium butyrate in 0.5ml of media were added on the third day to induce the lytic cycle. The fourth day, the cells were viewed under an inverted fluorescence microscope and the number of cells expressing eGFP were counted in each well. Green Raji Units (GRU) were determined by counting the number of Raji cells infected per µl of virus used.

2.9 EBV infection of B-cells

Due to the low titres of ER3B and FS3B viruses (refer to section 3.5 and 4.3), two million peripheral blood lymphocyte cells from donors 11 or 13 were resuspended in 2ml of either ER3B or FS3B virus supernatants. For primary B-cell infections, 1-2 million cells in 1.5ml were infected with 0.5ml of supernatant containing the viruses. Generally, the amount of viruses used to infect the cells were normalised according to the lowest titres of the viruses required for the experiment. The titres of virus supernatants used in primary B-cell infection time-courses can be found in Table 2.6 below, and all were obtained from either Dr Rob White or Adam Gillman.

For studies on RNA expression during B-cell outgrowth, three million uninfected cells were harvested for their baseline RNA levels on the day of infection. Thereafter, 0.5ml of cells were harvested every 5 days for RNA extraction and replaced by fresh medium. Similarly, one million cells were harvested pre-infection in the second study examining telomere length of chromosomes during primary B-cell outgrowth. In the initial stages of outgrowth, 0.5ml of cell media was removed every 3 days, and replaced with fresh media. Once the cell growth became robust, the cells were expanded to 20ml cultures in T75 flasks and split regularly every 3-4 days. For GH1+2 donor B-cells, 1 million cells were harvested on days 25, 40, 60, 80, 100 & 120 for genomic DNA extraction, whereas in the case of GH3+4 donor B-cells, cells were harvested on days 10, 35, 75, 95 & 115. Cells from donors GH5 and GH6 were harvested simultaneously on days 0, 30, 40, 67, 87 and 107.

Table 2.6. Titres in Green Raji Units (GRU)/ μ l of viruses used in this study

Virus	Clone	Virus titre (GRU)
3AKO	M16	940
3BKO	M14	230
3BKO	I3	26
3BRev	M8	231
3CKO	M12	520
EBNA3ARev	M4	250
EBNA3BRev	M5	135
EBNA3CRev	M14	320
WT HB9	A4	1095

2.10 Isolation of PBMCs

Peripheral blood mononuclear cells (PBMC) were isolated from buffy coat residues obtained from UK national blood transfusion service (NHS Blood and Transplant, Collindale). Blood was layered over an equal volume of Ficoll-Paque (GE Healthcare), and then spun at 1380rpm (Megafuge 40R, Thermo Scientific) for 30 min at 18°C with brake off. Lymphocytes were extracted from the grey-white interface and added to cooled wash buffer (RPMI + 1% FCS) in a Corning® 250ml bullet tube. The bullet tubes were washed twice with 200ml of wash buffer and centrifuged at 1380rpm for 5 min at 4°C with the brakes off. After removal of the supernatant the cell pellet was resuspended 50ml of growth medium (RPMI

supplemented with 15% FCS and penicillin/streptomycin). The resuspended PBLs were filtered through a 70µm Corning Falcon™ Cell Strainer (Thermo Scientific). The cells were kept on ice and counted. 5 ml of PBLs were kept at 2×10^6 cells/ml in 5ml for fluorescence-activated cell sorting (FACs) analyses.

2.11 Isolation of cells via magnetic sorting

2.11.1 Primary B-cells

Primary B-cells were isolated via magnetic cell sorting with AutoMACS Separator (Miltenyi Biotec). 1×10^9 cells were pelleted at 1100rpm for 5 min in a 50ml falcon tube and the supernatant was discarded. The cells were resuspended in 8ml of MACS running buffer and CD19-positive B-cells were magnetically labelled with 500µl of CD19 microbeads (Miltenyi Biotec). These were mixed well and incubated in a refrigerator for 15 min. The cell suspension was washed with 50ml of PBS/1%BSA and the cells were pelleted by centrifugation. The cells were resuspended in 5ml of MACS running buffer and magnetically labelled CD19+ cells were positively sorted using an AutoMACS Separator (Miltenyi Biotec), according to the manufacturer's protocol. CD19+ cells were resuspended at 2×10^6 cells/ml in growth medium and incubated at 37°C overnight. The next day, B-cell purity was determined using CD20-RPE (Thermo Scientific) antibody staining, followed by FACs analyses.

2.11.2 Dox-inducible mCherry positive cells

pRTS2/3-transfected LCLs were magnetically labelled using CD271 (NGFR) MicroBeads from Miltenyi Biotec according to the manufacturer's protocol. 1.5×10^7 cells were pelleted by centrifugation at 1200rpm for 10min and then resuspended in 90µl of PBS/1%BSA. 50µl each of FcR Blocking Reagent and CD271 MicroBeads were added and the mixture incubated in the refrigerator for 15 min. In the initial validation studies, cells were stained with CD271 (NGFR)-APC antibodies for another extra 5 min in the dark. Excess microbeads, blocking reagent and antibodies were removed by washing the cells twice with PBS/1%BSA. The cells were resuspended in 500µl of PBS/1%BSA. The cell suspension was passed through a 70µm cell strainer prior to magnetic sorting. Dox-induced cells were

positively selected based on the expression of NGFR on their cell membrane using the AutoMACS Separator. In addition, cells that were not treated with Dox (i.e. negative controls) were also incubated with CD271 MicroBeads and then subsequently depleted for NGFR-positive cells using the same procedure. Cells were seeded at 3×10^5 cells/ml and left to recover in the incubator at 37°C for at least two hour post-sort, before being harvested or stained for FACs analysis.

2.12 FACs analysis

2.12.1 Cell Staining protocol

2.12.1.1 Primary B-cells

B-cells were stained with CD20-RPE antibody (Thermo Scientific). In addition, DRAQ7 DNA intercalating stain (Biostatus) was used to exclude non-viable cells. Samples were prepared by washing 250µl of cells (5×10^6 cells) once with 1% BSA/PBS. After centrifugation at 2000rpm for 5 min, the supernatant was removed and the cells were resuspended in 50µl PBS/1%BSA. Human IgG (1ng) was added to the cells and the mixture was mixed and incubated at room temperature for 15 min. Cells were washed with 1ml PBS/1%BSA and after centrifugation, the supernatant was removed. Cell pellets were resuspended in 100µl PBS/1%BSA and 1µl of CD20-RPE antibody was added to each sample in the dark. The cell mixture was incubated for 15-30 min at 4°C in a dark refrigerator. Cells were then washed twice with 1ml PBS/1%BSA and resuspended in 500µl of 500µl PBS/1%BSA supplemented with 4µl of DRAQ7.

2.12.1.2 Dox-inducible mCherry positive cells

pRTS2/3-transfected 3BKO LCLs were stained with a final concentration of 4 µM SYTOX® blue dead cell stain (Thermo Scientific). To do this, 3×10^5 cells were pelleted and washed twice with PBS/1%BSA. The supernatant was removed, and the cells were resuspended in 500µl with appropriate amounts of SYTOX® blue stain added.

2.12.2 Gating Strategy

Flow cytometry was performed using the BD FACSDIVA™ software connected to a Beckton-Dickinson FACSsort flow cytometer and results were analysed with FlowJo™ software. First, gating was done to exclude events with extremes of forward scatter (FSC) and side scatter (SSC) in order to eliminate debris, dead cells and cell clumps. Next, singlets were defined using FSC-W and FSC-A. Live cells were gated using live and dead cell samples that were solely stained with DRAQ7 or SYTOX blue viability stains. 30,000 events of live cells were recorded. Primary B-cell purity was determined by the percentage of live cells that were stained with CD20-APC antibody. Purity of CD271(NGFR)-sorted cells was determined by the percentage of live cells that expressed mCherry.

2.13 Western Blot

Freshly harvested or frozen cell pellets (see section 2.5.5) were lysed using an equivalent volume of 1x RIPA buffer containing 1mM PMSF and cOmplete™ protease inhibitor cocktail (Roche). Cell lysates were incubated on ice for 10 minutes and then centrifuged at 14,000 rpm for 10 min at 4°C (5424R, Eppendorf). The supernatant was collected in a fresh tube and the protein concentration was determined using detergent compatible (DC) protein assay (Bio-Rad). The protein concentration was measured at 750nm on a Helios spectrophotometer (Thermo Scientific), using bovine serum albumin as standard. Between 20 to 40 µg of protein from each sample was added to 5x SDS-PAGE sample buffer, which was diluted to 1x final concentration with the appropriate amount of autoclaved milliQ-H₂O. The protein-SDS-PAGE mixture was heated at 100°C for 5 min to denature the proteins. The heated samples were loaded onto SDS-PAGE gels with the appropriate percentage of polyacrylamide required to analyse the relevant molecular weight of proteins (see Table 2.7 for the composition of each gel). The gels were placed in Mini-PROTEAN® II Electrophoresis Cell (Bio-Rad) containing 1x SDS-PAGE running buffer and run at 150V for 1-1.5h. Protein samples were separated based on their molecular weight and Spectra™ Multicolor Broad Range Protein Ladder (Thermo Scientific) was used as a reference.

Table 2.7. Recipe for stacking and resolving SDS-PAGE gels

*30 % Acrylamide/Bis solution 37.5:1 was obtained from Bio-Rad. ⁺TEMED was obtained from BDH chemicals (VWR). Recipe modified from Harlow and Lane (1988).

Resolving gel (10ml)			
Gel percentage	8%	10%	12%
H ₂ O	4.6ml	4.0ml	3.3ml
30% Acrylamide mix*	2.7ml	3.3ml	4.0ml
1.5M Tris (pH8.8)	2.5ml	2.5ml	2.5ml
10% SDS	100µl	100µl	100µl
10% ammonium persulphate	100µl	100µl	100µl
TEMED ⁺	6µl	4µl	4µl
Stacking gel (5ml)			
Gel percentage	5%		
H ₂ O	3.4ml		
30% Acrylamide mix*	830µl		
1.5M Tris (pH8.8)	630µl		
10% SDS	50µl		
10% ammonium persulphate	50µl		
TEMED ⁺	5µl		

After SDS-PAGE, the gel was overlaid with one layer of ProtranTM nitrocellulose membrane (Schleicher and Schuell) and packed between two triple layers of Whatman chromatography paper. This “sandwich” was held in place by transfer cassettes (Bio-Rad) and soaked in 1x Transfer Buffer. The resolved protein samples on the gel were transferred onto the nitrocellulose membrane at 100V for 1h using Mini Trans-Blot® Electrophoretic Transfer Cell (Bio-Rad). Membranes were blocked with blocking solution (5% milk powder in PBS-T) for 1h at room temperature and then incubated overnight with primary antibodies on a gentle shaker at 4°C (see Table 2.8). The next day, membranes were washed 3 times for 10 min with PBS-T and incubated at room temperature for 1h with the appropriate horseradish peroxidase (HRP)-conjugated secondary antibodies (Table 2.8). The membranes were washed as previously and incubated for 5 min with enhanced chemical luminescence (ECLTM; Amersham; GE healthcare). Finally the luminescence on the membranes was detected on Amersham HyperfilmTM (GE healthcare) autoradiography films which were subsequently developed with a JP-33 (JPI) film processor.

Table 2.8. Antibodies used for Western Blots

Primary antibodies					
Name	Species	Source	Catalogue/Clone number	Dilution Factor	References
EBNA1	Human serum	Prof. Paul Farrell	PF	1:100	
EBNA2	Mouse (monoclonal)	Abcam	ab90543	1:500	
EBNA3A	Sheep (polyclonal)	Exalpha Biologicals	F115P	1:1000	
EBNA3B	Rat (monoclonal)	Dr Elisabeth Kremmer	6C9	1:10	(White <i>et al.</i> , 2010)
EBNA3C	Mouse (monoclonal)	Prof. Martin Rowe	A10	1:10	(Mauders <i>et al.</i> , 1994)
EBNA-LP	Mouse (monoclonal)	Prof. George Klein	JF186	1:100	(Finke <i>et al.</i> , 1987)
LMP1	Mouse (monoclonal)	DAKO	M0897	1:500	
γ -tubulin	Mouse (monoclonal)	Sigma-Aldrich	T6557	1:8000	
Secondary antibodies					
Anti-Sheep (HRP conjugated)	Rabbit (polyclonal)	DAKO	P0163	1:2000	
Anti-Rat (HRP conjugated)	Rabbit (polyclonal)	DAKO	P0450	1:2000	
Anti-Mouse (HRP conjugated)	Sheep (polyclonal)	GE Healthcare	NA931V	1:2000	
Anti-Human (HRP conjugated)	Rabbit (polyclonal)	DAKO	P0214	1:2000	

2.14 Reverse transcription quantitative PCR (RT-qPCR)

RNA was extracted from freshly harvested cells or from cell lysates stored in RLT buffer at -80°C (see section 2.5.5). Generally, RNA was extracted using the RNeasy Mini Kit (QIAGEN) according to the manufacturer's protocol. Total RNA (including micro-RNA, miRNA) from cells harvested during the primary B-cell infections were extracted using the miRNeasy Mini Kit (QIAGEN) according to the manufacturer's protocol. On-column digestion of genomic DNA using RNase-free DNase set (QIAGEN) was performed for all RNA extractions. The RNA was quantified by measuring its absorbance at 260nm using a Nanodrop Lite spectrophotometer (Thermo Scientific). The RNA absorbance ratio at

wavelengths between 260 and 280nm (A260/A280) was used to determine RNA purity and all samples had an A260/A280 ratio between 1.8 and 2.1. Up to 1µg of RNA was reverse-transcribed into cDNA using SuperScript™ III First-Strand Synthesis SuperMix for qRT-PCR kit (Invitrogen), according to the manufacturer's protocol. cDNA synthesis reactions were performed on DNA Engine Dyad® Peltier Thermal Cycler System PTC-220 (MJ Research). For quantitative PCR (qPCR), cDNA was diluted to 2.5ng/µl with RNase-free water (assuming complete conversion of RNA into cDNA) and 10ng of cDNA was used for each qPCR reaction.

2.14.1 PCR quantitation using SYBR Green

The qPCR assays are set up using Platinum SYBR Green qPCR SuperMix-UDG kit (Invitrogen) and 200nM of primers (Table 2.9) in a final reaction volume of 10µl. Assays are set up in MicroAmp® Optical 384-Well Reaction Plates and the real time PCR performed using Fast Mode on QuantStudio 7 Flex (Life Technologies). The qPCR cycling conditions are as follows: one denaturation step at 95°C for 2 min, followed by 40 PCR cycles starting with denaturation at 95°C for 3 sec, followed by primer annealing and elongation at 60°C for 30 sec. Melt curve analyses were performed at the end of every qPCR run to ensure the absence of non-specific amplification. Samples are assayed in triplicates for each qPCR run. RNase-free water instead of cDNA was used as negative amplification control and at least 2 of 3 genes – *GNB2L1*, *RPLP0* and *ALAS1* – were used as endogenous controls. Results of the qPCR were analysed using DataAssist™ v3.01 Software and relative gene expression was calculated using the comparative Ct ($\Delta\Delta C_t$, Schmittgen and Livak, 2008). Error bars in graphs represent standard errors from triplicate qPCR reactions for each sample in a single experiment.

Primers used for the RT-qPCR assays were generally designed using Primer3Plus (publicly available online at <http://primer3plus.com/cgi-bin/dev/primer3plus.cgi>). RT-qPCR primer pairs are designed to amplify across 2 exons, and where possible, one primer sequence will span the exon-exon junction. The amplification efficiency of each primer set is

first determined by using standard curves produced by amplification of seven 5-fold serial dilutions of pooled cDNA samples. Generally, only primer pairs which amplify with efficiency between 90-105% were used.

Table 2.9. Primer sequences used in this study

Gene	Forward Primer (5'-3')	Reverse Primer (5'-3')	Source
ALAS1	GGATTCGAAACAGCCGAGTG	TGACATCATTGTGGCGGAAG	(Bazot <i>et al.</i> , 2015)
CD28	AACCCTAGCCCATCGTCAGG	GGCGACTGCTTCACCAAATC	Dr Rob White, unpublished
CXCL10	CTGCAAGCCAATTTGTCCACGTGT	AGCACTGCATCGATTTGCTCCCCTC	This study
GNB2L1	GCTTGCAGTTAGCCAGGTTT	GAGTGTGGCCTTCTCCTCTG	(Zhang <i>et al.</i> , 2005)
ITGAL	CCAAGTCAAGCACATGTACCA	ACGCTCCACTGGTGTGTG	(McClellan <i>et al.</i> , 2013)
p16INK4A	CCCCTTGCCTGGAAAGATAC	AGCCCCTCCTCTTTCTTCCT	(Kia <i>et al.</i> , 2008)
RPLP0	ACTCTGCATTCTCGCTTCCT	GGACTCGTTTGTACCCGTTG	(Skalska <i>et al.</i> , 2010)
SMAD1	CACAAACATGATGGCGCCT	TGTGGAGGAGGCATGGAACGC	Dr Lydia Eccersley, unpublished
TERT	ATGTCACGGAGACCACGTTT	GCACCCTCTTCAAGTGCTGT	This study
TNFRSF10A	GTACGCCCTGGAGTGACATC	AGCAACGGAACAACCAAAGT	
TNFRSF10D	GGTTGTGGTTGGCTTTTCAT	GTTCCGGACCTCCTCCAC	
TNFSF10	TGAGAACCTCTGAGGAAACCA	CCTCTGGTCCCAGTTATGTGA	

2.14.2 PCR quantitation using TaqMan® Gene expression assays

TaqMan® Gene expression assays (Applied Biosystems) were used to assess the expression of *LAIR1* and *FOXP1* (Table 2.10). For each qPCR reaction, 0.5µl of 20x TaqMan® Gene expression assays (Applied Biosystems) were added to 5µl of Takyon™ ROX Probe 2X MasterMix dTTP and made up to a final reaction volume of 10µl with nuclease-free water. The qPCR assays were performed using Fast Mode on QuantStudio 7 Flex (Life Technologies) with the following cycling conditions: one activation step (for DNA polymerase) at 95°C for 3 min, followed by 40 cycles starting with denaturation at 95°C for 3 sec, followed by primer annealing and extension at 60°C for 30 sec. Subsequent analyses were performed as described above (section 2.14.1).

Table 2.10. RT-qPCR probe assays used in this study

Gene	Probe assay	Source
LAIR1	Hs00253790_m1	Applied Biosystems
FOXP1	<u>Forward primer:</u> CTGGCGGCTACATGGAGAGTGTCA <u>Reverse primer:</u> GCTGCTGCTGTTGAAGCATGAG <u>Probe sequence:</u> CAGZGGCTAZGATGACACCTCAAGT CAGZGGCTAZGATGACACCTCAAGT [Z denotes locked nucleic acid (LNA)-based T]	Designed by Dr Rob White

2.15 TaqMan® Gene Expression Array Cards

Custom 384-well, 8-port Micro-Fluidic TaqMan® Low Density Array (TLDA) cards obtained from Applied Biosystems were used to assess the expression of 48 genes (Table 2.12) in LCLs transfected with Dox-inducible pRTS plasmids (see Chapter 3 section 3.6.6). For each sample, 350ng of RNA was reverse transcribed to cDNA and the resultant mixture made up to 50µl with nuclease-free water. The cDNA was added to 50µl of TaqMan® Universal PCR 2X Master Mix (Applied Biosystems), creating a total reaction volume of 100µl, which was loaded onto each port of the TLDA card according to the manufacturer's protocol. The array cards were inserted into an array card holder which was then placed into a Sorvall/Heraeus bucket for array cards. The array cards were centrifuged twice using Sorvall™ Legend RT centrifuge (Thermo Scientific) at 1200rpm for 1 min. Next, the cards were sealed using Low Density Array Sealer obtained from Applied Biosystems. Real-time PCR was performed using the standard mode on QuantStudio 7 Flex (Life Technologies) with the following cycling conditions: Uracil N-glycosylase (UNG) incubation at 50°C for 2 min and then polymerase activation at 95°C for 10 min, followed by 40 PCR cycles starting with denaturation at 95°C for 15 min and annealing and elongation at 60°C for 1 min. The results were analysed using DataAssist™ and relative gene expression was normalised to endogenous genes (*18S RNA* and *RPLP0*) as described above.

2.16 PCR amplification across splice junctions

PCR amplification across the splice junction between U-exon and exon 1 of *EBNA3B* was done to confirm the presence of mRNA transcripts for wild-type and recombinant ER3B proteins (see Chapter 3, section 3.4.1). RNA were extracted from cells and converted to cDNA as described above (section 0). For all samples, 50ng of cDNA was mixed with 1µM each of forward and reverse primers (Table 2.11) and 12.5µl of GoTaq® Green Master Mix (Promega). The PCR reaction mix was topped up to 25µl using nuclease-free water. Amplification was performed using DNA Engine Dyad® Peltier Thermal Cycler System PTC-220 (MJ Research) with the following cycling conditions: an initial denaturation step at 95°C for 2 min, followed by 25-35 cycles of denaturation at 95°C (30 sec), annealing at 58°C (30 sec) and extension at 72°C (1 min), with a final elongation step at 72°C for 5 min. The amplified products were run at on an agarose gel and visualised using the SYBR® Safe Gel Stain (Thermo Scientific) under a blue-light transilluminator (see section 2.2.3).

Table 2.11. Primers used for PCR amplification across the splice junction between U-exon and EBNA3B

Primer Name	Sequence (5'-3')	Source
U-exon (forward)	CTGGGATGAGCGTTTGGGAGAG	Dr Rob White, unpublished
EBNA3B (reverse)	GTCACATTACCTTGATCTCC	Dr Rob White, unpublished

Table 2.12. TaqMan® Assays (Applied Biosystems) used in this study

Candidate EBNA3B Target Genes		
Gene	Assay ID	Lot Number
ADARB1	Hs00953724_m1	A5917
ASCL1	Hs00269932_m1	A5917
BAG3	Hs00188713_m1	A5917
C13orf15	Hs00204129_m1	A5917
CD28	Hs00174796_m1	A5917
CD72	Hs00233564_m1	A5917
CDH1	Hs00170423_m1	A5917
CRLF2	Hs00845692_m1	A5917
CXCL10	Hs00171042_m1	A5917
CXCR4	Hs00976734_m1	A5917
CXCR7	Hs00604567_m1	A5917
FOXP1	Hs00212860_m1	A5917
GAS7	Hs00245902_m1	A5917
IL10	Hs00174086_m1	A5917
IL19	Hs00604657_m1	A5917
IL6R	Hs01075667_m1	A5917
ITGAL	Hs00158238_m1	A5917
ITGB1	Hs00236976_m1	A5917
ITGB3	Hs00173978_m1	A5917
LAG3	Hs00158563_m1	A5917
LAIR1	Hs00253790_m1	A5917
LILRA4	Hs00429272_g1	A5917
MAP4K4	Hs00377415_m1	A5917
MCAM	Hs00174838_m1	A5917
NR2F2	Hs00819630_m1	A5917
ONECUT2	Hs00191477_m1	A5917
RUNX1	Hs01021971_m1	A5917
SMAD1	Hs00195432_m1	A5917
SORT1	Hs00907094_m1	A5917
ST3GAL5	Hs00187405_m1	A5917
STAT4	Hs00231372_m1	A5917
TCL1A	Hs00951350_m1	A5917
TERT	Hs00972650_m1	A5917
TGFBR2	Hs00559660_m1	A5917
TGFBR3	Hs00234259_m1	A5917
TNF	Hs00174128_m1	A5917
TNFRSF10A	Hs00269492_m1	A5917
TNFRSF10D	Hs00174664_m1	A5917
TNFSF10	Hs00234355_m1	A5917
TRPS1	Hs00936363_m1	A5917
VAV3	Hs00196125_m1	A5917
ZMIZ1	Hs00277476_m1	A5917
Endogenous Control Genes		
Gene	Assay ID	Lot Number
RN18S1 (18S rRNA)	Hs99999901_s1	A5917
ACTB	Hs99999903_m1	A5917
GNB2L1	Hs00272002_m1	A5917
HMBS	Hs00609297_m1	A5917
RPLP0	Hs99999902_m1	A5917

2.17 Confocal microscopy

U2OS cells that were transfected with pGY015-ERT2-full3B-citrine expression construct were grown on cleaned cover slips in 6-well plates. The cells were left either treated or untreated with 4HT for 48h after transfection before being mounted onto slides for confocal analysis. Cell medium was removed from each well and the cells were washed 3 times with ice-cold PBS. 2ml of methanol (that was stored at -20°C) was added and the cells were incubated at -20°C for 10 min for fixation. The methanol was removed and left to evaporate at room temperature for 10 min. The cells were rehydrated by washing with PBS 3 times at room temperature. The cells were then stained at room temperature with 1:1000 dilution of DRAQ5 (Biostatus) in 2ml of PBS. They were then washed 3 times with PBS. 25µl of Mowiol® mounting medium (Sigma-Aldrich) was added onto a microscope slide. Coverslips were placed in water for 5 min to clear away any PBS crystals, and then removed and drip-dried on paper towel. The coverslips were carefully placed (cells side down) over the mounting fluid in order to ensure no air bubbles. This was left in the dark at room temperature for 1 day to dry, after which the slide was stored in the dark at 4°C. Visualisation and image capture was done by Mike Hollinshead, using a Zeiss LSM510 laser scanning confocal microscope.

2.18 Immunoprecipitation

For each immunoprecipitation (IP) reaction, 3×10^6 cells were harvested as described in section 2.5.5. The cells were lysed with 200µl IP lysis buffer at 4°C for 20 min on an orbital rotor. Cell debris was spun down by centrifugation at 14,000rpm for 20min at 4°C, and the supernatant was transferred to a fresh Eppendorf tube. The lysate was pre-cleared with protein G agarose beads (GE healthcare) for 1h at 4°C on an orbital rotor, after which the agarose beads were removed by centrifugation for 1 min at 1000 rpm. The lysate was transferred to fresh a fresh Eppendorf and was either kept as a sample for input, or used for IP reactions. Protein complexes within the lysate were precipitated with either 4µg of anti-FLAG antibody (mouse monoclonal anti-FLAG, M2, from Sigma-Aldrich), or 4µg of IgG

control (normal mouse IgG, sc-2025, Santa Cruz). These were left overnight on an orbital rotor at 4°C.

The next day, 30µl of protein G agarose beads were added and this mixture was incubated at 4°C for an hour with rotation. The beads were pelleted by centrifugation at 1000rpm for 1 min. The supernatant was removed and the beads were washed a further 4 times with IP lysis buffer. 30µl of SDS-PAGE sample buffer was added immediately to the agarose beads and 10µl of sample buffer was added to 20µl of input. The beads and input samples were boiled at 100°C for 5 min, and the beads were pelleted at 1000rpm for 1 min. All samples were loaded onto an SDS-PAGE gel for Western Blot, according to the protocols described in section 2.13.

2.19 ChIP and ChIP-seq analyses

2.19.1 Chromatin Immunoprecipitation – epitope-tagged EBNA3s

Cells were seeded at a density of 3×10^5 cells/ml one day prior to harvest. The next day, 20ml of cells were fixed in 1% paraformaldehyde (Sigma-Aldrich) for 10min at room temperature. Glycine was added to a final concentration of 125mM in order to quench the paraformaldehyde. Cells were immediately spun down at 1300rpm for 4min at 4°C. The supernatant was discarded and the cells were washed twice with ice-cold PBS. Next, the cells were transferred to a fresh 1.5ml Eppendorf tube and pelleted with centrifugation at 4000rpm for 1 min. The supernatant was aspirated and the cells were resuspended in 1ml ChIP swelling buffer. The cell suspension was left on an orbital rotor for 20 min at 4°C. The resultant nuclei were pelleted with a cold centrifuge for 5 min at 2000rpm. The supernatant was very gently removed, and the remaining nuclei were washed gently with ChIP sonication/IP buffer, and then resuspended in 1ml of the same buffer. The nuclei were incubated on ice for at least 30 min and then transferred to a 1ml milliTUBE (Covaris).

The cell lysate was sonicated for 1h using a M220 Focused-ultrasonicator with a milliTUBE holder with the following settings: 75W peak power, 26 duty factor, 200

cycles/burst, 6°C set point temperature. The sonicated lysate was centrifuged at 13,000rpm for 10 min at 4°C and the supernatant was diluted with 3ml of ChIP sonication/IP buffer. 5% of the input was kept at 4°C as a control, and the rest were incubated at 4°C with 16µg of anti-FLAG antibody (Cell Signalling) and 120µl of ChIP-grade protein G magnetic beads (NEB, #9006) in a 15ml Falcon tube, on rollers.

On the next morning, a magnetic separation rack was used to separate the beads from the lysate solution. The beads were washed in 4ml of wash buffers in the following order: Low Salt Immune Complex Wash Buffer, High Salt Immune Complex Wash Buffer, LiCl Immune Complex Wash Buffer, TE Buffer (2 washes). Each wash was performed for 15 min on rollers in a cold room at 4°C. The precipitated DNA and 5% input were eluted by adding 400µl of elution buffer and 100µg of RNase A (QIAGEN). The eluate was incubated at 37°C on a heat block for 1h. The beads were removed and the eluate was transferred into a fresh tube. 16µl of 5M NaCl was added to each tube of precipitated or input DNA. The mixture was incubated for at least 4h at 65°C on a heat block to reverse the cross-links between protein and DNA.

In order to facilitate protease activity, 8µl of 0.5M EDTA, 16µl of 1M Tris-HCl pH6.5 and 4µl of QIAGEN proteinase K were added to each tube and the samples were incubated on a heat block at 45°C for 1h. Finally, the DNA was purified using QIAGEN MinElute PCR cleanup kit, with some modifications. 3M of sodium acetate, pH5.2 was always added to each sample at the first step, and after the final spin-dry step, the columns were left for 5 min with the lids open for the ethanol to fully evaporate. The DNA was eluted into LoBind Eppendorf tubes with 20µl of autoclaved purified water.

2.19.2 Chromatin precipitation – other factors

ChIP for factors other than epitope-tagged EBNA3s was performed as above, although modifications were made to scale down the experiment. After dilution of the sonicated cell lysate to 4ml, the samples were separated equally into 4 tubes and 4µg of antibody or normal IgG were added to each tube for individual IP reactions. The reagents used for the subsequent steps were scaled down as appropriate, and the DNA was purified using QIAquick PCR purification kit (QIAGEN) in 80 µl of autoclaved purified water. The antibodies used for ChIP can be found in Table 2.13.

Table 2.13. Antibodies and IgG controls used for ChIP

Antibody/IgG	Species	Source	Catalogue number	Quantity
Anti-Flag (DYKDDDDK)	Rabbit (polyclonal)	Cell Signalling	2368S	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
H3K27me3	Rabbit (polyclonal)	Millipore	07-449	4 µg
SUZ12	Rabbit (polyclonal)	Abcam	ab12073	4 µg
Anti-RNA polymerase II CTD repeat YSPTSPS (phospho S5)	Rabbit (polyclonal)	Abcam	ab5131	4 µg
Normal Rabbit IgG	Rabbit	Abcam	ab171870	4 µg
Normal Mouse IgG	Mouse	Santa-Cruz	sc-2023	2 µg

2.19.3 ChIP-qPCR

RT-qPCR was performed as described in section 2.14.1 and was used to assess the enrichment of a DNA-bound protein at specific genomic regions. The 5% input DNA was diluted to 2.5% (of the original input) for the assay. 2µl of input or sample DNA was used for each PCR reaction, and the primer sequences used for each region examined can be found in Table 2.14. For every experiment, pooled input DNA were used to create standard curves of four 5-fold serial dilutions in order to ascertain the PCR efficiency and the relative amount of DNA in each sample. Relative ChIP signal was calculated based amount of signal relative

to 2.5% of the input DNA. Error bars represent one standard deviation calculated from technical triplicates.

Table 2.14. Primers used for ChIP-qPCR analysis

Name of primer pair	Forward primer (5'-3')	Reverse primer (5'-3')	Source
CTBP2 B	TTGCATATTTGGGATTTTCAGTTT	TTCTGTGTGAAACAGTTGTGGTT	(McClellan <i>et al.</i> , 2013)
ITGAL F	TGCTTACACTTCCTCCCTGAA	TTTCTCACAGAGGCAACAGG	
GAPDH	CTGAGCAGACCGGTGTCACATC	GAGGACTTTGGGAACGACTGAG	(Delbarre <i>et al.</i> , 2010)
MyoG	GGAGAAAGAAGGGGAATCACAT	GATAAATATAGCCAACGCCACA	
TERT +2kb	GTCGAGTGGACACGGTGAT	AAGTTTATGCAAACTGGACAGGA	(Hoffmeyer <i>et al.</i> , 2012)
TERT +1kb	CACACAGAAACCACGGTCAC	CAGTGCCAGCCGAAGTCT	This study
TERT -1kb	ACTTGGGCTCCTTGACACAG	CTCCATTTCCCACCCTTTCT	
TERT -2kb	AACACTTGAGTGGGGTGAGC	AGATGATCCACCTGCCTCTG	
TERTA	CAAAGAGAAGCCACCCAGAG	TCTGGTGGGTGTGTAGTGGA	
TERTB	TGGAAGAGCAGGTGCTATGA	GAGCTCCGTGGATAGCTTTG	

2.19.4 High-throughput ChIP sequencing

DNA pooled from 8 ChIP extractions was size-selected on an agarose gel (see section 2.2.3.2). At least 1ng of the DNA for each pooled Input and ChIP sample was sent to the Harvard Biopolymers Facility for library construction using the ChIP-seq Wafergen service. This involves end-repair of the fragmented DNA, phosphorylation of the 5' ends and addition of A-tails at the 3' ends to facilitate adapter ligation. Adapters compatible with the Illumina sequencing platform were ligated to the sequences, after which the samples were size-selected and gel purified. PCR was performed to enrich for sequences with adapters ligated to them. Next-Generation sequencing was performed with the Illumina HiSeq® 2500 Sequencing System with resultant single-end read lengths of 50bp. The sequences were mapped to *Homo sapiens* hg19 genome and analysed using Partek Genomics Suite (PGS) version 6.6 software.

2.19.4.1 ChIP-seq data analysis

Peak calling was done using Partek Genomics Suite (PGS)¹. Generally, enriched regions display two adjacent peaks of single-end reads. PGS extended each single-end read in the 3' direction by an estimated fragment length (145bp, determined based on strand cross-correlation profiles of the ChIP sample) in order to merge adjacent peaks into a single peak. Peak detection was done by dividing the genome into windows of 145bp and counting the number of midpoints of the reads within each window. PGS used a zero-truncated negative binomial model (Ji *et al.*, 2008; Spyrou *et al.*, 2009) to the read counts in each window and detected regions that were above a false discovery rate (FDR) of 0.001. A control sample that was processed in a similar manner as the EBNA3B ChIP sample was obtained from untagged wild-type LCL of the same donor background. Using a binomial test (p-value <0.05), PGS compared the peaks obtained from EBNA3B ChIP sample to those obtained from the control sample and filtered out peaks which were detected in both. The remaining filtered peaks were defined as EBNA3B peaks that would be used throughout this study.

Read pile-up from the ChIP-seq study was visualised using Integrative Genome Browser v2.3.59 (Broad Institute). In order to calculate the non-redundant fraction (NRF, see section 4.5.1 for details), Picard.MarkDuplicates v2 module was accessed via Gene Pattern 3.9.4 software (available online <https://genepattern.broadinstitute.org/gp/pages/index.jsf>) and used to generate a table detailing the total number of aligned single-end reads, the number of reads that were uniquely mapped, the number of unmapped reads and the number of reads mapped to duplicate locations on the genome. NRF was calculated using the ratio between the number of uniquely aligned reads over the total number of reads (Landt *et al.*, 2012). Peak co-localisations (defined as two peaks with at least one bp overlap) were determined using PGS. Publicly available peak and region tracks used in this study are described in Table 2.15. A full list of EBNA3B-regulated genes (p<0.05) was obtained from

¹ The exact algorithms are available online at http://www.partek.com/Tutorials/microarray/User_Guides/ChIP-Seq_White_Paper.pdf.

Dr Rob White (White *et al.*, 2010) and their gene annotations were acquired using UCSC Table Browser tool (available online at: <http://genome.ucsc.edu/cgi-bin/hgTables>). Genes were only described as being within contact domains if the transcription start site was found within a domain (see section 4.5.9).

Table 2.15. ChIP-seq and Hi-C data sets used in this study

Data set	Cell line	GEO accession number/Access link	Source
Contact Domains	GM12878 (LCL)	GSE63525	(Rao <i>et al.</i> , 2014)
EBNA2	Mutu III (BL)	GSE47630	(McClellan <i>et al.</i> , 2013)
EBNA3	Mutu III (BL)	GSE47629	
EBNA3A-TAP	LCL (Donor 11)	GSE88729	Dr Kostas Paschos
EBNA3C-TAP	LCL (Donor 11)	GSE88729	
FS-EBNA3B	LCL (Donor 11)	GSE88729	This study
RBPJ	Mutui (LCL)	GSE75503	(Lu <i>et al.</i> , 2016)
Genome segmentations	GM12878 (LCL)	ChIP-seq/Segmentation data available for download at: http://genome.ucsc.edu/ENCODE/downloads.html	ENCODE
Histone modification	GM12878 (LCL)		
Transcription factor binding sites	GM12878 (LCL)		

2.20 TRAP assay

TRAPeze® RT (S7710) kit from Merck Millipore was used to examine the enzymatic activity of telomerase in LCLs, with some modifications. 1×10^6 Cells were pelleted as described in section 2.5.5 and resuspended in 200µl CHAPS Lysis Buffer (supplied in kit). The suspension was incubated on ice for 30 min and spun in a centrifuge at 13,000rpm for 20 min at 4°C. 160µl of the supernatant was transferred into a fresh tube and the protein concentration determined as described in section 2.13. The samples were adjusted to a protein concentration of 750ng/µl using CHAPS lysis buffer. The lysates were aliquoted, snap frozen on dry ice and stored at -80°C. A standard curve was generated using four 1:10 serial dilutions (20 amol², 2 amol, 0.2 amol, 0.02 amol) of TSR8 control template (details in Chapter 5, section 5.3.4). Heat-treated telomerase negative controls were obtained by

² 1 amol = 10^{-18} moles

incubating 10µl of each sample at 85°C on a heat block for 10 min in order to inactivate the telomerase. 1µl each of the samples, telomerase positive cell extract (HeLa cells, provided in the kit), standard curve samples, CHAPS lysis buffer (telomerase-free control) or nuclease-free water (no template control) were used for each reaction. The reaction mixture was set up with the following constituents in a MicroAmp® Optical 384-Well reaction plate:

Sample	1 µl
5X TRAPEZE® RT Reaction Mix	2.5 µl
Titanium <i>Taq</i> Polymerase (Clontech)	0.2 µl (1 unit)
Nuclease Free water	6.3 µl
Total	10µl

Real-time PCR amplification was performed using an ABI 7900HT thermocycler with the following conditions: an incubation step for telomerase extension at 30°C for 30 min followed by 95°C for 2 min to activate the polymerase, and then 45 cycles of amplification starting with denaturation at 94°C for 15 sec, followed by annealing of primers at 59°C for 60 sec and then extension at 45°C for 10 sec. The data was analysed by performing a logarithmic plot of the amount of TSR8 (in amoles) against the Ct measurements of the standard curves in Microsoft Excel®. Arbitrary telomerase units relative to the TSR8 amplification was determined for each sample using the linear equation derived from the TSR8 standard curve.

2.21 Telomere length assay

Telomere lengths of chromosomes in cells were determined using an assay developed by Cawthon and colleagues (2009). Where possible, cells were seeded at 3×10^5 one day prior to harvesting the cells and approximately one million cells were used for this assay. DNA was extracted using DNeasy Blood & Tissue Kit (QIAGEN), according to manufacturer's protocol. The DNA was quantified using Nanodrop Lite spectrophotometer (Thermo Scientific) and the concentration adjusted to 20ng/µl. Primer sequences (Table 2.16) were obtained from Cawthon *et al.* (2009). This assay uses a SYBR green-based qPCR to quantify both telomeric DNA and a single copy gene as a normalisation control. The

primers are designed such that signals generated from amplifying the more abundant target sequence (telomere) is detected only at a lower melting temperature, where would cross the threshold at much earlier cycle numbers than the single copy gene control (scg, such as human beta-globin): the levels of the scg would be negligible relative to the telomere. The Ct values of the scg amplicon were collected at a higher temperature than the melting temperature of the telomere amplicon: hence the telomere amplicon was denatured and its signal returned to baseline, thereby allowing the level of the scg to be detected.

Table 2.16. Primers used for quantification of telomere length

Primer name		Primer sequence
Telomere primers	Telg:	ACACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT
	Telc:	TGTTAGGTATCCCTATCCCTATCCCTATCCCTATCCCTAACA
Single copy gene primers (Human beta-globin)	Hbgu:	CGGCGGCGGGCGGCGGCGGGCTGGGCGGCTTCATCCACGTTACCTTG
	Hbgd:	GCCCGGCCCGCCGCGCCCGTCCCGCCGAGGAGAAGTCTGCCGTT

The telomere primers (final concentration 900nM each) and the human beta-globin primers (final concentration 500nM each) were added to 80ng of DNA and 5µl of 2X KAPA SYBR® FastqPCR Master Mix (low rox) in a final volume of 10µl. Standard curves were made from pooled samples with four 1:5 serial dilutions starting at 400ng of DNA. RT-qPCR was performed using ABI QuantStudio™ 7 Flex Real-Time PCR System with the cycling conditions below:

Stage	Number of cycles	Temperature	Time
Stage 1	1 cycle	95°C	15 min
Stage 2	2 cycle	94°C	15 sec
		49°C	15 sec
Stage 3	32 cycles	94°C	15 sec
		62°C	10 sec
		74°C	15 sec (signal acquisition)
		84°C	10 sec
		88°C	15 sec (signal acquisition)

The relative amount of each amplicon was determined using standard curves for telomere and scg signals, using the same DNA sample (DNA from peripheral blood mononuclear cells) for normalizing all assays. The relative amount of standard DNA that corresponds to the same Ct value as the sample for the telomere template was denoted

T and the same for the scg template was denoted S. The T/S ratio is proportional to the average telomere length in each cell.

3 Strategies for validating EBNA3B regulated genes

3.1 Introduction

Initially, in order to study the effects of EBNA3B on host gene expression, I revisited the exon microarray published by White *et al.* (2010) (available at <http://www.epstein-barrvirus.org>) and chose 2 genes of interest (*CXCL10* and *CXCL9*) to study further. These genes are situated adjacent to each other at the same genomic locus and were picked based on our initial interest in the ability of *CXCL10* to attract T-cells (Groom and Luster, 2011; White *et al.*, 2012). I conducted preliminary analyses which included verifying these genes' differential expression in long-established wild-type (WT), EBNA3B knockout (3BKO) and EBNA3B-revertant (3BRev) cell lines, assessing the histone modifications at their transcriptional start sites and observing the kinetics of any regulation through primary B-cell infection and outgrowth experiments. Together, the experiments showed that the regulation of these genes may either be through an indirect action of EBNA3B or may be a result of clonal selection during outgrowth (refer to section 3.2). This led us to question whether all the genes identified in the microarray study are direct targets of EBNA3B or if the differences that were observed between two different cell lines (WT and 3BKO) might have arisen by selection pressure during outgrowth of LCLs after infection of primary B-cells. A system in which the presence of EBNA3B can be controlled in a single genetic background would be a better way to validate any EBNA3B-regulated genes identified by the microarray.

This chapter addresses the different methods with which we have tried to establish such a system: first by attempting to construct a conditional EBNA3B system using recombinant EBV virus and second by complementing an established 3BKO cell line with inducible EBNA3B expression by using stable transfection of a doxycycline-responsive RTS-vector encoding EBNA3B. By using a Taqman Low-Density Array (TLDA) card containing a panel of previously identified potential EBNA3B target genes (White *et al.*, 2010), we aimed to use these complementation systems to identify *bona fide* target genes of physiological interest which we could employ to study the function of EBNA3B. The end goal is to revise

the list of EBNA3B target genes to those which can be reliably shown to be regulated by EBNA3B within the same cellular background, thus removing any clonal selection effects resulting from LCL outgrowth which were not taken into account in our initial exon-microarray analyses.

3.2 Differential expression of CXCL10 & CXCL9 between WT and 3BKO

LCLs may be a result of clonal selection

Initial microarray studies performed by White *et al.* (2010) on WT and 3BKO LCLs derived from 3 different donor backgrounds (Donors 1,2 and 3) revealed a significant reduction ($p < 0.001$) of CXCL10 transcripts in 3BKO LCLs as compared to WT LCLs (Figure 3.1A). Overall, there were no significant changes ($p = 0.0674$) in the number of CXCL9 transcripts between 3BKO and WT LCLs, however, there appeared to be a 4-fold decrease in CXCL9 transcripts in 3BKO LCLs as compared to WT LCLs derived from Donor 2 (Figure 3.1B). RT-qPCR analyses on two additional sets of WT, 3BKO and 3BRev LCLs derived from Donors 11 and 13 (D11 and D13) showed that the regulation of CXCL10 in Donor 11 was not consistent with the findings from the microarray, in that although the mRNA expression of CXCL10 in 3BKO LCLs was repressed as compared to 3BRev LCLs, the level of CXCL10 transcripts was paradoxically lower in WT as compared to 3BKO LCLs (Figure 3.2A). In addition, the expression of CXCL9 is induced in D11 3BKO LCLs, but repressed in D13 3BKO LCLs, when compared to WT and 3BRev LCLs of similar donor backgrounds (Figure 3.2B).

Subsequent ChIP experiments found no significant differences in the levels of histone marks of activation (H3K4me3, Acetylated H3) and repression (H3K27me3), and in the levels of RNA polymerase (RNA PolII) and serine-5 phosphorylated RNA PolII (indicator of transcription initiation) between WT, 3BKO and 3BRev LCLs (Figure 3.3). Hence, EBNA3B is unlikely to have an effect on these epigenetic modifications, or on the recruitment of RNA PolII at the CXCL10 transcription start site. In addition, infection of

purified CD19+ B-cells pooled from two different donors (Donor 27 and 28) with recombinant EBV showed no significant changes in the expression of either CXCL10 or CXCL9 between WT, 3BKO and 3BRev LCLs during the first 42 days of outgrowth (Figure 3.4).

Put together, these early experiments show that there are no clear trends in the regulation of CXCL10 or CXCL9 by EBNA3B, despite the significant findings of the microarray studies done by White *et al.* (2010). CXCL10 is only clearly repressed when comparing established 3BKO and WT LCLs of D1, D2, D3 (used in the microarray), and D13 backgrounds. The regulation of CXCL10 in D11 3BKO is unclear because the levels of CXCL10 transcripts are higher in D11 3BKO as compared to the D11 WT LCLs, but lower as compared to the D11 3BRev LCLs. There is also no significant difference in the expression of CXCL10 during the first 42 days of outgrowth between EBNA3B-positive and EBNA3B-null cell lines. EBNA3B is also not directly effecting any epigenetic changes or PolII recruitment at the CXCL10 TSS. Hence, it is possible that the regulation of CXCL10 is donor dependent and therefore only evident in certain donors; however, it is also likely that the differences in expression detected in established WT and 3BKO LCLs is a result of clonal selection during the B-cell outgrowth process, or CXCL10 is an indirect target of EBNA3B. Similarly, any perceived EBNA3B regulation of CXCL9 is unreliable and changes with different donor backgrounds – this is already evident in the microarray studies (Figure 3.1). Hence, using CXCL10 as an example, the preliminary results demonstrate the importance of verifying putative EBNA3B-targets identified in the microarray study, and that some of these EBNA3B targets may have been erroneously identified due to differences in gene expression resulting from clonal selection during the outgrowth of LCLs. In order to verify the targets identified in the microarray, another system using inducible EBNA3B within a single cellular background would be a more reliable method of assessing EBNA3B regulation of host genes.

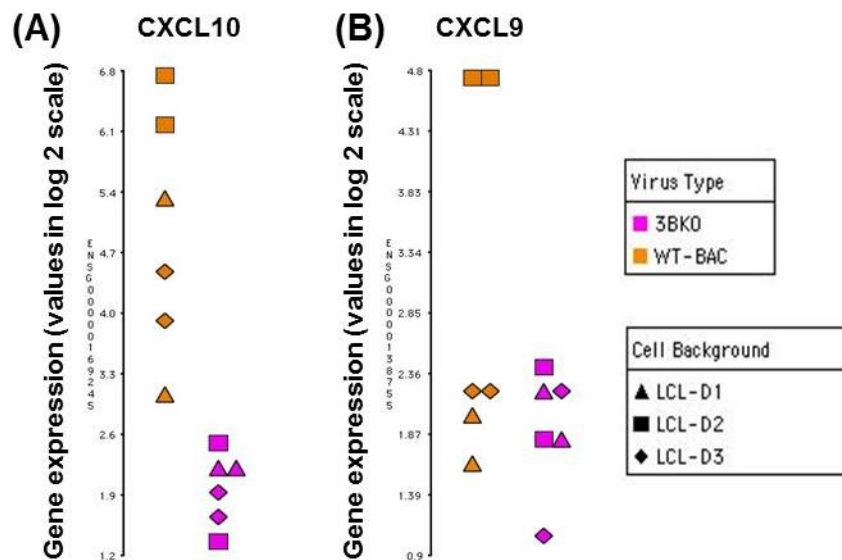


Figure 3.1. Differential expression of (A) CXCL10 and (B) CXCL9 between wild-type and 3BKO LCLs which are established from 3 independent donors (D1, D2 & D3). There is a significant 7-fold reduction ($p < 0.001$) of CXCL10 transcripts in 3BKO LCLs as compared to WT LCLs. No significant overall change ($p > 0.05$) is detected between the number of CXCL9 transcripts in 3BKO and WT LCLs. Data are obtained from the microarray analyses described by White *et al.* (2010) and are available on www.epstein-barrvirus.org.

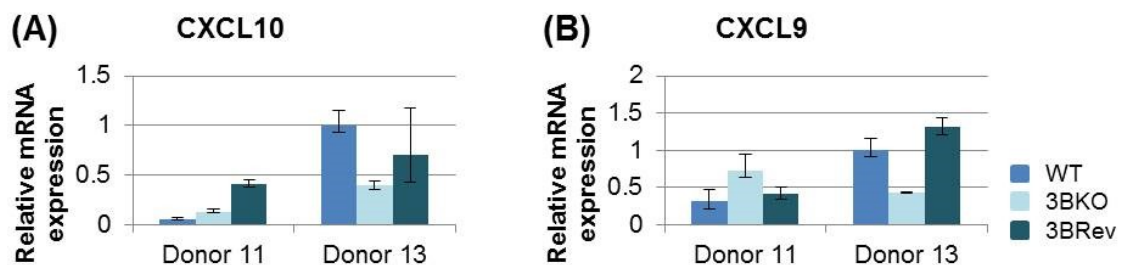


Figure 3.2. Relative gene expression of (A) CXCL10 & (B) CXCL9 in LCLs derived from Donor 11 and Donor 13. Data are normalised to endogenous control genes *GNB2L1*, *RPLP0* and *ALAS1*. Bars represent means in fold change compared to wild-type LCLs of Donor 13. Error bars represent technical triplicates from one experiment.

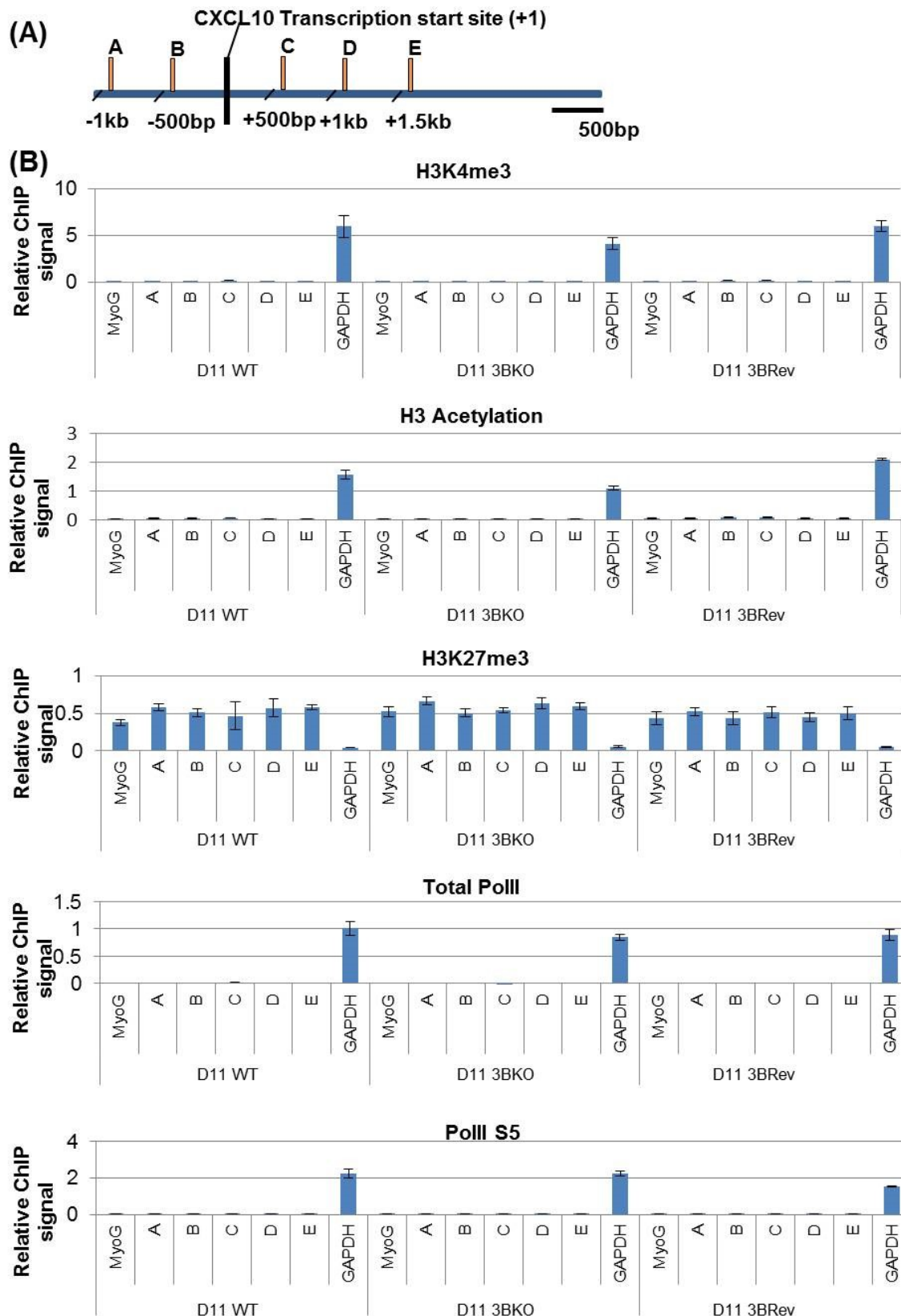


Figure 3.3. ChIP for histone modification changes at CXCL10 TSS in WT, 3BKO and 3BRev LCLs.

(A) Schematic showing where primers are amplifying at the CXCL10 locus in the ChIP-qPCR experiments. **(B)** ChIP using H3K4me3, global H3 Acetylation, H3K27me3, Total RNA Polymerase II (Total PolII) and Phosphorylated serine 5 residue of RNA PolII (PolII S5) at the CXCL10 TSS. Data are representative of 3 experiments using cell lines established from different donor backgrounds. Error bars represent technical replicates in one qPCR experiment.

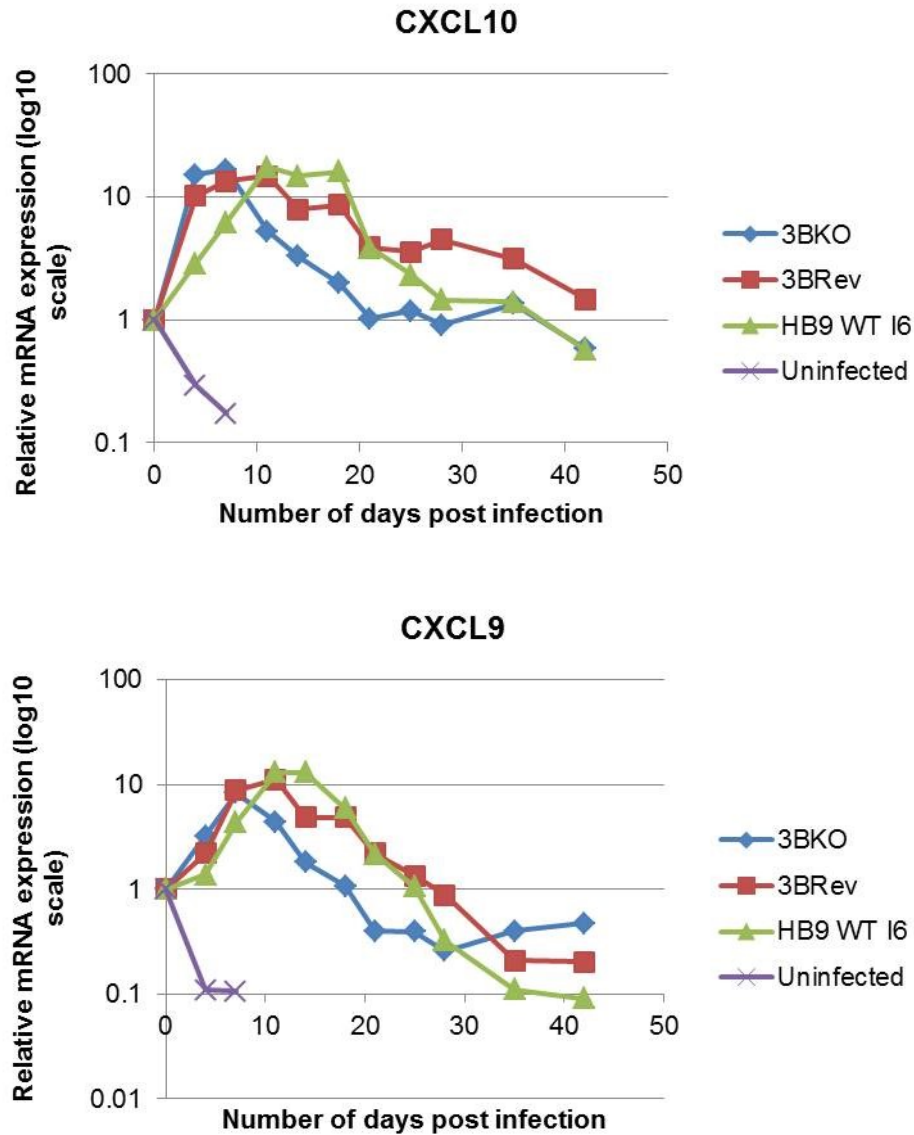


Figure 3.4. Timecourse following the expression of CXCL10 and CXCL9 over the first 40 days of LCL outgrowth.

CD19⁺-sorted primary B-cells pooled from 2 different donors (Donor 27 and 28) were infected with WT, 3BKO and 3BRev EBV. RNA was extracted from these cells over a period of 42 days (done by Dr Lydia Eccersley). RT-qPCR was performed with the RNA extracts. Data are represented relative to the uninfected sample at day 0.

3.3 Construction of a recombinant EBV BAC encoding 4HT-conditional EBNA3B (ER3B)

The cloning of the B95.8 EBV genome into a F factor-derived plasmid (Delecluse *et al.*, 1998; Feederle *et al.*, 2010) has allowed our lab and others to rapidly and precisely introduce mutations or fusion tags into the viral genome. Using this bacterial artificial chromosome (BAC) system, the viral genome can be propagated and manipulated in bacteria, after which the resultant recombinant EBV BAC may be purified in large quantities, ready for transfection into cells for virus production. In my first attempt at constructing a conditional EBNA3B system, I chose to fuse the 4-hydroxytamoxifen (4HT)-sensitive human oestrogen receptor (ERT2) (Feil *et al.*, 1997) to EBNA3B. This modified human oestrogen receptor ligand binding domain is responsive to 4HT but insensitive to oestrogen. Briefly, in the absence of 4HT, the ERT2 domain should target EBNA3B for degradation and removal from the cell nucleus. In the presence of 4HT, the fusion protein is stabilised and stays in the nucleus, where it can exert its gene-regulatory functions. C-terminal fusions of the 4-hydroxytamoxifen (4HT)-sensitive human oestrogen receptor (ERT2) to the other EBNA3s in our lab and others (Bazot *et al.*, 2015; Maruo *et al.*, 2003, 2006; Skalska *et al.*, 2010) have been shown to make them conditionally active in the presence of 4HT. We believe that a similar approach to EBNA3B could generate a cell line that expresses functionally conditional EBNA3B on 4HT.

Since the EBNA3s are alternatively spliced from long primary mRNA transcripts initiated at the Cp latency promoter, careful manipulation of the sequences in the genome is required to prevent disruption of splicing of the other EBNA3s. In particular, previous attempts by Dr Rob White to establish a C-terminally tagged EBNA3B in EBV compromised the splicing of EBNA3C transcripts (Appendix A), which was independent of their responsiveness to 4HT. As we were unsure of how the absence of EBNA3C would affect the function of EBNA3B in regulating its target genes, I attempted to modulate EBNA3B by fusing the ERT2 tag to the N-terminus of EBNA3B. This was done using the RecA-mediated

homologous recombination system adapted by White *et al.* (2003) from an original system developed by Lalioti & Heath (2001). The cloning strategy is discussed in detail in Chapter 2 (Materials and Methods). Briefly, the ERT2 sequence (Feil *et al.*, 1997) was cloned into a DNA sequence running from within the end of EBNA3A through the start of EBNA3B and on to the end of EBNA3B exon 1, thus creating the regions homologous with the EBV B95.8 BAC either side of the ERT2 insert (see Figure 2.3). The ERT2 tag was inserted immediately after the ATG start codon and the end of the domain is separated from EBNA3B by the peptide linker ASNGGSGEAS such that the entire ERT2-EBNA3B sequence is in frame and the start codon at the beginning of EBNA3B is omitted (Figure 3.5). RecA-mediated homologous recombination between the pKovKan Δ Cm shuttle plasmid (White, 2003) containing this insert and the B95.8 BAC in DH10B *E.coli* - either at the end of EBNA3A or the start of EBNA3B - results in co-integration at one of the two sites. The co-integrants were resolved to generate the required recombinant ER3B BAC.

Diagnostic digests of the EBV BACs isolated from both the co-integrant and resolution steps are illustrated in Figure 3.6. XbaI and BstBI (not shown) cut restriction sites within the ERT2 insert in the BACs which result in visible changes in band sizes on the pulsed field gel for the correct recombinant BAC (Figure 3.6). Other enzymes (EcoRI, NotI and AgeI) which cut throughout the EBV genome are used to identify any gross changes in the BAC integrity by comparison with the parental B95.8 BAC. Interestingly I was only able to obtain the co-integrant BAC that arose from homologous recombination at the start of *EBNA3B*. This suggests that the homologous recombination may be kinetically favoured for this co-integrant as compared to recombination at the end of *EBNA3A* gene, although no further investigations were done to ascertain if that is the case.

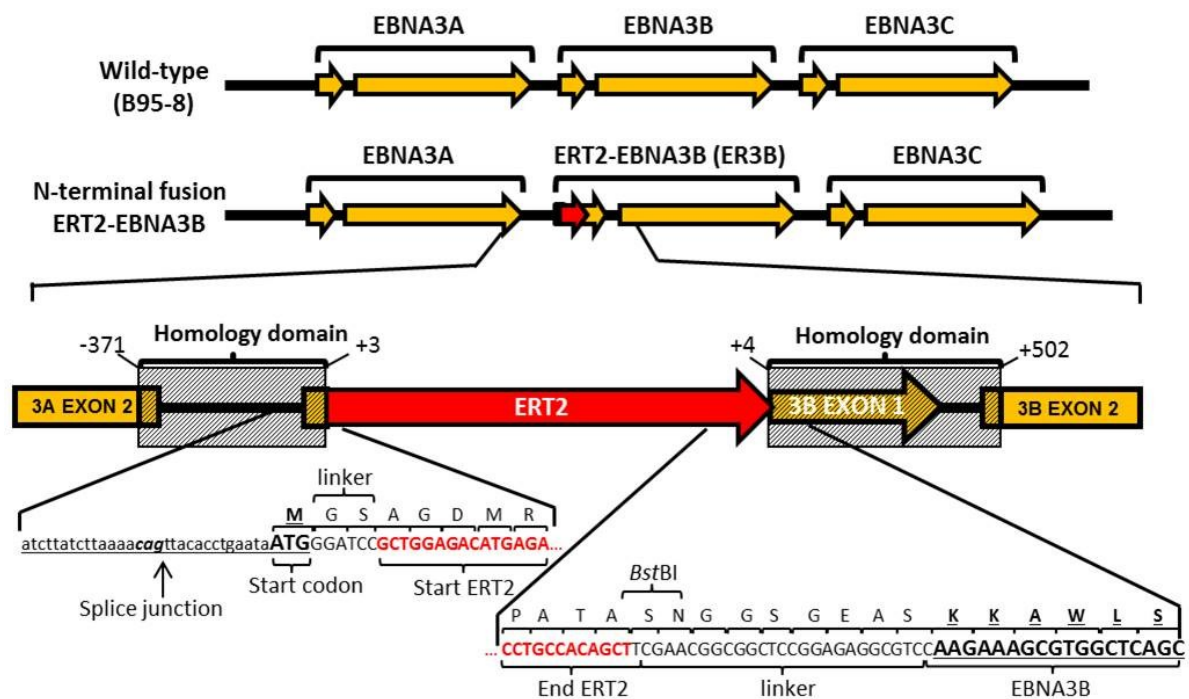


Figure 3.5. Schematic showing the insertion of ERT2 tag at the N-terminus of EBNA3B through homologous recombination.

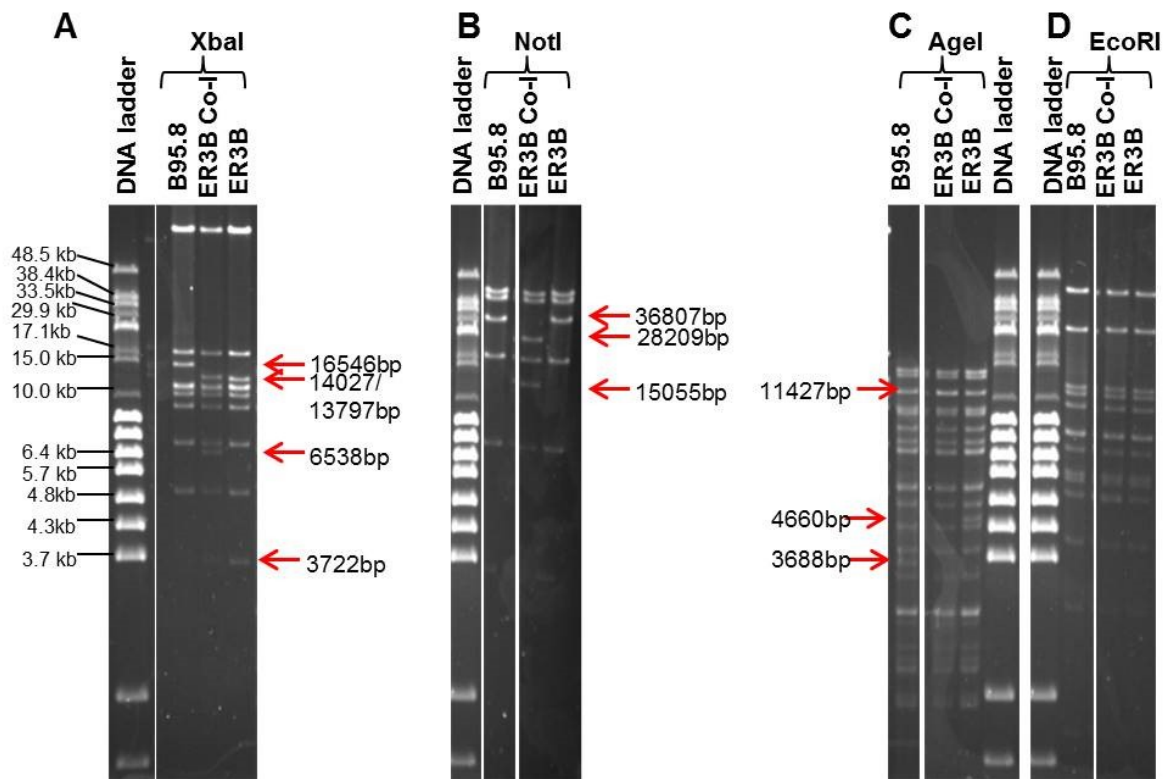


Figure 3.6. Generation of BAC DNA from wild-type (B95.8) to ER3B fusion protein.

Diagnostic digests followed by pulsed field gel electrophoresis (PFGE) are used to identify the BACs (*XbaI*) and validate the overall integrity of the BAC (*NotI*, *AgeI* and *EcoRI*). DNA ladder used is a mixture of λ DNA-BstEII digest and λ DNA-Monocut mix from New England Biolabs. Red arrows indicate bands that change during recombineering. **(A)** Insertion of the KovKan vector is expected to introduce 2 additional *XbaI* sites (one within KovKan backbone and one within ERT2 sequence), splitting the 16546bp band in B95.8 BAC into 14027bp, 6538bp and 3722bp bands in the co-integrand (ER3B Co-I). Similarly, the *XbaI* site within the ERT2 domain of the resolved ER3B BAC would produce 13797bp and 3722bp bands. **(B)** Likewise, presence of the KovKan backbone in ER3B Co-I adds 2 *NotI* sites (not within ERT2 region), and is predicted to split the 36807bp band into 28209bp, 15055bp and 1282bp bands. The resolved ER3B BAC reverts to 37779bp due to the additional ERT2 sequence that is indistinguishable from the 36807bp band in B95.8 by PFGE. **(C)** Addition of ERT2 and/or KovKan should not introduce any additional *AgeI* sites, but merely increase the size of the 3688bp band to 11427bp (Co-I) or 4660bp (ER3B BAC). **(D)** *EcoRI* digestion of the BACs should not cause any gross rearrangements of the band size distributions.

3.4 ER3B virus producer cell lines are generated by transfection into HEK-293 cells

Following the successful generation of ER3B BAC, I performed a maxiprep using CsCl₂ gradient separation and ethidium bromide to isolate the supercoiled BAC DNA. The DNA integrity was confirmed by diagnostic digest and PFGE (Figure 3.7). Using lipid-based transfection (Chapter 2, section 2.6.1), I transfected the recombinant ER3B BAC into HEK-293 cells to generate virus producer cell lines. Transfected HEK-293 cells were plated and grown in hygromycin to select for successfully transfected cells that retained the BAC, which would proliferate into discrete colonies over a period of 2-3 weeks. Several colonies were picked, expanded and the virus genomes were analysed by diagnostic digests and PFGE following episome rescue of the EBV-BAC. The episomes from these colonies were mostly intact, with no detectable mutations or major changes in EBV-BAC structure (Figure 3.8A).

As a quick diagnostic test to assess the functionality of the ER3B BAC, a whole flask of transfected 293 cells containing pooled colonies were analysed in bulk for the expression of ER3B with and without the addition of 4HT. Even after induction with 4HT after 21 days, ER3B was not detectable in the bulk population of 293 cells, whereas the wild-type B95.8 BAC expressed EBNA3B in the 293 cells. In contrast, the expression of other EBNA3 proteins was unaffected by the ERT2 fusion, showing that the N-terminal fusion at EBNA3B does not significantly affect the expression of the other EBNA3s in the 293 background (Figure 3.8B).

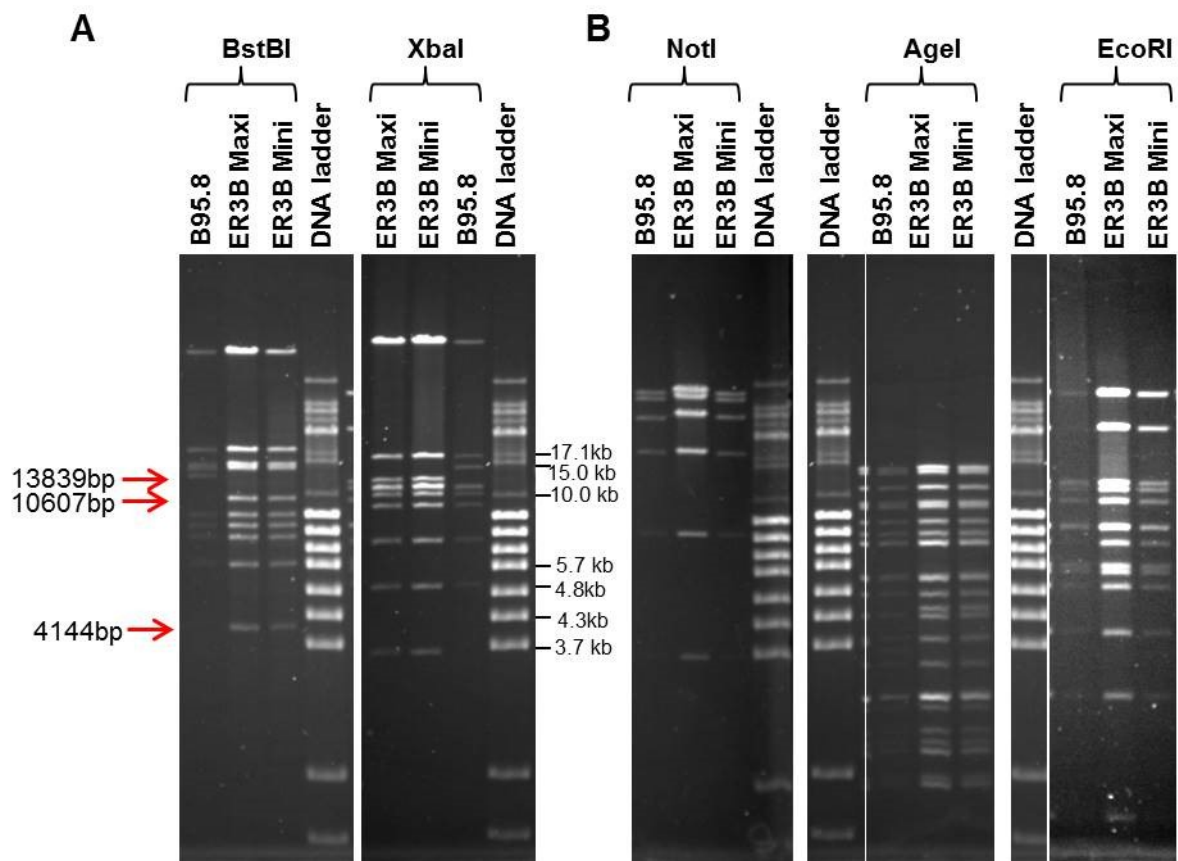


Figure 3.7. Verification of maxipreps of ER3B BAC before transfection into HEK-293 cells. Maxipreps of the ER3B BACs digested with **(A)** *Bst*BI and *Xba*I for identifying the ERT2 fusion in the BAC and **(B)** *Not*I, *Age*I and *Eco*RI for ensuring BAC integrity is intact. Wild-type B95.8 BAC and ER3B minipreps from the original clone are used as comparisons. N-terminal fusion of ERT2 to EBNA3B has introduced an additional *Bst*BI site, which should cause the 13839bp band from the B95.8 BAC to appear as two bands of 10607bp and 4144bp in ER3B BACs. Red arrows indicate bands that have changed during recombineering.

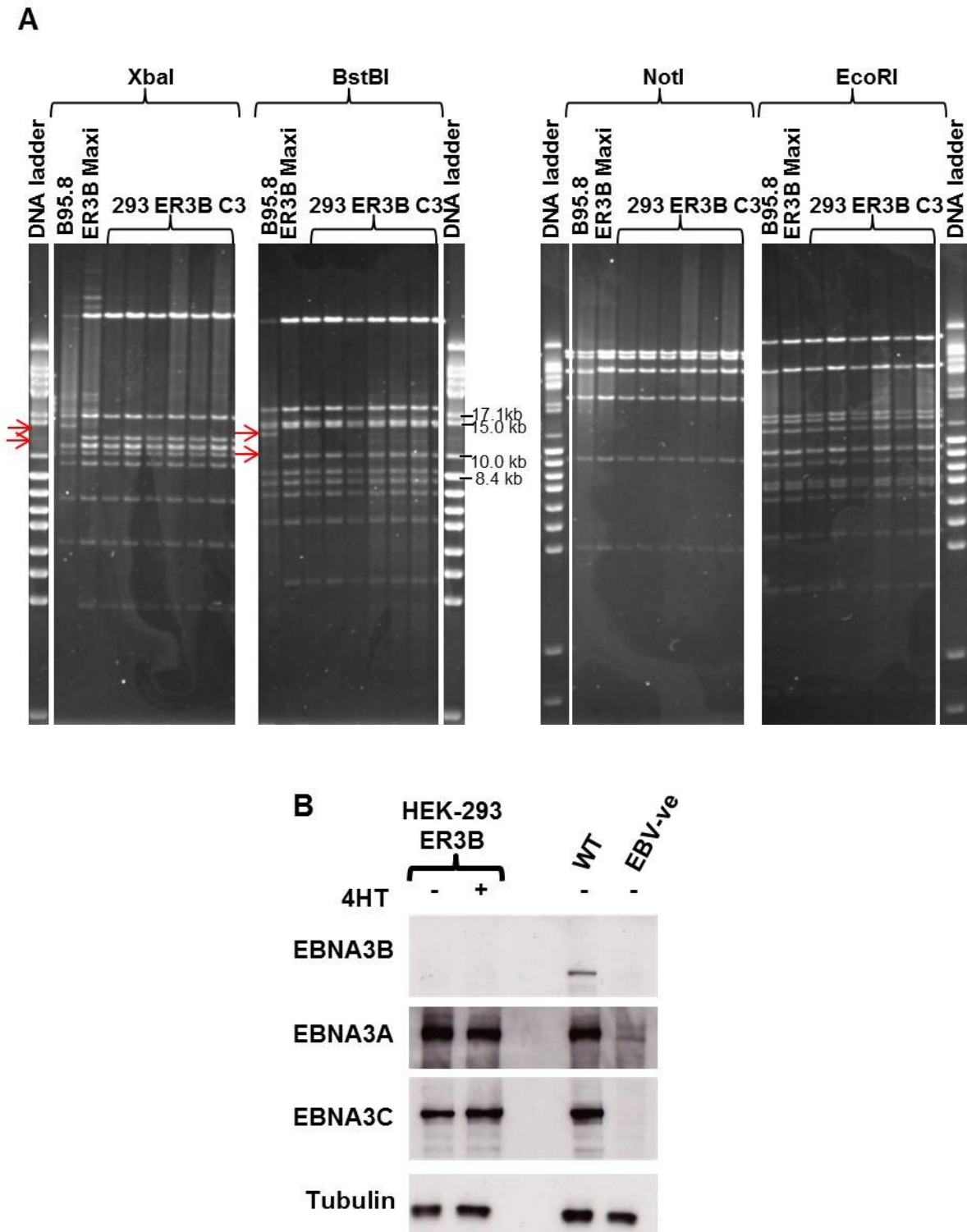


Figure 3.8. Verification of HEK-293 ER3B cell lines.

(A) Episomal DNA was recovered from the HEK-293 ER3B C3 cell line and BACs rescued into bacteria were screened for their integrity by digestion with *XbaI*, *BstBI*, *NotI* and *EcoRI* restriction enzymes, followed by PFGE. The B95.8 BAC and the ER3B BAC maxiprep were used as comparisons. Red arrows indicate bands that have changed during recombineering.

(B) Western blot showing three-week treatment of the transfected HEK-293 cells with 4HT. EBNA3B, EBNA3A and EBNA3C-probed blots are indicated. Tubulin is used as a loading control. Protein lysates from HEK-293 transfected with B95-8 BAC (WT) and untransfected HEK-293 cells (EBV -ve) were used as positive and negative controls for EBV protein levels.

3.4.1 mRNA transcripts for ER3B are present in HEK-293 ER3B clones.

We speculated that the failure to express the ER3B fusion protein could have been caused by a problem in splicing from the U-exon to the splice acceptor of *EBNA3B* which arose from the insertion of the ERT2 sequence. Hence, using primers that amplify across the splice junction between the U-exon and exon 1 of *EBNA3B* to 150 base-pairs downstream of the *EBNA3B* START codon, I PCR-amplified cDNA obtained from two of the HEK-293 ER3B clones (Figure 3.9). This showed that mRNA transcripts for recombinant fusion protein ER3B were present and that the splicing from U-exon to EBNA3B does not appear to be prevented by the insertion of the ERT2 tag.

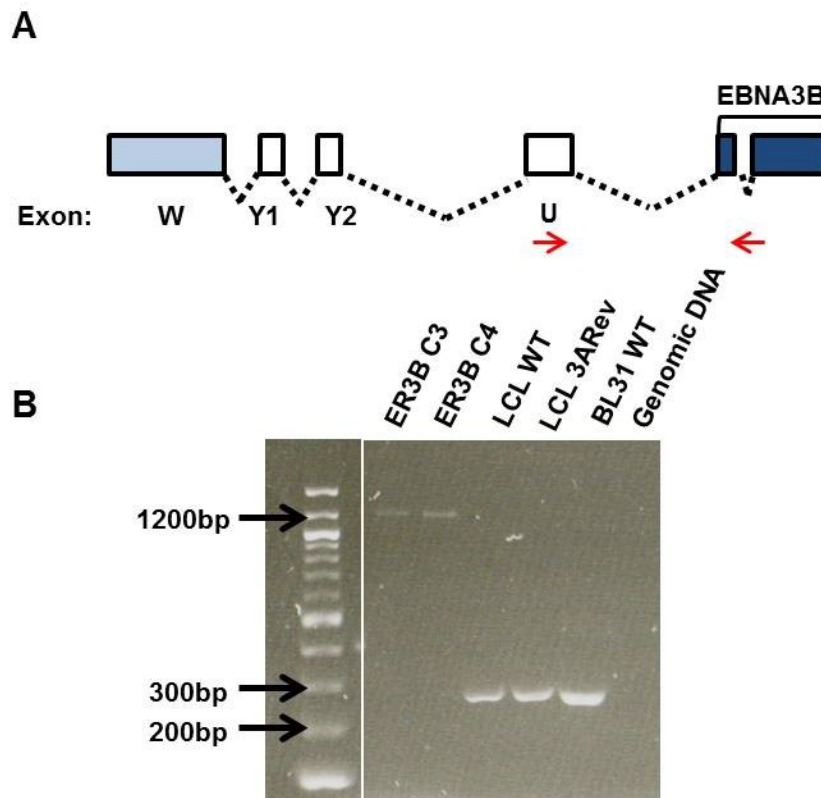


Figure 3.9. Accurately spliced transcripts for ER3B were detected by agarose gel electrophoresis.

(A) Schematic showing the splice regions of EBNA3B mRNA transcript. Red arrows indicate region of amplification for PCR. (B) Agarose gel electrophoresis of PCR-amplified cDNA across U-exon and exon 1 of EBNA3B. Predicted fragment sizes for ER3B as well as for wild-type controls (including EBNA3A-revertant, 3ARev) are 1226bp and 254bp respectively. No unspecific amplification of genomic DNA is expected.

3.4.2 ERT2-EBNA3B-citrine construct can be expressed in isolation.

In order to validate the function of the ERT2-EBNA3B constructs, I next tested whether the ERT2-EBNA3B fusion protein alone can be expressed and regulated by 4HT outside of the EBV BAC context. To do this, I fused a fluorescent protein (citrine) to the C-terminus of the ERT2-EBNA3B in a mammalian construct (pCDNA3), generating an ER3B-citrine mammalian expression vector (pGY15). Under this regulation, the ER3B-citrine protein should be constitutively transcribed and translated and fluoresce under confocal microscopy. However, in the absence of 4HT, the fusion protein would be targeted for degradation and should be detectable only at low levels under the microscope, and/or sequestered into the cytoplasm.

The plasmid pGY15 was transfected into the U2OS cells and viewed after 48 hours with the cells grown either with or without 4HT. The ER3B-citrine fusion protein was stabilised by the addition of 4HT to the culture medium after transfection and was localised mainly to the nucleus, but was undetectable in the absence of 4HT (Figure 3.10A). As a control, the parental pcDNA3 vector expressing EBNA3B-citrine without an N-terminal ERT2 tag was also transiently transfected into U2OS cells (Figure 3.10B) and showed that the parental EBNA3B-citrine fusion protein is unaffected by 4HT treatment. Together, these results showed that the ERT2-EBNA3B sequence can be expressed in isolation, and that the ERT2 domain modulates ER3B protein levels in response to 4HT.

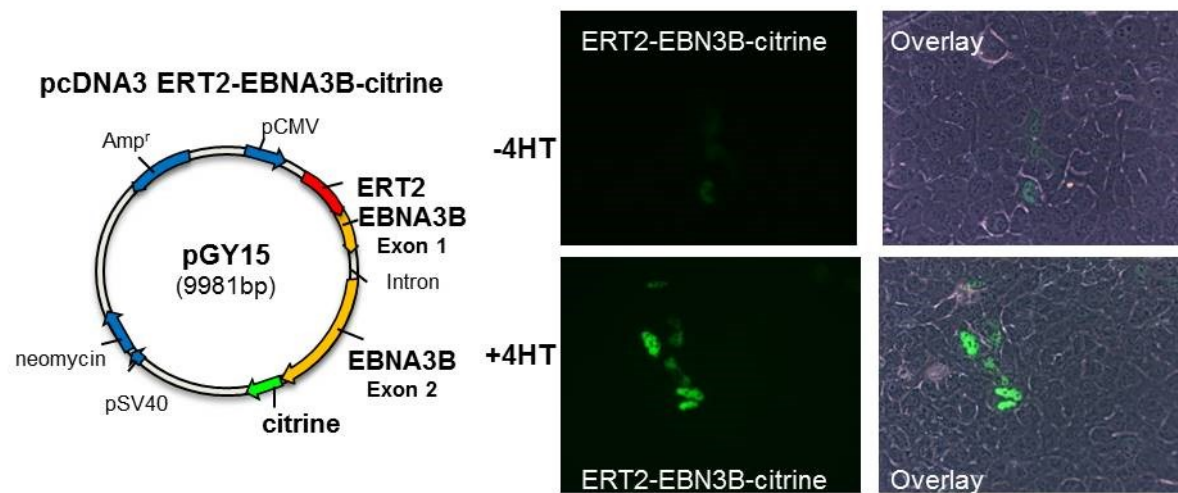
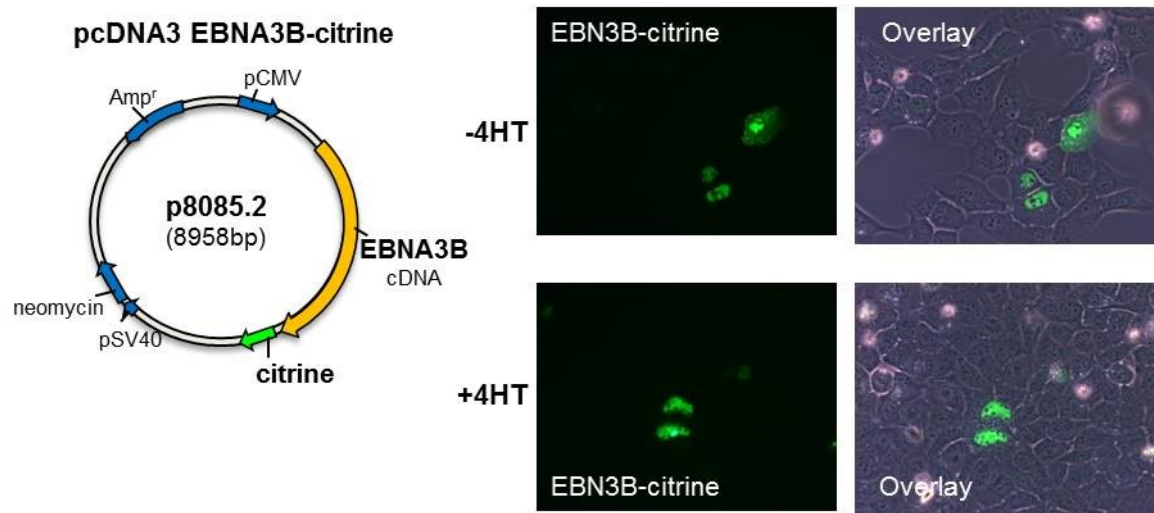
A**B**

Figure 3.10. ERT2-EBNA3B-citrine construct is stabilised with addition of 4HT 48h after transient transfection with pGY15.

(A) U2OS cells were transiently transfected with pGY15, which expresses the ERT2-EBNA3B-citrine fusion protein, and either left, (-4HT) or treated with 4HT (+4HT) for 48h before being analysed using Zeiss LSM510 laser scanning confocal microscope. **(B)** U2OS cells transfected with pcDNA plasmid coding for EBNA3B-citrine, treated with 4HT as in A. All cells are magnified 40x by the confocal microscope, using consistent exposure settings.

3.5 Recombinant ER3B viruses are produced and used to generate viable ER3B LCLs

Different EBV gene transcription programmes (latency 0-III) have been identified in various B-cell lines (LCL, BL, HL) (Deacon *et al.*, 1993; Rowe *et al.*, 1992) and epithelial (NPC) cells (Brooks *et al.*, 1992), whereby expression of EBV latent proteins are determined in part by the differential activation of Wp/Cp/Qps. This suggests possible differences in splicing, transcriptional, post-transcriptional and post-translational modifications arising from varying availability of cellular factors in different genetic backgrounds. Despite the presence of EBNA2 and EBNA-LP transcripts in EBV episome-transfected HEK-293 cells, both of these proteins are not usually detectable in the HEK-293 cells. We hoped that the phenomenon of undetectable ER3B protein expression is restricted to HEK-293 cells since our HEK-293 ER3B cell lines seemed to express correctly spliced transcripts of ER3B and possess ER3B episomes without detectable truncations or mutations. In addition, the ERT2-EBNA3B construct could be also be expressed in isolation. Because our primary interest is in the expression of ER3B protein in the more physiological B-cell background, we decided to use HEK-293 ER3B cell lines to produce recombinant EBV for the infection of primary B-cells.

The HEK-293 producer cells containing the recombinant EBV ER3B BAC were transiently transfected with BALF4 and BZLF1 expression plasmids to induce the lytic cycle and production of virions. Supernatants from these were assessed for infectivity by GFP expression in infected Raji cells. The titres from these producers were extremely low as compared to wild-type B95.8 virus obtained from Dr Rob White (Table 3.1). In order to generate LCLs expressing the recombinant ER3B protein, I infected peripheral blood lymphocytes (PBLs) isolated from buffy coat obtained from a single donor (D28) with various virus preparations.

Table 3.1. Table showing virus titres determined by Raji cell infections.
Green Raji Units (GR) are determined by counting the number of Raji cells infected per ul of virus used.

HEK-293 Cell line	Virus Titres (GRU)
ER3B C3	1.6
ER3B C18	1.1
WT I6	1034

PBLs infected with ER3B viruses from HEK-293 ER3B clones C18 and C3 transformed efficiently and LCLs grew out successfully in the absence of 4HT. Diagnostic digests of episomes rescued from these ER3B LCLs showed no sign of gross band size shifts on the pulsed field gel (other than changes in terminal repeat number, which is normal after an infection cycle), which is indicative of an intact ER3B BAC (Figure 3.11A).

I next tested whether 4HT was able to stabilise the expression of ER3B. Despite treating these LCLs with 4HT for up to 28 days, there were no detectable ER3B bands on western blot analysis, although as seen for the HEK-293 producer cell lines, EBNA3A and EBNA3C expression was not affected (Figure 3.11B). In order to verify that the splicing of ER3B mRNA transcripts has been successful in LCLs, I again PCR amplified across the U-exon to EBNA3B splice junction (Figure 3.12A). Faint bands are detected at around the predicted 1226bp area on the gel for both C3 and C18 clones, indicating that the correct splice products can be made, although the PCR efficiency may have been less than ideal since the primers have given off strong signal on the gel.

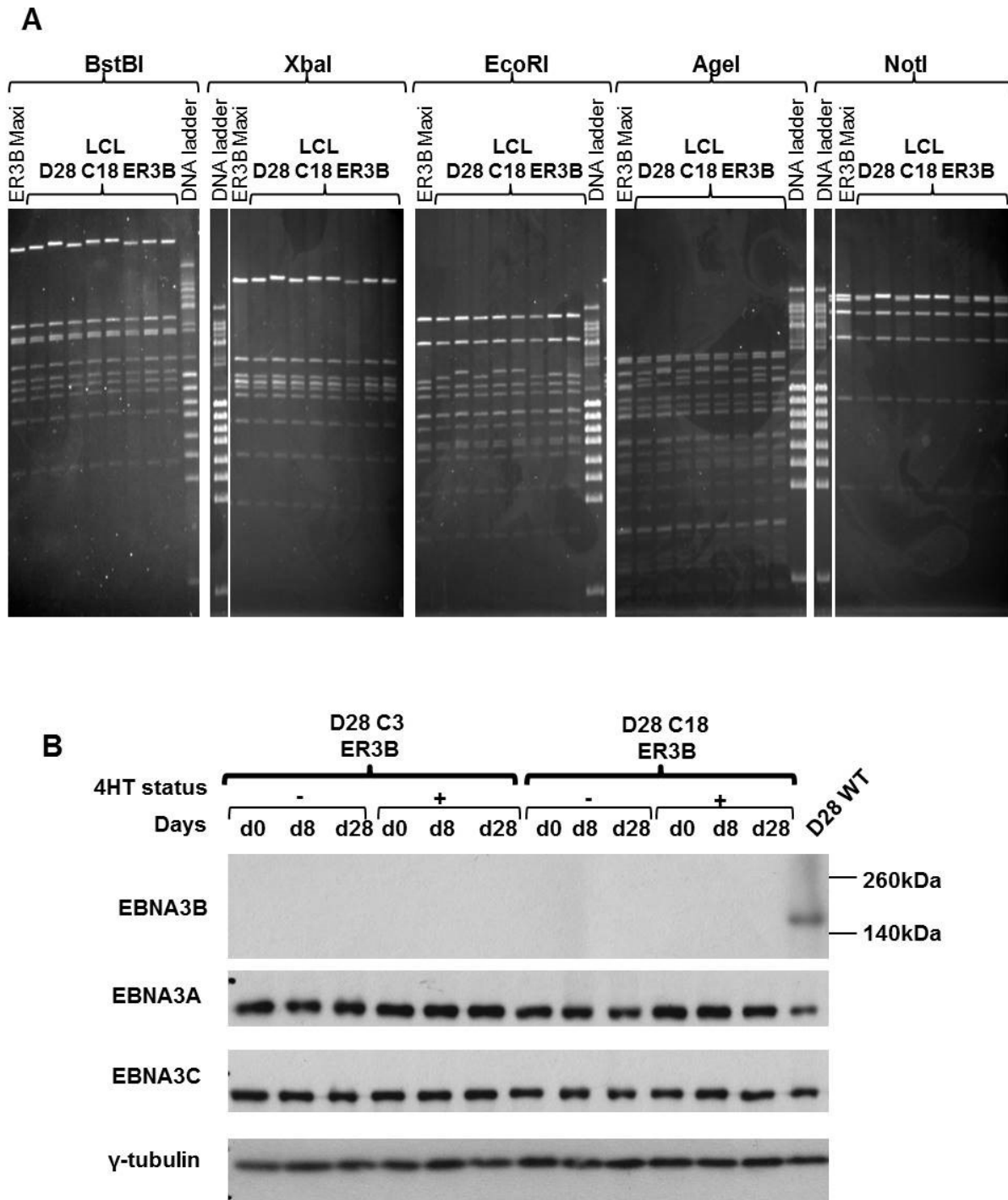


Figure 3.11. Verification of ER3B LCLs.

(A) Episomal BACs recovered from established D28 ER3B C18 LCL were assessed by restriction digestion and pulsed field gel electrophoresis. BACs rescued had undergone some alteration, but these are mainly limited to the terminal repeat regions (third band in *AgeI* and *EcoRI* digests). (B) Expression of EBNA3A, EBNA3B and EBNA3C proteins in ER3B C3 and C18 LCLs as detected by Western Blot after with or without treatment of 4HT for up to 28 days. The expected band size for recombinant ER3B protein is around 190 kDa and for WT EBNA3B is 150kDa. Tubulin is used as a loading control for the gel.

To verify the ER3B sequence, I used 7 sets of primers (previously described by White *et al*, 2012) to amplify the entire *ERT2-EBNA3B* domain (190bp upstream of *EBNA3B* through *EBNA3B* to 220bp downstream of the *EBNA3C* start codon) from low molecular weight DNA obtained from both the ER3B C3 and C18 LCLs (schematically shown in Figure 3.12B). The first set of primers would amplify from -190 to +209bp relative to the *EBNA3B* translational start site, which would include the inserted 945bp ERT2 domain. This set of primers failed to PCR-amplify in ER3B C3 LCL but was successful at amplifying the same region for the C18 clone (Figure 3.12C). The amplified fragments of ER3B C18 LCL were sequenced and the ERT2 domain sequence (and its junctions with EBNA3B) was verified to be as predicted (results not shown). The rest of *EBNA3B* (+177 to the 3' end of the gene) of the ER3B C18 LCL was also successfully amplified using the remaining 6 sets of PCR primers although these were not sent for further sequencing (Figure 3.12C).

Taken together, this has shown that ER3B protein is not produced from the viral genome, despite the sequence, transcription and splicing apparently being intact, and the protein (when expressed) being appropriately regulated by 4HT treatment. It appears that the lack of ER3B fusion protein expression is most likely due to a post-transcriptional or translational defect. However, since the main interest of my project was to focus on EBNA3B's function in regulating host genes, I decided to move on to another system that could be used to validate and study the kinetics of EBNA3B regulation of target genes within a uniform cellular background.

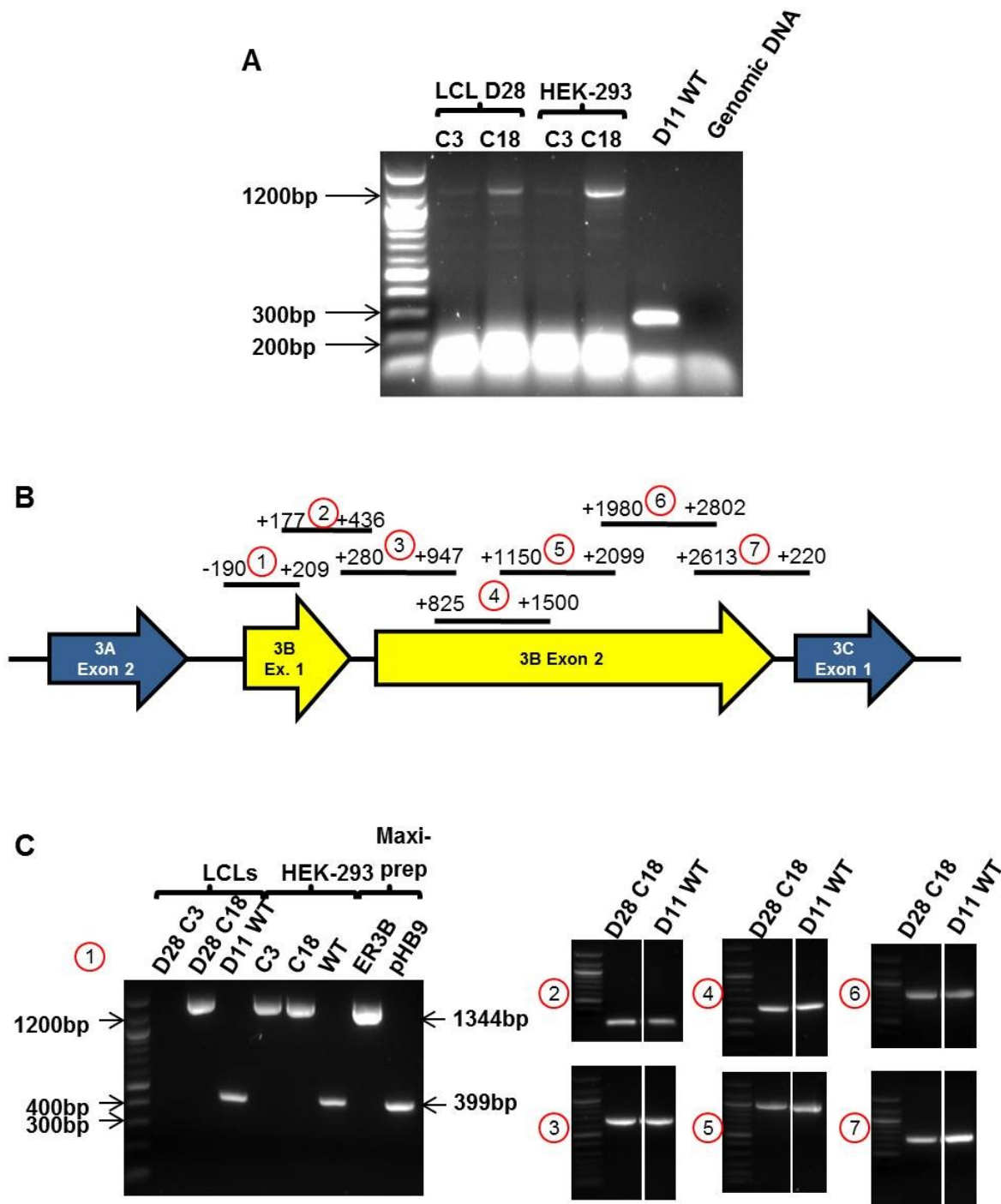


Figure 3.12. PCR amplification of cDNA or low molecular weight DNA recovered from ER3B LCLs and HEK-293 virus producers.

(A) cDNA from D28 C3 and D28 C18 ER3B LCLs and their corresponding HEK-293 producer cell lines are amplified by PCR across the U-EBNA3B splice junction. Expected amplified fragment size from ER3B and WT controls are 1226bp and 254bp respectively. (B) Schematic showing 7 sets of primers used to PCR amplify the whole of EBNA3B gene from cell-extracted virus genomes. Numbers indicate base pair position of the amplicon ends, relative to the EBNA3B start codon. Numbers within red circles indicate numbering of the primer sets. The first set of primers amplifies across the ERT2 region of ER3B BAC DNA (1344bp compared to WT 399bp). (C) Amplified DNA is shown on agarose gels alongside the 100bp ladder (New England Biolabs). Schematic of primer locations is adapted from White *et al.* (2012).

3.6 Generation and validation of doxycycline-inducible EBNA3B system in 3BKO LCLs.

In 2005, Bornkamm *et al.* described the pRTS-1 plasmid (Figure 3.13A), which generates a stable and stringent doxycycline-dependent inducible system for the expression of genes of interest in transfected B-cells (Bornkamm *et al.*, 2005). The group reported high inducibility of expression in the presence of doxycycline (Dox), low background activity and a dose-response to increasing concentrations of Dox. Significantly, the Kempkes lab successfully modified and used this plasmid to conditionally express EBNA3A in cell lines with an EBNA3A-null background (Harth-Hertle *et al.*, 2013). Hence, I chose to use a modified version of this plasmid in my next approach at establishing a conditional EBNA3B system.

Figure 3.13A illustrates the important features of the pRTS1 plasmid described by Bornkamm and colleagues (2005). Briefly, the chicken beta actin promoter and an element of the murine immunoglobulin heavy chain intron enhancer (Eu-CAGp) drive constitutive expression of a bicistronic cassette encoding the reverse tetracycline controlled transactivator rTA2S-M2 and the silencer Tet-repressor-KRAB fusion protein (tTSKRAB). Additionally, a bidirectional Dox-responsive promoter (Ptetbi-1) simultaneously drives transcription of the gene of interest (firefly luciferase) and eGFP, which is used as a marker of Ptetbi-1 activity. In the absence of Dox, the silencer tTSKRAB binds to the Ptetbi-1 promoter and the addition of Dox releases the silencer and activates the promoter by the activation and binding of transactivator rTTA2S-M2. This allows for a tight down-regulation of the Ptetbi-1 promoter to give low background expression of luciferase enzyme in the absence of Dox, and thus a highly sensitive dose-dependent response on the addition of Dox. In addition, the pRTS1 plasmid carries the EBV origin of replication (OriP) as well as *EBNA1* gene, which together would maintain episomal replication and ensure propagation of the episome into daughter cells (see section 1.2). The hygromycin resistance (hyg) gene –

under the control of SV40 promoter (SVp) – was used as a selection marker for successfully transfected cells.

Since our intention was to transfect the pRTS plasmid into LCLs that express EBNA1, our modified pRTS3 did not carry the *EBNA1* gene (Figure 3.13B). In addition, puromycin resistance gene (puro) was used as an alternative to hyg gene as a selection marker, since the 3BKO LCLs used in our lab already express hygromycin resistance. In our hands, SVp-driven puro did not confer puromycin resistance in B-cells (personal communication, Dr Rob White). To overcome this problem, the puro gene was inserted upstream of rtTA2S-M2, under the control of Eu-CAGp. In addition, Dr Rob White inserted a 2A-like peptide [derived from an insect virus, *Thosea asigna* virus (TAV)] between the puro and rtTA2S-M2 genes (Donnelly *et al.*, 2001; Pringle *et al.*, 1999). The 2A-like peptide mediates co-translational cleavage of polypeptides, hence facilitating the cleavage of puromycin resistance enzyme from rtTA2S-M2 transactivator (reviewed by Donnelly *et al.*, 2001 and Luke *et al.*, 2008). Another important feature of the pRTS3 is that the eGFP gene is replaced with a bicistronic expression cassette that encodes for two markers of Ptetbi-1 activity: a truncated version of nerve growth factor receptor (NGFR) and the fluorescent protein, mCherry. The NGFR can be used to magnetically sort for cells induced by Dox (see section 3.6.4). Since eGFP is already expressed in our 3BKO LCLs, the fluorescent mCherry is used as an alternative in flow cytometry to determine the percentage of Dox-responsive cells.

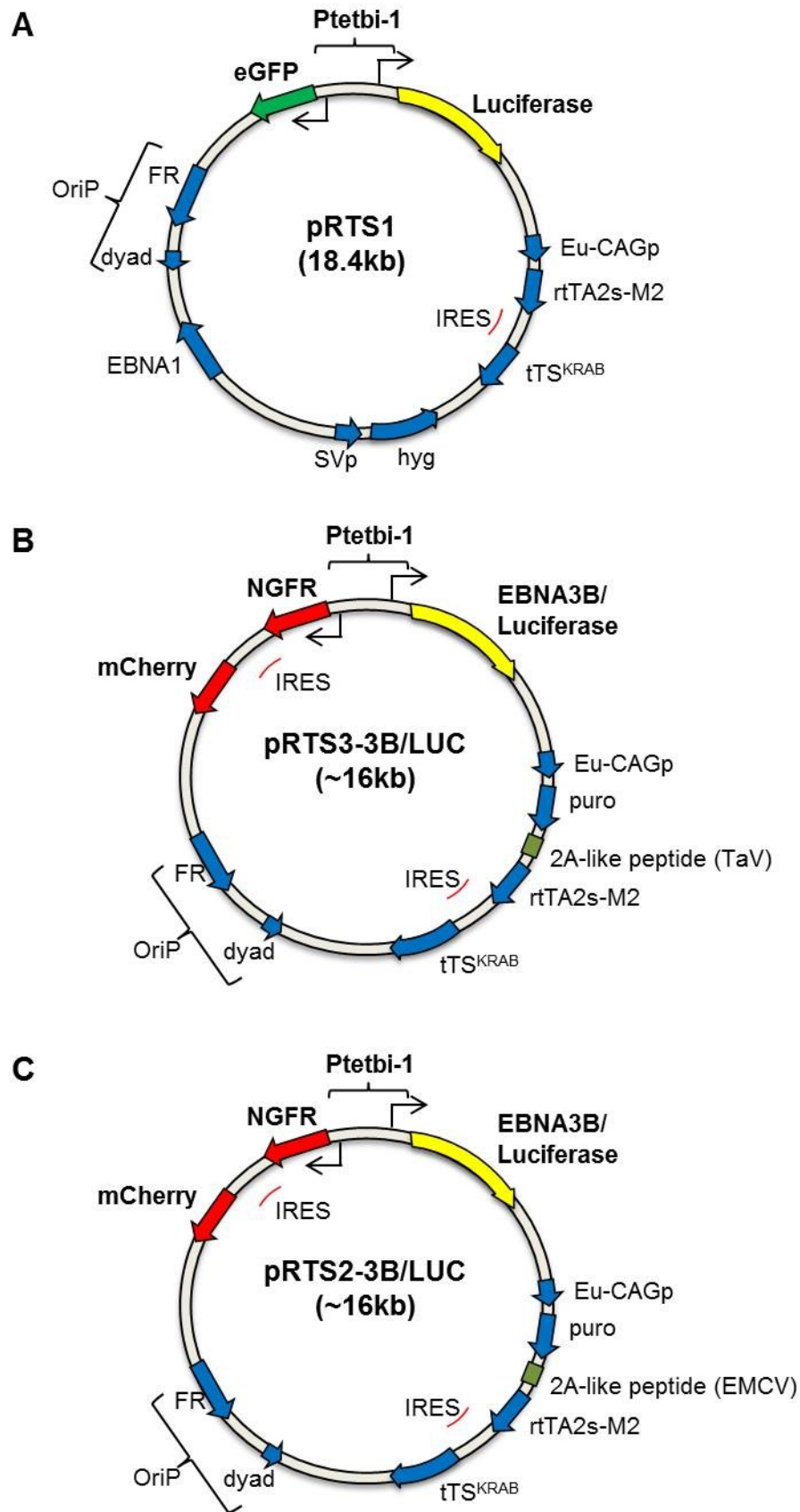


Figure 3.13. Schematic map of Doxycycline (Dox)-inducible RTS-vectors used to transfect 3BKO LCLs.

(A) pRTS1 vector described by Bornkamm et al (2005), and **(B)** pRTS3 and **(C)** pRTS2 vectors used in this study. Description of vectors found in section 3.6 (pRTS1 and pRTS3) and section 3.6.4 (pRTS2) of text.

Using a NEPA21 electroporator, I transfected the pRTS3 plasmid encoding conditional EBNA3B (pRTS3-3B) into 3BKO LCLs. At the same time, as a control for any plasmid-related effects seen in the established cell lines, the pRTS3 encoding for luciferase gene in place of EBNA3B (pRTS3-LUC) was transfected separately into these cell lines. Nine 3BKO cell lines established from a combination of various PBL donors and 3BKO virus clones (e.g. Donor 1 PBL infected with virus clone 5 is denoted as D1.5) were used for these transfections, and of these, 5 sets of 3BKO RTS3-3B (EBNA3B) LCLs and 3 sets of 3BKO RTS3-LUC (luciferase) LCLs grew out successfully under puromycin selection. The expression of EBNA3B and the responsiveness of these cell lines to Dox were further validated via western blots and flow cytometry (FACs).

3.6.1 Doxycycline successfully induced the expression of all RTS-vector components including EBNA3B in 3BKO LCLs.

As an initial validation of the RTS cell lines, I treated all 5 3BKO RTS3-3B LCLs with 50ng/ml of Dox over a period of 3 days. As shown in Figure 3.14A, all 5 cell lines can be induced and express EBNA3B at varying degrees after addition of Dox in the media. Importantly, the addition of Dox does not significantly alter the expression of other EBV latency III proteins, notably EBNA3A and EBNA3C, in these cell lines. Expression of mCherry is detectable by FACs (Figure 3.14B) in the Dox-treated cells, and the percentage of mCherry expressing cells appears to correlate directly with the amount of EBNA3B detectable by Western Blot in comparison between cell lines.

Next, FACs analysis of 2 sets of RTS cell lines (RTS3-3B and RTS3-LUC) from 2 different donors showed that mCherry and NGFR are expressed in a proportional manner within the same cell line (Figure 3.15). Together with the EBNA3B western blot data, this shows that the Ptetbi-1 bidirectional promoter is functional and responsive to Dox, and suggests that both NGFR and mCherry are good surrogates for EBNA3B expression level. In accordance to the observation made by Bornkamm *et al*, it appears that different cell lines respond differently to Dox; some are more responsive to the same concentration of Dox in

the media than others. It is unclear at present why this is the case. As Bornkamm and colleagues have suggested, this could be a result of rapidly evolving heterogeneity in the cell lines or due to outgrowth of multiple clones in the same transfected population of cells. At this point, these cannot be discriminated from each other.

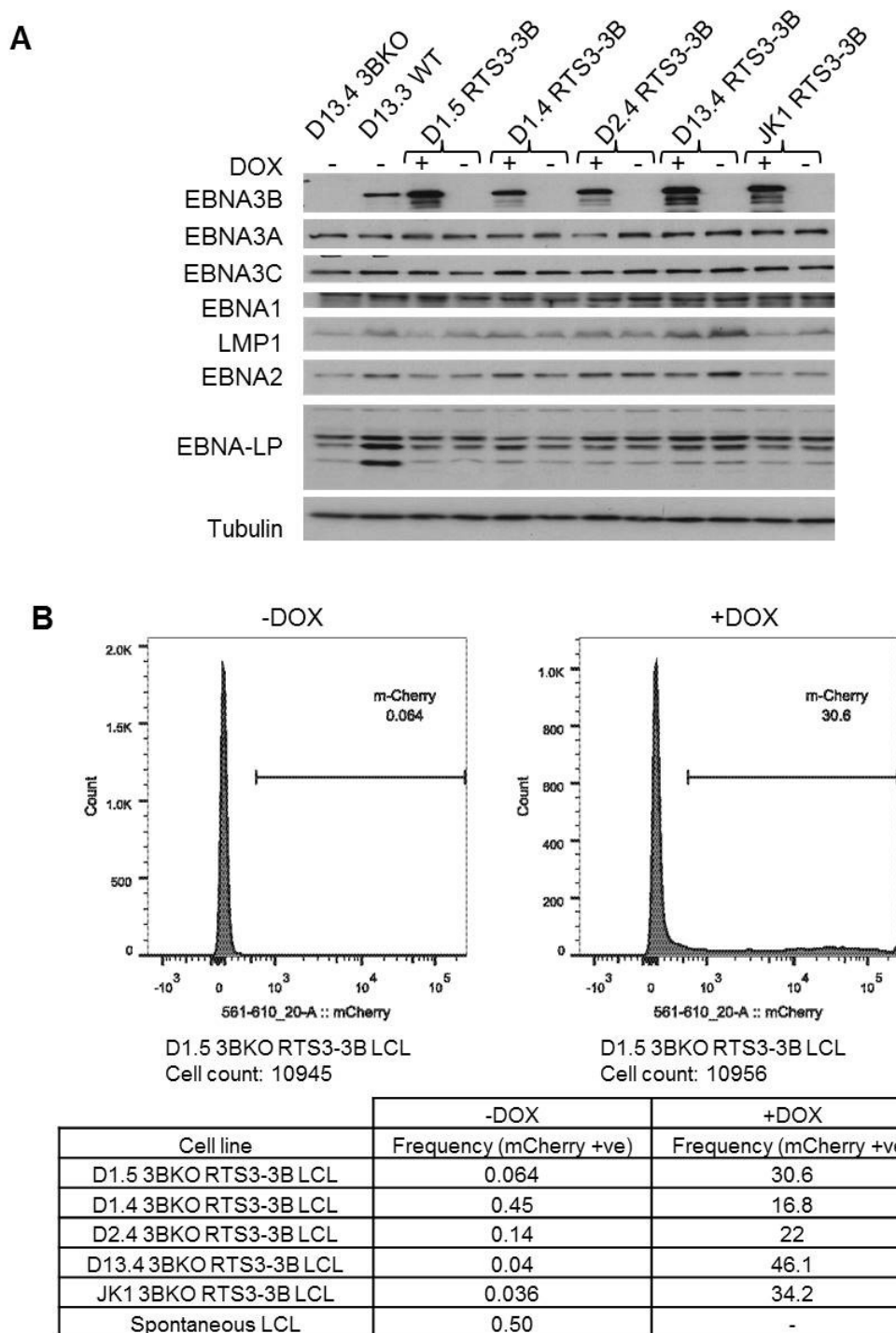


Figure 3.14. EBNA3B is expressed on the induction of RTS-3B 3BKO LCLs with 50ng/ml of doxycycline after 4 days.

(A) Expression of EBV latent proteins EBNA3B, EBNA3A, EBNA3C, EBNA1, EBNA2, EBNA-LP and LMP1 in 3BKO RTS3-3B cell lines treated with Doxycyclin (Dox). Variable expression of EBNA-LP, EBNA2 and LMP1 are sometimes reported in different cell lines. Protein extracts were taken from unsorted cells and 30ug of protein is loaded into each well. (B) FACS analyses of RTS3-3B cell lines to show the expression of mCherry. LCLs which grew out spontaneously without our recombinant EBV were used as negative gating controls.

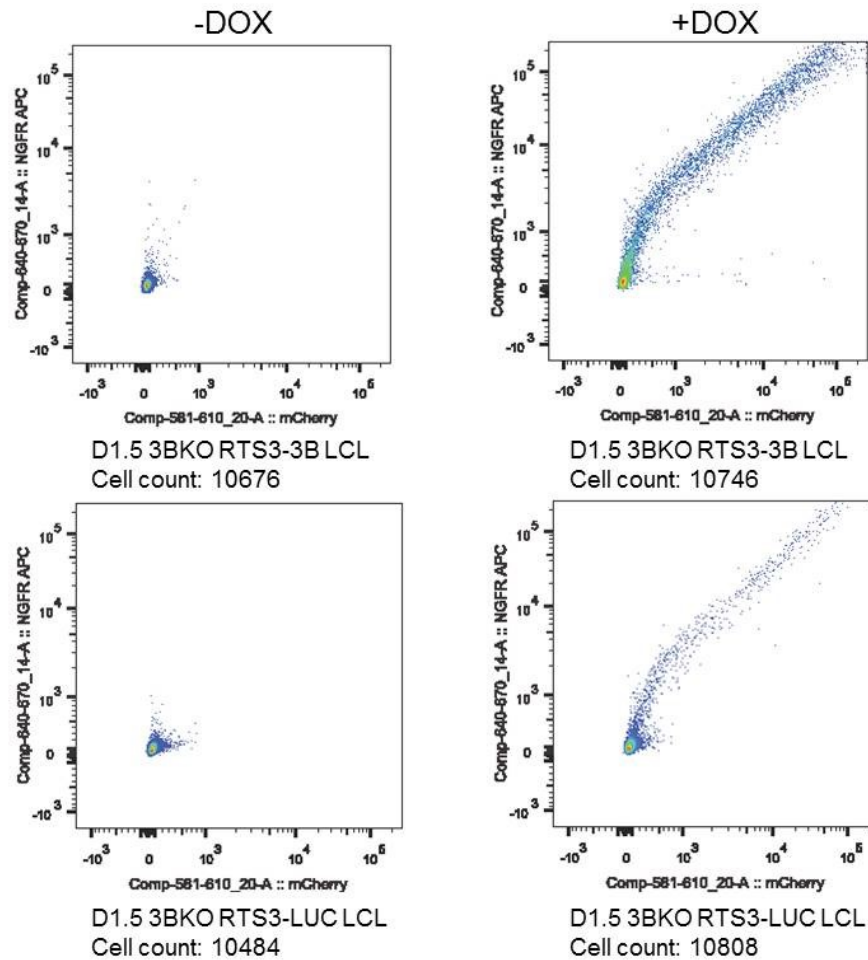


Figure 3.15. Expression of mCherry relative to truncated NGFR in RTS3-3B/LUC LCLs before and after induction with Dox.

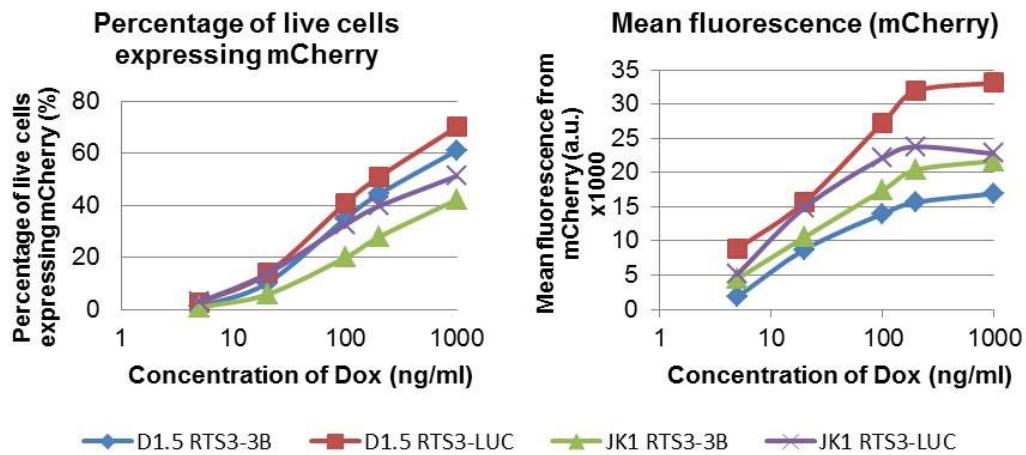
FACs showing the expression of mCherry (x-axis) in relation to truncated NGFR levels (y-axis) in RTS-3B and RTS-LUC cells after a 4-day induction with 50ng/ml of Dox. Results are representative of both D1.5 and JK1 RTS3-3B/LUC cell lines.

3.6.2 Increased dosage of doxycycline results in a dose-response increase in expression of EBNA3B.

In order to determine the best concentration to work with in my cell lines, I performed a dose titration using 0, 5, 20, 100, 200 and 1000ng/ml of Dox on 2 different sets of RTS3-3B/LUC cell lines (Figure 3.16) generated from donors D1 and JK1.

With an increasing concentration of Dox in the media, not only are higher percentages of cells induced and expressing detectable mCherry, the mean fluorescence intensity of mCherry-positive cells is also higher at higher dosages of Dox for all four inducible RTS3-3B/LUC cell lines (Figure 3.16A). Using D1.5 3BKO RTS3-3B LCL as an example to illustrate this, the histograms derived from FACs analyses of these inducible cells showed clearly that the overall proportion of mCherry-positive cells increased with increased dosages of Dox, and there was also a gradual shift in the mCherry fluorescence of the cells, with the cell count peaks starting at negligible mCherry fluorescent intensity at low dosages, then shifting to medium and high intensity with increased concentrations of Dox (Figure 3.16B). Consistent with these findings, higher levels of EBNA3B proteins were also detected with increasing dosages of Dox in both D1.5 and JK1 RTS-3B 3BKO LCLs (Figure 3.17). Hence, it appears that these cell lines combine both models of induction, with increased treatment of Dox corresponding to more cells being induced, and also higher induction at the single cell level.

A



B

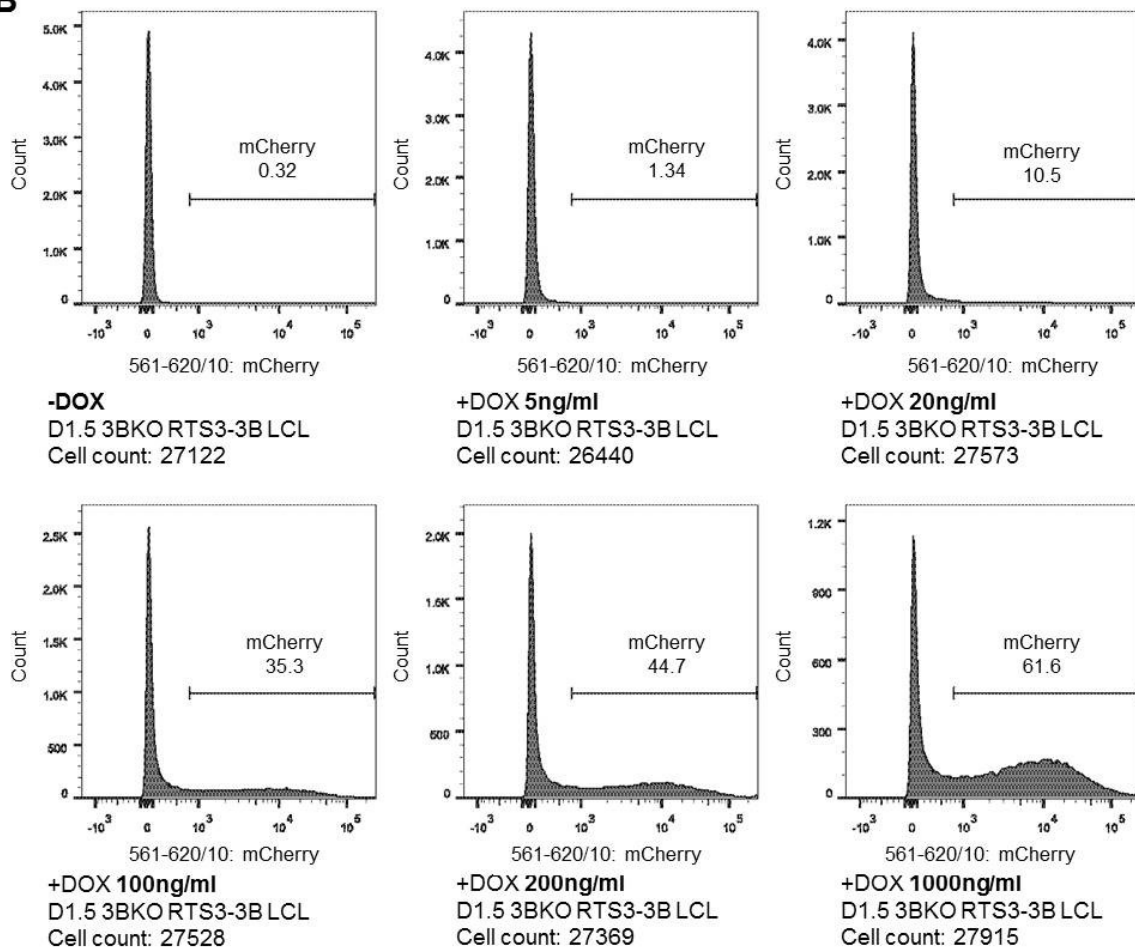


Figure 3.16. Dose-response induction of 3BKO RTS3-3B and RTS3-LUC LCLs.

(A) Graphs depicting the effect of increasing dosages of Dox on frequency of live cells expressing mCherry (left) and mean fluorescence of cells which are mCherry positive (right) in arbitrary units (a.u.). Dosages are plot on a log scale relative to the 0ng/ml treatment value (i.e. values from untreated populations are subtracted from the data points). (B) FACS data illustrating the frequency of cells expressing mCherry and their fluorescence intensity in response to increasing dosages of Dox in D1.5 3BKO RTS3-3B LCL. Data are representative of all four D1.5/JK1 3BKO RTS3-3B/LUC LCLs.

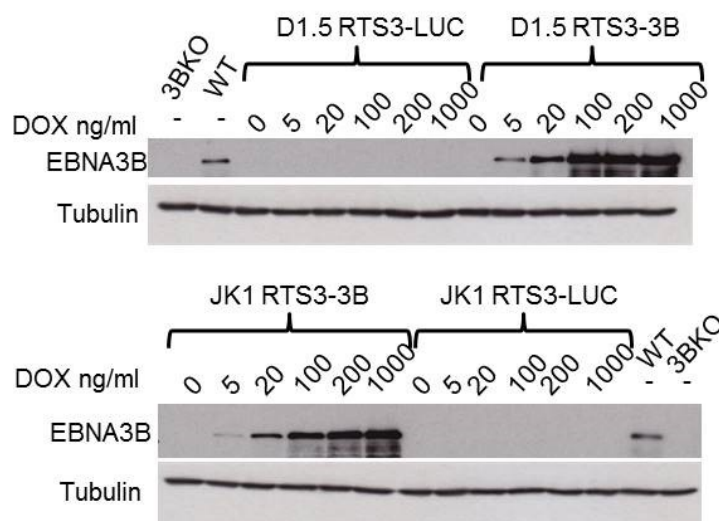


Figure 3.17. Western blot showing EBNA3B protein levels with increasing dosages of Dox. Western blot examining the expression of EBNA3B protein in D1.5 3BKO RTS3-3B/LUC and JK1 3BKO RTS3-3B/LUC LCLs with increasing dosages of Dox. Tubulin is used as loading control.

3.6.3 Long-term treatment of doxycycline maintains complementation of EBNA3B in transfected LCLs but shows no effect on selected predicted target genes.

To study the effects of long-term Dox exposure on the expression of EBNA3B, on the surrogate marker mCherry and on the effects on potential EBNA3B target genes, the 3BKO RTS3-3B LCLs were treated with Dox over a period of 23 days. Cells were grown with a relatively low level of Dox (20ng/ml) aiming to avoid excessively over-expressing EBNA3B.

The overall EBNA3B levels in the 3BKO RTS3-3B cell lines remained stable over the 23 days (Figure 3.18A). Three previously identified EBNA3B target genes (White *et al.*, 2010) were chosen based on available primer sets and used to determine whether this long term complementation of 3BKO cells with EBNA3B would affect their expression. Specifically, according to the previous microarray, *CD28* was induced by EBNA3B, and *FOXP1* and *LAIR1* were EBNA3B repressed genes. The basal expression levels of these genes in the LCLs varied greatly from the parental 3BKO LCLs, regardless of whether Dox had been added (Figure 3.18B), which seemed to suggest that the outgrowth of these cells have selected for specific clones. Interestingly, over the course of the 23 days, the expression of *CD28* and *LAIR1* in the pRTS-vector transfected 3BKO LCLs appeared to drift

(up for CD28 and down for LAIR1), regardless of Dox treatment (Figure 3.18B). This suggests that results obtained after long-term treatment of cells may not be reliable due to this sort of variation in the expression of some genes.

There appeared to be no significant difference in the expression levels of any of the three genes between Dox-treated and untreated D1.5 RTS-3B 3BKO LCLs, even after the 23 days (Figure 3.18B). There are various possible explanations for this observation: these genes may not necessarily be regulated by EBNA3B in this system; 23 days may be too short a time to detect any induction or repression of these particular genes; or the number of cells expressing EBNA3B in the population is so small that the overall effect on EBNA3B target genes is masked by higher numbers of non-responsive cells. This experiment highlighted the need to ensure a homogenous population of cells expressing EBNA3B, and the importance of having a good positive control EBNA3B target gene.

3.6.4 Magnetic sorting for NGFR+ cells increases yield of Dox-responsive cells but these cell populations tend to drift back to unsorted phenotypes over time.

Truncated NGFR has been used successfully to magnetically sort for cells that are expressing the pRTS vector (Harth-Hertle *et al.*, 2013). In order to eradicate any background expression of genes which could mask any actual EBNA3B regulation by the small percentage of the EBNA3B-expressing cells, the autoMACs cell separator (Miltenyl Biotec) was used to magnetically sort for NGFR-positive cells after 3 days of treatment with Dox (20ng/ml). The expression pattern of mCherry in these sorted cells was then monitored over the next 23 days (Figure 3.19). After sorting, more than 85% of all the 3BKO RTS3 LCLs expressed mCherry. These sorted cells were maintained in 20ng/ml Dox over the same period of 23 days to determine whether they maintain the high level of expression from the RTS vectors for this length of time.

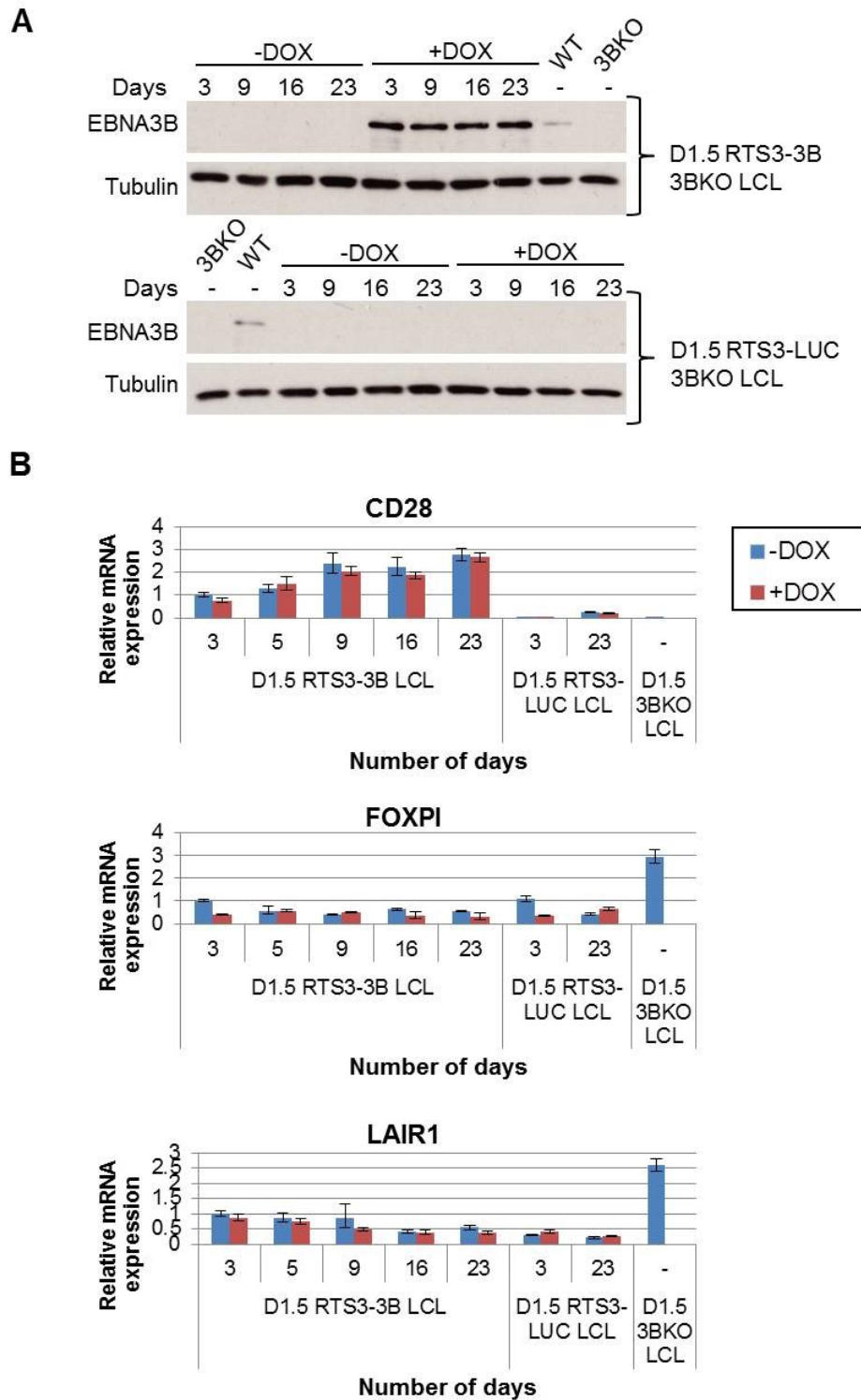


Figure 3.18. Expression of EBNA3B and its candidate regulated genes in D1.5 3BKO RTS-3B and RTS-LUC LCLs after a 23-day treatment with 20ng/ml of Dox.

(A) EBNA3B expression in D1.5 3BKO RTS3-3B and D1.5 RTS3-LUC LCLs, with and without Dox treatment. (B) RT-qPCR gene expression data for *CD28*, *FOXP1* and *LAIR1* comparing induced and un-induced D1.5 3BKO RTS3-3B cells over the 23 days. Expression of these genes in the parental cell line D1.5 3BKO LCLs (lacking an RTS vector) is also shown. The expression data are normalised to housekeeping genes *RPLP0* and *GNB2L1* and are reported relative to the untreated D1.5 3BKO RTS-3 3B cells at day 3. Error bars are derived from technical triplicates in one qPCR experiment. These results are representative of the 2 distinct experiments each conducted on cell lines derived from 2 donors (D1.5 and JK1).

Curiously, despite continual treatment with Dox, the mCherry expression of these cells drifted towards the unsorted phenotype over the 23 days (Figure 3.19A). In three of the four cell lines, this took a week or less from the time of sorting to matching the unsorted distribution of mCherry expression. Western blot analyses confirmed that the expression of EBNA3B followed this trend, and was apparent even in the early days after sorting (Figure 3.19B). This drift is observable more quickly in low purity populations (i.e. populations which started with lower percentage of mCherry-positive cells). At present, the explanation for this phenomenon remains unknown. It is, however, clear that sorting the Dox-treated cells at every time-point is necessary to obtain pure populations of fully-induced cells in order to detect changes in gene expression.

In summary, this 23-day long Dox treatment of the pRTS-transfected cell lines have drawn attention to certain challenges that need to be considered when using this system:

1. That the expression of some genes drifts over extended periods of time, regardless of Dox treatment or vector transfected.
2. Even populations purified to contain more than 90% of vector-expressing cells drift towards the unsorted phenotype over time.
3. It is essential to have a positive control gene which one is confident of being regulated by EBNA3B in order to properly validate this system and ensure the EBNA3B expressed is fully functional.

I initially generated this conditional-EBNA3B for the purposes of determining *bona fide* EBNA3B target genes and to study their expression kinetics over time. Given that the expression of genes may drift over long periods of time despite Dox treatment, and that introduction of EBNA3A using the pRTS vector was reported to robustly repress its target genes CXCL10 and CXCL9 within 24 hours (Harth-Hertle *et al.*, 2013), I decided to use the EBNA3B Dox-inducible system to generate a new list of genes that may be rapidly and robustly regulated upon complementation with EBNA3B within a shorter period of time.

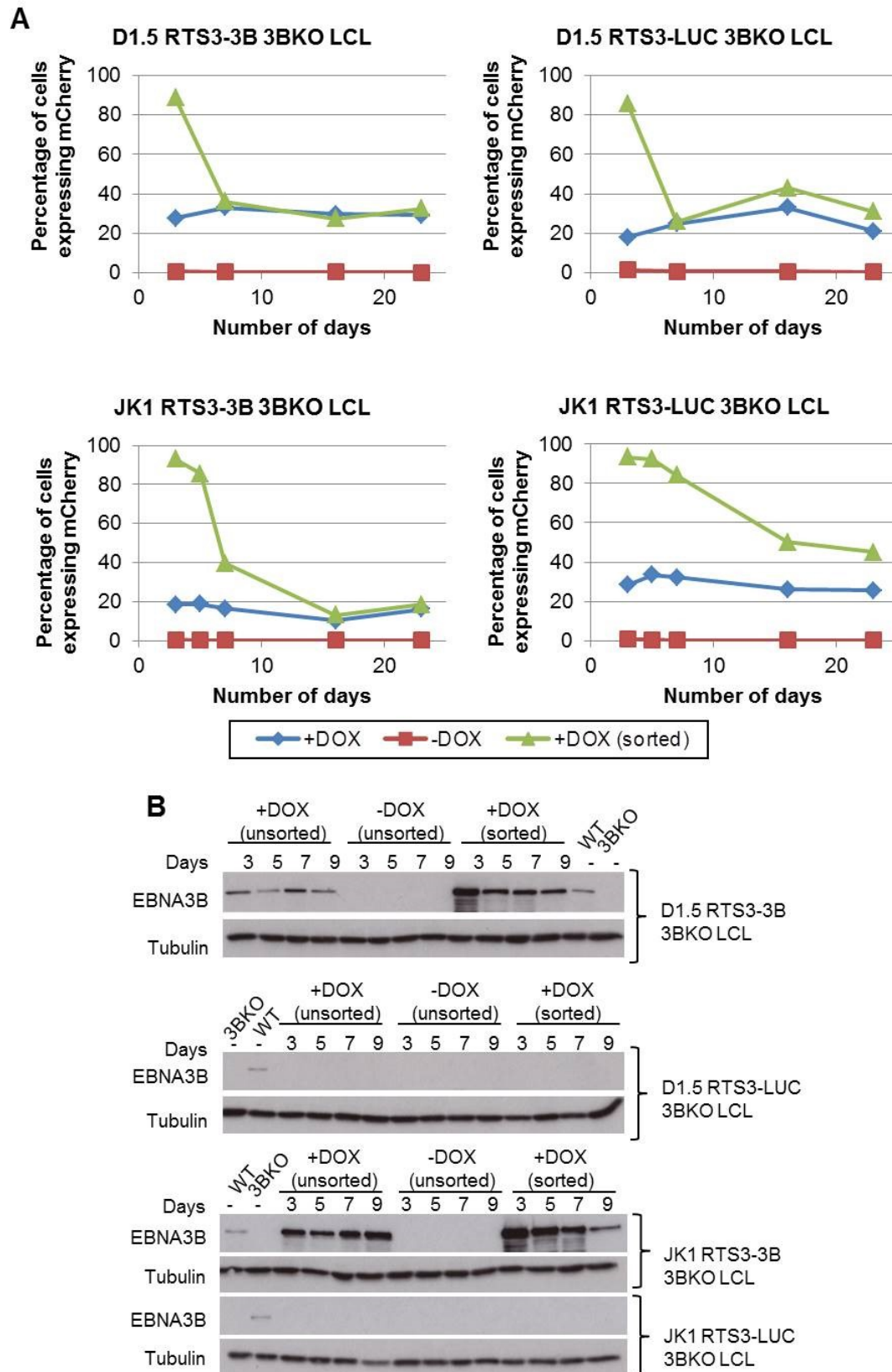


Figure 3.19. Magnetically-sorted RTS-3B/LUC cell populations from 2 different donors (D1.5 and JK1) drift to unsorted phenotype over time.

(A) The percentage of mCherry expressing cells over the 23 days for cell lines enriched for NGFR+ve cells at day 3 (first data point), and maintained in Dox, compared with the same cells grown continuously with or without Dox. X axis indicates time after initial induction with Dox **(B)** Protein levels of EBNA3B in the sorted and unsorted 3BKO RTS3-3B/LUC LCLs over the first 6 days after sorting.

At around this time, McClellan *et al.* (2013) demonstrated that EBNA3B plays a role in the repression of *ITGAL*, by competing with the viral transactivator EBNA2 for binding at its promoter region. Their findings concur with our microarray data, making *ITGAL* a potentially suitable control gene to validate the RTS system. With this in mind, I conducted a 7-day timecourse over which cells treated with Dox (20ng/ml) and subsequently NGFR-enriched were harvested on day 3 and day 7 (Figure 3.20). The LCLs were obtained from Dr Rob White, who had transfected another generation of pRTS vectors (p8304 and p8301 for RTS2-3B and RTS2-LUC respectively) into 3BKO cells derived from donor 13. The main difference between these and RTS3 vectors (p8317/p8316) is that a different 2A-like peptide derived from Encephalomyocarditis virus (EMCV) (Donnelly *et al.*, 2001; Hahn and Palmenberg, 1996) was inserted between the puro resistance gene and rtTA2S-M2 in the RTS2 (Figure 3.13C). This EMCV-derived 2A-like peptide does not introduce a translation-induced break between the two proteins as was intended, and results in a fusion protein puro-rTTA2S-M2. Although the reasons remain unknown, this does not appear to alter any aspect of the response to Dox in the cell lines or the efficacy of puromycin selection (results not shown).

RT-qPCR analyses revealed that sorting makes a substantial difference to our ability to detect effects on host gene regulation after complementation with EBNA3B. *ITGAL* expression is repressed just 3 days after induction with Dox, but this is only observable in the NGFR-sorted cells (Figure 3.20A).

During another 6-day timecourse, untreated cell populations were negatively sorted (i.e. cells were sorted for NGFR and only NGFR- cells were harvested) to avoid unwanted background gene expression changes caused by the processing and sorting of the cells, and to remove any spontaneously induced cells. By doing so, clean data for both treated and untreated cells were obtained, aiming for the most reliable comparisons of *ITGAL* expression. Again, the results suggest that 3-4 days of treatment is enough to induce the robust repression of *ITGAL* and that 6-7 day treatment is not more reliable (Figure 3.20B).

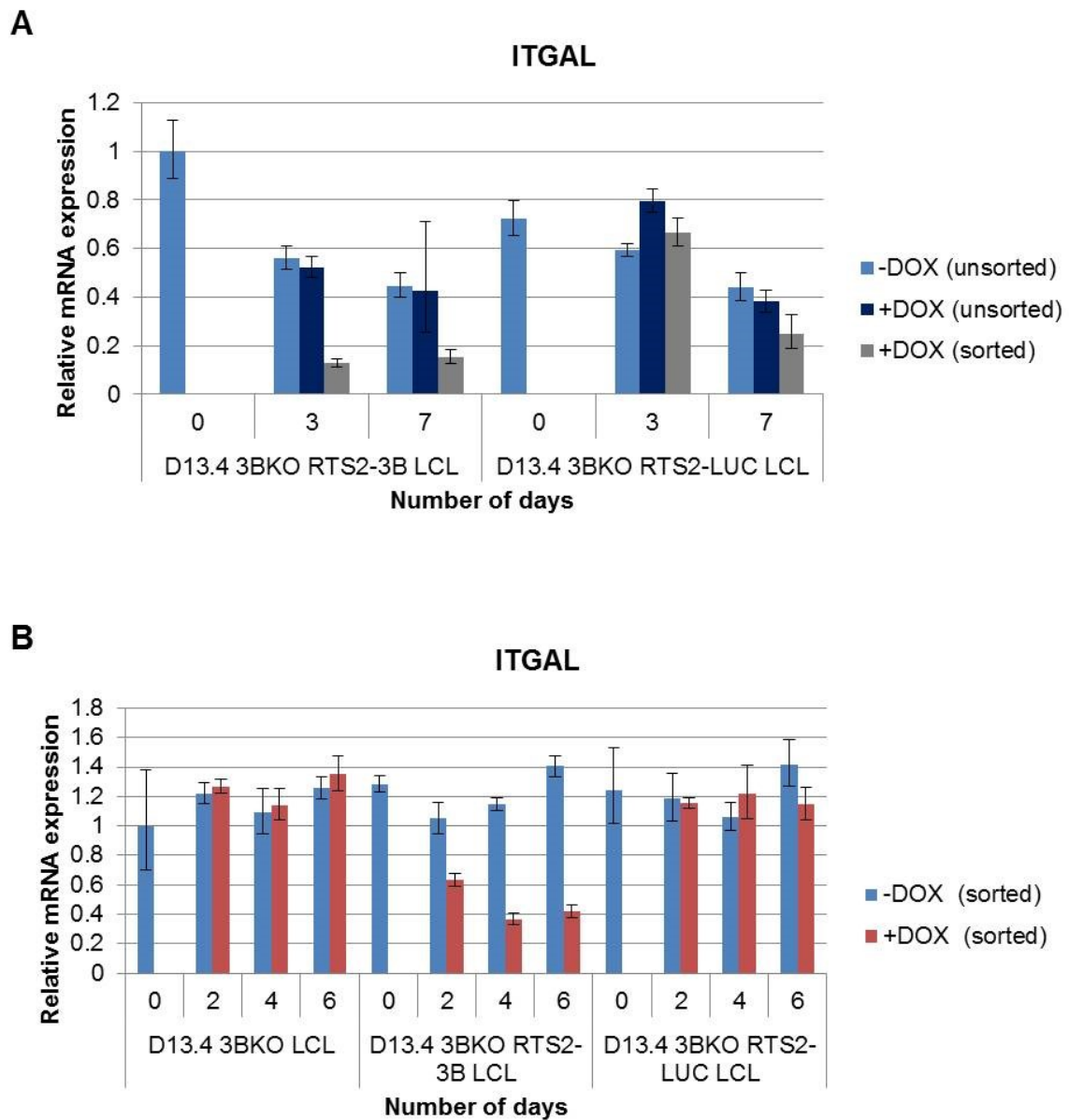


Figure 3.20. Short-term Dox induction (20ng/ml) of RTS2-3B/LUC cells using ITGAL expression as a positive control gene.

(A) Expression of *ITGAL* in D13.4 3BKO RTS2-3B/LUC LCLs pre and post-sort for each day 0, day 3 and day 7 of Dox treatment. (B) As in A, but untreated cells were additionally negatively sorted for NGFR (NGFR⁻) and time points taken at days 2, 4 and 6. The qPCR data were normalised to *GNB2L1* and *RPLP0* housekeeping genes and compared relative to the day 0 value of parental D13.4 3BKO LCLs. Error bars represent technical triplicates in one qPCR reaction.

3.6.5 Four-day dose treatment of Dox (with sorting for NGFR) consistently shows regulation of ITGAL by EBNA3B.

Encouraged by positive validation of this Dox-inducible EBNA3B system with *ITGAL*, I proceeded to induce EBNA3B expression in additional cell lines: D13.4 (RTS2) and D1.5 (RTS3) 3BKO RTS-3B/LUC LCLs with a short 4-day Dox treatment where cells were split into 4 groups and treated using 0, 10, 100 and 1000ng/ml of Dox. Cells were NGFR-sorted at day 0 (pre-treatment) and day 4 and their mCherry expression is illustrated in Figure 3.21.

Cells treated with low levels of Dox exhibit a broad range of mCherry expression post-sort, whereas at higher Dox concentrations, cells post-sort were disproportionately the higher mCherry expressing cells (Figure 3.21). Western blots confirmed a dose response of EBNA3B expression, to increasing dosage of Dox, and RT-qPCR confirmed that expression from *ITGAL* also decreased in a Dox dose-dependent manner in both of the 3BKO RTS-3B cell lines after 4 days' exposure to Dox (Figure 3.22). In order to broaden the range of genes tested, cDNAs from this experiment were used in a custom TaqMan low density array (TLDA) card (White *et al.*, 2010) containing primers and probes for TaqMan assays which would amplify cDNA of 48 genes: 42 of these had been chosen as candidate EBNA3B target genes – selected from our exon-microarray data described above – and 6 are housekeeping genes for use as normalising controls. This allowed us to perform a quick and efficient validation of a group of candidate genes.

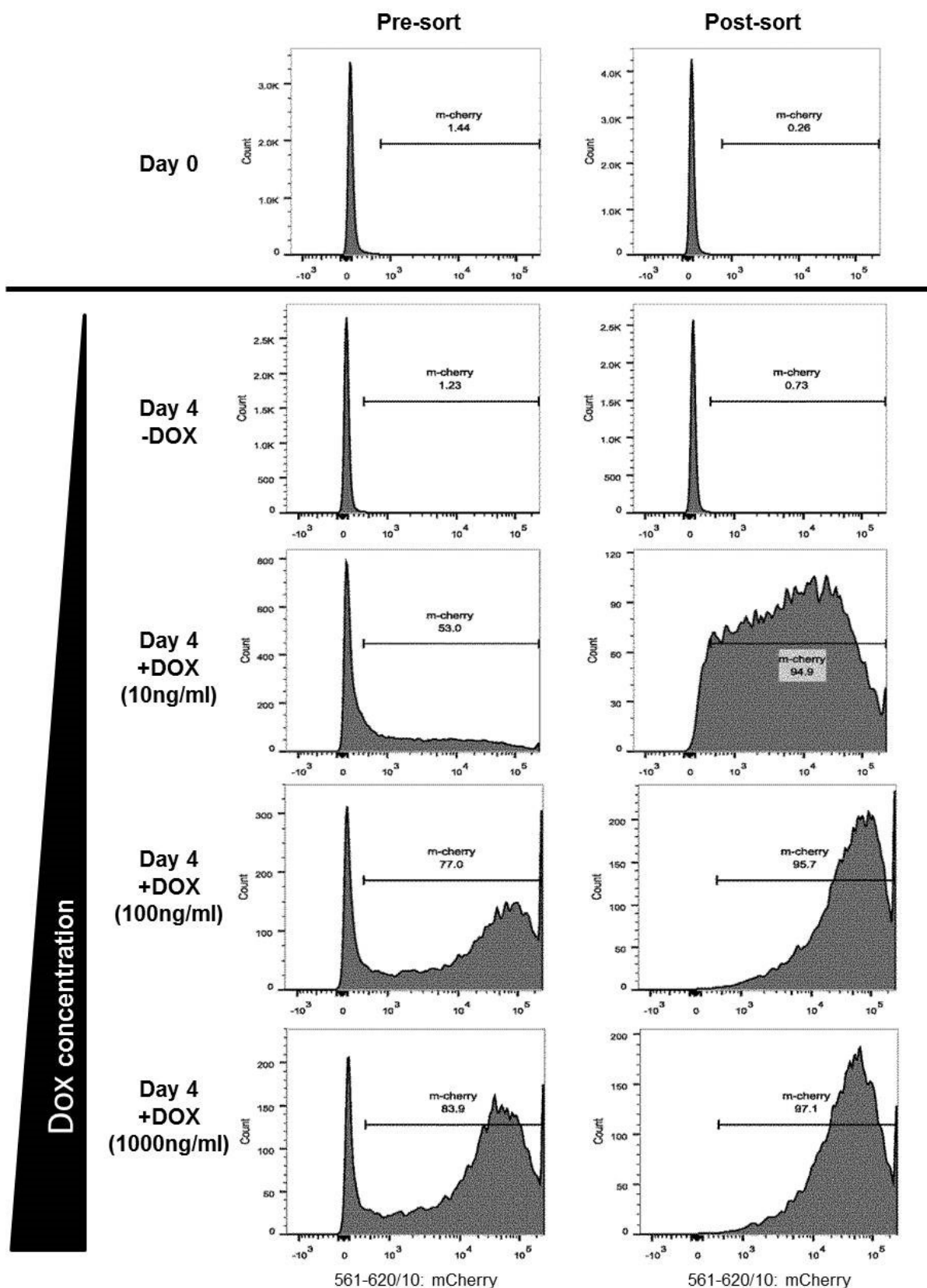


Figure 3.21. Flow cytometry histograms of D13.4 RTS-3B cells pre and post-sort for NGFR positivity, four days after Dox treatment.

Parental D13.4 3BKO and untreated RTS2-3B LCLs are negatively sorted for NGFR expression whereas all Dox treated cells are sorted for NGFR positivity.

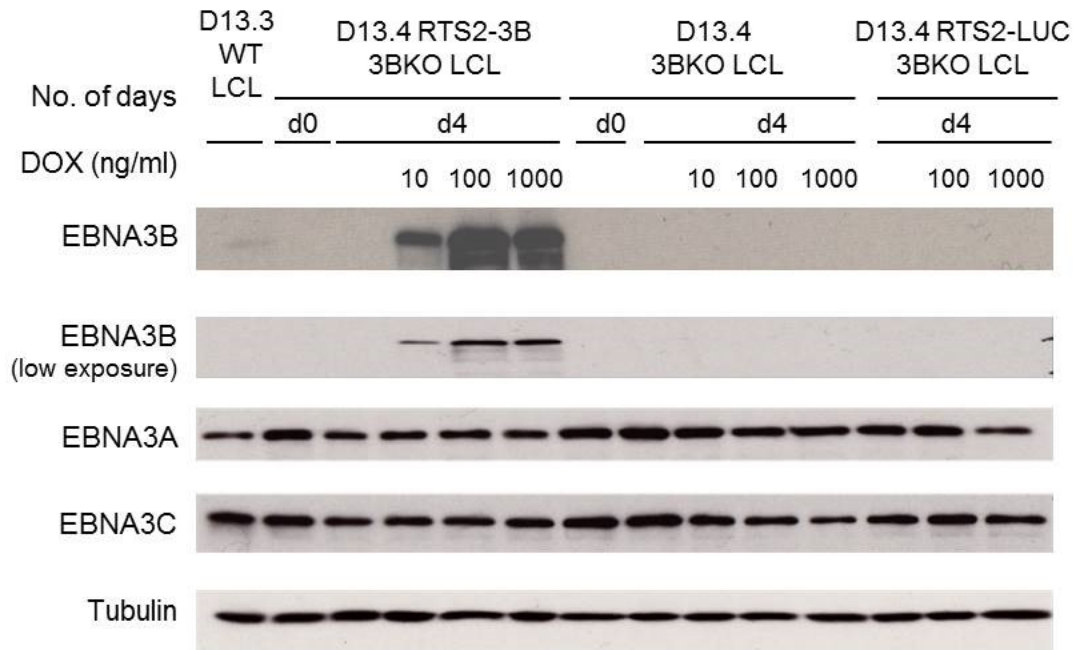
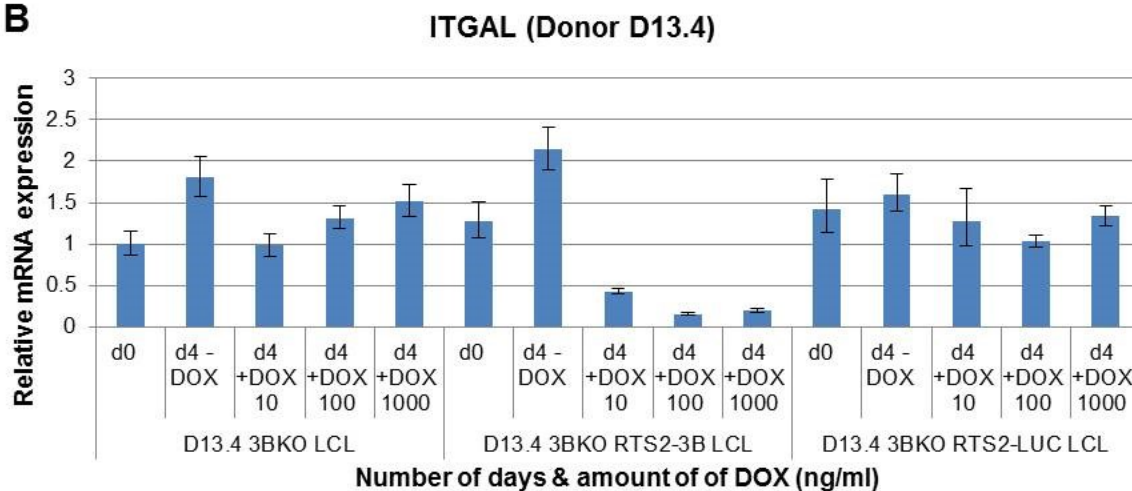
A**B**

Figure 3.22. Validation of 4-day Dox titration in NGFR-sorted D13.4 RTS-3B/LUC 3BKO LCLs using *ITGAL* as a positive control.

Parental D13.4 3BKO and untreated RTS2-3B/LUC LCLs are negatively sorted for NGFR expression whereas all Dox treated cells are sorted for NGFR positivity. **(A)** Western Blot showing levels of EBNA3B with increasing dosage of Dox, compared to wild-type LCLs. Tubulin is used as a loading control. **(B)** RT-qPCR analyses of *ITGAL* expression with increasing dosage of Dox. The data were normalised to *GNB2L1* and *RPLP0* housekeeping genes and compared relative to the day 0 value of parental D13.4 3BKO LCLs. Error bars represent technical triplicates in one qPCR reaction.

3.6.6 Analysis of 43 potential EBNA3B regulated genes using Taqman Low-Density Array card to identify possible rapidly regulated targets of EBNA3B.

Data from two array card experiments were analysed using DataAssist software provided by Life Technologies, and displayed using two different heat maps (Figure 3.23 & Figure 3.24). Samples from the parental 3BKO LCLs were used as comparisons for any gene modulation or toxicity due to Dox treatment alone. In the first heat map, the data are presented relative to the untreated sample in each set of 4 dosages for every cell line (Figure 3.23) – this is to show the variation within a single cell line of the addition of increasing dosages of Dox. In order to illustrate the differences between the baseline gene expression levels of the parental 3BKO and RTS-3B LCLs, data from each donor (including both the parental 3BKO and RTS-3B LCL) are presented relative to the untreated sample of corresponding the RTS-3B LCL in the second heat map (Figure 3.24). We find that variation in basal gene expression level is common between original pRTS-null 3BKO LCLs and 3BKO RTS-3B LCLs, which may be a result of clonal selection during the outgrowth process (Figure 3.24). In particular for four genes, *SMAD1*, *GAS7;MLL*, *SORT1* and *TGFBR3*, transfection of pRTS-3B vectors into the parental 3BKO LCLs appeared to have altered their expression in both donors (activation for *GAS7;MLL*, *SORT1* and *TGFBR3*, and in either directions for *SMAD1*) (Figure 3.24). This suggests that any regulatory effects on these genes with increasing dosages of Dox may likely be due to another component of pRTS vector.

Due to cost constraints, only 2 TLDA card experiments were performed and we were unable to perform any statistical analyses on the data. Hence, we devised a set of criteria used to assess whether the gene was regulated by EBNA3B, which is listed in Table 3.2. Of the 42 genes investigated, 26 were predicted EBNA3B-repressed genes and 17 were predicted to be EBNA3B-induced. Of the candidate EBNA3B-repressed genes, 7 genes (*ITGAL*, *LAIR1*, *TERT*, *LILRA4*, *MAP4K4*, *CDH1* and *ITGB3*) (Figure 3.23, Figure 3.24 & Table 3.3) fulfilled the criteria for definite EBNA3B-regulated genes, while 8 genes (*VAV3*,

CRLF2, *ZMIZ1*, *ST3GAL5*, *ASCL1*, *IL6R* and *FOXP1*) fulfilled the criteria for probable EBNA3B-regulated genes, suggesting that they are likely EBNA3B-repressed (Table 3.3). Finally, EBNA3B regulation was not confirmed, or below the threshold of detection for 11 of these genes (Table 3.3).

Table 3.2. Criteria used for determining EBNA3B regulation

Regulation	Criteria/Threshold
Definite	<u>Repressed genes (in both donors):</u> 1a. Compared to -Dox levels, less than 0.5-fold expression at all concentrations of Dox and/or 1b. A clear dose-response. <u>Induced genes (in both donors):</u> 1c. Compared to -Dox levels, more than 2-fold expression at all concentrations of Dox and/or 1d. A clear dose-response.
Probable	2a. Repression/Induction (dose response, <0.5-fold expression or >2-fold expression) only observed in one cell line. 2b. Regulation may be present, but there is an altered gene expression just on addition of RTS vector to the parental 3BKO cell line (in both donors). 2c. Regulation only at higher doses of Dox (100-1000ng/ml).
Not confirmed	3a. No clear trend of induction or repression. 3b. Dose response observed for the gene in 3BKO cells without pRTS vector. 3c. Expression of gene is undetectable in cell line. 3d. Regulation in the opposite direction between 2 3BKO RTS-3B LCLs.

Among the candidate EBNA3B-induced genes, only *TNFSF10* seemed to exhibit dose-dependent induction by EBNA3B in 3BKO RTS-3B LCLs from both donor backgrounds (Table 3.4). Six genes (*CD72*, *ITGB1*, *C13orf15*, *SMAD1*, *GAS7*; *MLL* and *SORT1*) (Table 3.4) fulfilled the criteria for probable EBNA3B-activated genes, and ten genes did not seem to be activated under these conditions. Thus I have shortlisted, using the Dox-inducible system, a list of genes that are robustly and rapidly regulated upon complementation of

EBNA3B-null cells with EBNA3B within a single cellular background. This paved the way for further investigation of how EBNA3B is able to regulate its host genes.

3.7 Overview

In order to study the function of EBNA3B on host gene regulation, I have attempted to establish a system whereby EBNA3B expression can be controlled in a single genetic background, by fusing a 4HT-conditional ERT2 tag to the N-terminus of EBNA3B (see section 3.3 - 3.5). Unfortunately, I was unable to detect the expression of this fusion protein. My next attempt at generating a conditional Dox-inducible EBNA3B system was more fruitful and returned a list of genes that have not only been shown to be regulated using exon-microarray studies in wild-type versus 3BKO LCLs (White *et al.*, 2012) but also exhibit a graded response to increasing amounts of EBNA3B in an EBNA3B-null LCL background (see section 3.6.6). While the lack of sophisticated statistical analyses, which caused difficulty in data interpretation due to the variation in responses between cell lines (discussed in Chapter 6), is a major limitation of the analyses, the TLDA card experiments have identified potential genes for further investigation.

The EBNA3s are known to be transcriptional regulators and have been shown to bind to several DNA binding factors. They have also been reported to play a role in chromatin remodelling to exert their gene-regulatory functions. Hence, the next step in investigating the function of EBNA3B was to determine whether it binds in the vicinity of any of these potential targets to provide further insight into the mechanism of EBNA3B-mediated gene regulation. This will bring us to the next chapter, where I will discuss the generation of an LCL expressing an epitope-tagged EBNA3B, its use in Chromatin Immuno-precipitation (ChIP) experiments and further ChIP linked to high throughput sequencing (ChIP-seq) to determine genome-wide localisation of EBNA3B during the latent infection of EBV transformed B-cells.

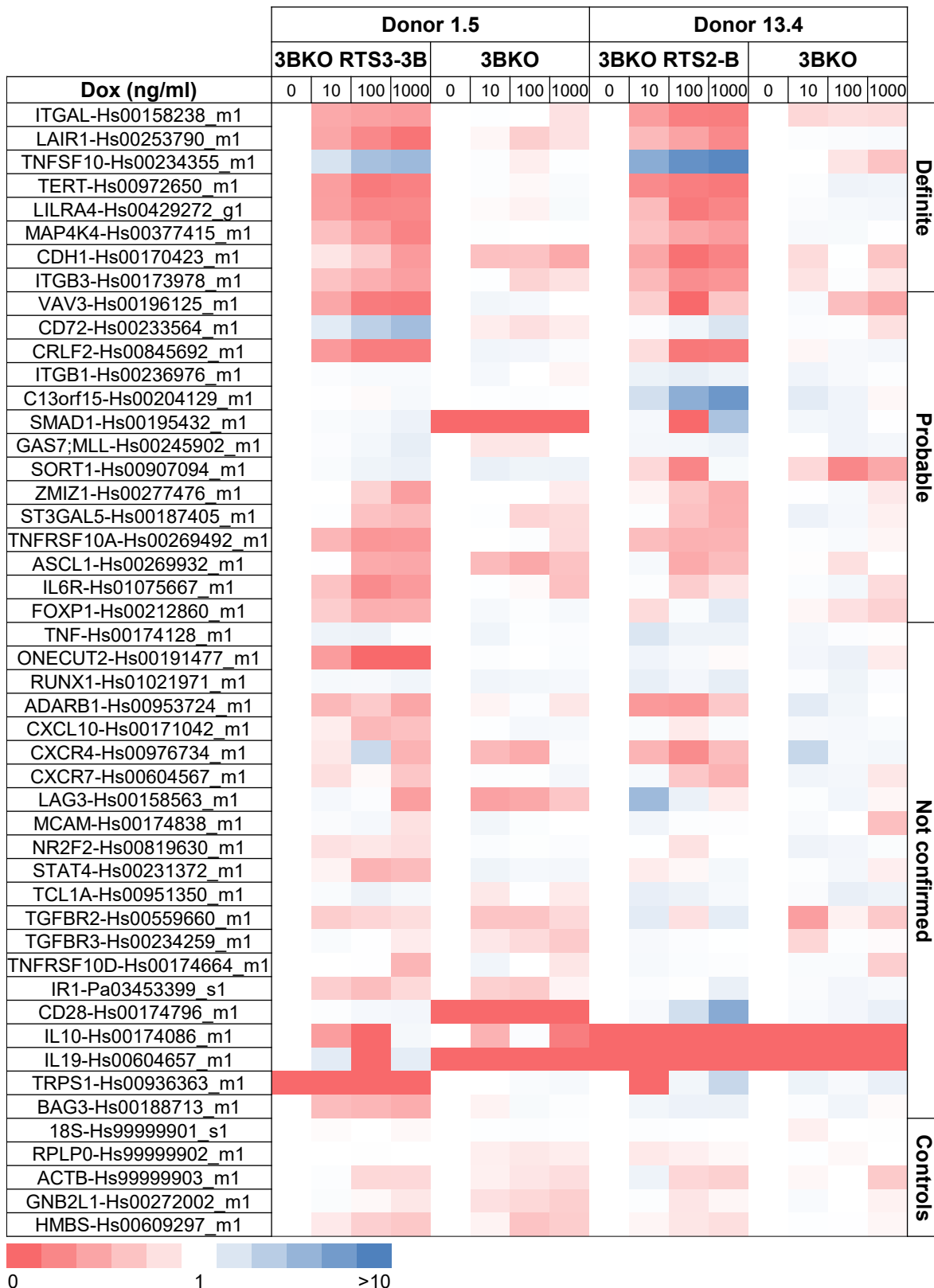


Figure 3.23. Heatmap showing the expression of candidate EBNA3B-regulated genes in D1.5 and D13.4 3BKO RTS-3B LCLs after induction with increasing dosages of Dox. After treatment of D1.5 and D13.4 3BKO and RTS-3B LCLs with increasing dosages of Dox, cDNA from these LCLs were analysed using TLDA cards containing Taqman probes for 48 pre-determined genes. Gene expression levels were normalised against *18S* & *RPLP0* housekeeping genes and reported relative to the untreated sample for each set of 4 doses. *ACTB*, *GNB2L1* & *HMBS* are omitted from the normalizing panel because they appear to exhibit some dose-dependent response to Dox in at least one cell line (See Table 3.4).

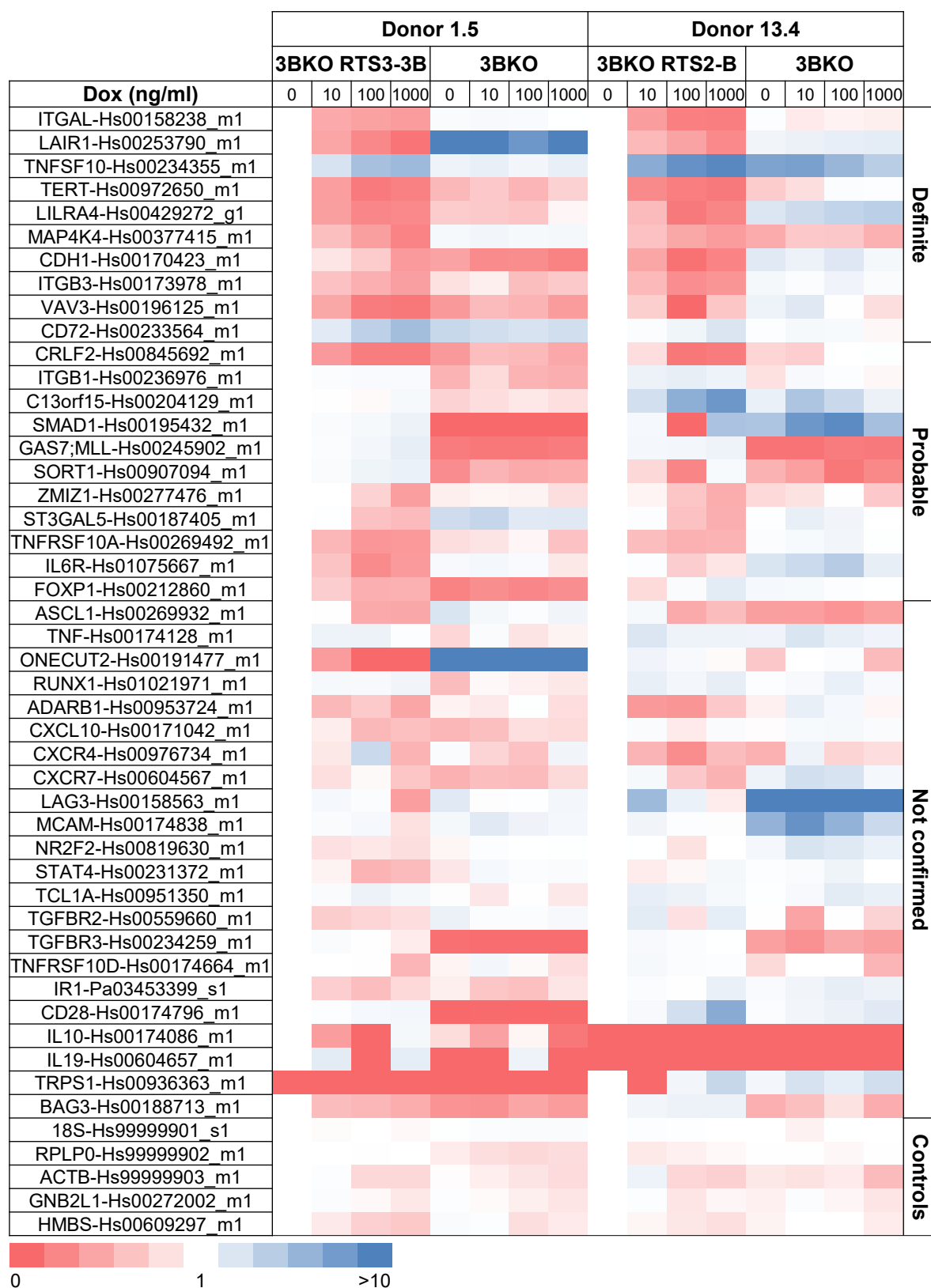


Figure 3.24. Heatmap highlighting any differences in gene expression of candidate EBNA3B-regulated genes between parental D1.5 and D13.4 3BKO LCLs and dox-inducible RTS-3B LCLs.

As in Figure 3.23, but all data from each donor (including 3BKO (pRTS-null) and RTS-3B LCLs) are represented relative to the untreated sample of RTS-3B LCLs to highlight any variation between original and vector-transfected cell lines.

Table 3.3. EBNA3B-repressed genes as determined by the Dox-inducible EBNA3B system.

Gene Symbol	Assay ID	Name	Category	Predicted regulation	Regulation by EBNA3B		Criteria used (refer to Table 3.2)
ITGAL	Hs00158238_m1	integrin, alpha L (antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide)	Cell adhesion molecule	Down	Down	Definite	1a; 1b
LAIR1	Hs00253790_m1	leukocyte-associated immunoglobulin-like receptor 1	Defense/immunity protein	Down	Down	Definite	1a; 1b
TERT	Hs00972650_m1	telomerase reverse transcriptase	Nucleic acid binding	Down	Down	Definite	1a
LILRA4	Hs00429272_g1	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 4	Defense/immunity protein	Down	Down	Definite	1b
MAP4K4	Hs00377415_m1	mitogen-activated protein kinase kinase kinase 4	Kinase	Down	Down	Definite	1b
CDH1	Hs00170423_m1	cadherin 1, type 1, E-cadherin (epithelial)	Cell adhesion molecule	Down	Down	Definite	1b
ITGB3	Hs00173978_m1	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	Extracellular matrix	Down	Down	Definite	1b
VAV3	Hs00196125_m1	vav 3 guanine nucleotide exchange factor	Select regulatory molecule	Down	Down	Probable	2a (D1.5 RTS3-3B LCL)
CRLF2	Hs00845692_m1	cytokine receptor-like factor 2	Molecular function unclassified	Down	Down	Probable	2a (D1.5 RTS3-3B LCL); 2c (D13.4 RTS2-3B LCL)
ZMIZ1	Hs00277476_m1	zinc finger, MIZ-type containing 1	Ligase	Down	Down	Probable	2c
ST3GAL5	Hs00187405_m1	ST3 beta-galactoside alpha-2,3-sialyltransferase 5	Transferase	Down	Down	Probable	2c
TNFRSF10A	Hs00269492_m1	tumor necrosis factor receptor superfamily, member 10a	Receptor	Down	Down	Probable	2c (both donors)
ASCL1	Hs00269932_m1	achaete-scute complex homolog 1 (Drosophila)	Transcription factor	Down	Down	Probable	2c (both donors)
IL6R	Hs01075667_m1	interleukin 6 receptor	Receptor	Down	Down	Probable	2c (D1.5 RTS3-3B LCL only)
FOXP1	Hs00212860_m1	forkhead box P1	Transcription factor	Down	Down	Probable	2c (D1.5 RTS3-3B LCL only)
ONECUT2	Hs00191477_m1	one cut homeobox 2	Transcription factor	Down	Not confirmed		3a
ADARB1	Hs00953724_m1	adenosine deaminase, RNA-specific, B1 (RED1 homolog rat)	Hydrolase	Down	Not confirmed		3a
CXCR4	Hs00976734_m1	chemokine (C-X-C motif) receptor 4	Receptor	Down	Not confirmed		3a
CXCR7	Hs00604567_m1	chemokine (C-X-C motif) receptor 7	Receptor	Down	Not confirmed		3a
MCAM	Hs00174838_m1	melanoma cell adhesion molecule	Cell adhesion molecule	Down	Not confirmed		3a
NR2F2	Hs00819630_m1	nuclear receptor subfamily 2, group F, member 2	Transcription factor	Down	Not confirmed		3a
STAT4	Hs00231372_m1	signal transducer and activator of transcription 4	Transcription factor	Down	Not confirmed		3a
TGFBR2	Hs00559660_m1	transforming growth factor, beta receptor II (70/80kDa)	Receptor	Down	Not confirmed		3a
TGFBR3	Hs00234259_m1	transforming growth factor, beta receptor III	Receptor	Down	Not confirmed		3a; 2b
TNFRSF10D	Hs00174664_m1	tumor necrosis factor receptor superfamily, member 10d, decoy with truncated death domain	Receptor	Down	Not confirmed		3a
BAG3	Hs00188713_m1	BCL2-associated athanogene 3	Chaperone	Down	Not confirmed		3d

Table 3.4. EBNA3B-induced genes as determined by Dox-inducible EBNA3B system.

Gene Symbol	Assay ID	Name	Category	Predicted regulation	Regulation by EBNA3B		Criteria used (refer to Table 3.2)
TNFSF10	Hs00234355_m1	tumor necrosis factor (ligand) superfamily, member 10	Signaling molecule	Up	Up	Definite	1c; 1d
CD72	Hs00233564_m1	CD72 molecule	Molecular function unclassified	Up	Up	Probable	2a (D1.5 RTS3-3B LCL)
ITGB1	Hs00236976_m1	integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	Receptor	Up	Up	Probable	2a (D1.5 RTS-3B :LCL) - slight dose response
C13orf15	Hs00204129_m1	chromosome 13 open reading frame 15	Select regulatory molecule	Up	Up	Probable	2a (D13.4 RTS2-3B LCL)
SMAD1	Hs00195432_m1	SMAD family member 1	Transcription factor	Up	Up	Probable	2b
GAS7;MLL	Hs00245902_m1	growth arrest-specific 7;myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	Transcription factor	Up	Up	Probable	2b (despite 1b)
SORT1	Hs00907094_m1	sortilin 1	Receptor	Up	Up	Probable	2b (despite 1b)
TNF	Hs00174128_m1	tumor necrosis factor (TNF superfamily, member 2)	Signaling molecule	Up	Not confirmed		3a
RUNX1	Hs01021971_m1	runt-related transcription factor 1	Transcription factor	Up	Not confirmed		3a
CXCL10	Hs00171042_m1	chemokine (C-X-C motif) ligand 10	Signaling molecule	Up	Not confirmed		3a
LAG3	Hs00158563_m1	lymphocyte-activation gene 3	Defense/immunity protein	Up	Not confirmed		3a
TCL1A	Hs00951350_m1	T-cell leukemia/lymphoma 1A	Select regulatory molecule	Up	Not confirmed		3a
IR1	Pa03453399_s1	Internal Repeat (EBV)	Molecular function unclassified	-	Not confirmed		3a
CD28	Hs00174796_m1	CD28 molecule	Defense/immunity protein	Up	Not confirmed		3b (D13.4 3BKO LCL)
IL10	Hs00174086_m1	interleukin 10	Signaling molecule	Up	Not confirmed		3c
IL19	Hs00604657_m1	interleukin 19	Signaling molecule	Up	Not confirmed		3c
TRPS1	Hs00936363_m1	trichorhinophalangeal syndrome I	Nucleic acid binding	Up	Not confirmed		3c
18S	Hs99999901_s1	Eukaryotic 18S rRNA	Molecular function unclassified	Control	Control		-
RPLP0	Hs99999902_m1	ribosomal protein, large, P0	Nucleic acid binding	Control	Control		-
ACTB	Hs99999903_m1	actin, beta	Cytoskeletal protein	Control	Excluded		3b (D1.5 3BKO LCL)
GNB2L1	Hs00272002_m1	guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1	Select regulatory molecule	Control	Excluded		3b (D1.5 3BKO LCL)
HMBS	Hs00609297_m1	hydroxymethylbilane synthase	Hydrolase	Control	Excluded		1b

4 Determination of Genome-wide EBNA3B binding

4.1 Introduction

ChIP experiments have been pivotal in contributing to our understanding of the function of EBNA3A and EBNA3C by detailing their binding regions relative to transcription start sites of putative regulated genes (Bazot *et al.*, 2015; Harth-Hertle *et al.*, 2013; McClellan *et al.*, 2012, 2013; Paschos *et al.*, 2012; Skalska *et al.*, 2013; Zhou *et al.*, 2015). Briefly, proteins bound to DNA are fixed using paraformaldehyde and sheared into small pieces (100bp-1kb) via sonication. Antibodies specific against the protein of interest or chromatin modification are traditionally used to immunoprecipitate these protein-bound short lengths of DNA which are then purified for interrogation. Specific sites can then be investigated directly using site-specific primers and quantitative PCR, or the ChIP material can be pooled for further analyses using high-throughput sequencing (ChIP-seq) (see Introduction and Chapter 2 for more details). Genome-wide comparisons of these ChIP-seq data with those of published transcription factor or histone modification sites have informed us of the putative interactions of EBNA3A and EBNA3C with transcription regulators and chromatin landscapes (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015). This breakthrough has allowed us to make deductions on the function of EBNA3A and EBNA3C, and paved the way for further investigations into their mechanisms of regulation for specific target genes (Bazot *et al.*, 2015; Jiang *et al.*, 2014; Kalchschmidt *et al.*, 2016a, 2016b; McClellan *et al.*, 2013; Schmidt *et al.*, 2015).

In order to further understand the contribution of EBNA3B to host cell gene regulation, I have cloned recombinant epitope-tagged EBNA3B (FS3B) EBV for use in a genome-wide interrogation of EBNA3B localization in LCLs. This chapter describes the cloning process, the preparation of the ChIP-seq samples and the eventual analyses of the results of high throughput sequencing – culminating in a list of potential EBNA3B-regulated genes which are not only associated with either proximal or distal EBNA3B binding but are also genes

which are induced or repressed in response to increasing amounts of EBNA3B in an EBNA3B-null LCL background (Chapter 3).

4.2 Construction of recombinant epitope-tagged EBNA3B (FS3B) BAC

In our hands, the commercially available antibody for EBNA3B (F120P, Exalpha biologicals) has been shown to cross-react with EBNA3A and EBNA3C in immunoprecipitation experiments (unpublished data; Watanatanasup, 2011), which would cause ambiguity in the interpretation of any ChIP data. Monoclonal rat antibodies raised against EBNA3B (obtained from Dr Elisabeth Kremmer; White *et al*, 2010) were able to immunoprecipitate EBNA3B without any cross-reactivity with EBNA3A or EBNA3C, however, these antibodies performed poorly in ChIP experiments, resulting in high background signal and poor differentiation between positive and negative controls (Appendix B).

To overcome the problem of cross-reactivity, EBNA3A and EBNA3C fused to a tandem affinity tag (Strep II-Strep II-FLAG) at the C-terminus have been shown to effectively isolate these two proteins in ChIP experiments using commercially available anti-FLAG antibody. This method can efficiently differentiate between the EBNA3s in pull-down experiments and avoids the issue of possible cross-reactivity of EBNA3 antibodies with other members of the EBNA3s. As highlighted in chapter 3, the tandem arrangement of EBNA3B and EBNA3C genes have made it difficult to retain stable expression of EBNA3C in C-terminally tagged EBNA3B cell lines. So, in order to generate a similar system which can be used to examine the role of EBNA3B in the transcriptional regulation of its target genes, I have inserted the tandem affinity tag (FLAG-Strep II-Strep II) (Gloeckner *et al.*, 2007; Schmidt and Skerra, 2007) at the N-terminus of EBNA3B to generate FS3B.

The cloning strategy of this N-terminus fusion is the same as for ERT2, and was done in parallel with the cloning of the ERT-2 tag (described in Chapter 3), with the only difference being that the epitope tag (FS) replaces the ERT-2 tag at the N-terminus of EBNA3B. Generation of the co-integrand and the final recombinant BAC from B95.8 BAC

was as described in Chapter 3, section 3.3. Interestingly, in contrast to the insertion of ERT-2 tag, where the integration of the KovKan vector seemed to be kinetically favoured at the start of EBNA3B, the insertion of the KovKan vector containing the smaller FS tag seemed to favour co-integration at the end of EBNA3A, since only co-integrant BACs in this form were recovered (Figure 4.1).

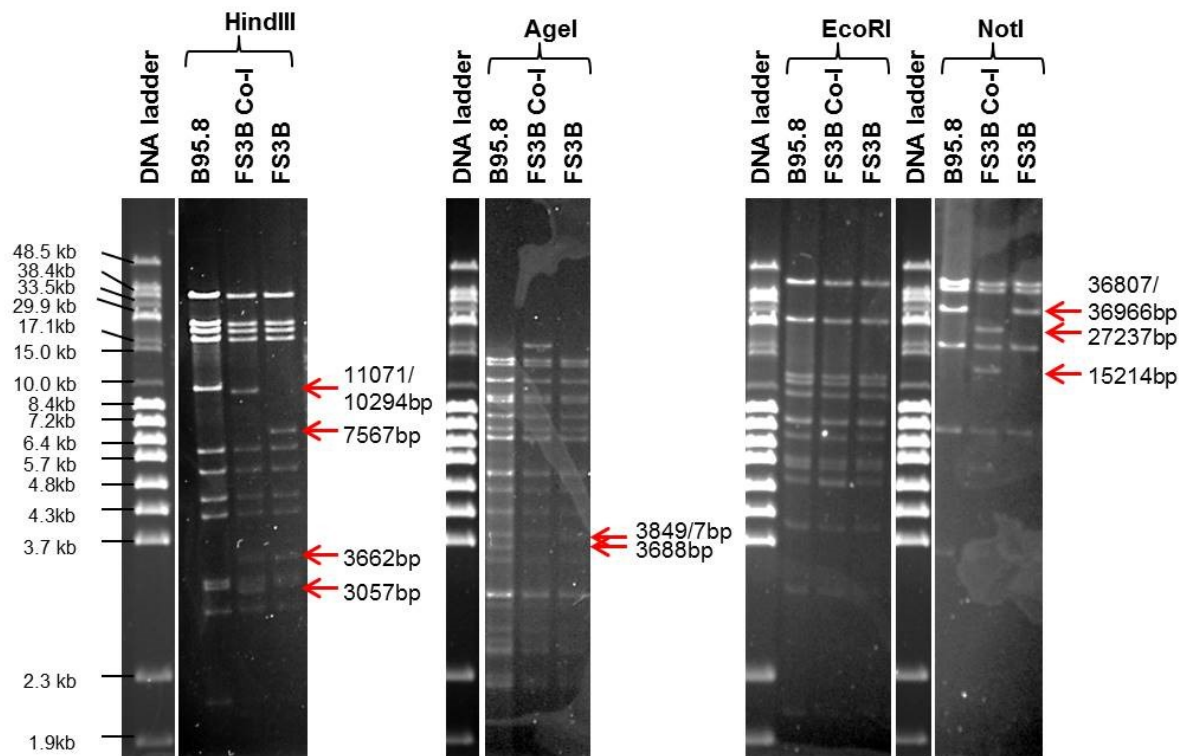


Figure 4.1. Generation of FS3B BAC from wild-type B95.8.

Pulse-field gel electrophoresis (PFGE) following diagnostic digests are used to identify the BACs (HindIII) and validate the integrity of the BAC backbone (NotI, AgeI and EcoRI). Insertion of the KovKan shuttle vector should add in 3 HindIII sites (two within KovKan backbone and one within the end linker of FS3B sequence) which should cause the 11071bp fragment in B95.8 BAC to be digested into 3 fragments (10294bp, 3662bp and 3057bp) in the co-integrant (Co-I) and 2 fragments (7567bp and 3662bp) in the final FS3B BAC. Likewise, addition of FS and/or KovKan sequence should not introduce any additional AgeI sites, and should merely increase the size of the 3688bp band to 10614bp and 3847bp in the Co-I and resolved FS3B BAC respectively. The band size changes for AgeI digests were not as expected for the Co-I but it went on to resolve to the correct FS3B. EcoRI digestion of the BACs should not result in any gross rearrangements of the band size distributions and changes should not be detectable on PFGE. Introduction of the KovKan vector into the FS3B Co-I introduces 2 NotI sites which should cleave the 36807 fragment into 27237bp, 15214bp and 1282bp. The resolved FS3B BAC should revert to 36966bp due to the additional FLAG-strep-strep (FS) sequence that is indistinguishable from the 36807 band in B95.8 by PFGE. The DNA ladder used is a combination of λ DNA-BstEII digest and λ DNA-Monocut mix from New England Biolabs.

4.3 Transfection of HEK-293 cells and generation of FS3B virus producer cell lines

Transfection of the recombinant BAC into HEK-293 cells and subsequent episome rescues from transfected cells were successful (Figure 4.2A), and recombinant FS3B EBV were produced and used to infect PBLs. At this stage, transfected HEK-293 producers were tested for expression of recombinant FS3B protein by western blot and the feasibility of immunoprecipitating (IP) EBNA3B using the anti-FLAG antibody for the pull-down. Both the immunoblot and IP experiments showed that the recombinant FS3B protein was expressed and could be efficiently immunoprecipitated by the anti-FLAG antibody (Figure 4.2B&C). Encouraged by these findings, these HEK-293 cells were used to produce viruses for infection of PBLs, but again, the virus titres were extremely low (Table 4.1).

Table 4.1. Table showing virus titres of FS3B HEK-293 cell lines as determined by Raji cell infections. Green Raji Units (GRU) are determined by counting the number of Raji cells infected per ul of virus used.

HEK-293 Cell line	Virus Titres (GRU)
FS3B C16	2.6
FS3B C8	1.69
FS3B E1	22.2

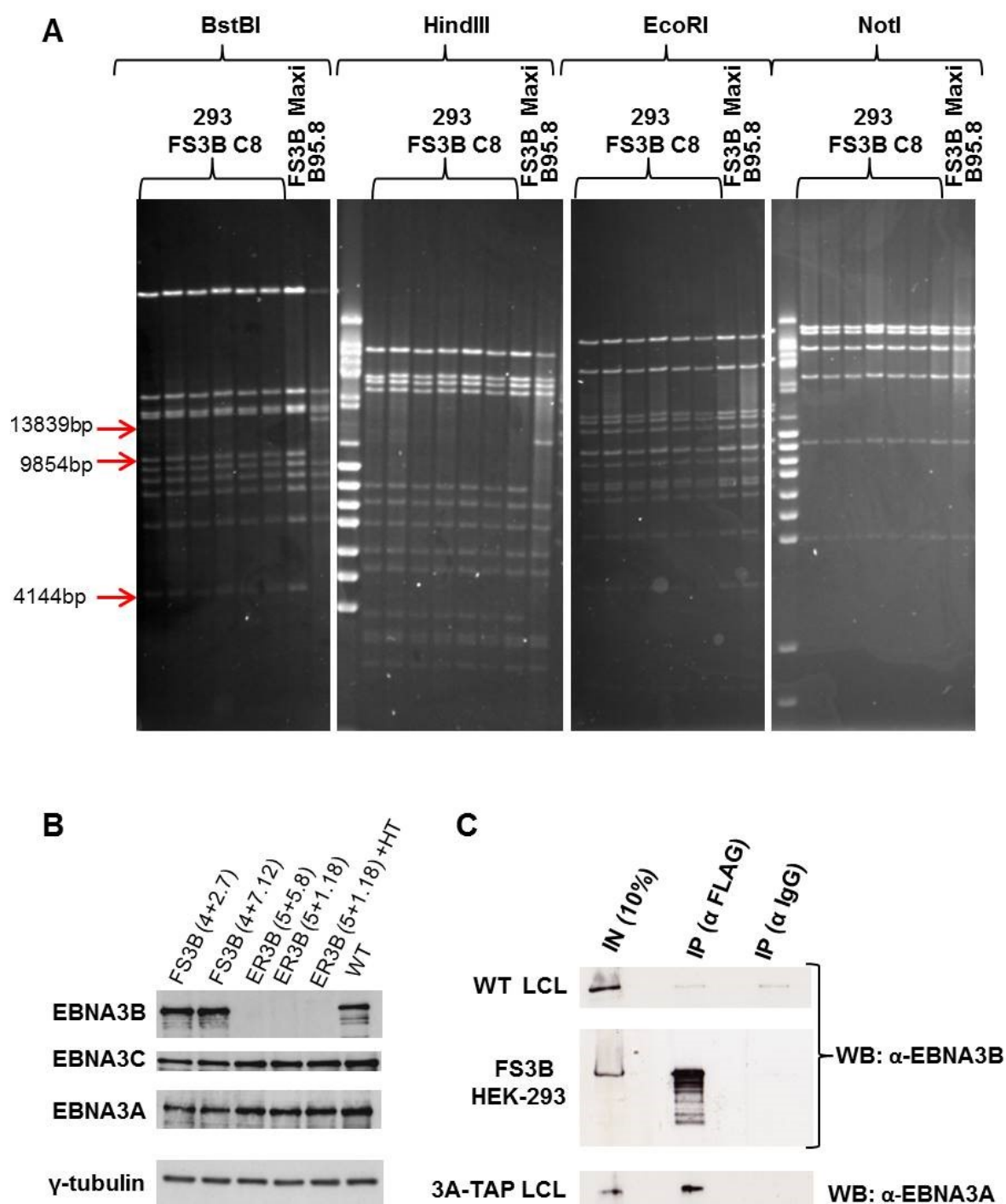


Figure 4.2. Validation of HEK-293 FS3B cell lines.

(A) *HindIII* and *BstBI* restriction digests are used to identify the BACs whereas *EcoRI* and *NotI* digests are used to validate the BAC integrity in HEK-293 cells. Fusion of FS sequence to EBNA3B introduces an additional *BstBI* site, which will result in two smaller fragments (9854 & 4144bp) instead of a large 13839bp fragment after digestion. Maxi-prepped FS3B and B95.5 BACs are used as comparisons. (B) Western blot showing the expression of EBNA3B, EBNA3A and EBNA3C in FS3B, ER3B (grown out in +/-HT) and WT LCLs. γ -tubulin is used as a loading control for the western blot. (C) Immunoprecipitation (IP) of EBNA3B using antibodies specific for FLAG in the FS3B HEK-293 cell lines. Western Blot using antibodies against EBNA3B or EBNA3A were used to confirm the pull-down. WT and 3A-epitope tagged (3A-TAP) cell lines are used as positive and negative controls respectively.

4.4 Validation of LCLs established using recombinant FS3B virus.

4.4.1 Episome rescued from FS3B LCLs showed correct integrity.

Three FS3B LCLs derived from 2 different donors (D11 and D13) were successfully established 2 months post infection. These are named D11 C8, D13 E1 and D13 C16 FS3B LCLs. The integrity of the episomal BAC DNA recovered from these LCLs were analysed by restriction digestion using HindIII, BstBI, AgeI, EcoRI & NotI restriction enzymes, followed by PFGE of the restricted fragments (Figure 4.3A). BACs rescued from these cell lines had mainly alterations in the BamW or the terminal repeat regions. However, since (with the exception of EBNA3B) the expression of several EBV latent genes (EBNA3s, EBNA2 & LMP1) seemed normal (Figure 4.3B), it was assumed that slight variations in these repeat regions would not adversely affect the function of the recombinant EBV in these LCLs.

4.4.2 FS3B protein is present at slightly higher levels than wild-type EBNA3B in established FS3B LCLs and BL31s.

In contrast to the ER3B LCLs, these FS3B LCLs expressed slightly higher than wild-type levels of the recombinant FS3B fusion protein, and the epitope tag is also functional in IP experiments (Figure 4.3C). Interestingly, the epitope-tagged EBNA3B seemed to be either expressed at higher levels, or was stabilised in the cells, since higher levels of EBNA3B were detected relative to wild-type levels from B95.8 BAC in similar LCLs (Figure 4.3B). In addition, the FS3B recombinant protein did not quite run at an expected size – that is slower than wild-type EBNA3B – on the SDS-PAGE gel, and it is possible that presence of the FS tag may have altered the protein electrophoretic mobility in some unknown way.

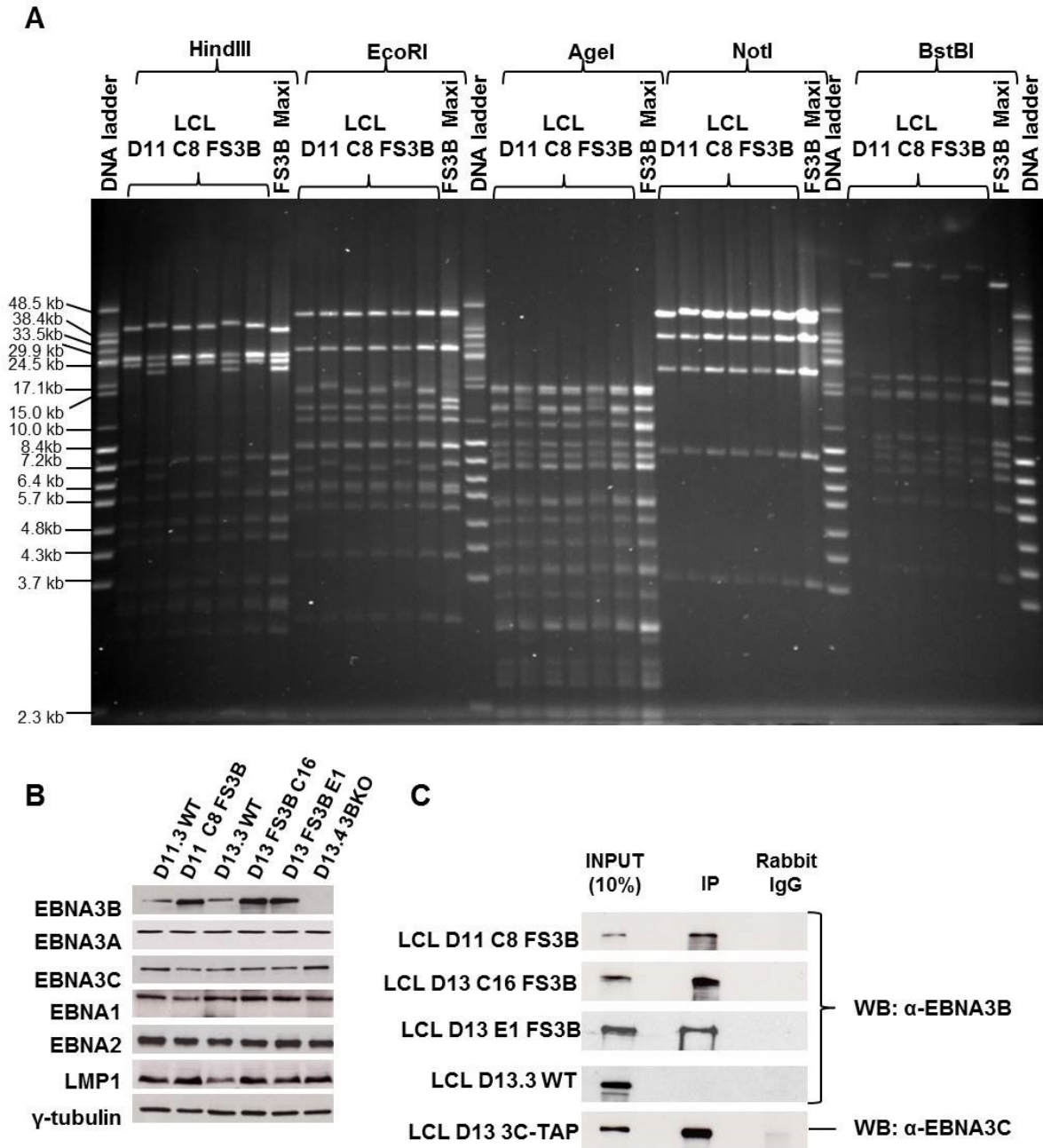


Figure 4.3. Validation of FS3B LCLs.

(A) Restriction digestion digests followed by PFGE were used to identify and validate FS3B episomes rescued from the FS3B LCL. Only gel runs of D11 C8 FS3B are shown. Maxi-prepped BACs used for transfection into HEK-293 cells were used as positive comparisons. Alterations were detected in the BamW or terminal repeat regions, with the FS3B regions remaining intact. (B) Western blot showing expression of EBNA3B and other latent proteins in WT, 3BKO and FS3B LCLs. γ -tubulin is used as a loading control. (C) Immunoprecipitation using anti-FLAG antibody followed by detection of EBNA3B with western blots in 3 FS3B cell lines. WT and C-terminal epitope-tagged EBNA3C are used as negative and positive controls respectively for the immunoprecipitation.

4.4.3 Low molecular weight DNA isolated from FS3B LCLs were verified by sequencing

Since the FS3B proteins did not show a size difference relative to non-tagged in the SDS-PAGE gel, episomal BAC DNA were rescued from the D11 C8 and D13 C16 LCLs and the entire *FS3B* domain was amplified, sequenced and verified (Figure 4.4). No pre-mature stop codon was found in these sequences, so it seems unlikely that the expressed FS3B is truncated.

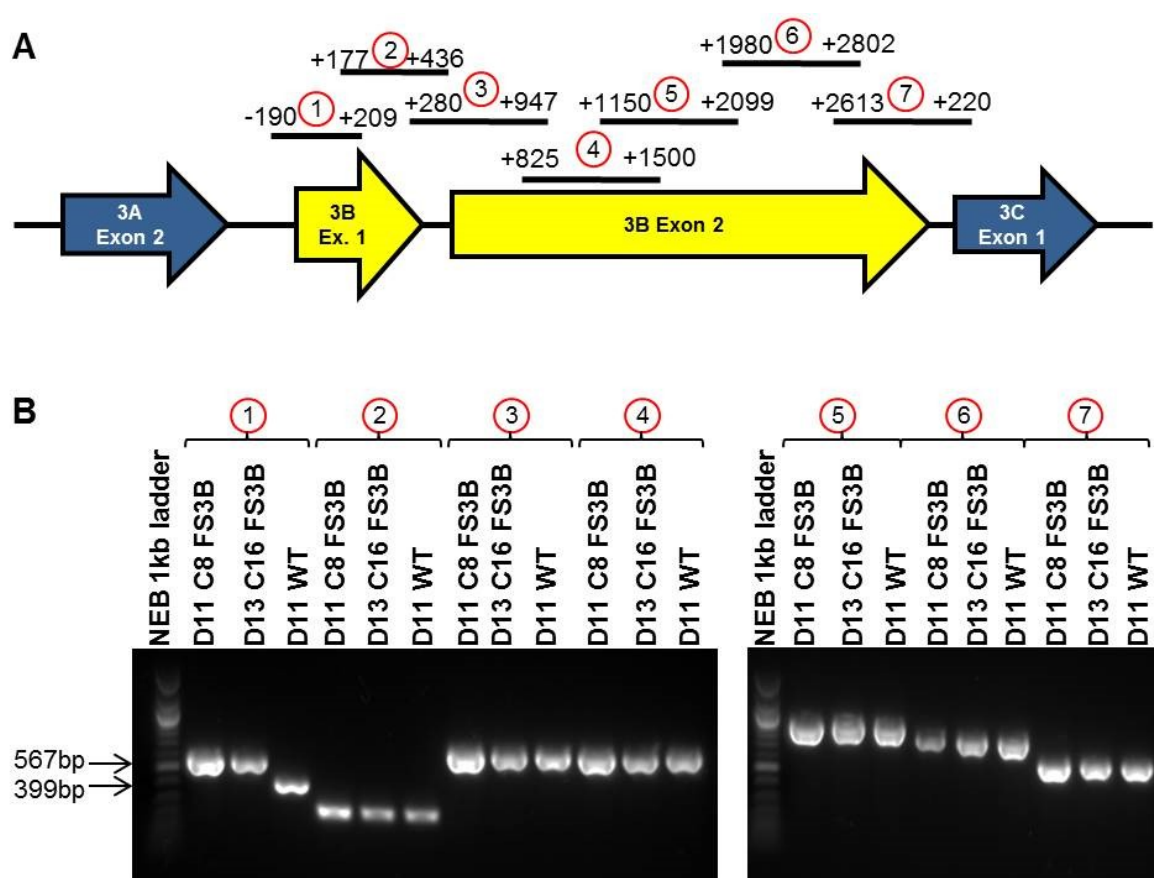


Figure 4.4. PCR amplification of the FS3B domain of low molecular weight DNA purified from FS3B LCLs.

(A) Schematic showing the position of primer pairs (relative to translational start site) used to amplify across the *EBNA3B* gene. Each pair is labelled with a number enclosed in a red circle. (B) The FS3B domains of D11 C8 FS3B and D13 C16 FS3B LCLs were PCR-amplified and products separated with agarose gel electrophoresis. PCR amplification using primer set 1 would result in band sizes of 567bp and 300bp for FS3B and WT LCLs respectively. The amplified DNA were sent for verification through sequencing.

4.4.4 Expression levels of several potential EBNA3B targets are closer to wild-type levels than those of 3BKO LCLs.

In order to ensure that the modification at the N-terminus did not affect the function of FS3B in regulating EBNA3B target genes, two potential EBNA3B-induced (*CD28* & *SMAD1*) and one candidate repressed gene (*ITGAL*) were picked for gene expression analysis using RT-qPCR. Comparing reverse-transcribed cDNA from these LCLs to that of WT and 3BKO LCLs, expression of all three genes in the FS3B LCLs were similar to their respective WT LCLs derived from the same donor, although the expression of *CD28* in D11 WT LCLs was already lower than that of D13 3BKO LCLs (Figure 4.5). In another run of the qPCR analysis (results not shown), expression of *CD28* in D11 WT LCL was verified to be higher than that of D11 3BKO LCL.

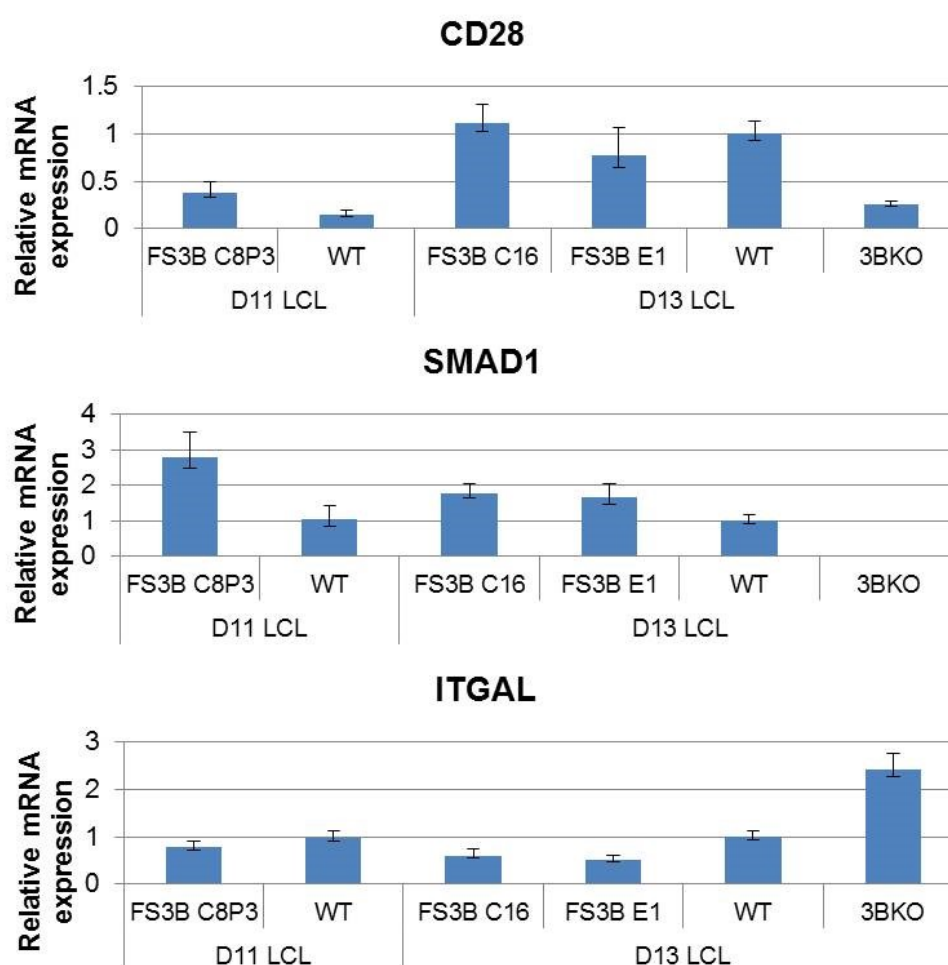


Figure 4.5. Expression of CD28, SMAD1 & ITGAL in FS3B LCLs.

Expression of *CD28*, *SMAD1* and *ITGAL* in FS3B, WT and 3BKO LCLs derived from the same donor PBLs (D11 and D13). The data are normalised to *GNB2L1* and *RPLP0* housekeeping genes and are reported relative to gene expression level of D13 WT LCLs. Error bars represent technical triplicates from one RT-qPCR experiment.

4.4.5 FS3B LCLs are validated for chromatin immunoprecipitation using ITGAL as a positive EBNA3B binding site

Next, western blot analyses of ChIP lysates showed that EBNA3B can be immunoprecipitated efficiently with the FLAG antibody in FS3B LCLs, whereas no EBNA3B was detected in WT lysates (Figure 4.6). In their recent paper, McClellan *et al.* (2013) have shown (using an EBNA3B-specific antibody) that EBNA3B binds at the promoter and enhancer of *ITGAL* and *CtBP2* respectively. Using the FLAG antibody and primer sequences described in this publication, I was able to replicate their results in ChIP-qPCR experiments with 2 FS3B LCLs (Figure 4.6). Encouraged by these promising results, I conducted more ChIP experiments with the D11 C8 FS3B LCL and pooled the samples to generate enough DNA for high-throughput sequencing.

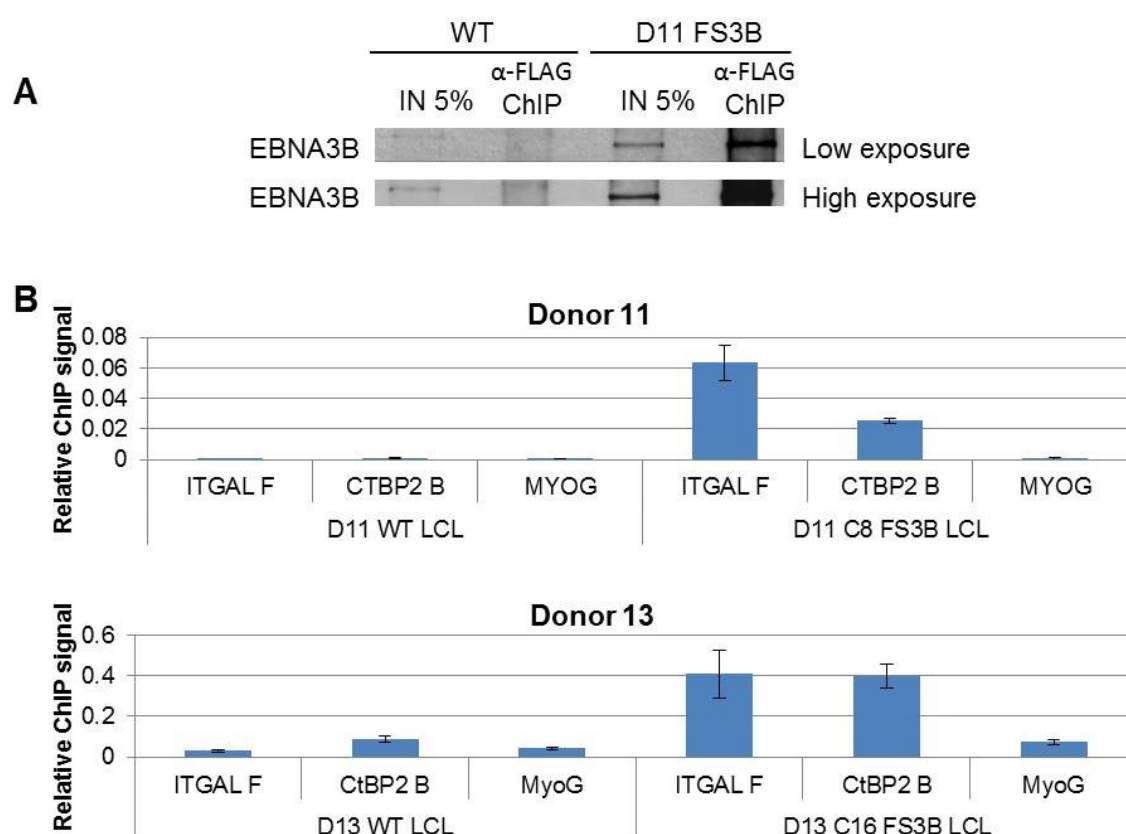


Figure 4.6. Validation of FS3B LCLs for use in ChIP experiments

(A) Western blot using rat monoclonal antibodies (6C9) to detect EBNA3B following immunoprecipitation of using α -FLAG antibodies in D11 C8 FS3B LCL. This is compared to 5% of the input lysate for the experiment. The untagged D11 WT LCL is used as control cell line. **(B)** ChIP-qPCR using α -FLAG antibodies in D11 and D13 FS3B and WT LCLs. ChIP signal is reported as enrichment over 0.5% of input DNA. Primers for the promoter/enhancer regions of *ITGAL* and *CtBP2* were obtained from McClellan *et al.* (2013). ITGALF and CtBP2B primer sets amplify the regions with the highest EBNA3B peaks in the study conducted by McClellan *et al.* (2013). The *MyoG* promoter is used as a negative control region. Results are representative of three independently done experiments and error bars represent technical triplicates from one ChIP-qPCR experiment.

4.5 ChIP coupled with high-throughput sequencing using FS3B LCLs

Briefly, material from 8 independent ChIP experiments were pooled and size-selected for fragments shorter than 500bp (Figure 4.7). As a control for any biases in chromatin fragmentation and sequencing efficiency, DNA that was cross-linked and sheared under the same conditions (“Input” DNA) as the immuno-precipitated DNA was also pooled from the same ChIP experiments and underwent the similar preparation for sequencing. The size-selected DNA were gel-purified and run on the Agilent 2100 Bioanalyzer to ensure that the samples are free of contaminants (Figure 4.8). These were sent off to the Biopolymers Facility (BPF) at Harvard Medical School for library preparation and sequencing (explained in detail in Chapter 2). Next-Generation sequencing was performed using the Illumina HiSeq2500 machine to obtain single-end³ read lengths of 50bp and the sequences were mapped to *Homo sapiens* hg19 genome and analysed using Partek Genomics Suite (PGS) version 6.6 software (further details in Chapter 2).

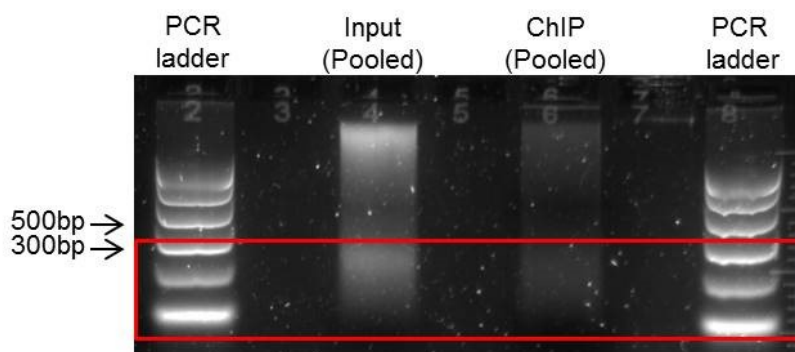


Figure 4.7. Agarose gel electrophoresis of pooled input and ChIP samples from D11 C8 FS3B LCLs.

Input and ChIP samples were pooled from 8 ChIP experiments, separated on an agarose gel by electrophoresis and visualised using the SYBR Gold (Invitrogen) stain under UV illumination. Fragments that were less than 500bp (boxed) were size selected and gel-purified.

³ DNA is sequenced from only one end of the prepared library fragments.

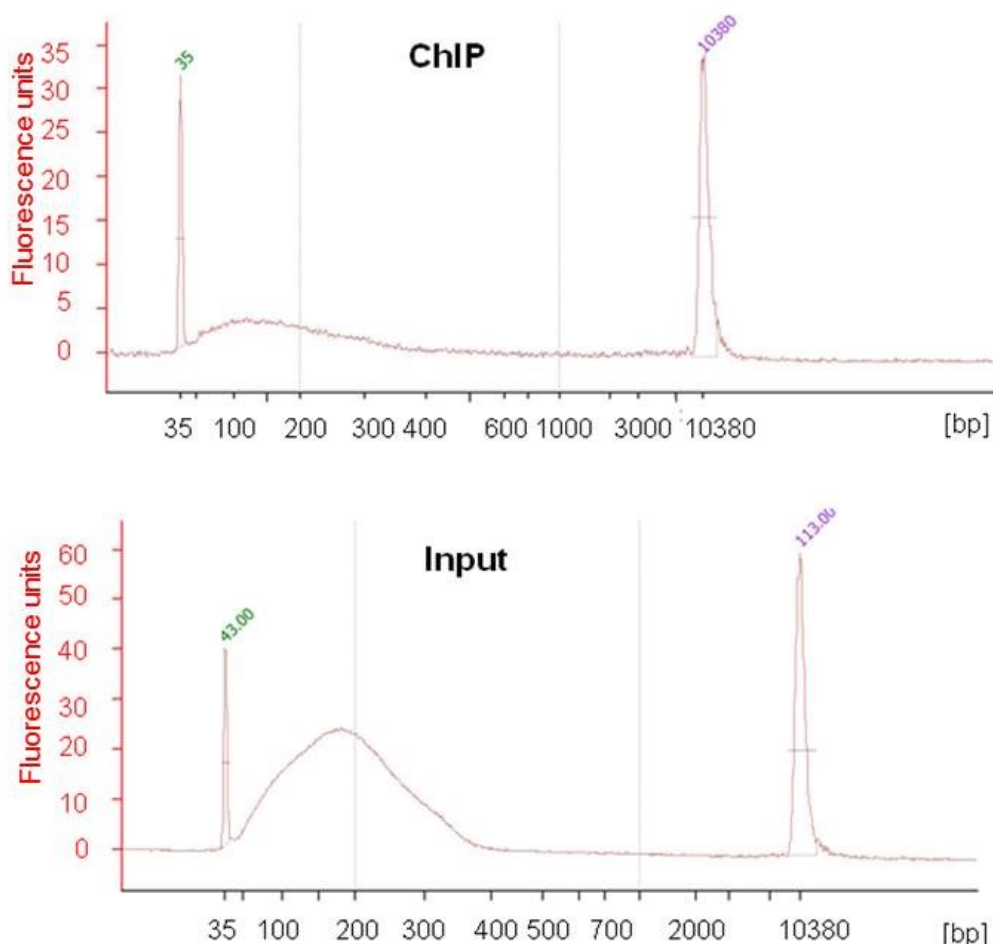


Figure 4.8. Bioanalyzer analysis (Agilent 2100 Bioanalyzer) of size-selected ChIP and Input samples from D11 C8 FS3B LCL.

Electropherogram showing the size distribution of DNA samples sent for next-generation sequencing. Peaks of the upper and lower markers are detected automatically by the machine and any contamination with salts or RNA will be represented by noise spikes.

In another independent ChIP experiment using FLAG antibodies performed by Dr Kostas Paschos, the same procedure was performed on WT (untagged EBNA3B) LCLs derived from Donor 11 as a comparative control for the antibody and background pull-down. Comparisons between the FLAG-tagged and untagged cell lines may be made using two different methods: peak calling for each of FS3B and WT ChIP may be done individually using their respective input as controls for the sonication process, or it can be done once by directly comparing the enrichment of FS3B ChIP sample over the WT ChIP sample, hence using the WT ChIP as background pull-down control. In the former method, peaks will be determined by the enrichment of FS3B or WT ChIP samples over their respective input DNA,

after which peaks that show up in the WT ChIP sample may be filtered away from the FS3B ChIP peaks in further analyses. In the latter method, EBNA3B binding peaks will be determined by the enrichment of FS3B ChIP sample over the ChIP sample of the untagged cell line, which is directly computed during the peak calling process and eliminates the use of either of their inputs.

Although using inputs as the control for peak calling may provide greater and more evenly distributed coverage of the genome, the Encyclopedia of DNA Elements (ENCODE) consortium recommends using material from similarly performed ChIP experiments in untagged, but otherwise identical cell lines as the control sample (Landt *et al.*, 2012). This is to remove any background pull-down of DNA or DNA-bound proteins by the FLAG antibody.

4.5.1 Quality control of sequencing data as defined by the ENCODE guidelines.

In 2012, the ENCODE consortium released a set of ChIP-seq guidelines which can be used to determine the adequacy of ChIP-seq data (Landt *et al.*, 2012). This is used as a reference in determining the adequacy of the sequencing data before commencing any analysis.

4.5.1.1 Sequencing depth and complexity of sequencing library

I began by verifying the adequacy of the sequencing depth and the complexity of the ChIP material. Overall, the majority of the sequencing reads (85-98%) obtained from input DNA and anti-FLAG ChIP samples of epitope-tagged FS3B LCLs and untagged WT LCLs were aligned to hg19. In all samples, the minimum recommended sequencing depth was obtained (Table 4.2) – for one biological replicate, more than 10 million reads are aligned exactly once to the human genome – as defined by the ENCODE guidelines (Landt *et al.*, 2012). While the number of reads was adequately high for each sample, a low library complexity (i.e. number of unique fragments present and sequenced in the prepared library) could mean that a majority of these reads are merely duplicates of the same PCR-amplified DNA fragment. As a measure of the library complexity, the ratio of the number of unique

start positions of aligned reads over total number of aligned reads [non-redundant fraction (NRF)] was calculated and shown to be more than the recommended ratio of 0.8 (Landt *et al.*, 2012) for all samples (Table 4.2). This implies that less than 20% of the aligned sequenced reads are duplicates arising from repeated sequencing of the same PCR-amplified fragments and indicates an acceptably diverse library for ChIP-seq analyses.

Table 4.2. Distribution of reads obtained from high-throughput sequencing of ChIP samples.

Sample	Unaligned	Single-end aligned reads: Number of reads that align exactly once to the genome	Number of unique positions that aligned reads are mapped to	Non-redundant fraction
FS3B D11 C8 LCL ChIP	3,820,040	54,231,981	44,272,259	0.82
WT D11 LCL ChIP	6,468,213	34,000,771	30097446	0.89
FS3B D11 C8 LCL Input	1,579,513	12,009,039	11,763,620	0.98
WT D11 LCL Input	619,218	17,348,816	16826206	0.97

4.5.1.2 Strand cross-correlation

Before performing any further analyses, I determined the quality of enrichment and the average size of the fragments by calculating the strand cross-correlation for the two samples (Figure 4.9). This is a quality metric that assesses the average fragment size and the signal to noise ratio of the ChIP-seq experiment. In a good quality ChIP experiment, there will be an increased density of aligned sequencing reads on the forward and reverse strands on either ends of the protein-binding site. The density of reads is computed separately at each position on the forward and reverse strands, and these are shifted incrementally towards or away from each other. The Pearson correlation co-efficient between the forward and reverse strand is calculated for each shift and the maximum co-efficient peak is detected at a shift size that corresponds to the predominant fragment length after sonication (Figure 4.9A) (Landt *et al.*, 2012). Hence, by plotting the correlation co-efficient against shift size for all of the reads, we'd expect to find in a ChIP experiment with more binding sites that there will be a pronounced correlation co-efficient peak at the

predominant fragment length. In the control (untagged) experiment, we'd expect to see a much smaller peak (if any) due to the low density of reads that should theoretically spread evenly across the genome. A peak corresponding to the read length (50bp) is also usually seen. By plotting the correlation co-efficient against shift size, the predominant fragment size is found to be 133bp and 144bp for the FS3B and untagged WT ChIP samples respectively, which fits to the sonication profile (Figure 4.9B).

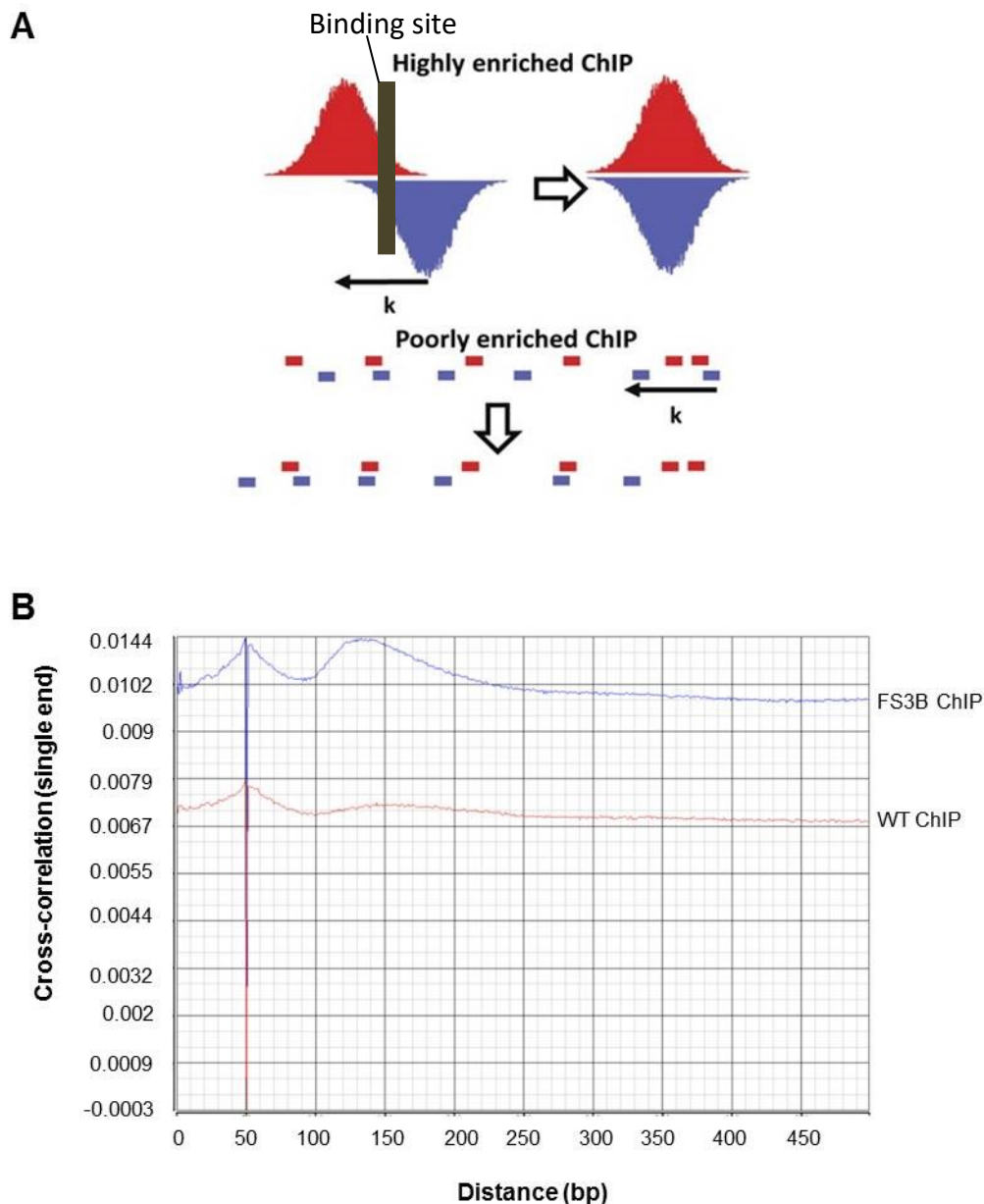
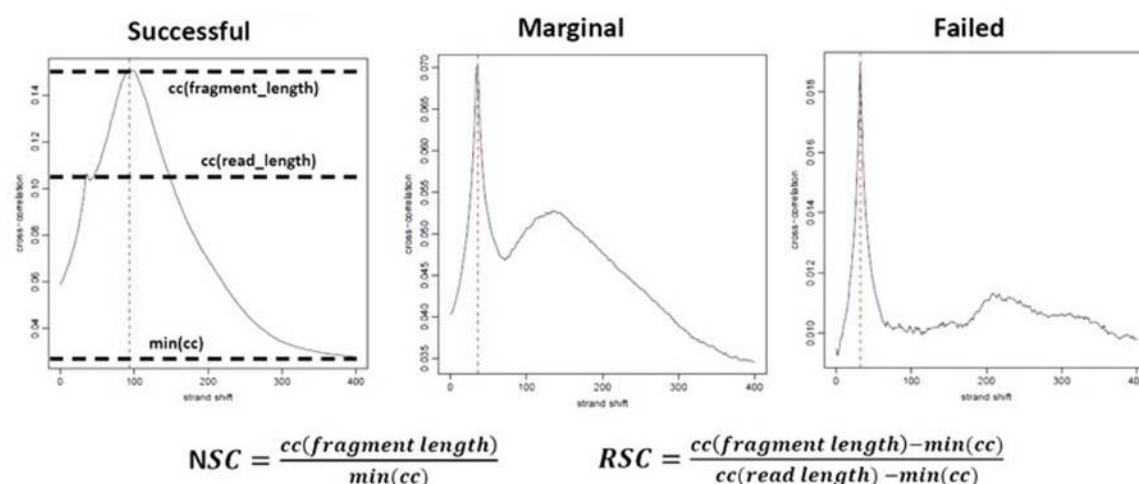


Figure 4.9. Cross-correlation of ChIP samples taken from D11 C8 FS3B and WT LCLs.

(A) Schematic showing the generation of cross-correlation profile. Ideally, read density should be the highest at each end of a ChIP-enriched DNA fragment. These reads are shifted downstream of the strand that they are mapped to and at each increasing basepair shift, the Pearson correlation between the forward and reverse is calculated. The highest cross-correlation is found when the shift size corresponds to the predominant fragment length (k). Figure taken from Landt *et al.* (2012). **(B)** Cross-correlation profile of D11 C8FS3B and WT ChIP samples

The normalised strand coefficient (ratio between the fragment length cross-correlation and the background cross-correlation peaks, NSC) and the relative strand coefficient (ratio between the fragment length and read length cross-correlation peaks, RSC) are used by ENCODE consortium as tools for assessing the adequacy of enrichment in a ChIP-seq experiment. Good quality ChIP-seq experiments with high signal to noise ratios have larger NSC and RSC values, whereas control samples and failed experiments tend to have values less than 1.1 and 0.8 for NSC and RSC respectively (Figure 4.10A). The FS3B ChIP sample appears to be adequately enriched since the both the NSC (1.17) and RSC (0.93) are above that of the threshold determined by ENCODE (Figure 4.10).

A



B

Sample	cc (fragment length)	cc (read length)	min (cc)	NSC	RSC
D11 C8 FS3B LCL ChIP	0.0098	0.0099	0.0084	1.17	0.93
D11 WT LCL ChIP	0.0073	0.0079	0.0068	1.07	0.45

Figure 4.10. Calculation of normalised strand co-efficient (NSC) and relative strand co-efficient (RSC).

(A) Figure showing the calculation for normalised strand co-efficient (NSC) and relative strand co-efficient (RSC). The NSC represents ratio between the fragment length and the background cross-correlation peaks and the RSC demonstrates the ratio between the fragment length and read length cross-correlation peaks. NSC>1.1 and RSC>0.8 usually indicate adequate enrichment in the samples. Figure taken from Landt *et al.* (2012). **(B)** Calculation of NSC and RSC for D11 C8 FS3B LCL ChIP and D11 WT LCL ChIP samples. Values are taken from the cross-correlation profile of the two samples as illustrated in Figure 4.9B.

4.5.1.3 Graphical pile-up of aligned sequencing reads.

As a brief evaluation of the sequencing data, I examined the distribution of aligned sequencing reads at known binding regions of EBNA3B at the *ITGAL* and *CtBP2* loci (McClellan *et al.*, 2013). The aligned pile-up of individual reads from the FS3B ChIP sample showed strong signals over background reads from FS3B Input and untagged wild-type anti-FLAG ChIP samples at the predicted EBNA3B-binding regions, which indicates that the ChIP-seq data are of adequate quality (Figure 4.11).

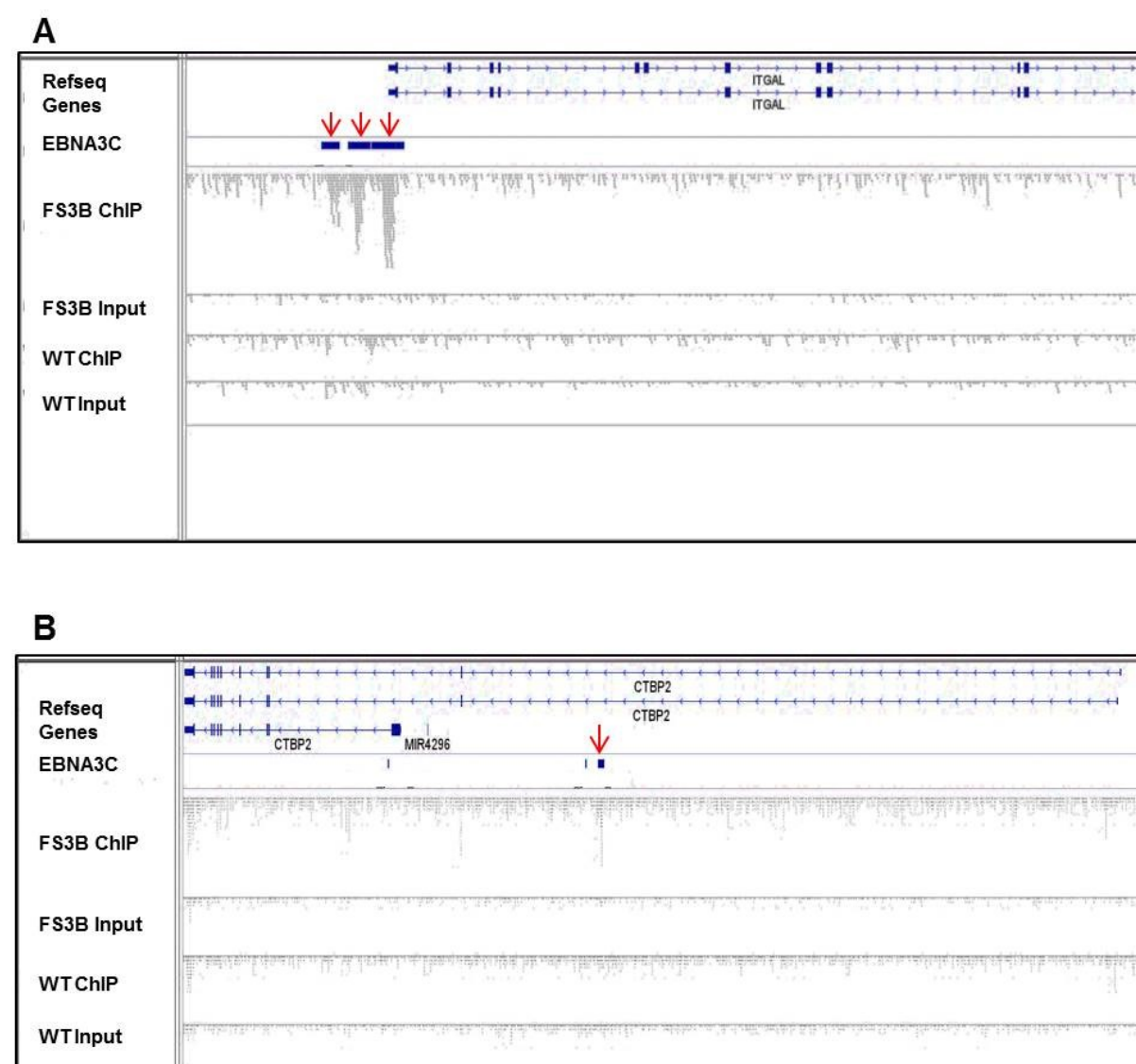


Figure 4.11. Aligned pile-up of reads from ChIP-seq analysis of FS3B and WT LCLs

Pile-up of sequencing reads of the ChIP and Input samples taken from FS3B and WT LCLs at (A) *ITGAL* and (B) *CtBP2* loci. This is compared to the EBNA3C ChIP-seq data (EBNA3C) obtained by McClellan *et al.* (2013). Red arrows indicate experimentally-verified EBNA3B binding peaks (McClellan *et al.*, 2013)

4.5.1.4 Peak-calling filtering control: ChIP sample prepared from untagged wild-type LCL

Next, I used PGS peak-calling algorithm (details in Chapter 2) to identify genome-wide EBNA3B-binding sites based on the aligned sequencing data from the ChIP-enriched epitope tagged FS3B sample. In accordance with the ENCODE guidelines (Landt *et al.*, 2012), the ChIP sample prepared from untagged WT LCLs was chosen as the appropriate control for peak calling since it has a sequencing depth that is comparable to that of the FS3B LCL ChIP sample, and also because it was prepared under the same conditions as the ChIP-enriched FS3B LCL sample, which would account for non-uniform fragmentation of DNA during sonication (Auerbach *et al.*, 2009), sequencing efficiency biases (Dohm *et al.*, 2008) and background binding of the antibody. As such, peak detection using PGS found significant enrichment (p-value <0.05) of EBNA3B at 1931 sites.

4.5.2 EBNA3B binds to fewer regions in the genome than EBNA3A, EBNA3C and EBNA2.

Since the EBNA3B might under some circumstances interact with EBNA3A, 3C and/or EBNA2, it seemed pertinent to study its genomic distribution in relation to these EBV latent proteins. Using separate LCLs encoding C-terminal epitope tagged EBNA3A or EBNA3C, Dr Kostas Paschos performed ChIP-seq experiments and peak calling in a similar manner as is done for EBNA3B. The ChIP-seq data for EBNA3A, EBNA3B and EBNA3C obtained from our lab were compared with previously identified global EBNA3 binding regions (using pan-specific EBNA3C antibody that also immunoprecipitates EBNA3A and EBNA3B in Mutu III BL cell lines) by McClellan *et al.* (2012). EBNA2 binding peaks, again in Mutu III BL cell lines, were also obtained from McClellan and colleagues (2013). Hence, the binding profile of EBNA3B will be compared to that of EBNA3A, EBNA3C, global EBNA3 and EBNA2.

Overall, EBNA3B targets fewer regions than EBNA3A, EBNA3C and EBNA2. Peak detection using PGS found significant enrichment (p-value <0.05) of EBNA3B at 1931 sites,

EBNA3A at 3854 sites and EBNA3C at 5044 sites in the genome (Figure 4.12). In addition, McClellan *et al.* have separately identified 7044 pan-specific EBNA3 peaks and 21605 EBNA2 binding peaks (Figure 4.12) (McClellan *et al.*, 2012, 2013). In total, 35% of our EBNA3C peaks, 30% of our EBNA3A peaks and only 19% of our EBNA3B peaks (Figure 4.13) overlapped with the EBNA3 peaks determined by McClellan and his colleagues (McClellan *et al.*, 2012). Of the EBNA3B peaks which overlapped with the EBNA3 pan-specific ChIP-seq, only 54 are unique EBNA3B peaks which do not overlap with either EBNA3A or EBNA3C (3% of our total EBNA3B peaks) (Figure 4.13). This further supports the need for an additional set of EBNA3B-specific ChIP-seq data. Comparing our ChIP-seq data with that of McClellan and colleagues, we have uncovered a substantial number of additional EBNA3 binding sites which could be used for future analyses. The discrepancies in our ChIP-seq data also imply that our varying approaches of immunoprecipitating EBNA3s may be sensitive to different binding regions in the genome. The varying genetic background of cells used for the ChIP-seq analyses by our lab and McClellan and his colleagues (LCLs vs. BLs respectively) may also account for the differences observed.

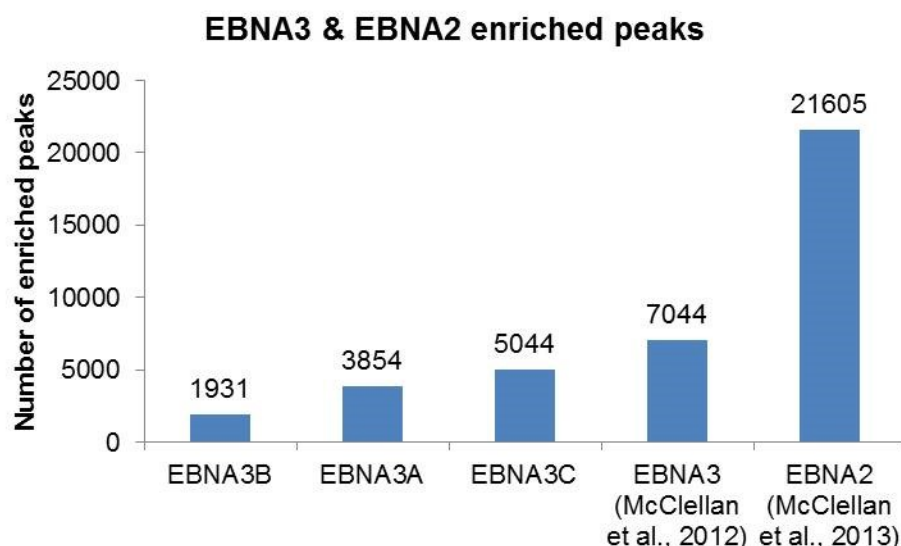


Figure 4.12. EBNA3A, EBNA3B, EBNA3C, pan-specific EBNA3 and EBNA2 enriched peaks in hg19.

Histogram showing number of enriched peaks as determined by our epitope-tagged EBNA3A, EBNA3B and EBNA3C ChIP-seq, and the EBNA2 and pan-specific EBNA3 ChIP-seq analyses done by McClellan and colleagues (McClellan *et al.*, 2012, 2013). **(B)** Overlap of EBNA3A, EBNA3B and EBNA3C ChIP-seq peaks (from our lab) with pan-specific EBNA3 ChIP-seq peaks (McClellan *et al.*, 2012).

Overlap of EBNA3B, EBNA3A or EBNA3C ChIP-seq peaks with pan-specific EBNA3 ChIP-seq peaks

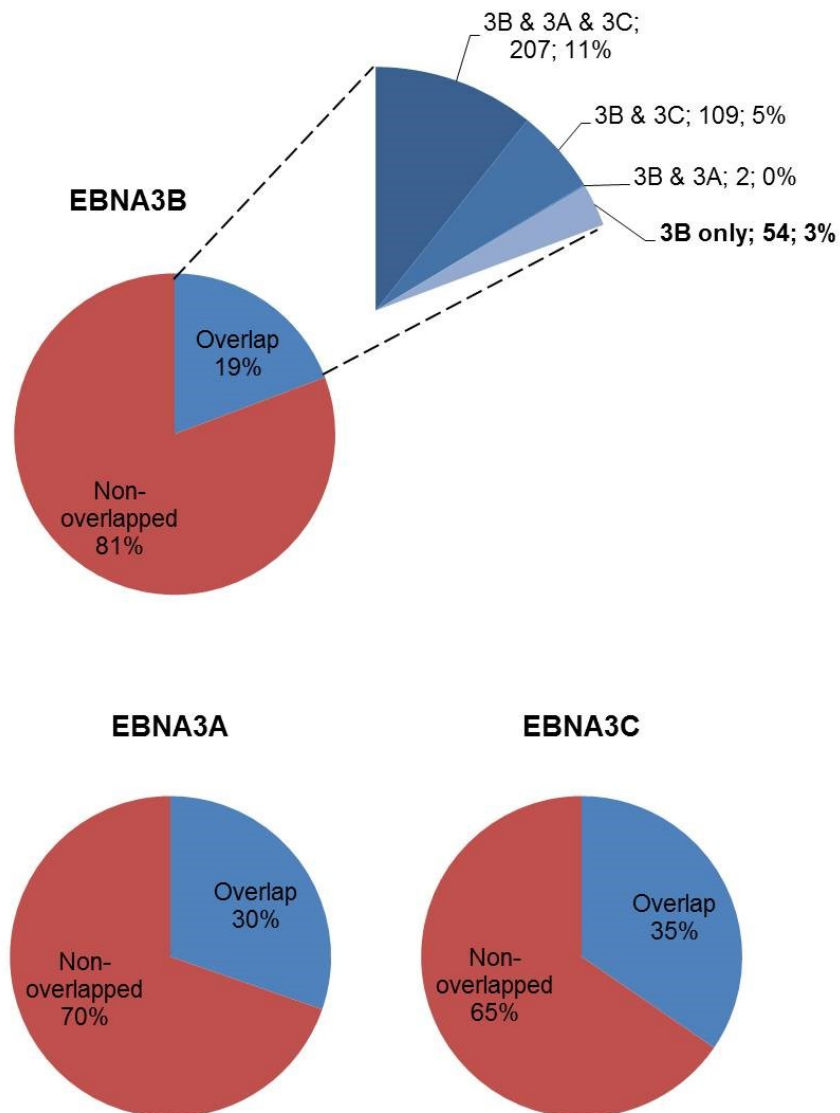


Figure 4.13. Overlap of EBNA3A, EBNA3B and EBNA3C ChIP-seq peaks (from our lab) with pan-specific EBNA3 ChIP-seq peaks (McClellan *et al*, 2012).

Enlarged insert for EBNA3B depicts any co-localisation of EBNA3B with EBNA3A and EBNA3C peaks within the 19% of the EBNA3B binding sites that overlapped with EBNA3 pan-specific peaks. The number of peaks and percentage relative to the total number of EBNA3B peaks are shown.

4.5.3 EBNA3B co-localises mainly with EBNA3C, but also with EBNA3A and EBNA2

Of the 1931 EBNA3B peaks, 592 (30%) of these peaks co-localise with EBNA3C and 388 (20%) peaks with EBNA3A. (Figure 4.15). Of the 21605 EBNA2 peaks identified by McClellan *et al.* (2013), 306 (16% of EBNA3B peaks) co-localised with EBNA3B (Figure 4.15). 197 (10%) of the EBNA3B peaks co-localised uniquely with EBNA3A & EBNA3C and 165 peaks (9%) overlapped with EBNA3A, 3C and EBNA2 (Figure 4.15).

4.5.4 Majority of the EBNA3B binding peaks do not co-localise with EBNA3A, 3C or EBNA2.

In order to evaluate the binding profile of EBNA3B in relation to the other EBNA3s and EBNA2, the EBNA3B binding peaks were classified into different groups depending on whether, and with which of the EBNA3s and EBNA2 the binding regions co-localise. A huge majority of the EBNA3B binding regions were unique to EBNA3B (65%) while the rest of the regions were mainly co-localizing with either all the EBNA3s (10% with EBNA3A and EBNA3C, 9% with EBNA3A, EBNA3C and EBNA2) or with only EBNA3C (7%) (Figure 4.14). Only a minority of the regions (1-3%) co-localised uniquely with either EBNA3A or EBNA2, and none with both EBNA3A and EBNA2 (Figure 4.14). In summary, very few (1%) of the EBNA3B peaks co-localised with EBNA3A without EBNA3C, whereas a much larger number of EBNA3B peaks overlap with EBNA3C without EBNA3A (12%, Figure 4.14). This suggests that EBNA3C is a core player in the co-ordinated gene regulation that involves EBNA3B and the other EBNA3s, whereas EBNA3A may have a minor role the function of EBNA3B as a transcriptional regulator. EBNA2 is unlikely to have a significant a role in the interaction between EBNA3B and EBNA3s, since its binding regions overlap mainly with EBNA3B where EBNA3C or where both EBNA3A and 3C are present. Despite the large numbers of EBNA2 binding sites genome-wide, only very few of these overlap uniquely with EBNA3B (54 peaks, 3% of EBNA3B peaks) (Figure 4.14). Put together, these data support the notion that EBNA3B interacts with other EBNA3s and EBNA2 to regulate a relatively small subset of its putative target genes (White *et al.*, 2010).

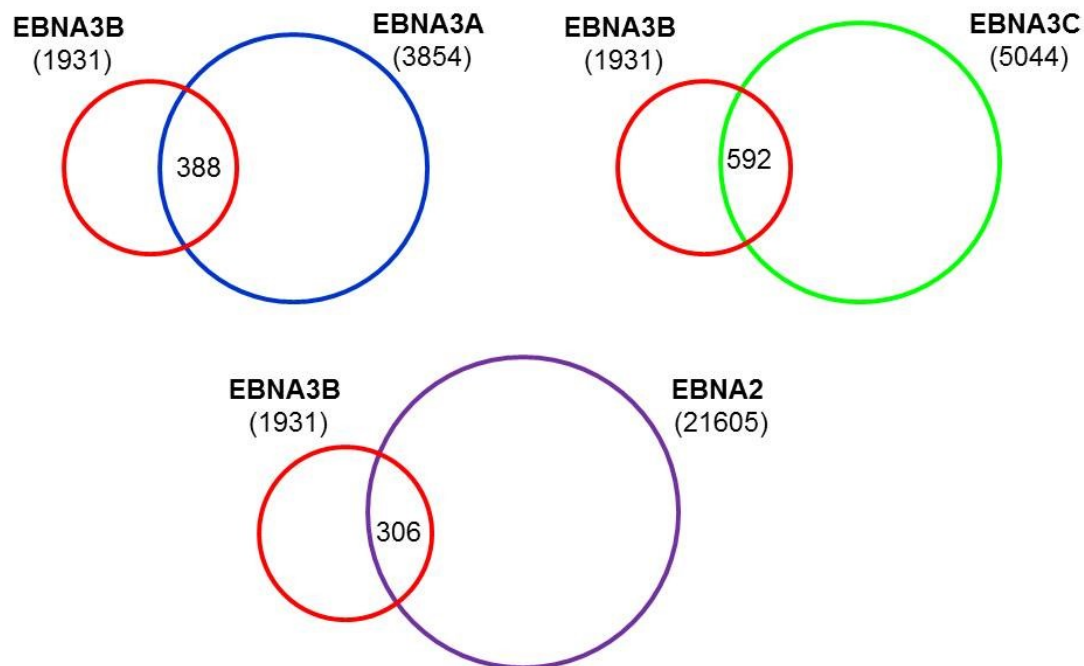


Figure 4.15. Overlap of EBNA3B peaks with EBNA3A, EBNA3C, & EBNA2 peaks (McClellan *et al.*, 2013).

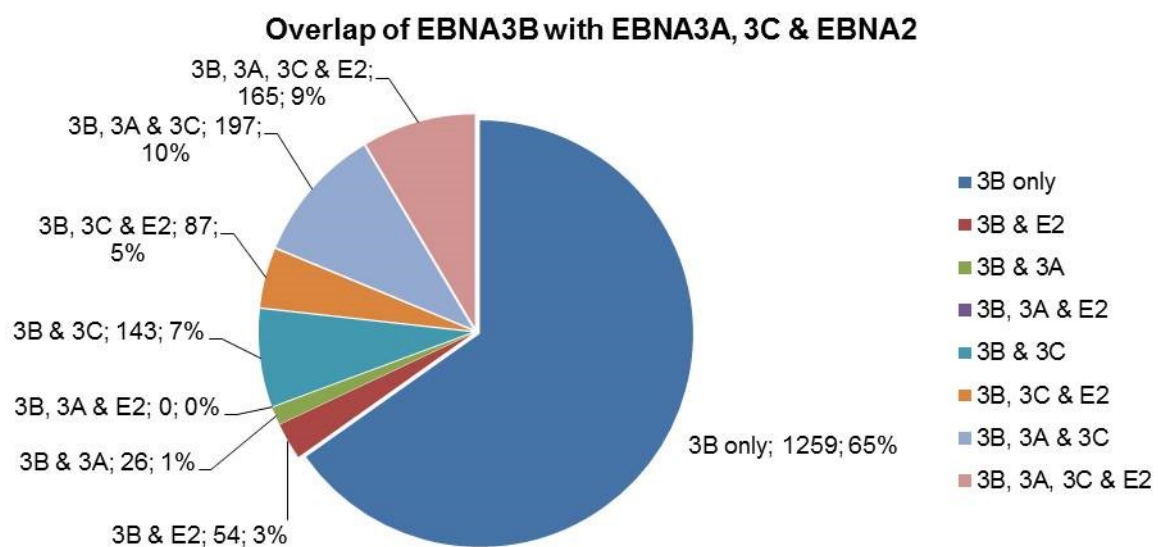


Figure 4.14. Distribution of EBNA3B peaks according to their co-localisation with EBNA3A, EBNA3C and/or EBNA2.

Pie chart displaying the distribution of EBNA3B binding peaks that uniquely overlap with EBNA3A, 3C & EBNA2, either individually or collectively.

4.5.5 EBNA3B binds distally to its putative regulated genes.

Next, I determined where EBNA3B is found relative to transcriptional start sites (TSS) of all genes, and of pre-determined EBNA3B-regulated genes. Less stringent criteria ($p < 0.05$ as opposed to $p < 0.001$ and fold change > 2 -fold between WT and 3BKO LCLs) is employed for selecting the EBNA3B-regulated genes using the data generated by White *et al.* (2010) (data available in a searchable form at <http://www.epstein-barrvirus.org>). Unless otherwise stated, this expanded set of genes was used throughout the ChIP-seq analyses. With a more inclusive threshold, we would be able to examine potential regulated genes associated with EBNA3B binding which would otherwise be ignored by the highly stringent selection.

Analyses using PGS revealed that 24% of EBNA3B peaks are found proximal to TSS (< 5 kb) of all genes but only 3% of the total EBNA3B peaks (i.e. 1 in 8 of peaks proximal to TSS) are located proximal to putative EBNA3B regulated genes (Figure 4.16). The highest number of EBNA3B binding peaks overlapping with genes was found between 5-50kb from the TSS of EBNA3B-regulated genes, with the median distance from TSS to be 23,732kb. This is consistent with the observations made by other groups that EBNA3A, EBNA3C and global EBNA3 binding peaks tend to be found at distal regulatory gene regions (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015)

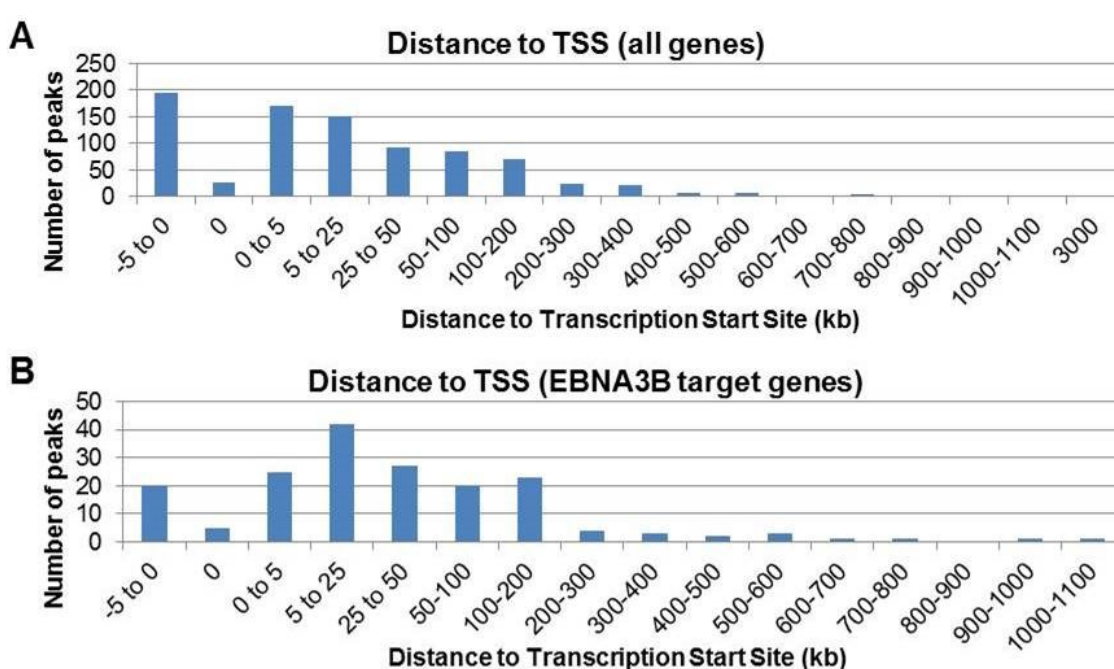


Figure 4.16. Distance of EBNA3B binding sites to transcriptional start sites of (A) all genes and (B) potential EBNA3B target genes (White *et al.*, 2010).

4.5.6 The distribution of EBNA3B across genome segments alters considerably depending on its co-localization with other EBNA3s and EBNA2

In order to identify EBNA3B association with functional regions genome-wide, I compared the binding regions of EBNA3B with publicly available combined genome-segmentation data available on the EBV-immortalised LCL GM12878 published by the ENCODE consortium (Dunham *et al.*, 2012; Ernst and Kellis, 2012; Hoffman *et al.*, 2013). Briefly explained, the genome is divided into discrete segments, which are assigned one of seven major chromatin states. These include predicted promoter regions (TSS and promoter flanking regions), transcribed regions, "active" distal states (enhancers, weak enhancers and CTCF-enriched elements), and repressed states. These chromatin states are defined by and limited to regions containing combinations of ChIP-seq signals of histone modifications (H3K4me1/2/3, H3K9Ac, H3K27Ac, H3K27me3, H3K36me3, H4K20me1) and transcription factors (such as RNA PolII and CTCF), and signals from DNase-seq and FAIRE-seq (which identify regions of open chromatin). A summary of the combined state types is described in Table 4.3.

Of the characterised genome segments, EBNA3B peaks associate mostly with enhancers and weak enhancers (24% in total), while the other peaks which localise with known, characterised genome segments are evenly distributed across TSS, transcribed regions and repressed regions (7-10% each) (Figure 4.17A). Interestingly, over half (56%) of the EBNA3B peaks do not associate with any of the pre-determined genome segments. While EBNA3A and EBNA3C are very similar to each other in their binding profile, comparatively, a higher percentage (8% vs 3-4%) of EBNA3B peaks are found within transcribed regions, and a lower percentage of them are localised within active regulatory regions (24 vs 74-76%) (Figure 4.17). The pan-specific EBNA3 binding profile appears to be highly similar to that of our EBNA3A and EBNA3C ChIP-seq, with a slightly higher percentage of EBNA3 sites localising within repressive segments (17% vs 8-9%) (Figure 4.17D). When compared to EBNA2, in addition to having fewer co-localisation sites with

enhancers (24 vs 70%), EBNA3B appears to have a lower percentage of sites located at TSS (10% vs 42%), and at repressive regions (9% vs 30%) (Figure 4.17E). Taken together, these results hint at EBNA3B having different specificities, and perhaps functional differences, to EBNA3A, 3C and EBNA2.

Table 4.3. Description of genome segments as described by the ENCODE consortium.
Table taken from Dunham *et al.* (2012) with some modifications.

Label	Description	Details
CTCF	CTCF-enriched element	Sites of CTCF signal lacking histone modifications, often associated with open chromatin. Also enriched for the cohesin components RAD21 and SMC3; CTCF is known to recruit the cohesin complex.
E	Predicted enhancer	Regions of open chromatin associated with H3K4me1 signal. Enriched for other enhancer-associated marks, including transcription factors known to act at enhancers.
PF	Predicted promoter flanking region	Regions that generally surround TSS segments (see below).
R	Predicted repressed or low-activity region	This is a merged state that includes H3K27me3 polycomb-enriched regions, along with regions that are silent in terms of observed signal for the input assays to the segmentations (low or no signal).
TSS	Predicted promoter region including TSS	Found close to or overlapping GENCODE TSS sites. Enriched for H3K4me3. Sites of open chromatin. Enriched for transcription factors known to act close to promoters and polymerases Pol II and Pol III.
T	Predicted transcribed region	Overlap gene bodies with H3K36me3 transcriptional elongation signal. Enriched for phosphorylated form of Pol II signal (elongating polymerase) and poly(A) ⁺ RNA, especially cytoplasmic.
WE	Predicted weak enhancer or open chromatin <i>cis</i> -regulatory element	Similar to the E state, but weaker signals and weaker enrichments.

Next I investigated whether co-localisation of EBNA3B with EBNA3A, EBNA3C or EBNA2 alters the distribution of EBNA3B across genome segments. EBNA3B binding peaks that overlapped with these other nuclear antigens (as illustrated in Figure 4.14) were inspected for association with the regulatory segments. Interestingly, peaks which are not associated with EBNA3A, EBNA3C or EBNA2 (i.e. unique EBNA3B peaks) tend to localise within transcribed or repressive segments, although a huge majority (74%) of these were found within genome locations that do not fit the characteristics of any predicted genome

segments (Figure 4.17A). A similar profile is also seen in the few EBNA3B peaks which associate solely with EBNA3A binding peaks (unique EBNA3B/3A peaks), although a higher percentage of these peaks (26% of 27 peaks) localise within transcribed regions. However the total number of peaks within this category may be too low for an adequate assessment of distribution (Figure 4.18B). When we examined the breakdown of each genome segment in detail, it is evident that unique EBNA3B binding regions make up the bulk of the association with repressed and transcribed segments (Figure 4.18C). These unique EBNA3B peaks also contribute to the bulk of the peaks that are localised within uncharacterised genome segments (Figure 4.18C).

In contrast, EBNA3B peaks which co-localise only with EBNA3C (i.e. unique EBNA3B/3C peaks) tend to associate with enhancer and weak enhancer segments (23 and 27% respectively) (Figure 4.18D). A large proportion of these peaks are also found in uncharacterised genome segments (27%). Similarly, this distribution is observed in EBNA3B peaks which co-localise uniquely with EBNA3A and EBNA3C (unique EBNA3B/3A/3C peaks) (Figure 4.18E).

While co-localisation of EBNA3B and EBNA3C (with or without EBNA3A) appears to favour association with enhancer and weak enhancer segments, co-localisation of EBNA3B with EBNA2 seems to be associated with an increased proportion of these peaks being found within TSS. EBNA3B peaks which co-localised with EBNA2, regardless of the presence of EBNA3A & 3C, tend to be found near TSS (30-42%) and strong enhancers (44-52%), which may reflect the function of EBNA2 as a transactivator of gene expression (Figure 4.18F-H).

In summary, the distribution of EBNA3B across genome segments varies greatly depending on its co-localization with EBNA3A, 3C and EBNA2. Intriguingly, EBNA3B on its own seems to localise pre-dominantly within transcribed, repressive and uncharacterised regions of the chromatin, whereas EBNA3B peaks which co-localise with EBNA3A, EBNA3C

and EBNA2 tend to be found in active regulatory regions and TSS. This seems to suggest in addition to having similar specificities to EBNA3A and 3C, EBNA3B alone may be involved in different regulatory processes. A closer look at the chromatin landscape around its binding regions may support these observations and provide further clues into its differential regulatory functions.

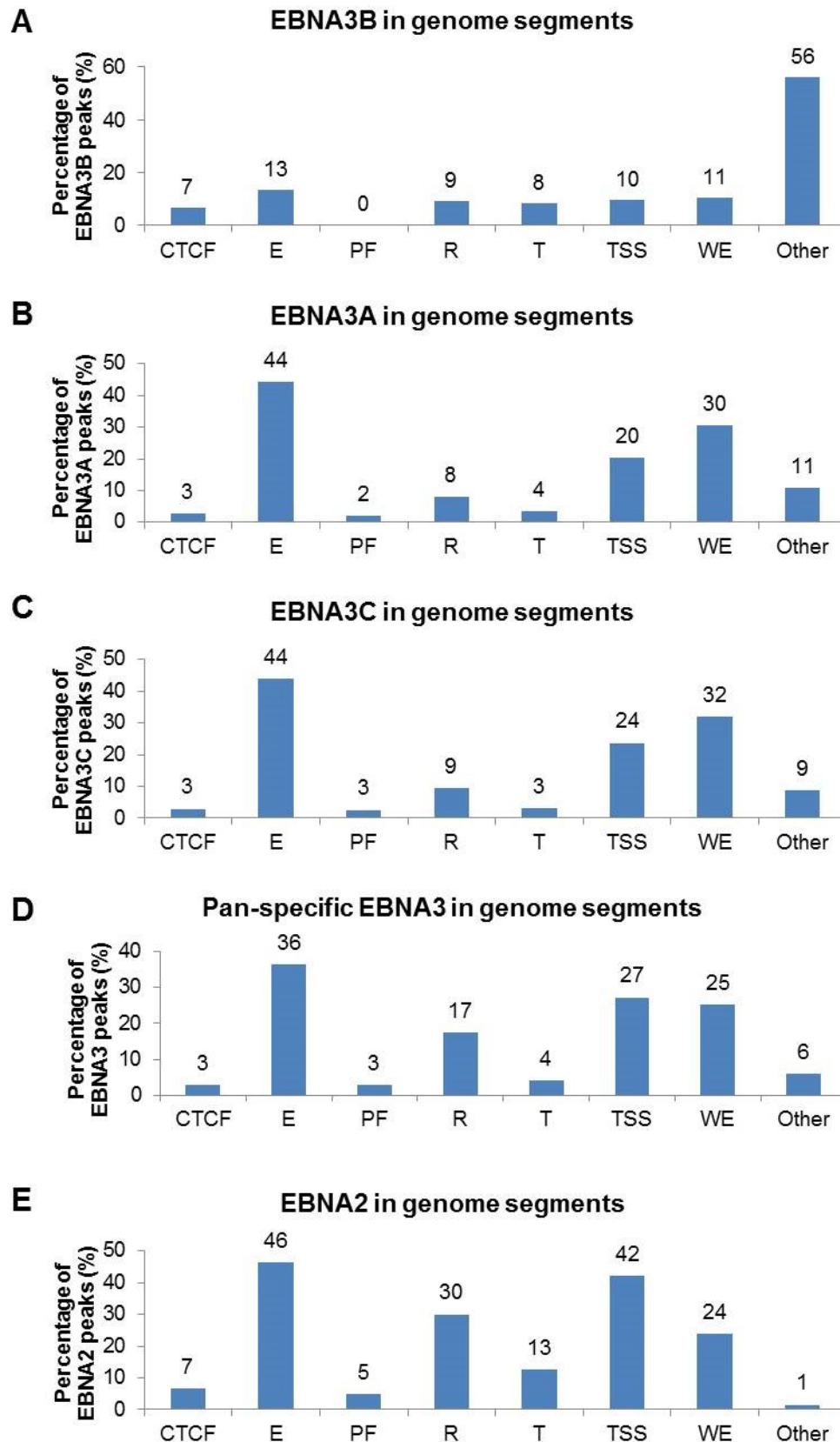


Figure 4.17. Distribution of (A) EBNA3B, (B) EBNA3A, (C) EBNA3C, (D) EBNA3 (pan-specific) and (E) EBNA2 ChIP-seq peaks across different genome segments.

Legend and descriptions of each genome segment are found in Table 4.3. Peaks which are not otherwise found in any of the described genome segments are collectively labelled as "Other".

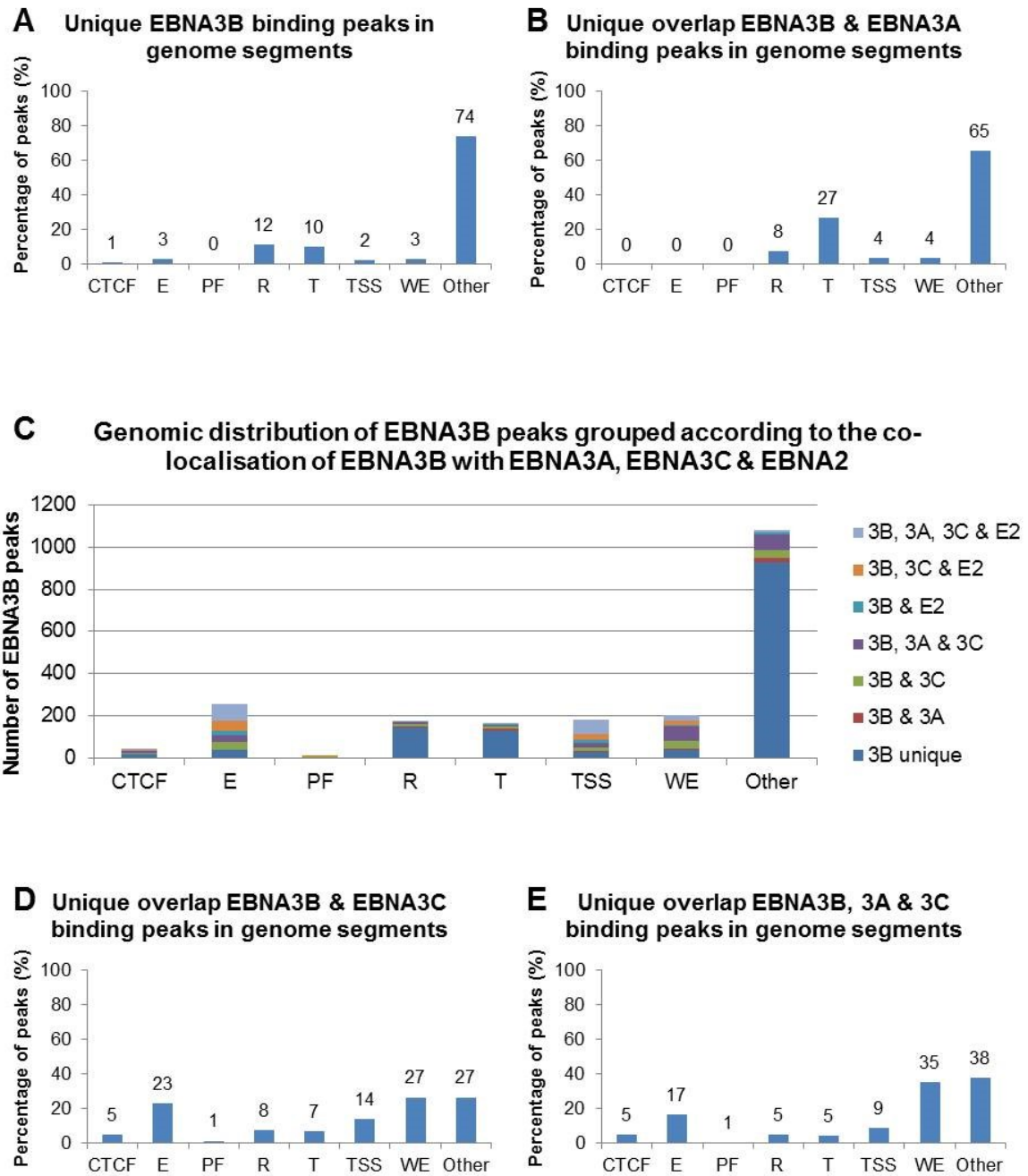


Figure 4.18. Distribution of EBNA3B binding peaks within different genome states. (part 1)

Distribution of (A) unique EBNA3B binding sites & (B) unique EBNA3B & EBNA3A binding sites across different genome states. (C) Distribution of EBNA3B (3B) peaks across genome states, grouped according to their co-localisation with EBNA3A (3A), EBNA3C (3C) and EBNA2 (E2). and EBNA3B & EBNA2 overlap regions across genome segments. Distribution of (C) unique EBNA3B & EBNA3C binding peaks and (E) unique EBNA3B, EBNA3A & EBNA3C binding sites in genome segments. (Figure continues next page)

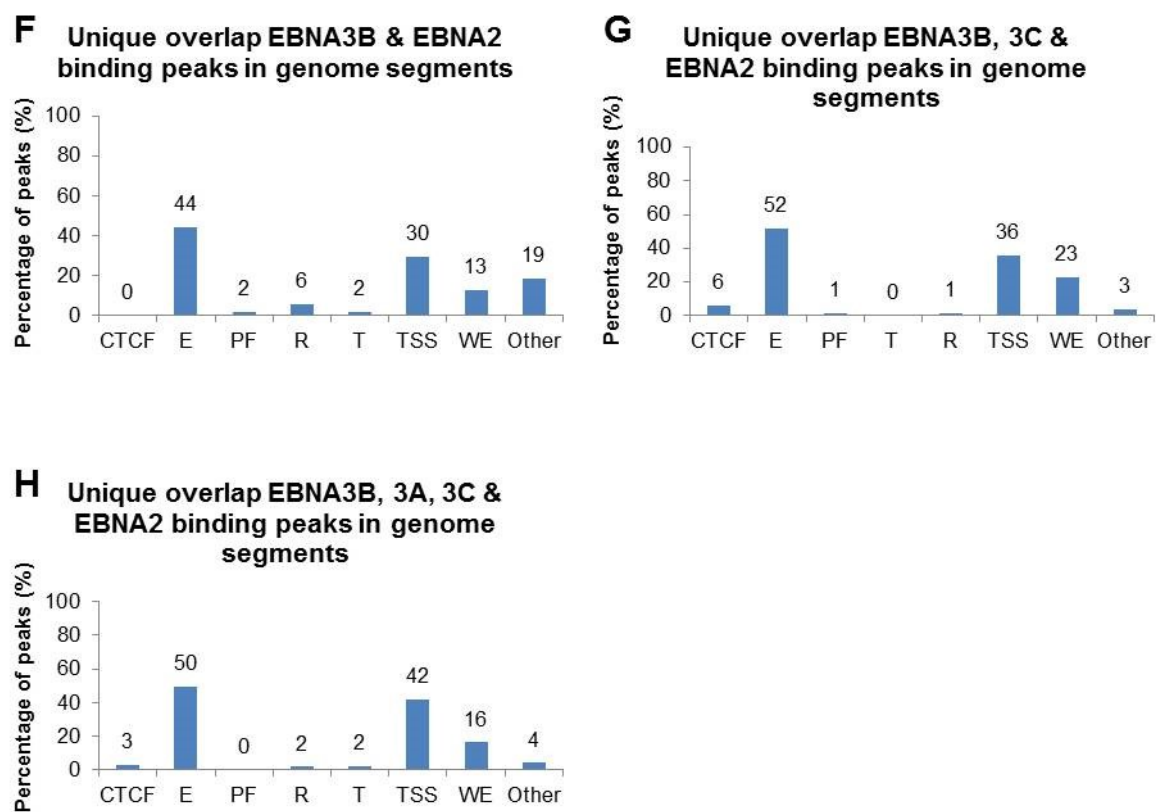


Figure 4.18. Distribution of EBNA3B binding peaks within different genome states. (part 2)
 Distribution of (F) unique EBNA3B & EBNA2 binding sites, (G) unique EBNA3B, EBNA3C & EBNA2 binding sites and (H) unique EBNA3B, EBNA3A, EBNA3C & EBNA2 binding sites across different genome states.

4.5.7 EBNA3B co-localise with histone modifications associated with active promoters/enhancers, when also co-localised with other EBNA3s

Since genome segments are largely determined by the chromatin landscape, I compared the binding regions of the EBNA3s with various histone modifications of the GM12878 LCL mapped out by the ENCODE consortium project to verify my findings. Of the modifications inspected, the top 4 modifications associated with EBNA3B, 3A, 3C and pan-specific EBNA3 peaks (McClellan *et al.*, 2012, 2013) are identical, namely H3K4me1, H3K4me2, H3K27Ac, and H3K4me3, which are all marks found near promoters and/or distal active enhancer regions (Figure 4.19A). Similarly, the top 3 histone marks found to associate with EBNA2 binding sites are also identical to those of the EBNA3s (individual or pan-specific). Modifications associated with repression (H3K27me3 and H3K9me3) are ranked low in the lists for all the EBNA3s and EBNA2 (Figure 4.19A).

A lower percentage of EBNA3B peaks (29% of EBNA3B peaks vs 69-75% in EBNA3A, EBNA3C, pan-specific EBNA3 and EBNA2 peaks for the top histone mark, H3K4me1) seemed to co-localise with histone modifications on our list, and a large proportion (66% of EBNA3B peaks vs. 16-19% in EBNA3A, EBNA3C, pan-specific EBNA3 and EBNA2 peaks) may co-localise with other histone marks that have not been included in this study (Figure 4.19). Additional examination and comparisons with other histone modification ChIP-seq data may be required to elucidate other major histone marks with which EBNA3B might be associated.

Further analyses found that the top three histone modifications associated with EBNA3B binding peaks remain unchanged regardless of their co-occupancy with other EBNA3s and EBNA2. With the exception of unique EBNA3B peaks and unique EBNA3B/3A peaks, the highest ranked modifications associated with EBNA3B peaks which co-localised with EBNA3A, EBNA3C and/or EBNA2 remained as H3K4me1, H3K4me2 and H3K27ac (Figure 4.19B). While H3K36me3 (usually found in gene bodies) modifications are not strongly associated with majority of the EBNA3B peaks, it appears to be the 3rd most common histone modification associated with unique EBNA3B peaks and 2nd for unique

EBNA3B/3A peaks (Figure 4.19B). This observation corresponds well with the data showing that unique EBNA3B and unique EBNA3B/3A binding peaks are frequently detected within genome segments containing transcribed regions.

Overall comparisons showed that a higher proportion (for H3K4me1, 65-95% of EBNA3B peaks which associate with EBNA2, vs 8-59% of EBNA3B peaks which do not associate with EBNA2) of EBNA3B binding sites which overlap with EBNA2 tend to co-localise with the histone modification marks that we have examined (Figure 4.19). This suggests strongly that EBNA2 may play a role in the deposition or alteration of these histone modification marks, which may be related to the gene regulatory function of EBNA2.

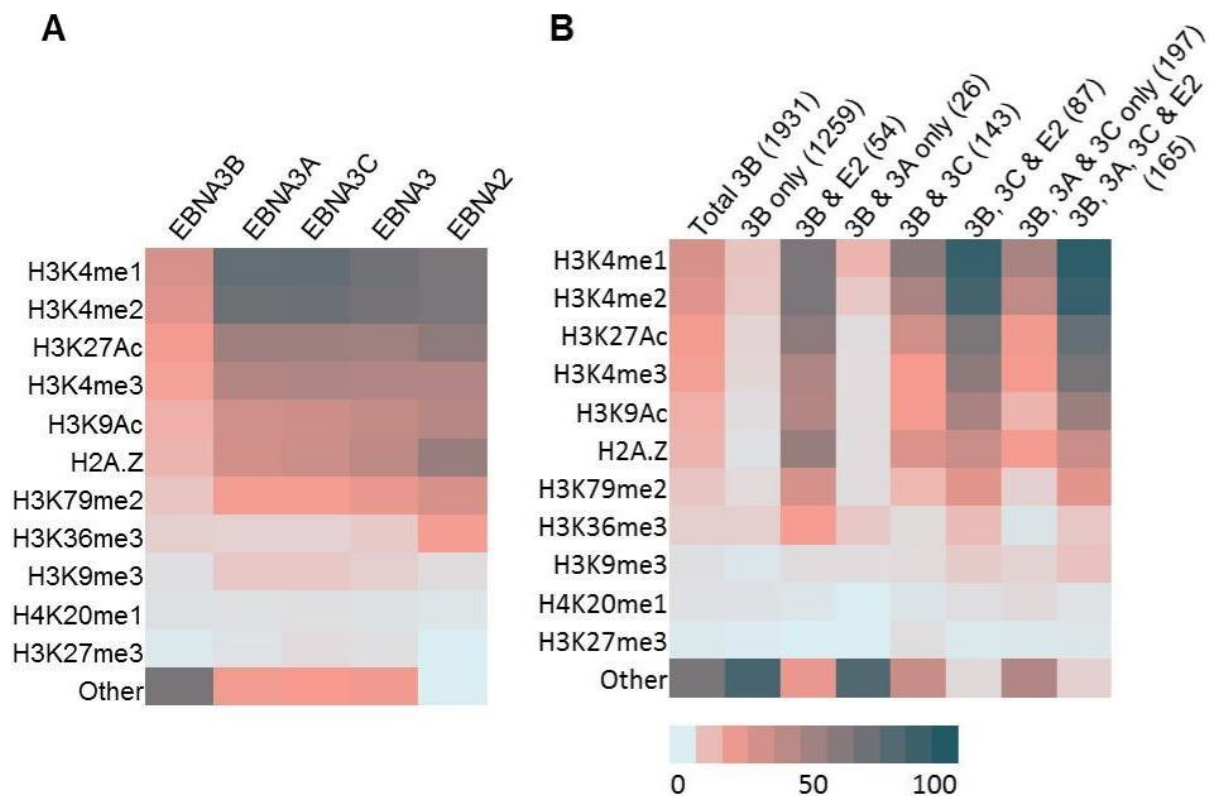


Figure 4.19. Heat map comparing the co-occupancy of EBNA3s and EBNA2 with histone marks.

(A) Percentage of peaks overlapping with histone modifications are reported for each EBNA3A, EBNA3B, EBNA3C (ChIP-seq by our group), EBNA3 (pan-specific ChIP-seq done by McClellan *et al*, 2012) & EBNA2 (ChIP-seq done by McClellan *et al*, 2013) protein and organized into a heat map for comparison. (B) Heat map showing percentage of EBNA3B peaks overlapping with histone marks, where EBNA3B peaks are grouped according to their co-localisation with EBNA3A/3C/EBNA2. The total number of peaks for each category is shown in brackets.

4.5.8 EBNA3B enriched sites are also bound by specific transcription factors

The EBNA3s are known to interact with other transcriptional regulators (Cotter and Robertson, 2000; Knight *et al.*, 2003; Radkov *et al.*, 1999; Touitou *et al.*, 2001b) to exert their gene regulatory functions. Most notably, in some circumstances the EBNA3s are able to exert their repressive function by binding to EBNA2 DNA-targeting partner RBPJk and inhibiting activation of its target genes (see section 1.7.1). Importantly, the EBNA3s, along with EBNA2 have also been reported to occupy sites bound by DNA-binding transcriptional regulators including RBP-Jk, BATF, BCL11A, BCL3, IRF4, PAX5, SP1, TCF12, PU.1 and EBF1 (McClellan *et al.*, 2013). EBNA3A and EBNA3C are also believed to be tethered to DNA via BATF-containing protein complexes (Jiang *et al.*, 2014; Schmidt *et al.*, 2015). In order to establish a set of transcriptional factors (TF) that may interact with EBNA3B at its binding sites, I compared the EBNA3B peaks to 90 transcription factor ChIP-seq data available for GM12878 LCLs taken from the UCSC Genome Browser. Unfortunately, at the time of this analyses, the only RBPJk ChIP-seq data available for B-lymphocytes were done on IB4 LCLs which lack EBNA3B (Zhao *et al.*, 2011a) hence it will not be included in the transcription factor analyses. Very recently, the Lieberman lab has published RBPJ binding sites based on a ChIP-seq done with LCLs (Lu *et al.*, 2016), and their data may be used for future comparisons with EBNA3B binding sites.

The top 5 transcription factors associated with EBNA3B binding peaks are RUNX3 (24%), NFIC (24%), ATF2 (20%), BATF (18%) and FOXM1 (17%) (Figure 4.20A). Of these, RUNX3, ATF2 and BATF have already been reported to have binding sites which coincide with EBNA3A and EBNA3C (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015). On closer inspection, a great number of the top 20 transcription factors (e.g. RUNX3, ATF2, BATF, IRF4, PAX5, BCL11A, MEF2A, POU2F2, EBF1, TCF12) (Figure 4.20A) have already been reported in literature to occupy the same binding sites as EBNA3A, EBNA3C and/or EBNA2 (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015; Zhao *et al.*, 2011a).

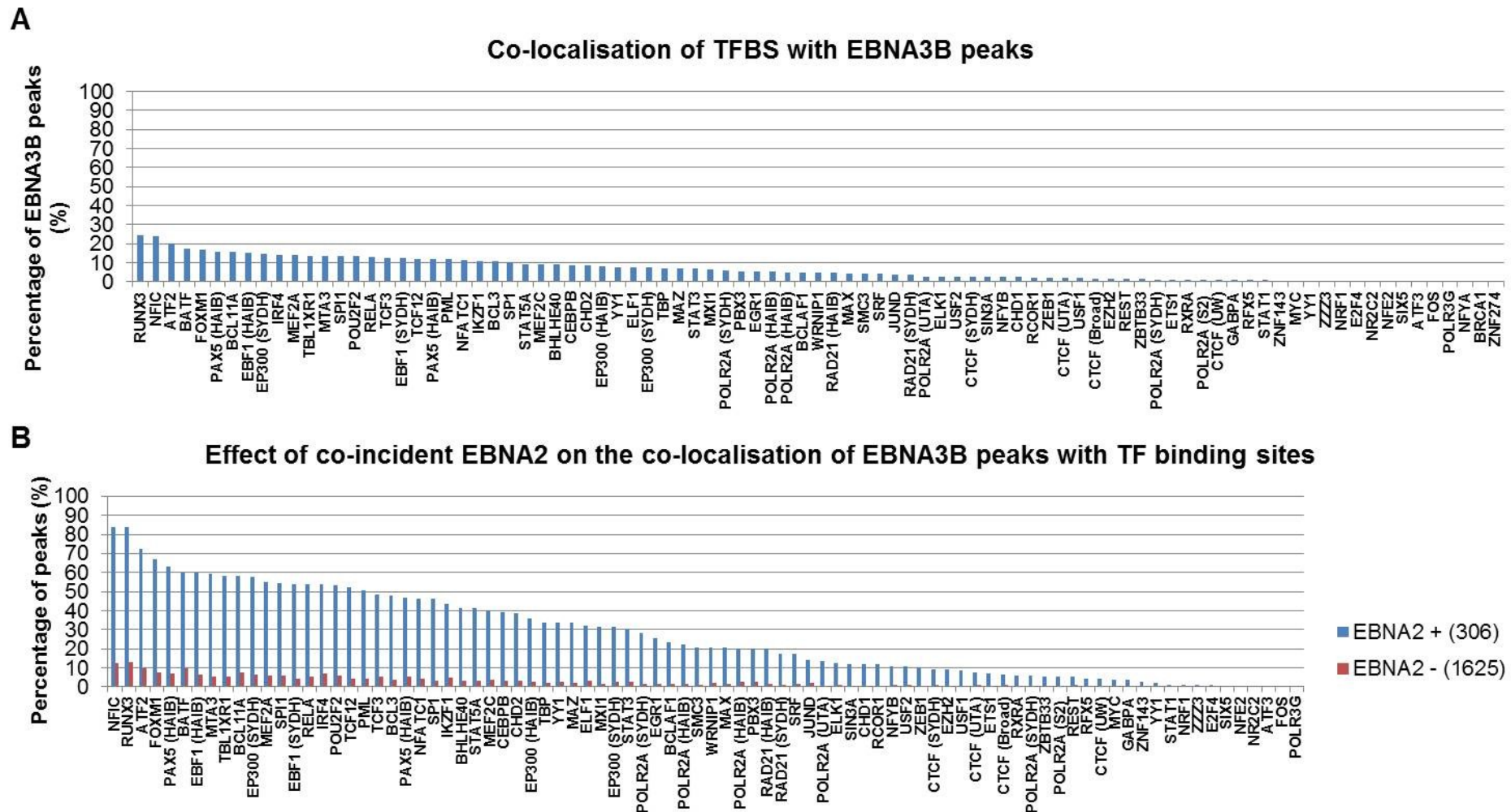


Figure 4.20. Co-localisation of transcription factor binding sites (TFBS) with EBNA3B peaks.

Annotations for TFBS are obtained from published ChIP-seq data available on the UCSC genome browser database. Where multiple ChIP-seq studies for the same transcription factor are available, the groups responsible for the data are indicated in brackets. These groups include Broad, Stanford/Yale/UC-Davis/Harvard (SYDH), HudsonAlpha Institute (HAIB), University of Texas-Austin (UTA) and University of Washington (UW). **(A)** Percentage of total EBNA3B peaks co-localising with each of the TFBS. **(B)** Histograms showing the effect of co-incident EBNA2 on the co-localisation of EBNA3B peaks with TFBS. EBNA3B peaks are grouped into two categories: those which co-localise with EBNA2 (EBNA2+) and those which do not co-localise with EBNA2 (EBNA2-). Percentages of EBNA3B peaks in each group which co-localise with TFBS are reported.

Since co-localisation with EBNA2 may be associated with an alteration of chromatin landscape at EBNA3B binding sites (section 4.5.6), I investigated whether co-occupancy with EBNA2 would be associated with any differences in the association of EBNA3B peaks with TF binding sites (TFBS). On inspection, I found that while co-incidence with EBNA2 does not change the top 20 TF which occupy the same sites as EBNA3B peaks, a very high percentage of EBNA3B peaks with co-incident EBNA2 binding (48-83% for top 20 TF) were associated with TF binding sites, whereas only a small proportion (4-14% for top 20 TF) of peaks which do not overlap with EBNA2 are be associated with transcription factors we are considering (Figure 4.20B). This may be a reflection of the gene regulatory activity of EBNA2, and suggest that EBNA3B may be interacting with EBNA2 in regions where they co-localise to co-regulate their target genes through the recruitment of transcriptional regulators.

Although comparisons of the overlap between TF binding sites and EBNA3B/EBNA2 peaks can be useful in identifying co-occupancy with potential co-regulators, it is important to note that the number of binding sites for each transcription factor can vary between 211 (POLR3G) and 314,874 [CTCF (Broad)]. Hence, transcription factors that bind in abundance across the genome (such as RUNX3: 67,965 sites) would have a higher likelihood of co-localising with EBNA3B peaks anyway. Using RUNX3 as an example, while 25% of EBNA3B peaks overlap with RUNX3, only 0.7% of RUNX3 peaks co-localised with EBNA3B. Conversely, overlap of EBNA3B with transcription factors that bind to fewer sites (such as STAT3: 6487 sites) could equally be significant despite fewer overlap regions. Using STAT3 as an example, although only 7% of EBNA3B peaks localised with STAT3-binding regions, a relatively high percentage (2%) of STAT3 binding sites (as compared to 0.7% of RUNX3) are co-occupied by EBNA3B. Hence, the caveat of this analysis is that it is influenced by the total number of binding sites across the genome for each transcription factor analysed. Careful consideration of the relative abundance of binding sites for each transcription factor should be done when interpreting the results.

Next, using Partek Genomics Suite, I conducted a search of known transcription factor binding motifs at all the EBNA3B binding peaks and compared the data to the top 20 transcription factors co-localizing with EBNA3B. Of the top 20 transcription factors which co-localised with EBNA3B peaks, the motif of 6 of these transcription factors (NFIC, PAX5, EBF1, SPI1, TCF3 & TCF12) were enriched within EBNA3B binding sites ($p < 0.05$) (Table 4.4). Future experiments could place more emphasis on these transcription factors to determine whether they play a major role in the gene regulatory functions of EBNA3B.

Table 4.4. Transcription factor motifs associated with EBNA3B binding regions.

Transcription factors that have a significant number of their motifs identified within the EBNA3B binding sites ($p < 0.05$) and are within the top 20 TF co-localising with EBNA3B.

Transcription factor motif	p-value
NFIC	6.4E-31
PAX5	0.00075
EBF1	5.72E-10
SPI1	0
TCF3	1.25E-55
TCF12	6.8E-11

4.5.9 EBNA3B regulated genes are found within the same Contact Domains as EBNA3B peaks

In order to provide insight into the function of EBNA3B, ChIP-seq data is most useful in identifying any association between its binding regions and its genes reported to be regulated by EBNA3B. In particular, looping mechanisms have been implicated in the regulation of *CTBP2* by EBNA3B and EBNA2 (McClellan *et al.*, 2013). To examine the genome-wide occurrence of this phenomenon, I interrogated chromatin conformation data available on the GM12878 LCL. In 2014, using Hi-C techniques, Rao and colleagues revealed that the genome is partitioned into contact domains comprising of looped chromatin which frequently bring promoters and distal enhancers into close proximity (Rao *et al.*, 2014). Using data obtained from their analyses of the GM12878 LCL cell line, I identified 951 contact domains (out of 9274) that contain EBNA3B peaks.

On close inspection, 522 contact domains which contained EBNA3B peaks overlapped with TSS of 641 (out of 4445) potential EBNA3B-regulated genes ($p < 0.05$ for differential gene expression between WT and 3BKO LCLs from the microarray done by White *et al.*, 2010) and there are considerably more contact domains containing genes predicted to be down-regulated by EBNA3B than up-regulated (Figure 4.21). 224 of these domains contain more than one EBNA3B regulated gene, which may suggest co-regulation. 70 of these contain both potentially repressed and activated genes – this is an interesting observation worthy of further study (Figure 4.21). Similarly, when the stringency for determining potential EBNA3B-regulated genes is reduced to $p < 0.001$, TSS of 128 potential EBNA3B-regulated genes were found in 170 contact domains which contains EBNA3B peaks (Figure 4.21). Of these, 41 contact domains contain more than one EBNA3B-regulated gene, although only 3 of these contain potential genes which are regulated in opposite directions. Regardless of the stringency, there are more down-regulated than up-regulated genes (2-fold more of down-regulated genes in $p < 0.05$ genes and 5-fold in $p < 0.001$ genes) found within the same contact domains as EBNA3B binding sites.

In addition to my analyses, Dr Kostas Paschos has independently generated a random set of peaks (using a random number generator in Microsoft Excel) and determined the co-localisation of these peaks with contact domains that contain EBNA3B-regulated genes (Paschos *et al.*, 2016). Using this method, Dr Paschos has found that the localisation of EBNA3B peaks within contact domains that harbour EBNA3B-induced genes is not more significant than the probability of finding randomly generated peaks within the same domains as EBNA3B-induced genes. The converse is true for EBNA3B-repressed genes: the co-localisation of EBNA3B peaks with contact domains that harbour EBNA3B-repressed genes is significantly higher than the co-localisation of random peaks with contact domains containing EBNA3B-repressed genes. Together, our findings support the hypothesis that EBNA3B is primarily a transcriptional repressor (Figure 4.21).

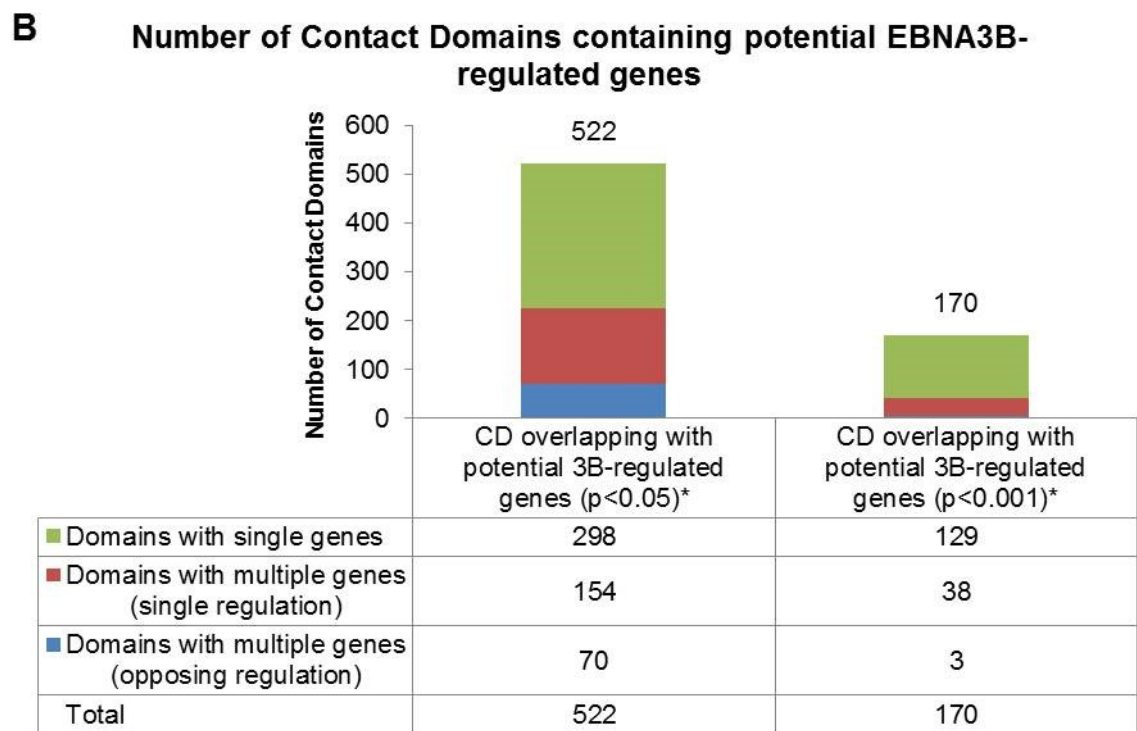
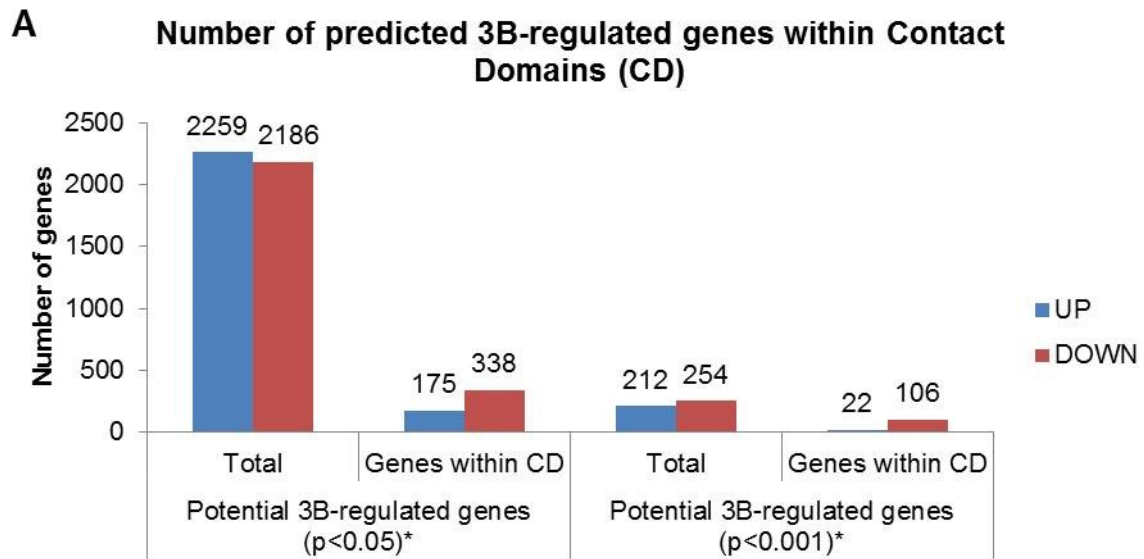


Figure 4.21. Contact Domains co-localising with both EBNA3B enriched peaks and potential EBNA3B-regulated genes.

(A) Total number of predicted EBNA3B-regulated genes* that are found within contact domains which contain EBNA3B enriched peaks. (B) Number of contact domains which contain EBNA3B enriched peaks and co-localise with potential EBNA3B-regulated genes. These are grouped into contact domains overlapping with a single predicted EBNA3B-regulated gene, contact domains containing multiple genes that may be regulated in the same direction (repressed or induced) and contact domains containing multiple genes that may be regulated in opposite ways (repressed and induced). *Predicted EBNA3B-regulated genes are split into 2 groups according to the p-value ($p < 0.05$ or $p < 0.001$) of the differences between gene expression in 3BKO and WT LCLs in the microarray done by White *et al.* (2010).

4.5.10 EBNA3B peaks which localise within the same Contact Domains as their target genes exhibit similar co-localisation patterns with histone modifications and transcription factors, regardless of the nature of the gene regulation.

Next, I examined the chromatin landscape of EBNA3B peaks that are associated with EBNA3B-regulated genes in order to ascertain whether any histone modifications can be linked to gene regulation. 269 peaks were found in contact domains comprising only repressed genes, 93 were in domains encompassing induced genes, and 75 peaks were in domains with both. The top 4 histone modifications associated with these peaks are identical to those of total EBNA3B peaks and no differences were detected between peaks found in contact domains comprising induced, repressed or both induced and repressed genes (Figure 4.22).

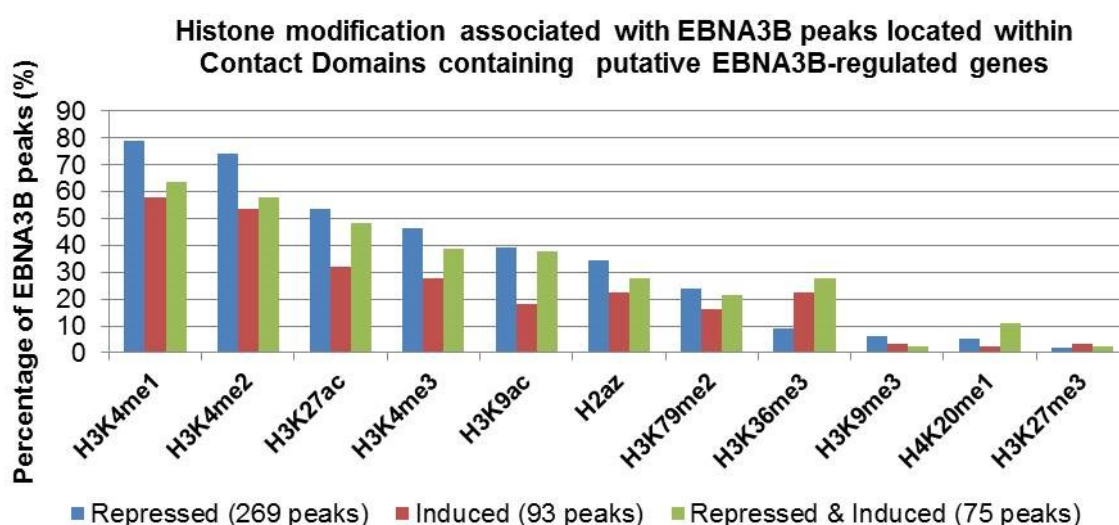


Figure 4.22. Histone modifications associated with EBNA3B peaks in Contact Domains.

In order to determine whether the regulation of EBNA3B target genes in the contact domains is influenced by transcription factors located at the sites of EBNA3B binding, I compared the TF binding at EBNA3B peaks located within these contact domains. The 20 TF with the highest co-occupancy with EBNA3B peaks in these domains are similar regardless of the type of regulation by EBNA3B (Figure 4.23).

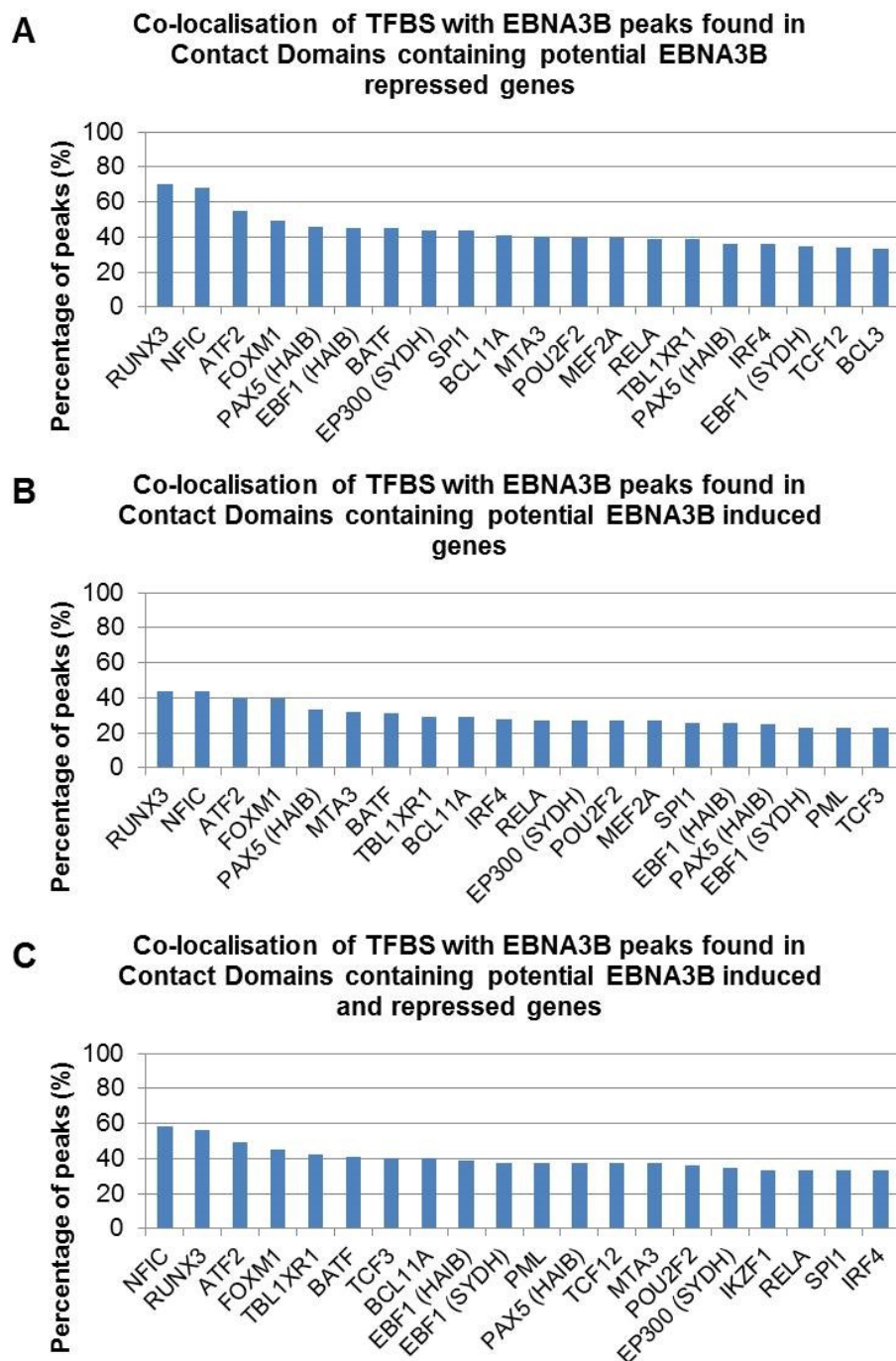


Figure 4.23. Co-localisation of TFBS with EBNA3B peaks found in Contact Domains
Co-localisation of TFBS with EBNA3B peaks found within Contact Domains containing only (A) potential EBNA3B-induced genes, (B) potential EBNA3B-repressed genes & (C) potential EBNA3B-induced and repressed genes ($p < 0.05$, White *et al.* 2010). A total of 93 peaks are found within Contact Domains containing only EBNA3B-induced genes, 269 peaks are found within Contact Domains containing only EBNA3B repressed genes and 75 peaks were found within Contact Domains containing both EBNA3B-repressed and induced genes. The percentages of the peaks found within these Contact Domains which co-localised with specific TFBS are reported. Only the top 20 TFBS are shown in each graph.

4.6 List of genes identified as rapidly targeted by EBNA3B and with EBNA3B binding sites within the same contact domain.

White *et al.* (2010) published a list of 54 genes (Appendix C) that have significantly altered expression – defined as $p < 0.001$ and more than 2-fold change – from EBNA3B deletion in both LCLs and EBV-infected BL31s. Of these genes, 32 are upregulated and 14 are downregulated in both 3BKO LCLs and 3BKO BL31s as compared to WT B95.8 infected LCLs and BL31s, whereas 8 are downregulated in 3BKO LCLs, but upregulated in 3BKO BL31s (Appendix C). No genes were found to be significantly upregulated in 3BKO LCLs and downregulated in 3BKO BL31s respectively. Where the genes are differentially regulated in opposite directions, it was proposed by White *et al.* (2010) that this is because LCLs and BL31s represent different B-cell populations – LCL representing activated naïve B-lymphoblasts and BL31s representing B-cells which have undergone germinal centre reactions. Since the 54 genes are selected with very high stringency, and their expression are affected by the absence of EBNA3B in two different cell types, there are strong indications that these are *bona fide* EBNA3B target genes. In total, the TSS of 25 of these genes are found within Contact Domains harbouring EBNA3B peaks, signifying that these genes are within the same looped chromatin as EBNA3B-bound regions (Table 4.5). Of these 25 genes, 21 have increased expression in 3BKO LCLs and 3BKO BL31s (hence are likely to be EBNA3B-repressed genes), 3 have reduced expression in 3BKO LCLs and 3BKO BL31s (EBNA3B-induced), and 1 gene (*TNFSF10*) is upregulated in 3BKO LCL but downregulated in 3BKO BL31s.

In order to study the gene regulatory functions of EBNA3B, and examine the biological consequences of this regulation, we have decided focus on specific *bona fide* EBNA3B-regulated genes which have known biological functions. In total, 7 of the 25 genes identified in Table 4.5 have been tested using the Dox-inducible EBNA3B system (see section 3.6.6), where 4 genes (*LILRA4*, *LAIR1*, *TERT*, *TNFSF10*) were deemed definite EBNA3B-regulated genes, and the regulation of the other 3 genes (*ADARB1*, *CXCR4*,

CXCL10) cannot be confirmed (Table 4.5). Of the definite EBNA3B-regulated genes, I have chosen to examine *TERT* and *TNFSF10* due to their potential physiological roles on cell senescence and apoptosis.

Table 4.5. EBNA3B-regulated genes which are found within the same contact domain (CD) as EBNA3B peaks.

Genes which have already been tested with the Dox-inducible EBNA3B system (see section 3.6.6) would be categorised as Definite, Probable or Not confirmed EBNA3B-regulated genes. Fold-difference and the associated p-value of relative gene expression in 3BKO and WT LCLs are reported. Data are retrieved from White and colleagues (2010).

EBNA3B-regulated genes within same CD as EBNA3B peaks			Response to EBNA3B-induction
Exon microarray: p<0.001 and >2-fold (25 genes)			(Dox-inducible system)
Regulation	Gene Symbol	ENSEMBL ID	Regulation
Up in LCL-3BKO and BL31-3BKO	ADARB1	ENSG00000197381	Not confirmed
	CXCR4	ENSG00000121966	Not confirmed
	ECE1	ENSG00000117298	-
	FAM107B	ENSG00000065809	-
	GPM6A	ENSG00000150625	-
	HHEX	ENSG00000152804	-
	INPP5F	ENSG00000198825	-
	KLHL6	ENSG00000172578	-
	LAIR1	ENSG00000167613	Definite
	LILRA4	ENSG00000239961	Definite
	MICAL3	ENSG00000243156	-
	PKIG	ENSG00000168734	-
	PRKCQ-AS1	ENSG00000237943	-
	PXK	ENSG00000168297	-
	RP11-13P5.1	ENSG00000224478	-
	SLC23A2	ENSG00000089057	-
	SLFN11	ENSG00000172716	-
	TAGAP	ENSG00000164691	-
	TERT	ENSG00000164362	Definite
	WWC3	ENSG00000047644	-
	XXbac-B461K10.4	ENSG00000093100	-
Down in LCL-3BKO and BL31-3BKO	CXCL10	ENSG00000169245	Not confirmed
	METTL7A	ENSG00000185432	-
	MIR22HG	ENSG00000186594	-
Down in LCL-3BKO and up in BL31-3BKO	TNFSF10	ENSG00000121858	Definite

4.7 Overview

To briefly summarize, I have generated epitope-tagged EBNA3B LCLs which have been used successfully to immunoprecipitate EBNA3B (section 4.2 - 4.4). Using one such LCL, ChIP-seq was performed and the results indicated that EBNA3B associates genome-wide in 1931 regions, which is considerably less than the number of EBNA3A and EBNA3C sites identified in LCLs (section 4.5). I found that where EBNA3B peaks co-localise with other EBNA3s, it is mainly with EBNA3C and EBNA2 (section 4.5.3). EBNA3A alone rarely occupies the same binding regions as EBNA3B, and any interactions with EBNA3B are mostly in the presence of EBNA3C (section 4.5.4). All these data suggest that EBNA3C often plays an important role in the gene regulatory function or recruitment of EBNA3B.

The percentages of peaks associated with histone modifications broadly used to define chromatin state were highest where EBNA3B is also co-localizing with all the other EBNA3s (section 4.5.6). These peaks were also most highly associated with enhancer elements. Furthermore, several of the transcription regulators which co-occupy binding regions of EBNA3B have also been shown to interact or co-localise with EBNA3A, EBNA3C and/or EBNA2 (section 4.5.8). All these observations are consistent with existing models that describe the EBNA3s and EBNA2 binding to distal enhancer regions and participating in an extensive interplay of interactions with transcription factor complexes which tether them to DNA, altering gene expression through histone modifications and chromatin remodelling (further discussed in Chapter 6).

My findings have also revealed that EBNA3B on its own binds mainly at transcribed or repressed segments of the genome, and it is also associated with transcriptional elongation-related H3K36me3 marks which are commonly found in gene bodies (sections 4.5.6 and 4.5.7). Comparatively, unique EBNA3B binding sites also have a lower tendency to reside in regions occupied by our list of histone modification marks or transcription factors. All these hint at EBNA3B having a separate mechanism of gene regulation in addition to its interactions with the other EBNA3s and demonstrates that the binding regions of EBNA3B are

both quantitatively and qualitatively different from the other EBNA3s. However, further investigations will have to be done to confirm these speculations.

ChIP-seq, Hi-C analyses and microarray studies inform us of possible interactions with chromatin and possible genes that are regulated. To extend these data and confirm some of the findings, further experiments need to be done to dissect the regions in question and try to answer the question of how EBNA3B can function and interact with the other EBNA3s in regulating specific genes. Comparing the data from Rao *et al.* (2014) with our lab's ChIP-seq analyses and the Dox-inducible EBNA3B system explored in Chapter 3, I compiled a list of genes that could be worthy of further investigations (section 4.6). In Chapter 5, with a particular focus on telomerase reverse transcriptase gene (*TERT*) – in order to move from speculation to experimentally testable hypotheses –, I describe my initial examination of the role of EBNA3B in the transcriptional regulation of its target genes.

5 Focused analysis of EBNA3B regulated genes to validate microarrays and ChIP-seq and provide some insight into mechanisms involved

5.1 Introduction

An intriguing property of EBNA3B is that phenotypically it appears to have the opposite effect on the host cell to the other members of the EBNA3 family. EBNA3A and EBNA3C are oncogenes which are essential for B-cell transformation *in vitro* – they increase the efficiency of proliferation and survival. In contrast, EBNA3B is not only inessential for transformation, it might even reduce proliferative capacity and *in vivo*, has tumour suppressive properties (White *et al.*, 2012). Moreover mutants of EBNA3B have been found in human EBV-associated B-cell lymphomas consistent with it acting as a tumour suppressor (White *et al.*, 2012). Based on the lists generated from the Dox experiments and ChIP-seq analyses in the previous two chapters, I selected several EBNA3B candidate target genes for further analyses. These genes have interesting physiological functions involved in genomic stability/cell senescence or apoptosis/cell survival which could contribute to EBNA3B's apparent role as a tumour suppressor. These include *TERT*, encoding telomere reverse transcriptase – the enzyme responsible for the maintenance and elongation of telomeres; *TNFSF10* (a.k.a. *TRAIL*), encoding an apoptosis-inducing cytokine; the genes encoding its receptor and decoy receptor, *TNFRSF10A* (a.k.a *TRAILR1*) and *TNFRSF10D* (a.k.a *TRAILR4*) respectively. Although *TRAILR1* and *TRAILR4* were tested, *TRAILR1* was found to be a probable EBNA3B-regulated gene and *TRAILR4* was not identified as an EBNA3B-regulated gene in the Dox system (section 3.6.6). However, my ChIP-seq analysis identified an EBNA3B binding site within the *TRAILR1* locus. In addition, the fact that *TRAILR4* has a genomic location just downstream of *TRAILR1* and within the same contact domain suggested they might be co-regulated and made both genes intriguing candidates to study.

A preliminary, but relatively detailed, examination of EBNA3B repression of *TERT* will be reported in this chapter. Time restrictions permitted only confirmation that EBNA3B is involved in the regulation of the TRAIL-family genes.

5.2 TLDA card study of the DOX system is validated using RT-qPCR – *TERT*, *TRAIL*, *TRAILR1* & *TRAIL R4*

To begin my analyses, I first confirmed the TLDA findings by designing primers for the genes of interest and running a qPCR reaction with the same cDNA samples (from donors D1 and D13) used for the array card. The results were largely consistent with the array card findings: that is *TERT* and *TRAILR1* were robustly repressed by EBNA3B whereas *TRAIL* was robustly induced. No apparent changes in gene expression were observed for *TRAILR4* (Figure 5.1). Further analyses from this point on are separately described in two sections: the first section will describe the experiments focusing on the regulation of *TERT*, and the next will touch on *TRAIL* and its associated receptors.

5.3 Detailed study of the regulation of *TERT* by EBNA3B

5.3.1 Primary B-cell infections demonstrated the kinetics of *TERT* regulation

In our Dox-inducible EBNA3B system, effects on gene repression/activation were determined from a 4-day time course with ectopic (over-)expression of EBNA3B (section 3.6.6). In order to determine the kinetics of EBNA3B regulation during the normal infection process, primary B-cells isolated from peripheral blood were infected with WT (B95.8 BAC), 3AKO, 3BKO, 3CKO recombinant EBVs and their respective revertants. Briefly, CD19+ primary B-cells were isolated from human peripheral blood using magnetic sorting. FACs analyses determined that more than 95% of these cells were CD20+ (Figure 5.3), indicating a high purity of B-cells in the population. In each case 3×10^6 cells were infected with 30 GRU of viruses and RNA were extracted every 5 days for a period of 30 days. The RNA was converted to cDNA. cDNA corresponding to *CXCL10*, *ITGAL* and *p16^{INK4A}* – which are known to be EBNA3A, EBNA3B and EBNA3C-regulated genes respectively – were assayed

as controls to verify that the viruses were behaving in an anticipated manner during the outgrowth.

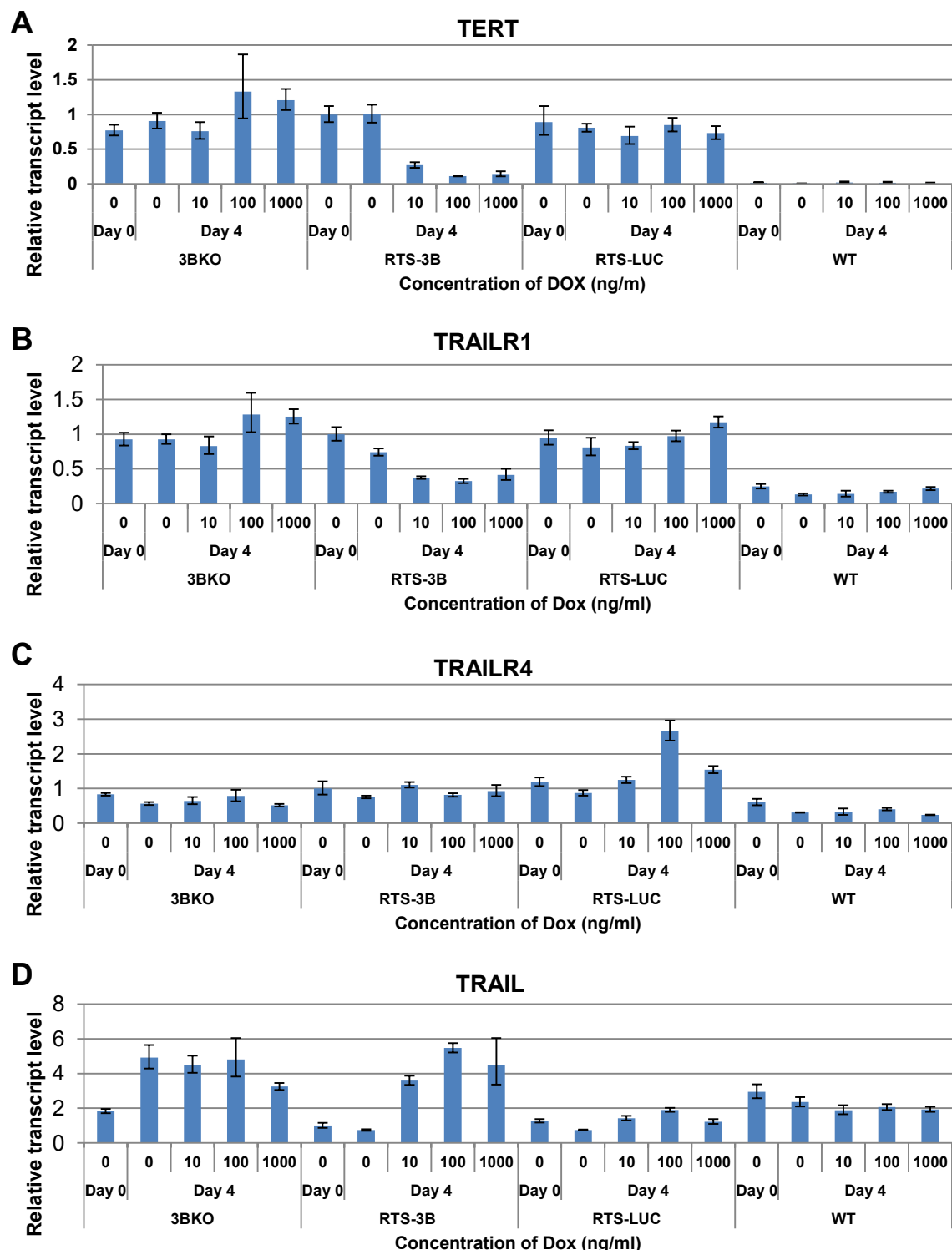


Figure 5.1. Verification of TLDA findings using qPCR analyses of (A) *TERT*, (B) *TRAILR1*, (C) *TRAILR4* and (D) *TRAIL* mRNA transcripts.

Quantitative-PCR of cDNA samples from the Dox induction experiments was done to verify the array card findings for *TERT*, *TRAILR1*, *TRAILR4* and *TRAIL*. The data are normalized to *GNB2L1* and *RPLP0* housekeeping genes and standardized to Day 0 transcript levels in RTS-3B EBNA3BKO LCLs. Mean $\pm$ SD of transcript levels from one qPCR experiment (Donor 13) are shown for each gene.

Expression of *TERT* is robustly activated in 3BKO LCLs as early as 5 days post infection, whereas in wild-type and revertant cell lines, this remained at around the levels of uninfected B-cells (Figure 5.2). This is consistent with results obtained from microarray (White *et al.*, 2010), and the array card experiments (section 3.6.6), showing that *TERT* is probably a *bone fide* EBNA3B-repressed gene. Although activation of *TERT* gene was seen in 3CKO LCLs, compared to 3BKO LCLs, this activation is very modest. This suggests that EBNA3C may be playing a minor role in the repression of *TERT* in LCLs; however the repression seems to be more robustly maintained by EBNA3B. Importantly, it is clear that other EBV-related factors (such as EBNA2 and LMP1) are involved in the activation of *TERT*, and EBNA3B plays a major role – either direct or indirect – in antagonising this activation (see Discussion, sections 6.3.1 and 6.3.5.1)

All the data reported above and in previous chapters have strongly and consistently supported the notion of *TERT* as an EBNA3B repressed gene. Furthermore, the *TERT* protein has an important physiological function in telomere maintenance (see section 1.8.1). Repression of this gene could have repercussions on the mortality of EBV-infected B-cells by promoting cell senescence and therefore limiting proliferative life-span. This could contribute towards EBNA3B's function as a tumour suppressor. For these reasons, I focused on the regulation of *TERT* in further ChIP experiments in order to elucidate the mechanism by which EBNA3B acts on this gene.

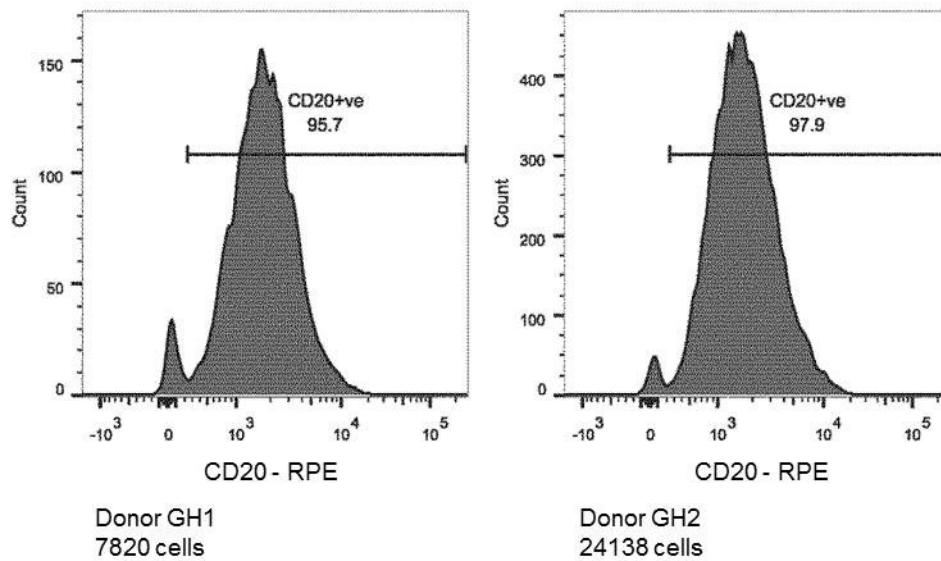


Figure 5.3. High yield of CD20+ B-cells was obtained using magnetic sorting for cells expressing CD19+ marker.

Primary B-cells were purified from blood obtained from two anonymous donors (GH1 & GH2) using magnetic sorting for CD19. FACS analysis was done to verify that more than 95% of these cells express the B-cell marker CD20.

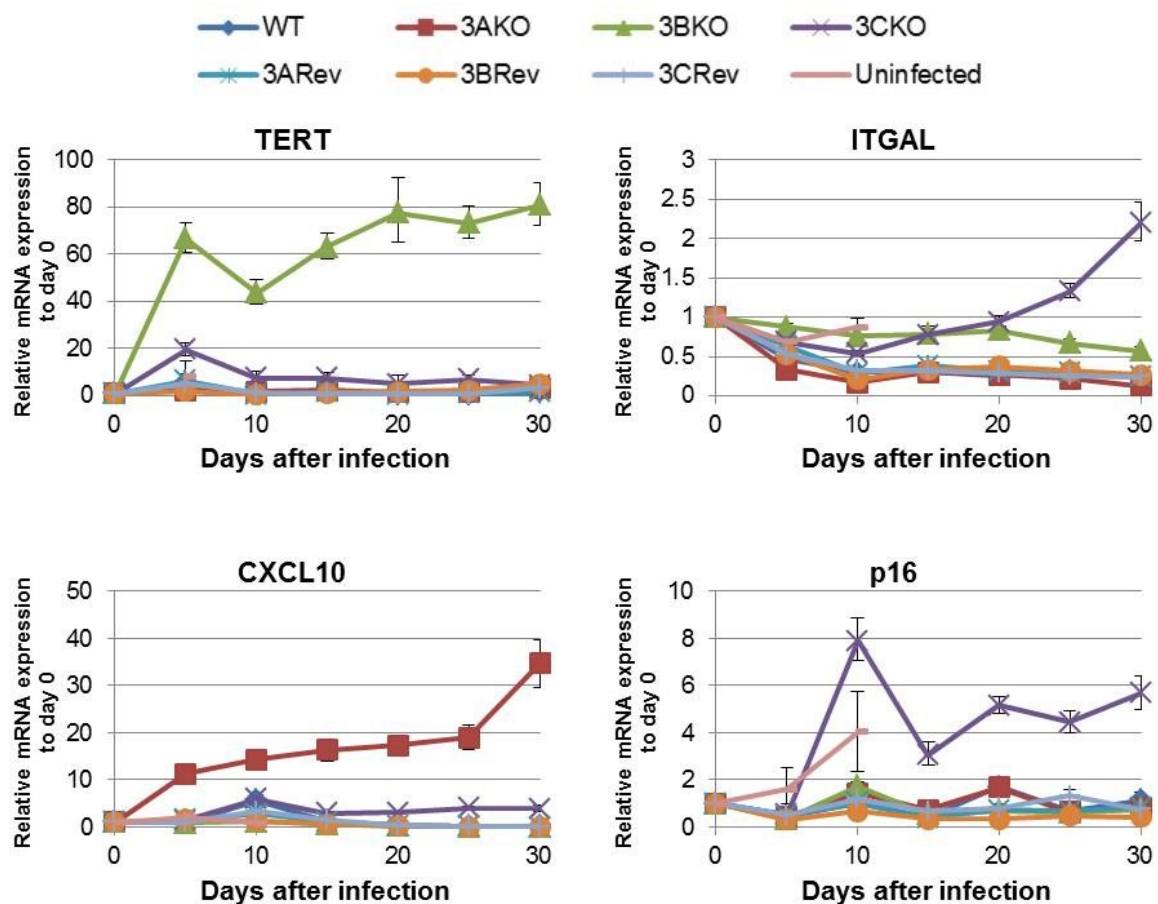


Figure 5.2. Outgrowth of recombinant EBV-infected primary B-cells.

Recombinant EBV (3AKO, 3BKO, 3CKO, and their respective revertants and WT) were used to infect primary B-cells pooled from two donors (GH1+2). The mRNA expression of *TERT* was tracked every 5 days post infection for 30 days. *CXCL10*, *ITGAL*, *p16* were used as positive controls for EBNA3A, EBNA3B and EBNA3C regulated genes respectively. Data were analysed in technical triplicates which were normalized to *GNB2L1* and *RPLP0* housekeeping genes and calibrated to day 0.

5.3.2 Binding of EBNA3s upstream of hTERT confirmed by ChIP-qPCR

In addition to the robust data showing that EBNA3B plays a role in modulating the expression of *TERT*, data from the ChIP-seq analyses have also revealed EBNA3B enrichment at a site that is approximately 17kb upstream of the *TERT* promoter (Figure 5.4A and Appendix D for pile-up of reads). This is found within the same contact domain [as defined by Rao *et al.* (2014)] as the *TERT* promoter (Figure 5.4A). On further examination, EBNA3A, EBNA3C and EBNA2 are also enriched at the same site as EBNA3B. There is also an EBNA3A, EBNA3C and EBNA2 enriched region upstream of this site (~20kb from *TERT* TSS), although EBNA3B does not seem to be enriched at this more distal site (Figure 5.4A). It is unclear whether either of these sites is involved in the regulation of *TERT* although these two regions coincided with a high incidence of H3K27Ac (marker of active enhancers) in GM12878 that expresses EBNA3B. This is consistent with observations from our ChIP-seq analyses that EBNA3B binds to enhancer sequences when it co-localises with the other EBNA3s and EBNA2 (Chapter 4, section 4.5.6). Consistent with these data, using each of our epitope-tagged EBNA3 LCLs derived from Donor 11, ChIP-qPCR confirmed binding for the EBNA3s at these regions (Figure 5.4B). This was repeated and validated in FS3B and 3C-TAP LCLs derived from Donor 13 (Appendix E).

5.3.3 Regulation of hTERT by EBNA3B may involve histone modifications

Next, I examined the histone modifications around the transcriptional start site (TSS) of *TERT* in order to determine whether the presence of EBNA3B alters the deposition of active or repressive histone marks. Due to the GC-rich region at the TSS, several sets of primers were designed for the TSS, however, none amplified the region efficiently in my hands. Hence, it was only possible to study the regions that are approximately 2kb upstream and downstream of the TSS.

It was previously reported that *TERT* transcripts are expressed at much higher levels in 3BKO LCLs than in WT and revertant cell lines derived from the same donor PBLs (White *et al.*, 2010). I repeated the experiment on LCLs obtained from Donors 11 and 13 and was able to reproduce these results (Figure 5.5). Correspondingly, the chromatin landscape at the TSS of *TERT* corresponds with its expression in these cell lines. In WT and 3BRev LCLs, the repressive H3K27me3 mark was enriched in regions 2kb both upstream and downstream from TSS of *TERT*, whereas histone modifications associated with gene activation (H3K4me3 and H3K9Ac) and active enhancers (H3K27Ac) were absent (Figure 5.6A-C). The reciprocal result was observed in 3BKO LCLs. In addition, slight enrichment of phosphorylated serine 5 residue of the C-terminal domain (CTD) of RNA polymerase II (PolII S5) – an indication of transcriptional initiation – was found 2kb and 1kb upstream of the TSS of 3BKO LCLs, but absent in WT and 3BRev cells, which is consistent with the observation that the transcription of *TERT* is repressed in the presence of EBNA3B (Figure 5.7A).

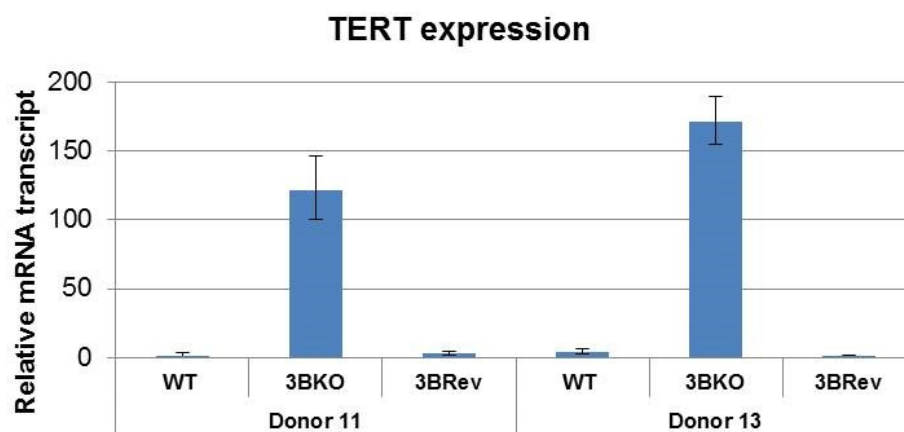


Figure 5.5. Levels of mRNA transcripts found in WT, EBNA3BKO and 3BRev LCLs derived from Donor 11 and Donor 13.

Data are normalised to *GNB2L1* and *RPLP0* housekeeping genes and expressed as fold-change over D11 WT LCLs. Error bars represent standard deviation from technical triplicates of one experiment. .

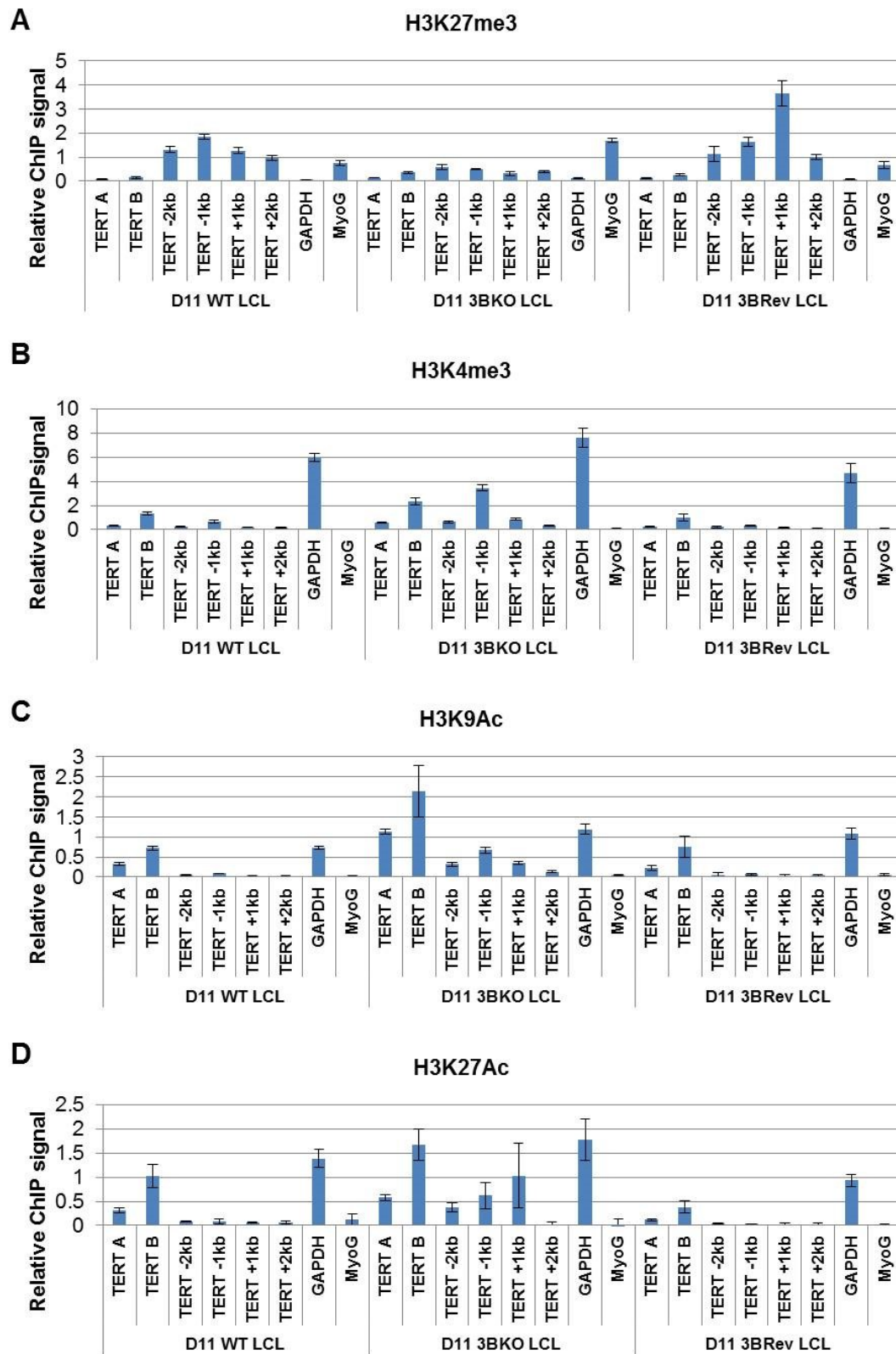


Figure 5.6. Enrichment of (A) H3K27me3 (B) H3K4me3 (C) H3K9Ac & (D) H3K27Ac at *TERT* promoter and EBNA3B binding sites.

ChIP for chromatin marker of repression, H3K27me3, and markers of activation H3K4me3, H3K9Ac and H3K27Ac near the transcriptional start site of *TERT* (± 1 or 2kb) and at the EBNA3 binding sites (TERT A and B). *GAPDH* and *MyoG* promoters are used as controls for active and repressed gene regulations respectively. ChIP signal is calculated relative to 2.5% input DNA. Error bars represent mean $\pm$ SD of technical triplicates from one experiment.

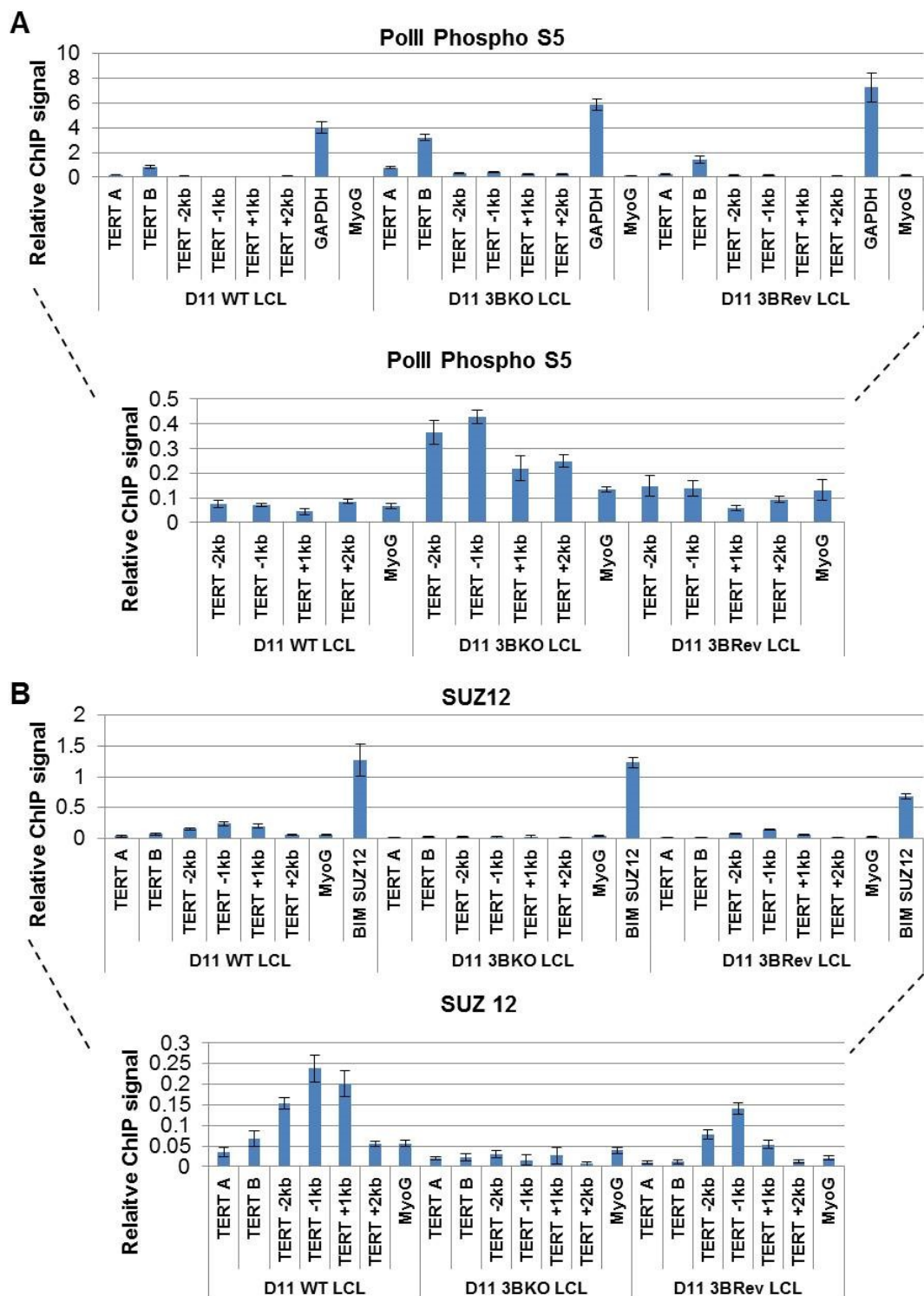


Figure 5.7. ChIP experiments showing (A) PolII S5 phosphorylation, and (B) SUZ12 at TERT promoter and EBNA3 binding sites.

(A) *GAPDH* promoter is used as a positive control for PolII S5 phosphorylation. In the sub-figure this sample is removed as *GAPDH* is highly actively transcribed and the enrichment of PolII S5 at this site dwarfs all signals at the highly repressed *TERT* promoter regions. **(B)** *BIM* promoter is used as a positive control for SUZ12 binding and the ChIP signal is removed in a second sub-figure for a clearer comparison at the *TERT* promoter regions between the cell lines. ChIP signal is relative to 2.5% input DNA.

It has been shown by our group and others that EBNA3A and EBNA3C can in some circumstances regulate target genes by recruiting polycomb repressive complex 2 (PRC2) components and directing the tri-methylation at H3K27 (Harth-Hertle *et al.*, 2013; McClellan *et al.*, 2012; Paschos *et al.*, 2012; Skalska *et al.*, 2010; see Introduction 1.7.2). In order to investigate whether EBNA3B-mediated repression of *TERT* involves the PRC2 complex, ChIP analyses using an antibody directed against SUZ12, an essential component of the PRC2 complex, were performed. As expected, SUZ12 is enriched around the TSS of *TERT* (-2kb to +1kb) in WT and 3BRev LCLs, but is absent in 3BKO LCLs (Figure 5.7B), directly reflecting the H3K27me3 levels. The presence of EBNA3B appears to be necessary for the deposition of H3K27me3 by the PRC2 complex at the *TERT* TSS. However, it should be noted that EBNA3B does not appear to bind at the TSS and neither SUZ12 nor H3K27me3 are found at the EBNA3B binding site (TERT A) – this will be further discussed below.

Examination of the EBNA3B binding region (TERT A) revealed that – in the absence of EBNA3B – activation-associated modifications H3K9Ac and H3K27Ac are enriched at this site (the latter only modestly) (Figure 5.6C & D). There appears to be no difference between the H3K4me3 enrichment at the EBNA3B binding region in 3BKO, WT and 3BRev LCLs. Similarly, H3K9Ac and H3K27Ac are also enriched at the EBNA3C/EBNA3A-only binding site further upstream (TERT B) in 3BKO LCLs (Figure 5.6C). The enrichment of these histone marks in 3BKO LCL suggests that the chromatin at these putative enhancer regions may be more open in the absence of EBNA3B, further relieving the repression of *TERT*. Hence, EBNA3B may not only have a role in recruiting PRC2 complexes to the TSS of *TERT*, but may be preventing the activation of the distal enhancers and their subsequent action in inducing the expression of *TERT*. To test this hypothesis, various other experiments, including chromatin conformation capture (3C) will have to be performed. These experiments will be performed by Dr Quentin Bazot in our lab.

5.3.4 Physiological implications of TERT regulation by EBNA3B

We were unable to identify any antibodies (commercial or otherwise) that could reliably detect the TERT protein in LCLs. However, one well-documented function of TERT is its catalytic activity in the telomerase complex to elongate telomeres and prevent telomere erosion and as a consequence, cell senescence (see section 1.8.1). Hence, in order to investigate the impact of *TERT* repression by EBNA3B, I examined the enzymatic activity of telomerase in 3BKO and 3BRev LCLs using the TRAPeze® RT (S7710) kit from Merck Millipore (Figure 5.8).

Briefly, nuclear extracts from cells are incubated at 30°C to allow the telomerase enzyme to add 6bp telomeric repeats onto a substrate oligonucleotide. Next, Taq polymerase enzyme in the reaction mixture amplifies the extended products using primers specific for the substrate and fluorescein-labelled primers specific for the telomeric repeats. These fluoresce upon incorporation into the PCR product and the fluorescence is directly proportional to the amount of product generated, which is detected by a real time PCR machine. Standard curves using a TSR8 quantitation control template – an oligonucleotide with an identical sequence to the substrate primer and extended with 8 telomeric repeats – allows for the quantification of telomerase activity relative to the TSR8 amplification.

Initial analyses of LCLs generated from Donors 11 and 13 revealed that the telomerase activity is substantially higher in 3BKO than in 3BRev LCLs (Figure 5.9). This demonstrates that the repression of TERT at the transcriptional level by EBNA3B ultimately results in a lower telomerase activity in the host cell, and that the TERT upregulation in 3BKO cells produces a functional telomerase enzyme.

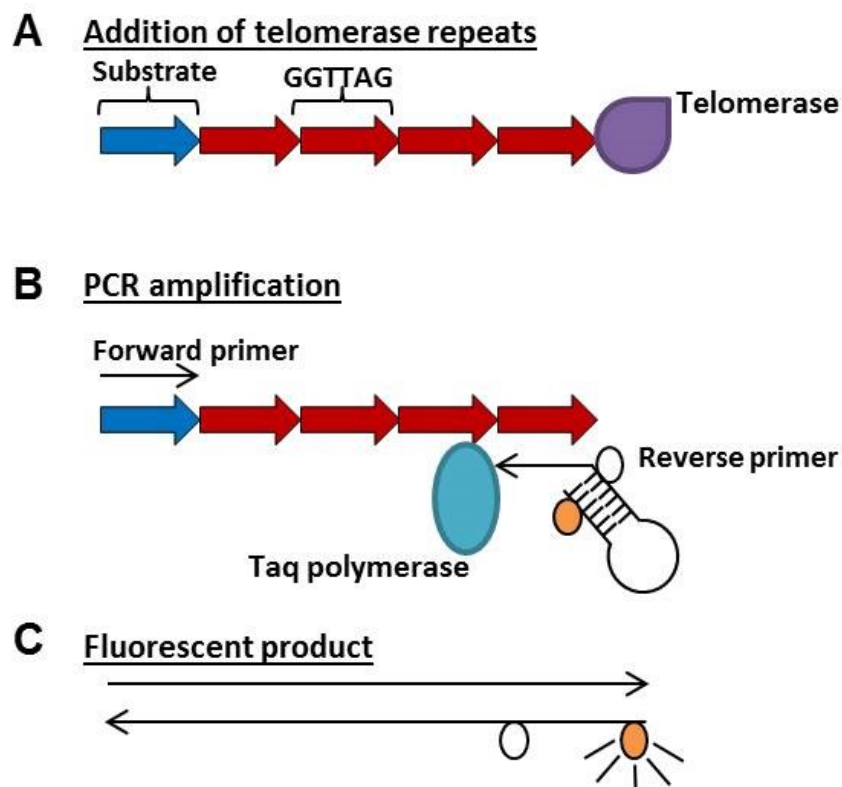


Figure 5.8. Detection of telomerase activity using the TRAPeze® RT (S7710) kit.

(A) Telomerase complex adds 6bp telomeric repeats (GGTTAG) to the substrate oligonucleotide at 30°C. **(B)** Taq polymerase performs PCR amplification by using forward primers complementary to the substrate and a reverse Amplifluor® primer consisting a sequence complementary to the telomeric repeats and an attached hairpin structure that contains a fluorophore (orange circle) which is brought into close proximity with the quencher (open circle). **(C)** On extension of the PCR product by Taq polymerase, the hairpin unfolds and the distance between the quencher and fluorophore increases, generating a fluorescent signal.

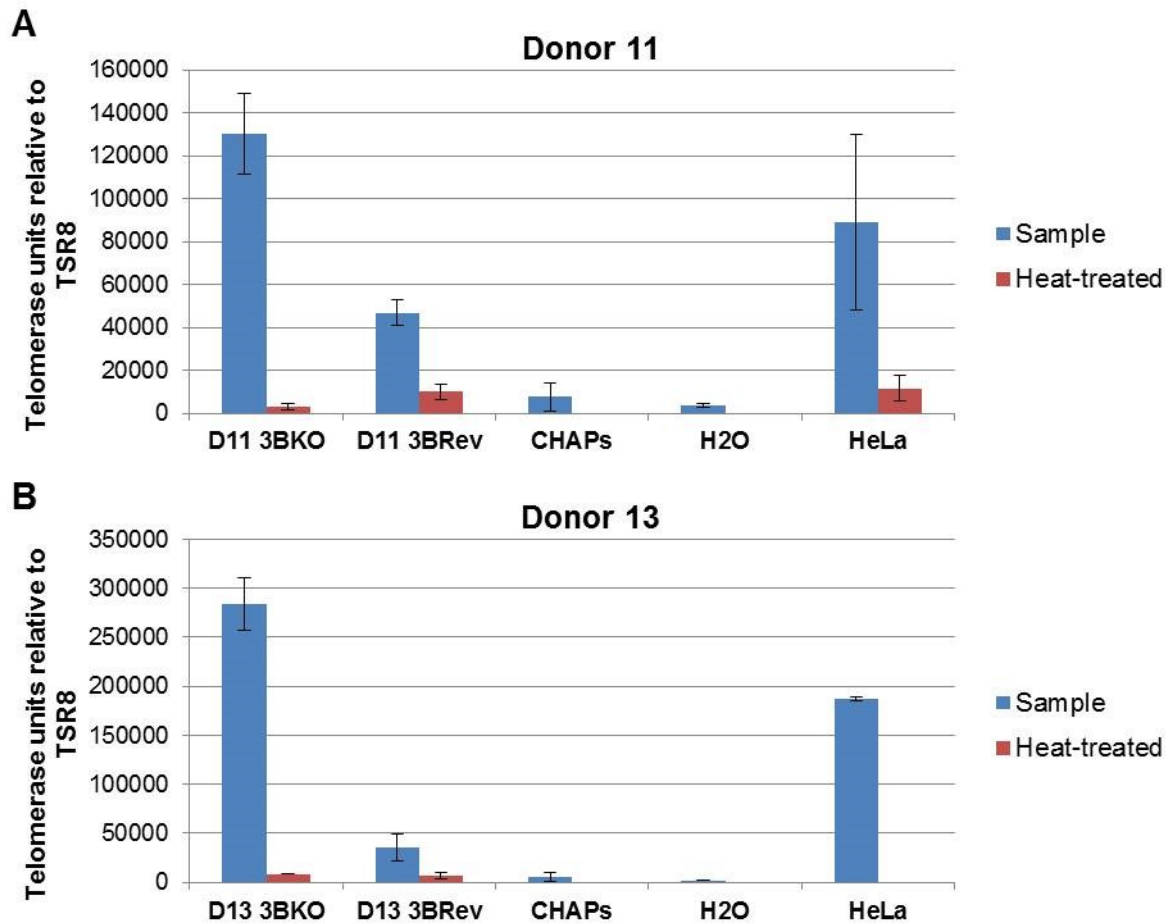


Figure 5.9. Telomerase activity in 3BKO and 3BRev LCLs of two different donor backgrounds.

Telomerase activity is measured by using the TRAPeze® RT Telomerase Detection Kit (S7710) from Merck Millipore. Heat inactivation of the telomerase complex in each sample was performed at 80°C as a negative control for non-specific activity in the telomerase assay. CHAPs lysis buffer and water were used as additional negative controls and HeLa cell extracts were used as positive controls for telomerase activity. Values are reported relative to the TSR8 quantitation control template and error bars indicate standard deviation of triplicate reactions.

In order to determine whether increased telomerase activity could have an impact on telomere length in dividing cells, in a preliminary experiment, I monitored the telomere length of WT, 3BKO and 3BRev LCLs for a period of more than 100 days. To do so, I used the multiplex q-PCR (MMQPCR) method developed by Cawthon *et al.* (2009) and optimised by Miss Lama Alzamil, a master's student from Dr White's lab. Briefly, primers designed to specifically amplify telomeres (tel) and a single copy gene (scg) – in this case human beta-globin – at different temperatures were used in a single qPCR reaction. The ratio of tel to scg (T/S ratio) is proportional to the average telomere length in the cell population.

Four outgrowth experiments were performed using WT, 3BKO and 3BRev EBV infections of primary B-cells. Two sets of LCLs (GH1+2 and GH3+4) were established from mixed donors (B-cells pooled from two different donors) and two sets (GH5 and GH6) were established from B-cells obtained from a single donor. In the single-donor experiments, recombinant 3BKO EBV obtained from two different producers (C14 and I3) were used in separate infections. Telomere length of LCLs increased over the first 20 days after infection, and appears to be maintained 100 days post infection in the mixed donor LCLs regardless of their EBNA3B status (Figure 5.10). Variations in telomere length between the single-donor cell lines appear between 40 and 60 days post infection, although there appear to be no clear correlation between the presence of EBNA3B and the maintenance or depletion of telomere length (Figure 5.10). To complete this experiment, a longer time-course experiment that also includes analyses of the *TERT* expression and telomerase activity over time is necessary. Moreover, it has been reported that the telomere length is maintained in the LCLs via the alternative lengthening of telomeres mechanism [(Kamranvar *et al.*, 2013); discussed further in Chapter 6]), which could be considered in future experiments.

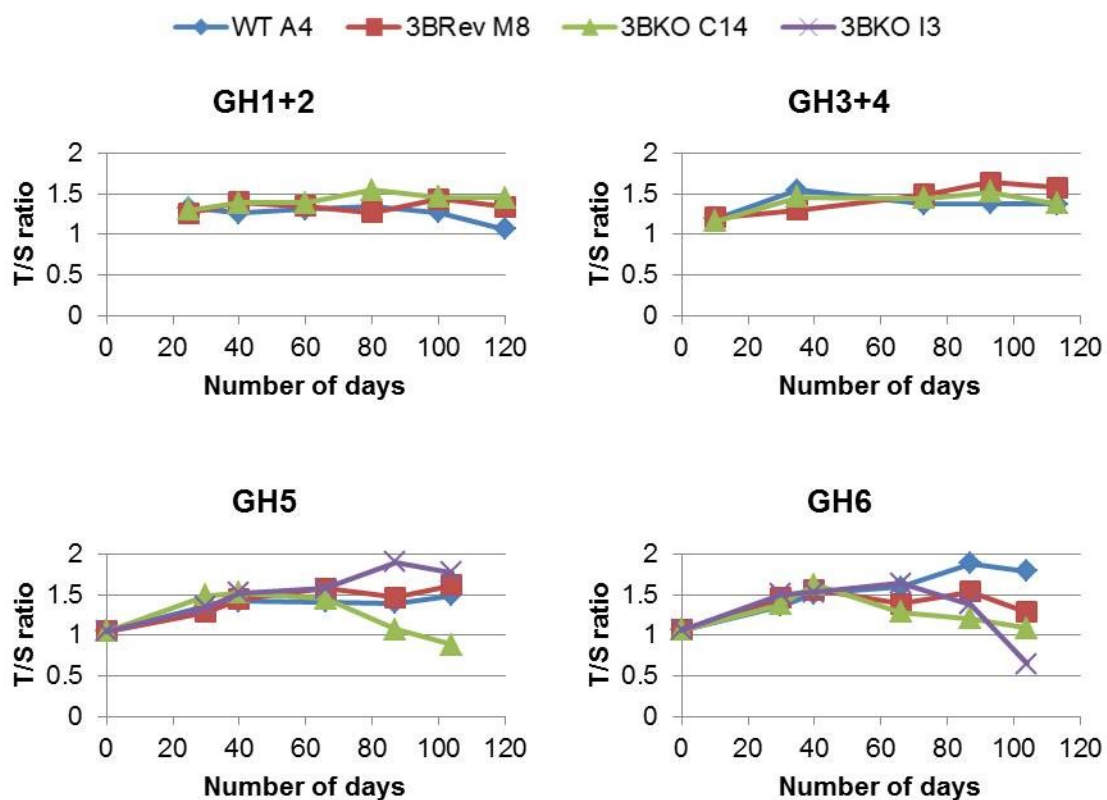


Figure 5.10. Telomere length during B-cell outgrowth of cells in 4 different genetic backgrounds.

B-cells pooled from either mixed donors (GH1+2 and GH3+4) or single donors (GH5 and GH6) were infected with WT, EBNA3BKO and 3BRev recombinant EBV and the telomere length tracked over a period of more than 100 days. For infection of GH1+2 donor B-cells, samples were taken on day 25, 40, 60, 80, 100 & 120. Similarly for the infection of GH3+4 donor B-cells, samples were taken on day 10, 35, 75, 95 & 115. Samples from the infection of B-cells obtained from donors GH5 and GH6 were done simultaneously on day 0, 30, 40, 67, 87 and 107.

5.4 Regulation of TRAIL, TRAILR1 & TRAILR4 by EBNA3B

In addition to studying the regulation of TERT by EBNA3B, I have also started examining the role of EBNA3B in regulating components of the TRAIL-regulated apoptotic pathway. RT-qPCR analyses confirmed our findings from the Dox-inducible EBNA3B experiments (section 5.2), that is, TRAIL mRNA is EBNA3B-induced and TRAILR1 mRNA is EBNA3B-repressed. No changes in TRAILR4 mRNA were detected on introduction of EBNA3B into the 3BKO LCLs.

Using material from the primary B-cell infection experiment described in section 5.3.1, I studied the kinetics of the regulation of *TRAIL*, its receptor *TRAILR1* and decoy receptor *TRAILR4* during the B-cell infection process. Interestingly, EBNA3B does not appear to play a vital role in regulating *TRAIL* expression in the first 30 days after infection, although EBNA3C may have an effect in its repression after 15 days (Figure 5.11). In contrast, EBNA3B appears to be involved in the repression of *TRAILR1* by EBV, since the expression of *TRAILR1* is activated by day 10 in 3BKO LCLs and continues on an upward trend until 30 days (Figure 5.11). A similar observation was made for *TRAIL R4*, where EBV represses the overall expression of this gene, but in 3BKO EBV infected B-cells, this repression is slightly relieved (Figure 5.11). This suggests that another EBV-related factor might contribute to the repression of this gene.

This primary B-cell infection experiment is largely consistent with results from the microarray (White *et al.*, 2010) and the array card experiments (described in Chapter 3) for *TRAILR1* and *TRAILR4* (microarray). This does not appear to be the case for *TRAIL*. Despite being a predicted EBNA3B-induced gene (both in the microarray and in the RTS Dox system), the differences observed between B-cells infected with 3BKO and others infected with wild-type phenotype EBV are not consistent across all of the infections, the caveat being that this experiment was only performed once with a single mixed donor background.

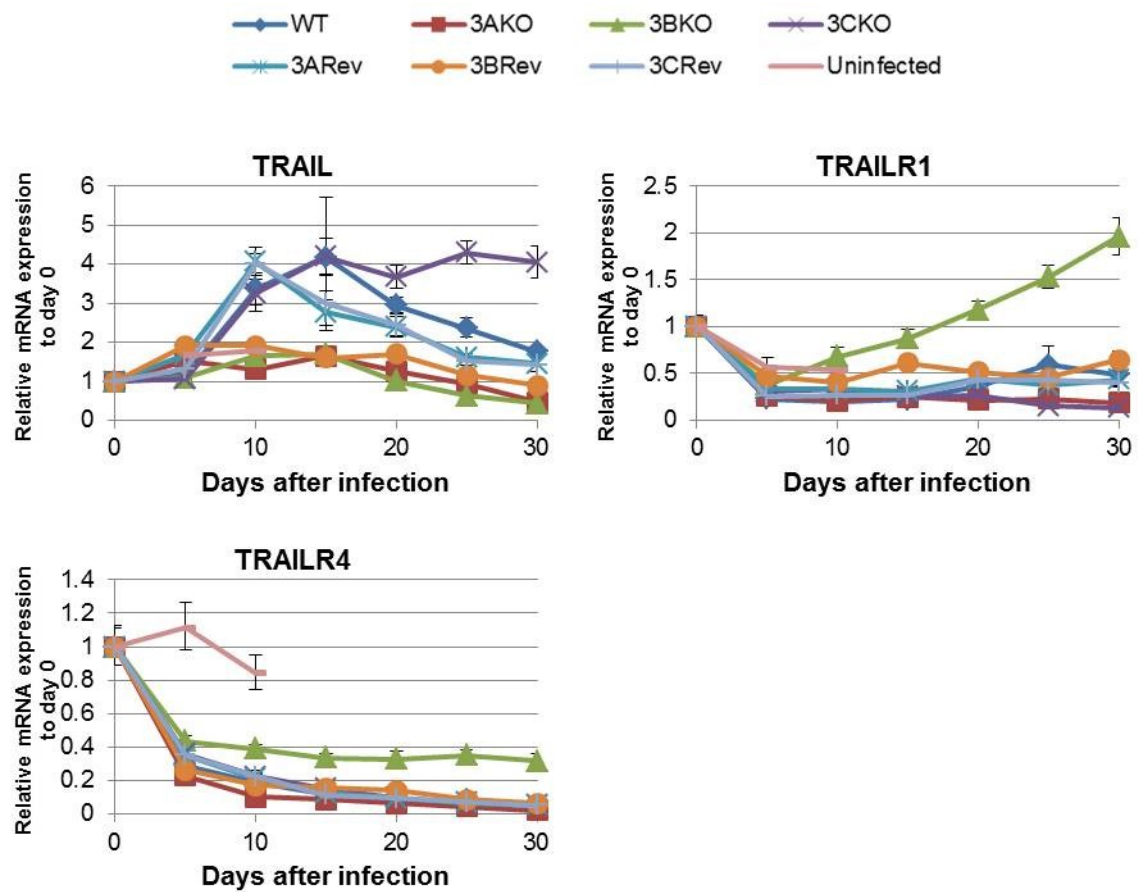


Figure 5.11. Outgrowth of recombinant EBV-infected primary B-cells –*TRAIL*, *TRAILR1*, *TRAILR4*.

The mRNA expression of *TRAIL*, *TRAILR1* & *TRAILR4* were tracked every 5 days post infection for 30 days. Experiment as described in Figure 5.2.

Examination of the *TRAIL* (a.k.a. *TNFSF10*) gene locus revealed that EBNA3B localises to a region within an intron of the gene, with two EBNA2 binding regions situated about 2kb and 4kb upstream of this EBNA3B binding site (Figure 5.12A and Appendix D). Similarly, EBNA3B (and also EBNA3A, EBNA3C and EBNA2) is enriched within the first intron of *TRAILR1* (*TNFRSF10A*) (Figure 5.12B and Appendix D). EBNA3B does not appear to localise at the *TRAILR4* (*TNFRSF10D*) gene, however this gene lies directly (~60kb) downstream of the TSS of *TRAILR1* (Figure 5.12B) and it is likely that this whole gene locus is co-ordinately regulated by EBNA3B through its binding site at *TRAILR1*. Interestingly, EBNA3C and EBNA2 both appear to bind in regions close to the TSS of *TRAILR4* and in the intergenic region between *TRAILR1* and *TRAILR4*, even though the B-cell outgrowth experiment does not suggest a role for EBNA3C in the regulation of this gene. Due to time constraints, these ChIP-seq binding data have not yet been experimentally verified by ChIP-qPCR in our lab. Future experiments will explore the EBNA3B binding, and any interaction with other EBNA3 and EBNA2 proteins at these loci. It will be interesting to determine whether EBNA3B regulates this *TRAILR1/R4* gene locus co-ordinately, and to elucidate the molecular mechanisms behind this putative regulation, which could be compared to the regulation of *TERT*.

5.5 Overview

To summarise, I have validated the results of the TLDA and ChIP-seq experiments, and found by infecting primary B-cells that *TERT* is very robustly repressed only when EBNA3B is expressed (section 5.3.1). I then investigated various histone modifications at the *TERT* promoter and the distal EBNA3B binding sites nearest to the locus in order to elucidate the function of EBNA3B in regulating the expression of *TERT*. I found that the repressive histone mark H3K27me3 is enriched at the promoter region of *TERT* in the presence of EBNA3B, which may be a result of the recruitment of PRC2 complex facilitated by EBNA3B (section 5.3.3). Similarly, a reduction of H3K9 acetylation could result from the recruitment of histone deacetylases (HDACs) by EBNA3B. Next, I began preliminary

experiments examining the biochemical impact of the repression of *TERT* by EBNA3B. I found that telomerase activity is higher in 3BKO LCLs than 3BRev LCLs generated from 2 independent donors (section 5.3.4). Finally, I explored the possibility of a repressed telomerase activity having an impact on telomere length and found that the telomere length appears be maintained up to 120 days post infection in mixed donor cell lines regardless of the presence of EBNA3B, and that variations in transformed cells derived from infection of single donor B-cells did not exhibit a clear trend that can be related to the EBNA3B status.

It is important to note that many of the results reported here are preliminary results and will have to be repeated before any firm conclusions can be made. The next chapter will discuss the implications of my findings, the limitations of the experiments and future projects that will examine the mechanism of action of EBNA3B on the regulation of *TERT* gene in greater detail.

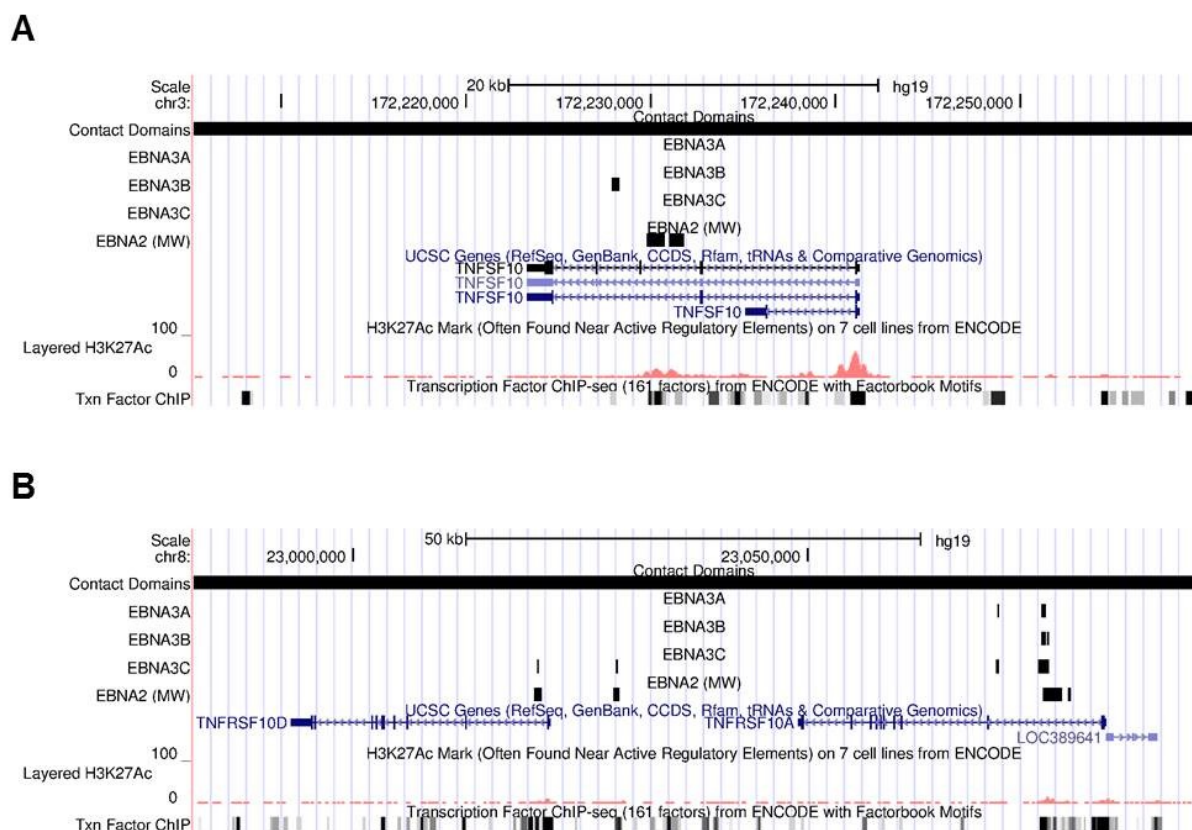


Figure 5.12. Figure showing the enrichment of EBNA3A, EBNA3B, EBNA3C and EBNA2 at the gene loci of (A) *TRAIL* and (B) *TRAILR1/R4*.

EBNA3A, 3B and 3C ChIP-seq data are obtained from the unpublished ChIP-seq analyses done by the Allday group. EBNA2 ChIP-seq data are obtained from McClellan *et al.* (2013) and the chromatin contact domains are defined by Rao and colleagues (2014).

6 Discussion

At the beginning of this study little was known of the function of EBNA3B in regulating host genes or its role in EBV-associated biology and lymphomagenesis. In Chapters 3 & 4, I have described different approaches adopted for the analysis of EBNA3B. This produced a list of genes that are not only differentially expressed with the complementation of EBNA3B-null cells with EBNA3B, but are also predicted to be in relatively close proximity to an EBNA3B binding site in the genome (section 4.6). While EBNA3B can bind directly within the promoter or even within its predicted target genes, it is more frequently found to bind at regions >100kb away from their TSS (our results in section 4.5.5; Wang *et al.*, 2015). We speculate that these two regions are brought into close proximity via complex chromosomal looping. A focused analysis of the regulation of *TERT* by EBNA3B is presented in Chapter 5, where emphasis is placed on the epigenetic changes at both the EBNA3B binding regions and at the TSS of *TERT*. In this chapter, I shall discuss the implications of my findings and the strengths and limitations of these analyses.

6.1 The Dox-inducible EBNA3B system can be used to identify bona fide EBNA3B target genes that are robustly and rapidly regulated

Since (in our hands) it was not possible to clone in the 4HT-sensitive ERT2 tag at either N- or C- terminals of EBNA3B without ablating EBNA3B or EBNA3C expression, we used the Dox-inducible RTS-3B system as an alternative to study the effects of complementing 3BKO LCLs with EBNA3B on possible EBNA3B target genes.

6.1.1 Magnetically sorted NGFR+ RTS-3B/LUC cells ensures high purity of cells expressing the RTS vector and have to be re-sorted at every time-point

Flow cytometric analysis showed that not all puromycin-selected RTS-3B cells responded to Dox by expressing the surrogate fluorescent marker, mCherry, and by extension, EBNA3B. Similar observations were made by Bornkamm *et al.* (2005) and Harth-Hertle *et al.* (2013) in their respective Dox-inducible pRTS cell lines. Magnetically sorting for

NGFR[±] cells partially overcomes this problem by ensuring that pure populations of either EBNA3B-expressing or EBNA3B-null cells are analysed, although it is interesting to note that the sorted populations of cells tend to fairly rapidly drift back to their original unsorted phenotypes (Figure 3.18) (i.e. they do not remain >90% responsive to Dox despite long-term maintenance in Dox). Since this phenomenon is more evident in populations which started with lower post-sort percentages of mCherry-positive cells, it is possible that the few Dox-unresponsive cells that survive post-sort may have selective advantage in proliferation and growth, hence overtaking the culture over time. This may not be surprising since the cellular machinery in Dox-induced cells will probably have to cope with over-expression of proteins, which in large amounts could be toxic. However, if this explanation were true, we would expect the culture to be completely dominated by Dox-unresponsive cells and not result in a return to stable levels of Dox-inducible cells.

The reversion of sorted cell populations towards a steady proportion that is Dox-sensitive suggests there is a constant and rapidly evolving heterogeneity in RTS-vector expression within the population which favours the shut-down of responsiveness to Dox within the sorted cell population until the population stabilises at a favourable “set-point”. This could be possible through a dynamic that resets regularly in any given cell and determines whether a cell is Dox-resistant or sensitive. Hence, for each of the cell lines, there would be a fixed chance that any cell will be Dox-responsive or non-responsive. While the mechanism of how this set point is determined and maintained is unclear, we speculate that a changing dynamic in the methylation of Ptetbi-1 or Eu-CAGp promoters during DNA replication may be involved. Whatever the reason(s) for this drift of RTS-vector expression, it is clear that the cells will have to be sorted at every time-point in order to harvest and analyse a “pure” population of EBNA3B-expressing cells and future experiments exploring regulation kinetics should take this into account.

6.1.2 The Dox-inducible RTS-3B experiments identifies genes that are robustly and rapidly repressed or induced in the presence of high levels of EBNA3B

While the exon-microarray studies provided a list of putative EBNA3B-regulated host cell genes (White *et al.*, 2010), the inducible RTS-3B system coupled with TLDA card analyses allowed us to refine this list by focusing on genes which additionally exhibit a rapid, robust and graded response to increasing amounts of EBNA3B in a single EBNA3B-null LCL background (see section 3.6.6). Thus, this system removes the potential confounding factor of differential expression in WT and 3BKO LCLs arising purely as an artefact of clonal selection (Heath *et al.*, 2012; Skalska *et al.*, 2010) during the outgrowth of the LCLs used in the exon-microarray studies. This additional information, in combination with preliminary data from our ChIP-seq analyses, gives us confidence in the choice of genes for further targeted analyses (in my case, *TERT*, *TRAIL*, *TRAILR1* and *TRAILR4*) and excludes genes in which regulation by EBNA3B is not confirmed.

Since *ITGAL* has been shown by our lab and others to exhibit features of a *bona fide* EBNA3B-regulated gene (McClellan *et al.*, 2013; White *et al.*, 2010, 2012) (see Introduction section 1.10), parameters of the experiment were determined and validated using *ITGAL* as a reference EBNA3B target gene, with the assumption that the regulation of most other genes by EBNA3B would follow similar kinetics. Although this might be the case for most EBNA3B target genes, genes which either are activated or repressed by EBNA3B via mechanisms that require a longer period than 4 days are excluded. Similarly, genes which have had been epigenetically silenced (during LCL outgrowth) through the formation of heterochromatin and/or DNA methylation (Ayyanathan *et al.*, 2003; Cameron *et al.*, 1999; Jones *et al.*, 1998) in such a way that cannot be readily reversed by complementation with EBNA3B are also excluded. Notably, Kalchschmidt *et al.* (2016) observed in 4HT-inducible EBNA3C-HT cell lines that although the initial repression of *COBLL1* gene by EBNA3C took place rapidly within 6 days, the subsequent reappearance of the COBLL1 protein after inactivation of EBNA3C by withdrawal of 4HT took at least 20 days. This suggests that

experiments with induction of EBNA3B for longer periods could provide further insights into the kinetics of gene regulation and identify genes that require longer periods for activation or repression.

During our initial validation of the system, we observed that the expression of two of the putative EBNA3B-regulated genes (*CD28* and *LAIR1*) appeared to drift in culture over a longer 23-day Dox induction of the RTS-3B 3BKO LCLs, regardless of the Dox treatment (Chapter 3, section 3.6.3). These analyses were done on unsorted cells, and a high degree of background noise from non-responsive cells within the Dox-treated population could be affecting the analyses. However, this does not explain the drift of gene expression in both populations of cells (treated and untreated). It is probable that continual puromycin selection for cells expressing the RTS-vector is promoting clonal expansion of some cells over others. In addition, Bornkamm *et al.* (2005) also noted that their pRTS-transfected BJAB cells lost responsiveness to Dox in long-term culture (5 months or more), and the background expression of the surrogate fluorescent marker of expression (eGFP) becomes higher in non-induced cells, which suggests that the bicistronic cassette expressing reverse tetracycline controlled transactivator rtTA2S-M2 and the silencer Tet-repressor-KRAB fusion protein tTSKRAB (section 3.6) could be unstable over time. Hence, our observations suggest that time-courses taking longer than a month (which is normally done in our lab for following expression kinetics in our conditional EBNA3A-HT or EBNA3C-HT cell lines) might be confounded by the instability of the responsiveness of the vector and gene expression in RTS-transfected cells, although this has not been formally tested.

With magnetic sorting, *LAIR1* is later shown to have a robust dose-response in a 4-day induction with Dox with the RTS-3B cells and response of *CD28* cannot be determined within 4 days. It could be worthwhile repeating the experiment with a longer time course to follow the kinetics of EBNA3B regulation of selected genes (such as *CD28*), but this time sorting the cells at every time point for a better representation of gene expression in Dox-responsive cells.

In summary, the RTS-3B system is an excellent tool for the preliminary selection of interesting target genes for further study, however it is important to note that genes that were disregarded as EBNA3B target genes in this study should not be immediately classified as non-target genes due to the reasons stated above. Future experiments including longer time-courses with lower dosages of Dox for expression of EBNA3B that more closely mimics physiological levels of EBNA3B could be considered for further analyses of the kinetics of EBNA3B-mediated gene regulation.

6.1.3 Variation exists between the response of different RTS-3B cell lines

The responses in gene expression after Dox induction can vary between RTS-3B cell lines, due to clonal differences or the different donor backgrounds. Cost and time constraints meant that we were only able to employ the RTS-system and array card analyses in cell lines derived from two donors – hence we were unable to perform rigorous statistical analyses to discriminate between changes in gene expression that were due to chance, or those that were mostly likely due to the effects of EBNA3B. Notably, there were several genes which responded to increasing levels of EBNA3B in only one of the two RTS-3B cell lines (*VAV3*, *CRLF2*, *CD72* & *ITGB1*) and one gene (*BAG3*) appeared to respond to EBNA3B in an opposing manner in each of the two RTS-3B cell lines (Chapter 3, section 3.6.6). Additionally, it appears that other genes (including *SMAD1*, *GAS7*, *MLL* and *SORT1*) cannot be properly analysed using this system, since the transfection and selection process seemed to have altered the expression of these genes even prior to the Dox induction. In order to overcome this problem, we have grouped the genes into 3 categories (Definite, Probable and Not Confirmed EBNA3B-regulated genes) using selection criteria as described in Chapter 3. In combination with our ChIP-seq data, the Dox-inducible EBNA3B system is a tool that permits us to rapidly analyse a list of possible EBNA3B target genes in order to hone in further analyses on those which exhibit striking features of being regulated by EBNA3B.

6.1.4 Conclusions and future work with the Dox-inducible RTS-3B system

We generated the Dox-inducible RTS-3B system with the intention of using it to study the kinetics of gene regulation by EBNA3B, and also to use it to identify genes that respond to complementation by EBNA3B in a single 3BKO cellular background. Although we have yet to validate and establish a reliable method for long-term observation of the kinetics of EBNA3B-mediated gene regulation, we have been able to use the RTS-3B system to identify genes that can be robustly regulated within 4 days of introducing EBNA3B into the 3BKO cells. Repeat experiments using cell lines from multiple donors and material from these could be used in future exon-microarray or RNA-seq analyses to generate a more comprehensive and reliable list of EBNA3B target genes. These can be compared to the data from existing microarray studies of WT and 3BKO LCLs. Using the RTS-3B system for longer-term analyses should also be done because the study of temporal sequence of events in gene regulation and their effects on cell proliferation have thus far revealed interesting models for the gene-regulatory functions of EBNA3A and EBNA3C (Bazot *et al.*, 2015; Harth-Hertle *et al.*, 2013; Kalchschmidt *et al.*, 2016b; Maruo *et al.*, 2003, 2006; Paschos *et al.*, 2012; Skalska *et al.*, 2010) (more details in Introduction section 1.7.4).

6.2 ChIP-seq has identified properties of EBNA3B binding regions that may be important for its function as a transcriptional regulator

In order to expand on our findings from the Dox-inducible RTS-3B system, and to shed light on the possible functions of EBNA3B as a transcriptional regulator, I generated LCLs expressing epitope-tagged EBNA3B and used these to perform ChIP linked to high-throughput sequencing (ChIP-seq) analyses. While this was in progress, Wang *et al.* (2015) performed similar analyses for FLAG-HA C-terminal epitope-tagged EBNA3A, EBNA3B and EBNA3C. C-terminal FLAG-HA tagged EBNA3B LCLs expressed lower levels of EBNA3B and struggled to transform and grow in culture as compared to other LCLs (Wang *et al.*, 2015; personal communication with Dr Eric Johannsen). Hence, to our knowledge, the FS3B

cell line is the first stable LCL expressing epitope-tagged EBNA3B and normal physiological levels of EBNA3C.

Wang *et al.* (2015) found 1640 EBNA3A, 3033 EBNA3B, 3588 EBNA3C and 8592 EBNA2 binding sites across the human genome. Both our group (and others) largely agree that the EBNA3s overlap substantially with each other, and consistently show that EBNA3 binding sites are predominantly found to be enriched with mono-, di-, and tri-methylated H3K4 histones, and acetylated H3K9 and H3K27 histones (Jiang *et al.*, 2014; Schmidt *et al.*, 2015; Wang *et al.*, 2015; and see section 4.5.7). This is in accordance to the earliest published work on pan-specific EBNA3 genome-wide binding, which again localised EBNA3s to histone modifications associated with active regulatory regions rather than silenced heterochromatin (H3K9me3 and H3K27me3) (McClellan *et al.*, 2013). Although we have yet to make direct comparisons between our data and the ChIP-seq from Wang *et al.* (2015), it is interesting to note that all three methods of ChIP-seq analyses appear to deliver varying number of binding peaks for each EBNA3, and only 29% and 21% of the pan-specific EBNA3 peaks overlapped with our and Wang *et al.*'s combined EBNA3 ChIP-seq peaks respectively. This shows that different ChIP-seq procedures, cell lines and analyses can produce quite different results. This is likely to be due to precise antibody specificities and affinities, software peak-calling algorithms and varying cellular backgrounds. Future analyses could focus on those peaks that are consistently identified in all the various studies, although this will never be completely satisfactory. Studies of cells with the same genetic background (ie single donors) using the same antibodies, protocols, algorithms etc are likely to produce the most reliable results – with the caveat that there might be minor donor differences. This has been done by Paschos *et al.* (submitted to Nucleic Acids Research).

6.2.1 EBNA3B is mainly a transcriptional repressor

Although there is evidence to suggest that EBNA3B can be involved in mediating both the activation and repression of cellular host genes (White *et al.*, 2010), results from my ChIP-seq (section 4.5.9) support the notion that EBNA3B is primarily a transcriptional

repressor. However, analyses from our lab and others revealed that EBNA3B (like EBNA3A and EBNA3C) tends to bind to transcriptional enhancers marked with H3K4me1 ± H3K27Ac, but very rarely occupies regions of closed chromatin with high levels repressive marks, such as H3K27me3 and H3K9me3 (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015; Wang *et al.*, 2015). This is despite observations made by our group (and others) that the recruitment of PRC2 complexes is necessary in the complete repression of EBNA3A/3C target genes (Harth-Hertle *et al.*, 2013; Kalchschmidt *et al.*, 2016b; Paschos *et al.*, 2012). It appears that the regulation of only a subset of the EBNA3B-repressed genes involves the recruitment of PRC2-complexes to their promoters, and perhaps more focus could be given to the interaction between the EBNA3s and complexes which affect the deposition or removal of enhancer-linked histone marks like H3K4me1, H3K4me3 and H3K27Ac, such as histone methyltransferases, acetyltransferases or HDACs. That said, although the presence of histone marks may provide an indication of the state of the chromatin and transcription status of the genes in question, whether that is a direct result of these regions interacting with EBNA3B, or a consequence of this interaction requires further study. Importantly, it is necessary to verify the ChIP-seq findings with experimental data and to test our hypotheses with detailed analyses at specific gene loci.

6.2.2 EBNA3C might be an important mediator in the co-ordinated gene regulation with EBNA3B

Preliminary analyses of the EBNA3B ChIP-seq data support our speculation that EBNA3B jointly regulates a subset of its target genes with EBNA3C and EBNA3A. EBNA3C could be a major player in this co-operative gene regulation since EBNA3C is almost always present when EBNA3B co-localises with EBNA3A and/or EBNA2. A relatively large subset (12%) of EBNA3B binding peaks co-localise with EBNA3C alone, and a further 19% of the peaks are co-incident with both EBNA3A and EBNA3C (Figure 4.14). In contrast, only a small percentage (1%) of the EBNA3B binding peaks overlapped solely with EBNA3A. Although Wang *et al.* (2015) found a more even distribution of EBNA3B peaks co-localising

with EBNA3A and EBNA3C (10% each), previous exon-microarray analyses of BL31s infected with WT, 3AKO, 3BKO & 3CKO EBV, are consistent with a large subset of EBNA3B-regulated genes (289 of 598) also being co-regulated by EBNA3C (White *et al.*, 2010). In contrast, very few genes (9 of 598) appear to be regulated by the combination of EBNA3A and EBNA3B alone, and 56 genes appear to be regulated by all three EBNA3s (White *et al.*, 2010) (Figure 6.1). Hence, both our ChIP-seq and exon-microarray concur that EBNA3C plays significant role and EBNA3A a more minor role, in the co-regulation of genes with EBNA3B.

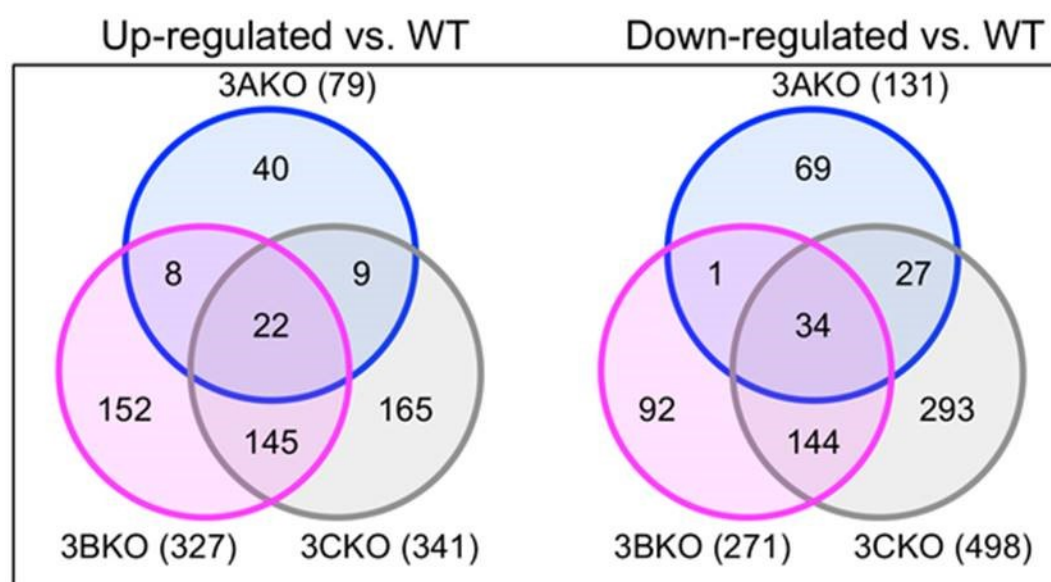


Figure 6.1. Overlap of genes which are differentially expressed in 3A/3B/3CKO recombinant EBV-infected BL31 cells as compared to WT EBV-infected BL31s.
Figure retrieved from White *et al.* (2010).

Our analyses show that EBNA3B binding peaks that co-localise with EBNA3C are most often found within genome segments classified as enhancers or weak enhancers (refer to section 4.5.6). This is consistent with the EBNA3B ChIP-seq analyses done by Wang *et al.* (2015). Similar analyses done by other groups have revealed that EBNA3C and EBNA3A are mainly enriched at strong and weak enhancers within the genome (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015; Wang *et al.*, 2015). EBNA3C and EBNA3A have

both been shown to bind to distal enhancer elements and are involved in either preventing or stabilising looping of the bound regions to TSS of specific genes, thereby facilitating the activation or repression of gene expression (see Introduction, section 1.7.3) (Bazot *et al.*, 2015; McClellan *et al.*, 2013). EBNA3B, when bound to regions that coincide with EBNA3A and EBNA3C, could be involved in a similar mechanism of gene regulation (i.e. facilitating or preventing chromatin looping) as the other EBNA3s. Our analyses show that EBNA3B binds within the same looped chromatin domains as its regulated genes (see section 4.5.9), indicating that its binding region could be brought into close proximity with the regulated gene promoter. How the EBNA3s may interact at their binding regions and exactly how this interaction can affect chromatin looping and gene regulation remains unclear, although we speculate that this is done through physical interactions with each other, cellular factors and recruitment of cellular co-repressors and co-activators such as CtBP, HDACs, components of the PRC complexes and histone acetyl transferase p300 (Cotter and Robertson, 2000; Hickabottom *et al.*, 2002; Knight *et al.*, 2003; Paschos *et al.*, 2012; Radkov *et al.*, 1999; Skalska *et al.*, 2010; Touitou *et al.*, 2001a).

6.2.3 EBNA3B might function in a different manner to EBNA3A and EBNA3C

More than half of the EBNA3B peaks revealed in my analysis are unique sites that do not co-localise with EBNA3A, EBNA3C (see section 4.5.4 and Paschos, unpublished) or EBNA2 (section 4.5.4; McClellan *et al.*, 2013). EBNA3B, when not associated with EBNA3A and/or EBNA3C, localises mainly to transcribed, repressed or uncharacterised regions of the chromatin (section 4.5.6). In addition, H3K36me3 (usually found in gene bodies of active genes) is the 3rd most common histone modification co-localised with unique EBNA3B peaks (section 4.5.7). This again suggests that in addition to having similar regulatory functions to EBNA3A and 3C, EBNA3B might be involved in different regulatory processes. A large majority of the unique EBNA3B peaks localise within regions that do not contain histone marks that I have examined, so it will be interesting to investigate whether these regions associate with other specific histone marks that have been omitted in my analyses

(e.g. less extensively studied histone marks such as tri-methylation at lysine residue 20 of histone H4, a histone mark commonly associated with heterochromatin). Of course it is possible that “uncharacteristic” EBNA3B-bound sites could be an artefact of using the N-terminal epitope-tagged FS3B cells in our ChIP experiments (due to either the FLAG-Strep-Strep tag or slight overexpression/stabilisation of the EBNA3B in our FS3B cells). In order to exclude this possibility, further comparison of our ChIP-seq data with that from other groups can be done. Although Wang and colleagues (2015) identified more (3033) EBNA3B binding regions than our group (1931), we have found similar numbers of unique EBNA3B binding peaks (1285 and 1259 respectively). Future analyses should examine whether these two sets of unique EBNA3B peaks overlap and whether they are associated with similar histone modifications, transcription factors and gene segments. In the end it will require a detailed analysis of single sites and genes as done for TERT.

6.2.4 Data analyses using high stringency produces similar results as this study

While the main focus of my study was to examine the properties of EBNA3B-bound regions, Dr Kostas Paschos was involved in interrogating all EBNA3 binding sites and making comparisons across the three epitope-tagged viruses in one single donor. Using a different peak calling algorithm (MACS, q value cut-off of 0.05, Zhang *et al.*, 2008) with a higher stringency, he has uncovered 1715 peaks for EBNA3A, 454 for EBNA3B and 3835 for EBNA3C (Paschos *et al.*, 2016). Specifically for EBNA3B, the MACS algorithm was used with the broad peak setting because of 64 peaks which would have been excluded otherwise. Despite the fewer numbers of peaks being revealed by MACS, the general findings Dr Paschos has made were consistent with my study: that the EBNA3s largely associate with enhancers and active promoters, whereas EBNA3B on its own has a greater association with H3K36me3 than other EBNA3s (section 4.5.6, 4.5.7 and Appendix F). In addition, EBNA3B on its own localised more frequently with quiescent regions and areas of active transcription but less with enhancers and TSS than other EBNA3s (section 4.5.6 and Appendix F). Dr Paschos has similarly found that EBNA3B rarely associates with the same

regions as EBNA3A, unless EBNA3C is present (section 4.5.4, Appendix F), again this suggests that either EBNA3C has a role in recruiting the other two EBNA3s, or that different recruitment models exist for EBNA3A and EBNA3B.

Dr Paschos has also examined the association between contact domains and EBNA3-regulated genes. Here, he has used a higher stringency to select for EBNA3-regulated genes (differential expression of at least 2-fold and/or with $p < 0.001$ between the individual EBNA3 knock-out and WT LCLs), and found that when compared to randomly generated “peaks”, the EBNA3s have significantly higher association with contact domains that contain their regulated genes, with the exception of EBNA3B and its upregulated genes (Appendix F). This supports our view that EBNA3B is mainly a transcriptional repressor and is likely to be indirectly activating cellular genes. With higher stringency applied for both peak calling and gene selection, Dr Paschos has also revealed that all EBNA3-regulated genes that fall within the same EBNA co-localised contact domain were regulated in the same direction (either activated or repressed), with the exception of one domain for EBNA3B, strongly supporting the view that these can be co-ordinately regulated by EBNA3s bound at a region distal to the genes. This regulation is likely to be mediated through chromosomal looping. Finally, Dr Paschos revealed that a large majority of the top 20 transcription factors associating with each EBNA3 were very similar, hence supporting the notion that the EBNA3s can interact with the similar cellular factors. Similar observations were made about EBNA3B co-localising with low percentages of transcription factors (Appendix G), although the significance of this remains unclear. He has also similarly observed that although the EBNA3s are collectively very similar in their binding profile, EBNA3B on its own behaves differently to EBNA3A or EBNA3C. Hence, comparisons of this study with the one performed by Dr Paschos have shown that even with more stringent peak-calling, the conclusions we arrive at are generally very similar.

6.2.5 EBNA3B co-localises with EBNA2 at strong enhancers and TSS of regulated genes

A large proportion of the EBNA3B peaks that co-localise with EBNA2 are found at strong enhancers and TSS of genes, which is mirrored by the high proportion of EBNA2+EBNA3B peaks which occupy sites that contain histone marks associated with positive regulatory regions (e.g. H3K4me3, H3K9Ac) (section 4.5.7). This reflects the role of EBNA2 as a strong viral transactivator that is often localised within distal enhancers and tethered to promoter and regulatory regions of genes by binding to RBPJ (Zhao *et al.*, 2011a; Zhou *et al.*, 2015; Zimmer-Strobl and Strobl, 2001), and suggests that some functions of EBNA3B may be associated with this property of EBNA2 at these regions. The molecular details of how the two proteins may interact – whether in a competitive, antagonistic or agonistic manner remains to be determined. The widely accepted model of EBNA3s competing with EBNA2 for binding to DNA-bound RBPJ in order to antagonise the transactivation of EBNA2-regulated (see Introduction section 1.7.1) is one explanation for the presence of both EBNA2 and EBNA3B at these loci. A dynamic equilibrium between EBNA2 and EBNA3B binding could have resulted in a proportion of cells having one or the other bound at these locations. A similar observation was made by McClellan *et al.* (2013) regarding EBNA3C and EBNA2 binding at the WEE1 distal enhancer. Re-ChIP analyses, or ChIP experiments done on 3BKO and WT LCLs examining specific EBNA2 binding sites could be done to further investigate this hypothesis.

6.2.6 EBNA3B is involved in chromatin looping mechanisms between TSS and distal regulatory elements

Consistent with current understanding of EBNA3A and EBNA3C-bound loci, EBNA3B binding regions are mainly found distal to transcription start sites (Jiang *et al.*, 2014; McClellan *et al.*, 2013; Schmidt *et al.*, 2015). Around half of the contact domains (of around 10kb long) identified by Rao and colleagues that harbour EBNA3B peaks additionally contain the TSS of one or more regulated genes (section 4.5.9). Whether EBNA3B plays a role in

affecting the 3D topology of the chromatin by recruiting chromatin remodelling factors (for example the SWI/SNF complex) cannot be determined by these analyses alone, and further analyses and experiments should be done to directly assess the EBNA3B-driven chromatin remodelling processes. I have found that more than one potential EBNA3B-regulated promoter can occur in the same contact domain and these promoters might therefore be coordinately regulated by EBNA3B. Further analyses of the data with follow-up experiments, such as 3C and gene expression analyses for specific gene loci will be done to determine whether this is the case.

6.2.7 Overview and future work on EBNA3B ChIP-seq

In conclusion, *in silico* analyses of our EBNA3B sequencing data have revealed that EBNA3B binding exhibits several similarities with EBNA3A and EBNA3C localisation within the genome, and is consistent with the hypothesis that EBNA3B can interact with other EBNA3s and perhaps EBNA2 in the regulation of its target genes. A large proportion of the EBNA3B peaks were found within genome segments that, to date, have no well-defined characteristics, and a more focussed biochemical investigation is required.

6.3 Regulation of *TERT* by EBNA3B

6.3.1 EBNA3B regulates *TERT* expression by mediating epigenetic modifications at the *TERT* promoter

A robust de-repression of *TERT* is evident within 5 days of infecting primary B cells with 3BKO virus (sections 3.6.6 and 5.2). During the outgrowth of LCLs, the *TERT* gene is robustly induced in the absence of EBNA3B, which suggests that a component of EBV might be activating the gene – the timing and presence of upstream EBNA2 binding sites suggest that this is likely to be either direct activation by EBNA2 or indirect activation via MYC, since MYC is activated by EBNA2, and *TERT* is known to be a MYC target (Kaiser *et al.*, 1999; Wu *et al.*, 1999). The EBV latent protein LMP1 has already been reported to activate the *TERT* promoter in transient transfection assays in B-lymphocytes via the NFκB, MAPK and

ERK1/2 signalling pathways (Terrin *et al.*, 2008) and is also a likely candidate for this activation (see Introduction, section 1.8.2).

6.3.2 EBNA3B may facilitate TERT regulation by mediating chromatin looping of a distal enhancer to the promoter region

EBNA3B, EBNA3A and EBNA3C bind to a region ~17kb upstream of the *TERT* TSS (TERT A) and a second region ~20kb upstream of the TERT TSS (TERT B) is bound by only EBNA3A and EBNA3C. Both binding sites are found within a wider region (Figure 5.4A) that is enriched with H3K27Ac, suggesting that they are within an active enhancer in GM12878 WT LCLs. Additionally, I have found higher enrichment of H3K9Ac and H3K27Ac at both TERT A and TERT B binding sites in 3BKO LCLs, which suggests activation and increased availability of this proposed enhancer region when EBNA3B is absent. Since the distal enhancer could already be active in WT LCLs, the action of EBNA3B would be to prevent access of the enhancer to the TSS of *TERT*, hence preventing activation of the gene. Here, two modes of action may be possible, as suggested by McClellan *et al.* (2013): EBNA3B could repress TERT by preventing the formation of enhancer-promoter loops, or paradoxically, could be repressing the gene despite the formation of such enhancer-promoter loops. Since both the EBNA3B-bound site and *TERT* promoter are within the same looped chromatin contact domain (Rao *et al.*, 2014) in WT GM12878 LCLs, it is likely that the second mechanism occurs in the regulation of *TERT*. Enhancer-promoter looping could prevent further acetylation of H3K26 and H3K9 and enable EBNA3B to recruit components of the PRC2 complex in order to facilitate deposition of H3K27me3 at the TSS. This probably involves EBNA3B-mediated prevention of histone acetyltransferases and SET-domain-containing H3K4 methyltransferases from gaining access to the *TERT* TSS or recruitment of cellular factors such as PRC2 complexes and HDACs. This is consistent with reports from our group and others who have shown that EBNA3A and EBNA3C orchestrate gene repression by first facilitating the removal of histone marks of activation at the EBNA3-bound site and/or gene promoter, followed by the recruitment of PRC2 complex components and

the deposition of H3K27me3 (Harth-Hertle *et al.*, 2013; Kalchschmidt *et al.*, 2016b; see Introduction section 1.7.4). Together, the loss of activating histone marks coupled with increased deposition of repressive H3K27me3 would inhibit transcription initiation. In order to confirm this model of regulation, 3C experiments comparing chromatin conformation at the *TERT* locus in both EBNA3B-positive and negative cell lines are required. These experiments are in progress (Quentin Bazot, personal communication).

Thus far, microarray analyses done on WT/3AKO and EBNA3C-HT cell lines (available on www.epstein-barrvirus.org) and our primary B-cell outgrowth experiments have shown that EBNA3A exerts no effect and EBNA3C has a minimal effect on the repression of *TERT* expression (Hertle *et al.*, 2009; Skalska *et al.*, 2013; White *et al.*, 2010). Although EBNA3A and EBNA3C are apparently both present at the same locus as EBNA3B, it is likely they do not have an obvious role in the regulation of *TERT* transcription. Similar observations have been made by Kalchschmidt *et al.* (2016b) and Wang *et al.* (2015), where despite co-incident binding of EBNA3A and EBNA3C at *CXCR5*, *COBLL1*, *ADAM28* and *ADAMDEC1* loci, EBNA3A does not appear to have an effect on the expression of *CXCR5* and *COBLL1*, and has a subsidiary effect on the repression of *ADAM28* and *ADAMDEC1* during B-cell outgrowth.

6.3.3 Heterogeneity in telomere length in cell cultures could mask subtle changes early in the B-cell outgrowth

It has been reported that LCLs undergo a proliferative crisis during their maintenance in culture, and the few cell lines that are truly immortalised are able to overcome this crisis (Counter *et al.*, 1994; Sugimoto *et al.*, 1999; Tahara *et al.*, 1997). True immortalisation of EBV-transformed B-cells is associated with activation of *TERT* expression and sustained telomerase activity in the cells at the point of the proliferative crisis. Hence, B-cells which do not express *TERT* arrest in culture when they senesce due to the erosion of telomeres (Counter *et al.*, 1994; Sugimoto *et al.*, 2004; Terrin *et al.*, 2007). In order to determine whether the repression of *TERT* by EBNA3B has an effect on telomere maintenance after

multiple passages, I studied the telomere length of B-cells that were infected with WT, 3BKO and 3BRev EBV and found that the telomere lengths are maintained in all LCLs established from pooled mixed-donor B-cells after 120 days, regardless of the presence of EBNA3B. Shortening of telomeres was observed after 40 days in some LCLs established from single-donor B-cell cultures, but again this appears to be independent of their EBNA3B status.

According to the Sugimoto group, population doubling for LCLs grown at 1×10^6 cells/ml occurs every 3-4 days, therefore 40 and 160 PDLs would approximately be equivalent to 140 and 560 days respectively. We cultured our cells under different conditions (3×10^5 cells/ml, passaged every 3-4 days), which could affect any direct comparisons in terms of days post infection. However, despite the differences in culture conditions, it is likely that the shortening of the telomere length becomes evident earlier in our single-donor cultures compared to pooled-donor cultures because of their reduced heterogeneity in the baseline telomere length, basal expression of *TERT* and activity of telomerase, and an earlier establishment of monoclonality of the culture. Heterogeneity in the telomere length of B-cells in mixed-donor cultures would have masked any subtle changes in the telomere lengths in the first 120 days. In the single-donor cultures, differences would be observed earlier and be more pronounced because of reduced heterogeneity between cells in the culture, and any effects of clonal selection during outgrowth would become more evident. Indeed, studies have shown that *in vitro* EBV infection of purified naïve B-cells frequently selects for clones and drives the population towards monoclonality after around 4 weeks (28 days) of outgrowth (Heath *et al.*, 2012). In addition, while Sugimoto and colleagues have also noted that the proliferative crisis of LCLs established from PBLs happens at around 160 population doubling levels (PDLs), Counter *et al.* observed that this crisis appears earlier at 40 PDLs in populations of LCLs which were cloned by sub-culturing the LCLs on a layer of human fibroblast feeder cells (Counter *et al.*, 1994; Sugimoto *et al.*, 2004). Sugimoto *et al.* (2004) postulated that this difference is due to the non-cloned LCL populations taking a

longer time to acquire clonal characteristics via the selection for cells which are able to adapt to long term *in vitro* culture.

6.3.4 Telomeres of EBV-infected B-cells may shorten until they arrive at their proliferative crisis despite the *TERT* status of the cells

It has been reported by several groups that the telomeres of newly EBV-infected B-cell cultures continue to shorten with cellular division until the cells arrive at their proliferative crisis, regardless of *TERT* expression and telomerase activity in these cells (Counter *et al.*, 1994; Guerrini *et al.*, 1993; Tahara *et al.*, 1997; Terrin *et al.*, 2007). The activation of *TERT* and increased telomerase activity become important factors for survival only when cells have arrived at this critical point, after which the telomere lengths become stabilised. The TRAP assay showed differential telomerase activity in long-established WT and 3BKO LCLs (section 5.3.4), but higher telomerase activity would not necessarily correlate to the maintenance of telomeres at the original length in 3BKO LCLs. Therefore, although we expected to observe enhanced telomerase activity associated with the de-repression of *TERT* in 3BKO LCLs, it is perhaps not surprising that shortening of telomeres occur in both 3BKO and WT LCLs before they arrive at proliferative crisis. Repression of *TERT* by EBNA3B could contribute to one of the barriers that cells have to overcome in order to survive the crisis. Given that the cells in our cultures had probably not have arrived at a crisis state before 120 days, it is not possible to determine whether the absence of EBNA3B can confer survival advantage to the cells through de-repression of *TERT*. Future studies could examine longer experiments and focus on whether increased telomerase activity would aid the stabilisation of telomere lengths in 3BKO LCLs on arrival at the critical point.

To further confound the issue of telomere length in proliferating B-cells, the telomerase-independent alternative lengthening of telomere (ALT) pathway has been implicated in maintaining and modestly increasing the length of telomeres in newly established EBV-infected B-cells which exhibit low telomerase activity during the first 4 weeks of culture (Kamranvar *et al.*, 2013). Although the exact mechanisms of how it is done

are unclear, the ALT pathway is believed to be able to maintain and extend the telomeres through homologous recombination between telomeric sequences in the cell (reviewed in Cesare and Reddel, 2010). This raises the question of whether the cell lines which have managed to maintain telomere lengths in our 120-day experiment do so by exploiting the ALT pathway in a situation with repressed *TERT* levels. In order to answer this question, a study of any changes in the phenotypic markers of the ALT pathway, including extra-chromosomal telomeric DNA and the accumulation of promyelocytic leukaemia nuclear bodies (PMB) containing telomeric chromatin (reviewed in Cesare and Reddel, 2010) could be done.

6.3.5 Future work on the repression of *TERT* by EBNA3B and its physiological implications

6.3.5.1 Investigations on the mechanisms of repression of *TERT* by EBNA3B

Further to the 3C experiments discussed in section 6.3.2, a study of the kinetics of epigenetic regulation of *TERT* using the Dox-inducible EBNA3B system or outgrowth in primary B-cells could be done in order to analyse temporally and spatially the molecular details of EBNA3B function in regulating *TERT*. Examination of the deposition of histone marks and enrichment of cellular factors such as components of the PRC2 complex and perhaps RBPJ at both the promoter and distal EBNA3B binding regions over time would shed light on the temporal sequence of events that occurs on introduction and removal of EBNA3B into the B-cells. This has been done with 4HT-inducible EBNA3C-HT and Dox-inducible EBNA3A RTS cell lines by our group and others with illuminating results on the role of EBNA3C and EBNA3A in mediating gene regulation (Bazot *et al.*, 2015; Harth-Hertle *et al.*, 2013; Maruo *et al.*, 2003, 2006; Skalska *et al.*, 2010) (See Introduction, section 1.7).

In addition, further investigations into the possible involvement of EBNA2 and RBPJ in mediating the regulation of *TERT* by EBNA3B can be performed. As discussed in section 1.7.1, the interactions of the EBNA3s with RBPJ and EBNA2 are important in the regulation

of a subset of cellular genes. Importantly, EBNA2, and EBNA3B have reported co-incident binding at the distal enhancer from the *TERT* promoter (Figure 6.2). Previously, the relationship between EBNA3B and RBPJ could not be reliably assessed since the RBPJ ChIP-seq available for B-cells was done on IB4 cells that do not express EBNA3B. Very recently, however, a RBPJ ChIP-seq was performed by Lu *et al.* (2016) in LCLs expressing all EBNA3 proteins. This RBPJ ChIP-seq has revealed that RBPJ binds to the same genomic loci as the EBNA3s upstream of the *TERT* gene (Figure 6.2). It is thus necessary to determine if their co-incident binding is a result of their interactions with EBNA3B and the *TERT* promoter. Co-incident binding of RBPJ, EBNA2 and EBNA3 could indicate dynamic equilibrium of competition between EBNA2 and EBNA3B binding the enhancer, perhaps recruiting different transcriptional regulators such as RBPJ to activate or repress the gene respectively. This is a similar mechanism of action proposed by McClellan and colleagues (2013) for EBNA3C and its regulated gene, *WEE1*. This hypothesis will have to be rigorously tested by further gene expression analyses, ChIP and re-ChIP, and Chromatin conformation capture (3C) experiments.

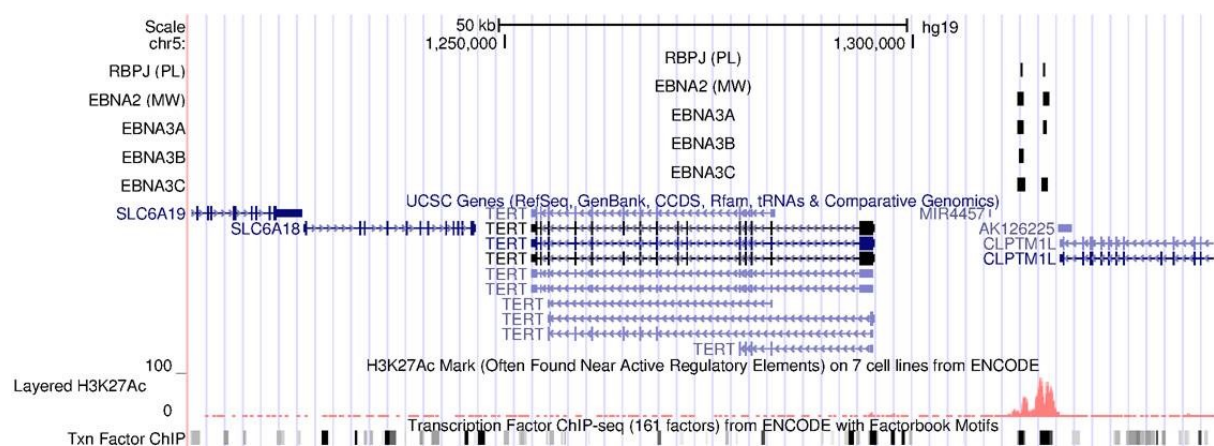


Figure 6.2. View of RBPJ, EBNA2, and EBNA3A/3B/3C binding at the *TERT* locus using Genome Browser (accessible at genome.ucsc.edu.)

The RBPJ binding peaks are obtained from Lieberman group (Lu *et al.*, 2016) and EBNA2 peaks are taken from the West group (McClellan *et al.*, 2013). The rest, including EBNA3A, 3B and 3C ChIP-seq data are provided by the Allday group (unpublished).

6.3.5.2 Further work on studying the implications of TERT repression by EBNA3B

Terrin and colleagues (2007) noted that the onset of *TERT* induction and activation of telomerase during the outgrowth of WT EBV-infected primary B-cell cultures differed greatly and ranged from 3 culture passages to 50 culture passages (cultures were passaged every 3 to 5 days). There could have been differences in the onset of activation of *TERT* in all of the 3BKO EBV-infected B-cell cultures that we have used in the telomere length experiments. However, the expression of TERT in these WT/3BKO/3BRev cell lines was also not formally investigated. Due to limited number of purified B-cells obtained from each donor, I was unable to harvest cells for RNA simultaneously throughout the initial outgrowth process, with the exception of cell cultures established from pooled donor GH1+2 B-cells. Although the presence of detectable *TERT* transcripts before the proliferative crisis does not appear to have an immediate effect on the telomere length (Counter *et al.*, 1994; Tahara *et al.*, 1997; Terrin *et al.*, 2007), future experiments should include RNA and protein analyses so that we could investigate whether there is a correlation between the presence of EBNA3B and the timing of activation of TERT, telomerase activity and the effects on telomere lengths of the cells. If time and funding permits, these experiments could also be repeated for longer durations and in larger and diverse sample sizes.

Increasingly, there have been reports on non-canonical functions of TERT including modulation of the expression of factors involved in NF- κ B and Wnt/ β -catenin signalling pathways (reviewed by Li and Tergaonkar, 2014) – signalling cascades that govern inflammatory processes and cell proliferation – and dysregulation of these pathways is often associated with carcinogenesis. Importantly, in B-cells, TERT is reported to activate *NOTCH2* (Giunco *et al.*, 2015), a member of the NOTCH gene family of receptors which regulate cellular differentiation, proliferation and programmed cell death. This TERT-induced activation of NOTCH2 is reported to involve the NF- κ B signalling pathway. NOTCH2 in turn activates BATF, a transcription factor which represses the expression of the BZLF1, the main transcription factor in triggering the EBV lytic cycle. Hence, TERT may play a

role in inhibiting the activation of the virus lytic cycle, thereby indirectly promoting the extended proliferation of primary B-cells, and maintaining the latency state of EBV-infected cells (Giunco *et al.*, 2013, 2015). Future work could also explore the implications of EBNA3B on these non-canonical functions of TERT and their effects on cell proliferation and survival. Since EBV has not evolved to be a tumour virus, it is likely that the reason for the co-evolution of EBNA3B repression of TERT relates to transient B cell proliferation rather than that seen in extended culture. However, in the aetiology of B cell lymphomas, presumably having an active TERT is likely to be pro-oncogenic – hence the selection for loss of EBNA3B function in DLBCL/PTLD (White *et al.*, 2012).

6.3.6 Overview on repression of *TERT* by EBNA3B

Using both the Dox-inducible EBNA3B and epitope-tagged EBNA3B cell lines I have shown, using *TERT* as an example, that EBNA3B can regulate a host cell genes by effecting epigenetic changes at the gene loci, and at distal binding regions. It is important that future experiments examine any direct evidence of EBNA3B (and any other EBNA3s, EBNA2 and RBPJ) actively involved in the recruitment of cellular factors which affect the deposition or removal of histone marks, or the looping of chromatin to bring distal enhancers closer to the promoter of *TERT*. The consequence of the repression of TERT is that the cell lines can have reduced telomerase activity, although the impact on telomere length or other non-canonical functions of TERT cannot yet be established based on our results thus far, and more work will have to be done to ascertain this.

6.4 Regulation of *TRAIL*, *TRAILR1* and *TRAILR4* by EBNA3B

In addition to TERT, we briefly examined the role of EBNA3B in the regulation of the tumour necrosis factor-related apoptosis-inducing ligand (TRAIL), and its receptor TRAILR1 and decoy receptor TRAILR4. TRAIL is part of the tumour necrosis factor (TNF) family of cytokines that induces apoptosis upon binding to its functional receptors such as TRAILR1. Binding of TRAIL to its decoy receptors (e.g. TRAILR4), however, would not induce apoptosis and high levels of TRAILR4 may protect the cell against TRAIL-induced

programmed cell death (Abdulghani and El-Deiry, 2010; Mahalingam *et al.*, 2009). Data from our EBNA3B ChIP-seq showed EBNA3B binding within the intronic region of the *TRAIL* and to another binding site within the intron of *TRAILR1*, which is located immediately upstream of *TRAILR4*. Furthermore, primary B-cell outgrowth experiments were consistent with our WT/3BKO microarray (White *et al.*, 2012) that *TRAILR1* and *TRAILR4* are likely to be EBNA3B-repressed genes, but in contrast to our previous experiments, the expression of *TRAIL* during outgrowth gave equivocal results (Figure 5.11) which does not support EBNA3B having an effect on TRAIL transcription early during outgrowth. Discrepancy also exists between our Dox-inducible EBNA3B experiments and the primary B-cell outgrowth, since the regulation of *TRAILR4* by EBNA3B could not be confirmed by the RTS system, but the outgrowth of 3BKO versus WT LCLs seemed to suggest that *TRAILR4* is suppressed in WT cells. It is therefore important to repeat these experiments in order to ensure that the results are reproducible and to determine whether these genes are indeed regulated by EBNA3B. Further investigations should be done to determine the role of EBNA3A, 3C and EBNA2 in modulating these genes and to elucidate the underlying molecular mechanisms behind this regulation. This is especially interesting given that therapeutic approaches aiming to down-regulate TRAIL-induced apoptosis are being studied in clinical trials as cancer treatments (Abdulghani and El-Deiry, 2010; Micheau *et al.*, 2013), and may be potential agents against EBNA3B-deficient lymphomas which could over-express TRAIL receptors.

6.5 Final remarks

Even though EBNA3B is non-essential for B-cell transformation *in vitro*, the fact that it has not been counter-selected over years of host and virus co-evolution supports the general view that it has an essential role *in vivo*. The limited available data indicate that unlike EBNA3A and EBNA3C that behave as oncoproteins and enhance B-cell transformation, EBNA3B is a tumour suppressor reducing oncogenic potential. Hence, given that little is known about how this EBV protein functions, it was necessary to initiate studies

on its molecular mechanisms of action. The key outcomes of my investigations are summarised below:

1. I have established two systems that can be used to examine the function of EBNA3B – the Dox-inducible RTS-3B system and the epitope-tagged EBNA3B LCLs.
2. TLDA analyses using the Dox-inducible RTS-3B system have identified 8 definite and 14 probable EBNA3B-regulated genes, including *TERT* as a proven EBNA3B-repressed gene, *TRAIL* and *TRAILR1* as probable EBNA3B-regulated genes. The regulation of *TRAILR4* by EBNA3B could not be confirmed by this system.
3. ChIP-seq analyses done on the epitope-tagged EBNA3B LCLs revealed:
 - a. That EBNA3B alone may have a different binding profile to EBNA3A and EBNA3C
 - b. That EBNA3C may be a common co-factor in the co-ordinated regulation of gene expression with EBNA3B
 - c. That EBNA3B with EBNA2 are co-localised within active regulatory regions and TSS of genes, which corresponds to EBNA2 as a transactivator.
 - d. EBNA3B binds at distal gene regulatory regions that may be looped to come into contact with gene promoters, including *TERT*.
4. Focussed analyses of *TERT* found that:
 - a. EBNA3B may regulate *TERT* expression by initiating the de-acetylation of H3K9 and H3K27 at the promoter and the distal upstream enhancer, and mediating the deposition of H3K27me3 by recruiting PRC2 complexes at these loci.
 - b. Regulation could be facilitated by the looping of the distal enhancer to the promoter of *TERT*, although further 3C experiments are necessary to formally demonstrate this.

- c. The repression of TERT does not have an observable effect on the maintenance of telomere length in the first 120 days after EBV-infection of primary B-cells.
- 5. Preliminary analyses of TRAIL, TRAILR1 and TRAIL4 revealed that there may be some discrepancies between the Dox-inducible system and our primary B-cell outgrowth experiments and these will have to be repeated to ensure reproducibility.
 - a. TRAIL is a probable EBNA3B-induced gene as determined by the Dox-inducible EBNA3B system but shows no real indications of EBNA3B regulation in infection and outgrowth experiments.
 - b. TRAILR1 is classed as a probable EBNA3B-repressed gene with the Dox-inducible EBNA3B system and shows repression in WT LCLs as compared to 3BKO LCLs in early EBV-driven B-cell outgrowth. This needs to be explored in multiple donors.
 - c. Despite falling into the category of “not confirmed” regulation by EBNA3B, TRAILR4 shows robust repression by EBNA3B in early B-cell outgrowth. Again, this requires exploration using various donors.

7 Bibliography

- Abdulghani, J., and El-Deiry, W.S. (2010). TRAIL receptor signaling and therapeutics. *Expert Opinion on Therapeutic Targets* 14, 1091–1108.
- Allday, M.J. (2009). How does Epstein-Barr virus (EBV) complement the activation of Myc in the pathogenesis of Burkitt's lymphoma? *Seminars in Cancer Biology* 19, 366–376.
- Allday, M.J., and Farrell, P.J. (1994). Epstein-Barr virus nuclear antigen EBNA3C/6 expression maintains the level of latent membrane protein 1 in G1-arrested cells. *Journal of Virology* 68, 3491–3498.
- Allday, M.J., Crawford, D.H., and Griffin, B.E. (1988). Prediction and demonstration of a novel Epstein-Barr virus nuclear antigen. *Nucleic Acids Research* 16, 4353–4367.
- Allday, M.J., Bazot, Q., and White, R.E. (2015). The EBNA3 family: two oncoproteins and a tumour suppressor that are central to the biology of EBV in B cells. In *Epstein Barr Virus Volume 2*, C. Munz, ed. pp. 61–117.
- Almasan, A., and Ashkenazi, A. (2003). Apo2L/TRAIL: Apoptosis signaling, biology, and potential for cancer therapy. *Cytokine and Growth Factor Reviews* 14, 337–348.
- Anderton, E., Yee, J., Smith, P., Crook, T., White, R.E., and Allday, M.J. (2008). 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.
- Auerbach, R.K., Euskirchen, G., Rozowsky, J., Lamarre-Vincent, N., Moqtaderi, Z., Lefrançois, P., Struhl, K., Gerstein, M., and Snyder, M. (2009). Mapping accessible chromatin regions using Sono-Seq. *Proceedings of the National Academy of Sciences of the United States of America* 106, 14926–14931.
- Autexier, C., and Lue, N.F. (2006). The structure and function of telomerase reverse transcriptase. *Annual Review of Biochemistry* 75, 493–517.
- Ayyanathan, K., Ayyanathan, K., Lechner, M.S., Lechner, M.S., Bell, P., Bell, P., Maul, G.G., Maul, G.G., Schultz, D.C., and Schultz, D.C. (2003). Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes & Development* 18, 1855–1869.
- Bain, M., Watson, R.J., Farrell, P.J., and Allday, M.J. (1996). Epstein-Barr Virus Nuclear Antigen 3C Is a Powerful Repressor of Transcription when Tethered to DNA. *Journal of Virology* 70, 2481–2489.
- Bazot, Q., Deschamps, T., Tafforeau, L., Siouda, M., Leblanc, P., Harth-Hertle, M.L., Rabourdin-Combe, C., Lotteau, V., Kempkes, B., Tommasino, M., Gruffat, H., and Manet, E. (2014). Epstein-Barr virus nuclear antigen 3A protein regulates CDKN2B transcription via interaction with MIZ-1. *Nucleic Acids Research* 42, 9700–9716.
- Bazot, Q., Paschos, K., Skalska, L., Kalchschmidt, J.S., Parker, G. a., and Allday, M.J. (2015). Epstein-Barr Virus Proteins EBNA3A and EBNA3C Together Induce Expression of the Oncogenic MicroRNA Cluster miR-221/miR-222 and Abrogate Expression of Its Target p57KIP2. *PLOS Pathogens* 11, e1005031.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S.L., and Lander, E.S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326.

- Bornkamm, G.W., Berens, C., Kuklik-Roos, C., Bechet, J.-M., Laux, G., Bachl, J., Korndoerfer, M., Schlee, M., Hölzel, M., Malamoussi, A., Chapman, R.D., Nimmerjahn, F., Mautner, J., Hillen, W., Bujard, H., and Feuillard, J. (2005). Stringent doxycycline-dependent control of gene activities using an episomal one-vector system. *Nucleic Acids Research* 33, e137.
- Borza, C.M., and Hutt-Fletcher, L.M. (2002). Alternate replication in B cells and epithelial cells switches tropism of Epstein-Barr virus. *Nature Medicine* 8, 594–599.
- Bourillot, P.Y., Waltzer, L., Sergeant, A., and Manet, E. (1998). Transcriptional repression by the Epstein-Barr virus EBNA3A protein tethered to DNA does not require RBP-Jkappa. *The Journal of General Virology* 79, 363–370.
- Brookes, E., and Pombo, A. (2009). Modifications of RNA polymerase II are pivotal in regulating gene expression states. *EMBO Reports* 10, 1213–1219.
- Brooks, L., Yao, Q.Y., Rickinson, A.B., and Young, L.S. (1992). Epstein-Barr virus latent gene transcription in nasopharyngeal carcinoma cells: coexpression of EBNA1, LMP1, and LMP2 transcripts. *Journal of Virology* 66, 2689–2697.
- Brou, C., Logeat, F., Gupta, N., Bessia, C., Lebail, O., Doedens, J.R., Cumano, A., Roux, P., Black, R.A., and Israel, A. (2000). A Novel Proteolytic Cleavage Involved in Notch Signaling: The Role of the Disintegrin-Metalloprotease TACE. *Molecular Cell* 5, 207–216.
- Buck, M., Burgess, A., Stirzaker, R., Krauer, K., and Sculley, T. (2006). Epstein-Barr virus nuclear antigen 3A contains six nuclear-localization signals. *Journal of General Virology* 87, 2879–2884.
- Burgess, A., Buck, M., Krauer, K., and Sculley, T. (2006). Nuclear localization of the Epstein-Barr virus EBNA3B protein. *The Journal of General Virology* 87, 789–793.
- Burkitt, D.P. (1958). A sarcoma involving the jaws in African children. *The British Journal of Surgery* 46, 218–223.
- Caldwell, R.G., Wilson, J.B., Anderson, S.J., and Longnecker, R. (1998). Epstein-Barr virus LMP2A drives B cell development and survival in the absence of normal B cell receptor signals. *Immunity* 9, 405–411.
- Cameron, E.E., Bachman, K.E., Myöhänen, S., Herman, J.G., and Baylin, S.B. (1999). Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nature Genetics* 21, 103–107.
- Cesare, A.J., and Reddel, R.R. (2010). Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet* 11, 319–330.
- Chen, A., Divisconte, M., Jiang, X., Quink, C., and Wang, F. (2005). Epstein-Barr virus with the latent infection nuclear antigen 3B completely deleted is still competent for B-cell growth transformation in vitro. *Journal of Virology* 79, 4506–4509.
- Chen, A., Zhao, B., Kieff, E., Aster, J.C., and Wang, F. (2006). EBNA-3B- and EBNA-3C-regulated cellular genes in Epstein-Barr virus-immortalized lymphoblastoid cell lines. *Journal of Virology* 80, 10139–10150.
- Cheung, A., and Kieff, E. (1982). Long internal direct repeat in Epstein-Barr virus DNA. *Journal of Virology* 44, 286–294.
- Cheutin, T., and Cavalli, G. (2014). Polycomb silencing: from linear chromatin domains to 3D chromosome folding. *Current Opinion in Genetics & Development* 25C, 30–37.

Cordier, M., Calender, A., Billaud, M., Zimmer, U., Rousselet, G., Pavlish, O., Banchereau, J., Tursz, T., Bornkamm, G., and Lenoir, G.M. (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 Robertson, E.S. (2000). Modulation of histone acetyltransferase activity through interaction of Epstein-Barr nuclear antigen 3C with prothymosin alpha. *Molecular and Cellular Biology* 20, 5722–5735.

Counter, C.M. (1996). The roles of telomeres and telomerase in cell life span. *Mutation Research/Reviews in Genetic Toxicology* 366, 45–63.

Counter, C.M., Avilion, A.A., LeFeuvre, C.E., Stewart, N.G., Greider, C.W., Harley, C.B., and Bacchetti, S. (1992). Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. *The EMBO Journal* 11, 1921–1929.

Counter, C.M., Botelho, F.M., Wang, P., Harley, C.B., and Bacchetti, S. (1994). Stabilization of Short Telomeres and Telomerase Activity Accompany Immortalization of Epstein-Barr Virus-Transformed Human B Lymphocytes. *Journal of Virology* 68, 3410–3414.

Di Croce, L., and Helin, K. (2013). Transcriptional regulation by Polycomb group proteins. *Nature Structural & Molecular Biology* 20, 1147–1155.

Deacon, E.M., Pallesen, G., Niedobitek, G., Crocker, J., Brooks, L., Rickinson, A.B., and Young, L.S. (1993). Epstein-Barr virus and Hodgkin's disease: transcriptional analysis of virus latency in the malignant cells. *The Journal of Experimental Medicine* 177, 339–349.

Degli-Esposti, M.A., Dougall, W.C., Smolak, P.J., Waugh, J.Y., Smith, C.A., and Goodwin, R.G. (1997). The Novel Receptor TRAIL-R4 Induces NF- κ B and Protects against TRAIL-Mediated Apoptosis, yet Retains an Incomplete Death Domain. *Immunity* 7, 813–820.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. *Science (New York, N.Y.)* 295, 1306–1311.

Delbarre, E., Jacobsen, B.M., Reiner, A.H., Sørensen, A.L., and Collas, P. (2010). Chromatin Environment of Histone Variant H3 . 3 Revealed by Quantitative Imaging and Genome-scale Chromatin and DNA Immunoprecipitation. *Molecular Biology of the Cell* 21, 1872–1884.

Delecluse, H.J., Hilsendegen, T., Pich, D., Zeidler, R., and Hammerschmidt, W. (1998). Propagation and recovery of intact, infectious Epstein-Barr virus from prokaryotic to human cells. *Proceedings of the National Academy of Sciences of the United States of America* 95, 8245–8250.

Dohm, J.C., Lottaz, C., Borodina, T., and Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research* 36.

Donnelly, M.L.L., Hughes, L.E., Luke, G., Mendoza, H., Dam, E., Gani, D., and Ryan, M.D. (2001). The “cleavage” activities of foot-and-mouth disease virus 2A site-directed mutants and naturally occurring “2A-like” sequences. *Journal of General Virology* 82, 1027–1041.

Draborg, A., Izarzugaza, J.M.G., and Houen, G. (2016). How compelling are the data for Epstein-Barr virus being a trigger for systemic lupus and other autoimmune diseases? *Current Opinion in Rheumatology* 1.

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C. a., Doyle, F., Epstein, C.B.,

Frietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B.R., Birney, E., *et al.* (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.

Dunham, M.A., Neumann, A.A., Fasching, C.L., and Reddel, R.R. (2000). Telomere maintenance by telomerase and by recombination can coexist in human cells. *Nature Genetics* 26, 447–450.

Ehlers, B., Spieß, K., Leendertz, F., Peeters, M., Boesch, C., Gatherer, D., and McGeoch, D.J. (2010). Lymphocryptovirus phylogeny and the origins of Epstein-Barr virus. *Journal of General Virology* 91, 630–642.

Epstein, M.A. (1975). Recent studies with EB virus. *Yale Journal of Biology and Medicine* 48, 431–437.

Epstein, M.A., and Achong, B.G. (1968). Observations on the Nature of the Herpes-Type EB Virus In Cultured Burkitt Lymphoblasts , Using a Specific Immunofluorescence Test. *Journal of the National Cancer Institute* 40, 609–621.

Epstein, M.A., Achong, B.G., and Barr, Y.M. (1964a). Virus particles in cultured lymphoblasts from Burkitt's Lymphoma. *The Lancet* 283, 702–703.

Epstein, M.A., Henle, G., Achong, B.G., Ch, B., and Barr, Y.M. (1964b). Morphological and biological studies on a virus in cultured lymphoblasts from Burkitt's Lymphoma. *The Journal of Experimental Medicine* 121, 761–770.

Ernst, J., and Kellis, M. (2012). ChromHMM: automating chromatin-state discovery and characterization. *Nature Methods* 9, 215–216.

Faulkner, G.C., Krajewski, a S., and Crawford, D.H. (2000). The ins and outs of EBV infection. *Trends in Microbiology* 8, 185–189.

Feederle, R., Bartlett, E.J., and Delecluse, H.-J. (2010). Epstein-Barr virus genetics: talking about the BAC generation. *Herpesviridae* 1, 6.

Feil, R., Wagner, J., Metzger, D., and Chambon, P. (1997). Regulation of Cre recombinase activity by mutated estrogen receptor ligand-binding domains. *Biochemical And Biophysical Research Communication* 237, 752–757.

Finke, J., Rowe, M., Kallin, B., Ernberg, I., Rosen, A., Dillner, J., and Klein, G. (1987). Monoclonal and Polyclonal Antibodies against Epstein-Barr Virus Nuclear Antigen 5 (EBNA-5) Detect Multiple Protein Species in Burkitt ' s Lymphoma and Lymphoblastoid Cell Lines. *Journal of Virology* 61, 3870–3878.

Gahn, T.A., and Sugden, B. (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.

Giunco, S., Dolcetti, R., Keppel, S., Celeghin, A., Indraccolo, S., Dal Col, J., Mastorci, K., De Rossi, A., Col, J.D., Mastorci, K., Rossi, A. De, Dal Col, J., Mastorci, K., and De Rossi, A. (2013). hTERT inhibition triggers Epstein-Barr virus lytic cycle and apoptosis in immortalized and transformed B cells: A basis for new therapies. *Clinical Cancer Research* 19, 2036–2047.

Giunco, S., Celeghin, a, Ganesin, K., Dolcetti, R., Indraccolo, S., and De Rossi, a (2015). Cross talk between EBV and telomerase: the role of TERT and NOTCH2 in the switch of latent/lytic cycle of the virus. *Cell Death and Disease* 6, e1774.

Gloeckner, C.J., Boldt, K., Schumacher, A., Roepman, R., and Ueffing, M. (2007). A novel

tandem affinity purification strategy for the efficient isolation and characterisation of native protein complexes. *Proteomics* 7, 4228–4234.

Groom, J.R., and Luster, A.D. (2011). CXCR3 ligands: redundant, collaborative and antagonistic functions. *Immunology and Cell Biology* 89, 207–215.

Grossman, S.R., Johannsen, E., Tong, X., Yalamanchili, R., and Kieff, E. (1994). The Epstein-Barr virus nuclear antigen 2 transactivator is directed to response elements by the J kappa recombination signal binding protein. *Proceedings of the National Academy of Sciences of the United States of America* 91, 7568–7572.

Guerrini, a M., Camponeschi, B., Ascenzioni, F., Piccolella, E., and Donini, P. (1993). Subtelomeric as well as telomeric sequences are lost from chromosomes in proliferating B lymphocytes. *Human Molecular Genetics* 2, 455–460.

Gunven, P., Klein, G., Henle, G., Henle, W., and Clifford, P. (1970). Epstein-Barr Virus in Burkitt's Lymphoma and Nasopharyngeal Carcinoma. *Nature* 228, 1053–1056.

Hahn, H., and Palmenberg, A.N.N.C. (1996). Mutational Analysis of the Encephalomyocarditis Virus Primary Cleavage. *Journal of Virology* 70, 6870–6875.

Hammerschmidt, W., and Sugden, B. (2013). Replication of Epstein-Barr Viral DNA. *Cold Spring Harbor Perspectives in Biology* 5, a013029–a013029.

Harada, S., and Kieff, E. (1997). Epstein-Barr Virus Nuclear Protein LP Stimulates EBNA-2 Acidic Domain-Mediated Transcriptional Activation. *Journal of Virology* 71, 6611–6618.

Harlow, E., and Lane, D. (1988). *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory Press).

Hart, S.L., Arancibia-Carcamo, C. V, Wolfert, M.A., Mailhos, C., Reilly, N.J.O., Ali, R.R., Coutelle, C., Seymour, L.W., Thrasher, A.J., and Kinnon, C. (1998). Lipid-mediated Enhancement of Transfection by a Nonviral Integrin-Targeting Vector. *Human Gene Therapy* 9, 575–585.

Harth-Hertle, M.L., Scholz, B. a, Erhard, F., Glaser, L. V, Dölken, L., Zimmer, R., and Kempkes, B. (2013). Inactivation of intergenic enhancers by EBNA3A initiates and maintains polycomb signatures across a chromatin domain encoding CXCL10 and CXCL9. *PLoS Pathogens* 9, e1003638.

zur Hausen, H., Schulte-Holthausen, H., Klein, G., Henle, W., Henle, G., Clifford, P., and Santesson, L. (1970). EBV DNA in Biopsies of Burkitt Tumours and Anaplastic Carcinomas of the Nasopharynx. *Nature* 228, 1056–1058.

Hayward, S.D. (2004). Viral interactions with the Notch pathway. *Seminars in Cancer Biology* 14, 387–396.

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

Henle, G., and Henle, W. (1966). Immunofluorescence in Cells Derived from Burkitt's Lymphoma. *Journal of Bacteriology* 91, 1248–1256.

Henle, G., Henle, W., and Diehl, V. (1968). Relation of Burkitt's tumor-associated herpes-type virus to infectious mononucleosis. *Proceedings of the National Academy of Sciences of the United States of America* 59, 94–101.

Henle, G., Henle, W., Clifford, P., Diehl, V., Kafuko, G.W., Kirya, B., Klein, G., Morrow, R.H., Munube, G.M.R., Pike, P., Tukei, P.M., and Ziegler, J.L. (1969). Antibodies to Epstein-Barr Virus in Burkitt's Lymphoma and Control Groups. *JNCI Journal of the National Cancer Institute* 43, 1147–1157.

Henle, W., Diehl, V., Kohn, G., Zur Hausen, H., and Henle, G. (1967). Herpes-Type Virus and Chromosome Marker in Normal Leukocytes after Growth with Irradiated Burkitt Cells. *Science* 157, 1064–1065.

Hennessy, K., Fennewald, S., and Kieff, E. (1985). A third viral nuclear protein in lymphoblasts immortalized by Epstein-Barr virus. *Proceedings of the National Academy of Sciences of the United States of America* 82, 5944–5948.

Hertle, M.L., Popp, C., Petermann, S., Maier, S., Kremmer, E., Lang, R., Mages, J., and Kempkes, B. (2009). Differential gene expression patterns of EBV infected EBNA-3A positive and negative human B lymphocytes. *PLoS Pathogens* 5, e1000506.

Hickabottom, M., Parker, G. a, Freemont, P., Crook, T., and Allday, M.J. (2002). Two nonconsensus sites in the Epstein-Barr virus oncoprotein EBNA3A cooperate to bind the co-repressor carboxyl-terminal-binding protein (CtBP). *The Journal of Biological Chemistry* 277, 47197–47204.

Hislop, A.D. (2015). Early virological and immunological events in Epstein-Barr virus infection. *Current Opinion in Virology* 15, 75–79.

Hislop, A.D., Taylor, G.S., Sauce, D., and Rickinson, A.B. (2007). Cellular responses to viral infection in humans: lessons from Epstein-Barr virus. *Annual Review of Immunology* 25, 587–617.

Hoffman, M.M., Ernst, J., Wilder, S.P., Kundaje, A., Harris, R.S., Libbrecht, M., Giardine, B., Ellenbogen, P.M., Bilmes, J. a., Birney, E., Hardison, R.C., Dunham, I., Kellis, M., and Noble, W.S. (2013). Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Research* 41, 827–841.

Hoffmeyer, K., Raggioli, A., Rudloff, S., Anton, R., Hierholzer, A., Valle, I. Del, Hein, K., Vogt, R., and Kemler, R. (2012). Wnt/B-Catenin Signaling Regulates Telomerase in Stem Cells and Cancer Cells. *Science* 336, 1549–1554.

Hsu, J.L., and Glaser, S.L. (2000). Epstein – Barr virus-associated malignancies : epidemiologic patterns and etiologic implications. *Critical Reviews in Oncology/Haematology* 34, 27–53.

Hu, B.T., and Insel, R.A. (1999). Up-regulation of telomerase in human B lymphocytes occurs independently of cellular proliferation and with expression of the telomerase catalytic subunit. *European Journal of Immunology* 29, 3745–3753.

Hung, S.C., Kang, M.S., and Kieff, E. (2001). Maintenance of Epstein-Barr virus (EBV) oriP-based episomes requires EBV-encoded nuclear antigen-1 chromosome-binding domains, which can be replaced by high-mobility group-I or histone H1. *Proceedings of the National Academy of Sciences of the United States of America* 98, 1865–1870.

Hutt-Fletcher, L.M. (2007). Epstein-Barr Virus Entry. *Journal of Virology* 81, 7825–7832.

Isaksson, A., Berggren, M., and Ricksten, A. (2003). Epstein-Barr virus U leader exon contains an internal ribosome entry site. *Oncogene* 22, 572–581.

Jarrett, R.F., Gallagher, A., Jones, D.B., Alexander, F.E., Krajewski, A.S., Kelsey, A., Adams, J., Angus, B., Gledhill, S., Wright, D.H., and *et al.* (1991). Detection of Epstein-Barr

virus genomes in Hodgkin's disease: relation to age. *J Clin Pathol* 44, 844–848.

Ji, H., Jiang, H., Ma, W., Johnson, D.S., Myers, R.M., and Wong, W.H. (2008). An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nature Biotechnology* 26, 1293–1300.

Jiang, S., Willox, B., Zhou, H., Holthaus, A.M., Wang, A., Shi, T.T., Maruo, S., Kharchenko, P. V., Johannsen, E.C., Kieff, E., and Zhao, B. (2014). Epstein-Barr virus nuclear antigen 3C binds to BATF/IRF4 or SPI1/IRF4 composite sites and recruits Sin3A to repress CDKN2A. *Proceedings of the National Academy of Sciences of the United States of America* 111, 421–426.

Jiménez-Ramírez, C., Brooks, A.J., Forshell, L.P., Yakimchuk, K., Zhao, B., Fulgham, T.Z., and Sample, C.E. (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.

Joab, I., Rowe, D.T., Bodescot, M., Nicolas, J.C., Farrell, P.J., and Perricaudet, M. (1987). Mapping of the gene coding for Epstein-Barr virus-determined nuclear antigen EBNA3 and its transient overexpression in a human cell line by using an adenovirus expression vector. *Journal of Virology* 61, 3340–3344.

Jochum, S., Moosmann, A., Lang, S., Hammerschmidt, W., and Zeidler, R. (2012). The EBV immunoevasins vIL-10 and BNLF2a protect newly infected B cells from immune recognition and elimination. *PLoS Pathogens* 8.

Johannsen, E., Miller, C.L., Grossman, S.R., and Kieff, E. (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., Luftig, M., Chase, M.R., Weicksel, S., Cahir-McFarland, E., Illanes, D., Sarracino, D., and Kieff, E. (2004). Proteins of purified Epstein-Barr virus. *Proceedings of the National Academy of Sciences of the United States of America* 101, 16286–16291.

Jones, P.L., Veenstra, G.J.C., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *19*, 187–191.

Kaiser, C., Laux, G., Eick, D., Jochner, N., Bornkamm, G.W., and Kempkes, B. (1999). The Proto-Oncogene c-myc Is a Direct Target Gene of Epstein-Barr Virus Nuclear Antigen 2. *Journal of Virology* 73, 4481–4484.

Kalchschmidt, J.S., Rogers, R.B., Paschos, K., Gillman, A.C.T.T., Styles, C.T., Kellam, P., Allday, M.J., Bashford-Rogers, R., Paschos, K., Gillman, A.C.T.T., Styles, C.T., Kellam, P., and Allday, M.J. (2016a). Epstein-Barr virus nuclear protein EBNA3C directly induces expression of AID and somatic mutations in B cells. *The Journal of Experimental Medicine* 213, 921–928.

Kalchschmidt, J.S., Gillman, A.C.T., Paschos, K., and Bazot, Q. (2016b). EBNA3C Directs Recruitment of RBPJ (CBF1) to Chromatin during the Process of Gene Repression in EBV Infected B Cells. *PLoS Pathogens* 12.

Kamranvar, S. a, and Masucci, M.G. (2011). The Epstein-Barr virus nuclear antigen-1 promotes telomere dysfunction via induction of oxidative stress. *Leukemia : Official Journal of the Leukemia Society of America, Leukemia Research Fund, U.K* 25, 1017–1025.

Kamranvar, S. a, Chen, X., and Masucci, M.G. (2013). Telomere dysfunction and activation of alternative lengthening of telomeres in B-lymphocytes infected by Epstein-Barr virus.

Oncogene 32, 5522–5530.

Kanda, T., Yajima, M., Ahsan, N., Tanaka, M., and Takada, K. (2004). Production of High-Titer Epstein-Barr Virus Recombinants Derived from Akata Cells by Using a Bacterial Artificial Chromosome System Production of High-Titer Epstein-Barr Virus Recombinants Derived from Akata Cells by Using a Bacterial Artificial Chromosome. *Journal of Virology* 78, 7004–7015.

Kataoka, H., Tahara, H., Watanabe, T., Sugawara, M., Ide, T., Goto, M., Furuichi, Y., and Sugimoto, M. (1998). immortalization of immunologically committed Epstein-Barr virus-transformed human B-lymphoblastoid cell lines accompanied by a strong telomerase activity. *Differentiation* 62, 203–211.

Kempkes, B., and Ling, P.D. (2015). Epstein Barr Virus Volume 2: One Herpes Virus: Many Diseases. C. Münz, ed. (Cham: Springer International Publishing), pp. 35–59.

Khan, G., Miyashita, E.M., Yang, B., Babcock, G.J., and Thorley-lawson, D. a (1996). Is EBV Persistence In Vivo a Model for B Cell Homeostasis ? *Immunity* 5, 173–179.

Khanna, R., Burrows, S.R., Kurilla, M.G., Jacob, C. a, Misko, I.S., Sculley, T.B., Kieff, E., and Moss, D.J. (1992). Localization of Epstein-Barr virus cytotoxic T cell epitopes using recombinant vaccinia: implications for vaccine development. *The Journal of Experimental Medicine* 176, 169–176.

Kia, S.K., Gorski, M.M., Giannakopoulos, S., and Verrijzer, C.P. (2008). SWI / SNF Mediates Polycomb Eviction and Epigenetic Reprogramming of the INK4b-ARF-INK4a Locus . *Molecular and Cellular Biology* 28, 3457–3464.

Kim, N.W., Piatyszek, M.A., Prowse, K.R., Harley, C.B., West, M.D., Ho, P.L.C., Coviello, G.M., Wright, W.E., Weinrich, S.L., and Shay, J.W. (1994). Specific association of human telomerase activity with immortal cells and cancer. *Science* 266, 2011–2015.

Kintner, R., and Sugden, B. (1979). The Structure of the Termini of the DNA of Epstein-Barr Virus. *Cell* 17, 661–671.

Klein, G. (1983). Specific chromosomal translocations and the genesis of B-cell-derived tumors in mice and men. *Cell* 32, 311–315.

Klein, E., Kis, L.L., and Klein, G. (2007). Epstein-Barr virus infection in humans: from harmless to life endangering virus-lymphocyte interactions. *Oncogene* 26, 1297–1305.

Knight, J.S., and Robertson, E.S. (2004). Epstein-Barr Virus Nuclear Antigen 3C Regulates Cyclin A / p27 Complexes and Enhances Cyclin A-Dependent Kinase Activity. *Journal of Virology* 78, 1981–1991.

Knight, J.S., Lan, K., Subramanian, C., Robertson, E.S., Erle, S., and Robertson, E.S. (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.

Kopan, R., and Ilagan, M.X.G. (2009). The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell* 137, 216–233.

Krauer, K., Buck, M., Flanagan, J., Belzer, D., and Sculley, T. (2004a). Identification of the nuclear localization signals within the Epstein–Barr virus EBNA-6 protein. *Journal of General Virology* 85, 165–172.

Krauer, K.G., Buck, M., Belzer, D.K., Flanagan, J., Chojnowski, G.M., and Sculley, T.B.

(2004b). The Epstein-Barr virus nuclear antigen-6 protein co-localizes with EBNA-3 and survival of motor neurons protein. *Virology* 318, 280–294.

Laichalk, L.L., and Thorley-Lawson, D. a (2005). Terminal Differentiation into Plasma Cells Initiates the Replicative Cycle of Epstein-Barr Virus In Vivo Terminal Differentiation into Plasma Cells Initiates the Replicative Cycle of Epstein-Barr Virus In Vivo. *Journal of Virology* 79, 1296–1307.

Lalioti, M.D., and Heath, J.K. (2001). A new method for generating point mutations in bacterial artificial chromosomes by homologous recombination in *Escherichia coli*. 29, 1–6.

Landt, S., Marinov, G., Kundaje, A., Kheradpour, P., Pauli, F., Batzoglou, S., Bernstein, B., Bickel, P.J., Brown, J.B., Cayting, P., Chen, Y., DeSalvo, G., Epstein, C.B., Fisher-Aylor, K.I., Euskirchen, G., Gerstein, M., Gertz, J., Hartemink, A.J., Hoffman, M.M., Iyer, V.R., Jung, Y.L., Karmakar, S., Kellis, M., Kharchenko, P. V., Li, Q., Liu, T., Liu, X.S., Ma, L., Milosavljevic, A., Myers, R.M., Park, P.J., Pazin, M.J., Perry, M.D., Raha, D., Reddy, T.E., Rozowsky, J., Shores, N., Sidow, A., Slattey, M., Stamatoyannopoulos, J. a., Tolsorukov, M.Y., White, K.P., Xi, S., Farnham, P.J., Lieb, J.D., Wold, B., and Snyder, M. (2012). ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome ...* 1813–1831.

De Lange, T. (2005). Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes and Development* 19, 2100–2110.

Laux, G., Dugrillon, F., Eckert, C., Adam, B., Zimmer-strobl, U., and Bornkamm, G.W. (1994). Identification and Characterization of an Epstein-Barr Virus Nuclear Antigen 2-Responsive cis Element in the Bidirectional Promoter Region of Latent Membrane Protein and Terminal Protein 2 Genes. *Journal of Virology* 68, 6947–6958.

LeBlanc, H.N., and Ashkenazi, A. (2003). Apo2L/TRAIL and its death and decoy receptors. *Cell Death and Differentiation* 10, 66–75.

Lee, S., Sakakibara, S., Maruo, S., Zhao, B., Calderwood, M.A., Holthaus, A.M., Lai, C.-Y., Takada, K., Kieff, E., and Johannsen, E. (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.

Levine, M. (2011). Paused RNA polymerase II as a developmental checkpoint. *Cell* 145, 502–511.

Li, Y., and Tergaonkar, V. (2014). Noncanonical functions of telomerase: Implications in telomerase-targeted cancer therapies. *Cancer Research* 74, 1639–1644.

Li, Q., Spriggs, M.K., Kovats, S., Turk, S.M., Comeau, M.R., Nepom, B., and Hutt-Fletcher, L.M. (1997). Epstein-Barr virus uses HLA class II as a cofactor for infection of B lymphocytes. *Journal of Virology* 71, 4657–4662.

Lieberman, P.M. (2013). Keeping it quiet: chromatin control of gammaherpesvirus latency. *Nature Reviews. Microbiology* 11, 863–875.

Long, H.M., Taylor, G.S., and Rickinson, A.B. (2011). Immune defence against EBV and EBV-associated disease. *Current Opinion in Immunology* 23, 258–264.

Longnecker, R.M., Kieffe, E., and Cohen, J.I. (2013). Epstein-Barr Virus. In *Fields Virology*, D.M. Knipe, and P.M. Howley, eds. (Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins), pp. 1898–1959.

Lu, F., Chen, H.-S., Kossenkov, A. V., DeWisleare, K., Won, K.-J., and Lieberman, P.M.

(2016). EBNA2 Drives Formation of New Chromosome Binding Sites and Target Genes for B-Cell Master Regulatory Transcription Factors RBP-jk and EBF1. *PLOS Pathogens* 12, e1005339.

Luke, G.A., Felipe, P. De, Lukashev, A., Kallioinen, S.E., Bruno, E.A., and Ryan, M.D. (2008). Occurrence, function and evolutionary origins of “2A-like” sequences in virus genomes. *Journal of General Virology* 89, 1036–1042.

Luzuriaga, K., and Sullivan, J.L. (2010). Infectious mononucleosis. *The New England Journal of Medicine* 362, 1993–2000.

Mahalingam, D., Szegezdi, E., Keane, M., Jong, S. de, and Samali, A. (2009). TRAIL receptor signalling and modulation: Are we on the right TRAIL? *Cancer Treatment Reviews* 35, 280–288.

Maruo, S., Johannsen, E., Illanes, D., Cooper, A., and Kieff, E. (2003). Epstein-Barr Virus nuclear protein EBNA3A is critical for maintaining lymphoblastoid cell line growth. *Journal of Virology* 77, 10437–10447.

Maruo, S., Johannsen, E., Illanes, D., Cooper, A., Zhao, B., and Kieff, E. (2005). Epstein-Barr Virus Nuclear Protein 3A Domains Essential for Growth of Lymphoblasts: Transcriptional Regulation through RBP-J κ / CBF1 Is Critical Epstein-Barr Virus Nuclear Protein 3A Domains Essential for Growth of Lymphoblasts: Transcriptional Regulat. *Journal of Virology* 79, 10171–10179.

Maruo, S., Wu, Y., Ishikawa, S., Kanda, T., Iwakiri, D., and Takada, K. (2006). Epstein-Barr virus nuclear protein EBNA3C is required for cell cycle progression and growth maintenance of lymphoblastoid cells. *Proceedings of the National Academy of Sciences of the United States of America* 103, 19500–19505.

Maruo, S., Wu, Y., Ito, T., Kanda, T., Kieff, E.D., and Takada, K. (2009). Epstein-Barr virus nuclear protein EBNA3C residues critical for maintaining lymphoblastoid cell growth. *Proceedings of the National Academy of Sciences of the United States of America* 106, 4419–4424.

Maruo, S., Zhao, B., Johannsen, E., Kieff, E., Zou, J., and Takada, K. (2011). Epstein-Barr virus nuclear antigens 3C and 3A maintain lymphoblastoid cell growth by repressing p16^{INK4A} and p14^{ARF} expression. *Proceedings of the National Academy of Sciences* 108, 1919–1924.

Maunder, M.J., Pettit, L., and Rowe, M. (1994). Precipitation of the Epstein-Barr virus protein EBNA 2 by an EBNA 3c-specific monoclonal antibody. *Journal of General Virology* 75, 769–778.

McClellan, M.J., Khasnis, S., Wood, C.D., Palermo, R.D., Schlick, S.N., Kanhere, A.S., Jenner, R.G., and West, M.J. (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.

McClellan, M.J., Wood, C.D., Ojieniyi, O., Cooper, T.J., Kanhere, A., Arvey, A., Webb, H.M., Palermo, R.D., Harth-Hertle, M.L., Kempkes, B., Jenner, R.G., and West, M.J. (2013). Modulation of enhancer looping and differential gene targeting by Epstein-Barr virus transcription factors directs cellular reprogramming. *PLoS Pathogens* 9, e1003636.

Micheau, O., Shirley, S., and Dufour, F. (2013). Death receptors as targets in cancer. *British Journal of Pharmacology* 169, 1723–1744.

Mouzakiti, A., and Packham, G. (2003). Regulation of tumour necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced apoptosis in Burkitt's lymphoma cell lines. *British Journal of Haematology* 122, 61–69.

Mumm, J.S., Schroeter, E.H., Saxena, M.T., Griesemer, A., Tian, X., Pan, D.J., Ray, W.J., Kopan, R., Program, N., and Louis, S. (2000). A Ligand-Induced Extracellular Cleavage Regulates γ -Secretase-like Proteolytic Activation of Notch1. *Molecular Cell* 5, 197–206.

Münz, C., Bickham, K.L., Subklewe, M., Tsang, M.L., Chahroudi, A., Kurilla, M.G., Zhang, D., O'Donnell, M., and Steinman, R.M. (2000). Human CD4(+) T lymphocytes consistently respond to the latent Epstein-Barr virus nuclear antigen EBNA1. *The Journal of Experimental Medicine* 191, 1649–1660.

Murray, R.J., Kurilla, M.G., Brooks, J.M., Thomas, W.A., Rowe, M., Kieff, E., and Rickinson, A.B. (1992). Identification of target antigens for the human cytotoxic T cell response to Epstein-Barr virus (EBV): implications for the immune control of EBV-positive malignancies. *The Journal of Experimental Medicine* 176, 157–168.

Nakayama, J., Tahara, H., Tahara, E., Saito, M., Ito, K., Nakamura, H., Nakanishi, T., Tahara, E., Ide, T., and Ishikawa, F. (1998). Telomerase activation by hTERT in human normal fibroblasts and hepatocellular carcinomas. *Nature Genetics* 18, 65–68.

Nitsche, F., Bell, A., and Rickinson, A. (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.

Ohashi, M., Holthaus, A.M., Calderwood, M.A., Lai, C.-Y., Krastins, B., Sarracino, D., and Johannsen, E. (2015). The EBNA3 family of Epstein-Barr virus nuclear proteins associates with the USP46/USP12 deubiquitination complexes to regulate lymphoblastoid cell line growth. *PLoS Pathogens* 11, e1004822.

Oswald, F., Tauber, B., Dobner, T., Bourteele, S., Kostezka, U., Adler, G., and Liptay, S. (2001). p300 Acts as a Transcriptional Coactivator for Mammalian Notch-1. *Molecular and Cellular Biology* 21, 7761–7774.

Pallesen, G., Hamilton-Dutoit, S.J., Rowe, M., and Young, L.S. (1991). Expression of Epstein-Barr virus latent gene products in tumour cells of Hodgkin's disease. *The Lancet* 337, 320–322.

Panagopoulos, D., Victoratos, P., Alexiou, M., Kollias, G., and Mosialos, G. (2004). Comparative analysis of signal transduction by CD40 and the Epstein-Barr virus oncoprotein LMP1 in vivo. *Journal of Virology* 78, 13253–13261.

Parker, G.A., Crook, T., Bain, M., Sara, E.A., Farrell, P.J., and Allday, M.J. (1996). Epstein-Barr virus nuclear antigen (EBNA)3C is an immortalizing oncoprotein with similar properties to adenovirus E1A and papillomavirus E7. *Oncogene* 13, 2541–2549.

Paschos, K., Parker, G. a., Watanatanasup, E., White, R.E., and Allday, M.J. (2012). BIM promoter directly targeted by EBNA3C in polycomb-mediated repression by EBV. *Nucleic Acids Research* 40, 7233–7246.

Paschos, K., Bazot, Q., Ho, G., Parker, G.A., Lees, J., Barton, G., and Allday, M.J. (2016). Core binding factor (CBF) is required for Epstein-Barr virus EBNA3 proteins to regulate target gene expression. *Nucleic Acids Research* 1–16.

Peng, R., Tan, J.I.E., and Ling, P.D. (2000). Conserved Regions in the Epstein-Barr Virus Leader Protein Define Distinct Domains Required for Nuclear Localization and

Transcriptional Cooperation with EBNA2. *Journal of Virology* 74, 9953–9963.

Petrara, M.R., Giunco, S., Serraino, D., Dolcetti, R., and De Rossi, A. (2015). Post-transplant lymphoproliferative disorders: From epidemiology to pathogenesis-driven treatment. *Cancer Letters* 369, 37–44.

Petti, L., and Kieff, E. (1988). A sixth Epstein-Barr virus nuclear protein (EBNA3B) is expressed in latently infected growth-transformed lymphocytes. 62, 2173–2178.

Petti, L., Sample, J., Wang, F., and Kieff, E. (1988). A Fifth Epstein-Barr Virus Nuclear Protein (EBNA3C) Is Expressed in Latently Infected Growth-Transformed Lymphocytes. *Journal of Virology* 62, 1330–1338.

Pope, J.H., Horne, M.K., and Scott, W. (1968). Transformation of foetal human leukocytes in vitro by filtrates of a human leukaemic cell line containing herpes-like virus. *International Journal of Cancer* 3, 857–866.

Price, A.M., and Luftig, M.A. (2014). Dynamic Epstein-Barr virus gene expression on the path to B-cell transformation (Elsevier Inc.).

Pringle, F.M., Gordon, K.H.J., Hanzlik, T.N., Kalkmakoff, J., Scotti, P.D., and Ward, V.K. (1999). A novel capsid expression strategy for *Thomaspoxvirus* (Tetraviridae). *Journal of Virology* 80, 1855–1863.

Raab-Traub, N., and Flynn, K. (1986). The structure of the termini of the Epstein-Barr virus as a marker of clonal cellular proliferation. *Cell* 47, 883–889.

Radkov, S.A., Touitou, R., Brehm, A., West, M., Kouzarides, T., Allday, M.J., Rowe, M., West, M., Kouzarides, T., and Allday, M.J. (1999). Epstein-Barr Virus Nuclear Antigen 3C Interacts with Histone Deacetylase To Repress Transcription. *Journal of Virology* 73, 5688–5697.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., and Aiden, E.L. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159, 1665–1680.

Rastelli, J., Hömig-Hölzel, C., Seagal, J., Müller, W., Hermann, A.C., Rajewsky, K., and Zimmer-Strobl, U. (2008). LMP1 signaling can replace CD40 signaling in B cells in vivo and has unique features of inducing class-switch recombination to IgG1. *Blood* 111, 1448–1455.

Rickinson, A.B. (1990). Epstein-Barr Virus Nuclear Antigen 2 Induces Expression of the Virus-Encoded Latent Membrane Protein. *Journal of Virology* 64, 2126–2134.

Rickinson, A.B., and Moss, D.J. (1997). Human cytotoxic T lymphocyte responses to Epstein-Barr virus infection. *Annual Review of Immunology* 15, 405–431.

Ricksten, A., Kallin, B., Alexander, H., Dillner, J., Fåhræus, R., Klein, G., Lerner, R., and Rymo, L. (1988). BamHI E region of the Epstein-Barr virus genome encodes three transformation-associated nuclear proteins. *Proceedings of the National Academy of Sciences of the United States of America* 85, 995–999.

Ring, C.J.A. (1994). The B cell-immortalizing functions of Epstein-Barr virus. *Journal of General Virology* 75, 1–13.

Robbiani, D.F., Deroubaix, S., Feldhahn, N., Oliveira, T.Y., Callen, E., Wang, Q., Jankovic, M., Silva, I.T., Rommel, P.C., Bosque, D., Eisenreich, T., Nussenzweig, A., and Nussenzweig, M.C. (2015). Plasmodium Infection Promotes Genomic Instability and AID-

Dependent B Cell Lymphoma. *Cell* 162, 727–737.

Robertson, E.S., Grossman, S., Johannsen, E., Miller, C., Lin, J., Tomkinson, B., and Kieff, E. (1995). Epstein-Barr Virus Nuclear Protein 3C Modulates Transcription through Interaction with the Sequence-Specific DNA-Binding Protein Jk. *Journal of Virology* 69, 3108–3116.

Robertson, E.S., Lin, J., and Kieff, E. (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.

Roughan, J.E., and Thorley-Lawson, D.A. (2009). The intersection of Epstein-Barr virus with the germinal center. *Journal of Virology* 83, 3968–3976.

Le Roux, A., Berebbi, M., Moukaddem, M., Perricaudet, M., and Joab, I. (1993). Identification of a Short Amino Acid Sequence Essential for Efficient Nuclear Targeting of the Epstein-Barr Virus Nuclear Antigen 3A. *Journal of Virology* 67, 1716–1720.

Le Roux, A., Kerdiles, B., Walls, D., Dedieu, J.-F., and Perricaudet, M. (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.

Rowe, M., and Zuo, J. (2010). Immune responses to Epstein-Barr virus: molecular interactions in the virus evasion of CD8+ T cell immunity. *Microbes and Infection / Institut Pasteur* 12, 173–181.

Rowe, M., Young, L.S., Crocker, J., Stokes, H., Henderson, S., and Rickinson, A.B. (1991). Epstein-Barr virus (EBV)-associated lymphoproliferative disease in the SCID mouse model: implications for the pathogenesis of EBV-positive lymphomas in man. *The Journal of Experimental Medicine* 173, 147–158.

Rowe, M., Lear, A.L., Croom-Carter, D., Davies, A.H., and Rickinson, A.B. (1992). Three pathways of Epstein-Barr virus gene activation from EBNA1-positive latency in B lymphocytes. *Journal of Virology* 66, 122–131.

Saha, A., and Robertson, E.S. (2011). Epstein-Barr virus-associated B-cell lymphomas: pathogenesis and clinical outcomes. *Clinical Cancer Research* 17, 3056–3063.

Saha, A., Murakami, M., Kumar, P., Bajaj, B., Sims, K., and Robertson, E.S. (2009). Epstein-Barr virus nuclear antigen 3C augments Mdm2-mediated p53 ubiquitination and degradation by deubiquitinating Mdm2. *Journal of Virology* 83, 4652–4669.

Schlager, S., and Speck, S.H. (1996). Transcription of the Epstein-Barr Virus Nuclear Antigen 1 (EBNA1) Gene Occurs before Induction of the BCR2 (Cp) EBNA Gene Promoter during the Initial Stages of Infection in B Cells. *Journal of Virology* 70, 3561–3570.

Schmidt, T.G.M., and Skerra, A. (2007). The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nature Protocols* 2, 1528–1535.

Schmidt, S.C.S., Jiang, S., Zhou, H., Willox, B., Holthaus, A.M., Kharchenko, P. V, Johannsen, E.C., Kieff, E., and Zhao, B. (2015). Epstein-Barr virus nuclear antigen 3A partially coincides with EBNA3C genome-wide and is tethered to DNA through BATF complexes. *Proceedings of the National Academy of Sciences of the United States of America* 112, 554–559.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols* 3, 1101–1108.

Shannon-Lowe, C., and Rowe, M. (2011). Epstein-Barr Virus Infection of Polarized Epithelial Cells Via the Basolateral Surface By Memory B Cell-Mediated Transfer Infection. *PLoS Pathogens* 7.

Shannon-Lowe, C., and Rowe, M. (2014). Epstein Barr virus entry; Kissing and conjugation. *Current Opinion in Virology* 4, 78–84.

Shay, J.W., Shay, J.W., Zou, Y., Zou, Y., Hiyama, E., Hiyama, E., Wright, W.E., and Wright, W.E. (2001). Telomerase and cancer. *Human Molecular Genetics* 10, 677–685.

Sheridan, J.P., Marsters, S.A., Pitti, R.M., Gurney, A., Skubatch, M., Baldwin, D., Ramakrishnan, L., Gray, C.L., Baker, K., Wood, W.I., Goddard, A.D., Godowski, P., and Ashkenazi, A. (1997). Control of TRAIL-Induced Apoptosis by a Family of Signaling and Decoy Receptors. *Science* 277.

Shirley, S., and Micheau, O. (2013). Targeting c-FLIP in cancer. *Cancer Letters* 332, 141–150.

Simon, J. a, and Kingston, R.E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature Reviews. Molecular Cell Biology* 10, 697–708.

Sjöblom, a, Nerstedt, a, Jansson, a, and Rymo, L. (1995). Domains of the Epstein-Barr virus nuclear antigen 2 (EBNA2) involved in the transactivation of the latent membrane protein 1 and the EBNA Cp promoters. *The Journal of General Virology* 76 (Pt 11, 2669–2678.

Skalska, L., White, R.E., Franz, M., Ruhmann, M., and Allday, M.J. (2010). Epigenetic repression of p16(INK4A) by latent Epstein-Barr virus requires the interaction of EBNA3A and EBNA3C with CtBP. *PLoS Pathogens* 6, e1000951.

Skalska, L., White, R.E., Parker, G.A., Sinclair, A.J., Paschos, K., and Allday, M.J. (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 Pathogens* 9, e1003187.

Snow, A.L., Lambert, S.L., Natkunam, Y., Esquivel, C.O., Krams, S.M., and Martinez, O.M. (2006). EBV can protect latently infected B cell lymphomas from death receptor-induced apoptosis. *Journal of Immunology (Baltimore, Md. : 1950)* 177, 3283–3293.

Spyrou, C., Stark, R., Lynch, A.G., and Tavaré, S. (2009). BayesPeak : Bayesian analysis of ChIP-seq data. *BayesPeak: Bayesian Analysis of ChIP-Seq Data* 10.

Straus, S.E., Cohen, J.I., Tosato, G., and Meier, J. (1993). Epstein-Barr Virus Infections : Biology, Pathogenesis, and Management. *Annals of Internal Medicine* 118, 45–58.

Strooper, B. De, Annaert, W., Cupers, P., Saftig, P., Craessaerts, K., Mumm, J.S., Schroeter, E.H., Schrijvers, V., Wolfe, M.S., Ray, W.J., Goate, A., and Kopan, R. (1999). A presenilin-1-dependent γ -secretase-like protease mediates release of Notch intracellular domain. *Nature* 398, 518–522.

Styczynski, J., Velden, W. Van Der, Fox, C.P., Engelhard, D., de la Camara, R., Cordonnier, C., and Ljungman, P. (2016). Management of Epstein-Barr Virus infections and post-transplant lymphoproliferative disorders in patients after allogeneic hematopoietic stem cell transplantation: Sixth European Conference on Infections in Leukemia (ECIL-6) guidelines. *Haematologica* 101, 803–811.

Sugden, B., and Warren, N. (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.

- Sugimoto, M., Ide, T., Goto, M., and Furuichi, Y. (1999). Reconsideration of senescence, immortalization and telomere maintenance of Epstein-Barr virus-transformed human B-lymphoblastoid cell lines. *Mechanisms of Ageing and Development* 107, 51–60.
- Sugimoto, M., Tahara, H., Ide, T., and Furuichi, Y. (2004). Steps involved in immortalization and tumorigenesis in human B-lymphoblastoid cell lines transformed by Epstein-Barr virus. *Cancer Research* 64, 3361–3364.
- Sung, N.S., Kenney, S., Gutsch, D., and Joseph, S. (1991). EBNA-2 Transactivates a Lymphoid-Specific Enhancer in the BamHI C Promoter of Epstein-Barr Virus. *Journal of Virology* 65, 2164–2169.
- Tafuku, S., Matsuda, T., Kawakami, H., Tomita, M., Yagita, H., and Mori, N. (2006). Potential mechanism of resistance to TRAIL-induced apoptosis in Burkitt's lymphoma. *European Journal of Haematology* 76, 64–74.
- Tahara, H., Tokutake, Y., Maeda, S., Kataoka, H., Watanabe, T., Satoh, M., Matsumoto, T., Sugawara, M., Ide, T., Goto, M., Furuichi, Y., and Sugimoto, M. (1997). Abnormal telomere dynamics of B-lymphoblastoid cell strains from Werner's syndrome patients transformed by Epstein-Barr virus. *Oncogene* 15, 1911–1920.
- Tao, Q., Young, L.S., Woodman, C.B., and Murray, P.G. (2006). Epstein-Barr virus (EBV) and its associated human cancers - Genetics, epigenetics, pathobiology and novel therapeutics. *Frontiers in Bioscience* 11, 2672–2713.
- Temple, R.M., Zhu, J., Budgeon, L., Christensen, N.D., Meyers, C., and Sample, C.E. (2014). Efficient replication of Epstein-Barr virus in stratified epithelium in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 111, 16544–16549.
- Terrin, L., Dolcetti, R., Corradini, I., Indraccolo, S., Dal Col, J., Bertorelle, R., Bonaldi, L., Esposito, G., and De Rossi, A. (2007). hTERT inhibits the Epstein-Barr virus lytic cycle and promotes the proliferation of primary B lymphocytes: Implications for EBV-driven lymphomagenesis. *International Journal of Cancer* 121, 576–587.
- Terrin, L., Dal Col, J., Rampazzo, E., Zancai, P., Pedrotti, M., Ammirabile, G., Bergamin, S., Rizzo, S., Dolcetti, R., and De Rossi, A. (2008). Latent membrane protein 1 of Epstein-Barr virus activates the hTERT promoter and enhances telomerase activity in B lymphocytes. *Journal of Virology* 82, 10175–10187.
- Thomas, A.J., Hotchin, N.A., Allday, M.J., Amlot, P., Rose, M., Yacoub, M., and Crawford, D.H. (1990). Immunohistology of Epstein-Barr Virus-associated antigens in B cell disorders from immunocompromised individuals. *Transplantation* 49, 944–953.
- Thorley-Lawson, D.A. (2001). Epstein-Barr virus: exploiting the immune system. *Nature Reviews. Immunology* 1, 75–82.
- Thorley-Lawson, D.A., and Allday, M.J. (2008). The curious case of the tumour virus: 50 years of Burkitt's lymphoma. *Nature Reviews. Microbiology* 6, 913–924.
- Thorley-Lawson, D.A., and Gross, A. (2004). Persistence of the Epstein-Barr virus and the origins of associated lymphomas. *The New England Journal of Medicine* 350, 1328–1337.
- Thorley-Lawson, D.A., Hawkins, J.B., Tracy, S.I., and Shapiro, M. (2013). The pathogenesis of Epstein-Barr virus persistent infection. *Current Opinion in Virology* 3, 227–232.
- Tomkinson, B., and Kieff, E. (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., Robertson, E., and Kieff, E. (1993). Epstein-Barr virus nuclear proteins EBNA-3A and EBNA-3C are essential for B-lymphocyte growth transformation. *Journal of Virology* 67, 2014–2025.

Torgbor, C., Awuah, P., Deitsch, K., Kalantari, P., Duca, K.A., and Thorley-Lawson, D.A. (2014). A Multifactorial Role for *P. falciparum* Malaria in Endemic Burkitt's Lymphoma Pathogenesis. *PLoS Pathogens* 10, e1004170.

Toutou, R., Hickabottom, M., Parker, G.A., Crook, T.I.M., Allday, M.J., and Crook, T.I.M. (2001a). Physical and Functional Interactions between the Corepressor CtBP and the Epstein-Barr Virus Nuclear Antigen EBNA3C. *Journal of Virology* 75, 7749–7755.

Toutou, R., Hickabottom, M., Parker, G.A., Allday, M.J., and Crook, T.I.M. (2001b). Physical and Functional Interactions between the Corepressor CtBP and the Epstein-Barr Virus Nuclear Antigen EBNA3C. *Journal of Virology* 75, 7749–7755.

Vockerodt, M., Cader, F.Z., Shannon-Lowe, C., and Murray, P. (2014). Epstein-Barr virus and the origin of hodgkin lymphoma. *Chinese Journal of Cancer* 33, 591–597.

Walling, D.M., Andritsos, L.A., Etienne, W., Payne, D.A., Aronson, J.F., Flaitz, C.M., and Nichols, C.M. (2004). Molecular markers of clonality and identity in Epstein-Barr virus-associated B-cell lymphoproliferative disease. *Journal of Medical Virology* 74, 94–101.

Waltzer, L., Logeat, F., Brou, C., Israel, A., Sergeant, A., and Manet, E. (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. *The EMBO Journal* 13, 5633–5638.

Waltzer, L., Perricaudet, M., Sergeant, a, and Manet, E. (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., Welch, R., Zhao, B., Ta, T., Keleş, S., and Johannsen, E. (2015). EBNA3 proteins regulate EBNA2 binding to distinct RBPJ genomic sites. *Journal of Virology* JVI.02737-15.

Wang, F., Gregory, C.D., Rowe, M., Rickinson, A.B., Wang, D., Birkenbach, M., Kikutani, H., Kishimoto, T., and Kieff, E. (1987). Epstein-Barr virus nuclear antigen 2 specifically induces expression of the B-cell activation antigen CD23. *Proceedings of the National Academy of Sciences* 84, 3452–3456.

Wang, F., Tsang, S., Kurilla, M.G., Cohen, J.I., and Kieff, E. (1990). Epstein-Barr Virus Nuclear Antigen 2 Transactivates Latent Membrane Protein LMP1. *Journal of Virology* 64, 3407–3416.

Wang, J., Xie, L.Y., Allan, S., Beach, D., and Hannon, G.J. (1998a). Myc activates telomerase. *Genes & Development* 12, 1769–1774.

Wang, X., Kenyon, W.J., Li, Q., Müllberg, J., and Hutt-Fletcher, L.M. (1998b). Epstein-Barr virus uses different complexes of glycoproteins gH and gL to infect B lymphocytes and epithelial cells. *Journal of Virology* 72, 5552–5558.

Watanatanasup, E. (2011). Interactions between the Epstein-Barr virus (EBV) nuclear proteins EBNA3A , EBNA3B , and EBNA3C. Imperial College London.

Weinrich, S.L., Pruzan, R., Ma, L., Ouellette, M., Tesmer, V.M., Holt, S.E., Bodnar, A.G., Lichtsteiner, S., Kim, N.W., Trager, J.B., Taylor, R.D., Carlos, R., Andrews, W.H., Wright, W.E., Shay, J.W., Harley, C.B., and Morin, G.B. (1997). Reconstitution of human telomerase

with the template RNA component hTR and the catalytic protein subunit hTRT. *Nature Genetics* 17, 57–61.

West, M.J., Webb, H.M., Sinclair, A.J., and Woolfson, D.N. (2004). Biophysical and Mutational Analysis of the Putative bZIP Domain of Epstein-Barr Virus EBNA 3C. *Journal of Virology* 78, 9431–9445.

White, R.E. (2003). Generation and precise modification of a herpesvirus saimiri bacterial artificial chromosome demonstrates that the terminal repeats are required for both virus production and episomal persistence. *Journal of General Virology* 84, 3393–3403.

White, R.E., Groves, I.J., Turro, E., Yee, J., Kremmer, E., and Allday, M.J. (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.

White, R.E., Rämer, P.C., Naresh, K.N., Meixlsperger, S., Pinaud, L., Rooney, C., Savoldo, B., Coutinho, R., Bödör, C., Gribben, J., Ibrahim, H.A., Bower, M., Nourse, J.P., Gandhi, M.K., Middeldorp, J., Cader, F.Z., Murray, P., Münz, C., and Allday, M.J. (2012). EBNA3B-deficient EBV promotes B cell lymphomagenesis in humanized mice and is found in human tumors. *Journal of Clinical Investigation* 122, 1487–1502.

Woisetschlaeger, M., Yandava, C.N., Furmanski, L.A., Strominger, J.L., and Speck, S.H. (1990). Promoter switching in Epstein-Barr virus during the initial stages of infection of B lymphocytes. *Proceedings of the National Academy of Sciences* 87, 1725–1729.

Woisetschlaeger, M., Jin, X.W., Yandava, C.N., Furmanski, L.A., Strominger, J.L., and Speck, S.H. (1991). Role for the Epstein-Barr virus nuclear antigen 2 in viral promoter switching during initial stages of infection. *Proceedings of the National Academy of Sciences* 88, 3942–3946.

Wood, C.D., Veenstra, H., Khasnis, S., Gunnell, A., Webb, H.M., Shannon-low, C., Andrews, S., Osborne, C.S., and West, M.J. (2016). MYC activation and BCL2L1 silencing by a tumour virus through the large-scale reconfiguration of enhancer-promoter hubs. *Elife* 5, e18270.

Wu, K.J., Grandori, C., Amacker, M., Simon-Vermot, N., Polack, a, Lingner, J., and Dalla-Favera, R. (1999). Direct activation of TERT transcription by c-MYC. *Nature Genetics* 21, 220–224.

Yenamandra, S.P., Sompallae, R., Klein, G., and Kashuba, E. (2009). Comparative analysis of the Epstein-Barr virus encoded nuclear proteins of EBNA-3 family. *Computers in Biology and Medicine* 39, 1036–1042.

Yi, F., Saha, A., Murakami, M., Kumar, P., Knight, J.S., Cai, Q., Choudhuri, T., and Robertson, E.S. (2009). Epstein-Barr virus nuclear antigen 3C targets p53 and modulates its transcriptional and apoptotic activities. *Virology* 388, 236–247.

Young, L.S., and Murray, P.G. (2003). Epstein-Barr virus and oncogenesis: from latent genes to tumours. *Oncogene* 22, 5108–5121.

Young, L.S., and Rickinson, A.B. (2004). Epstein-Barr virus: 40 years on. *Nature Reviews. Cancer* 4, 757–768.

Young, K.A., Chen, X.S., Holers, V.M., and Hannan, J.P. (2007a). Isolating the Epstein-Barr virus gp350/220 binding site on complement receptor type 2 (CR2/CD21). *Journal of Biological Chemistry* 282, 36614–36625.

Young, L.S., Arrand, J.R., and Murray, P.G. (2007b). EBV gene expression and regulation

(Cambridge University Press).

Young, L.S., Sung, N.S., and Pagano, J.S. (2009). *Epstein-Barr Virus* (Chichester, UK: John Wiley & Sons, Ltd).

Young, P., Anderton, E., Paschos, K., White, R., and Allday, M.J. (2008). Epstein–Barr virus nuclear antigen (EBNA) 3A induces the expression of and interacts with a subset of chaperones and co-chaperones. *Journal of General Virology* 89, 866–877.

Zhang, X., Ding, L., and Sandford, A.J. (2005). Selection of reference genes for gene expression studies in human neutrophils by real-time PCR. *BMC Molecular Biology* 6.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* 9, R137.

Zhao, B., and Sample, C.E. (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., Marshall, D.R., and Sample, C.E. (1996). A Conserved Domain of the Epstein-Barr Virus Nuclear Antigens 3A and 3C Binds to a Discrete Domain of Jk. *Journal of Virology* 70, 4228–4236.

Zhao, B., Zou, J., Wang, H., Johannsen, E., Peng, C.-W., Quackenbush, J., Mar, J.C., Morton, C.C., Freedman, M.L., Blacklow, S.C., Aster, J.C., Bernstein, B.E., and Kieff, E. (2011a). Epstein-Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. *Proceedings of the National Academy of Sciences of the United States of America* 108, 14902–14907.

Zhao, B., Mar, J.C., Maruo, S., Lee, S., Gewurz, B.E., Johannsen, E., Holton, K., Rubio, R., Takada, K., Quackenbush, J., and Kieff, E. (2011b). Epstein-Barr virus nuclear antigen 3C regulated genes in lymphoblastoid cell lines. *Proceedings of the National Academy of Sciences of the United States of America* 108, 337–342.

Zhou, H., Schmidt, S.C.S., Jiang, S., Willox, B., Bernhardt, K., Liang, J., Johannsen, E.C., Kharchenko, P., Gewurz, B.E., Kieff, E., and Zhao, B. (2015). Epstein-Barr virus oncoprotein super-enhancers control B cell growth. *Cell Host & Microbe* 17, 205–216.

Zimber-Strobl, U., and Strobl, L.J. (2001). EBNA2 and Notch signalling in Epstein–Barr virus mediated immortalization of B lymphocytes. *Seminars in Cancer Biology* 11, 423–434.

Zimmermann, J., and Hammerschmidt, W. (1995). Structure and role of the terminal repeats of Epstein-Barr virus in processing and packaging of virion DNA. *Journal of Virology* 69, 3147–3155.

Appendix A: Expression of EBNA3C is compromised when EBNA3B is tagged at the C-terminus.

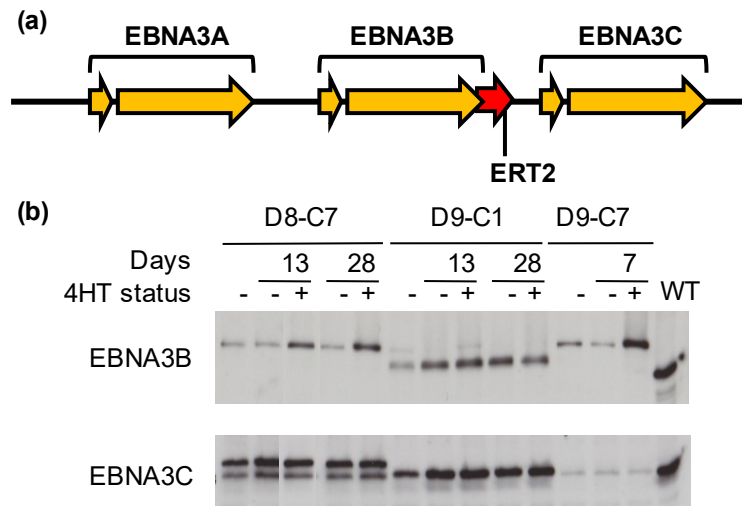


Figure 8. Expression of EBNA3C is compromised in LCLs generated with EBNA3B-ERT2 viruses regardless of their functionality in response to 4HT. (a) Schematic showing the location of the ERT2 tag at the C-terminus of EBNA3B. (b) Western blot of cell extracts taken from EBNA3B-ERT2 LCLs which were generated using two different virus strains (C7 and C1) and PBLs from two different donor backgrounds (D8 and D9). Figure provided by Dr Rob White.

Appendix B: ChIP experiments using monoclonal rat antibodies

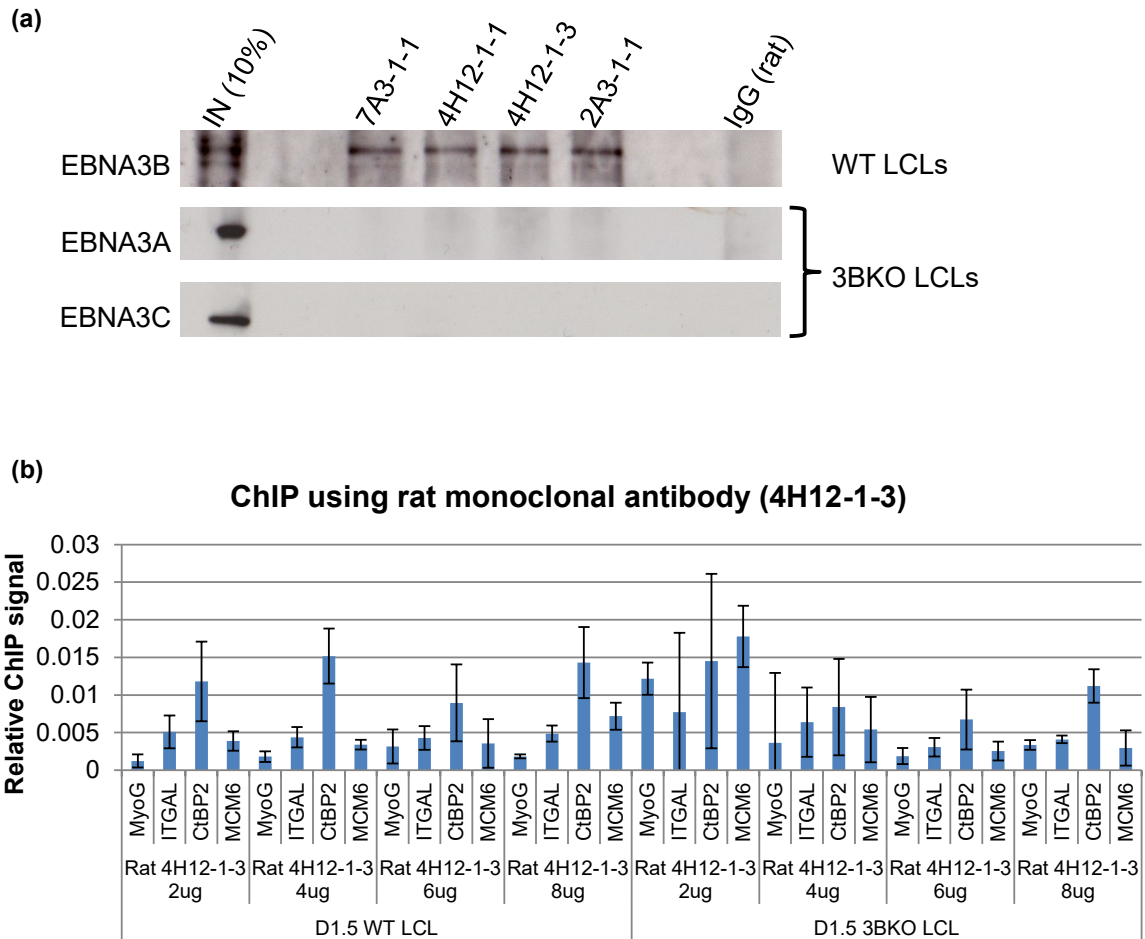


Figure 9. ChIP using rat monoclonal antibodies (a) Western blot of protein extracts taken from ChIP experiments using rat monoclonal antibodies (7A3-1-1, 4H12-1-2, 4H12-1-3 and 2A3-1-1) in WT and 3BKO LCLs derived from Donor 1.3. Rat IgG is used as a negative control for the immunoprecipitation. Exalpha EBNA3B antibody (F120P) is used to detect EBNA3B in the western blot. (b) ChIP experiment showing the ChIP signal relative to 1% input at each putative EBNA3B binding site (*ITGAL* and *CtBP2* promoter/enhancer regions) (McClellan *et al*, 2013). *MyoG* and *MCM6* promoter sequences are used as negative control regions for EBNA3B binding.

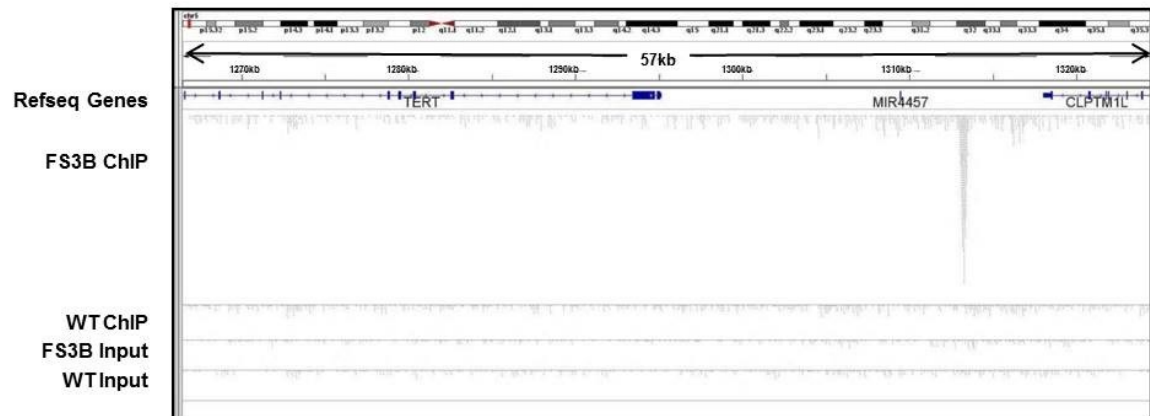
Appendix C: Table of genes that are differentially regulated by EBNA3B in both LCLs and BL31

	ENSEMBL gene ID	Gene Name	LCL		BL31	
			3BKO vs WT	Fold-Change	3BKO vs WT	Fold-Change
Up in LCL-3BKO and BL31-3BKO	ENSG00000197381	ADARB1	0.0005489	6.49	0.0000000	6.79
	ENSG00000237943	AL158043.1	0.0000894	4.81	0.0002166	2.65
	ENSG00000224478	AL356417.1	0.0003367	2.55	0.0000221	3.34
	ENSG00000107863	ARHGAP21	0.0003488	2.79	0.0000057	4.62
	ENSG00000151929	BAG3	0.0000045	8.60	0.0002022	2.66
	ENSG00000060982	BCAT1	0.0000182	2.80	0.0003050	2.30
	ENSG00000146592	CREB5	0.0000003	11.14	0.0000000	5.63
	ENSG00000121966	CXCR4	0.0007225	5.51	0.0000000	3.32
	ENSG00000117298	ECE1	0.0001549	4.08	0.0000001	8.27
	ENSG00000065809	FAM107B	0.0000398	2.46	0.0000022	4.45
	ENSG00000158089	GALNT14	0.0000012	6.66	0.0000008	4.02
	ENSG00000150625	GPM6A	0.0000016	17.38	0.0005127	3.42
	ENSG00000152804	HHEX	0.0000072	19.30	0.0000001	7.62
	ENSG00000130787	HIP1R	0.0003070	2.31	0.0000065	2.07
	ENSG00000198825	INPP5F	0.0000005	23.56	0.0000000	21.38
	ENSG00000107719	KIAA1274	0.0000052	2.84	0.0000794	2.65
	ENSG00000172578	KLHL6	0.0000037	9.06	0.0000000	7.66
	ENSG00000167613	LAIR1	0.0000008	11.72	0.0000000	9.32
	ENSG00000239961	LILRA4	0.0003930	5.15	0.0000000	7.94
	ENSG00000243156	MICAL3	0.0000297	3.04	0.0000000	6.82
	ENSG00000168734	PKIG	0.0002587	3.32	0.0000284	2.25
	ENSG00000105854	PON2	0.0007890	4.56	0.0000120	13.11
	ENSG00000168297	PXK	0.0000004	4.55	0.0000000	3.53
	ENSG00000185989	RASA3	0.0000056	2.55	0.0000000	4.79
	ENSG00000126821	SGPP1	0.0000129	2.27	0.0000074	2.97
	ENSG00000089057	SLC23A2	0.0000042	8.46	0.0000000	5.16
	ENSG00000172716	SLFN11	0.0005754	2.93	0.0000784	4.53
	ENSG00000008513	ST3GAL1	0.0001883	3.27	0.0006242	2.35
	ENSG00000164691	TAGAP	0.0008287	5.13	0.0000003	12.04
	ENSG00000164362	TERT	0.0000067	2.83	0.0000000	5.89
	ENSG00000047644	WWC3	0.0000148	3.92	0.0000022	4.06
	ENSG00000093100	XXbac-B461K10.1	0.0000028	7.15	0.0000000	16.17
Down in LCL-3BKO and BL31-3BKO	ENSG00000231312	AC007246.1	0.0000285	-2.15	0.0000197	-2.10
	ENSG00000234663	AC013473.1	0.0005655	-2.31	0.0000000	-2.91
	ENSG00000151320	AKAP6	0.0002469	-6.43	0.0005032	-3.94
	ENSG00000133805	AMPD3	0.0004407	-2.18	0.0000002	-5.71
	ENSG00000100852	ARHGAP5	0.0000147	-3.53	0.0000017	-2.64
	ENSG00000186594	C17orf91	0.0007530	-2.13	0.0006659	-2.89
	ENSG00000178562	CD28	0.0000036	-11.85	0.0000000	-17.20
	ENSG00000169245	CXCL10	0.0009958	-7.89	0.0000120	-9.04
	ENSG00000160219	GAB3	0.0009302	-2.37	0.0005647	-4.24
	ENSG00000011566	MAP4K3	0.0001949	-4.28	0.0000477	-2.95
	ENSG00000185432	METTL7A	0.0000215	-3.50	0.0000000	-6.56
	ENSG00000111859	NEDD9	0.0000245	-2.54	0.0000007	-2.47
	ENSG00000109756	RAPGEF2	0.0000189	-2.73	0.0000527	-2.41
	ENSG00000134243	SORT1	0.0000121	-25.70	0.0000000	-4.04
Down in LCL-3BKO and up in BL31-3BKO	ENSG00000241990	AL121672.2	0.0002773	-4.47	0.0004624	2.86
	ENSG00000150347	ARID5B	0.0000637	-2.71	0.0008092	2.46
	ENSG00000179630	C13orf31	0.0007877	-2.06	0.0000955	5.55
	ENSG00000151773	CCDC122	0.0006050	-2.04	0.0000066	5.61
	ENSG00000137101	CD72	0.0004458	-2.91	0.0000086	2.38
	ENSG00000159216	RUNX1	0.0001786	-2.72	0.0005120	2.22
	ENSG00000100721	TCL1A	0.0000402	-30.77	0.0000481	4.21
	ENSG00000121858	TNFSF10	0.0002546	-2.28	0.0001763	2.65

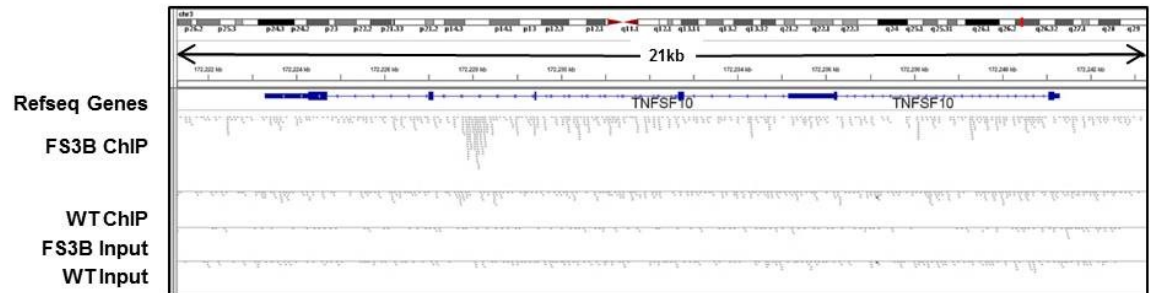
Figure 10. Table retrieved from White *et al.* (2010) showing genes differentially regulated by EBNA3B in LCLs and BL31

Appendix D: Pile-up of ChIP-seq reads at the *TERT*, *TRAIL* & *TRAILR1/R4 loci*

(a) *TERT*



(b) *TRAIL*



(c) *TRAILR1/R4*

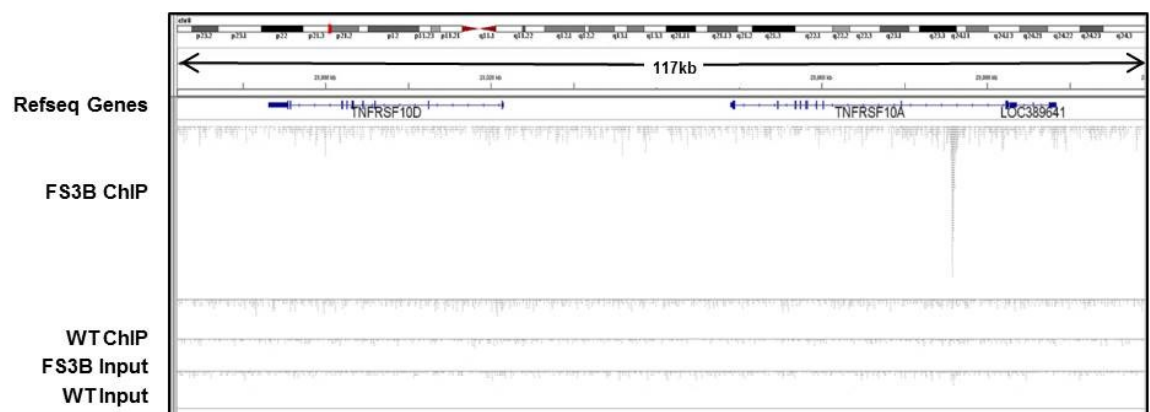


Figure 11. Pile-up of EBNA3B ChIP-seq reads at the *TERT*, *TRAIL* & *TRAILR1/R4* loci.

Appendix E: α -FLAG ChIP using Donor 13 LCLs expressing epitope-tagged EBNA3B or EBNA3C

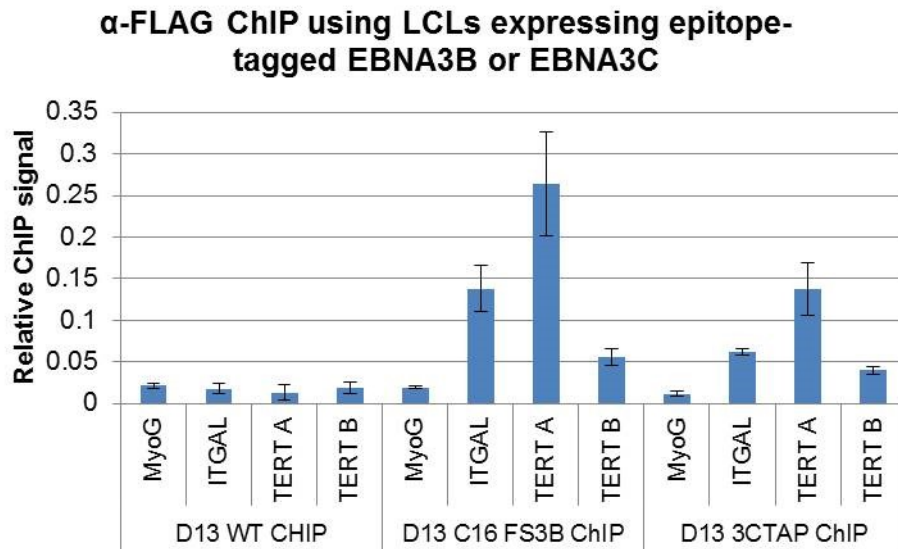


Figure 12. EBNA3B and EBNA3C binding at distal regulatory regions (TERTA & TERTB) upstream of TERT TSS. ChIP-qPCR in Donor 13 LCLs expressing epitope-tagged EBNA3B and EBNA3C was done to confirm the findings from the ChIP-seq (see section 5.3.2), using *MyoG* and *ITGAL* promoters as negative and positive controls respectively. Signal relative to 0.5% input is reported in a single qPCR experiment. Error bars represent mean $\pm$ SD of three technical triplicates.

Appendix F: Comparisons of EBNA3 binding profile

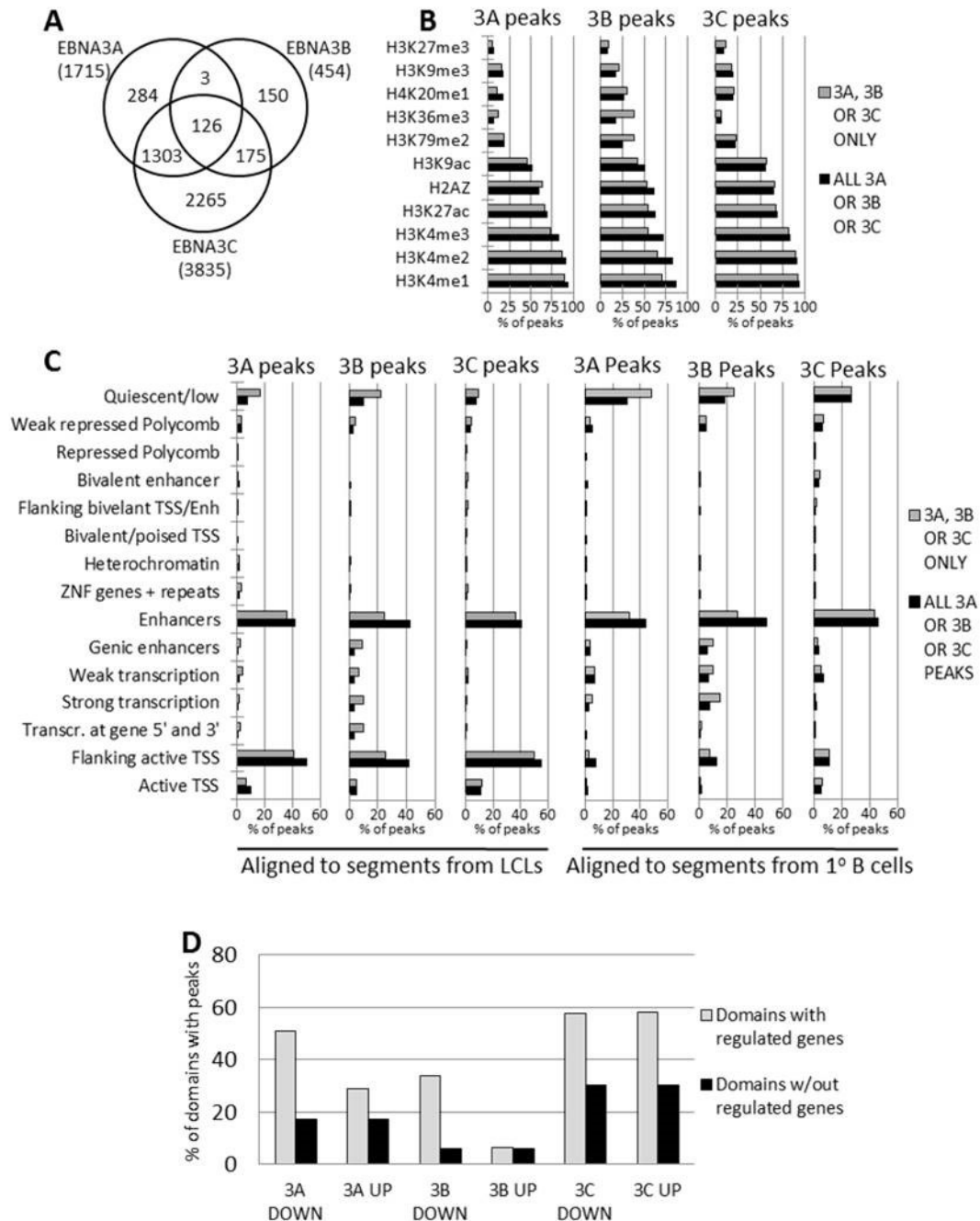


Figure 13. Co-localisation of EBNA3A, EBNA3B and EBNA3C peaks with (A) other EBNA3s (B) histone modifications, (C) genome segments and (D) contact domains containing EBNA3-regulated genes of LCL GM12878.

(A) Overlap of EBNA3 peaks called by MACS algorithm (B) Overlap of histone modifications with all peaks for each EBNA3 and with binding sites where only one EBNA3 localises (EBNA3A-, EBNA3B- or EBNA3C-only) are shown. (C) Co-localisation of EBNA3 peaks found in LCLs with different chromatin segmented states found in LCL GM12878 and CD19+ve 1° B cells (D) Percentage of total contact domains which overlap with EBNA3-regulated genes and also EBNA3 peaks (grey bars) and percentage of contact domains which overlap with EBNA3 peaks but do not contain EBNA3-regulated genes (black bar). Modified from Paschos *et al.* (submitted to Nucleic Acids Research)

Appendix G: EBNA3 co-localisation with TFBS

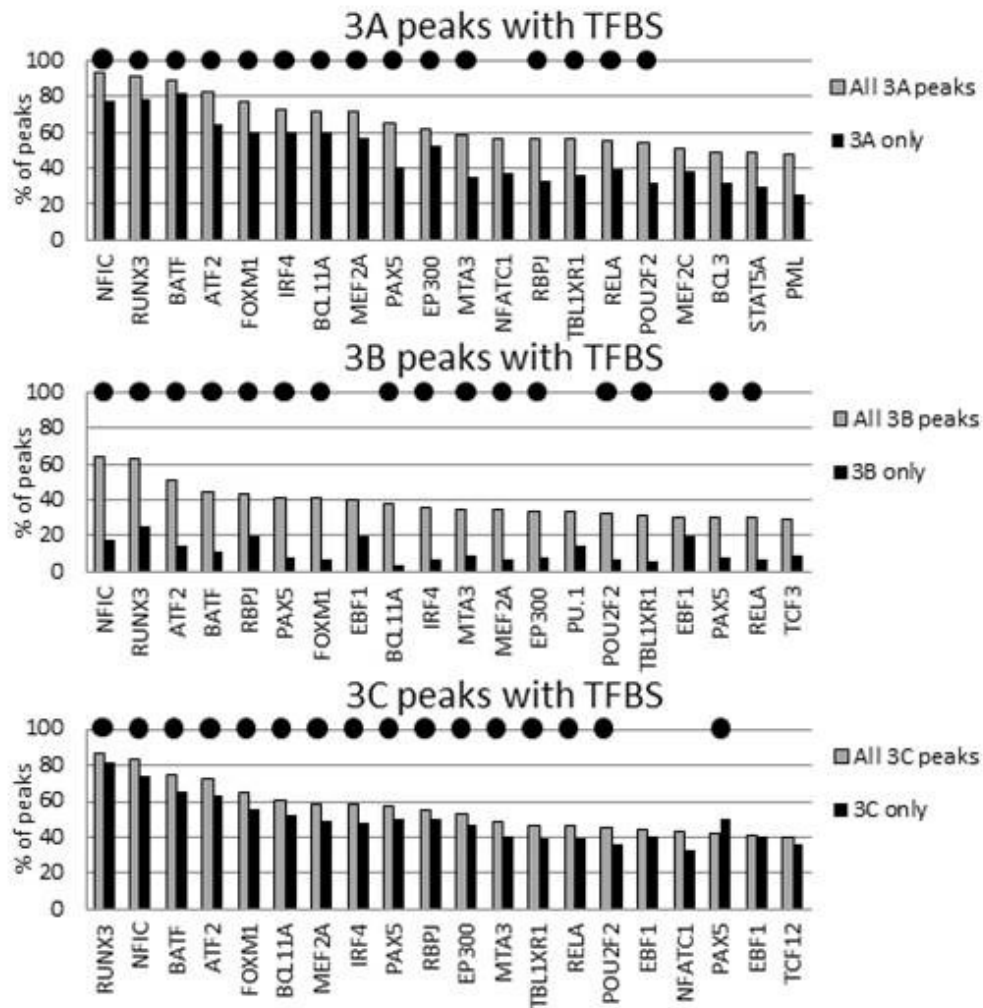


Figure 14. Overlap of EBNA3s with TFBS available from UCSC genome browser.

Overlap of all peaks for each EBNA3 (grey bars), and binding sites with only one EBNA3 (EBNA3A, -3B and -3C only peaks; black bars) with transcription factor binding sites (TFBS) determined by the ENCODE project. The top 20 TFBS co-localising with each EBNA3 are shown. Black dots above bars indicate transcription factors that are in the top 20 for all three EBNA3s. Figure adapted from Paschos *et al.* (2016, submitted to Nucleic Acids Research)