Cellular reprogramming by Epstein-Barr virus nuclear antigens

By

Michael J. McClellan

Submitted in total fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

School of Life Sciences

University of Sussex

September 2014

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Acknowledgements

I would like to thank my supervisor Professor Michelle West for her direction, support and patience. I would also like to thank all past and present members of the West-Sinclair labs who have made my time here so enjoyable.

Special thanks must go to Richard Palermo for passing on the baton of ChIP to me, Sharada Ramasubramanyan for holding my hand throughout the traumatic ordeal of ChIP sequencing and David Wood for his friendship and limitless technical expertise.

I would further like to thank the PhD student community here at Sussex University many of whom I have become great friends with through the course of this PhD.

Publications arising from this work

Tzellos S, Correia PB, Karstegl CE, Cancian L, Cano-Flanagan J, **McClellan MJ**, West MJ, Farrell PJ (2014) *A single amino acid in EBNA-2 determines superior B lymphoblastoid cell line growth maintenance by Epstein-Barr virus type 1 EBNA-2*. J Virol. Aug;88(16):8743-53.

McClellan MJ, Wood CD, Ojeniyi O, Cooper TJ, Kanhere A, Arvey A, Webb HM, Palermo RD, Harth-Hertle ML, Kempkes B, Jenner RG, West MJ (2013) Modulation of enhancer looping and differential gene targeting by Epstein-Barr virus transcription factors directs cellular reprogramming. PLoS Pathog. Sep;9(9):e1003636.

McClellan MJ, Khasnis S, Wood CD, Palermo RD, Schlick SN, Kanhere AS, Jenner RG, West MJ (2012) *Downregulation of integrin receptor-signaling genes by Epstein-Barr virus EBNA 3C via promoter-proximal and -distal binding elements.* J Virol. May;86(9):5165-78.

Cellular reprogramming by Epstein-Barr virus nuclear antigens

Thesis Abstract

Epstein-Barr virus (EBV) is a widespread human B cell virus that is linked to many malignancies. EBV modulates the transcriptome of B lymphocytes to drive immortalisation and viral persistence. This is primarily coordinated by the EBV nuclear antigens (EBNA) 2 and the EBNA 3 family (3A, 3B and 3C), which regulate overlapping sets of cellular genes. Using Chromatin immunoprecipitation (ChIP) coupled to next generation sequencing we found >21000 EBNA 2 and >7000 EBNA 3 binding sites in the human genome, providing the first evidence of EBNA 3 association with the human genome *in vivo*. Binding sites were predominantly distal to transcription start sites (TSS) indicating a key role in long-range gene control. This was especially pronounced for EBNA 3 proteins (84% of sites over 4kb from any TSS). 56% of genes previously reported to be regulated by these EBNA proteins in micro array experiments were bound by an EBNA. Using ChIP-QPCR we confirmed EBNA 3C bound to and promoted epigenetic silencing of a subset of integrin receptor signalling genes (*ITGA4*, *ITGB1*, *ADAM28*, *ADAMDEC1*). Indirect silencing of *CXCL10* and *CXCL11* chemokines by EBNA 3C was also demonstrated. 75% of sites bound by EBNA 3 were also bound by EBNA 2 implicating extensive interplay between EBNA proteins in gene regulation. By examining novel (*WEE1*, *CTBP2*) and known (*BCL2L11*, *ITGAL*) targets of EBNA 3 proteins bound at promoter-proximal or distal binding sites, we found both cell-type and locus-specific binding and transcriptional regulation. Importantly, genes differentially regulated by a subset EBNA 3 proteins were bound by the same subset, providing a mechanism for selective regulation of host genes by EBNA 3 proteins. In summary, this research demonstrates that EBNA proteins primarily act through long-range enhancer elements and regulate gene expression in a locus and gene-specific manner through differential binding.

Table of Contents

Common Abbreviations

ADAM	A disintegrin and metalloprotease
AID	activation-induced cytidine deaminase
BAC	bacterial artificial chromosome
BCR	B-cell receptor
BL	Burkitts lymphoma
BRD4	dual bromodomain protein
CBP	CREB-binding protein
ChIP	Chromatin Immunoprecipitation
CLP	common lymphoid progenitor
Cp	C promoter
CTBP1	C-terminal binding protein 1
CTBP2	C-terminal binding protein 2
CTD	Carboxy-terminal domain
EBNA	EBV nuclear antigen
EBV	Epstein-Barr virus
ENCODE	Encyclopaedia of DNA elements
EZH2	enhancer of zeste 2
GC	Germinal centre
GSB	Gel sample buffer
H3	Histone 3
H3ac	acetylated histone H3
H3K27ac	acetylation of lysine 27 on histone H3
H3K27me3	tri-methylation of lysine 27 on histone H3
H3K4me1	mono-methylation of lysine 4 on histone H3
H4	Histone 4
H4ac	acetylated histone H4
HAT	histone acetyl transferase
HDAC	Histone deacetylases
HL	Hodgkins lymphoma
HMT	histone methlytransferases
HRS	Hodgkin/Reed-Sternberg
HSC	Hematopoietic stem cells
IM	infectious mononucleosis
IP	immuno-precipitation
LCL	lymphoblastoid cell line
LMP	latent membrane protein
NK	natural killer
NPC	nasopharyngeal carcinoma
OriP	origin of replication
PcG	polycomb group proteins
PIC	pre-initiation complex
PRC1	polycomb repressive complex 1
PRC2	polycomb repressive complex 2

pTEFb	positive transcription elongation factor b
PTLD	Post-transplant lymphoprolierative disease
Qp	Q promoter
RNAPII	RNA polymerase II
TAD	transactivation domain
TF	transcription factor
TFBS	transcription factor binding sites
TSS	Transcription start site
WB	western blot
Wp	W promoter

1.Introduction

Epstein-Barr virus (EBV) is an almost ubiquitous gamma herpes virus that is capable of asymptomatically infecting the vast majority of the worldwide adult population *in vivo*. It is also associated with numerous human malignancies and was discovered through its association with Burkitts lymphoma (BL), the most common childhood lymphoma in sub Saharan Africa. EBV can immortalise resting B-cells forming permanently proliferating lymphoblastoid cell lines (LCLs). During LCL outgrowth EBV virus expresses only 9 proteins, of which 6 are nuclear. The Epstein-Barr virus nuclear antigen (EBNA) 2, EBNA 3A, EBNA 3B and EBNA 3C play a key role in reprogramming host cells to enable the virus to persist and survive. At the start of this PhD project, we had little understanding of how these proteins target and deregulate host cell genes.

1.1 Regulation of transcription

1.1.1 RNA polymerase II recruitment and regulation

Eukaryotes have three RNA polymerases, RNAPI, RNAPII and RNAPIII. RNAPII is primarily responsible for creating mRNA which can be translated into protein. RNAPII usually initiates transcription with the additional binding of general transcription factors at a TATA box, typically located 25bp upstream of the transcription start site; which when bound causes a distortion in DNA structure which may act as a physical landmark for the location of an active promoter (Kim et al. 1993). The subsequent recruitment of further general transcription factors and RNAPII then forms the complete transcription pre-initiation complex (PIC). This includes a DNA helicase TFIIH, which unwinds DNA and allows RNAPII access to the template strand. RNAPII initiation and elongation involves phosphorylation of the RNAPII largest subunit (Rpb1) Carboxy-terminal domain (CTD) by members of the transcription initiation complex. The CTD has an essential role in regulating efficient transcription, facilitating events in initiation, elongation and RNA processing through modifications by regulatory proteins. In humans it consists of 52 heptapeptide repeats ($Y_1S_2P_3T_4S_5P_6S_7$). The CTD 'code' describes RNAPII regulation by transient modifications of the CTD, including serine 5 phosphorylation at initiation and serine 2 phosphorylation during elongation (Egloff and Murphy 2008). As well as phosphorylation the CTD can also be methylated, glycosylated, ubiquitinated and undergo changes in proline isomerisation, all of which have been ascribed particular biological functions in relation to RNAPII transcription (reviewed in (Eick and Geyer 2013)). For example, serine 2 phosphorylation by positive transcription elongation factor b (pTEFb) is required for productive

elongation (Marshall and Price 1995). pTEFb also phosphorylates DRB sensitivity-inducing factor (DSIF) and negative elongation factor (NELF) relieving promoter proximal pausing of RNAPII (Wada et al. 1998; Fujinaga et al. 2004). Interestingly, it appears that promoter proximal pausing is required for efficient transcription of many genes by occluding nucleosome assembly around the promoter (Gilchrist et al. 2008). RNAPII and general transcription factors must also access DNA tightly associated with histones to bind promoter sequences, and post translational modifications of histone tails allow direction of RNAPII to target genes. A great deal of cross talk exists between RNAPII interacting proteins and histone modifying enzymes. For example the dual bromodomain protein, BRD4, associates with acetylated histone tails of histones H3 and H4 which are often found at sites of active transcription. BRD4 can then recruit p-TEFb and induce RNAPII CTD phosphorylation facilitating high level gene specific transcription elongation (Jang et al. 2005; Yang et al. 2005).

1.1.2 Chromatin

DNA in the cell is contorted and compacted into chromatin, the basic unit of which is the nucleosome. A single nucleosome contains an octamer of histone proteins H3, H4, H2A, H2B and has ~147bp of DNA completing nearly two full turns around the nucleosome. Each nucleosome is separated by 10-60bp linker piece of DNA which is commonly bound by histone H1 (reviewed in (Peterson and Laniel 2004). The first evidence for the manner of DNA packaging came from electron microscope images (Oudet et al. 1975) which showed the genome existed in uniformly sized particles. Each of these repeating bead like structures was called a nucleosome and formed the classic ‘beads on a string’ packaging of DNA. Chromatin is further condensed by internucleosomal attraction forming a zig-zag pattern of DNA and nucleosomes which further folds in to a ~30nm thick fibre (reviewed in (Woodcock and Dimitrov 2001). This fibre is further condensed within the nucleus, partly by interaction with the condensing complex (Swedlow and Hirano 2003). Whilst the core histones packed into chromatin form highly structured fibres, the individual histones have long unstructured histone tails, that protrude from the core fibre. These comprise about 30% of the mass of each individual histone and can be post translationally modified at specific residues. Observed modifications include acetylation, methylation, ubiquitination and ADP ribosylation (Luger and Richmond 1998; Wolffe and Hayes 1999).

Oudet *et al* also observed that DNA from actively transcribing cells had the same sized nucleosomes but the spaces between them were less uniform and generally larger when compared to those of cells undergoing little transcription; this therefore provided the first

evidence that nucleosomes are part of the transcriptional regulatory network (Oudet et al. 1975). Experiments on the structure of ectopic nucleosomes with DNA bound from preparations made in *Escherichia coli* (*E.coli*) and Chicken erythrocytes gave clues about how this may happen. Both sets of nucleosomes had very similar core structure, but histones prepared in *E.coli* had no post-translational modifications on the histone tails (Luger et al. 1997; Harp et al. 2000). This suggests that post translational modification of the histone tail does not exert biological function by significantly reshaping the nucleosome but by recruitment of specific proteins. Purely considering electrostatic interactions at least acetylation and phosphorylation are predicted to aid transcription by chromatin decondensation via neutralisation of a positive charge and addition of a negative charge respectively (Roth and Allis 1992). This process appears to be dynamic as phosphorylation of H1 and H3 has been implicated in chromatin condensation during mitosis (Koshland and Strunnikov 1996) but specific phosphorylation of serine 10 on histone 3 is directly related to the active transcription of *MYC*, *JUN* and *FOS* (Mahadevan et al. 1991; Thomson et al. 1999).

1.1.2.1 Epigenetics and the regulation of chromatin structure

Chromatin can be altered in several interconnected ways: through covalent modification of histones, exchange of core histones with histone variants, disruption of the nucleosome and DNA modification. An essential component in gene silencing by DNA modification and chromosome compaction is the polycomb group proteins (PcG), which are conserved in all higher eukaryotes. The tri-methylation of lysine 27 on histone H3 (H3K27me3) is thought to be the operative chromatin mark that accompanies PcG mediated gene silencing (reviewed in (Simon and Kingston 2009)). PcG proteins were originally discovered in Drosophila as regulators of development by repression of *hox* genes. PcG proteins primarily exert their function as part of a complex, either Polycomb repressive complex 1 or 2 (PRC1 or PRC2), although in mammals multiple forms of PRC1 and PRC2 exist with different subunits. It has therefore proved difficult to determine the biological function of individual subunits and specific complexes. A known core function of PRC2 is the deposition of H3K27me3 (Cao et al. 2002), whereas PRC1 is known to ubiquitylate histone H2A and cause nucleosome compaction (Francis et al. 2004; Cao et al. 2005). In cancer PcG gain of function mutants appear to be key in altering cancer cell expression profiles perhaps by hyper-silencing of tumour suppressor genes, especially in aggressive tumours (Yu et al. 2007; Ben-Porath et al. 2008). Interestingly whilst this remains true in B-cell lymphomas, PcG mutations found in T cell and myeloid leukemia are generally loss of function (Ernst et al. 2010; Ntziachristos et al. 2012).

In Drosophila, PcG proteins can be targeted to DNA by polycomb response elements (PREs) which bind sequence specific DNA binding proteins in PRC1 and PRC2 (Ringrose and Paro 2007; Schuettengruber and Cavalli 2009). PREs are depleted of nucleosomes and can be many kilobases from the genes under their control. One important transcription factor involved in this process in Drosophila is PHO, the principal polycomb recruitment protein, the human homologue of which is YY1, a human polycomb associated transcription factor (Oktaba et al. 2008). In the case of PHO and YY1 this process appears to be mediated by the coactivator/corepressor C-terminal binding protein (CTBP)(Basu and Atchison 2010). Whilst PREs appear not to be as important in mammals, functional YY1 dependant PREs have been shown to repress human *hox* genes (Woo et al. 2010). Human PcG recruitment is primarily dependent on the presence of CpG islands in a relatively sequence independent manner. This was demonstrated by the findings that PRC2 components were recruited to CpG rich bacterial artificial chromosome (BAC) derived DNA and CpG rich bacterial DNA when integrated into a mouse genome (Mendenhall et al. 2010). Methylation of CpG islands usually results in gene silencing by directly preventing transcription factor binding and recruiting methyl-binding domain proteins and Histone deacetylases (HDACs) (Taby and Issa 2010). CpG islands incapable of binding PcG proteins generally contain binding sites for transcriptional activators or are methylated. Methylation dependant binding of PRC2 to DNA is supported by data showing that full PRC2 recruitment requires tet methylcytosine dioxygenase 1 (TET1) which can convert 5-methyl cytosine to 5-hydroxy methyl cytosine, which can be further processed to unmodified cytosine (Wu et al. 2011; Tan and Shi 2012). This is a dynamic interaction as PcG proteins appear to also recruit TET1 to H3K27me3 sites to maintain hypomethylation of DNA, perhaps to maintain bivalent chromatin at a subset of sites (Neri et al. 2013). Interestingly a PRC1 variant complex containing lysine (K)-specific demethylase 2B (KDM2B) can also be recruited to unmethylated CpG islands (Wu et al. 2013) and PRC1 can be recruited to DNA via cellular transcription factors such as RUNX1 (Yu et al. 2012).

1.1.2.2 Histone tail modification

Epigenetics describes stable changes in gene expression that do not change the genetic code itself, allowing multiple cell types to diverge despite a common genetic code. Whilst this term first applied to the truly heritable modification of DNA methylation at CpG islands it is now used in the context of histone modifications, and as discussed, these processes are related. Early findings in the role of histone tail modifications lead to the proposal of a histone code: ‘multiple histone modifications, acting in a combinatorial or sequential fashion on one or

multiple histone tails, specify unique downstream functions' (Strahl and Allis 2000). Since then the true complexity of the histone code has been subject to intense study. In recent years many ChIP sequencing experiments mapping the genomic locations of different marks have been performed in a variety of cell lines. This data is publicly accessible via the Encyclopaedia of DNA elements (ENCODE) project allowing users to align their own sequencing data against this resource (http://genome.ucsc.edu/ENCODE/). Some histone modifications studied are described in Table 1-1 along with their proposed function by the ENCODE project. Whilst discussion of all histone modifications is beyond the scope of this work I shall discuss a few pertinent modifications in more detail.

1.1.2.2.1 Acetylation

As discussed previously, acetylation of histones is involved in reducing DNA histone contacts but also play a role in transcriptional activation, DNA replication, histone deposition and DNA repair via recruitment of proteins with acetyl binding domains. In the case of acetyl-lysines this is the bromodomain (Kurdistani and Grunstein 2003). Acetyl groups are deposited on histones by histone acetyl transferases (HATs) and removed by histone de-acetylases (HDACs), a process that requires careful regulation and balance (Marmorstein 2001). The linker histone H1 has been shown bind with a lower affinity to core histones with hyperacetylated tails contributing to chromatin decondensing (Misteli et al. 2000). Acetylation of histones in general may then contribute to the unfolding of nucleosomes at transcriptionally active promoters (Boeger et al. 2003). However the chronology of these events is less well understood. A great deal of cross talk exists between histone modifications; for example, phosphorylation of serine 10 in the histone H3 tail causes a conformational change creating additional interaction sites with HATs enhancing H3K14 acetylation and transcription (Clements et al. 2003). Apart from facilitating open chromatin permissive to transcription (euchromatin), acetylation also plays a role in the recruitment of basic transcription machinery. The largest subunit of TFIID binds to the di-acetylated H4 tail with 40 fold higher affinity than the mono-acetylated tail (Jacobson et al. 2000).

The deposition of post-translational modifications is site-specific in many cases. For example, on histone H3 most acetylation of lysines occurs on K9, K14, K18 and K23, which are differentially acetylated during different biological processes (Kuo et al. 1996). Studies on the yeast GCN5 HAT revealed it may be directed to H3K14 by surrounding residues, in this case a glycine and a proline, suggesting that substrate specificity is in part due to short preferred consensus sites (Rojas et al. 1999; Cieniewicz et al. 2014). Post-translational modifications of

Histone modification or variant	Signal characteristics	Putative functions
H2A.Z	Peak	Histone protein variant (H2A.Z) associated with regulatory elements with dynamic chromatin
H3K4me1	Peak/region	Mark of regulatory elements associated with enhancers and other distal elements, but also enriched downstream of transcription starts
H3K4me2	Peak	Mark of regulatory elements associated with promoters and enhancers
H3K4me3	Peak	Mark of regulatory elements primarily associated with promoters/transcription starts
H3K9ac	Peak	Mark of active regulatory elements with preference for promoters
H3K9me1	Region	Preference for the 5′ end of genes
H3K9me3	Peak/region	Repressive mark associated with constitutive heterochromatin and repetitive elements
H3K27ac	Peak	Mark of active regulatory elements; may distinguish active enhancers and promoters from their inactive counterparts
H3K27me3	Region	Repressive mark established by polycomb complex activity associated with repressive domains and silent developmental genes
H3K36me3	Region	Elongation mark associated with transcribed portions of genes, with preference for 3′ regions after intron 1
H3K79me2	Region	Transcription-associated mark, with preference for 5′ end of genes
H4K20me1	Region	Preference for 5′ end of genes

Table 1-1 Histone modifications and putative function. (adapted from 'an integrated encyclopaedia of DNA elements in the human genome' (2012)) Many labs have conducted ChIP sequencing experiments for Histone modifications. Here are listed the known histone modifications for which there is publicly available ChIP sequencing data, the characteristics of the mark and the proposed putative function.

histone tails also inhibit protein-protein interactions which modulate transcription. TUP1, a protein involved in transcriptional silencing in yeast was shown to bind with greater affinity to unacetylated or monoacetylated histone H3 and H4 than hyperacetylated forms (Edmondson et al. 1996).

The acetylation of histone H3 on lysine 27 (H3K27ac) is of particular interest as it is widely regarded as a marker for active enhancer sites (Creyghton et al. 2010). One mechanism of the removal of H3K27ac is via ZEB1 interaction with CTBP proteins which subsequently recruit HDAC activity to ZEB1 binding sites, repressing genes such as *CDH1* (Postigo and Dean 1999; Aghdassi et al. 2012). In some lung cancers *ZEB1* is repressed leading to a local and global decrease in H3K27ac (Roche et al. 2013). The bulk of H3K27ac is deposited by CREB-binding protein (CBP) and in drosophila has been shown to block the PRC2 mark H3K27me3 (Tie et al. 2009; Pasini et al. 2010). CBP also physically associates with UTX (KDM6A) and BRM (a H3K27 specific demethylase and a chromatin-remodelling ATPase Brahma respectively) and directs them to CBP sites including many polycomb response elements, removing H3K27me3 and simultaneously adding H3K27ac. RNAi knock down of UTX and BRM reduces H3K27ac at target sites (Tie et al. 2012). H3K27ac exhibits a great deal of crosstalk with other histone modifications.

One intriguing situation concerns a switch from poised chromatin to active chromatin. H3K4 methylation is indicative of accessible chromatin and is deposited by lysine specific methyltransferases such as KMT2 (Allis et al. 2007; Zentner and Henikoff 2013). H3K27me3 is associated with repressed chromatin and is deposited by the PRC2 protein EZH2 (enhancer of zeste); together they signify poised chromatin, the defining feature of which is rapid transitions between transcriptional activation and repression, and is important in development (reviewed in (Simon and Kingston 2013; Voigt et al. 2013)). ASH2L-KMT2 complexes have been shown to bind to methylated H3K4 with a preference for di and tri methylated forms, this has the net effect of removing H3K27me3 and spreading H3K4 methylation (Wang et al. 2010). Recently the transcriptional regulator MYC was shown to interact with this complex and also promote H3K27ac at these sites, thus facilitating a switch from bivalent to fully active chromatin state (Ullius et al. 2014).

1.1.2.2.2 Methylation

H3K27me3 is deposited by the core PRC2 protein EZH2, which can perform each of the three sequential methylation events but only in complex with two of the other three core PRC2

subunits, SUZ12 and EED (Cao and Zhang 2004; Ketel et al. 2005). This activity can be modulated by mutation (McCabe et al. 2012), phosphorylation of EZH2 (Chen et al. 2012) or other PRC2 components, differential binding of EED isoforms and alternate PRC2 compositions (reviewed in (O'Meara and Simon 2012)). PRC2 can also methylate itself and other proteins including histone H1 (Kuzmichev et al. 2004; He et al. 2012); furthermore there is evidence that EZH2 outside of the PRC2 complex can activate transcription in a methylation dependent manner (Xu et al. 2012). PRC2 interacts with many other proteins which modulate its activity, for example, Polycomb like proteins (PLCs) aid PRC2 to complete the final methylation of H3K27 and facilitate recruitment of PRC2 to CpG islands (Sarma et al. 2008; Hunkapiller et al. 2012). JARID2, a member of the jumonji family of histone demethylases associates with PRC2 although the loss of this protein has been reported to only moderately change H3K27me3 in both a positive and negative fashion; thus, the biological importance of JARID2 binding remains controversial. It may fine tune PRC2 signalling and also PRC2 targeting (reviewed in (Herz and Shilatifard 2010)).

PRC2 positively regulates its own activity as the presence of H3K27me3 has been shown to stimulate PRC2 activity. This function may serve to spread H3K27me3 across regulatory domains, and indeed, H3K27me3 presents as a broad mark rather than a distinct peak (Margueron et al. 2009; Xu et al. 2010). Furthermore, PRC2 preferentially methylates densely packed nucleosomes via interactions with histone tails from neighbouring histones; potentially serving as a mechanism for linking multiple forms of epigenetic repression (Yuan et al. 2012). The presence of various positive regulatory marks and elements can inhibit PRC2 activity. H3K4me3 and H3K36me2 or me3 can prevent PRC2 activity although this appears to be via allosteric effects rather than a substrate affinity affect (Yuan et al. 2011). PRC2 dependent H3K27me3 is mutually exclusive with H3K27ac, with H3K27ac enzymatically occluding PRC2 activity. H3S28 phosphorylation also prevents PRC2 activity as this residue is adjacent to H3K27 and allosterically inhibits methylation (reviewed in (O'Meara and Simon 2012)).

Due to the complex cross-talk between histone modifications, it is extremely difficult to determine the output of H3K27me3. In fact, there is still debate as to whether this mark has a direct repressive effect or if it is just a by-product of gene repression; an argument supported by data showing that gene repression precedes H3K27me3 deposition (Henikoff and Shilatifard 2011; Yuan et al. 2012). However, loss of PRC2 in drosophila results in 25% (~3500) of genes being derepressed with coincident H3K27me3 reduction and increased PolII recruitment. Of these the 2100 genes that lacked PolII in wild type lines cells gained 5’ paused PolII in the PRC2

mutant, the rest showed increased levels of PolII at the promoter (Chopra et al. 2011). It has been reported that in some cases demethylases such as JMJD3 are required to release paused PolII into productive elongation (Chen et al. 2012). These promoter specific effects are unlikely to be the only output as H3K27me3 exists in broad peaks. The repressive output of H3K27me3 (if any) appears to be through the recruitment of PRC1. PRC1 CBX subunits have a high affinity for H3K27me3, PRC1 then causes histone ubiquitination, nucleosome compaction and directly interacts with transcriptional machinery, finally facilitating DNA methylation. It further appears likely that PRC2 mediated H3K27me3 deposition may inhibit transcription in other ways including: recruitment of non PRC1 proteins, impeding binding of activators or initiation/elongation factors, impacting nucleosome packaging, preventing interactions with chromatin remodellers and excluding H3K27ac (reviewed in (Simon and Kingston 2009, 2013). Remarkably, PcG mediated silencing can be retained post mitosis in daughter cells, although the molecular mechanisms of this remains largely unknown. It may be that whilst H3K27me3 is reduced post mitosis, enough exists to allow PRC2 to 'fill in' the mark and perpetuate repression (Hansen et al. 2008).

1.1.3 Transcription factors

The character of a cell is determined by its specific transcriptional program, which is primarily defined by what subset of transcription factors (TFs) it expresses. TFs bind to DNA to activate or repress a set of genes under its control. Transcription factors are often directed to the genes under their control via DNA binding motifs, although a good deal of deviation is observed with many of these consensus motifs and it is likely that other chromatin or protein-protein interactions aids transcription factor recruitment (reviewed in (Li et al. 2000; Georges et al. 2010; Sive and Gottgens 2014)). The B cell-development pathway has long been used as a model for transcription factor interplay and is of vital importance to understanding mechanisms of EBV induced immortalisation. Here I will describe the importance of transcription factors in B-cell development and focus on a subset of these pertinent to this work.

1.1.3.1 Transcriptional regulation of B-cell development

B-cell development is one of the best studied in terms of temporal transcriptional control due to the ease of identifying subsets of cells and their natural proliferative abilities (Hardy et al. 2007)(Fig 1-1). How a given cell develops into a B cell rather than a T cell, natural killer cell, macrophage or erythrocyte depends on a series of developmental branching points, resolved

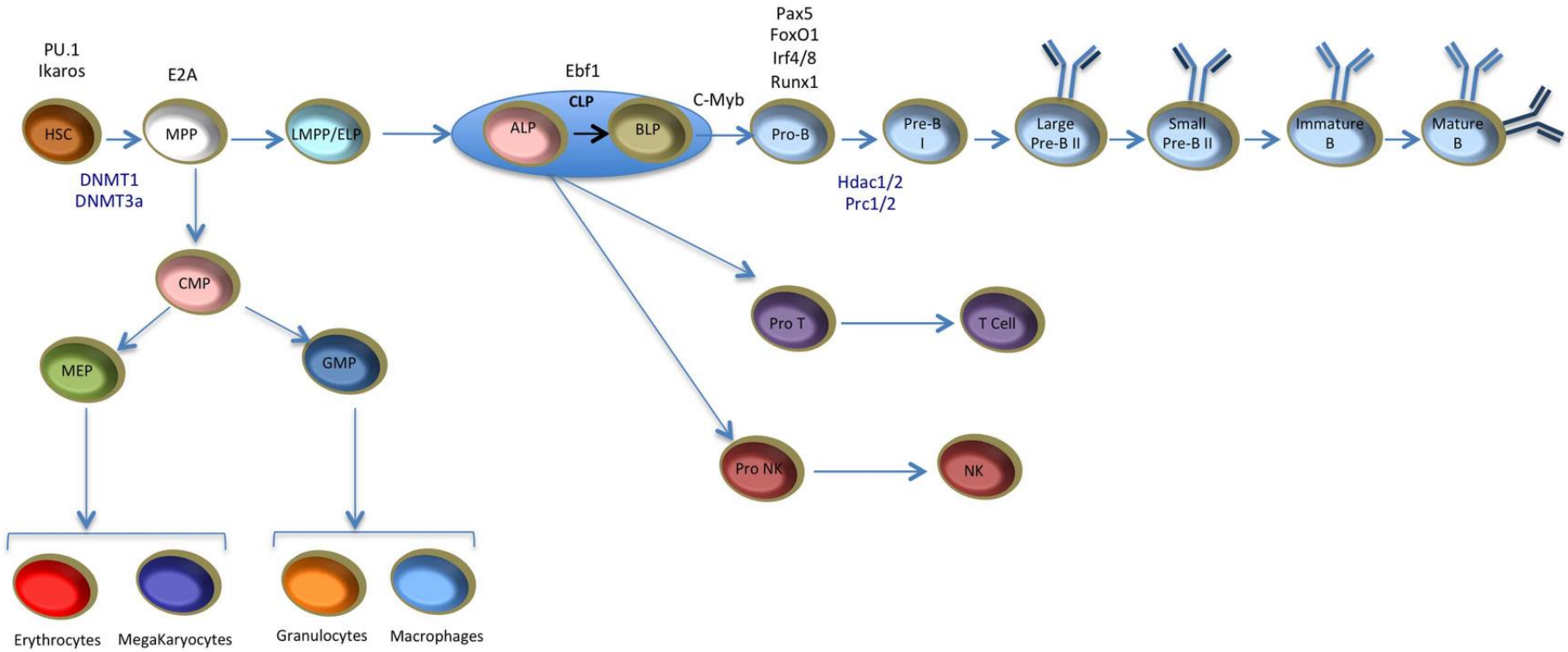


Figure 1-1 Schematic for B cell development. Successive developmental branching points for B cell differentiation are indicated along with other cell fates. Transcription factors and chromatin modifying enzymes are shown according to when their activity is detected or required. If a transcription factor is required at multiple points in this pathway only the earliest event is depicted. HSC, hematopoietic stem cell; MPP, multipotent progenitors; MEP megakaryocyte-erythrocyte progenitors; LMPP, lymphoid-primed multipotent progenitors; ELP, early lymphoid progenitors; CMP, common myeloid progenitors; GMP, granulocyte-macrophage progenitors; CLP, common lymphoid progenitors; ALP, all lymphoid progenitors; BLP, B cell biased lymphoid progenitor; Pro NK, progenitor natural killer cells, NK, natural killer cells. (Choukrallah and Matthias 2014)

by the balance of transcription factors, their dosage, chromatin modifications (enhancer or promoter priming) and stochastic events. Figure 1-1 shows this process along with when relevant transcription factors are expressed. At the DNA level all of classes of cell are identical, their fate and function is decided by the repertoire of transcription factors it expresses, and transcription factor binding profile, which is dependent on chromatin structure and histone modification. The generation of mature B cells is a progressive process, each step characterised by a specific transcription program. There is clearly temporal regulation; factors such as PU.1, E2A and Ikaros are expressed very early, before the developmental branch for lymphoid or myeloid fates. Others such as EBF1 are expressed only during/after commitment to the lymphoid branch in common lymphoid progenitors (CLPs). This suggests a transcriptional hierarchy (reviewed in (Choukrallah and Matthias 2014)).

The first progenitor from hematopoietic stem cells (HSCs) are multipotent progenitors (MPPs), these have lost the capacity for self-renewal but still have multilineage differentiation potential (Adolfsson et al. 2001). The loss of myeloid potential *en route* to lymphoid cell fate is gradual and is associated with increased expression of the flt3 receptor (Adolfsson et al. 2005). The developmental branch point of B-lymphocyte or macrophage (CMPs) is primarily resolved by the dosage of PU.1 (DeKoter and Singh 2000). The inactivation of PU.1 in sorted CLP cells allows B-cell differentiation, primarily towards a lymphoid cell fate, suggesting that PU.1 exerts its role in B-cell development early and is not required for later stages (Iwasaki et al. 2005). The basic helix-loop-helix proteins E12 and E47 heterodimerise to create E2A which is expressed in MPPs and is required for CLP formation. E2A knock out cells do not create any B cell progenitors due to the absence of EBF1 (Borghesi et al. 2005). E2A upregulates *EBF1* which is expressed in CLPs and these proteins cooperatively regulate most B-cell specific genes facilitating their conversion into B cells (Kee and Murre 1998). Mice lacking EBF1 fail to express important B-cell genes such as *CD79a*, *CD79b*, *IGLL1* and do not undergo any Immunoglobulin recombination (Lin and Grosschedl 1995). Furthermore, ectopic expression of EBF1 in HSCs skews differentiation towards B-cell lineage and the expression of EBF1 in E2A deficient mice rescues B-cell differentiation although these cells are unable to proliferate following IL-7 induction (Bain et al. 1997; Zhang et al. 2003; Seet et al. 2004).

PAX5 is a multifunctional transcriptional regulator which is induced by EBF1 in pro-B cells. It is expressed at a relatively stable level throughout the B-cell lineage until differentiation into plasma cells where its expression is reduced (Fuxa and Busslinger 2007). Loss of PAX5 arrests cells at the pro-B stage and its expression is regarded as being required for initiating and

maintaining B-cell commitment (Nutt et al. 1999; Rolink et al. 1999). PAX5 not only activates many B-cell specific genes but represses essential receptors for other differentiation pathways (Delogu et al. 2006). The later stages of B-cell development occur away from the bone marrow and immature B cells can be found in peripheral lymphoid organs. Mature, naïve B cells account for ~65% of all B cells in the peripheral blood (Perez-Andres et al. 2010). Naive B cells (that is B cells that have not encountered an antigen) created in this way circulate the blood stream until an antigen is encountered.

The process of maturation of an antigen stimulated mature naïve B-cell to a memory B-cell (or another fate) is also tightly controlled at a transcriptional level. Upon antigen stimulation naïve B-cells are driven to specific microenvironments in the lymphoid organs where they form germinal centres (GCs); here they proliferate and form a large clone of cells and undergo somatic hypermutations in their Igg genes such that the exact specificity of each cell for the activating antigen is slightly different (Jackson et al. 2013). Progression from this stage is thought to proceed through centroblasts (CD77+), centrocytes (CD77-), then on to memory B-cells or plasma cells. Klein *et al* (Klein et al. 2003) sorted these four cell fractions and performed micro-array experiments to characterise transcriptional alterations in these cell subtypes. Comparing naïve B-cells to centroblasts they found that 457 genes were differentially regulated. Cell proliferation, chromatin remodelling, checkpoint transition and pro apoptotic genes were generally all up regulated whereas negative cell cycle regulators, anti-apoptotic, cytokine and chemokine receptor and adhesion genes were strongly repressed in this transition. Interestingly, centroblasts and centrocytes were found to have little significant differences in gene expression although this may be due to the heterogeneity of the centrocyte subpopulation as these cells give rise to both memory and plasma cells. Importantly they found that memory B-cells generally represent a reversion to a naïve B-cell pattern of gene expression. Centrocyte to memory B-cell transition was shown to involve 267 genes, whereas comparing differentially expressed genes between the naïve and memory subpopulations involved only 62 genes. Of note, it was found that memory B-cells expressed significantly higher levels of *CD27*, which upon encountering its T-cell cell surface ligand CD70, promotes differentiation into antibody secreting plasma cells (Jang et al. 2013; Klein et al. 2003).

The purpose of B-cell development is to create an effective immune system, capable of defending the host from foreign pathogens and to develop a memory to cope with future insults whilst not targeting any host tissue (Fig 1-2). To achieve this, throughout this process B-

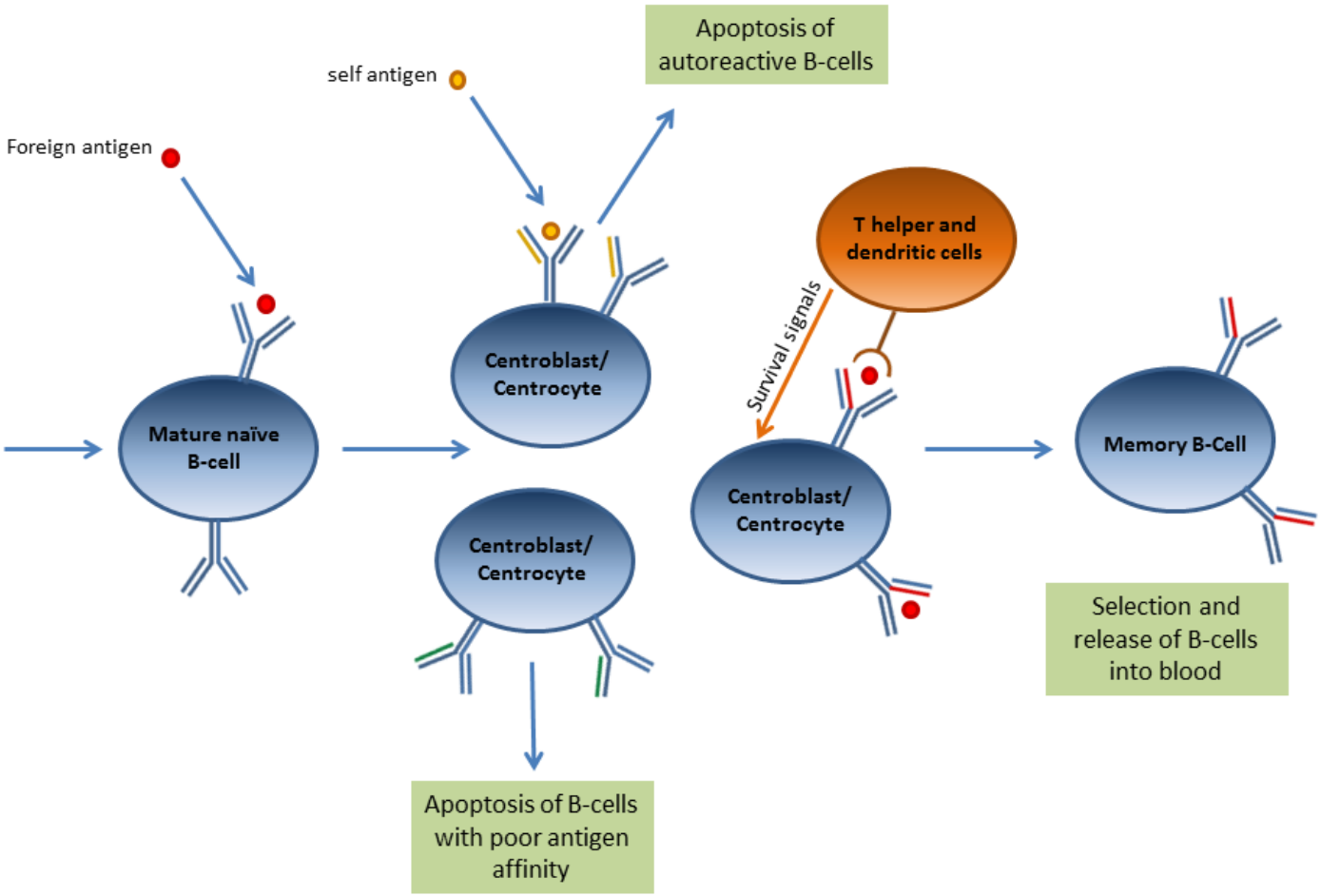


Fig 1-2 Schematic of B-Cell development following antigen stimulation. Naïve or mature naïve B-cells encountering forgein antigen (red circle) upon leaving the bone marrow progress to the germinal centre where they undergo proliferation and changes in gene expression (Centroblasts and Centrocytes). Here cells also aquire somatic hypermutations and class switch recombination (shown by changing green, red and yellow variation on surface receptor). Cells expressing receptors with superior affinity for the imminuzing antigen will outcompete others for T helper survival signals and progress to memory B-cells in the blood.

cells are subject to multiple checkpoints to assess their antigen specificity, signalling strength and autoreactivity, the result of which is the apoptosis or developmental regression of a large proportion of developing B-cells (reviewed in Rowland et al. 2013; Scholz and Cancro. 2012). It has been estimated that approximately 50% of B-cell Receptors (BCRs) of Immature B-cells in the BM are autoreactive (Wardemann et al. 2003). These are eliminated from the naïve repertoire by the process of tolerance and only non autoreactive B-cells are permitted to enter the periphery. Essentially BCR signalling alone (tonic signalling) is required for cells at this stage of development, and those that are 'antigen' stimulated revert to a more pro or pre B-cell state and undergo further V(D)J recombination to alter their BCR (Nossal 1983; Osmond 1991; Schram et al. 2008; Verkoczy et al. 2007). The reason for this is likely that antigen stimulated cells in the BM are almost certainly recognising a self-antigen. In the periphery naïve to mature antigen unstimulated B-cells are further selected for based on cytokines such as BAFF, which control the fate and quantity of pre-immune cells in the periphery as well as selecting for cells that have functional cytokine receptors (Stadanlick et al 2008).

In the GC, the induction of proliferation is counterbalanced by the induction of pro-apoptotic genes (Klein et al. 2003) such that only cells receiving adequate survival signals are permitted to progress, this process selects for B-cells which have higher antigen affinity than their near clonal neighbours by a process of competition for limited amounts of antigen and T-follicular helper cell interaction (Goodnow et al. 2010; MacLennan 1994). Follicular helper T-cells produce the IL-21 cytokine which can activate STAT3 and STAT5 in B-cells which in turn activates both BCL6 and BLIMP which promote and inhibit a memory B-cell fate respectively (Schmidlin et al. 2009). It is unknown whether B-cells encountering IL-21 signalling are predisposed to the memory or Plasma cell fate or if kinetic or stochastic mechanisms resolve this developmental branch point, however the BCL6 and BLIMP expression appear to be mutually exclusive (Cattoretti et al. 2005). Interestingly both PU.1 and PAX5 were shown to repress BLIMP1 suggesting that these TFs promote a memory B-cell fate in the GC (Shaffer et al. 2002; Su et al. 1997).

There are many more consensus binding sites for TFs than sites that are actually bound. This suggests that co-factors and chromatin marks are also crucial to the direction of TFs to target genes. However TFs have also been shown to induce changes in the chromatin landscape. We are starting to develop an understanding about the chronology of these events and much of this work has used the B-cell differentiation pathway as a model due to the hierarchical nature of TF expression. One model suggests that regulatory elements are in a pre-active state in early

progenitor cells before transcriptional initiation occurs. This 'gene priming' has been proposed to be conducted by 'pioneer TFs' (Zaret and Carroll 2011). The requirements for a 'pioneer TF' are that it must bind to a regulatory element prior to other factors, it must be able to target repressed sites such as those with compacted chromatin and it must induce chromatin remodelling to a more permissive state (reviewed in (Zaret and Carroll 2011)). Such sites would often present as poised enhancers or promoters prior to further TF binding and gene activation and such sites are abundant throughout the hematopoietic system (Ghisletti et al. 2010; Mercer et al. 2011). An example candidate for a 'pioneer TF' is PU.1 which can bind its cognate site in condensed chromatin, can recruit histone methyl transferases and chromatin remodelling complexes resulting in sites that have the mark of a poised site; H3K4me1 but no H3 or H4 acetylation. Subsequent collaborative binding of EBF1, E2A and HATs could then lead to enhancer activation and gene transcription (Fig 1-3) (Zaret and Carroll 2011).

In a study in which *E2A* expression was conditional, Mercer *et al* examined histone modifications before and after *E2A* expression and subsequent differentiation (Mercer et al. 2011). Many enhancers important to B cell differentiation were shown to be pre marked with H3K4me1 in MPPs, preceding commitment to myeloid/lymphoid lineage. These sites were enriched for PU.1 motifs in pre-*E2A* expressing cells whereas H3K4me1 sites after *E2A* expression were enriched for E2A, EBF1 and PU.1 motifs. These data suggests that TFs can be recruited by chromatin modifications and can create them. PU.1 binding can directly induce H3K4me1 at certain loci, yet many bound sites lack this mark, suggesting a requirement for additional factors in a locus specific manner (Heinz et al. 2010). Interestingly, at sites where PU.1 binds to existing H3K4me1 sites PU.1 was shown to induce chromatin remodelling (Heinz et al. 2010). At the *PAX5* locus it appears that PU.1 and IRF4 binding precedes gene activation and can induce active chromatin marks, priming the gene for subsequent activation by EBF1 (Decker et al. 2009). However EBF1 can also cause positive epigenetic changes in B-cell progenitors (such as at the *CD79a* promoter where EBF1 can bind despite non permissive chromatin) and PAX5 can induce both activation or repression of genes via the recruitment of histone modifications (Maier et al. 2004; Cho et al. 2007; Gao et al. 2009; McManus et al. 2011). It is therefore likely that the term 'pioneer TF' is not useful unless used in a specific context. Perhaps all TFs can display the previously mentioned qualities given the right context and those that do display these qualities may not possess them at other loci. For example at some loci PU.1 cannot bind without factors such as E2A and EBF1. What determines which loci are permissible to an individual TF acting as a 'pioneer' likely depends on protein-protein

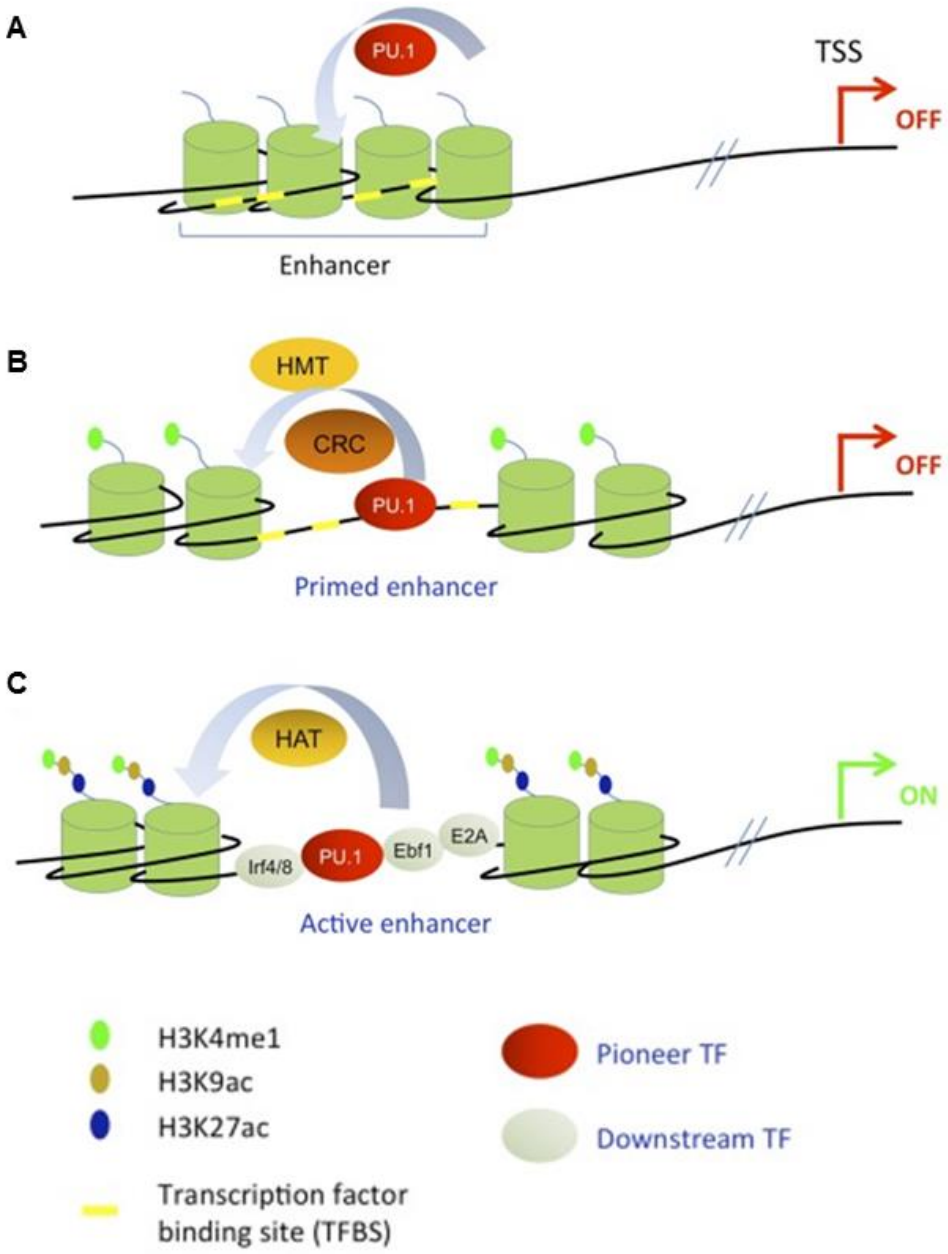


Figure 1-3 Model of pioneer transcription factor activity. (Choukrallah and Matthias 2014) Gene priming by a 'pioneer TF. (A) In this model an inactive gene with an enhancer site in compacted chromatin is bound by PU.1 before any other factors. (B) PU.1 then recruits histone methlytransferases (HMT) and chromatin remodelling complexes (CRCs) to create a more permissive chromatin environment and deposition of markers of enhancers (H3K4me1). (C) subsequent association with other transcription factors such as IRF4, EBF1 or E2A causes further activating histone modifications by the recruitment of Histone acetyl transferases (HATs), resulting in H3K9ac, H3K27ac and gene activation.

interactions, chromatin modifications, DNA modifications and even long non coding RNAs (reviewed in Choukrallah and Matthias 2014).

1.1.3.2 PU.1

PU.1 belongs to the ETS family of TFs and is essential for the development of lymphoid cells, macrophages and neutrophils (McKercher et al. 1996). PU.1 is expressed in HSCs, lymphoid and myeloid progenitors and in mature B-cells (Klemsz et al. 1990; Akashi et al. 2000; Iwasaki et al. 2005). PU.1 is not required for later differentiation steps in B-cells despite the fact that it is still expressed, suggesting that it has important functions other than specifying B cell fate although there is limited evidence of what these roles might be. ChIP sequencing of PU.1 showed that it primarily associates with intra and inter-genic sites suggesting a role in transcriptional regulation via enhancer elements. Furthermore, PU.1 binding sites at promoters were highly conserved throughout various HSC derived cells whereas enhancer binding sites were highly cell type specific (Heinz et al. 2010). Analysis of enriched motifs at PU.1 binding sites revealed that in B-cell progenitors E2A (TCF3), EBF1, OCT and NFKB motifs were enriched whereas in macrophages, C/EBP and AP1 were enriched. This suggests that cell-type specific binding is at least partly dependent on collaborative binding or interaction with other TFs. Importantly the dosage of PU.1 was found to be critical to lineage specification at developmental branches. High PU.1 expression was shown to correlate with macrophage and neutrophil cell types, intermediate with B-cells and low with erythroid and T-cell types (Fig 1-4) (DeKoter and Singh 2000; Back et al. 2004). The mechanism of dose dependence of PU.1 remains poorly understood.

1.1.3.3 EBF1

EBF1 was first identified as a DNA binding protein which could activate *CD79a*, an essential component of the B-cell receptor (BCR) and its signalling functions (Hombach et al. 1990; Gold et al. 1991; Hagman et al. 1991). It can form stable homodimers in the absence of DNA and has an unusual DNA binding motif (its closest relative is NFKB with 14% sequence identity), which associates with a long palindromic consensus motif with a 2bp spacer (Hagman et al. 1995; Siponen et al. 2010). EBF1 can activate B-cell specific genes including *Pax5* which causes B-cell lineage commitment (O'Riordan and Grosschedl 1999). Further lineage commitment is also provided directly by EBF1 as it downregulates expression of genes driving alternate lineages (Thal et al. 2009). At target loci EBF1 can initiate epigenetic changes but often requires existing chromatin modifications directed by upstream TFs such as PU.1. For example at *CD79a*, EBF1

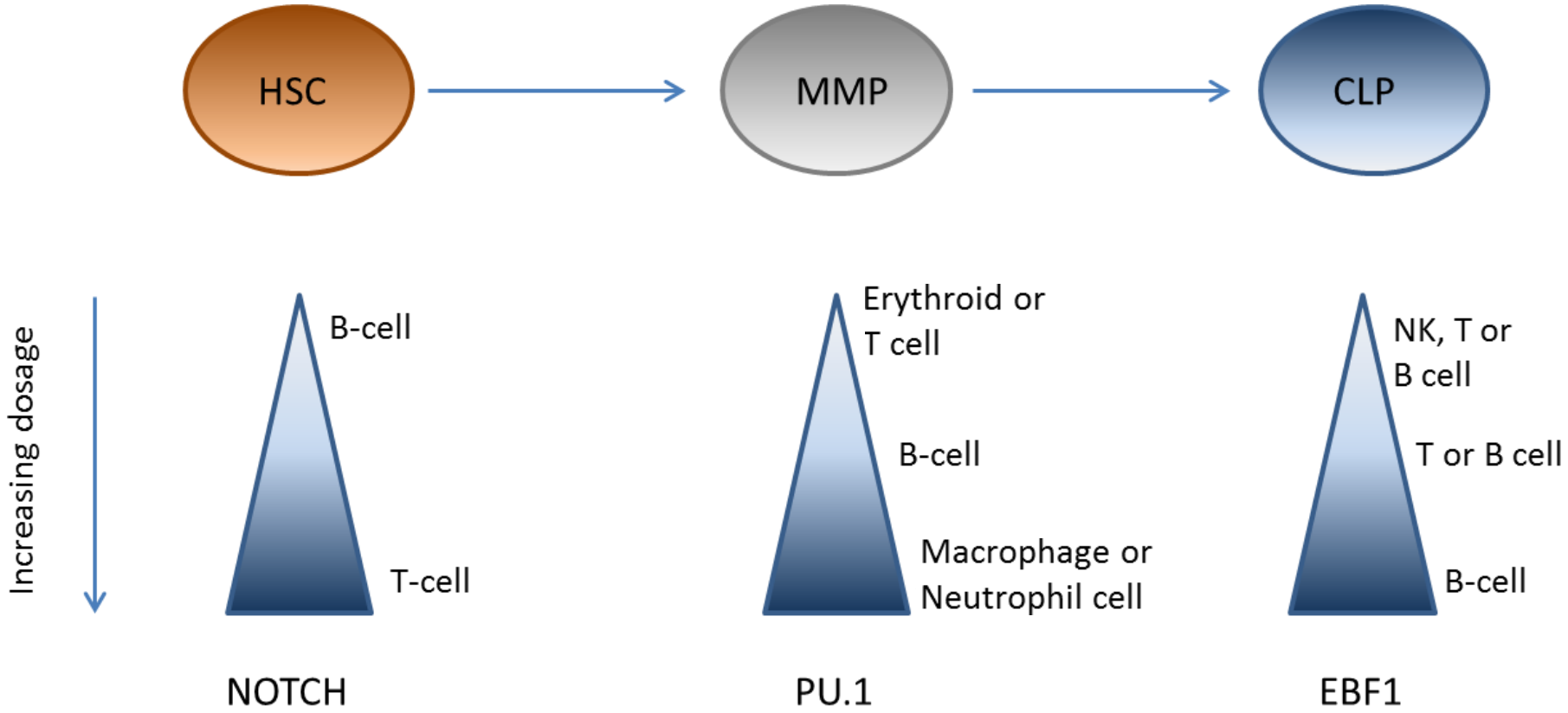


Figure 1-4 Transcription factor dosage is important to B cell development signalling. Cell types in which key transcription factors modulate cell fate and function are displayed at the top of the figure. Below is the transcription factor which is expressed and exerts its function on that cell type. Dosage of signalling from that transcription factor is schematically represented with increase of dosage as indicated. The cell types that can develop from cells expressing a particular dosage of TF are displayed on the right of each gradient.

binding occurs after PU.1 and RUNX1 associate with an intragenic enhancer and subsequently causes increased chromatin accessibility and decreased DNA methylation, facilitating *CD79a* expression (Gao et al. 2009). *EBF1* is transcriptionally upregulated by PU.1, E2A and FOXO1 and the loss of EBF1 results in complete developmental arrest at the CLP stage of development (reviewed in (Hagman et al. 2012)). Furthermore, ectopic expression of EBF1 in non-lymphoid cells that lack upstream factors involved in B-cell development causes activation of at least part of the B-cell developmental program (Romanow et al. 2000; Medina et al. 2004). Similarly to PU.1 the dosage of EBF1 is important for its function. CLP cells expressing low EBF1 can differentiate into NK, B or T cells, those with intermediate EBF1 expression differentiate into B or T cells whereas those with a high dosage of EBF1 are committed to a B-cell fate only (Fig 1-4) (Mansson et al. 2010). This developmental branch point may be decided by stochastic mechanisms as EBF1 upregulates its own expression, as well as repressing suppressors of *E2A* and upregulating *PAX5*, both of which would result in increased *EBF1* expression (Smith et al. 2002; Pongubala et al. 2008; Thal et al. 2009).

Less is known about the role of EBF1 after progenitor B cells leave the bone marrow as immature or mature B cells. However, there is emerging evidence that it has a role as a tumour suppressor as *EBF1* mutations are commonly found in acute lymphoblastic leukaemia (ALL) cases and 25% of relapsed ALL cases show mutations at this locus (Mullighan et al. 2007; Yang et al. 2008). Furthermore *EBF1* mutations have been shown to be implicated in solid tumours including breast and adenocarcinoma cells (Neve et al. 2006; Jones et al. 2008). This may be due to its role in maintenance of cell fate and cell differentiation. ChIP-sequencing of EBF1 has confirmed its role as one of the primary regulators of B-cell genes and its ability to recruit, and be recruited by histone modifications (Lin et al. 2010; Treiber et al. 2010).

1.1.3.4 NOTCH/RBPJκ

NOTCH signalling is highly conserved and is involved in a vast array of biological processes, including development, differentiation, lymphogenesis and regulation of B-cell activity (reviewed in (Radtke et al. 2010; Greenwald 2012)). NOTCH is a membrane protein which upon ligand interaction undergoes cleavage by A disintegrin and metalloprotease (ADAM) proteins. Cleaved intracellular NOTCH is then transported to the nucleus where it interacts with recombination signal binding protein – Jκ (RBPJκ or CBF-1) and recruits coactivators such as mastermind proteins (MAML) at RBPJκ consensus DNA binding sites. This nucleates a transcriptional activation complex (Gordon et al. 2008; Kopan and Ilagan 2009). Some known targets of NOTCH signalling include *MYC*, *HEY1*, *HES1* and chemokine receptors (Weng et al.

2006). Humans have 4 NOTCH proteins (NOTCH1-4) which use Delta like 1, 3, 4 and Jagged 1, 2 as ligands (Bray 2006). These are non-redundant and whilst there is overlap in function between different NOTCH-ligand interactions they fulfil distinct biochemical processes. For example, NOTCH1 signalling has been implicated in T-cell development. Hematopoietic stem cells (HSC) are not generated in cells lacking NOTCH1 or RBPJκ. However, NOTCH/RBPJκ signalling is not required for HSC maintenance and NOTCH signalling must be controlled to prevent ectopic differentiation (Kumano et al. 2003; Maeda et al. 2007). Like PU.1 and EBF1, the dose of NOTCH signalling is important, with HSCs undergoing more NOTCH signalling more likely to become T cells rather than B cells (Fig 1-4) (Dallas et al. 2005). Inducible inactivation of NOTCH1, RBPJκ or MAML1 results in a block in T cell development and ectopic B cell development, whereas overexpression of intracellular NOTCH results in ectopic T-cells at the expense of the B-cell population (Pui et al. 1999; Radtke et al. 1999; Han et al. 2002). *NOTCH1* is also upregulated in mature T-cells after T-cell receptor activation and is responsible for promoting proliferation and IFN-γ production (Palaga et al. 2003). *NOTCH2* on the other hand is preferentially expressed in B-lymphocytes and is essential for the development of marginal zone B cells (MZB) along with ADAM proteins and RBPJκ (Saito et al. 2003; Gibb et al. 2010). NOTCH signalling in mature B-lymphocytes can cause differentiation into antibody secreting cells and enhanced BCR activation (Santos et al. 2007; Thomas et al. 2007). Importantly, NOTCH signalling in GC B cells after exposure to antigen protects them from apoptosis by interacting with ligands expressed on dendritic cells (Yoon et al. 2009).

1.1.4 Role of DNA elements

1.1.4.1 Transcriptional regulation via promoter elements

Much of the regulation of transcription is initially coordinated by DNA elements, including the core promoter, promoter-proximal elements and distal sites such as enhancers, silencers, insulators and locus control regions. These DNA elements are usually modular, allowing sets of cellular factors to bind in a coordinated fashion to create a vast number of unique expression patterns. Combinatorial control of gene expression by factors bound at multiple DNA elements also allows cells to rapidly respond to environmental or developmental stimuli. It is noteworthy that of the ~25,000 human protein coding genes only ~1850 are DNA binding transcription factors (Venter et al. 2001).

The core promoter is the site at which RNAPII and general transcription factors bind, and defines the transcription start site (TSS) and direction of transcription (Smale and Kadonaga

2003). The core initiation complex is sufficient to transcribe a gene but generally only produces low levels of mRNA. The core promoter is recognised through DNA sequences that aid in recruiting members of the RNAPII initiation complex (Table 1-2). However, statistical analysis of ~10,000 known promoters revealed that only one in eight contain a TATA box, and one quarter possess none of the proposed core promoter DNA elements (Gershenzon and Ioshikhes 2005). This suggests that there may be other core promoter elements which have yet to be described or the known sequences can be much more degenerate than originally thought. Another hypothesis is that the exact sequence is secondary to the DNA secondary structure at the core promoter ((Florquin et al. 2005). There is also evidence that the composition of core promoter elements contributes the regulatory specificity of distal regulatory inputs (Morris et al. 2004).

For 60% of human genes the promoter is near a CpG island, an area with a high G and C content typically 500bp-2kb in length, which can be methylated to modulate binding of activators (Venter et al. 2001). Methylated DNA is indicative of gene repression as many activators are not able to bind to methylated forms of their DNA recognition sequences. Furthermore proteins such as MeCP2 can specifically bind methylated CpG nucleotides and recruit histone modifying complexes to establish a repressive chromatin structure (Jones et al. 1998).

Basal transcription can be stimulated by the action of activators which are broadly defined as sequence-specific DNA binding proteins. Activators generally bind to promoter-proximal elements within a few hundred base pairs of the core promoter (Ptashne and Gann 1997) and can categorised as belonging to one of the following families: zinc finger, homeobox, helix-loop-helix, basic leucine zipper, fork head, ETS and POU DNA binding domain proteins (Pabo and Sauer 1992). The DNA binding site for transcription factors is sequence specific but many are degenerate. Transcription factor consensus binding motifs are therefore created by examining all known binding sites and determining the most represented nucleotides (Claessens and Gewirth 2004). The degeneracy of transcription factor binding sites may in fact have biological significance as the precise sequence can modulate activator binding strength and affinity (Yang et al. 2014), dimerization partners (Claessens and Gewirth 2004) and even activator structure (Lefstin and Yamamoto 1998). Importantly, it has also been shown that PU.1 and Ets-1 (ETS family proteins with little amino acid identity in their ETS domain binding site) both bind to the same ETS consensus site but via different mechanisms (Wang et al. 2014).

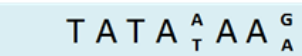
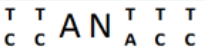
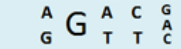
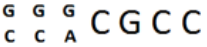
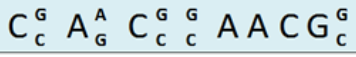
core promoter element	abbreviation	consensus motif
TATA box	TFIID	
initiator element	Inr	
downstream promoter element	DPE	
TFIIB recognition element	BRE	
motif ten element	MTE	

Table 1-2 promoter proximal elements (adapted from(Maston et al. 2006)). Current known core promoter elements, their common abbreviation and consensus motif are shown.

Transcription factors (activators) can stimulate transcription in numerous ways. In some cases TFs increase PIC formation via direct interactions with components of the transcriptional machinery (Orphanides et al. 1996). Further roles in aiding PIC assembly, initiation, elongation and reinitiation have also been described (Lee and Young 2000). Furthermore transcription factors can play a role in modulating local chromatin structure to facilitate transcription by recruiting chromatin structure modifiers (de la Serna et al. 2005). Indeed, recent work suggests that the binding of a particular transcription factor to a piece of DNA allows for accurate prediction of histone modifications present at that site (Benveniste et al. 2014). The effects of activators on transcription can be synergistic; that is, a gene targeted by more than one activator can have an effect on its transcription greater than the sum of each by itself. This is likely due to activators aiding recruitment of different parts of the PIC or altering different aspects chromatin structure and histone modification (Green 2005). Activators and transcription factors can have their activity further controlled by the action of co-activators. These do not exhibit intrinsic sequence specific DNA binding activity but instead use protein-protein interactions with DNA bound activators. Co-activators function in many of the same ways as activators, modifying chromatin landscape and altering the composition of the core transcriptional machinery. Co-activators/repressors can positively or negatively affect transcription dependant on the gene in question and the proteins it is associated with (Lonard and O'Malley 2005).

1.1.4.2 Long-range transcriptional regulatory elements

The first vertebrate distal DNA element was described at the immunoglobulin heavy chain locus where it was shown to dramatically increase the expression of a number of nearby genes. It was able to do this regardless of the orientation of the element and acted over distances of several hundred to several thousand base pairs (Banerji et al. 1983). These are the features of an enhancer element, these element characteristically function independently of orientation and distance from the promoter. A single promoter can be acted upon by distinct enhancers at different times or in different tissues, allowing more unique expression patterns than could be gained from promoter-proximal elements alone (Atchison 1988). Enhancers typically contain a cluster of transcription factor binding sites (TFBS), the spatial organisation and orientation of which is important to its function. Strikingly, it has been shown that even inserting 6bp of random DNA between two TFBS in an enhancer reduces its activating ability by ~17 fold (Thanos and Maniatis 1995). Enhancers are functionally similar to promoter proximal elements; in fact their only defining feature may be the distance over which they act. It appears protein-

bound enhancers use many of the same mechanisms to stimulate transcription as promoter proximal elements and can physically contact the promoter in question via DNA looping (Fig 1-5) (Vilar and Saiz 2005). Interestingly, contacts have also been detected between enhancers and promoters on different chromosomes. In fact one enhancer was shown to interact with multiple promoters on different chromosomes, and it is likely that some promoters are contacted by multiple enhancers from different chromosomes (Spilianakis et al. 2005; Lomvardas et al. 2006). How enhancers specifically target promoters is not well understood.

Silencers share many of the same attributes as enhancers and promoter-proximal elements only their effect on transcription is primarily negative. Silencers typically act independent of orientation or distance and are binding sites for negative transcription factors or repressors and co repressors. Indeed some activators can become repressors by interactions with particular co-factors (Murayama et al. 2004). The mechanism of transcriptional repression can be as straightforward as allosteric inhibition of activator binding to a nearby site, or by direct competition for a binding site. Some can establish repressive chromatin structures by recruiting histone modifying or chromatin stabilising factors, thereby inhibiting general transcription factor binding (Srinivasan and Atchison 2004). Unlike enhancers, there is little evidence for silencers repressing transcription via looping

Another important DNA element involved in the regulation of transcription the insulator element. Generally, these elements block genes from being affected by the transcriptional activity of neighbouring genes through association with proteins or via chromatin structure. Insulators have been reported to prevent enhancer-promoter communication and to prevent the spread of chromatin modifications, delineating chromatin domains such as heterochromatin or euchromatin. These two mechanisms appear to be separable activities, ie not all insulators prevent both (Recillas-Targa et al. 2002). The transcription factor CTCF appears to be one protein that can mediate insulator activity, reviewed in (Holwerda and de Laat 2013). The mechanism of insulator function is not well understood. One model proposes that by interacting with transcriptional activators and histone modifiers, insulators create areas of chromatin permissive to transcription thereby blocking the spread of heterochromatin whilst preventing bound activators from interacting with target promoters (Defossez et al. 2005). A second model is that insulators physically separate areas of DNA by interacting with each other or nuclear attachment proteins thereby forming large scale chromatin secondary structures. This model is supported by recent genome wide chromosome conformation

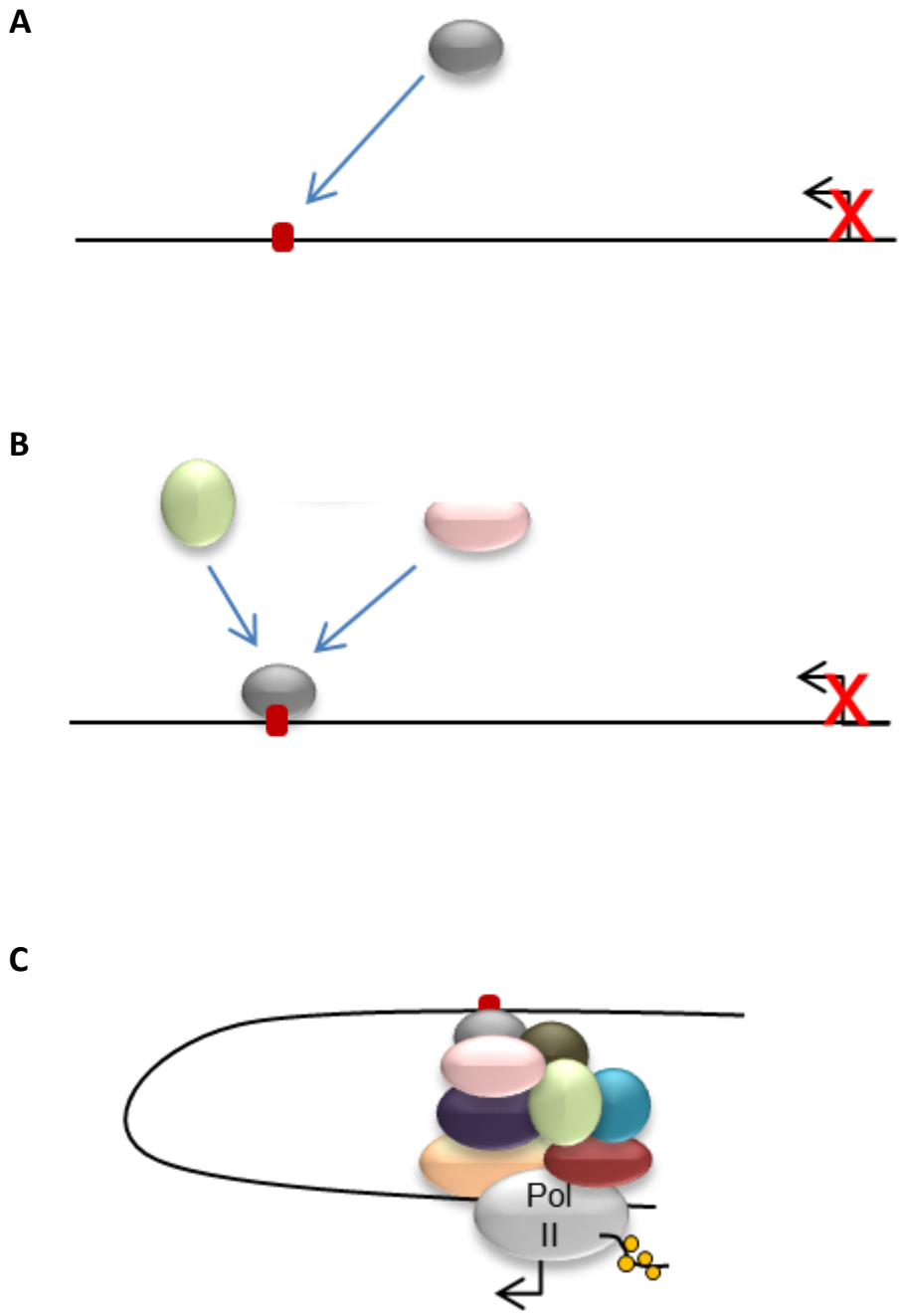


Figure 1-5 Model for gene activation via enhancer looping. (A) A repressed promoter with a downstream unoccupied enhancer element indicated in red. In this situation DNA would likely be associated with repressive histone modifications and condensed chromatin. A 'pioneer TF' indicated in grey can associate with this enhancer element. (B) Association of the 'pioneer TF' with the enhancer facilitates the recruitment of other TFs, co factors and histone modifying enzymes. (C) The enhancer complex physically associates with TFs at the promoter facilitating the recruitment of the PIC and transcriptional initiation.

capture (HiC) studies which observe large-scale domain separation, a process possibly mediated by CTCF induced DNA looping (Lieberman-Aiden et al. 2009).

CTCF is a transcription factor which can act as a repressor (Lobanenkov et al. 1990) or activator (Klenova et al. 1993) of transcription. It acts at insulator elements to prevent enhancer, promoter communication and insulates chromatin modifications (Bell et al. 1999; Recillas-Targa et al. 2002). CTCF can also promote promoter enhancer looping (Handoko et al. 2011). At the β-globin locus CTCF insulates the activity of a local locus control region (LCR) and prevents heterochromatin spreading into the region using a mechanism that involves looping between two CTCF sites either side of the locus in a cell type specific manner (Farrell et al. 2002; Bulger et al. 2003; Palstra et al. 2003; Splinter et al. 2006). CTCF is a DNA binding protein with a large and degenerate consensus motif. This has been shown to be functionally relevant to CTCF function with changes in which sites (degree of difference from consensus) are bound during the course of mouse embryonic stem cell development (Plasschaert et al. 2014). The methylation of CTCF binding sites prevents CTCF binding which in turn allows allele specific gene expression via the removal of a CTCF bound insulator element and subsequent promoter enhancer looping (Bell and Felsenfeld 2000; Hark et al. 2000). CTCF itself can influence DNA methylation via interactions with PARP1 mediated DNMT1 inactivation resulting in methyl free DNA at CTCF binding sites (Zampieri et al. 2012), but has also been shown to increase the repressive H3K27me3 mark as knock down of CTCF results in less H3K27me3 in repressed domains (Van Bortle et al. 2012). This process is perhaps mediated via concentrating polycomb group proteins via the insulator activity of CTCF. ChIP sequencing experiments reveal that CTCF associates with tens of thousands of sites in the human genome, only ~30% of which are conserved across cell types (Kim et al. 2007; Wang et al. 2012). At *MYC* and other genes, CTCF has been shown to bind to promoter proximal elements, and the removal of a CTCF site from the *MYC* promoter induced gene repression via DNA methylation (Barski et al. 2007; Gombert and Krumm 2009). CTCF mediated looping appears in part to be dependent on its association with cohesin (Rubio et al. 2008) and at the *MHCII* locus CTCF, cohesin and CIITA can induce promoter enhancer loops increasing gene expression (Majumder and Boss 2011). It can also recruit the basal transcription factor TAF3 to intergenic sites in embryonic stem cells allowing TAF3 dependant enhancer looping and gene activation (Liu et al. 2011). CTCF is even implicated in RNAPII stalling potentially aiding the incorporation of weak exons or RNAPII processivity (Shukla et al. 2011). CTCF activity is likely decided by which other transcription factors it associates with at a given binding site in a tissue and genomic context specific manner, reviewed in (Holwerda and de Laat 2013).

LCRs are regions of the genome with clusters of enhancers, silencers and insulators bound by transcription factors, coactivators, repressors, co repressors and chromatin modifiers can be found and they typically regulate entire gene clusters or loci (Li et al. 2002). One of the best described LCRs is involved in the regulation of the β-globin locus. Here five globin genes that are differentially expressed during development are arranged in order of developmental expression. The distal LCR confers a high level of activity to all genes in this cluster, a function which was shown to be orientation dependant as inversion of the LCR ablated transcriptional activation (Tanimoto et al. 1999). If LCRs are functionally distinct from their individual components remains to be determined.

The *in silico* prediction of DNA elements such as transcription factor binding sites, promoter-proximal elements, enhancers, silencers and others is challenging. Even high throughput experiments such as ChIP-sequencing are difficult to analyse as function cannot readily be dissected from binding data. Programmes such as MEME-ChIP (http://meme.nbcr.net/meme/doc/meme-chip.html) are able to use consensus binding motifs to predict potential transcription factor binding sites, and this coupled with ChIP sequencing data is improving the *in silico* process. But how transcription factors bind to sites with no consensus motif, or why they don't bind to perfect consensus motifs is a paradox that so far only gene by gene analysis can answer. It may eventually be possible to incorporate all histone modification, transcription factor and gene expression (RNA sequencing data) to provide accurate predictions for novel enhancers, silencers, insulators and LCRs. This is likely to be cell type and environment specific, and is the ultimate goal of the ENCODE project.

1.2 Epstein-Barr virus

1.2.1 Discovery

Epstein-Barr virus (EBV) is a gamma herpesvirus first discovered in 1964 by Anthony Epstein, Yvonne Barr and Bert Achong (Epstein et al. 1964). Unlike most viruses which are identified due to an acute illness in the patient, EBV was discovered due to its association with Burkitt's lymphoma (BL), a B-cell lymphoma first described in the 1950s by the Irish surgeon Dennis Burkitt (Burkitt 1958). BL manifests as a malignant tumour of the jaw, eye or abdomen with high occurrence in young children in Africa. In a subsequent 'tumour safari' Burkitt noticed that the cancer had a particular geographical distribution: firstly, BL was found to be endemic only in the malaria belt of equatorial Africa, where it is the most common childhood cancer; secondly, within central Africa a higher incidence of BL was observed in the lowlands with

tropical climate compared to higher-level areas with less rainfall (Burkitt 1962). This led Burkitt to believe that the tumour had infectious agent involvement, possibly transmitted by the mosquito as malaria is.

EBV was confirmed as the first human tumour virus in the 1970s when culture media in which BL cells had been growing was found to be able to immortalise B cells (reviewed in (Andersson 2000; Bornkamm 2009)). More recently it has been discovered that nasopharyngeal carcinoma (NPC), a cancer found most commonly in southeast Asia and southern China is also significantly linked with EBV infection (Chang et al. 2009). The association between EBV and these endemic cancers is strong, almost 100% of endemic BL cells are infected with EBV. (Kutok and Wang 2006). Early studies into EBV prevalence in BL patients and control groups discovered that approximately 95% of the adult population worldwide is infected with EBV and in the vast majority of cases such an infection is asymptomatic (Henle et al. 1969). Epstein-Barr virus has become one of the best studied herpes viruses and is a model for viral latent infection. Its study may aid the treatment of numerous EBV-associated malignancies and progress our general understanding of tumourigenesis, viral infection and B-cell biochemistry.

1.2.2 Initial infection

In common with other members of the gammaherpesvirinae family, EBV most commonly infects B cells, although in the case of EBV it has also been shown to productively infect epithelial and T cells. The virus exists as a 172kb double stranded linear DNA molecule encased in a protein envelope. EBV gains entry to B cells by the interaction between the cellular compliment receptor 2 (CR2 or CD21) protein and a viral membrane fusion complex containing gp350/220 (Tanner et al. 1988). EBV proteins gH/gL and gp42 then bind human leukocyte antigen (HLA) class II which triggers viral entry into the B cell (Li et al. 1997). gH, gL, gB and gp42 are indispensable for EBV infection of B cells whilst only gH, gL and gB are indispensable for epithelial cell infection (Li et al. 1995; Li et al. 1997; Haan et al. 2001). How EBV enters epithelial and T-cells is less well understood. It is likely that EBV can infect B cells at multiple stages of B-cell development but only immature and mature infected B-cells are found in the peripheral blood *in vivo* and lymph nodes.

Upon internalisation the terminal repeats at either end of the viral genome fuse forming a circular episome which can then be transcribed by cellular transcription machinery to enable expression of the range of EBV proteins required to establish viral persistence (Fig 1-6). The circularisation of the EBV genome is detectable 20 hours post infection and is required for

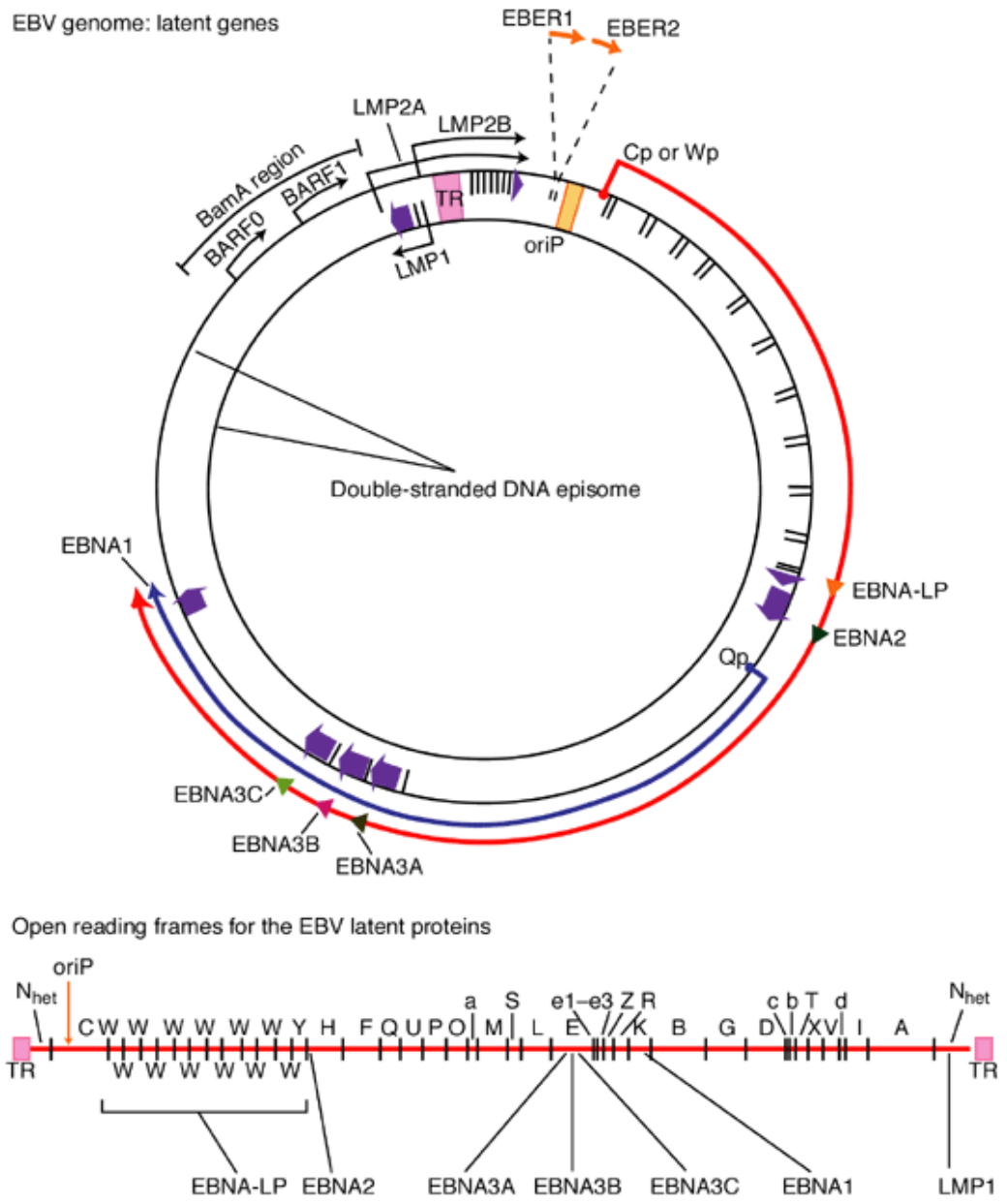


Figure 1-6 Epstein Barr Virus genome. (Murray and Young 2001) Diagram showing the location of latency genes within the viral genome. Origin of replication (OriP) is shown in orange and the primary message produced from Wp or Cp is shown in red. This is then alternatively spliced to produce the latent genes shown with purple arrows. The transcript from Qp encoding EBNA 1 is shown in blue. Terminal repeats where circularisation takes place is shown in pink. The location and details of the LMP promoters and locations of the EBV encoded RNAs are also displayed.

immortalisation. B-cell activation is indicated by a coincident increase in CD23 expression, a marker of activated B cells (Hurley and Thorley-Lawson 1988). B cells infected by EBV *in vitro* in the absence of an immune response are transformed into indefinitely proliferating lymphoblastoid cell lines (LCLs).

The initial infection of epithelial cells in the nasopharynx is followed by further lytic activation creating infected cell plaques of up to 10^5 cells and amplifying viral output before plaque growth is halted by the immune system. During this time virus is continually shed into the saliva such that mouth washing every two minutes fails to reduce the quantity of EBV viral progeny in the saliva (Hadinoto et al. 2009). Viral infection of epithelial cells occurs in a non CR2 (CD21) dependant manner and likely occurs through alternative viral glycoproteins gH and gL with an unknown epithelial cell receptor (gHgLR). Viral progeny created in epithelial cells are better able to infect B cells due to higher levels of viral glycoprotein gp42 and viral progeny created in B cells are better at infecting epithelial cells. This suggests that an epithelial amplifying infection is likely to occur prior to initial infection and on virion release (reviewed in (Hutt-Fletcher 2007)). This process is vigorously controlled by the immune response at all stages resulting in persistent infection, continual production of viral progeny and host-virus balance.

Upon infection EBV expresses a 'pre latent' gene profile including some genes normally associated with lytic cycle, the effects of which are controversial. One of these genes is *BZLF1* which encodes Zta (ZEBRA or Z), the master regulator of the switch from latency to lytic cycle, which is expressed alongside a set of latency genes described as the latency III pattern of gene expression. Other lytic associated genes such as *BRLF1*, another 'immediate early' lytic gene which encodes Rta (which is directly transactivated by Zta) are also expressed. Late lytic genes are not expressed in the pre-latent stage and no viral progeny are released (Kalla et al. 2010). However Halder *et al* have showed in a similar system that a more extensive set of lytic genes are expressed in the pre-latent stage of infection including late lytic genes such as the major capsid protein Bclf1, which is detectable as early as 6 hours post infection, and that such an infection does produce infectious viral progeny (Halder et al. 2009). They postulate that this strategy allows the expansion of the number of infected B cells on initial infection and that either some cells survive lytic replication and return to a latent form or that some cells become lytic upon initial infection and die whereas others express a latent pattern of gene expression. It was subsequently proven that Zta has a much higher affinity for some binding sites with methylated CpG dinucleotides and that the incoming viral DNA is unmethylated and does not

produce viral progeny (Kalla et al. 2012). This may occur because PCNA, which is required to tether cellular DNA methyltransferase to newly replicating DNA, does not participate in herpes viral DNA amplification (Muylaert and Elias 2010). Discrepancies between the two studies may be due to the preparation of EBV virions. Halder *et al* transfected 293t cells with EBV GFP BAC to produce viral progeny whereas Kalla *et al* used HEK293 cell transfected with maxi-EBV plasmid p2098 (Halder et al. 2009; Kalla et al. 2010). This could result in differences between methylation states of either viral progeny.

The expression of lytic genes early in infection may facilitate immune evasion strategies. *BCRF1* has also been shown to be expressed in the pre-latent stage and encodes vIL-10 (viral IL-10 homolog) which has been shown to protect EBV-infected cells from NK cell mediated elimination and is secreted to prevent release of antiviral cytokines abrogating the CD4+ T cell response (Jochum et al. 2012a). *BNLF2a* is also expressed early in infection and impairs recognition of virally infected cells by EBV specific CD8+ T lymphocytes (Jochum et al. 2012a). The expression of this restricted lytic subset of genes has also been shown to support the proliferation of resting B cells (Halder et al. 2009) and protect against apoptosis (Altmann and Hammerschmidt 2005). Intriguingly it has recently been reported that viral mRNA for *BZLF1*, *BNLF2a*, *BCRF1*, *LMP1*, *LMP2A/B* and *EBNA 2* has been found to be present within the virion and is transduced and translated immediately upon infection which may contribute to the immune evasion and rapid proliferation provided by the pre latent stage of expression (Jochum et al. 2012b). The pre-latent phase is transient and appears to completely cease 1-2 weeks post infection as a result of extensive CpG methylation of the viral genome and is followed by a strictly latent gene profile (Kalla et al. 2010).

1.2.3 EBV latent gene transcription

On initial infection of B cells by EBV *in vitro* and most likely *in vivo,* a latency III pattern of gene expression is adopted resulting in the expression of nine viral proteins, including the Epstein-Barr nuclear antigen (EBNAs), latent membrane proteins (LMPs) and non-coding viral RNAs such as the Epstein-Barr virus-encoded small RNAs (D'Souza et al. 2006) and BART miRNAs. Initially the virus uses the BamHI W promoter (Wp) which drives expression of *EBNA 2* and EBNA leader protein (*EBNA-LP*). The latent W promoters consist of tandemly arrayed W repeats and is dependent on cellular factors such as CREB/ATF (Kirby et al. 2000). The production of EBNA 2 drives a switch to the C promoter (Cp) from which a long primary transcript encoding *EBNA 2*, *EBNA 3A*, *EBNA 3B* and *EBNA 3C* is transcribed. Wp is then largely silenced by epigenetic changes (Hutchings et al. 2006) such as DNA methylation of CpG islands

(Tierney et al. 2000). EBNA 2 binding to Cp results in a significant local increase in H4 and H3 acetylation by EBNA 2 mediated p300 recruitment. This recruits bromodomain protein 4 (BRD4) which in turn allows recruitment of positive transcription elongation factor b (pTEFb). At Cp there is significant promoter proximal stalled RNAPII, negative elongation factor (NELF) and DRB sensitivity inducing factor (DSIF)(Palermo et al. 2011). Promoter proximal stalling is required at Cp for keeping the promoter free of nucleosomes and to necessitate the recruitment of pTEFb thereby maintaining serine 2 CTD phosphorylation (Palermo et al. 2011).

The single transcript produced from Cp in a latency III pattern of gene expression is alternatively spliced to produce *EBNA 1*, *EBNA 2*, *EBNA 3A*, *EBNA 3B* and *EBNA 3C*. *LMP1* and and *LMP2B* are expressed from their own bidirectional promoter. The *LMP2A* promoter is in the opposite orientation to that of *LMP1* and the two genes transcripts overlap (Kerr et al. 1992) (Fig 1-6). In all forms of latency the noncoding Epstein Barr encoded RNA (EBER) 1 and 2 and the Bam H1 A rightward transcripts (*BARTs*) are expressed (Kim do et al. 2011). Latency III gene products promote EBV-induced cell growth, B cell transformation, immune evasion and resistance to apoptosis and also induces a vigorous cytotoxic T cell response (Middeldorp and Pegtel 2008). This is normally sufficient to eliminate rapidly proliferating cells *in vivo*, however in cases of immunodeficient individuals this can result in life threatening neoplasias. *In vitro,* and in the case of post-transplant lymphoproliferative disease (PTLD), the absence of an immune response results in infected B cells persisting as EBV immortalised lymphoblastoid cell lines (LCLs) continuously expressing the latency III pattern of gene expression. *In vivo* it is likely that latency III is only adopted during primary infection and perhaps sporadically during persistent infection but such cells are destroyed by the host immune system.

1.2.4 Course of a latent infection in a healthy host

The latency III gene products drive resting naïve B cells to proliferate as active b-blasts, it is these cells that then migrate to the GC. It is generally agreed that EBV gene expression is gradually shut down, going through latency II in the GC, and latency I patterns of gene expression in the final memory B cell (Fig 1-7)(reviewed in (Roughan and Thorley-Lawson 2009; Roughan et al. 2010)). Infection of resting B-cells *ex vivo*, without a GC or immune response instead gives rise to permanently proliferating lymphoblastoid cell lines (LCLs) continuously producing latency III gene products (Young et al. 2000). There is general agreement that the process of LCL outgrowth resembles the process of initial infection of naïve B-cells prior to differentiation and long term persistence *in vivo*. Therefore the initiation and sustaining of proliferation at this early stage in the host is a required step in the EBV life cycle. Latency III

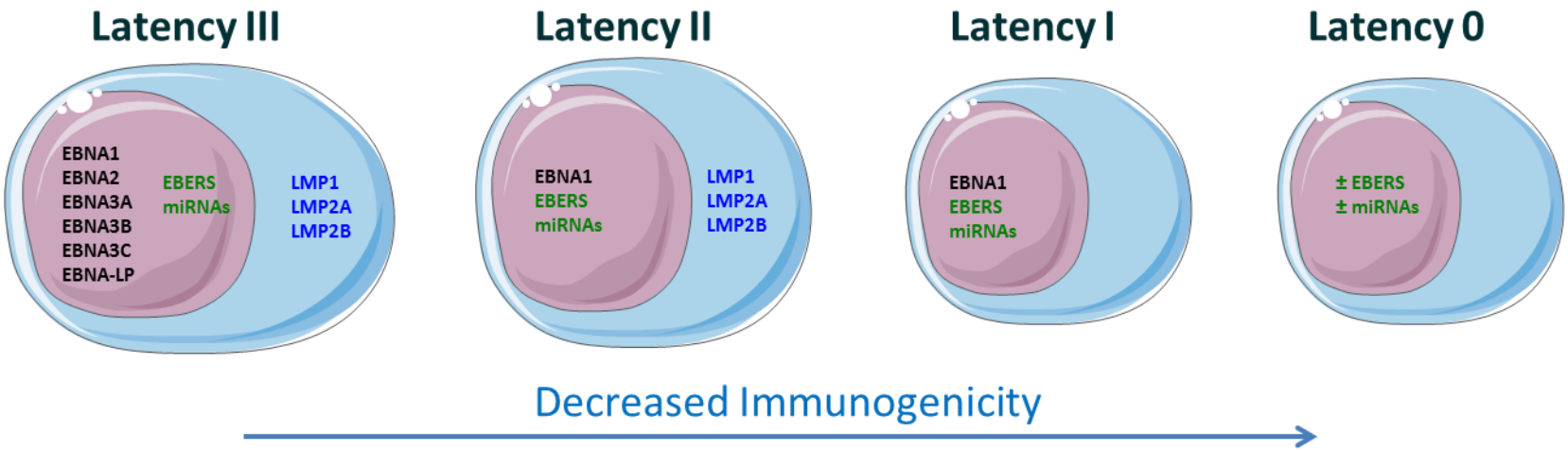


Figure 1-7 EBV latent gene expression patterns. Latency III, II, I and 0 patterns of gene expression are listed. The EBV nuclear antigens are displayed in black and the EBV encoded RNAs are displayed in green. The viral membrane proteins are displayed in blue. Genes are annotated in the subcellular compartment in which they exert the majority of their function. The decrease in immunogenicity that accompanies latency type gene expression patterns is indicated below.

gene products induce this hyperproliferative state, and maintain it despite intrinsic barriers to aberrant replication and a strong immune response. Nikitin *et al* showed that EBV infection caused a transient period of hyperproliferation 3-4 days post infection which correlated with an increased DNA damage response pathway activation. Analysis of transcript levels in resting, hyperproliferating and outgrown LCLs revealed that a number of genes involved in pathways including cell proliferation and DNA damage response were upregulated from resting to proliferating cells and subsequently attenuated in LCLs (Nikitin et al. 2010). This hyperproliferative phase may be similar to that induced by in vivo infection and, intriguingly, proliferation rates observed were similar to those of B cells in the GC (MacLennan 1994). The roles of individual latency III gene products in this process will be discussed later in this introduction.

The initial burst of viral induced proliferation in infected naive B cells is rapidly curbed by the developing EBV specific CD8+ T cell response (reviewed in (Odumade et al. 2011)). However during the period of initial latent gene expression EBV infected activated B cells migrate to the GC (Laudencia-Chingcuanco et al. 2011) where a latency II pattern of gene expression is observed. In latency II EBV expresses EBNA1 and LMP1, 2A and 2B using the viral Q promoter and LMP specific and bidirectional promoters (see Fig 1-6). In the GC activated B cells first rapidly proliferate in the 'dark zone' accumulating somatic hypermutations (SHM) in their immunoglobulin genes which results in increased affinity of the B cell receptor for its antigen in some of the proliferating cells. Further differentiation causes migration to the 'light zone' where B cells with the highest affinity antigen outcompete others for the survival signal produced by T helper cells and undergo class switching by DNA recombination (CSR) to alter the function of their immunoglobulins. EBNA 2, which is not expressed in latency II, inhibits GC reactions, partly by inhibiting activation-induced cytidine deaminase (AID), which is essential for both CSR and SHM (Tobollik et al. 2006). It also downregulates *BCL6* and *TCL1*, both germinal centre associated genes, of which *BCL6* is absolutely indispensable for the GC reaction and SHM (Boccellato et al. 2007). The few B cells that survive the GC reaction differentiate to become either plasma cells or memory B cells. Cells that do not receive adequate survival signals rapidly undergo apoptosis (reviewed in (Spender and Inman 2011)).

EBV infected cells surviving the GC reaction display all the hallmarks of a normally selected memory B cell. Such cells have accumulated SHM, have undergone CSR and the resulting immunoglobulins appear to have experienced normal antigen selection pressure as proper protein fold and increased antigen binding affinity are displayed (Souza et al. 2005). Also,

infected as well as uninfected B cells express GC transcription factor BCL6 and upon receiving survival signals uptake the BCL2 survival protein, suggesting normal B cell maturation (Roughan et al. 2010). It is also observed that infected B cells surviving the process express the expected chemokine profile and are therefore phenotypically indistinguishable from normally produced GC memory B cells (Roughan and Thorley-Lawson 2009). EBV therefore is likely to play a role in improving the survival chances of its host B cell in the GC rather than altering the process dramatically. LMP1 and 2 possess CD40 and BCR signalling functions providing pro survival signalling. Levels of LMP2A have been discovered to correlate with those of integrin-alpha-6 (ITGα6), which is involved in cell migration and improved cell-cell contact, perhaps further contributing to survival signalling (Pegtel et al. 2005). LMP1 has been implicated in both increased cell proliferation and immune system evasion by regulating various chemokines (eg. *CCL17*), cytokines (eg. *IL6*) and adhesion molecules (eg. *CD18*)(Middeldorp and Pegtel 2008). The fate of B cells leaving the GC appears to be dependent on the cytokine environment and relative strength and duration of interaction with antigen and T helper cells (Bernasconi et al. 2002). Prolonged T helper cell exposure may be more likely to lead to a memory B cell fate which is the desired destination for EBV infected cells as it allows persistent infection in the peripheral blood. It is therefore likely that EBV exploits the normal GC reaction rather than overriding it to produce latently infected memory B cells which can persist in the peripheral blood.

Upon entering the peripheral blood as a memory B cell EBV assumes a latency I pattern of gene expression in which only EBNA 1, BARTs and the EBERs are expressed. EBNA 1 appears to induce no dramatic immune response *in vivo* due to purine bias in the glycine alanine repeat region forming G-quadruplexes. This results in reduced EBNA 1 translation and antigen presentation (Murat et al. 2014). However, whilst EBNA 1 is protected from presentation to cytotoxic T cells via the MHC class I pathway it does appear that it is presented to CD8+ and CD4+ helper T cells *in vivo*, perhaps via exogenous uptake of antigen via dentritic cells and subsequent ‘cross priming’ in both primary and memory stages of infection (Blake et al. 2000; Leen et al. 2001). Infected memory cells are phenotypically indistinguishable from uninfected cells and the viral episome mimics the host DNA; it is packaged in cellular histones, co-replicates with cellular DNA once during S phase and is split into two daughter cells during mitosis. Latency 0 is also observed *in vivo* as EBV infected cells expressing no viral protein and only EBERs. This however may be due to EBNA 1 expression correlating with sporadic memory cell division requiring EBV to preserve its own genome (Hochberg et al. 2004). The switching

between the various forms of latency and lytic cycles of EBV is largely controlled by particular promoters within its genome being activated at specific times.

In latency I circulating memory B-cells the Qp promoter is active, producing only the EBNA 1 nuclear antigen, which along with oriP is sufficient for the virus to maintain its own genome (Leight and Sugden 2000; Fejer et al. 2008). Six EBNA 1 binding sites within the oriP region have also been shown to increase the activity of EBNA 2 induced transcriptional activation from Cp (Zetterberg et al. 2004). In order for EBV to infect new hosts, infected circulating memory B cells return to the GC in the tonsils and undergo periodic reactivation to produce infectious virus. As circulating memory B cells express an extremely limited set of genes the signal to do this likely comes from host factors during memory cell differentiation. The process of memory cell differentiation upon antigen stimulation usually results in plasma cells which are essentially biochemical factories for antibody production. In infected cells the terminal differentiation into a plasma cell and the associated changes in cellular gene expression and epigenetic environment results in the expression of Zta resulting in lytic phase activation (Laichalk and Thorley-Lawson 2005; Murata et al. 2012). In the lytic phase the virus arrests the cell cycle in G0/G1 thereby inhibiting growth whilst it amplifies its own genome 100-1000 fold; producing the components of its viral progeny before lysing the cell and releasing viral progeny (Tsurumi et al. 2005). Viruses released from B-cells then infects epithelial cells as virions in the saliva appear to originate from a non B cell (Jiang et al. 2006).

1.2.5 The host immune response to EBV

EBV infection is tightly controlled by the host immune system. Perturbing the balance of EBV infection with the host immune response is a common cause of EBV related diseases. If initial infection occurs in adolescence 25% of individuals develop infectious mononucleosis (IM), which is an acute but self-limiting disease with symptoms including fever, sore throat, lymphadenopathy and is characterised by a large quantity of atypical lymphoblasts of CD8+ T cell origin in the blood (Crawford et al. 2006). IM is not entirely caused by late exposure to EBV as it sometimes occurs in childhood infection and may in fact be underdiagnosed (Chan et al. 2003). In acute IM up to 10% of all B-cells in the blood are EBV positive. Natural killer (NK) cells are key components of the early innate immune response to microbial infection. NK cells have been shown to inhibit EBV induced B-cell transformation if added within a few days post infection, an effect that is at least in part due to the release of IFN-ϒ (Lotz et al. 1985). However, this is unlikely to be the dominant pathway for controlling EBV infection as individuals undergoing T cell transplantation are prone to EBV driven lymphoproliferative

disease despite the fact that NK cell levels are restored rapidly post transplantation (O'Reilly et al. 1997).

The primary host response to EBV comes from CD8+ and CD4+ T-cells. In response to acute IM there is a dramatic proliferation of EBV epitope specific CD8+ T-cells responses, accounting for 40% of total T-cells for lytic epitopes. There is also a CD8+ T-cell latent epitope response that accounts for up to 5% of the total CD8+ T-cell response. This is primarily directed against the EBV EBNA 3 family (reviewed in (Hislop et al. 2007)). Lytic specific CD8+ T-cells in IM blood rapidly undergo apoptosis without antigen stimulation due to low levels of expression of anti-apoptotic BCL2 and BclX proteins and high expression of the pro-apoptotic Bax (Tamaru et al. 1993; Callan et al. 2000). This helps to explain the rapid shut down of the CD8+ T cell response to IM as the infection is brought under control. In contrast the CD8+ T-cells with specificity to latent epitopes are less down regulated and as a result undergo a relative increase in the percentage of total circulating CD8+ T-cells (Hislop et al. 2005). After initial infection only 1-50 per million B cells are infected with EBV. These sporadically undergo lytic reactivation when an antigen is encountered. Still epitope specific CD8+ T-cells account for up to 2% of the total pool for lytic epitopes and up to 1% for latent epitopes (Benninger-Doring et al. 1999). This stays relatively stable throughout the host's life; however individuals over 60 tend to show increased levels of EBV epitope specific CD8+ T cells, reaching 14% of circulating T-cells in some cases (Ouyang et al. 2003).

Less is known about the CD4+ host response, and whilst it is crucial for control of EBV infection, CD4+ cells do not rapidly expand in IM; however CD4+ epitope specific responses are detectable in IM blood for viral epitopes (Maini et al. 2000; Precopio et al. 2003). There is also detectable immune responses from memory CD4+ cells against EBV epitopes including EBNA 1, LMP proteins and BZLF1 resulting in IFNϒ production (Dukers et al. 2000; Bickham et al. 2001; Adhikary et al. 2006). Different EBV encoded proteins have different immunogenic properties. The CD8+ response is primarily composed of cells specific to BZLF1 and BRLF1 for lytic cycle proteins and the EBNA 3 family for latent proteins. The CD4+ response generally targets epitopes from BZLF1 for lytic cycle proteins and primarily EBNA 1 for latent proteins although EBNA 2 and EBNA 3C are also targeted (reviewed in (Hislop et al. 2007). This is not consistent across all infected hosts, for example in some cases CD8+ T-cells can be primarily composed of cells targeting epitopes from EBNA 2, EBNA-LP or LMP2 (Lee et al. 1997; Chapman et al. 2001). EBNA 1 has long been considered invisible to the CD8+ T-cell response due to the presence of a glycine-alanine repeat preventing its processing by the proteasome; however in some donors a

CD8+ IFNϒ type response is mounted against EBNA 1 epitopes suggesting the glycine-alanine repeat region offers protection rather than immunity (Lee et al. 2004). To counteract the host immune response EBV has evolved a plethora of immune evasion strategies in both lytic and latent stages of infection. The role of EBNA 2 and EBNA 3 family proteins in immune evasion will be discussed later in this introduction.

1.2.6 EBV and disease

Whilst it does not appear that EBV benefits from killing or compromising its host it is associated with approximately 1% of all tumours worldwide (Delecluse et al. 2007). If primary infection is delayed to adolescence EBV can cause infectious mononucleosis (IM) resulting in rapid proliferation of infected B cells and high viral titres. This is normally controlled by the cytotoxic T cell response but fatal forms of IM do occur where proliferation continues uncontrolled (Williams and Crawford 2006).

EBV was initially discovered through its association with BL, which is extremely aggressive with one of the fastest doubling times of any human cancer (Burkitt 1958, 1962; de Leval and Hasserjian 2009). There are in fact three subtypes of BL, endemic (eBL), sporadic (sBL) and HIV associated. Whilst nearly 100% of eBL are EBV positive only 5-15% of sBL and 40% of HIV associated BL are EBV positive (Wright 1999; van den Bosch 2004; Young and Rickinson 2004). The unifying feature of BL is a translocation of the oncogene *MYC* to an immunoglobulin locus; this translocation is dependent on activation-induced cytidine deaminase (AID) which is highly expressed in GC B cells (Filipovich et al. 1992; Dorsett et al. 2007; Pasqualucci et al. 2008; Allday 2009). Furthermore BL predominantly express a latency I pattern of gene expression. Whilst the EBERs have been shown to have anti-apoptotic potential, EBNA1 has been shown to have limited oncogenic capabilities, raising the question of how EBV is involved in BL (Kitagawa et al. 2000). Hyper activation of MYC signalling as a result of its translocation to the immunoglobulin locus normally results in apoptosis. MYC induced apoptosis is mediated by p53 and Bim and the repression of either of these genes can allow a cell to survive a *MYC* translocation (Allday 2009). Importantly, *Bim* is one of the best characterised targets for EBV induced repression. EBNA 3A and EBNA 3C have been shown to promote H3K27me3 at this site resulting in eventual CpG methylation and persistent gene silencing (Paschos et al. 2012).

Hodgkins lymphoma (HL) is characterised by disrupted lymph node architecture and the presence of Hodgkin/Reed-Sternberg (HRS) cells surrounded by T and B-lymphocytes. Hodgkins lymphomas account for 30% of all lymphoid malignancies and, depending on the

subtype of the disease, is up to 95% associated with EBV (Harris et al. 1999). A common feature of HL is that cells have lost a functional BCR during the GC reaction and should undergo apoptosis due to lack of survival signals. HRS cells express a latency II pattern of gene expression, including LMP1 and LMP2A which mimic CD40 and BCR signalling respectively (reviewed in (Kapatai and Murray 2007)). Interestingly, HRS cells have been reported to lack critical B-cell transcription factors such as PU.1 (Torlakovic et al. 2001). Furthermore the inhibitor of DNA binding family protein ID2, has been shown to be over expressed in HRS cells (Renne et al. 2006). ID2 specifies an erythroid fate for MLL cells and does so by negatively regulating *E2A* (Mathas et al. 2006; Ji et al. 2008); in turn E2A activates *EBF1* which represses *ID2*, a process that is essential for B-cell development (Thal et al. 2009). It has been suggested that the repression of B-cell lineage transcription factors and subsequent T-cell like phenotype of HRS cells is important for escaping apoptosis induced by lack of a BCR.

Nasopharyngeal carcinoma (NPC) is an epithelial cell tumour most commonly found in southern china in 20-50 per 100,000 individuals. NPC is up to 100% associated with EBV depending on subtype and location and in southern china was discovered to be in early, preinvasive lesions indicating the virus in the initiation of tumours (Wolf et al. 1975; Raab-Traub et al. 1987; Pathmanathan et al. 1995). NPCs also express a latency II pattern of gene expression and EBNA 1, LMPs and EBV encoded RNAs have been shown to play a role in its carcinogenesis (Niedobitek 2000; Zheng et al. 2007; Marquitz and Raab-Traub 2012). EBV associated tumours expressing a latency II pattern of gene expression have also been found in T/NK cell lymphomas. These are prevalent in Asia and China and are characterised by an absence of T cell antigens and the expression of the NK cell maker CD56 (Jaffe et al. 1996; Wensing and Farrell 2000).

In immunocompromised individuals EBV can cause numerous diseases. EBV is present in ~60% of HIV related lymphomas, with prevalence varying in subtype. These include: BL, diffuse large B cell lymphoma, extranodal marginal zone lymphoma, peripheral T-cell lymphoma, HL, primary effusion lymphoma, plasmablastic lymphoma of the oral cavity and post-transplant lymphoproliferative disease (PTLD) like lymphoma (Patte et al. 2001). Individuals who have undergone transplantation almost always require the artificial knock down of the immune system to prevent graft vs host disease. A major risk factor in this procedure is the development of PTLD which is nearly 100% associated with EBV. These cells express a latency III pattern of gene expression, the products of which are generally considered to be the primary effector in tumour development (Brink et al. 1997).

1.3 Epstein-Barr virus nuclear antigens

EBV expresses nine proteins during the latency III pattern of gene expression, LMP1, LMP2A, LMP2B and the Epstein-Barr virus nuclear antigens (EBNA) EBNA 1, EBNA 2, EBNA LP, EBNA 3A, EBNA 3B and EBNA 3C. Of these EBNA 2, EBNA 3A, EBNA 3C and LMP1 are essential for in vitro infection and immortalisation.

1.3.1 EBNA 1

EBNA 1 is a homo-dimeric DNA binding protein that binds specific sequences in the host and viral genome. EBNA 1 binding sites within the EBV OriP mediate EBV genome synthesis and non-random partitioning of EBV episomes into daughter cells by tethering the episome to the host genome (reviewed in (Westhoff Smith and Sugden 2013). EBNA 1 can also operate as a transcriptional activator of Cp and *LMP1*, and as a transcriptional repressor at Qp, negatively regulating its own expression via CTCF mediated looping interactions from OriP (Sample et al. 1992; Gahn and Sugden 1995; Altmann et al. 2006; Tempera et al. 2011). EBNA 1 ChIP sequencing studies have revealed extensive interaction with the host genome, one reported 247 significant binding sites at human gene promoter proximal elements and another reported 4785 binding sites throughout the human genome. Interestingly the DNA binding consensus for EBNA 1 is much more variable in the human genome than in the viral genome (Dresang et al. 2009; Lu et al. 2010). Some bound genes such as *STAT1* and *PBX2* have previously been reported to be regulated by EBNA 1, however, the mechanism of EBNA 1 transcriptional regulation remains largely undefined and there is generally a poor correlation between regulated and bound genes suggesting EBNA 1 is not the primary transcriptional regulator of host genes (Wood et al. 2007; Lu et al. 2010). Lu *et al* performed ChIP sequencing analysis of EBNA1 binding sites in the viral and human genome. Surprisingly they found 4785 candidate sites which were enriched for promoter regions of cellular genes. Potential targets of EBNA1 include *MAP3K7IP2*, *CDC7* and *HDAC3* as well as smaller peaks close to *MYC* and *IgH*, *IgD* regions. The result of this binding is as yet unconfirmed (Lu et al. 2010).

1.3.2 EBNA-LP

EBNA LP is coordinately expressed with EBNA 2 within hours of infection of resting B-cells from Wp and is important to EBV mediated transformation (Mannick et al. 1991). EBNA LP commonly co-regulates genes with EBNA 2 and can stimulate EBNA 2 transactivation domain mediated transcriptional activation (Harada and Kieff 1997). EBNA LP can also activate transcription by interactions with Hsp72 to repress histone modifying proteins such as HDACs

and can decrease RBPJκ binding to its repressive partner NCOR (Peng et al. 2007; Portal et al. 2011).

The primary modulators of cellular transcription are EBNA 2 and the EBNA 3 family of proteins which are all incapable of binding DNA themselves and must associate with cellular transcription factors to direct them to DNA.

1.3.3 EBNA 2

EBNA 2 is essential for persistent infection, immortalisation and continued proliferation of EBV infected B cells (Cohen et al. 1989; Tomkinson et al. 1993; Kempkes et al. 1995). EBNA 2 is the master controller of latent viral gene transcription and drives transcription from Cp, producing the long alternatively spliced transcript from which all EBNA proteins are translated (Woisetschlaeger et al. 1991). Furthermore, expression of LMP1, LMP2A and LMP2B depend on EBNA 2 mediated transcriptional activation via association with promoter proximal elements (Wang et al. 1990b; Zimber-Strobl et al. 1993; Laux et al. 1994). These interactions depend on EBNA 2 association with RBPJκ at *LMP2A* and *LMP2B*, and PU.1 at *LMP1*, both of which directs EBNA 2 to target promoters (Ling et al. 1993; Grossman et al. 1994; Henkel et al. 1994; Waltzer et al. 1994; Robertson et al. 1996). Strikingly, EBNA 2 can mimic the effects of intracellular (active) NOTCH in its association with RBPJκ and can functionally replace the intracellular region of NOTCH (Sakai et al. 1998). Furthermore, activated NOTCH1 can transactivate some EBNA 2 regulated promoters (Hofelmayr et al. 1999; Strobl et al. 2000). EBNA 2 transcriptional activation of *LMP1* is further mediated by its interaction with PU.1 and ATF; removal of the PU.1 binding site at the *LMP1* promoter completely eliminates EBNA 2 responsiveness (Johannsen et al. 1995; Sjoblom et al. 1995; Sjoblom et al. 1998) whereas EBNA 2 activation of Cp requires the cellular transcription factor AUF1 (Fuentes-Panana et al. 2000). LCLs appear to ubiquitously express PU.1 which targets many genes essential to B-cell function (Siemer et al. 2008). More recently both EBF and RUNX proteins have been implicated in targeting EBNA 2 to DNA (Zhao et al. 2011).

Once associated with DNA through a cellular binding partner EBNA 2 can activate transcription through numerous mechanisms. EBNA 2 contains an acidic transactivation domain (TAD) which has shown to be essential for transformation and transactivation (Cohen et al. 1991). EBNA 2 transactivation domain fused to GAL4 was shown to upregulate expression from plasmids containing GAL4 binding sites 125 fold compared to GAL4 only (Cohen and Kieff 1991). Multiple general transcription factors including TFIIB, TAF40, TFIIH and TFIIE have been shown

to interact with EBNA 2 implicating it with the recruitment and activation of RNAPII to target promoters (Tong et al. 1995a; Tong et al. 1995b; Tong et al. 1995c). Further transactivation potential derives from EBNA 2 TAD association with HATs including p300, CREB binding protein (CBP) and lysine acetyltransferase 2B (KAT2B or PCAF) (Wang et al. 2000). EBNA 2 also induces permissive chromatin formation at target genes by recruiting chromatin remodelling via RBPJκ, including *LMP2A*. This association could be prevented by deletion of the TATA box at the promoter suggesting that EBNA 2 requires other cellular factors to recruit chromatin remodelling activity (Wu et al. 2000). EBNA 2 also activates repressed subsets of RBPJκ targeted genes by binding to and masking the RBPJκ repressive domain whilst recruiting transcriptional activators to the same sites (Hsieh and Hayward 1995).

Numerous direct cellular targets of EBNA 2 transcriptional activation have been reported including *MYC*, *FGR*, *RUNX3*, *CD23* and *CR2* (*CD21*, the cellular surface antigen through which viral entry is facilitated). The upregulation of MYC causes a dramatic increase in cellular activation and proliferation; furthermore RUNX3 represses *RUNX1*, preventing RUNX1 mediated growth inhibition (Cordier et al. 1990; Knutson 1990; Kaiser et al. 1999; Spender et al. 2002). Genes involved in survival are also affected by EBNA2, such as the p55α subunit of PI3K, which is involved in signal transduction from chemokines, growth factors and cytokines. The PI3K subunit is up-regulated by EBNA2 and subsequent down-regulation by RNAi causes LCLs to undergo apoptosis (Spender et al. 2006). EBNA 2 also activates cyclin D2 expression in resting B cells causing G0 to G1 cell cycle transition (Sinclair et al. 1994).

To date, five gene expression microarrays using cells that conditionally express EBNA 2 have been published. In total they have identified 654 unique cellular genes regulated at the RNA level by EBNA 2 (Spender et al. 2002; Maier et al. 2006; Spender et al. 2006; Zhao et al. 2006b; Lucchesi et al. 2008). All arrays used a chimeric form of EBNA 2 fused to an estrogen receptor in different cell background. These data confirmed many known targets of EBNA 2 regulation including *MYC*, *CR2* and *RUNX3*. However, this technique has identified many novel targets of EBNA 2. Whilst the mechanism of regulation for most cellular genes is unknown and could include downstream targets of EBNA 2, it is clear that EBNA 2 has an important and extensive role in regulating the host transcriptome.

1.3.4 EBNA 3 family

The EBNA 3 family (EBNA 3A, EBNA 3B and EBNA 3C) are expressed following EBNA 2 activation of Cp. They likely arose from gene duplication events as they are arrayed in tandem

in the EBV genome and share a similar gene structure with a short 5' exon and a long 3' exon; however they share little amino acid identity and perform overlapping and distinct functions in EBV biology (Young et al. 2000; Touitou et al. 2005). EBNA 3A and 3C are essential for *in vitro* EBV induced transformation and immortalisation whilst EBNA 3B was shown to be entirely dispensable for this process (Tomkinson and Kieff 1992; Tomkinson et al. 1993; Young and Murray 2003). EBNA 3A deficient virus is capable of immortalising resting B-cells and creating LCLs but with greatly reduced efficiency (Hertle et al. 2009). Like EBNA 2, EBNA 3A, 3B and 3C are incapable of binding to DNA directly but can bind via RBPJκ. Competition for the same cellular TF as EBNA 2 results in EBNA 3 repression of EBNA 2 activities at certain targets. Co-transfection of all EBNA 3 proteins with EBNA 2 inhibited EBNA 2 activation of a chloramphenicol acetyltransferase (CAT) reporter gene plasmid attached to the *LMP1* promoter (Le Roux et al. 1994). EBNA 3A and EBNA 3C were also shown to induce similar inhibitory effects on EBNA 2 activation on the cellular thymidine kinase promoter attached to CAT and four upstream RBPJκ binding sites (Waltzer et al. 1996). However, promoter-reporter constructs were not sensitive to the activity of EBNA 3 proteins in these assays. It was suggested from this and other data that EBNA 3 proteins were not recruited to DNA via RBPJκ but instead inhibited EBNA 2 activation by binding to, and destabilising RBPJκ.

EBNA 3C has also been shown to also interact with PU.1, and co activates *LMP1* though PU.1 binding sites at the promoter (Zhao and Sample 2000). EBNA 3 proteins are both activators and repressors of transcription and the domains responsible within EBNA 3A and 3C have been mapped by GAL4 fusion constructs (Marshall and Sample 1995; Bain et al. 1996; Bourillot et al. 1998; Cludts and Farrell 1998; White et al. 2010). The repression and activation of target genes by EBNA 3 proteins is at least in part mediated by protein-protein interactions allowing EBNA 3 mediated recruitment of co-repressors and co-activators to EBNA 3 sites. All EBNA 3 proteins can interact with CTBP proteins, which are potent oncogenes and transcriptional co-repressors and co-activators (Touitou et al. 2001; Hickabottom et al. 2002; Chinnadurai 2009). EBNA 3A and 3C have also been shown to recruit histone deacetylases and PcG proteins for part of their repressive function (Radkov et al. 1999; Knight et al. 2003; Paschos et al. 2012). The mechanism of transcriptional activation by EBNA 3 proteins is less well known but at least EBNA 3C can interact with the histone acetyltransferase p300, mediating gene activation by histone acetylation (Subramanian et al. 2002).

Much of the current understanding of the cellular targets of EBNA 3 protein regulation comes from gene expression microarray analysis. mRNA levels have been documented in the

presence or absence of EBNA 3A in LCLs and in BL31 and similar data exists for EBNA 3B (Hertle et al. 2009; White et al. 2010). One gene expression micro array dataset exist for cells that are entirely lacking EBNA 3B and have decreased expression of EBNA 3C (Chen et al. 2006). Most studies have examined mRNA levels in various cell lines in the presence or absence of EBNA 3C (Zhao et al. 2006a; White et al. 2010; McClellan et al. 2012; Skalska et al. 2013). Finally data exists for cells infected with a BAC derived virus lacking all three EBNA 3 proteins (White et al. 2010).

It has been established that EBNA 3A and 3C cooperatively repress the key tumour suppressor *p16*^{INK4A} through recruitment of PcG proteins, a process that is dependent on CTBP proteins (Skalska et al. 2010). This has been proposed to be the primary barrier to proliferation as cells without functional EBNA 3C undergo growth arrest following *p16* derepression whereas cells also lacking *p16* do not (Skalska et al. 2013). *Bim* (*BCL2L11*) is another key target for EBNA 3 proteins. Bim is a potent initiator of apoptosis due to its ability to inhibit pro-survival proteins such as BCL2 and its activation of the pro-apoptoic protein BAX which causes cytochrome c release from mitochondria (Enders et al. 2003; Fischer et al. 2007). In the germinal centre Bim is a major regulator of life and death signals involved in B-lymphogenesis and mice lacking Bim accumulate excess lymphoid cells (Egle et al. 2004). Bim is directly activated by MYC which is activated by EBNA 2 and translocated to the immunoglobulin locus in all BL cells (Dang et al. 2005). EBNA 3A and EBNA 3C cooperate to repress *Bim* in a PcG dependent manner suggesting that they are able to counteract the apoptotic effects of a hyperactive MYC (reviewed in (Allday 2013)).

White *et al* performed a gene expression micro array experiment in which BL cells were infected with wild type, or 3A, 3B or 3C knock out or revertant BAC derived virus. 36 genes were shown to be regulated by all 3 EBNA 3 proteins. Gene ontology analysis of EBNA 3C regulated genes revealed that genes involved in cell cycle and mitosis were most significantly enriched as well as genes involved in chromosome organisation. Cells infected with a virus lacking all 3 EBNA 3 proteins had altered transcript levels of 1008 genes which were enriched for gene ontology clusters including cell migration, lymphocyte activation, differentiation and apoptosis. They further discovered that 17 TF associated with B-cell development were altered at the transcript level by EBNA 3 proteins including NOTCH2, EBF1, PU.1 and PAX5 (White et al. 2010). These genes were repressed by different subsets of EBNA 3 proteins. Investigation of the promoters of *NOTCH2* (repressed by 3A and 3C) *RASGRP1* (repressed by 3B and 3C) and *TOX* (repressed by all) revealed that H3K27me3 was decreased in knock out lines relative to

wild type, correlating with gene derepression. H3K4me3 however was unchanged creating a poised chromatin state at these genes in the absence of one or more EBNA 3 proteins. H3K9Ac increases correlated with increased TOX and RASGRP1 expression but was unchanged at NOTCH2 suggesting that the mechanism of regulation by EBNA 3 proteins is context specific. Importantly genes regulated by EBNA 3B in an LCL background revealed only modest overlap with genes regulated by EBNA 3B in a BL background. Only 14/89 EBNA 3B upregulated genes in LCLs were also upregulated in BLs, and only 32/108 EBNA 3B repressed genes were commonly repressed in both cell lines (White et al. 2010). This suggests that cell context is crucial for directing EBNA 3 proteins to different targets and modulating transcription. This may be due to different chromatin contexts or abundance of different TFs and cofactors. Importantly, work from this paper implicates EBNA 3B as an essential EBV gene in vivo as many genes regulated by it are essential in the GC reaction which is not reproduced in *in vitro* immortalisation.

Together the EBNA 3 proteins are a highly transforming family of proteins and are some of the main drivers of EBV induced transcriptional regulation, transformation and apoptosis resistance. This is emphasised by Wp restricted BL cell lines. Wp restricted BL cell lines appear to have arose from *in vivo* cells as cells expressing this pattern of genes have been found in original tumour biopsies. Such cells always involve an EBNA 2 deletion but retain the classical *MYC* translocation associated with BL suggesting that EBNA 2 activity is not compatible with *MYC* translocation. The lack of EBNA 2 results in sole usage of the W promoter and prevents normal promoter switching to Cp (Woisetschlaeger et al. 1991). Unlike classical BL these cells express EBNA 3A, 3B and 3C and are highly resistant to apoptosis even compared to classically transformed BL cell lines, making such cells highly prone to tumour development (Kelly et al. 2005).

1.4 Summary

The transcriptome defines the fate and function of a cell. The cellular transciptome is modulated by chromatin accessibility, proximal or distal DNA elements and the binding of transcription factors and cofactors. EBV can alter the fate and function of resting B-cells with a restricted set of viral proteins. The viral proteins EBNA 2, 3A, 3B, 3C are known to alter the transcript levels of thousands of cellular genes, via association with cellular transcription factors; however, remarkably little is known about the mechanism of EBV transcriptional regulation by these factors.

1.5 Aims of my project

To elucidate the mechanism of transcriptional regulation by EBNA 2 and EBNA 3 proteins, we determined to perform chromatin immunoprecipitation coupled with next generation sequencing to generate a library of all DNA elements bound by these factors. Using this data we would examine known regulated genes and examine if EBNA 2 and EBNA 3 binding targeted these genes, as well as elucidate new cellular target genes. As EBNA 2 and EBNA 3 proteins share cellular TFs as binding partners we also resolved to study the interactions between these proteins at potential shared sites and how this impacted gene regulation.

2. Materials and Methods

All reagents were purchased from Fisher Scientific unless otherwise stated

2.1 Tissue Culture

2.1.1 Tissue culture media and supplements

100x Penicillin-Streptomycin-Glutamine (PSG) (Life Technologies)

Contains 100 units/ml penicillin G, 100 µg/ml streptomycin sulphate and 29.2 mg/ml L-glutamine with 0.85% saline and 10 mM citrate buffer. Stored in 5 ml aliquots -20°C. Used at 1x in all media

β-estradiol (Sigma)

Dissolved in DMSO and stored as 100mM aliquots at 4°C.

Dimethyl Sulphoxide (DMSO) (Sigma)

Dulbecco's Phosphate Buffered Saline without $CaCl_2$ and $MgCl_2$ (PBS) (Life Technologies)

Fetal Bovine serum (FBS) (Life Technologies)

Pre-screened for endotoxins (≤5 EU/ml), haemoglobin (≤10 mg/dl) levels and heat inactivated at 56°C for 30 mins. Stored in 50ml aliquots at -20°C.

Freezing mix

80% RPMI media (supplemented with 10% FBS and PSG), 10% FBS and 10% DMSO.

G418 (life technologies)

Dissolved in 100mM HEPES and stored as 200mg/ml aliquots at 4°C.

Hygromycin B (Life Technologies)

Stored as 50mg/ml stock at 4°C.

Sodium Pyruvate (Sigma)

Dissolved in PBS and stored as 100mM aliquots at 4°C.

RPMI 1460 media without L-glutamine (Life Technologies)

RMPI 1460 media phenol red free (Life Technologies)

Thioglycerol (Sigma)

Diluted in PBS and stored as 100mM aliquots at 4°C.

2.1.2 Maintenance of cell lines

All cells were incubated at 37°C with 5% CO_2.

2.1.2.1 Mutu I and III

Mutu I (clone 179) is an EBV latency I BL cell line derived from a 7-year old African black male with Burkitt's Lymphoma (Gregory et al. 1990). Mutu III (clone 48) is a clonal cell-line that arose as a result of drifting of Mutu I in culture to a latency III pattern of gene expression (Gregory et al. 1990). Both cell lines were kindly provided by Prof. Martin Rowe (University of Birmingham) and were passaged 2-3 times weekly in RPMI media supplemented with 10% FBS and PSG.

2.1.2.2 BL31, BL31 wtBAC, BL31 3CKO, BL31 3CREV and BL31 E2KO

BL31 is an EBV negative BL cell line. BL31 wtBAC, BL31 3CKO, BL31 3CREV and BL31 E2KO are BL31 cell lines infected with wild-type (BL31 wtBAC 2, BL31 wtBAC 3), EBNA 3C knock out (BL31 3CKO 2, BL31 3CKO 4), EBNA 3C revertant (BL31 3CREV 3, BL31 3CREV 6) or EBNA 2 knock out (BL31 E2KO) recombinant EBV BACs. EBNA 2 or Individual EBNA3 genes were deleted from the EBV genome by RecA-based homologous recombination between the B95.8 EBV BAC. The deleted gene was then restored in each case to generate revertant BACs (White et al. 2003; Kelly et al. 2005; Anderton et al. 2008). Cell lines were kindly provided by Dr. Rob White and Prof. Martin Allday (Imperial College London) and were passaged 2-3 times weekly in RPMI media supplemented with 10% FBS,PSG, 1mM Na Pyruvate, 50µM thioglycerol and in the case of all cell lines other than BL31, 100µg/ml hygromycin B for BAC selection.

2.1.2.3 BJAB PZ1, PZ3, E3C3, E3C7, E3A1 and E3B2

BJAB is an EBV negative B cell Lymphoma which lacks translocated MYC. PZ1 and PZ3 are BJAB clones transfected with the empty pZipNEOSV(X) expression vector containing neomycin phosphotransferase providing resistance to the neomycin analogue G418. E3C3, E3C7 are BJAB clones transfected with pZipNEOSV(X) expression vector also containing the open reading

frame for EBNA 3C (Wang et al. 1990a). E3A1 and E3B2 are transfected with plasmids expressing EBNA 3A and EBNA 3B respectively. Cell lines were kindly provided by Prof. Alan Rickinson (University of Birmingham) and were passaged 2-3 times weekly in RPMI media supplemented with 10% FBS, PSG and 2mg/ml G418.

2.1.2.4 PER253 B95.8, PER142 B95.8, PER253 3BKO and PER142 3BKO LCLs

PER253 and PER142 LCLs are generated by infection of primary B-cells from donor PER253 and PER142 with the B95.8 strain of EBV. PER253 3BKO and PER142 3BKO LCL are B-cell lines from the same donors transformed with B95-8 EBNA 3B-knock-out recombinant virus (Chen et al. 2005). These cell lines were kindly provided by Dr. Heather Long (University of Birmingham) and were passaged 2-3 times weekly using RPMI media supplemented with 10% FCS and PSG.

2.1.2.5 EREB 2.5 LCL

EREB 2.5 is an LCL transformed by co-transfection of EBV viral strain P3HR1 which lacks EBNA2 and a viral mini plasmid containing an estrogen receptor tagged EBNA2 fusion protein (ER-EBNA 2). ER-EBNA 2 is conditionally active as it is sequestered in the nucleus and released on hormone binding. β-estradiol causes ER-EBNA 2 translocation to the nucleus and EBNA 2 activity (Kempkes et al. 1995). EREB 2.5 cells were kindly provided by Prof. Paul Farrell (Imperial College London) and passaged 2-3 times weekly in phenol red free RMPI media supplemented with 10% FCS, PSG and 1μM β-estradiol. For β-estradiol withdrawal and add-back experiments ER/EB 2.5 cells were incubated in the absence of β-estradiol for 4 days and 1mM β-estradiol was re-added for 6 or 17 hrs prior to cell harvest.

2.1.2.6 D2 and D3 wild type (wt) and EBNA 3A (mtB) LCLs

D2wt1, D2wt2, D2wt3, D2mtB1, D2mtB2, D2mtB3 and D3wt1, D3wt2, D3wt3, D3mtB1, D3mtB2, D3mtB3 are B cells from two donors (D2 and D3) infected with the wild type EBV strain B95-8 (wt1, wt2, wt3) or B95.8 EBNA 3B-knock-out recombinant virus (mtB1, mtB2, mtB3) (Hertle et al. 2009). Cells were kindly provided by Prof. Bettina Kempkes (Department of Gene Vectors, Munich) and were passaged 2-3 times weekly in RMPI media supplemented with 10% FCS and PSG.

2.1.3 Freezing cells

200 mls of cells were pelleted by centrifugation at 1300rpm for 10 minutes at 4°C in a sorvall legend centrifuge and resuspended in 5 mls of freezing mix. Cells were divided between 5

cryogenic vials in 1ml aliquots, and frozen at -80°C in a freezing container (Nalgene) with isopropanol. Vials were transferred to liquid nitrogen storage after at least 24 hrs at -80°C.

2.1.4 Thawing cells

Cells were transferred from liquid nitrogen storage to a 37°C water bath. Once thawed, the 1ml of cells were added to 10mls pre-warmed RPMI media in a 25ml flask and incubated overnight at 37°C with 5% CO_2.

2.1.5 Cell counting

15 µl of cell suspension in DPBS was diluted 1:1 in trypan blue (HyClone) which stains dead cells, and added to a Neubauer haemocytometer by capillary uptake under a coverslip. Live cells in the 4 corner squares were counted. The average cell counts of the four squares were used to calculate the cell culture concentration using the formula:

Average cell count= Count x $1x10^4$ cells/ml

This was doubled to take account of the addition of trypan blue.

2.2 Biochemical reagents and methods

BSA (Sigma)

Dissolved in PBS and stored as 100mg/ml aliquots at -20°C.

Cell lysis buffer

85 mM KCl, 0.5% NP-40, 5 mM PIPES pH 8.0. Before use buffer was chilled on ice and made to 1 mM PMSF (Sigma). 1 EDTA-free protease inhibitor cocktail tablet (Roche)/10mls was added.

ECL solution I

125 µM Luminol (250 mM stock in DMSO), 20 µM coumaric acid (Sigma)(90 mM stock in DMSO) and 5 mM Tris pH 8.5 in 1 ml Milli-Q water

ECL solution II

0.0075% Hydrogen Peroxide (H_2O_2) and 5 mM Tris pH 8.5 in 1 ml Milli-Q water.

Elution Buffer

10 mM Tris pH 8.0, 5 mM EDTA, 1% SDS

Gel sample buffer

50 mM Tris, 4% SDS (w/v), 5% 2-Mercaptoethanol (Sigma), 10% Glycerol, 1 mM EDTA and 0.01% Bromophenol blue.

High salt buffer

0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.0, 500 mM NaCl. Chilled on ice before use.

Immunoprecipitation dilution buffer (IP)

0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris pH 8.0,167 mM NaCl. Chilled on ice before use and made to 1 mM PMSF. 1 EDTA-free protease inhibitor cocktail tablet /10 mls was added.

Lithium Chloride buffer (LiCL)

250 mM LiCl, 1% NP40, 1% Na deoxycholate, 1 mM EDTA, 10 mM Tris pH 8.0. Chilled on ice before use.

Low salt buffer

0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris pH 8.0, 150 mM NaCl. Chilled on ice before use.

PBS-Tween

100 PBS tablets (Oxoid) and 10mls Tween-20 made up in 10L dH_20.

Proteinase K (Sigma)

Dissolved in water and stored as 20mg/ml aliquots at -20°C

SDS lysis buffer

1% SDS, 10 mM EDTA, 50 mM Tris pH 8.0. Chilled on ice before use and made to 1 mM PMSF. 1 EDTA-free protease inhibitor cocktail tablet /10 mls was added.

TE buffer

10 mM Tris pH 8.0, 1 mM EDTA. Chilled on ice before use.

Western blot Transfer buffer

15g Tris, 72g Glycine, 4L dH_20 and 1L Methanol

2.2.1 Preparation of whole cell lysates

Cells were washed in PBS, counted and resuspended in 100 µl GSB/1x10^6 cells. Cells were sonicated on ice using the Vibra-Cell VC 750 sonicator (Sonics) for 7 pulses at 25% amplitude for 10 seconds with 10 second gaps. Samples were boiled at 95°C for 10 mins, vortexed, briefly centrifuged for 30 seconds and stored at -20°C or loaded directly onto SDS-PAGE gels.

2.2.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

10µl samples were loaded into pre-poured 4-12% Bis-Tris gel (Life Technologies) using a gel loading tip. Lysates were resolved using 1x MOPS running buffer (Life Technologies). 15 µl of SeeBlue Plus2 pre-stained standard marker (Life Technologies) was also loaded and electrophoresis carried out for 50 mins at 200 V.

2.2.3 Immunoblotting

After separation by SDS-PAGE, proteins were transferred onto Protran nitrocellulose membranes (Whatman) in transfer buffer at 85 V for 90 mins using a blotting cell (Bio-Rad). Membranes were incubated with Ponceau stain (Sigma) for one minute to examine loading and verify a successful transfer. Membranes were cut as required to probe for multiple proteins on one blot. Membranes were washed three times for 5 minutes in PBS-Tween on a shaker and blocked for 1hr using 2% (w/v) milk powder (Marvel) in PBS-Tween. Primary antibodies (Appendix B) were added to membranes in a 2% (w/v) milk PBS-Tween solution and incubated overnight at 4°C with rocking. After further washing with PBS-Tween (3x10 mins), secondary antibodies conjugated to a horse-radish peroxidase (HRP) enzyme (Appendix A) made up in 2% (w/v) milk PBS-Tween solution were added to membranes and incubated with rocking at room temperature for 1 hr. Final washing (3x10 mins) in PBS-Tween was performed and equal volumes (1ml each) of ECL solutions I & II were added to membranes and mixed. Membranes were imaged on a LiCor Odyssey imaging system or Membranes were exposed to Fuji medical X-ray film (Fisher) for varying time periods and developed using the Konica SRX-101A film processor.

2.2.4 Chromatin Immunoprecipitation (ChIP)

2.2.4.1 Bead preparation

2.2.4.1.1 Protein A sepharose bead preparation

Protein A sepharose beads (Sigma) in aqueous ethanol solution were washed 3 times in PBS. 500µl Protein A sepharose beads were blocked using 350 µg of single-stranded sonicated salmon testes DNA (Sigma) and were rotated for 1hr at 4°C. Beads were then washed 3 times in PBS and stored as a 1:1 slurry with IP dilution buffer at 4°C prior to use.

2.2.2.4.2 Protein A agarose bead preparation

Protein A agarose beads (Merck Millipore) in aqueous ethanol solution were washed 3 times in PBS. 400 µl Protein A agarose beads were blocked using 5mg of BSA diluted in PBS and rotated for 1hr at 4°C. Beads were then washed 3 times in PBS and stored as a 1:1 slurry with IP dilution buffer at 4°C prior to use.

2.2.4.2 Chromatin preparation

Cells were resuspended to $5x10^5$ cells/ml 24 hrs prior to chromatin preparation. Cells were then diluted to $1x10^7$cells/ml in RPMI 1640 media and crosslinked using 1% formaldehyde (Sigma) for 15 mins on a rocker at RT. Cross-linking was halted by incubation with 125 mM glycine for 10 mins at room temperature on a rocker. Cells were pelleted by centrifugation at 1300 rpm for 10 mins at 4°C in a Sorvall legend centrifuge and washed in PBS. Cell pellets were lysed on ice for 10 mins in 300µl cell lysis buffer per 10^7cells. Nuclei were pelleted at 8000 rpm for 5 mins at 4°C in accuspin benchtop centrifuge and resuspended in 200µl SDS lysis buffer per $1x10^7$cells mixing well. Chromatin was then sonicated using the Vibra-Cell VC 750 sonicator (Sonics) for 10 pulses at 30% amplitude for 10 seconds with 10 second gaps. and stored as aliquots at -80°C.

2.2.4.3 ChIP from cross-linked chromatin

100 µl ($5x10^6$ cells) of chromatin was diluted ten-fold in IP dilution buffer and pre-cleared by adding 45 µl of blocked protein A Sepharose or protein A agarose beads and rotating at 4°C for an hour. After removal of beads by pelleting at 4000rpm 4°C in an accuspin benchtop centrifuge for 1 minute, an input control (40 µl) was removed from the pre-cleared chromatin and stored at -20°C. The remaining chromatin was incubated with or without (no antibody control) primary antibodies (see Appendix A) overnight at 4°C with rotation. Secondary antibodies were added where required (Appendix A) and samples rotated for a further 2 hrs 4°C. Immune complexes were collected by rotating for 3 hrs at 4°C with 45 µl of 1:1 pre-

blocked protein A sepharose bead or protein A agarose bead slurry. Samples were then washed for 10 mins with rotation at 4°C in 1ml of each of the following wash buffers; 1x low salt, 1x high salt, 1x LiCl and 2x TE. Immune complexes were eluted in 150 µl elution buffer at 65°C for 20 mins, resuspending beads every 5 minutes by agitation. Beads were removed by pelleting at 4000rpm at 4°C in an accuspin benchtop centrifuge for 1 minute and DNA-protein immuno complexes were incubated at 65°C overnight to reverse crosslinking. Input controls were also incubated overnight in 150 µl elution buffer to reverse crosslinks. Proteinase K (50 µg) and TE buffer (150 µl) were added to samples and protein digested for 2 hrs at 50°C. DNA was purified using a QIAquick® Gel extraction kit (Qiagen) according to the manufacturer's instructions (QG and propan-2-ol were substituted for PB in most cases) and DNA then eluted in 110 µl sterile millipore water.

2.2.4.4 ChIP from non-cross linked chromatin

ChIP protocol was performed on chromatin that had been prepared without formaldehyde. After washing Immune complexes were eluted in 35 µl GSB and boiling. Beads were pelleted and supernatant loaded directly onto SDS-PAGE gels.

2.2.4.5 ChIP Sequencing DNA isolation and library preparation

600 µl ($30x10^6$ cells) of Mutu III chromatin was diluted ten-fold in IP dilution buffer and pre-cleared by adding 270 µl of BSA blocked protein A Sepharose beads and rotating at 4°C for an hour. After removal of beads by pelleting an input control (40 µl) was removed from the pre-cleared chromatin and stored at -20°C. The remaining chromatin was incubated with primary antibodies (see Appendix A) overnight at 4°C with rotation. Secondary antibodies were added where required (Appendix A) and samples rotated for a further 2 hrs 4°C. Immune complexes were collected by rotating for 3 hrs at 4°C with 270 µl of 1:1 pre-blocked protein A sepharose bead or protein A agarose bead slurry. Samples were then washed for 10 mins with rotation at 4°C in 6 ml of each of the following wash buffers; 1x low salt, 1x high salt, 1x LiCl and 2x TE. Immune complexes were eluted in 900 µl elution buffer at 65°C for 20 mins, resuspending beads every 5 minutes by agitation. Beads were removed by pelleting and DNA-protein immuno complexes were incubated at 65°C overnight to reverse crosslinking. Input controls were also incubated overnight in 150 µl elution buffer to reverse crosslinks. Samples were incubated with RNAse A (Sigma) at 37°C for 1 hour (30ng for input and 0.18 µg for IP samples). Proteinase K (50 µg for input and 300 µg for IP samples) and TE buffer (150 µl for input and 900 µl IP samples) were added to samples and protein digested for 2 hrs at 50°C.

Input DNA was purified using a QIAquick® Gel extraction kit (Qiagen) according to the manufacturer's instructions (replacing QG and propan-2-ol with 750 µl PB (Qiagen)) and eluted in 55 µl sterile millipore water. 4.5 mls of PB was added to IP samples and samples mixed by pipetting. 750 µl samples were added to three QIAGEN columns from the QIAquick® Gel extraction kit (Qiagen) and subjected to centrifugation at 13000rpm at 4°C in an accuspin benchtop centrifuge for 1 minute. Flow through was discarded and the process repeated until all of the IP samples were added to the three columns. Columns were washed according to the manufacturer's instructions. DNA was eluted by the addition of 55 µl sterile millipore water to one column, incubating at room temperature for 1 minute and subjected to subjected to centrifugation at 13000rpm at 4°C in an accuspin benchtop centrifuge for 1 minute. Flow through was then re-added to the same column and process repeated. After removing the first column, flow through was then added to the second column and process repeated such that each of the 3 columns undergoes 2 DNA elutions. Used columns each had 110 µl water to them and residual DNA eluted as above.

ChIP and input DNA (10 ng) were used to generate sequencing libraries with a ChIP-Seq sample preparation kit (Illumina). DNA fragments were end repaired and phosphorylated, 3'-dA overhangs were added, and adapters were ligated according to the manufacturer's instructions. PCR-amplified samples (16 cycles) were separated on a 2% agarose gel in Tris-acetate–EDTA (TAE) buffer and visualized using SYBR-safe stain and a Dark Reader transilluminator instrument (Clare Chemical Research). The region of the gel containing 150- to 350-bp DNA fragments was excised, and DNA was purified using a min-elute gel extraction kit (Qiagen). The library generated was quantified using an Agilent bioanalzer instrument and subjected to 35-bp single-end read sequencing with an Illumina IIx genome analyzer.

2.2.4.6 ChIP-reChIP

600 µl ($30x10^6$ cells) of Mutu III chromatin was diluted ten-fold in IP dilution buffer and pre-cleared by adding 270 µl of BSA blocked protein A Sepharose beads and rotating at 4°C for an hour. After removal of beads by pelleting an input control (40 µl) was removed from the pre-cleared chromatin and stored at -20°C. The remaining chromatin was incubated with primary antibodies (see Appendix A) overnight at 4°C with rotation. Secondary antibodies were added where required (Appendix A) and samples rotated for a further 2 hrs 4°C. Immune complexes were collected by rotating for 3 hrs at 4°C with 270 µl of 1:1 pre-blocked protein A sepharose

bead or protein A agarose bead slurry. Samples were then washed for 10 mins with rotation at 4°C in 6 ml of each of the following wash buffers; 1x low salt, 1x high salt, 1x LiCl and 2x TE. Immune complexes were eluted in 900 µl elution buffer at 65°C for 20 mins, resuspending beads every 5 minutes by agitation. Beads were removed by pelleting and each IP split into 6 aliquots each of 150 µl.

With one aliquot the ChIP was continued as described. On the 5 other aliquots a fresh ChIP was performed, using the protocol described in 2.3.4 with the exception that instead of diluting 100 µl of chromatin in 900 µl IP dilution buffer, 150 µl of eluted sample was diluted in 850 µl IP dilution buffer.

2.2.5 RNA isolation

RNA was extracted from cells using TRI Reagent (Sigma) and purified using the RNeasy Mini Kit (Qiagen) according to the manufacturer's guidelines.

2.2.6 Determining RNA concentration

Samples were diluted 1 in 50 (2µl in 98µl sterile Millipore water) and quantified using the eppendorf BioPhotometer using a UV-Cuvette micro (Brand). RNA concentration (µg/ml) was calculated using the formula A_{260} x 40µg/ml x dilution factor (50). The purity was measured by the ratio of the two absorbance's between 260nm and 280nm (1.8 to 2 for pure RNA).

2.2.7 cDNA synthesis

cDNA was synthesised from purified RNA using the ImProm-II Reverse Transcription system (Promega) based on the manufacturers specifications. 1µg of RNA was incubated at 70°C for 5 mins with 0.5µg of random primer and Millipore sterile water to a final volume of 5µl. Samples were left on ice for 5 mins and spun in the Techico mini microfuge.

2.2.8 Q-PCR

Quantitative PCR (QPCR) was performed using an Applied Biosystems step one plus real-time PCR machine. For ChIP analysis, 3µl DNA was added to a SYBR green master mix containing 7.5µl 2xGoTaq QPCR master mix (Promega), 150nM forward and 150nM reverse primers (Appendix B) and sterile Millipore water to a final volume of 15µl. Samples were heated to 95°C for 10 mins, followed by 40 cycles at 95°C for 15 sec and 60°C for 1 min and dissociation curve analysis. Input controls were serially diluted to generate a standard curve for each

primer set (Appendix B). A percentage input value was obtained by measuring the crossing threshold (C_t) value in relation to each standard.

For cDNA analysis, samples were analysed using the absolute quantification method. Standard curves were generated from cDNA from either BL31, EREB 2.5 +est 17hrs, D2mtB1 LCL or PER142 3BKO LCL. Transcript levels were determined using cDNA-specific primers for CTBP2, ITGAL or WEE1 (Appendix B) and values obtained were made relative to the signal from GAPDH cDNA specific primers.

2.2.9 Determination of DNA fragment size after sonication

Mutu III chromatin was prepared as normal. After resuspension of nuclear pellet in SDS lysis buffer 1/3rd of the chromatin was removed and stored at -80°C. Remaining chromatin was then sonicated using the Vibra-Cell VC 750 sonicator (Sonics) for 10 pulses at 30% amplitude for 10 seconds with 10 second gaps at 4°C. This aliquot was split into 2 and one half frozen. Remaining half was subjected to an additional 10 pulses of sonication and stored at -80°C. At time of use for ChIP a 5 µl aliquot of each was taken and diluted in 92.5 µl Tris pH7.5. 50 µg of proteinase K was added and diluted chromatin incubated at 65°C. 20 µl aliquots with 5 µl DNA loading dye were loaded into a 1% TBE agarose gel alongside a 1kb DNA ladder and subjected to gel electrophoresis at 90V for 1 hr. Gel was imaged on a LiCor Odyssey imaging system.

2.3 Bioinformatics analysis

2.3.1 Comparison of ChIP sequencing and gene expression microarray data

Lists of genes bound in ChIP sequencing and genes regulated in gene expression microarray data sets were assembled and analysed in Microsoft EXCEL. Gene expression microarray gene nonclamenture were converted to official gene symbols using the Database for Annotation Visualization and Integrated Discovery (DAVID) gene ID conversion tool. These were then compared to lists of genes either from other gene expression microarray data sets or lists of genes bound in ChIP sequencing experiments provided by Dr. Aditi Kanhere (University College London). Lists were compared by using a COUNTIF function (=COUNTIF(range,criteria)) where any cell in the range (eg all official gene symbols of upregulated genes in a dataset) was analysed to see if matched the criteria (eg an official gene symbol from a gene shown to be bound by ChIP sequencing data).

2.3.2 ChIP sequencing data analysis (performed by Dr Aditi Kanhere and Dr Richard Jenner, University College London)

ChIP sequencing data from our lab and from the Kieff lab for EBNA 2 and RBPJκ ChIP sequencing experiments in IB4 cells were processed and analysed, Initial processing of sequencing images generated by the Illumina IIx genome analyzer was performed using the CASAVA pipeline. Polony identification, base calling, and quality control statistics were performed using GOAT and Bustard modules. Thirty-six-bp short reads were aligned to the hg19 build of the human genome using ELAND software. Only reads with zero to two mismatches were aligned, and only uniquely aligned reads passing the quality threshold were retained. When multiple reads matched to the same position, only a single read was considered for further analysis. Sequence reads were extended to 150 bp. WIG files were generated by calculating tag density in 10-bp windows and were normalized to reads per million total reads using in-house R scripts. The data for the EBNA 2 and EBNA 3 sequencing run were then background corrected using data from input DNA. Significant peaks of EBNA 2 and EBNA 3 binding were identified with MACS ($P < 10^{-7}$) (Zhang et al. 2008). Encyclopedia of DNA elements (ENCODE) ChIP-Seq data for transcription factors and histone modifications in the EBV-immortalized LCL GM12878 were accessed through the human genome browser at http://genome.ucsc.edu/cgi-bin/hgGateway and compared with our data. RBPJκ data from a recently published ChIP sequencing experiment was compared sites of significant EBNA 2 and EBNA 3 binding (Zhao et al. 2011).

2.3.3 Motif searching

Motif searching was performed by generation of FASTA files from ChIP sequencing MACS peaks data and used as an input to MEME http://meme.nbcr.net/meme/cgi-bin/meme-chip.cgi. Default settings were used (normal mode, zero or one occurrence per sequence).

2.3.4 DAVID gene ontology analysis

Official gene symbols were submitted to DAVID as a gene list in the background of homo sapiens. Functional annotation clustering selecting only GOTERM-BP (biological process) or Kegg pathway analysis was performed with default significance settings http://david.abcc.ncifcrf.gov/.

3. Mapping EBNA 2 and EBNA 3C binding in the human genome

At the start of this project most information on the regulation of transcription by EBNA 2 and 3 proteins came from studies of EBV promoters. EBNA 2 was known to activate the C, *LMP1/2B* and *LMP2A* promoters through RBP-J sites. PU.1 sites in the *LMP1* promoter were also shown to be required for full activation by EBNA 2. EBNA 3 proteins were documented as competitive binders to RBP-Jκ, inhibiting EBNA 2 activation in reporter assays, although independent repressive effects of EBNA 3C on Cp and coactivation of *LMP1* transcription with EBNA 2 via the PU.1 site were also reported. Direct interactions between EBNA 2 and EBNA 3C and PU.1 were also reported. In contrast, no EBNA 3 binding sites in the human genome were known and information on EBNA 2 activation of cellular genes was restricted to a few targets where RBP-Jκ sites could be identified in the target gene promoters and shown to confer responsiveness in promoter-reporter assays. However, many known regulated genes were not responsive to EBNA 2, EBNA 3A, EBNA 3B or EBNA 3C in these assays. In order to gain important insights into the mechanism of cellular gene regulation by these key EBNAs, Chromatin immunoprecipitation coupled to next generation sequencing was performed to map the genome-wide binding profiles of EBNA 2 and EBNA 3C, the best characterised of the EBNA 3 proteins.

3.1 Validation of EBNA 3C antibody for ChIP

EBNA 2 binding to Cp, LMP1 and LMP2A had already been confirmed by ChIP using the EBNA 2 mouse monoclonal antibody PE2 by others in our laboratory (Palermo et al. 2011). The -310 primer set amplifies DNA containing two RBPJκ binding sites at -298 to -290 and -223 to -212. This site had also previously been reported as an EBNA 3C binding site and was the only site known in the viral or human genome where EBNA 3C bound (Jimenez-Ramirez et al. 2006). This study used a sheep polyclonal anti-EBNA 3C antibody (exalpha) which was not validated for use in ChIP. We had previously used an anti-EBNA 3C sheep polyclonal antibody (abcam) for immunoprecipitations and western blotting. To examine if this antibody could successfully precipitate DNA in ChIP conditions, and to confirm EBNA 2 binding to *LMP1* in other EBV immortalised cell lines, ChIP was conducted using PE2 and the anti-EBNA 3C sheep polyclonal antibody (abcam) in Mutu III BL cells and PER253 LCL, both expressing all latency III gene products. ChIP DNA was analysed using primers at the *LMP1* locus including sites previously identified as EBNA 2 binding sites in Mutu III cell lines. We were able to confirm EBNA 2 association with *LMP1* in Mutu III (Fig 3-1 C), and showed that EBNA 2 also binds to the -310 region in the PER253 LCLs (Fig 3-1 A). Importantly, QPCR analysis of DNA precipitated by the

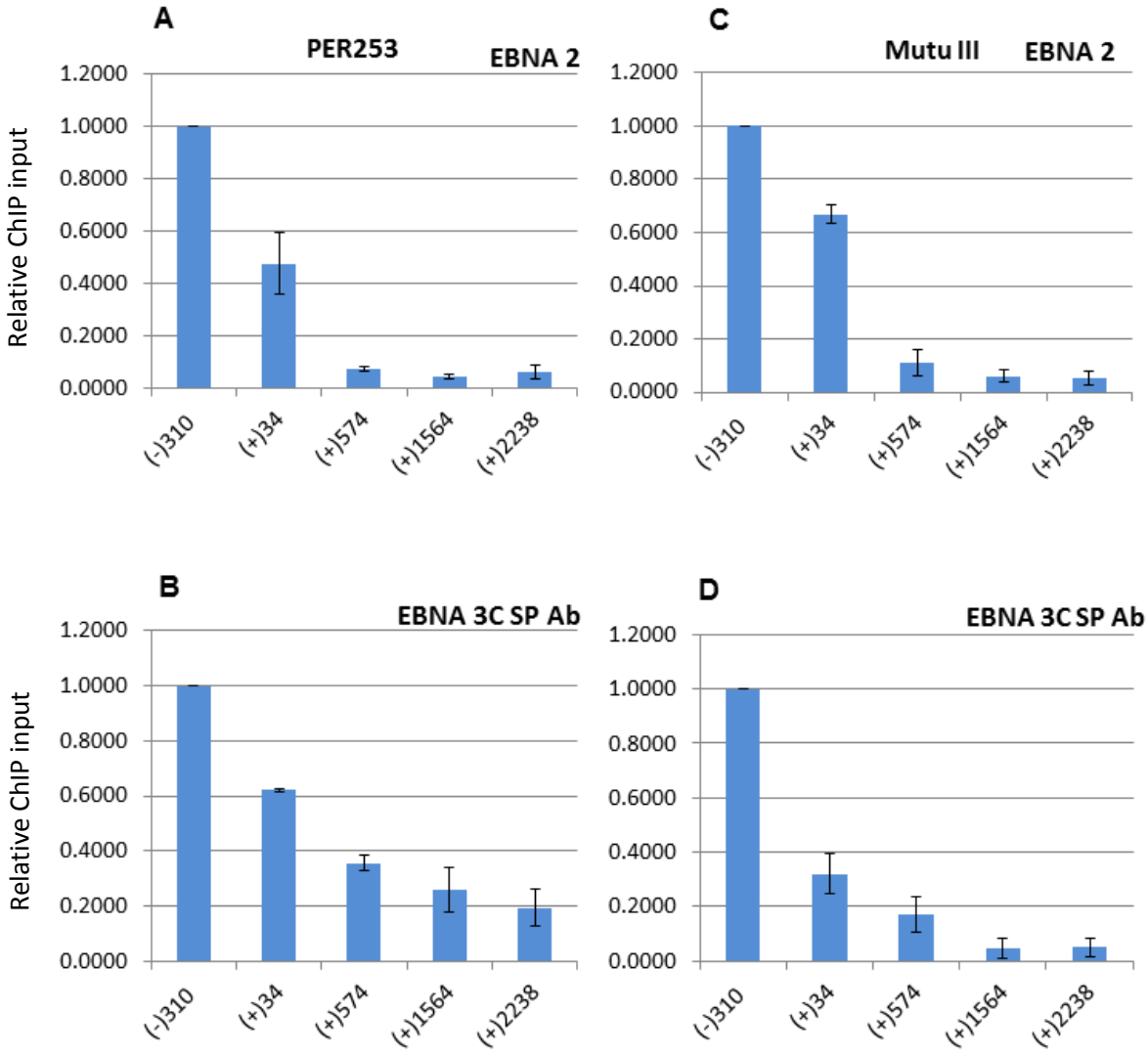


Figure 3-1 EBNA 3C sheep polyclonal antibody ChIP precipitated DNA is enriched at *LMP1* regulatory elements. ChIP experiments conducted in PER253 LCL and Mutu II cell lines. Location of primer sets used are indicated as distance from the 5' end of the forward primer to the *LMP1* TSS. Percentage input signals, after subtraction of signals of controls with no antibody, are expressed relative to the highest signal obtained. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 2 in PER253 (A), EBNA 3C sheep polyclonal antibody in PER253 (B), EBNA 2 in Mutu III (C), EBNA 3C sheep polyclonal antibody in Mutu III (D).

EBNA 3C sheep polyclonal antibody revealed increased quantities of DNA using the -310 and +34 primer sets relative to those within the *LMP1* gene in both PER253 (Fig 3-1 B) and Mutu III cell lines (Fig 3-1 D). These data confirm that the EBNA 3C antibody used can precipitate DNA in ChIP conditions.

3.2 Optimisation of ChIP sequencing protocol

ChIP sequencing required high yields of good quality ChIP DNA eluted in a small volume. Before conducting this experiment, optimisation and adaptation of the current ChIP protocol was required. After cross linking with formaldehyde, cells are lysed and then subjected to sonication to randomly break DNA into small fragments. ChIP sequencing involves a size selection step where DNA between 150-300 bp is selected before being sequenced. If DNA fragments precipitated in a ChIP are too large it will result in a reduced yield after this stage as DNA will be removed. To determine if yield could be improved by altering the pulses of sonication that chromatin is subjected to, Mutu III chromatin, either un-sonicated or subjected to 10 and 20 pulses was resolved using agarose gel electrophoresis. This revealed that DNA fragment size is reduced by increasing rounds of sonication (Fig 3-2 A). We also used the same chromatin in ChIP assays using antibodies against EBNA 2 and PU.1 and quantified precipitated DNA at the *LMP1* locus where both are known to bind. These data reveal no significant changes in the ratio of bound and unbound sites but a loss of input DNA precipitated was observed with 20 pulses of sonication compared to 10 (Fig 3-2 B vs C and Fig 3-2 D vs E). In EBNA 2 precipitated DNA samples there was ~15% less DNA amplified by the -310 primer set in the 20 pulsed aliquot compared to the 10 pulsed aliquot. From this data it was decided to use 10 rounds of sonication in future experiments.

The ChIP protocol for QPCR includes a blocking step where beads are incubated with salmon sperm DNA. This step is required for reducing non-specific binding of DNA and proteins in the cell lysate to the beads. However, if salmon sperm DNA is eluted from the beads at the elution step of ChIP then this DNA may contaminate our sample for sequencing. This would reduce sequencing reads as a proportion of the DNA submitted for sequencing would not map to the human genome. To test whether salmon sperm DNA would be eluted from the beads at the end of the ChIP experiment, we performed ChIP in Mutu III cells for EBNA 2 using beads blocked with either salmon sperm DNA or BSA. We designed primers at the salmon insulin gene to amplify a site that has no homology to human insulin. QPCR on DNA precipitated from beads blocked with salmon sperm DNA or BSA confirmed that salmon insulin DNA was enriched in EBNA 2 ChIP DNA blocked with salmon sperm compared to EBNA 2 ChIP DNA

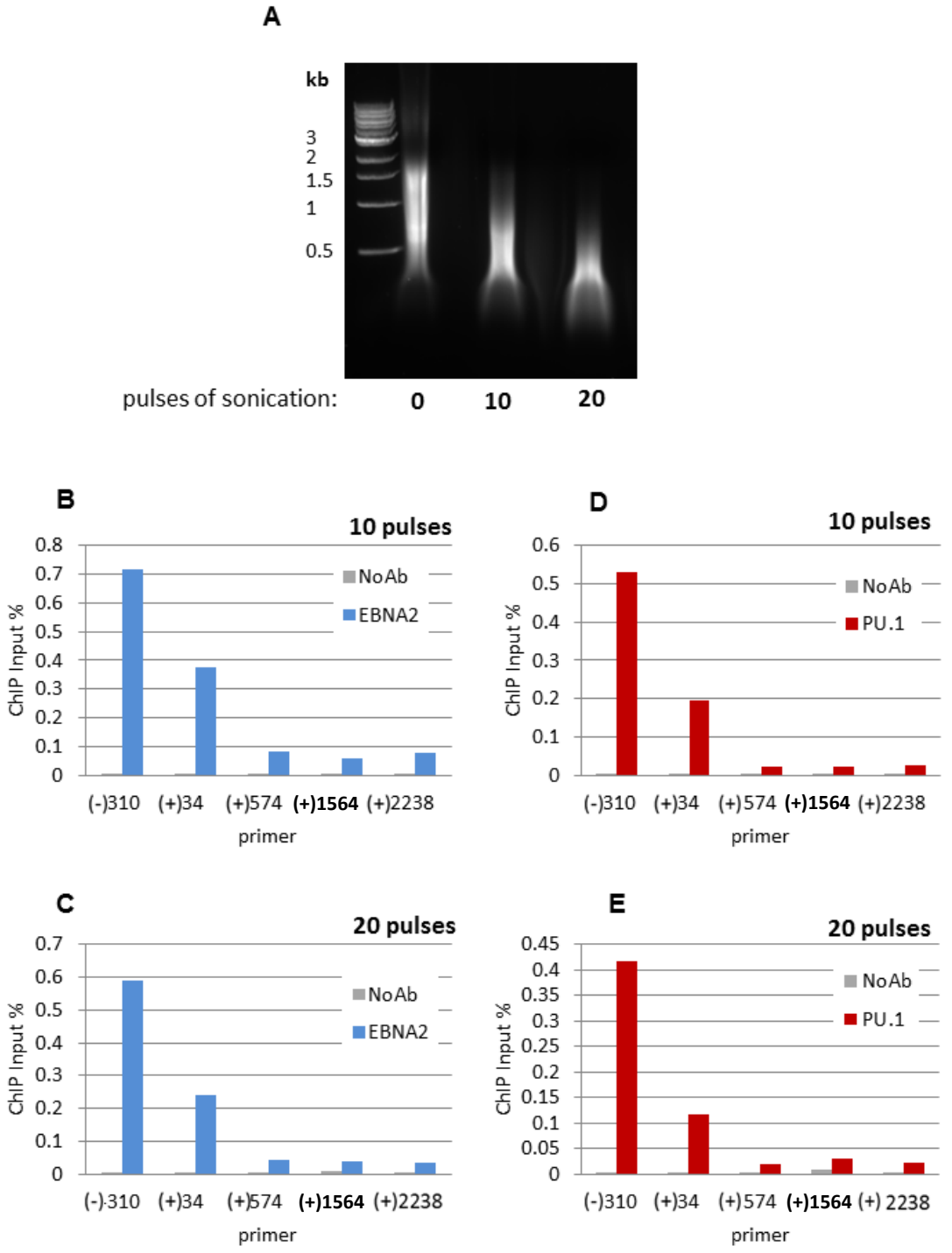


Figure 3-2 Optimisation of sonication step of chromatin preparation Mutu III chromatin either un-sonicated or sonicated for 10 or 20 pulses was examined. (A) Size of DNA fragments sonicated 0, 10 or 20 times. Leftmost lane was loaded with a 1kb ladder. ChIP experiments performed on chromatin sonicated with 10 or 20 pulses using antibodies for EBNA 2 and PU.1. Location of primer sets used are indicated as distance from the 5' end of the forward primer to the *LMP1* TSS. Signal for no antibody are displayed. ChIP QPCR for EBNA 2 with 10 pulses (B), EBNA 2 with 20 pulses (C), PU.1 with 10 pulses (D) and PU.1 with 20 pulses (E).

blocked with BSA (Fig 3-3 A). To test if this enrichment was significant or at background levels, we also used primer sets at the *LMP1* locus to analyse salmon sperm and BSA blocked EBNA 2 ChIP DNA. Data was plotted alongside the percentage ChIP input obtained using salmon insulin primer sets and the no antibody negative control. These data show that in EBNA 2 ChIP DNA salmon DNA is present at a level significantly higher than that of the no antibody and is equivalent to EBNA 2 background binding in the *LMP1* gene (Fig 3-3 B). No salmon DNA was detected in BSA blocked beads, and BSA blocked beads were able to precipitate EBNA 2 as shown by enrichment of EBNA 2 precipitated DNA at the -310 and +34 primer set of *LMP1* (Fig 3-3 C). Due to this finding we subsequently blocked all beads in BSA.

3.3 ChIP sequencing library preparation

A large-scale ChIP protocol was designed and performed on Mutu III chromatin using EBNA 2 specific and anti-EBNA 3C sheep polyclonal antibodies. This used six times more starting chromatin per antibody ($3x10^7$ cells) and DNA was diluted in half of the standard volume of water creating a 12-fold concentrated DNA sample. This was not necessary for the input DNA as this is already sufficiently concentrated in a standard ChIP experiment. After elution of DNA from DNA purification columns an extra elution step was carried out using the standard volume of water in a ChIP experiment. This was performed in order to test if the large scale ChIP method did not overload the DNA purification columns and leave large quantities of residual DNA. An aliquot of the concentrated ChIP samples was diluted 12 fold and analysed alongside residual DNA by QPCR using primers at the *LMP1* locus. These data show that the large scale ChIP successfully enriched DNA bound by EBNA 2 (Fig 3-4 A) and EBNA 3C sheep polyclonal antibody (Fig 3-4 B) as a higher signal was obtained using the -310 and +34 primer set than the +574 primer set. It also demonstrates that residual DNA from EBNA 2 (Fig 3-4 A) and EBNA 3 (Fig 3-4 B) was detectable from the second elution step. Given that the concentrated DNA sample was diluted 12-fold, residual DNA accounts for ~ 8-15% of the total DNA precipitated by each antibody (Fig 3-4 A, B).

ChIP sequencing library preparation was carried out using an Illumina kit but on advice of our collaborator, Dr. Richard Jenner (University College London), we altered the protocol to conduct PCR enrichment of DNA prior to gel extraction and size selection. After library amplification by PCR, DNA samples were size selected by agarose gel electrophoresis and the 150-350bp fragment was purified. As QG buffer was necessary for gel extraction it was necessary to re-purify DNA as guanidium salts in the QG buffer absorb strongly at 230nm.

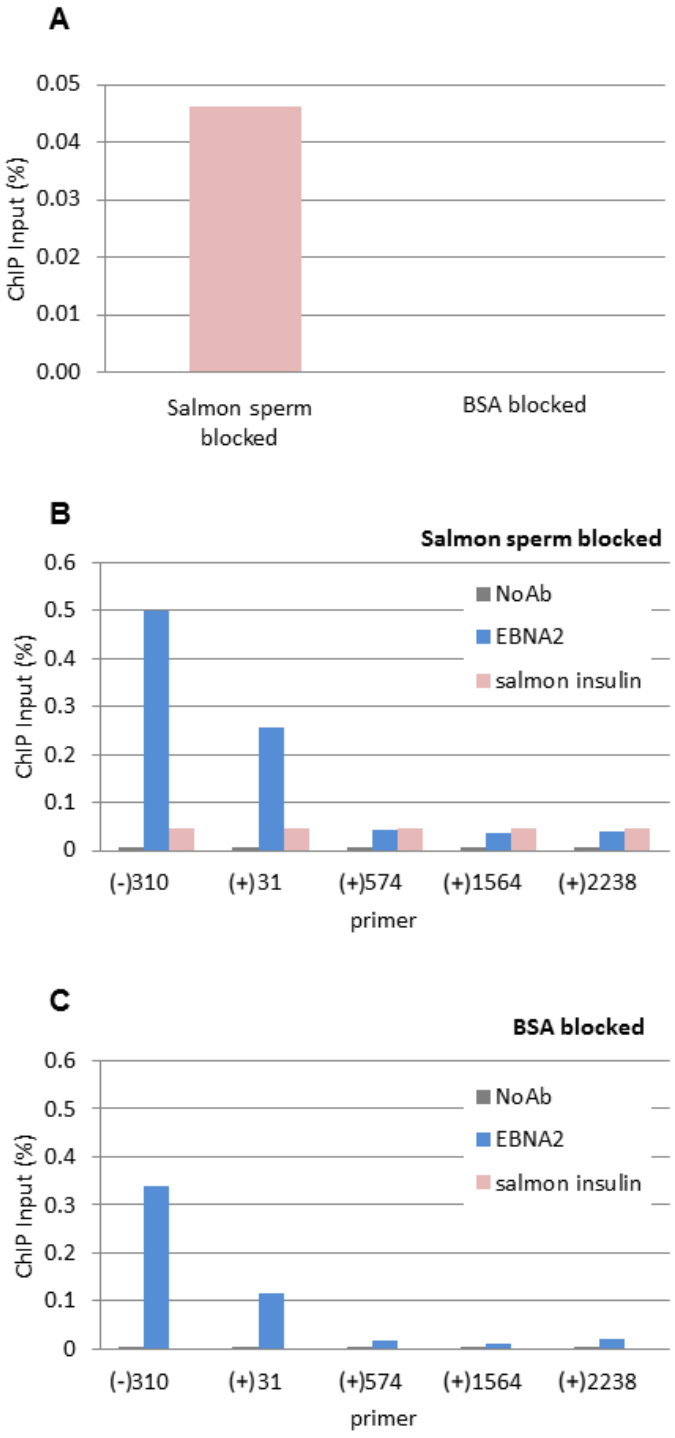


Figure 3-3 Salmon DNA is eluted from salmon sperm blocked beads in ChIP conditions. (A) ChIP QPCR from Mutu III chromatin using salmon sperm and BSA blocked beads and EBNA 2 antibody. Primers were designed for a region of the salmon insulin gene with no homology to humans. (B) ChIP in Mutu III using EBNA 2 antibody and salmon sperm blocked beads (B) and BSA blocked beads (C) analysed by QPCR at *LMP1*. Location of primer sets used are indicated as distance from the 5' end of the forward primer to the *LMP1* TSS. Signal for no antibody and signal obtained from salmon insulin primers are shown.

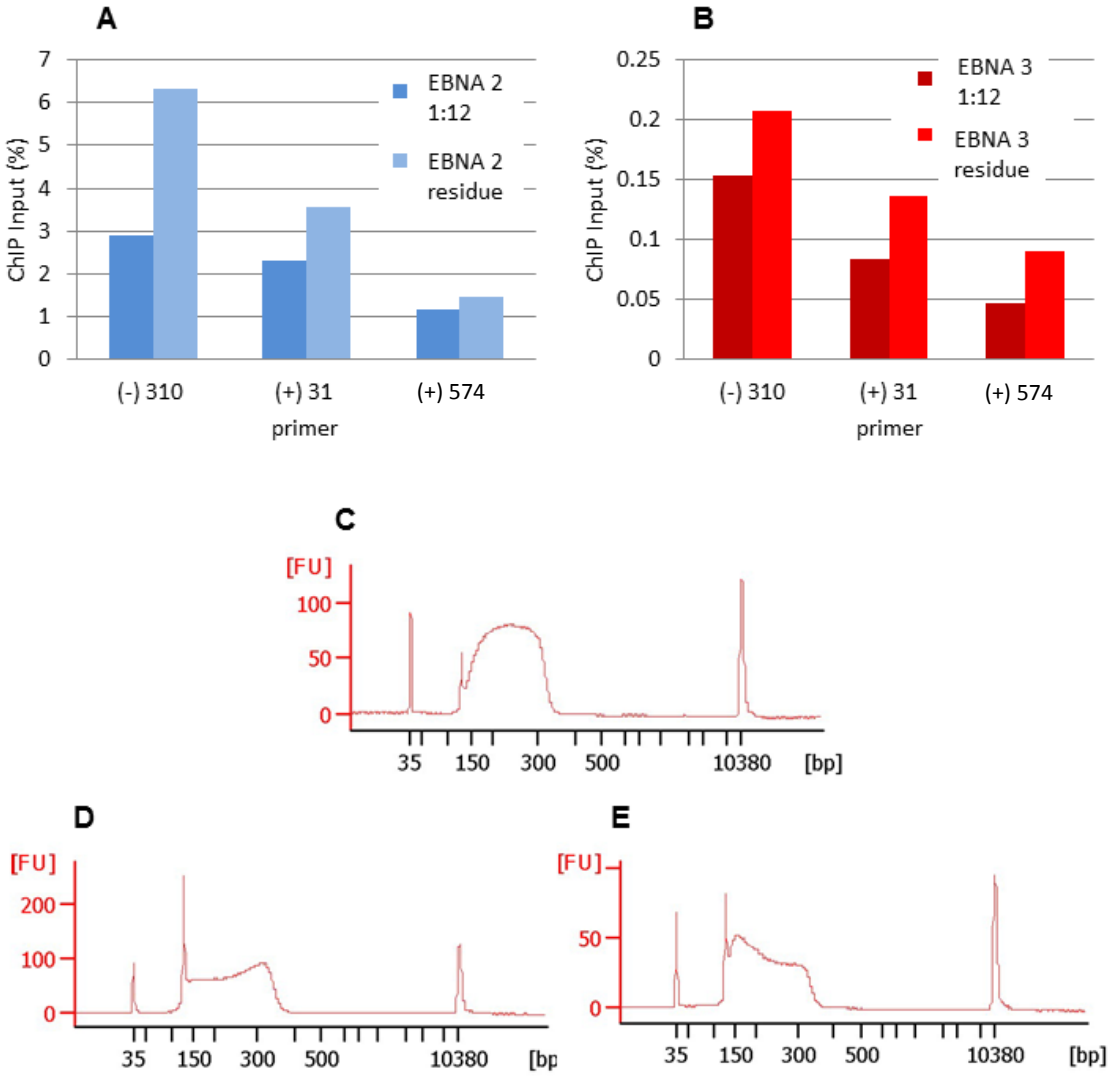


Figure 3-4 Quality control of ChIP DNA. EBNA 2 (A) and EBNA 3 (B) ChIP QPCR from Mutu III chromatin using *LMP1* primers. DNA eluted from QIAGEN column was diluted 1:12 and residual DNA left on the column was eluted in water. Bioanalyser data for input (C), EBNA 2 (D) and EBNA 3 (E) ChIP experiments. X axis displays fluorescent units used to measure quantity of DNA and Y axis displays the fragment sizes is base pairs. Markers at 35bp and 10380bp are also displayed.

Samples were then examined on a bioanalyser. The plots show the size of the purified DNA on the x-axis, and the quantity of DNA measured in fluorescent units on the y-axis for input DNA (Fig 3-4 C), EBNA 2 precipitated DNA (Fig 3-4 D) and EBNA 3C sheep polyclonal antibody precipitated DNA (Fig 3-4 E). These were analysed alongside a DNA ladder of known DNA concentration and sized fragments allowing quantification of library prepared DNA. Having shown that ChIP DNA had been successfully purified, samples were submitted for illumina next generation sequencing at UCL genomics.

3.4 EBNA 3C sheep polyclonal antibody co-precipitates EBNA 3A and EBNA 3B

After submitting samples for sequencing we received a personal communication from a member of the Allday lab suggesting that the EBNA 3C sheep polyclonal antibody co-precipitates EBNA 3A and EBNA 3B. To determine the specificity of the EBNA 3C antibody we performed immunoprecipitations (IPs) on non-crosslinked chromatin from the BJAB stable transfectants E3A1, E3B2 and E3C3; EBV negative B cell lymphomas which express only EBNA 3A, EBNA 3B and EBNA 3C respectively. IPs were conducted using anti-EBNA 3A, anti-EBNA 3B, EBNA 3C sheep polyclonal antibodies and IgG as a negative control. Western blotting analysis of whole cell lysates from these cells confirmed that each cell line expresses the correct EBV nuclear antigen (Fig 3-5 left column). Probing blots with anti-EBNA 3C antibody showed that EBNA 3C was successfully immunoprecipitated by the EBNA 3C sheep polyclonal antibody. However, blots probed with either anti-EBNA 3A or anti-EBNA3B antibodies showed that these proteins were also precipitated by the EBNA 3C SP antibody (Fig 3-5 right). This effect was specific as blots probed for TATA box binding protein (TBP) showed that this was not precipitated by the EBNA 3C antibody (Fig 3-5 bottom panel). This blot also shows that the EBNA 3A and EBNA 3B antibodies used did not cross react with the other EBNA 3 proteins as only the correct antigen was precipitated in each IP (Fig 3-5 lanes 1 and 3 of IPs). From these data we concluded that the EBNA 3C antibody was in fact pan specific to all EBNA 3 proteins, as it could precipitate EBNA 3A and EBNA 3B, albeit at reduced levels to EBNA 3C when compared to the amount of protein in the whole cell lysate.

3.5 Distribution of EBNA 2 and EBNA 3 binding sites

This analysis identified 21,605 EBNA 2 and 7,044 EBNA 3 binding sites in the human genome. The data for EBNA 2 and EBNA 3 sequencing runs were background corrected using data from input DNA and aligned with the human genome. Significant peaks of EBNA 2 and EBNA 3

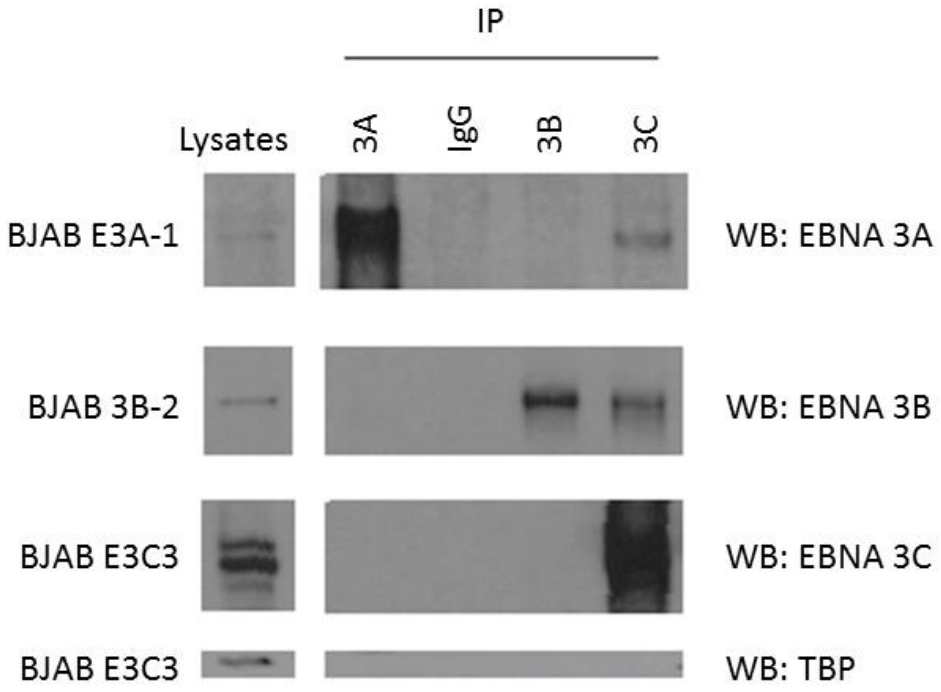


Figure 3-5 EBNA 3C sheep polyclonal antibody co-precipitates EBNA 3A and EBNA 3B under ChIP conditions. Immunoprecipitations performed under ChIP conditions. IPs were performed with anti-EBNA 3A, anti-EBNA 3B, anti-EBNA 3C sheep polyclonal antibodies and IgG as a negative control. IPs were performed in E3A1, E3B2 and E3C3 which express only EBNA 3A, EBNA 3B and EBNA 3C respectively. Precipitates were subjected to western blotting (WB) using anti-EBNA 3A, anti-EBNA 3B and anti-EBNA 3C antibodies. Leftmost column shows whole cell lysates. Top panel: IPs in E3A1 probed for EBNA 3A, second panel: IPs in E3B2 probed for EBNA 3B, third panel: IPs in E3C3 probed for EBNA 3C, bottom panel: IPs in E3C3 probed for TBP.

binding were identified with MACS (P-value < 10^{-7})(Zhang et al. 2008). By analysing the distribution of significant EBNA 2 and EBNA 3 binding sites we found that 75% of EBNA 2 (Fig 3-6 A), and 84% (Fig 3-6 B) of EBNA 3 binding sites were more than 4Kb away from any gene TSS, suggesting these proteins primarily target enhancer elements. Further analysis examined the distribution of binding sites from any TSS, which revealed that EBNA 2 had a bimodal distribution with high frequencies of binding sites both 0.2-0.5 kb and 20-50kb from any TSS (Fig 3-6 C). Interestingly EBNA 3 peaks did not share the same pattern of binding sites and predominantly targeted regions 20-50kb from any gene TSS (Fig 3-6 C). This analysis was gene centric and only the nearest peak to a gene was recorded. For each significant EBNA 2 binding site the average number of sequencing reads per million was analysed over a 2 kb window, centred on the centre of each binding site (Fig 3-6 D). The same analysis was also performed on EBNA 3 binding sites (Fig 3-6 E). This revealed that EBNA 2 and EBNA 3 binding profiles around the centre of each peak are similar. Together this data reveals extensive association of EBNA 2 and EBNA 3 proteins via distinct binding sites primarily found at distal regions. The abundance of EBNA 2 binding sites at 0.2-0.5 kb from gene TSSs may reflect the known role of EBNA 2 recruitment of general TFs to promoter proximal domains, mediated by the EBNA 2 TAD. This data also explains why known EBNA 2 and EBNA 3 regulated genes were not often responsive to promoter reporter assays as EBNA 2 and EBNA 3 binding sites are not commonly promoter proximal. We were also able to map our sequencing reads to the EBV genome. Reassuringly, this confirmed binding sites at Cp, *LMP1* and *LMP2A* for EBNA 2, and *LMP1* for EBNA 3 (Fig 3-7 A, B). EBNA 3 association with *LMP2A* is a novel finding.

To identify cellular genes targeted by the binding of EBNA 2 and 3 proteins, each binding site was analysed and the closest gene was termed as bound by that site. We first analysed genes that were bound by promoter proximal peaks (<2kb from a TSS). This revealed that 3554 genes were bound at a promoter proximal element by EBNA 2 and 664 were bound by EBNA 3 proteins. Of the EBNA 3 targeted genes 62% (412/664) were also bound by EBNA 2 (Fig 3-8 A). we then expanded this analysis to examine genes bound at any distance from a TSS. Remarkably this revealed that 80% (3157/3937) of genes targeted by EBNA 3 proteins were also targeted by EBNA 2 (Fig 3-8 B). These data suggests a high degree of cross talk between the EBNA 2 and 3 proteins and that they likely co-regulate a subset of cellular genes. EBNA 2 and EBNA 3 cross talk has already been implicated by gene expression microarray data but these data suggest the overlap in gene regulation is even more pronounced than expected (Hertle et al. 2009; White et al. 2010). We next analysed whether genes were being targeted

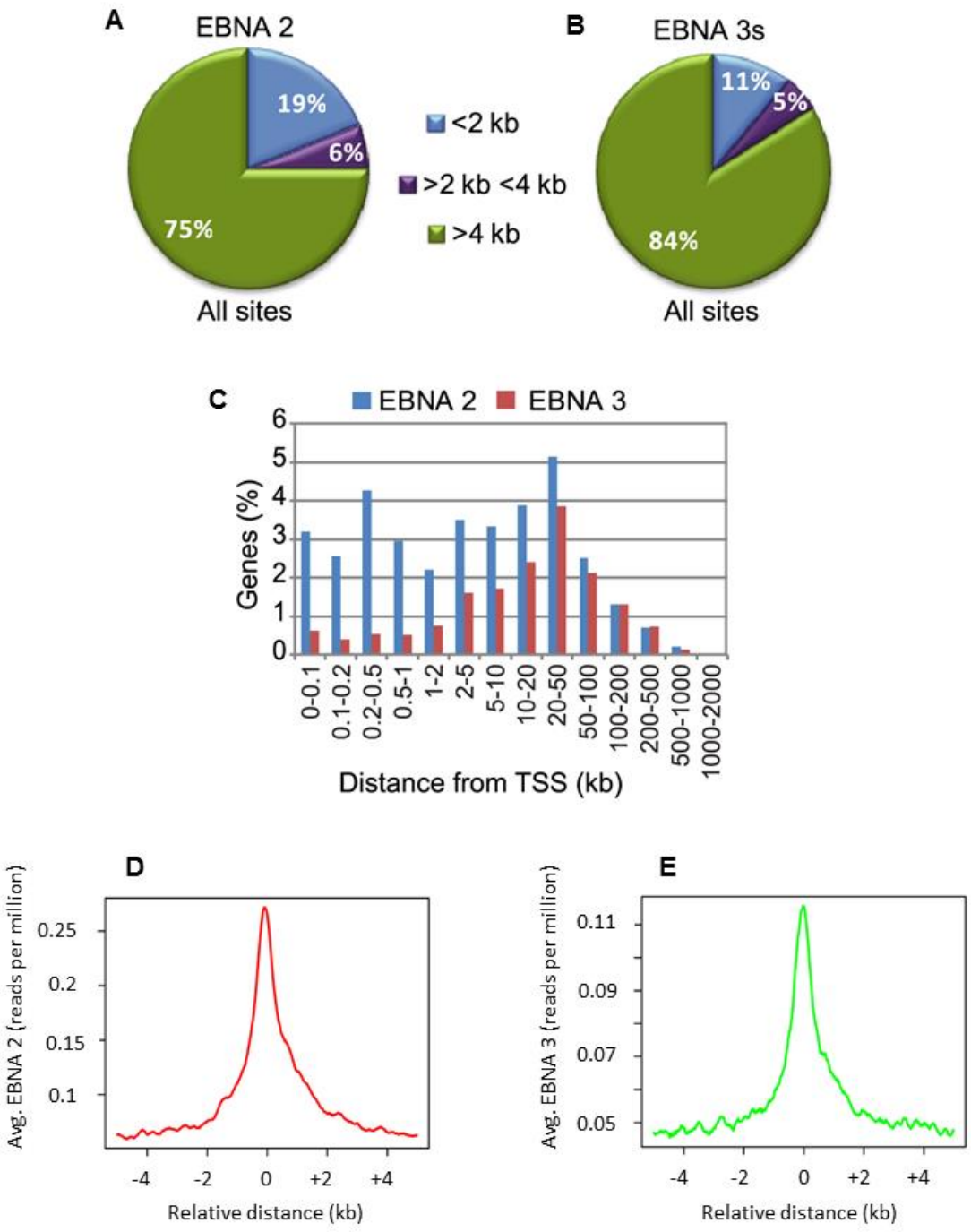


Figure 3-6 Distribution of EBNA 2 and EBNA 3 binding site (A) Pie chart showing the distribution of all significant binding sites for EBNA 2 relative to any gene TSSs. (B) Distribution of EBNA 3 family binding sites. (C) The frequency of EBNA 2 or EBNA 3 protein binding sites plotted as distance from the TSS of the closest gene. (D) average signal from every EBNA 2 binding site, centred on the middle of the binding site. (E) average signal from every EBNA 3 binding site.

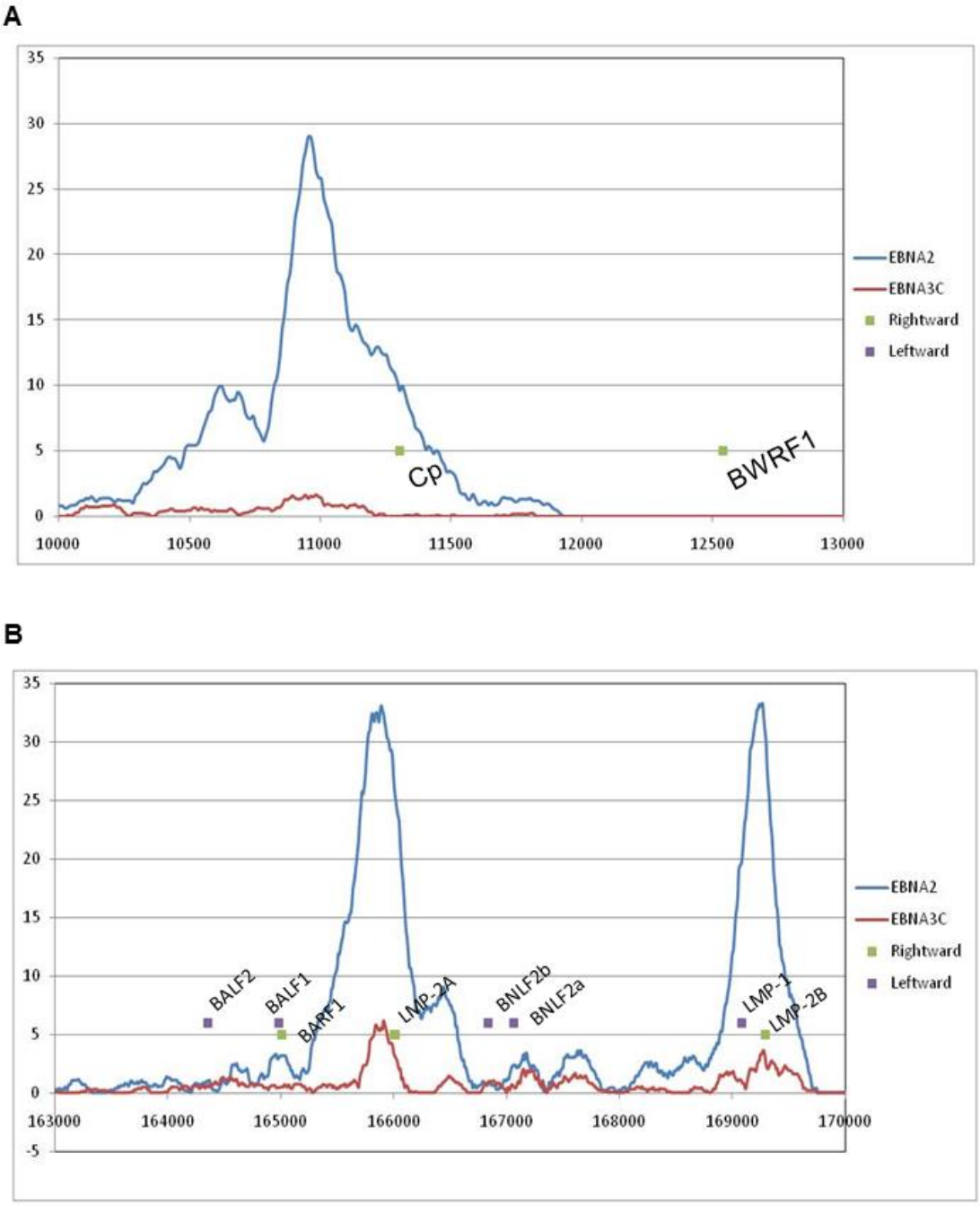


Figure 3-7 EBNA 2 and EBNA 3 proteins map to known binding sites in the viral genome (A) alignment of EBNA 2 and EBNA 3 sequencing reads to the EBV virus genome at Cp. Viral genes are annotated along with direction of transcription of each gene. X-axis shows distance from OriP. (B) alignment of EBNA 2 and EBNA 3 sequencing reads to the viral genome at LMP1 and LMP2A.

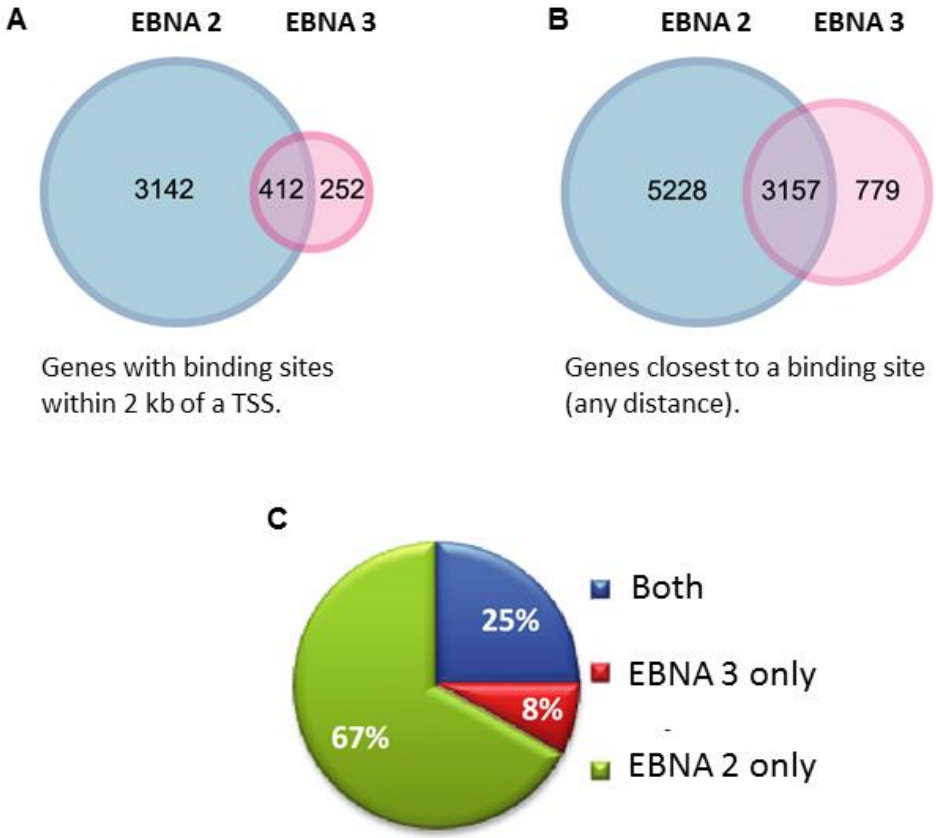


Figure 3-8 Coincident binding of EBNA 2 and 3 in the human genome. (A) Venn diagram of all genes with a promoter proximal EBNA 2 or EBNA 3 peak with the overlap showing genes bound by both. (B) Venn diagram of all genes most proximal to any peak for EBNA 2 or 3.(C) Pie chart showing the proportion of sites identified for EBNA 2 and EBNA 3 family proteins that are shared or unique.

by individual EBNA 2 and EBNA 3 sites or if these proteins were binding to shared sites. We found that only 8% of all EBNA 2 and 3 binding sites in the human genome were bound by EBNA 3 proteins alone and 25% of sites were bound by both EBNA 2 and EBNA 3 proteins (Fig 3-8 C). This analysis was conducted using a significance 'not bound' for potentially shared peaks. Crucially, this discovery shows that not only do EBNA 2 and EBNA 3 proteins target overlapping sets of cellular genes, but that they do so via the same enhancer elements.

3.6 Colocalisation of EBNA 2 and EBNA 3 binding sites with histone modifications

The Encyclopedia of DNA elements (ENCODE) project provides freely-available ChIP sequencing data from numerous cell lines and primary cells on a wide range of transcription factors, histone modifications and other proteins involved in transcription. Our EBNA 2 and EBNA 3 ChIP sequencing data tracks were aligned against published ChIP sequencing tracks showing the location of various histone modifications from the EBV immortalised LCL GM12878. The most significant 1000 EBNA 2 and EBNA 3 peaks (+/-1 1kb from the peak centre) were analysed and colocalised peaks of histone modifications from GM12878 ChIP-sequencing data by hierarchical clustering and displayed as a heat map. From these data we found that EBNA 2 bound sites are commonly enriched for histone modifications including H3K27ac, H3K4me and H3K9ac in GM12878 LCLs (Fig 3-9). Active enhancer clusters were defined by high levels of acetylated lysine 27 on histone H3 (H3K27ac) and mono-methylated lysine 4 on histone H3 (H3K4me1) whereas the poised clusters are defined by the presence of H3K4me1 but a lack of H3K27ac. EBNA 2 binding sites appeared not to co-localise with the repressive mark tri-methylation on lysine 27 histone H3 (H3K27me3) or the PRC2 protein EZH2 which is responsible for depositing this mark (Fig 3-9). Analysis of EBNA 3 binding sites also revealed a co-localisation with active enhancer elements; however the most active sites were also strongly bound by EBNA 2 (Fig 3-10). The remaining EBNA 3 most significant peaks were clustered with poised enhancers as shown by an enrichment of H3K4me1 (Fig 3-10).

Further analysis was conducted by taking an average signal across a 2kb window centred on the middle of the top 1000 binding sites for EBNA 2. This was compared to the average ChIP sequencing data for GM12878 histone modifications at these sites (Fig 3-11 A). These data confirms peaks of H3K27ac, H3K4me and H3K9ac are abundant in GM12878 at sites where EBNA 2 is bound significantly. The same analysis was performed on EBNA 3 binding sites and the same marks were enriched (Fig 3-11 B). In both cases it is striking that the top 1000 EBNA 2 peaks are also commonly bound by EBNA 3 and vice versa. The lack of enrichment at the top 1000 peaks for histone modifications at the middle of the binding site is a common feature of

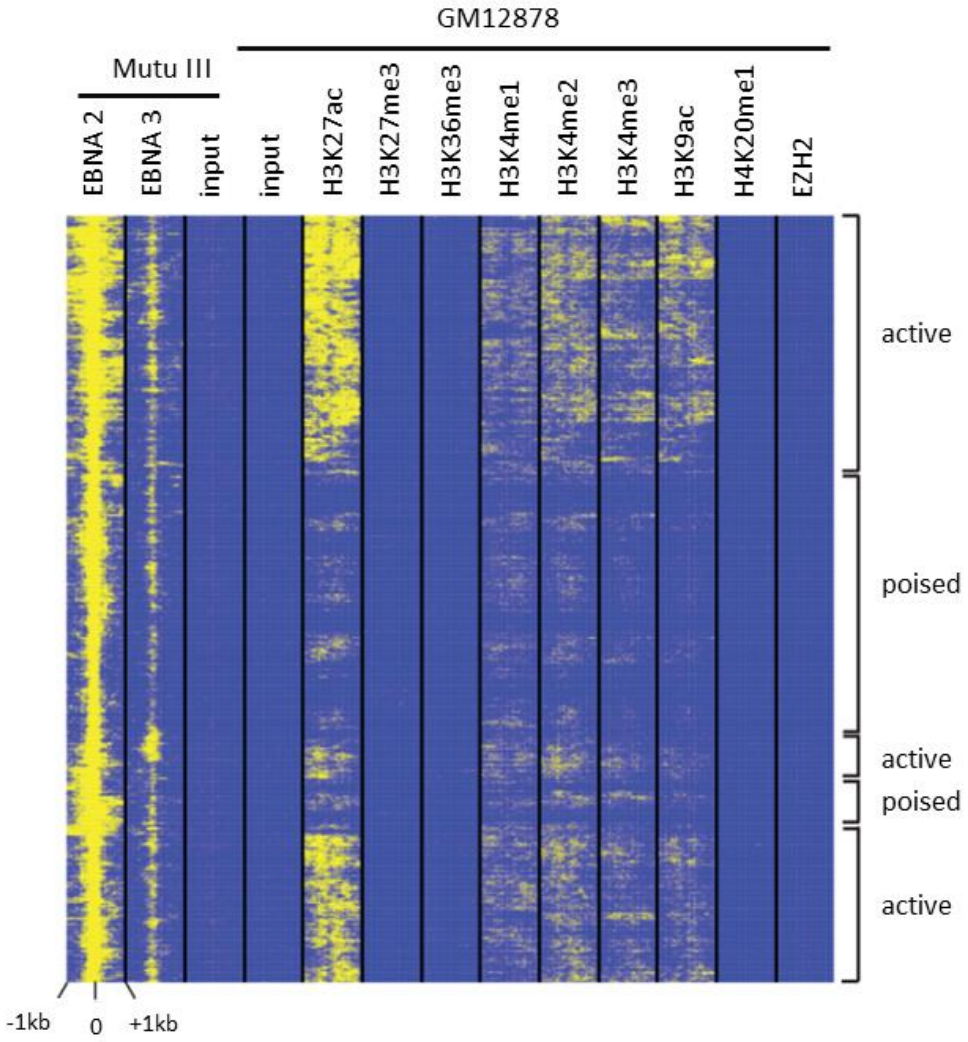


Figure 3-9 Colocalisation of histone modifications at the top 1000 most significant EBNA 2 binding sites. Heatmap of EBNA2, EBNA 3 and histone modification ChIP-sequencing signals at the top 1000 EBNA 2 binding sites. EBNA 2 and 3 ChIP-sequencing data from Mutu III BL cells was aggregated with ENCODE histone modification ChIP-seq data from the GM12878 LCL using hierarchical clustering. Each window displays the ChIP-seq signal -/+1 kb around the EBNA 2 binding site midpoint. Clusters of active enhancers (H3K4me1+, H3K27ac+) and poised enhancers (H3K4me1+, H3K27ac-) are indicated.

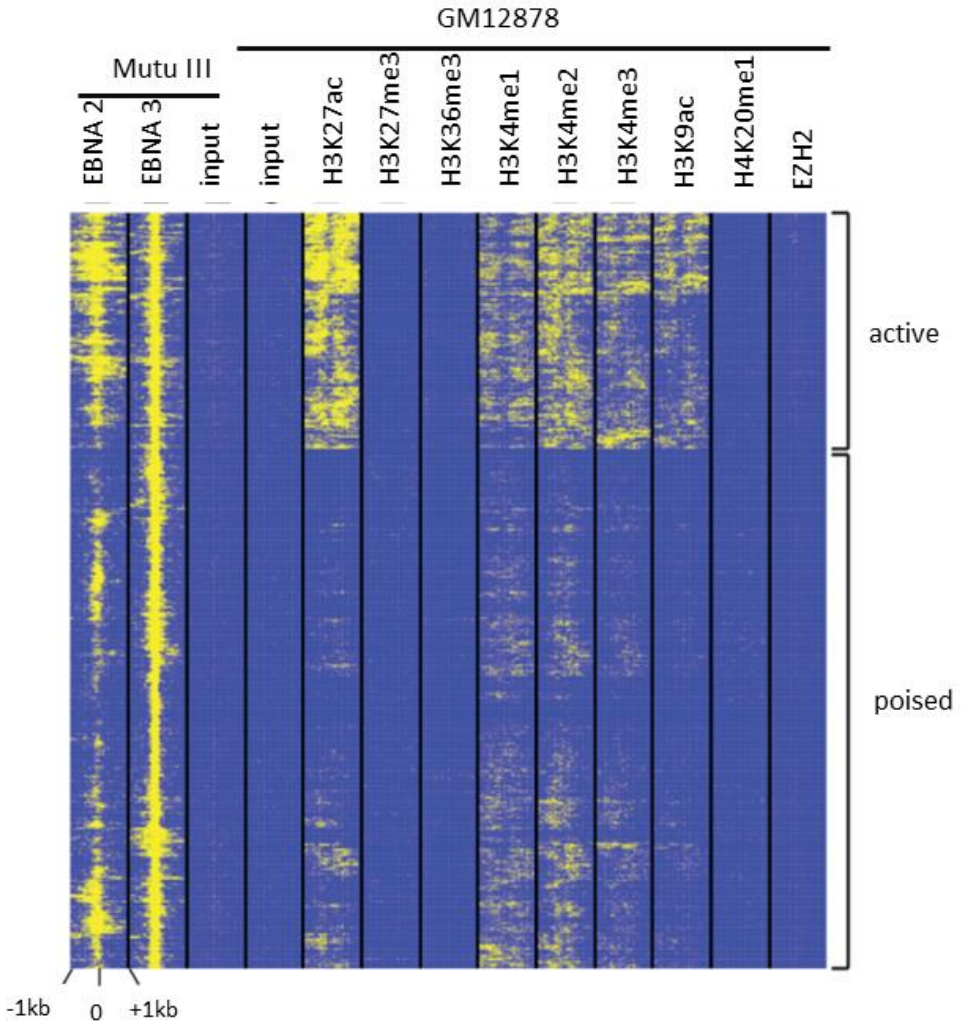


Figure 3-10 Colocalisation of histone modifications at the top 1000 most significant EBNA 3 binding sites. Heatmap of EBNA2, EBNA 3 and histone modification ChIP-sequencing signals at the top 1000 EBNA 3 binding sites. EBNA 2 and 3 ChIP-sequencing data from Mutu III BL cells was aggregated with ENCODE histone modification ChIP-seq data from the GM12878 LCL using hierarchical clustering. Each window displays the ChIP-seq signal -/+1 kb around the EBNA 2 binding site midpoint. An active enhancer (H3K4me1+, H3K27ac+) and poised enhancer (H3K4me1+, H3K27ac-) clusters are indicated.

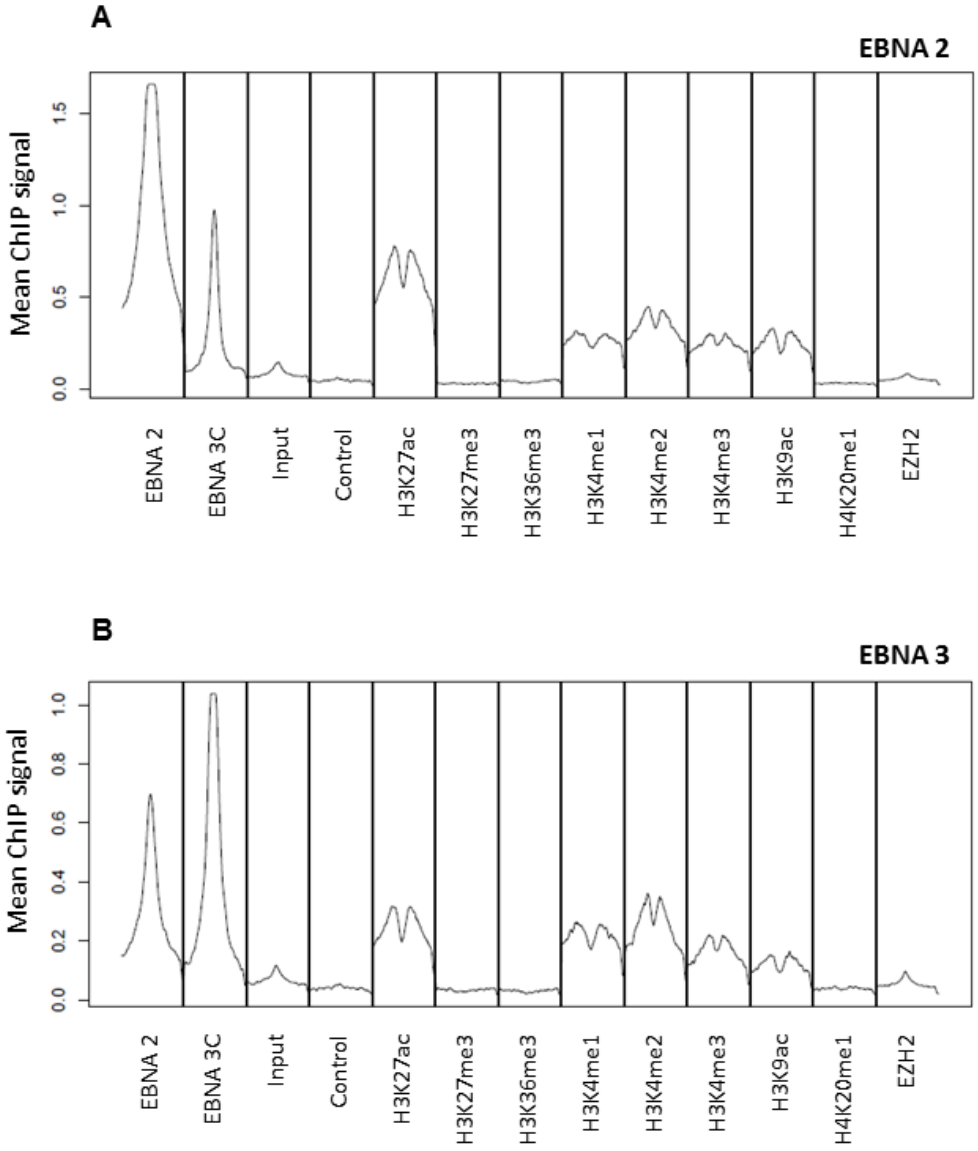


Figure 3-11 Average ChIP sequencing signal of histone modifications at the top 1000 most significant EBNA 2 and EBNA 3 binding sites. (A) Average EBNA2, EBNA 3 and histone modification ChIP-sequencing signals from ENCODE histone modification ChIP-seq data for the GM12878 LCL at the top 1000 EBNA 2 binding sites. (B) Average signals for the top 1000 EBNA 3 binding sites.

histone modifications and can readily be observed on the ENCODE database at known enhancer sites. This likely represents a local depletion of histones at sites where proteins bind in abundance. Together these data may suggest that EBNA 2 and EBNA 3 proteins are targeted to active and poised enhancers by cellular TFs. This may serve as another explanation of why promoter reporter assays had failed to be responsive for EBNA 2 and EBNA 3 proteins as these are conducted out of chromatin context. In summary it appears that histone modifications have a prominent role to play in gene regulation by EBNA 2 and EBNA 3 bound enhancers.

3.7 Colocalisation of EBNA 2 and EBNA 3 biding sites with cellular transcription factors

As it is known that EBNA 2 and EBNA 3 proteins cannot bind DNA directly our next approach was to analyse which cellular transcription factors co-localised with peaks of binding. This was performed to confirm known binding partners of EBNA 2 and 3 proteins (RBPJκ and PU.1) and could potentially identify unknown candidate partners. We analysed GM12878 LCL ChIP sequencing data for cellular TFs (ENCODE database) and recently published EBNA 2 and RBPJκ ChIP sequencing data from the EBV immortalised LCL IB4 (Portal et al. 2011), at the top 1000 EBNA 2 binding sites. We found that EBNA 2 co-localised with a plethora of B cell transcription factors, many of which have not previously been described as EBNA associated proteins (Fig 3-12). EBNA 2 significantly colocalised with its known binding partner RBPJκ and other B cell transcription factors including: BATF, BCL11A, Bcl3, IRF4, PAX5, SP1 and TCF12. EBNA 2 binding sites also co-localised with the histone acetyl transferase p300, a co-activator usually associated with enhancer elements (Erwin et al. 2014)(Fig 3-12). The most significant 1000 EBNA 3 binding sites also co-localised with the same set of B cell transcription factors (Fig 3-13). We also took the average signal for transcription factor binding signal in GM12878 and IB4 LCLs at all 1000 EBNA 2 sites (Fig 3-14 A). This analysis confirmed co-localisation of B-cell transcription factors at EBNA 2 sites and in this case the signal for PU.1 was included, revealing co-localisation. A similar result was obtained from analysis of the average ChIP sequencing signal of cellular TFs at the top 1000 EBNA 3 peaks (Fig 3-14 B). It is noteworthy that of all the B-cell transcription factors in GM12878 on which there is ChIP-seq data, the highest signal at EBNA 2 peaks was EBF1.

To further examine potential cellular TFs involved in targeting EBNA 2 and EBNA 3 proteins to DNA we performed unbiased analysis using MEME ChIP (motif analysis of large DNA datasets) to identify any enriched motifs in EBNA 2 and EBNA 3 precipitated DNA. Enriched motifs were then compared for matches to known TF consensus binding sites. Using DNA sequences at the top 300 most significant EBNA 2 binding sites we found that the most enriched (p-value 5E-161)

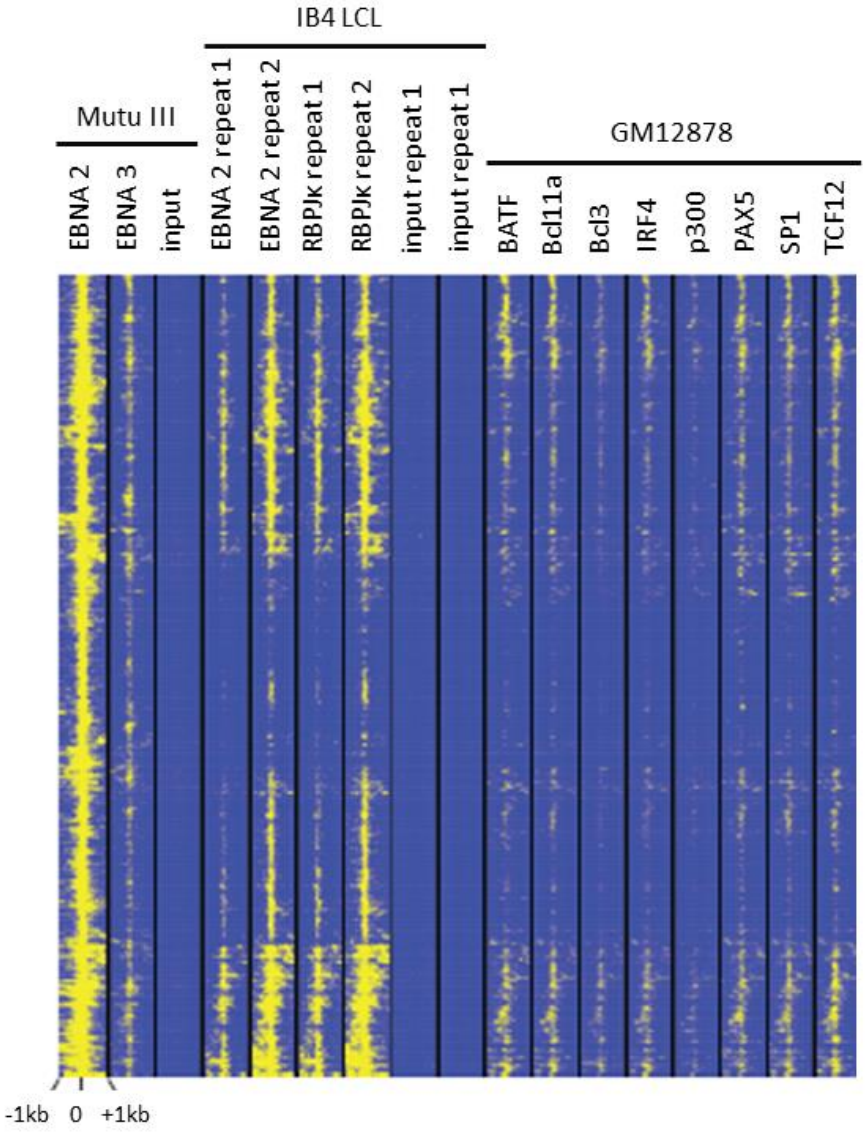


Figure 3-12 Colocalisation of TFBS at the top 1000 most significant EBNA 2 binding sites. Heatmap of EBNA 2, EBNA 3 and transcription factor ChIP-sequencing signals at the top 1000 EBNA 2 binding sites. EBNA 2 and 3 ChIP-sequencing data from Mutu III BL cells was aggregated with ENCODE histone modification ChIP-seq data from the GM12878 and IB4 LCLs. Each window displays the ChIP-seq signal -/+1 kb around the EBNA 2 binding site midpoint.

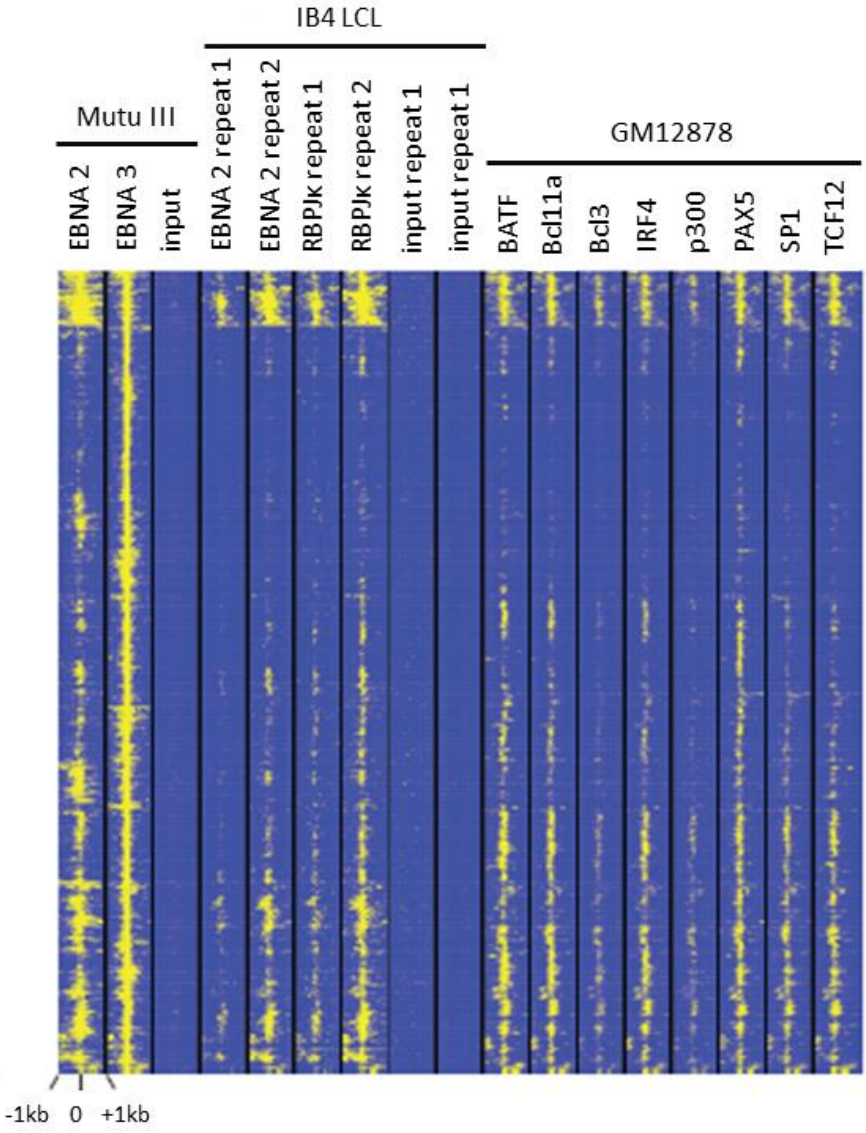


Figure 3-13 Colocalisation of TFBS at the top 1000 most significant EBNA 3 binding sites. Heatmap of EBNA 2, EBNA 3 and transcription factor ChIP-sequencing signals at the top 1000 EBNA 3 binding sites. EBNA 2 and 3 ChIP-sequencing data from Mutu III BL cells was aggregated with ENCODE transcription factor ChIP-seq data from the GM12878 and IB4 LCLs. Each window displays the ChIP-seq signal -/+1 kb around the EBNA 2 binding site midpoint.

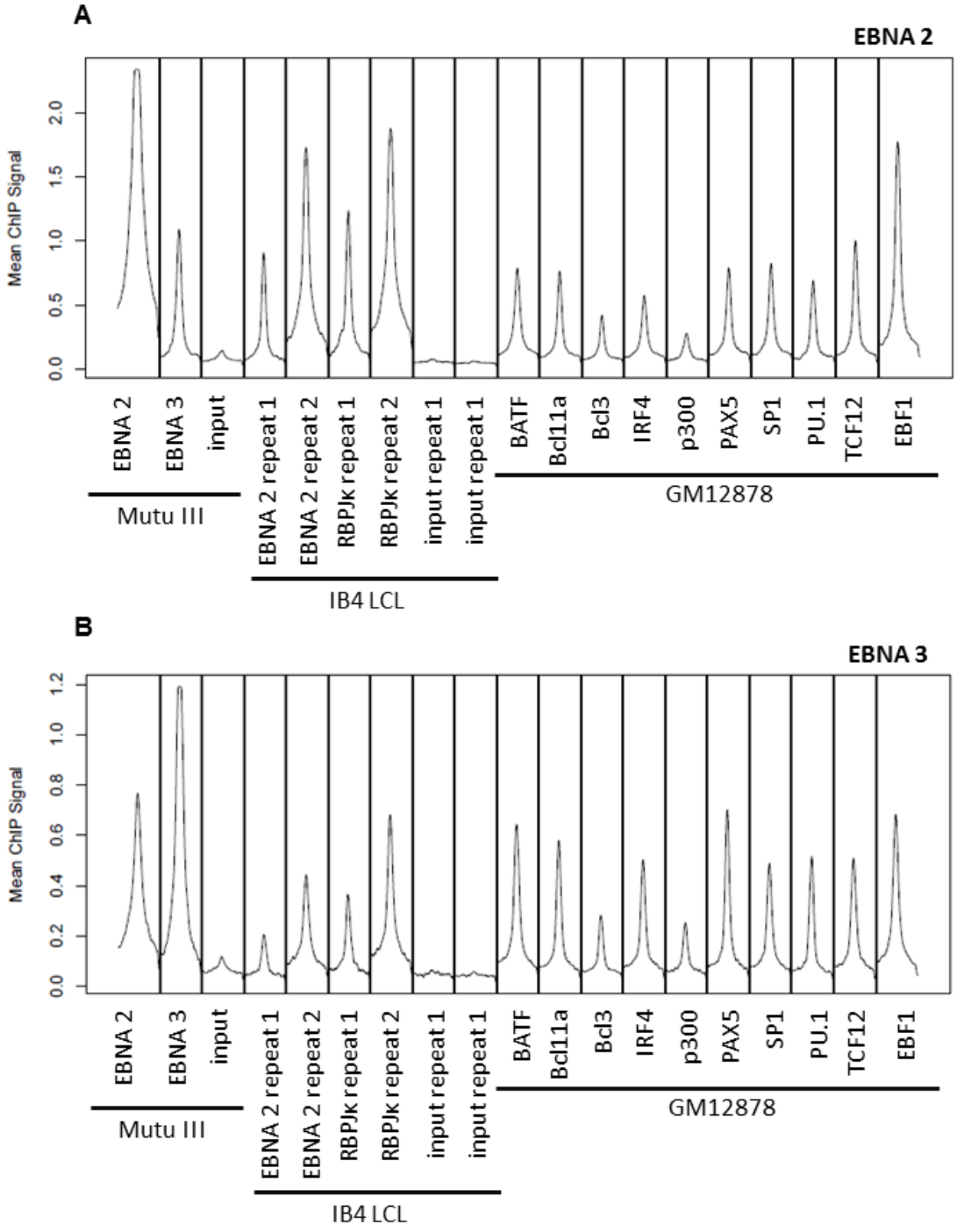


Figure 3-14 Average ChIP sequencing signal of transcription factors at the top 1000 most significant EBNA 2 and EBNA 3 binding sites. (A) Average EBNA 2, EBNA 3 and transcription factor ChIP-sequencing signals from ENCODE histone modification ChIP-sequencing data for the GM12878 and IB4 LCLs at the top 1000 EBNA 2 binding sites. (B) Average signals for the top 1000 EBNA 3 binding sites.

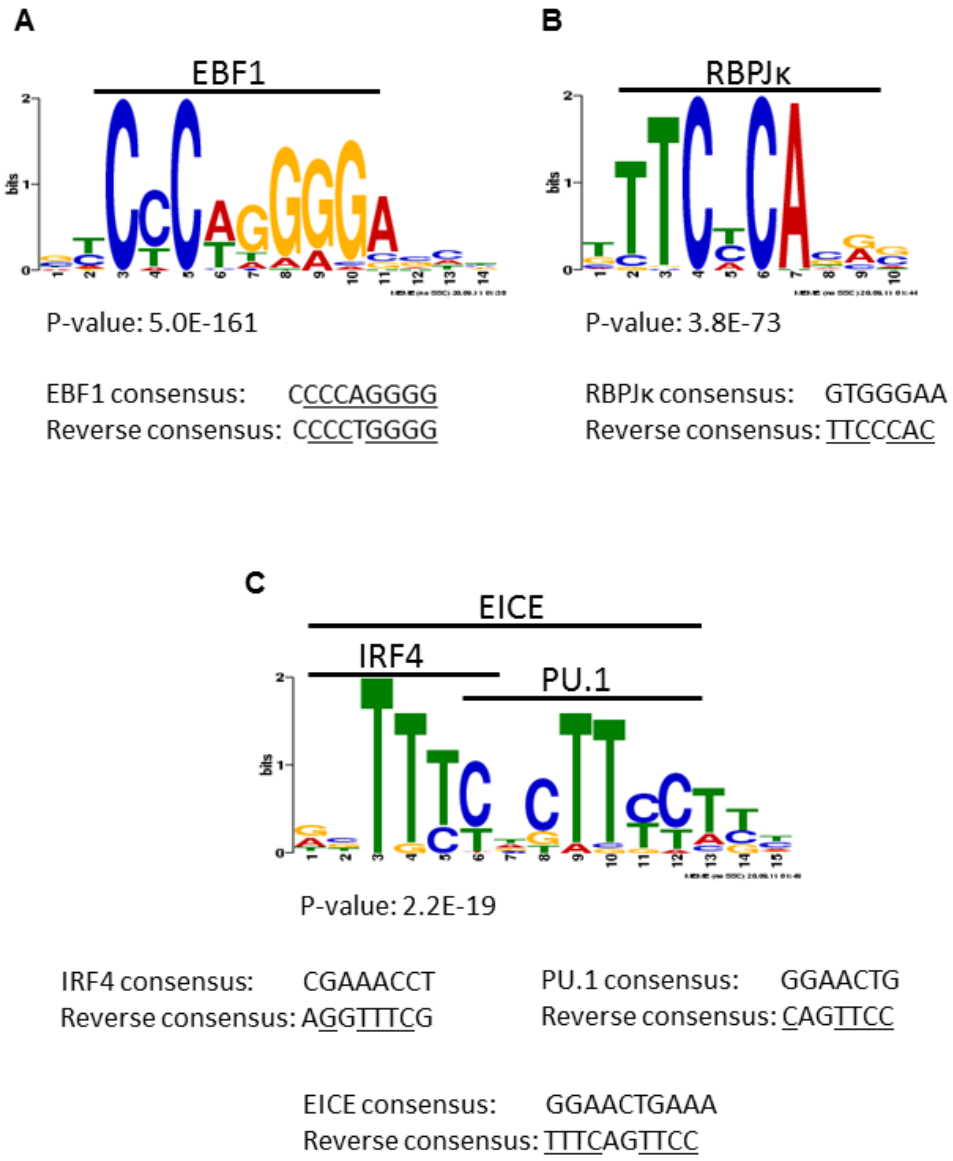


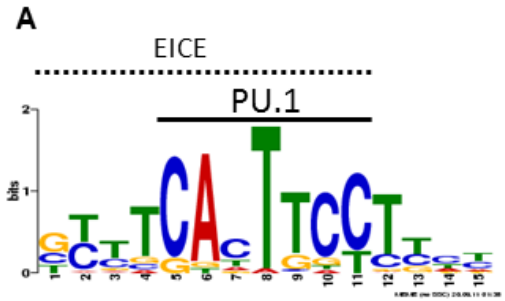
Figure 3-15 Enriched motifs at the top 300 most significant EBNA 2 binding sites. The top three most enriched motifs from the DNA sequences of the top 300 EBNA 2 binding sites generated by unbiased motif searching. All sequences are displayed 5' to 3'. Letter height represents frequency of a position being that nucleotide within the motif-like sequences. These were compared to consensus motifs for cellular TFs from the ENCODE project and JASPAR database. Areas within an enriched motif that are similar to a known TF consensus site are indicated. Consensus and reverse consensus sites are displayed below the enriched motif and part of consensus or reverse consensus found in the enriched motif are underlined. Numbers show the p-value for site enrichment. (A) EBF1 like motif, (B) RBPJκ like motif, (C) EICE (PU.1 and IRF4) like motif.

motif by EBNA 2 binding was an almost exact match for EBF1 (Fig 3-15 A). A motif similar to the RBPJκ consensus was also significantly enriched (Fig 3-15 B). The third most enriched motif was found to be similar to a composite PU.1, IRF4 site or EICE (Fig 3-15 C). This enriched motif was not highlighted as aligned against any known TF consensus site by MEME as it does not examine composite overlapping sites. EICEs are however a well categorised subclass of TF binding sites and combine the core PU.1 consensus (GGAA) with that of IRF4. The same analysis of the DNA sequence of the top 300 most significant EBNA 3 peaks revealed that the most enriched sequence was almost identical to that of the consensus binding site for PU.1 (Fig 3-16 A). The enriched sequence also bore similarity to the EICE consensus sequence but the IRF4 portion of this was not as consistently found at the top 300 EBNA 3 binding sites as the core PU.1 consensus was. The next two most significantly enriched consensus motifs (p-value 9.9E-8 and 3.2E-6) did not align to any known cellular TF consensus motif (Fig 3-16 B, C). They may represent complex composite sites or may be an artefact, as the P-value of these enriched motifs is lower than other EBNA 2 and EBNA 3 enriched motifs. As we had found that a large proportion of our binding sites were bound by both EBNA 2 and EBNA 3 proteins we also analysed DNA sequences from all shared EBNA 2 and EBNA 3 binding sites. We did this to determine EBNA 2 and EBNA 3 targeting to a shared site was dependent on their joint recruitment by a particular cellular TF. We found that DNA sequences bound by both EBNA 2 and EBNA 3 contained significantly enriched motifs, the top 3 of which were shown to be similar to the EICE (Fig 3-17 A), EBF1 (Fig 3-17 B) and PU.1 (Fig 3-17 C) consensus sequences respectively.

3.8 Alignment of genes bound by EBNA 2 and EBNA 3 with genes regulated in published micro array data sets

To determine if EBNA 2 and 3 family binding correlated with gene regulation we assembled a list of all the genes most proximal to a binding site (8386 genes for EBNA 2 (Table 3-1) and 3937 genes EBNA 3 (Table 3-2)). We compared these list with all genes known to be regulated by EBNA 2, EBNA 3A, EBNA 3B and EBNA 3C from gene expression microarray data (Spender et al. 2002; Chen et al. 2006; Maier et al. 2006; Spender et al. 2006; Zhao et al. 2006a; Zhao et al. 2006b; Lucchesi et al. 2008; Hertle et al. 2009; White et al. 2010; McClellan et al. 2012; Skalska et al. 2013).

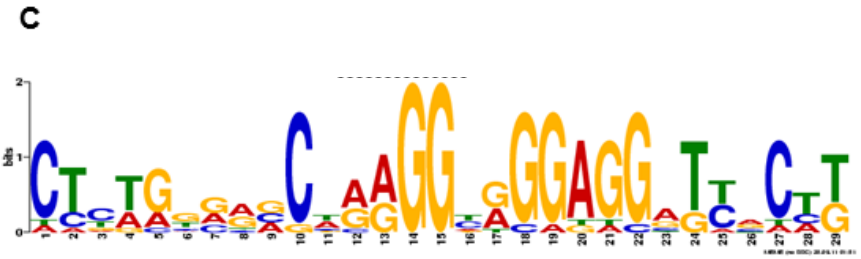
1. **3A KO LCL** (Hertle et al. 2009). Compared expression data from wild type infected LCLs vs LCLs infected with a virus lacking EBNA 3A.



P-value: 1.1E-29

PU.1 consensus: GGAACTG
Reverse consensus: CAGTTCC

P-value: 9.9E-8



P-value: 3.2E-6

Figure 3-16 Enriched motifs at the top 300 most significant EBNA 3 binding sites. The top three most enriched motifs from the DNA sequences of the top 300 EBNA 3 binding sites generated by unbiased motif searching. All sequences are displayed 5' to 3'Letter height represents frequency of a position being that nucleotide within the motif-like sequences. These were compared to consensus motifs for cellular TFs from the ENCODE project and JASPAR database. Areas within an enriched motif that are similar to a known TF consensus site are indicated. Consensus and reverse consensus sites are displayed below the enriched motif and part of consensus or reverse consensus found in the enriched motif are underlined. Numbers show the p-value for site enrichment. (A) PU.1 like motif, (B) unaligned motif, (C) unaligned motif.

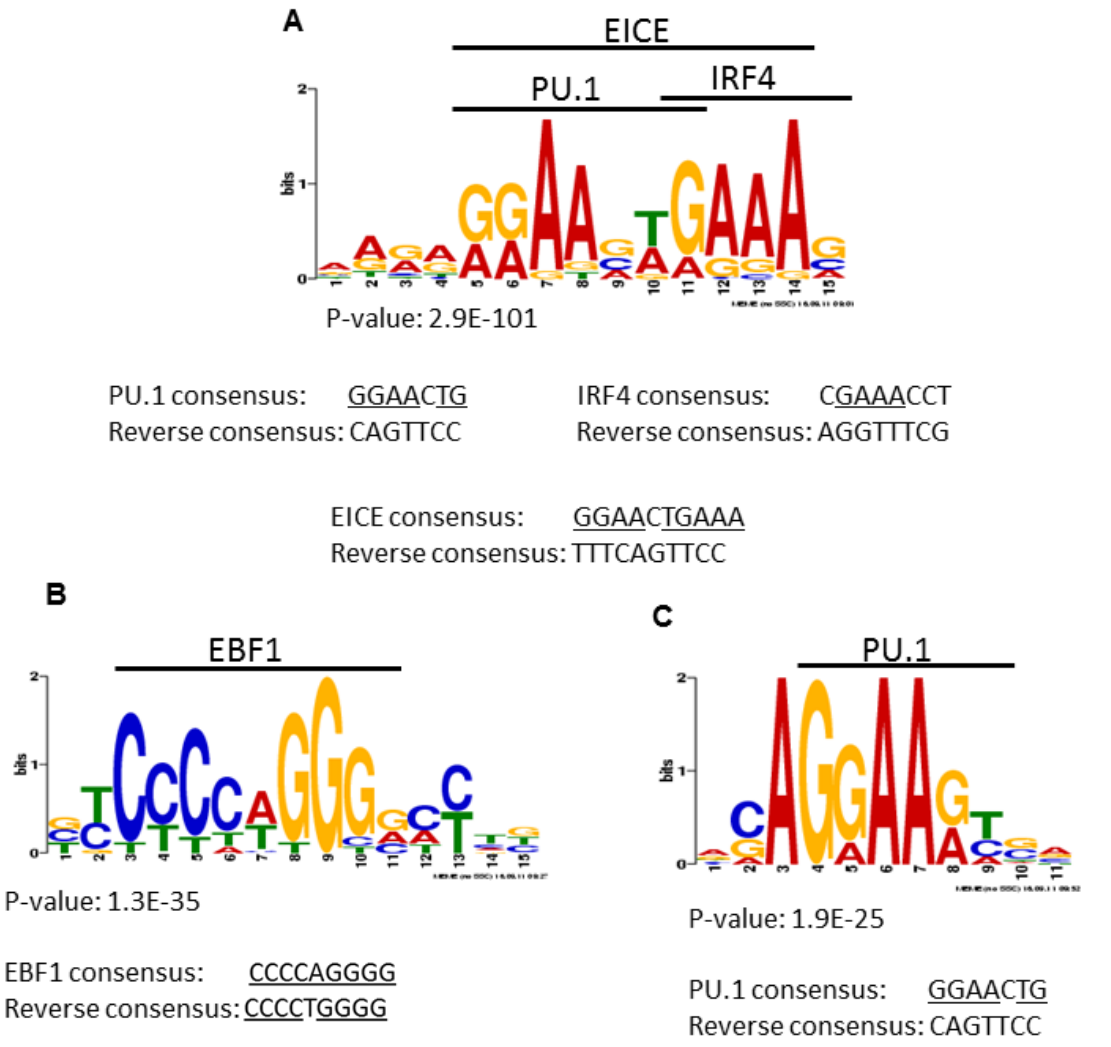


Figure 3-17 Enriched motifs at all EBNA 2 and EBNA 3 shared binding sites. The top three most enriched motifs from the DNA sequences of all EBNA 2 and EBNA 3 binding sites generated by unbiased motif searching. Letter height represents frequency of a position being that nucleotide within the motif-like sequences. All sequences are displayed 5' to 3'. These were compared to consensus motifs for cellular TFs from the ENCODE project and JASPAR database. Areas within an enriched motif that are similar to a known TF consensus site are indicated. Consensus and reverse consensus sites are displayed below the enriched motif and part of consensus or reverse consensus found in the enriched motif are underlined. Numbers show the p-value for site enrichment. (A) EICE (PU.1 and IRF4) like motif, (B) EBF1 like motif, (C) PU.1 like motif.

	3A		3B		3B/c		3C				E3	E2					
EBNA 2 bound genes	1. 3A LCL	2. 3A BL31	3. 3B LCL	4. 3B BL31	5. 3B/3C LCL	6. BJAB 3C	7. 3C LCL	8. 3C BL31	9. 3C LCL	10. 3C LCL	11. 3E BL31	12. EB2 BL41	13. EB2 BJAB	14. EB2 LCL	15. EB2 Ak31	16. EB2 LCL	17. EB2 LCL
total genes in array	300	210	197	601	16	170	542	864	81	430	1009	325	376	76	23	64	7
bound genes	183	111	110	317	11	74	333	408	43	260	567	225	239	42	22	56	5
% of array bound	61%	53%	56%	53%	69%	44%	61%	47%	53%	60%	56%	69%	64%	55%	96%	88%	71%
up reg. bound genes	105	61	42	141	3	43	177	253	26	170	372	168	86	42	22	56	5
down reg. bound genes	78	50	68	176	8	31	156	155	17	90	195	57	153	0	0	0	0
total percentage of ChIP seq hit by arrays: 18.9% (1591/8386)																	

Table 3-1 Comparison of EBNA 2 bound genes to genes regulated by EBNA 3A, EBNA 3B, EBNA 3C and EBNA 2. All published regulated genes were compared with a list of genes bound by EBNA 2. Chart shows the EBNA factors investigated and cell line in which the experiment was conducted. Data sets are clustered colour coded and labelled. The total number of genes reported to be regulated, the number of these that are bound in our ChIP sequencing data by EBNA 2 and the percentage which this represents are all indicated. Total number of bound regulated genes is further divided into genes that bound in our ChIP sequencing and upregulated in the microarray dataset and genes that are bound and downregulated. At the bottom of the chart is the total number of genes bound by EBNA 2 and how many of those have been shown, in any array, to be regulated.

	3A		3B		3B/c	3C					E3	E2					
EBNA 3C bound genes	**1. 3A LCL**	**2. 3A BL31**	**3. 3B LCL**	**4. 3B BL31**	**5. 3B/3CLCL**	**6. BJAB 3C**	**7. 3C LCL**	**8. 3C BL31**	**9. 3C LCL**	**10. 3C LCL**	**11. 3E BL31**	**12. EB2 BL41**	**13. EB2 BJAB**	**14. EB2 LCL**	**15. EB2 Ak31**	**16. EB2 LCL**	**17. EB2 LCL**
total genes in array	300	210	197	601	16	170	542	864	81	430	1009	325	376	76	23	64	7
bound genes	121	82	87	234	6	56	168	263	35	195	362	159	170	21	17	39	5
% of array bound	40%	39%	44%	39%	38%	33%	31%	30%	43%	45%	36%	49%	45%	28%	74%	61%	71%
up reg. bound genes	58	36	29	90	1	31	77	146	19	125	211	122	68	21	17	39	5
down reg. bound genes	63	46	58	144	5	25	91	117	16	70	151	37	102	0	0	0	0
total percentage of ChIP seq hit by arrays: 24.7% (974/3937)																	

Table 3-2 Comparison of EBNA 3 bound genes to genes regulated by EBNA 3A, EBNA 3B, EBNA 3C and EBNA 2. All published regulated genes were compared with a list of genes bound by EBNA 3. Chart shows the EBNA factors investigated and cell line in which the experiment was conducted. Data sets are clustered colour coded and labelled. The total number of genes reported to be regulated, the number of these that are bound in our ChIP sequencing data by EBNA 3 and the percentage which this represents are all indicated. Total number of bound regulated genes is further divided into genes that bound in our ChIP sequencing and upregulated in the microarray dataset and genes that are bound and downregulated. At the bottom of the chart is the total number of genes bound by EBNA 3 and how many of those have been shown, in any array, to be regulated.

2. **3A KO BL31** (White et al. 2010). Compared expression data from EBV negative BL31 cells infected with wild type BAC derived virus with BL31 cells infected with BAC derived virus lacking EBNA 3A.
3. **3B KO LCL** (White et al. 2010). Compared expression data from wild type infected LCLs vs LCLs infected with a virus lacking EBNA 3B.
4. **3B KO BL31** (White et al. 2010). Compared expression data from EBV negative BL31 cells infected with wild type BAC derived virus with BL31 cells infected with BAC derived virus lacking EBNA 3B.
5. **3B-/3C low LCL** (Chen et al. 2006). Compared expression data from wild type infected LCLs vs LCLs infected with a virus lacking EBNA 3B and with low EBNA 3C expression.
6. **BJAB 3C +/-** (McClellan et al. 2012). Compared expression data from EBV negative B-cell lymphoma BJAB lines transfected with a control plasmid, or a plasmid stably expressing EBNA 3C.
7. **3CHT LCL** (Zhao et al. 2006a). Data from an EBV immortalised LCLs expressing a hydroxy-tamoxifen (HT) conditional EBNA 3C; compared expression data from LCLs growing in HT to LCLs grown for 7 days without HT, corresponding to LCL levels of EBNA 3C vs depleted levels of EBNA 3C but prior to growth arrest. Cut off 1.5 fold.
8. **3C KO BL31** (White et al. 2010). Compared expression data from EBV negative BL31 cells infected with wild type BAC derived virus with BL31 cells infected with BAC derived virus lacking EBNA 3C.
9. **3CHT LCL** (Skalska et al. 2013). Data from an $p16^{INK4A}$ deleted EBV immortalised LCLs expressing a hydroxy-tamoxifen (HT) conditional EBNA 3C; compared expression data from LCLs growing in HT to cells after HT removal, and those growing without HT to those to which HT has been re added.
10. **3CHT LCL** (Skalska et al. 2013). Data from an $p16^{INK4A}$ deleted EBV immortalised LCLs expressing a hydroxy-tamoxifen (HT) conditional EBNA 3C; compared expression data from LCLs growing in HT to cells after HT removal, and those growing without HT to those to which HT has been re added. (1.4 fold)
11. **E3 KO BL31** (White et al. 2010). Compared expression data from EBV negative BL31 cells infected with wild type BAC derived virus with BL31 cells infected with BAC derived virus lacking EBNA 3A, EBNA 3B and EBNA 3C.
12. **ER/EB2 BL41** (Maier et al. 2006). Data from EBV negative BL41 lines transfected with a plasmid expressing a β-estradiol conditional EBNA 2. Compared expression data from cells with active EBNA 2 to those without.

13. **ER/EB2 BJAB** (Maier et al. 2006). Data from EBV negative B-Cell lymphoma lines transfected with a plasmid expressing a β-estradiol conditional EBNA 2. Compared expression data from cells with active EBNA 2 to those without.
14. **EB2 LCL** (Zhao et al. 2006b). Data from an EBV immortalised LCLs expressing a hydroxy-tamoxifen (HT) conditional EBNA 2; compared expression data from LCLs growing in HT to LCLs grown for 24 hours without HT, corresponding to EBNA 2 cytoplasmic re-localisation. Cut off 1.5 fold.
15. **EB2 Ak31** (Lucchesi et al. 2008). Compared expression data from EBV negative Ak31 cells transfected with a control plasmid, or a plasmid containing type 1 or type 2 EBNA 2.
16. **EB2 EREB 2.5 (A)** (Spender et al. 2006). Data from an EBV immortalised LCLs expressing a β-estradiol conditional EBNA 2; compared expression data from LCLs growing without β-estradiol for 5 days to LCLs which have had β-estradiol re-added for 6 hours.
17. **EB2 EREB 2.5 (B)** (Spender et al. 2002). Data from an EBV immortalised LCLs expressing a β-estradiol conditional EBNA 2; compared expression data from LCLs growing without β-estradiol for 5 days to LCLs which have had β-estradiol re-added for 6 hours.

Using the 8386 EBNA 2 bound genes we found that 18.9% have previously been shown to be regulated by at least one of EBNA 2, EBNA 3A, EBNA 3B or EBNA 3C. 24.7% of 3937 genes bound by EBNA 3 have been shown to be regulated in at least one data set. The percentage of bound genes regulated in individual arrays was striking. In one data set 96% of genes shown to be regulated by EBNA 2 were bound by EBNA 2 in our ChIP sequencing data set. Furthermore, in an EBNA 3C array, 45.3% of EBNA 3C regulated were bound by EBNA 3 proteins. As EBNA 2 is regarded as an activator of transcription and EBNA 3 proteins generally as repressors of transcription; we next examined if EBNA 2 bound genes more likely to be activated by EBNA 2 (Fig 3-18) and EBNA 3 bound genes were more likely to be repressed by EBNA 3 proteins (Fig 3-19). Data was plotted as percentage of total genes. Whilst we could observe limited correlations within individual data sets it appears in general that EBNA 2 and EBNA 3 proteins activate and repress similar ratios of cellular genes. This data strongly suggests that EBNA 2 and EBNA 3 binding to the human genome is functionally relevant to their transcriptional effects on nearby genes.

3.9 Discussion

EBV is capable of immortalising resting B cells through the concerted effect of its latent gene products. Latency III proteins have been shown to perturb cell cycle progression, immune

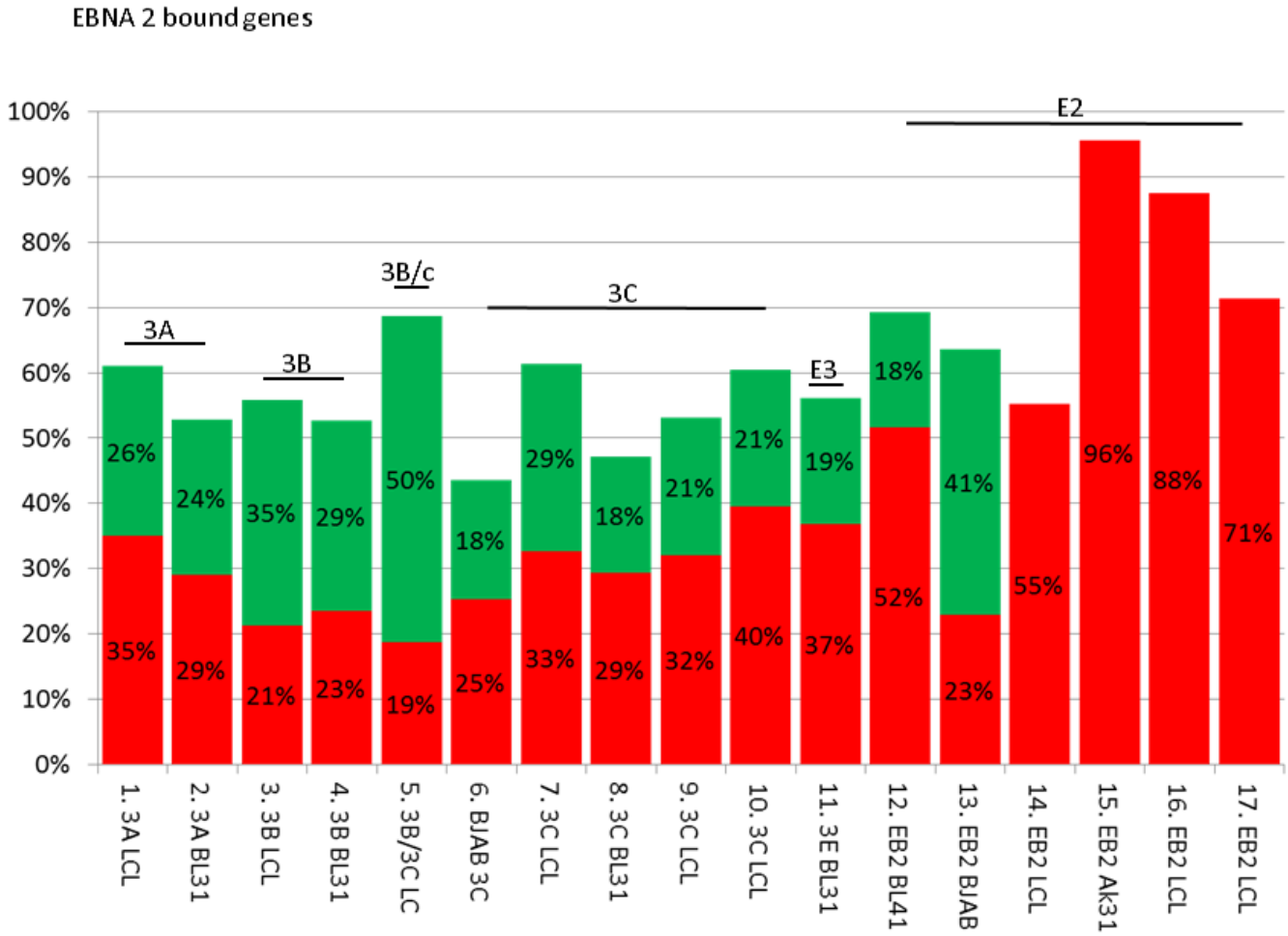


Figure 3-18. Comparison of EBNA 2 bound genes to genes regulated by EBNA 3A, EBNA 3B, EBNA 3C and EBNA 2. Graphical representation of the data in Table 3-1. Y-axis is % of genes in an array that are bound by EBNA 2. A value of 100% would indicate every gene reported to be regulated was bound by EBNA 2. These are sub divided into up (red) and down (green) regulated genes.

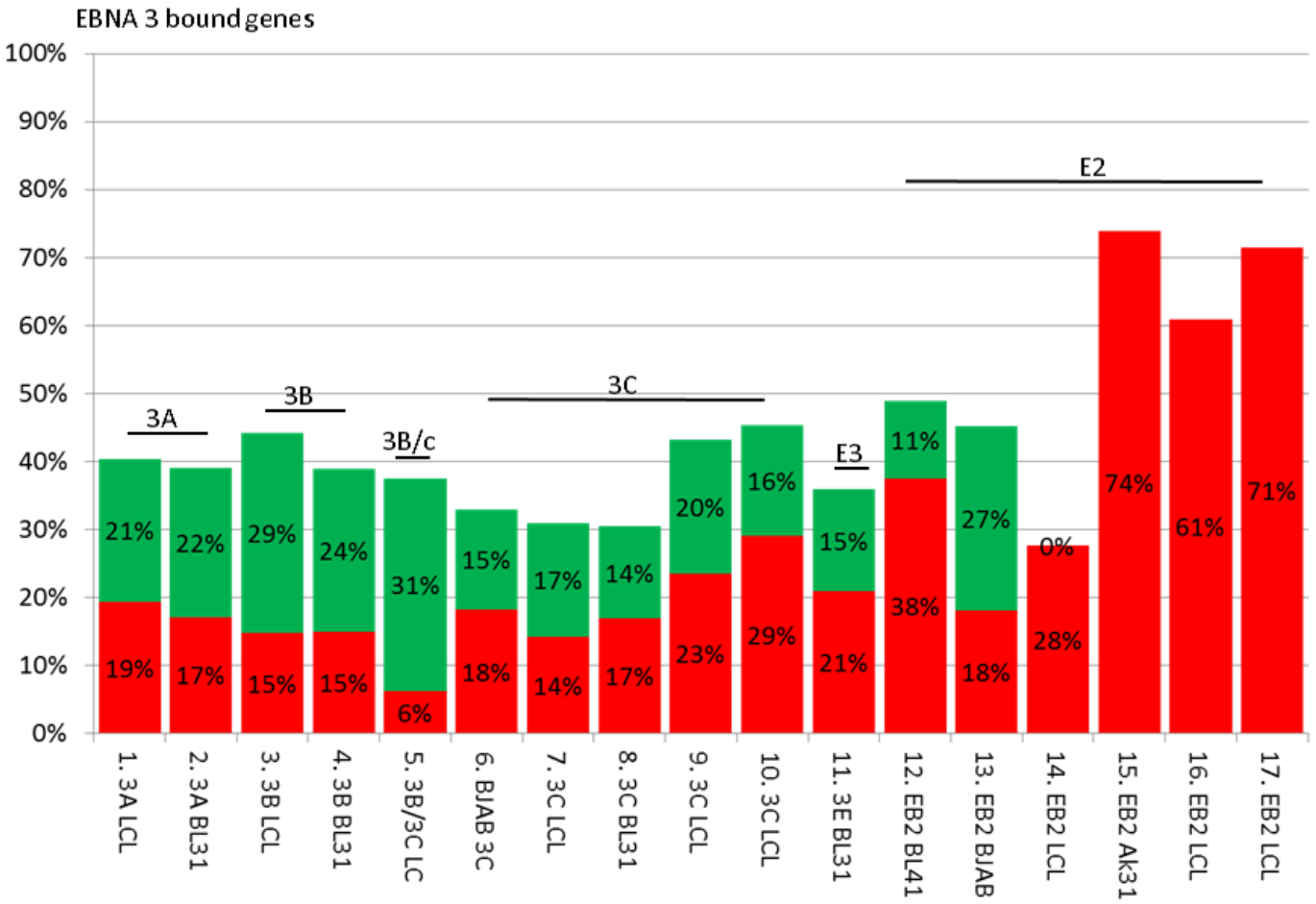


Figure 3-19. Comparison of EBNA 3 bound genes to genes regulated by EBNA 3A, EBNA 3B, EBNA 3C and EBNA 2. Graphical representation of the data in Table 3-1. Y-axis is % of genes in an array that are bound by EBNA 3. A value of 100% would indicate every gene reported to be regulated was bound by EBNA 3. These are sub divided into up (red) and down (green) regulated genes.

response, cell migration and growth control through both protein-protein interactions and transcriptional effects. Many cellular genes regulated by EBNA 2, 3A, 3B and 3C have been previously identified by gene expression microarray data but very little was known about how regulation was mediated. Here we show that EBNA proteins likely exert a large proportion of their influence by extensive interactions with the human genome. EBNA 2 and EBNA 3 binding sites do not seem to be spread randomly across the genome as genes most proximal to these binding sites are also likely known targets of EBNA proteins from expression data (Table 3-1, Table 3-2, Fig, 3-18, Fig 3-19). These include extensively studied targets of EBV such as *CR2* (*CD21*)(Calender et al. 1987) and *BCL2L11* (*Bim*)(Clybouw et al. 2005). Furthermore the EBNA binding sites cluster such that only a limited number of genes are most proximal to a site. The 21605 EBNA 2 and 7044 EBNA 3 sites are most proximal to 8385 and 3936 genes respectively. A high proportion of these sites are shared between EBNA 2 and EBNA 3 proteins suggesting extensive interplay between them, although both also have unique binding sites (Fig 3-7). We also show that EBNA 2 and EBNA 3 binding sites are commonly found distal to any gene TSS suggesting that enhancers play a key role in EBNA 2 and EBNA 3 mediated gene regulation (Fig 3-6). However, the data presented here does not show the complete mechanism of transcriptional regulation by EBNA 2 and EBNA 3 proteins, or indeed if a particular EBNA 2 or EBNA 3 bound site is directly responsible for the transcriptional regulation of the nearest gene; as recent chromatin looping experiments have shown that enhancer-promoter interactions can occur between extremely distal genes, and even between chromosomes (Spilianakis et al. 2005; Lomvardas et al. 2006).

The data obtained also provided information on the relationship of histone modifications with EBNA 2 and EBNA 3 binding. The histone modifications present at the peaks of binding were enriched for active and poised chromatin marks (Fig 3-9 to 3-11). EBNA 3C and EBNA 2 can transactivate *LMP1* but cellular gene activation by EBNA 3C is less well characterised, so this result was unexpected. Also, although EBNA 3 proteins have also been implicated in H3K27me3 mediated repression, there was a lack of enrichment at EBNA 3C binding sites for this repressive mark H3K27me3 or for EZH2 (Fig 3-10). The reason for this may be cell type specific differences in histone modifications between Mutu III and GM12878 LCLs (the cell line histone modification ChIP sequencing was examined). Another explanation could be that H3K27me3 appears to be found in broad peaks across large segments of a chromosome between insulating sites, distinguished by CTCF and other insulator protein binding sites (Lhoumaud et al. 2014). Lastly this analysis could be complicated by the fact that a high

proportion of the top 1000 EBNA 3 binding sites were also bound by EBNA 2 which is generally considered to be an activator of transcription.

Analysis of EBNA 2 binding site DNA sequences revealed enriched motifs which corresponded to known cellular interaction partners; RBPJκ and PU.1 (Fig 3-15). The most enriched motif at significantly bound EBNA 2 sites was EBF1, which was also the strongest average signal at the top 1000 EBNA 2 binding sites (Fig 3-14). To date, EBF1 has not been shown to target EBNA 2 to DNA but there is increasing evidence that EBNA 2 and EBF1 at least share a great number of binding sites (Zhao et al. 2011). These data add to this finding. EBNA 3 bound DNA sequence analysis revealed an enriched motif which was highly similar to the consensus PU.1 binding site, suggesting that EBNA 3 proteins are brought to cellular sites by this factor (Fig 3-16). The role of EICEs in targeting EBNA 2 and EBNA 3 proteins to the human genome is one that is only just being investigated. A ChIP sequencing experiment using antibodies against EBNA 2 also found EICE consensus sites to be commonly found in EBNA 2 precipitated DNA (Zhao et al. 2011). Furthermore a recent study found that this consensus motif was enriched at genes differentially regulated by EBNA 2 type 1 and 2 and is found at the *LMP1* promoter where both EBNA 3C and EBNA 2 can bind (Tzellos et al. 2014). It is therefore noteworthy that an EICE was found in the most enriched motif in DNA bound by both EBNA 2 and EBNA 3. What role EICEs have in recruitment of EBNA 2 and EBNA 3 to shared sites remains to be discovered.

This study was based on the knowledge that EBNA 2 and EBNA 3 proteins are important regulators of cellular transcription. Much of this knowledge was gained from gene expression microarray data sets. Here we show that a remarkably high proportion of these previously reported genes are bound by EBNA 2 and EBNA 3 proteins. This suggests that EBNA 2 and EBNA 3 proteins exert their influence on host transcription by extensive binding to the human genome.

In summary we have developed a ChIP method that precipitates sufficient amounts of native, endogenous EBNA 2 and EBNA 3 protein and associated DNA to perform ChIP-sequencing experiments. This data shows that EBNA 2 and the EBNA 3 proteins associate extensively with the human genome and that EBNA binding sites are significantly enriched for histone modifications indicative of open or poised chromatin. It also appears that the EBNA proteins preferentially target long-range enhancer elements to perturb host transcription. Furthermore a large number of B cell transcription factors are enriched at these sites although which of these transcription factors are targeting the EBNAs to the host genome has yet to be elucidated. We also found a striking overlap between genes that have been previously

reported as regulated by EBNA 2 and EBNA 3 proteins and our ChIP sequencing data. Importantly, this data suggests that EBNA 2 and EBNA 3 proteins are targeted to enhancers by B-cell transcription factors to regulate the transcription of cellular genes.

4. EBNA 3C represses integrin receptor signalling genes via promoter proximal and distal elements

The West lab had previously carried out gene expression analysis using Affymetrix microarray chips to determine which cellular genes were regulated by EBNA 3C to gain functional insight into how EBNA 3C contributes to B-cell immortalisation. EBNA 3C is a potent repressor of transcription when tethered to DNA via GAL4 fusion proteins (Bain et al. 1996). It is known to also co-activate *LMP1* with EBNA 2 via PU.1 sites (Lin et al. 2002). At the time this work was conducted this in fact was the only EBNA 3C DNA associated site in the viral or human genome. Other gene expression microarray experiments have revealed that EBNA 3C regulates over 1500 cellular genes (Chen et al. 2006; Zhao et al. 2006a; White et al. 2010; McClellan et al. 2012; Skalska et al. 2013). Our ChIP sequencing data suggests that much of the regulatory effects of EBNA 3C occur via EBNA 3C associations with DNA. We wanted to find out if any of the highly repressed genes highlighted by expression data were directly bound by EBNA 3C. Furthermore we wished to determine the mechanism of repression via EBNA 3C DNA interactions and resolved to investigate changes in histone modifications at repressed genes as EBNA 3C dependant changes in histone methylation and acetylation have been observed at other loci (White et al. 2010).

4.1 EBNA 3C represses chemokine integrin receptor signalling genes

To investigate potential direct targets of EBNA 3C on host cell gene transcription we analysed gene expression microarray data from our lab (McClellan et al. 2012). Total RNA was extracted from stable transfectants of the EBV-negative B-cell lymphoma BJAB cell lines E3C4 and PZ2 (described in (Wang et al. 1990a). E3C4 stably expresses EBNA 3C in the absence of other EBV proteins whereas PZ2 contains a control plasmid and is both EBV and EBNA 3C negative. Two independent RNA preparations were used on Affymetrix GeneChip arrays and normalised data was analysed using four different pairwise comparisons of signals from E3C4 and PZ2. Data obtained in this way revealed 47 genes were downregulated and 122 genes were upregulated by 2 fold or more. We decided to focus on the genes repressed by EBNA 3C as more is known about how EBNA 3C may repress transcription. We found that of the top 20 most downregulated genes there were a number of chemokine and integrin receptor signalling genes (Table 4-1). We then used all 47 repressed genes as an input to DAVID which highlighted that several pathways were statistically enriched in this data set including chemokine signalling and focal adhesion. 8 out of the 47 genes were annotated as being involved in the chemokine

Up regulated by EBNA 3C

Gene Symbol	Signal log ratio
CADM1	8.06
LUM	7
PELI2	5.56
CTNND2	4.28
TNNI1	4.25
SNAI2	4.18
PAPSS2	4.1
LGSN	4.07
HIST1H2BG	4.05
CR2	4
CSRP2	3.89
IGHM	3.65
MYO1D	3.48
SSBP2	3.42
HIST1H2AE	3.41
LSAMP	3.38
MME	3.34
NREP	3.33
KLRC1	3.32
RAPGEF4	3.29

Down regulated by EBNA 3C

Gene Symbol	Signal log ratio
ADAM28	-5.66
CXCL11	-5.03
CXCL13	-3.65
ITK	-3.38
C5AR1	-3.27
ITGB1	-3.26
KLF6	-3.09
ITGA4	-3.08
CELF2	-2.97
ADAMDEC1	-2.82
LRIG1	-2.68
TTC38	-2.41
TCTN1	-2.4
CASP4	-2.34
CXCL10	-2.18
VAV3	-2.17
NAAA	-2.16
TIMP1	-2.08
CCL3 + CCL3L1/3	-2.07
CCL4	-1.94

Table 4-1 Top 20 most down regulated and upregulated genes in BJAB EBNA 3C positive and control cells. Gene expression microarray data conducted in BJAB EBNA 3C positive cell line E3C3 and BJAB EBNA 3C negative cell line PZ2 was analysed and the top 20 most upregulated and downregulated were examined. Genes are listed by the average log ratio signal from two independent RNA extractions analysed using four different pairwise comparisons, only genes counted 3 or more times are listed. Log 2 ratio was used to linearise data and ratios of +1 and -1 would be equivalent to a two fold up regulation or two fold down regulation respectively. Microarray experimetns were performed by Sarika Khasnis and table is adapted from supplementary table 1 in McClellan et al 2012 (McClellan et al. 2012).

signalling pathway and subsets of these 8 genes were also involved in cytokine-cytokine receptor interaction and toll-like receptor signalling (Table 4-2). The focal adhesion and leukocyte activation pathways were also significantly enriched, each with 4 out of the 47 genes annotated as being part of that pathway. These two pathways are highly related as ITGA4, ITGB1 and VAV3 are annotated as being in both and will be collectively referred to as integrin receptor signalling pathway. The significance value for the enrichment of this pathway is likely an underestimate as the top 20 most downregulated genes included the a disintegrin and metalloproteases ADAM28 and ADAMDEC1. These were not annotated in DAVID for focal adhesion despite the fact that ADAM28 has published roles in adhesion and ADAMDEC1 is a related gene and has a role in adhesion (Bridges et al. 2002; Bridges et al. 2005; Lund et al. 2013). We therefore included ADAM28 and ADAMDEC1 in our future analysis and concluded that EBNA 3C, when expressed without other EBV proteins, specifically and significantly represses chemokine signalling and integrin receptor signalling genes.

In the chemokine signalling pathway EBNA 3C appears to target multiple points primarily regulating the levels of chemokines and therefore the input signal into the pathway (*CXCL10*, *CXCL11*, *CCL3*, *CCL4*, *CCL3L1* and *CCL3L3*)(Fig 4-1). Secreted chemokines fundamentally function as chemoattractants. C-X-C motif ligand (CXCL) 10 and 11 along with CXCL9 are involved in leukocyte trafficking, are co-located in the genome and secreted by leukocytes and tissue cells (such as epithelial cells). Interestingly they all share the CXCR3 receptor. CXCR3 is rapidly induced on naïve T-cells following stimulation and is highly expressed on CD4+, CD8+ T-cells and NK cells. CXCL9, 10 and 11 have overlapping, antagonistic, distinct and redundant roles depending on the cellular and environmental context. In general however they function in the generation of T-cells and the interactions and migration behaviour of effector T-cells (reviewed in (Groom and Luster 2011; van den Borne et al. 2014). CCL3, CCL4 and the CCL3 like proteins CCL3L1/2 are also chemoattractants with overlapping roles to CXCL proteins. CCL3 and CCL4 produced by NK cells has been shown to inhibit HIV replication and provide protection against human cytomegalovirus by recruitment of other effector immune cells (Robertson 2002). Chronic lymphocytic leukaemia cells are characterised by high levels or production and secretion of CCL3 and CCL4 which appear to provide survival advantages by the recruitment of helper T-cells and accessory factors (Stratowa et al. 2001; Davids and Burger 2012). Whilst the function of chemokine signalling appears to be context specific it is generally true that it is an important pathway by which the body communicates with the immune system, facilitating recruitment, migration and survival of immune cells.

enriched pathway	number of genes	% of total genes	P-Value	Benjamini
chemokine signaling pathway	8	17	1.20E-06	4.10E-05
cytokine-cytokine receptor interaction	6	12.8	1.30E-03	2.20E-02
Toll-like receptor signaling pathway	4	8.5	4.20E-03	4.80E-02
leukocyte transendothelial migration	4	8.5	6.50E-03	5.60E-02
focal adhesion	4	8.5	2.70E-02	1.80E-01
small cell lung cancer	3	6.4	3.10E-02	1.70E-01
T-cell receptor signaling pathway	3	6.4	4.90E-02	2.20E-01

Table 4-2 DAVID analysis of all down regulated genes in BJAB EBNA 3C +/- array. All 47 repressed genes from a BJAB EBNA 3C+/- data set (McClellan et al. 2012) were entered into DAVID as an official gene symbol list. Kegg pathways enriched are shown and the count of the number of genes from the input in each pathway, the percentage of total genes this represents and measures of statistical significance are displayed. The p-value is a modified Fisher Exact P-Value, for gene-enrichment analysis. It ranges from 0 to 1. A Fisher Exact P-Value = 0 represents perfect enrichment. Usually a P-Value equal or smaller than 0.05 is considered strongly enriched in the annotation categories.

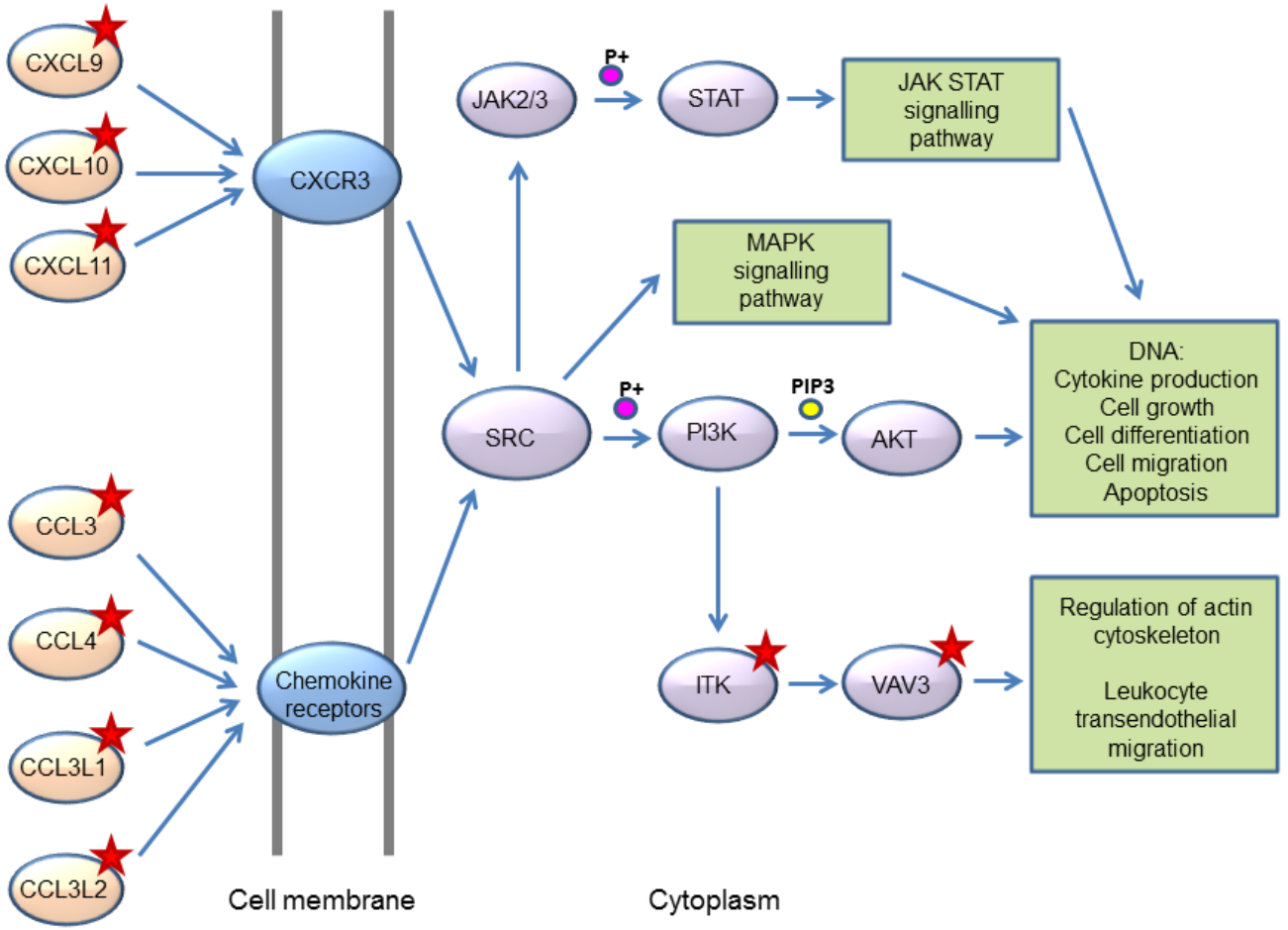


Figure 4-1 Chemokine signalling pathway genes are enriched in genes downregulated by EBNA 3C. EBNA 3C down regulated gene annotated as part of the chemokine signalling pathway are indicated by a red star. Orange: extra cellular proteins, blue: receptors or membrane proteins, purple: cytoplasmic proteins, green: pathways or results of signalling. Cell membrane is represented by grey lines. Small molecule signalling or involvement is displayed in circles, pink for phosphorylation and yellow for PIP3 signalling. Arrows without extra annotation show activation via protein interactions. Environment is shown at the bottom of the figure. This data is adapted from DAVID pathway output and many intermediate interactions are not shown here

EBNA 3C also downregulates *VAV3* and *ITK* which are downstream targets in this pathway and affect the regulation of actin cytoskeleton. VAV3 had also been shown to be important in acute lymphoblastic leukaemia with its loss resulting in increased apoptosis (Chang et al. 2012). Interestingly VAV proteins are also targeted by LMP2 which blocks the effect of surface immunoglobulins through the LYN and SRC family protein kinases thereby repressing lytic activation (Miller et al. 1995). Both ITK and VAV3 have been highlighted as key genes downregulated during LCL generation with mRNA reduction detectable just 4 weeks after exposure to EBV (Dai et al. 2012). Furthermore mutations in ITK have been shown to be directly related to EBV associated lymphoproliferation (Huck et al. 2009; Linka et al. 2012).

The integrin receptor signalling pathway is also regulated at multiple points by EBNA 3C including the integrin receptor α4β1 expressed from *ITGA4* and *ITGB1*. There is a large degree of cross talk between the chemokine receptor signalling pathway and integrin receptor signalling and both signal through proteins such as SRC (Fig 4-2). ADAM28 is a ligand for α4β1 integrin (expressed from *ITGA4* and *ITGB1* genes), with binding of soluble ADAM28 mediating adhesion and lymphocyte infiltration of tissue through VCAM-1 (Bridges et al. 2002; Bridges et al. 2005). VCAM-1 signalling in endothelial cells causes changes in cell morphology and gene expression facilitating transendothelial migration. This process is dependent on ezrin, radixin and moesin (ERM) signalling which is required for T-cell adhesion and homing to lymphoid organs (Chen et al. 2013). Communication between endothelial cells and leukocytes during migration results in α4β1 binding to JAM2 which allows penetration through the epithelial cell layer (Cunningham et al. 2002). Less is known about ADAMDEC1 although it is expressed in B cells, is down regulated during tumourigenesis and is found on the same gene cluster as *ADAM28* and *ADAM7*. It also shares 47% sequence identity to *ADAM28* and has recently been shown to be a secreted, active enzyme capable of cleaving α_2-macroglobulin and casein (Lund et al. 2013).

We next investigated whether these genes were bound by EBNA 3 proteins in our ChIP-sequencing analysis to try to gain functional insights into their mechanism of repression. Interestingly many of the most downregulated genes in these pathways were also bound by EBNA 3 proteins in our Mutu III ChIP-sequencing data. To investigate the mechanism of repression by EBNA 3C and to validate EBNA 3C binding in BJAB background we used ChIP-QPCR to investigate a subset of these genes in more detail, focussing on *ITGA4*, *ITGB1*, *ADAM28*, *ADAMDEC1*, *CXCL10* and *CXCL11*.

4.2 EBNA 3C binds to and represses ITGB1 via promoter-proximal and distal elements

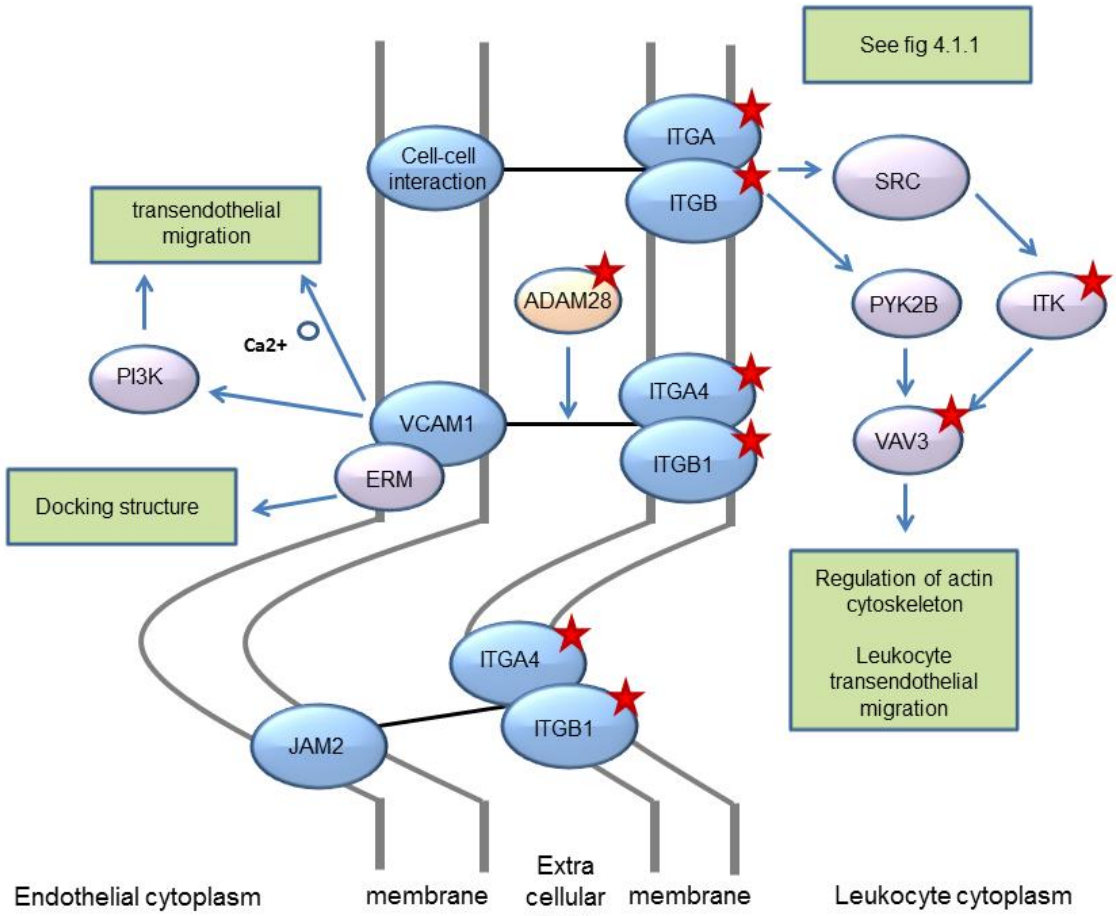


Figure 4-2 Integrin receptor signalling pathway genes are enriched in genes downregulated by EBNA 3C. EBNA 3C down regulated genes annotated as part of the integrin receptor signalling pathway are indicated by a red star. Orange: extra cellular proteins, blue: receptors or membrane proteins, purple: cytoplasmic proteins, green: pathways or results of signalling. Cell membrane is represented by grey lines. Intracellular protein interactions are shown in black lines. Activation via calcium signalling is indicated by a white circle. Arrows without extra annotation show activation via protein interactions. Environment is shown at the bottom of the figure. This data is adapted from DAVID pathway output and many intermediate interactions are not shown here

To confirm binding of EBNA 3C to repressed genes which had been shown to have proximal peaks of EBNA 3 binding in our ChIP sequencing data we performed ChIP-QPCR using the EBNA 3 pan specific antibody in BJAB control cell lines PZ1, PZ3 and the EBNA 3C-only-expressing BJAB cell lines E3C3 and E3C7. As PZ2 and E3C4 had previously been used in gene expression micro array experiments we used the alternative lines to attempt to control for clonal differences between the cell lines. We also also examined levels of histone modifications to see if EBNA 3C binding and repression cause enrichment of repressive marks (H3K27me3) or loss of active marks (H3 acetylation and H4 acetylation). Our ChIP sequencing data revealed that there are two EBNA 3 binding sites at *ITGB1*, one promoter-proximal and one distal. A visualisation of our ChIP sequencing data in the University of California Santa Cruz (UCSC) genome browser is shown in Figure 4-3 A. This is annotated with scale (top panel), our ChIP sequencing track data (EBNA 3 panel), H3K27Ac (layered H3K27Ac panel) , the *ITGB1* gene and TSS (for ease of graphical representation all exons are included regardless of isoform dependant inclusion) and positions of QPCR primers used at this locus. The layered H3K27Ac panel shows published ChIP sequencing data using antibodies against H3K27Ac from the EBV infected LCL GM12878; this mark is indicative of active enhancer elements. The y-axis of the EBNA 3 panel depicts reads per million background subtracted total reads (Fig 4-3 A). This is a measure of how many fragments of DNA from that site in the human genome were precipitated by the EBNA 3 antibody. Binding is only considered if the significance of the enrichment of fragments from that region passes a stringent threshold (p-value less than $1x10^{-7}$). Peak height is therefore generally well correlated with significance; for comparison, the 'smallest' peak called as significant in our data set is ~1.5 reads per million and the 'highest' observed peak is ~40 reads per million.

ChIP QPCR using primers at both peaks and in the gene as well as a control site where no EBNA is bound (*PPIA*) confirmed that EBNA 3C binds to both EBNA 3 binding sites in both E3C3 and E3C7 (Fig 4-3 B). H3K27me3 significantly increased across the promoter and at sites ~1kb upstream and downstream from the TSS but not at the upstream EBNA 3C binding site in EBNA 3C expressing cell lines compared to EBNA 3C negative ones (Fig 4-3 C). Increasing H3K27me3 levels suggests that EBNA 3C may be coordinating epigenetic silencing through polycomb group proteins as previously demonstrated at the *Bim* locus (Paschos et al. 2009). It is known that H3K27me3 exists in broad areas of the human genome and is not confined to distinct peaks. Acetylation of histone H3 and H4 is generally decreased in the presence of EBNA 3C compared to control cell lines although this is more apparent for H3ac than H4ac (Fig 4-3 D, E).

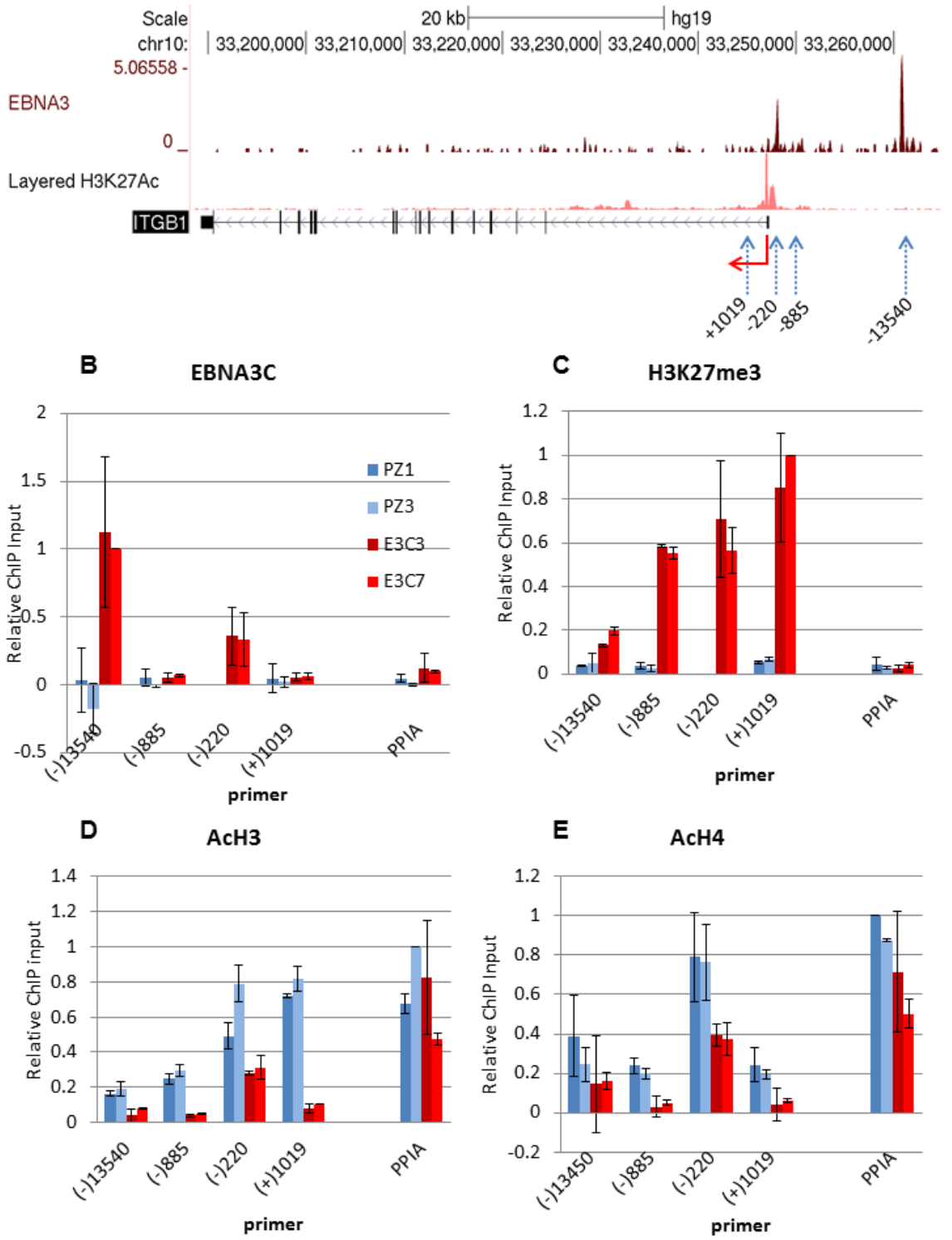


Figure 4-3 EBNA 3C binds to *ITGB1* and promotes epigenetic changes associated with gene repression. ChIP experiments conducted in the EBNA 3C positive cell lines E3C3, E3C7 and the EBNA 3C negative cell lines PZ1, PZ3. (A) EBNA 3 binding at the *ITGB1* locus. Top panel; scale, EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Numbers indicate the 5' end of the forward primers used for ChIP-QPCR analysis relative to the TSS. Control primers spanned the TSS of PPIA. Percentage input signals, after subtraction of signals of controls with no antibody, are expressed relative to the highest signal obtained. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 3C (B), H3K27me3 (C), H3ac (D) and H4ac (E).

This indicates a loss of active chromatin marks at *ITGB1* which is also consistent with gene repression.

4.3 EBNA 3C binds to and represses *ITGA4* via a promoter-proximal site

Along with *ITGB1*, *ITGA4* gene products forms α4β1 integrin. This molecule is able to interact with VCAM-1 and promote adhesion and cell migration (Garmy-Susini et al. 2013). *ITGA4* is bound by a promoter proximal EBNA 3 peak (Fig 4-4 A). Using ChIP QPCR EBNA 3C binding to this site was confirmed (Fig 4-4 B) and H3K27me3 levels were found to increase within the *ITGA4* but not at the promoter or upstream regions in the presence of EBNA 3C (Fig 4-4 C). In PZ1 and PZ3 control cells the levels of H3ac were already low compared to the constitutively active control gene (*PPIA*); however levels of H3ac were detectably reduced at this gene (Fig 4-4 D). Acetylation of H4 was reduced at primer sets for the peak and in the gene (Fig 4-4 E). Similarly to *ITGB1* this suggests that EBNA 3C is able to bind to this locus and downregulate expression through epigenetic mechanisms.

4.4 Binding of EBNA 3C to an intergenic peak at the *ADAM* locus causes epigenetic repression of *ADAM28* and *ADAMDEC1*

The *ADAM28* gene was the most downregulated gene in BJAB EBNA 3C expressing cells compared to control cells (Table 4-1). *ADAMDEC1* is related *to ADAM28*, is found in the same gene cluster and was also in the top 20 most downregulated genes. At the *ADAM* locus our ChIP sequencing data revealed a large peak of EBNA 3 binding between *ADAM28* and *ADAMDEC1* (Fig 4-5 A). As this was the most proximal EBNA 3 peak to both genes we used ChIP-QPCR to investigate whether EBNA 3C was capable of binding to this site. We used primer sets at the peak of binding as well as two flanking primer sets just outside the area of the genome significantly enriched by EBNA 3. The effect of EBNA 3C on levels of H3K27me3, H3ac and H4ac across this locus were also investigated by comparing ChIP QPCR data from control and EBNA 3C expressing cell lines. These data confirmed EBNA 3C was capable of binding to this site in E3C3 and E3C7 (Fig 4-5 B). Cells expressing EBNA 3C showed a dramatic increase in the polycomb associated mark H3K27me3 at the *ADAMDEC1* promoter and within *ADAM28* compared to control cell lines. Interestingly this mark did not increase at the *ADAM28* promoter (primer sets -587. -87, +1523) and only subtly increased at the peak itself (primer sets 5', mid, 3') (Fig 4-5 C). H3ac was low over this locus in control cell lines compared to the control gene (*PPIA*); however this mark was decreased to background levels in the presence of EBNA 3C, indicative of repression (Fig 4-6 B). H4ac levels were higher than those of H3ac

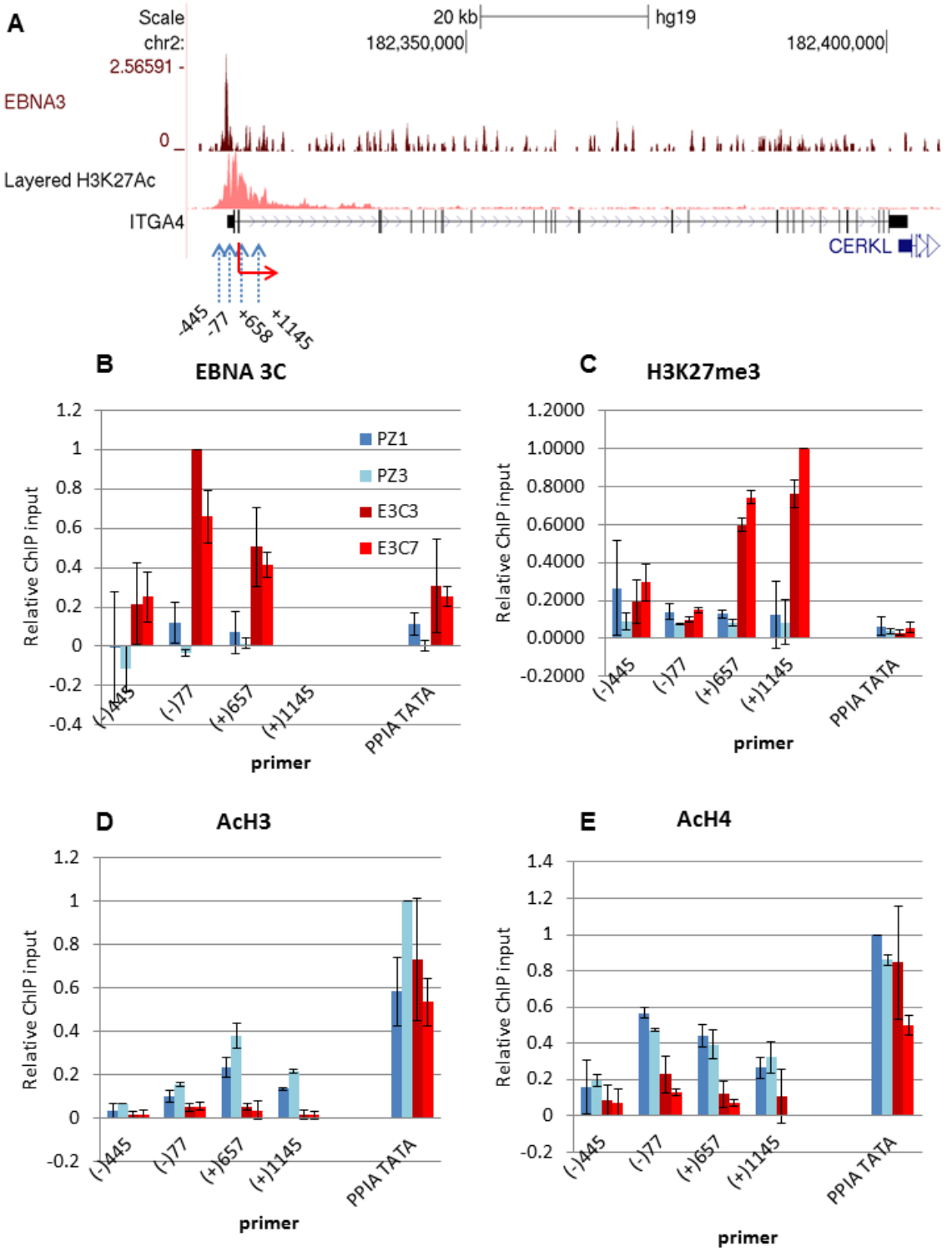


Figure 4-4 EBNA 3C binds to *ITGA4* and promotes epigenetic changes associated with gene repression. ChIP experiments conducted in the EBNA 3C positive cell lines E3C3, E3C7 and the EBNA 3C negative cell lines PZ1, PZ3. (A) EBNA 3 binding at the *ITGA4* locus. Top panel; scale, EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Numbers indicate the 5' end of the forward primers used for ChIP-QPCR analysis relative to the TSS. Control primers spanned the TSS of PPIA. Percentage input signals, after subtraction of signals of controls with no antibody, are expressed relative to the highest signal obtained. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 3C (B), H3K27me3 (C), H3ac (D) and H4ac (E).

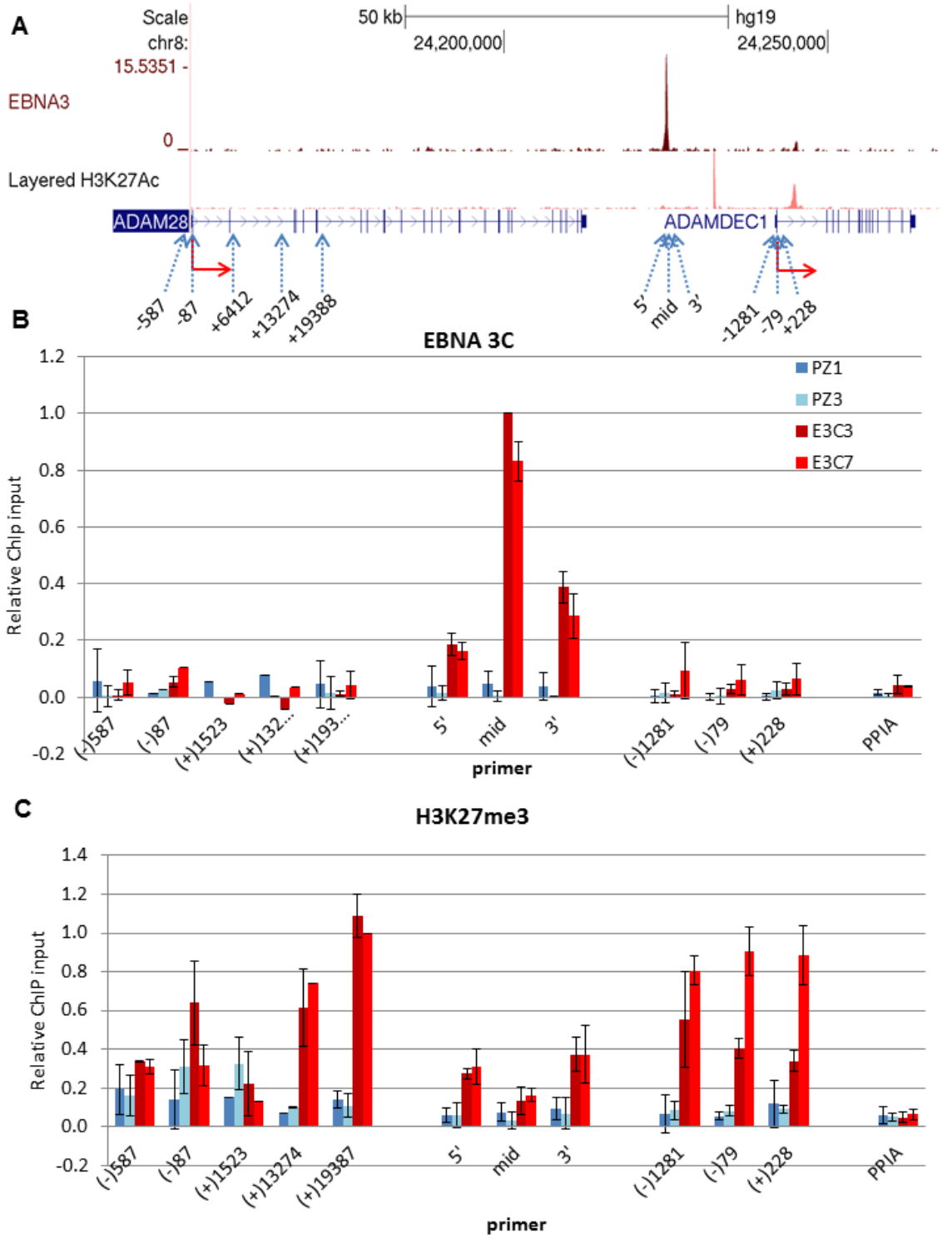


Figure 4-5 Binding of EBNA 3C to an intragenic peak at the *ADAM* locus causes epigenetic repression of *ADAM28* and *ADAMDEC1*. ChIP experiments conducted in the EBNA 3C positive cell lines E3C3, E3C7 and the EBNA 3C negative cell lines PZ1, PZ3. (A) EBNA 3 binding at the *ADAM* locus. Top panel; scale, EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Numbers indicate the 5' end of the forward primers used for ChIP-QPCR analysis relative to the TSS. Control primers spanned the TSS of PPIA. Percentage input signals, after subtraction of signals of controls with no antibody, are expressed relative to the highest signal obtained. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 3C (B), H3K27me3 (C).

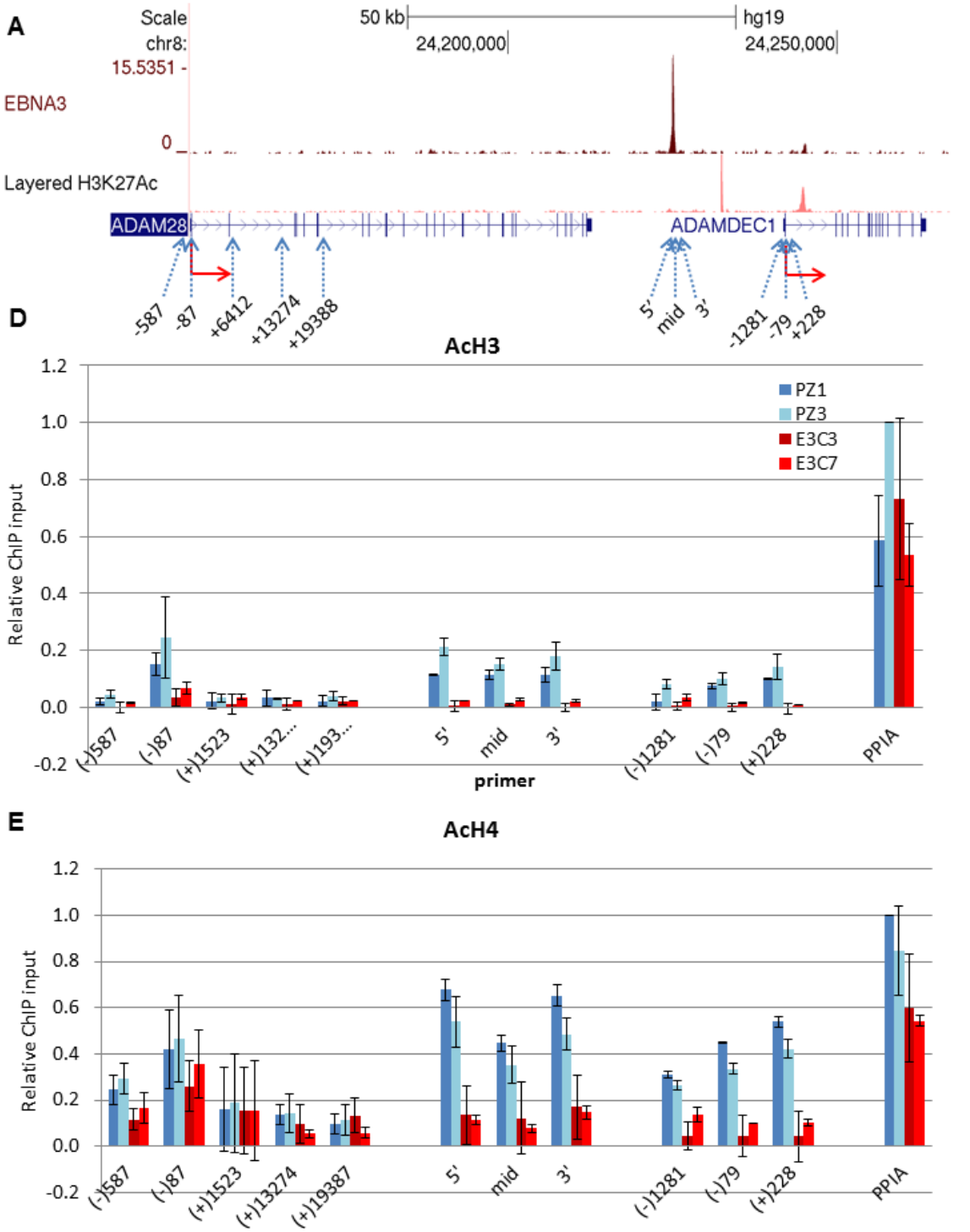


Figure 4-6 Binding of EBNA 3C at the *ADAM* locus correlates with a decrease in acetylation at *ADAM28* and *ADAMDEC1*. ChIP experiments conducted in the EBNA 3C positive cell lines E3C3, E3C7 and the EBNA 3C negative cell lines PZ1, PZ3. (A) EBNA 3 binding at the *ADAM* locus. Top panel; scale, EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Numbers indicate the 5' end of the forward primers used for ChIP-QPCR analysis relative to the TSS. Control primers spanned the TSS of PPIA. Percentage input signals, after subtraction of signals of controls with no antibody, are expressed relative to the highest signal obtained. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for H3ac (B), H4ac (C).

relative to the control gene but were only found to be significantly altered at the EBNA 3C binding site and at *ADAMDEC1*. At these sites the expression of EBNA 3C was correlated with a reduction in H4ac, further implicating epigenetic mechanisms in EBNA 3C mediated repression (Fig 4-6 C).

4.5 Epigenetic repression of *CXCL10* and *CXCL11* in cells expressing EBNA 3C

Chemokine signalling pathway genes were the most enriched cluster of repressed genes in the BJAB EBNA 3C -/+ gene expression microarray analysis with a total of 8 out of 47 repressed genes being involved in this pathway. Six of these genes are chemokine signalling molecules and are amongst the most consistently regulated genes across all EBNA 2 and 3 microarray expression data sets. Expression of *CCL3L1/3* has been shown to be repressed by EBNA 3C in a BJAB background, repressed by EBNA 3A in an LCL background, activated by EBNA 3B and EBNA 3C in a BL31 background and activated by EBNA 2 in BL41, BJAB, LCL and Akata backgrounds (Chen et al. 2006; Maier et al. 2006; Spender et al. 2006; Zhao et al. 2006a; Zhao et al. 2006b; Lucchesi et al. 2008; Hertle et al. 2009; White et al. 2010; McClellan et al. 2012). Of the eight chemokine signalling genes repressed by EBNA 3C only *CXCL11* is not regulated in any other array. *CXCL10* and *CXCL11* are located in a gene cluster along with *CXCL9* and coordinate lymphocyte migration (Groom and Luster 2011).

Low levels of EBNA 3 binding to this locus was detected (Fig 4-7 A). We attempted to confirm EBNA 3C binding to these sites and investigate changes in histone modifications across this locus with and without the presence of EBNA 3C. Primer sets were designed to amplify DNA at each of the EBNA 3 binding sites (primer sets 1 and 3) and a region in between the two peaks where no EBNA 3 was shown to bind (primer set 2). We could not detect significant EBNA 3C binding at either site as QPCR signals obtained were similar to that of the EBNA 3C negative control cell lines and were significantly below the signal obtained at the ADAM binding site (Fig 4-7 B). Following publication of our study the Kempkes lab reported binding of EBNA 3A to an intergenic region between *CXCL9* and *CXCL10*. The binding site was marked by high levels of H3K27ac in the EBV infected LCL GM12878 in ChIP sequencing data from the ENCODE project. Both *CXCL9* and *CXCL10* had been previously identified as targets of EBNA 3A regulation in LCLs by gene expression microarray analysis. To determine whether EBNA 3C targeted *CXCL10* and *CXCL11* via this site we used the published primer sets used to detect EBNA 3C binding; R1, R2 and R3. ChIP-QPCR data from EBNA 3C expressing and control cell lines revealed that EBNA 3C does not associate with this site in a BJAB background (Fig 4-7 C). In contrast the Kempkes lab

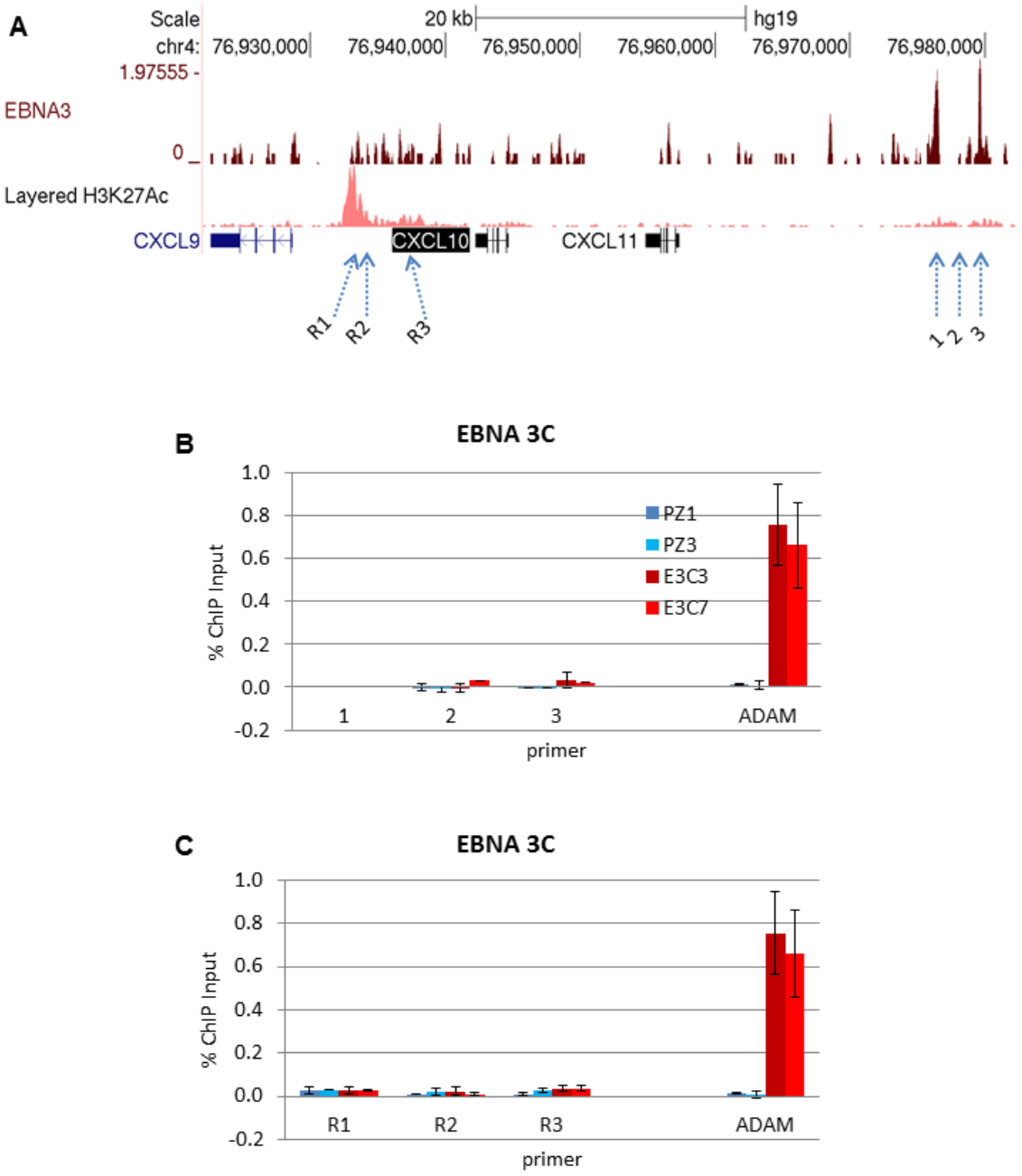


Figure 4-7 EBNA 3C does not bind to regulatory sites at the *CXCL* locus ChIP experiments conducted in the EBNA 3C positive cell lines E3C3, E3C7 and the EBNA 3C negative cell lines PZ1, PZ3. (A) EBNA 3 binding at the *ADAM* locus. Top panel; scale, EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Location of primer sets designed to amplify peaks from our ChIP sequencing data (1 and 3) and an intervening primer set (2) are shown. Location of primer sets designed to amplify a reported region of EBNA 3A binding (R1, R2 and R3) are also indicated). Positive control of ChIP QPCR signal at ADAM binding site is shown. Percentage input signals, after subtraction of signals of controls with no antibody, are expressed relative to the highest signal obtained. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 3C at *CXCL11* upstream sites (B) and *CXCL10* downstream sites (C).

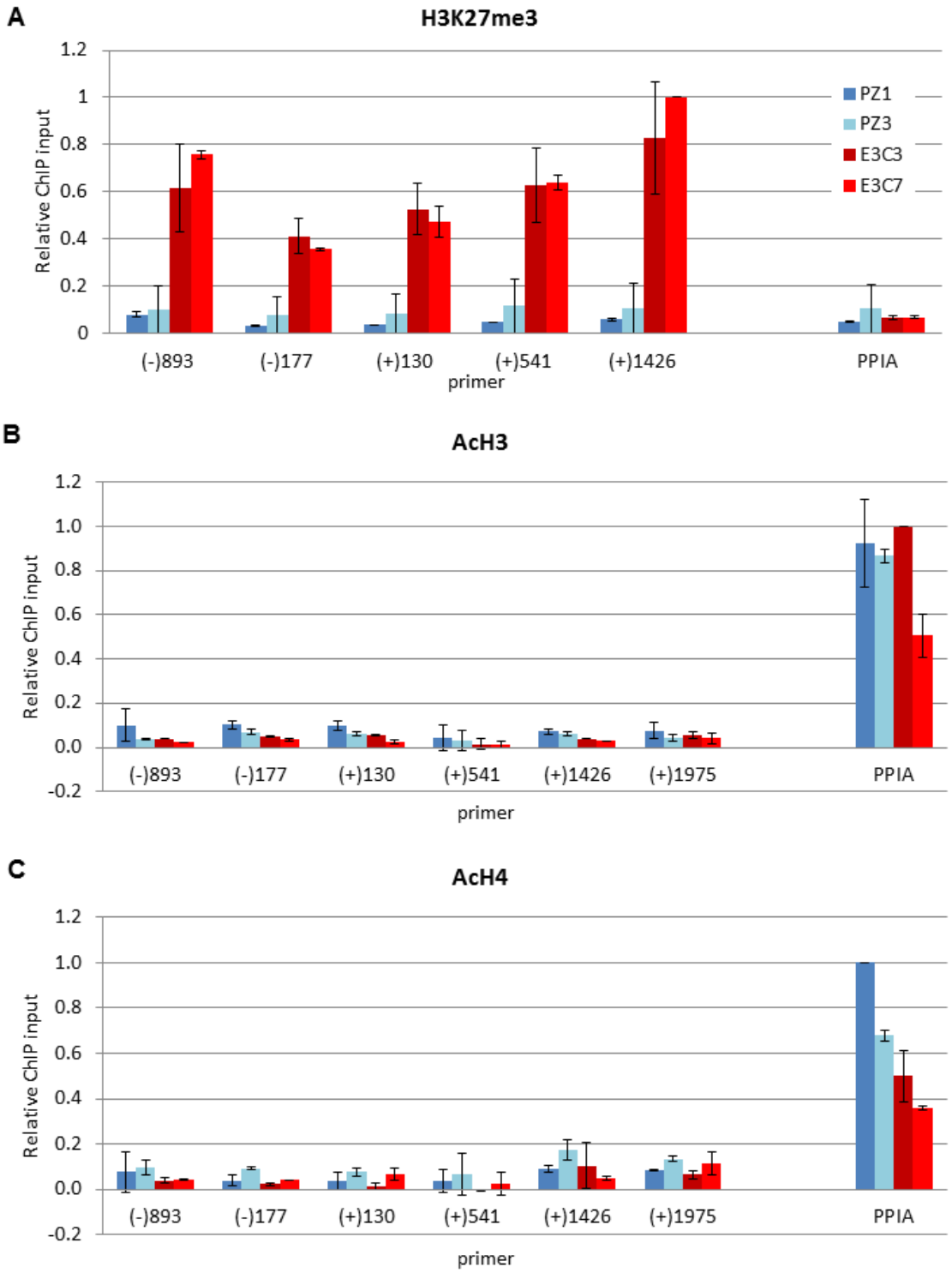


Figure 4-8 Epigenetic repression of *CXCL10* in cells expressing EBNA 3C ChIP experiments conducted in the EBNA 3C positive cell lines E3C3, E3C7 and the EBNA 3C negative cell lines PZ1, PZ3 at *CXCL10*. Numbers indicate the 5' end of the forward primers used for ChIP-QPCR analysis relative to the TSS. Control primers spanned the TSS of *PPIA*. Percentage input signals, after subtraction of signals of controls with no antibody, are expressed relative to the highest signal obtained. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for H3K27me3 (A), H3ac (B), H4ac (C).

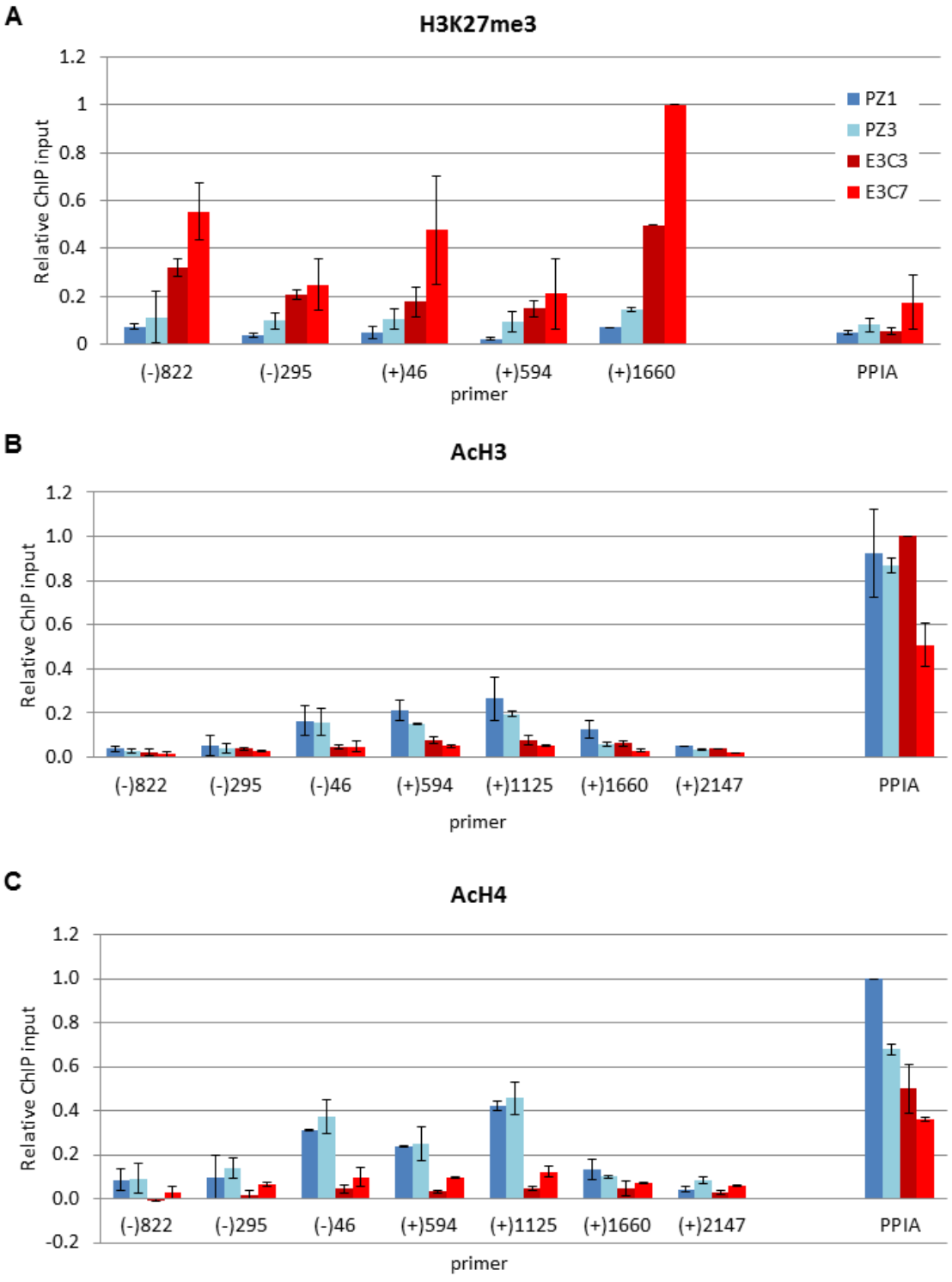


Figure 4-9 Epigenetic repression of *CXCL11* in cells expressing EBNA 3C ChIP experiments conducted in the EBNA 3C positive cell lines E3C3, E3C7 and the EBNA 3C negative cell lines PZ1, PZ3 at *CXCL11*. Numbers indicate the 5' end of the forward primers used for ChIP-QPCR analysis relative to the TSS. Control primers spanned the TSS of *PPIA*. Percentage input signals, after subtraction of signals of controls with no antibody, are expressed relative to the highest signal obtained. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for H3K27me3 (A), H3ac (B), H4ac (C).

observed a broad peak of EBNA 3A at this site with significant binding at all three primer sets (Harth-Hertle et al. 2013).

In the EBNA 3C expressing cell lines H3K27me3 was observed to be significantly higher than that of the control cell lines across the whole of *CXCL10* and 3' end of *CXCL11* indicative of gene repression by polycomb group proteins (Fig 4-8 A, Fig 4-9 A). *CXCL10* was not associated with the active transcription marks H3ac and H4ac in either control cells or EBNA 3C expressing cells (Fig 4-8 B, Fig 4-9 B). In contrast H3ac and H4ac were evident at the promoter of *CXCL11* in control cells and reduced in cells expressing EBNA 3C (Fig 4-8 B, Fig 4-9 B). As we could not detect EBNA 3C binding at this locus these data suggest that EBNA 3C is indirectly repressing transcription of *CXCL10* and *CXCL11*. However, changes in histone modifications suggests that the mechanism of indirect repression may be similar to that at directly bound genes such as *ITGA4*, *ITGB1* and at the *ADAM* locus.

4.6 Discussion

Our data demonstrate that *ITGA4*, *ITGB1*, *ADAM28* and *ADAMDEC1* identified as highly repressed EBNA 3C target genes are bound at proximal and distal sites by EBNA 3C. This implicates binding of EBNA 3C to these elements in repression of these genes.

Chemokines are an integral part of the immune response performing an important role in guiding the migration of immune response cells. Secreted chemokines and other alarmones can attract immune cells by increasing their adhesion to local endothelial cells via integrin signalling molecules such as integrin β2 and integrin αL (ITGAL)(collectively known as LFA1) interaction with ICAM1 (reviewed in (Newton and Dixit 2012)). It is therefore possible that the EBNA 2 and EBNA 3 proteins have a role in directly regulating these responses as the EBNA 2 and 3 proteins themselves are highly immunogenic. Despite this the CD4+ and CD8+ T cell response very effectively controls EBV infection *in vivo* although they cannot clear the virus and an infected host will remain infected for the rest of their life (reviewed in (Landais et al. 2005)). Chemokine and integrin signalling genes were known to be targets of EBNA 2 and 3 proteins prior to this investigation through gene expression microarray analysis. However the mechanism of regulation was completely unknown.

We previously identified chemokine and integrin signalling genes as some of the most significantly downregulated genes in EBV –ve EBNA 3C +ve B cell lymphoma BJAB compared to control cells. Using ChIP QPCR and guided by our ChIP sequencing data we were then able to confirm that EBNA 3C bound to sites proximal and distal to *ITGB1*, *ITGA4*, *ADAM28* and

ADAMDEC1 in these cells. This is the first demonstration of cellular gene targeting by EBNA 3C that correlates with gene regulation. Furthermore we were able to show that the repression of these genes was likely mediated through a combination of H3K27me3 (a mark associated with the polycomb group protein EZH2) and modulating histone acetylation. Subsequent experiments by Sarika Khasnis in our laboratory further validated the repression of these genes. cDNA levels in EBNA 3C expressing cell lines (E3C3, E3C4 and E3C7) were compared to control cell lines (PZ1, PZ2 and PZ3) revealing cDNA levels of all of studied genes here (as well as *CCL3* and *CCL4*) was reduced or ablated in the presence of EBNA 3C (McClellan et al. 2012).

CXCL10 and 11 play key roles in the attraction of leukocytes to sites of infection and inflammation. This property makes CXCL10 and 11 potent anti-tumour proteins as they can recruit cytotoxic T lymphocytes to malignant tissue (Sgadari et al. 1996; Hensbergen et al. 2005). CXCL10 and 11 are both ligands for CXCR3 chemokine receptor along with CXCL9, which is in the same gene cluster. CXCR3 is found on CD8+ cytotoxic T cells, CD4+ type-1 helper T cells, natural killer (NK) cells, NKT cells and subsets of B cells and its ability to interact with this many ligands likely reflects the complex nature of its various roles in various cell types. Indeed, CXCL10, CXCL11 and CXCL9 have been shown to have redundant, synergistic and even antagonistic effects on CXCR3 *in-vivo* (reviewed in (Groom and Luster 2011)). Chemotaxis assays conducted by Sarika Khasnis using the BJAB series used in this study revealed that cells expressing EBNA 3C (E3C3, E3C4 and E3C7) caused reduced migration of CXCR3 expressing cells compared to control cells (PZ1, PZ2 and PZ3)(McClellan et al. 2012). This implicates EBNA 3C mediated repression of *CXCL10* and *CXCL11* in immune evasion by EBV. The mechanism of regulation of *CXCL10* and *CXCL11* appears to be via indirect or long range mechanisms as no EBNA 3C peak could be verified at this locus. Two EBNA 3 peaks were detected upstream of *CXCL11* by ChIP sequencing in Mutu III cells which could represent either an EBNA 3A or 3B binding sites. *CXCL10* and *CXCL11* cDNA levels have been shown to be reduced in an EBV negative EBNA 3A expressing cell line compared to controls (McClellan et al. 2012). However, only one of these peaks passed the significance threshold of $1x10^{-7}$ and the other was not highly significant. Data from the Kempkes lab reported an EBNA 3A binding site in an intergenic region between *CXCL9* and *CXCL10* which was proposed to mediate the observed repression of these genes by EBNA 3A. We did not detect a binding site for EBNA 3 proteins at this site, which is likely to be a result of cell-type specific EBNA 3 binding. Furthermore we did not detect EBNA 3C association with this site in a BJAB background.

Together, these data suggest that repression of *CXCL10* and *CXCL11* by EBNA 3C in E3C3, E3C4 and E3C7 cell lines is an indirect effect of the presence of EBNA 3C in these cell lines. In this case it is possible that the peak of binding in the ChIP sequencing data represents a false positive and EBNA 3C is repressing this locus through other mechanisms. It is also possible that our ChIP-QPCR technique was incapable of detecting a site with such a low level of binding by ChIP sequencing data, and that at least some of the EBNA 3C mediated repression of *CXCL10* and *CXCL11* is via interactions with binding sites observed via ChIP sequencing. These data further implicate that EBNA 3C is epigenetically repressing *CXCL10* and *CXCL11* via an unknown mechanism in order to evade to CD8+ and CD4+ T cell immune response.

The integrin signalling pathway was also highlighted as an enriched pathway in our DAVID analysis of EBNA 3C downregulated genes. These were highlighted as part of the focal adhesion pathway but did not include ADAM28 despite the fact that this soluble protein is a ligand for α4β1 integrin. ADAMDEC1 was also not included, probably because very little is known about this protein and it has only recently been described as a soluble, active enzyme capable of proteolytic cleavage (Lund et al. 2013). This is a limitation of DAVID analysis as the pathway clustering only clusters and annotates well-described genes. Many human genes due to lack of study have a limited number of gene ontology terms ascribed to them meaning they will not appear in many clustering functions or pathways. The focal adhesion pathway is the subject of much research due to its role in cell migration and therefore tumour metastasis. The migration of immune cells through tissue is an essential role of combating infection and cancers. This is a highly regulated complex process and involves four major steps; rolling, activation, firm adhesion and transmigration. Only cells expressing the correct surface molecules can undergo this process, which in the case of T cells may involve activation via the T cell receptor or chemokines. Some of the key surface molecules involved in this process are LFA-1 (ITGAL-β2) and α4β1 which bind ICAM-1 and VCAM-1 respectively. Rolling occurs when immune cells acquire a weak adhesion to a cell surface (e.g. endothelial surface) and requires expression of selectins, chemokine receptors and integrins. This greatly increases the chances of encountering other ligands that further activate the immune cell resulting in firm adhesion. LFA-1 is especially important in this process. Transmigration is less well understood as it requires de-adhesion and re-adhesion to new cells. In this situation α4β1 has been reported to be both a positive and negative regulator of the LFA-1 ICAM-1 adhesion (reviewed in (Denucci et al. 2009)). LFA-1 has been previously described to be regulated by EBV upon infection mimicking B-cell activation, an effect that has been attributed to the activity of the viral membrane protein LMP1 (Wang et al. 1987; Wang et al. 1990a).

To determine the relevance of this pathway to EBV immortalisation it is important to consider when EBNA 2 and the EBNA 3 family are expressed *in vivo*. In a healthy host this appears only to occur upon initial infection when the naive B cell takes on an activated phenotype and migrates to the germinal centre (GC). Once there EBV assumes a latency II phenotype and the EBNA 2 and EBNA 3 proteins are repressed. It is therefore possible that the regulation of this pathway aids EBV infected B cells to assume an activated phenotype and migrate to the germinal centre. Furthermore regulation of integrin receptor signalling genes may play a role in immune evasion as T cell recognition of infected cells requires adhesion via VCAM-1, either by evading contact or breaking contact once it has been established (reviewed in (Wu 2007)). The data presented here suggests that the regulation of focal adhesion and integrin genes by EBNA 3C is established through epigenetic mechanisms.

Polycomb mediated gene repression has been shown to cause CpG methylation which results in long lasting gene repression, suggesting that the repressive effects of EBNA 3C may persist long after the transition of EBV from a latency III to a latency II phenotype. Such mechanisms of persistent epigenetic repression by EBNA 3C have already been reported at *Bim* (Paschos et al. 2009). If this is the case at these genes then EBNA 3C repression may persist into the GC and beyond. In the GC many of the survival signals required to avoid apoptosis are provided by contact with T cells and follicular dendritic cells which is mediated through VCAM-1 integrin interactions (Taylor et al. 1999; Yoon et al. 2014). It is important to note that whilst EBNA 3C binding to the loci described does correlate with repression, enrichment of H3K27me3 and a reduction of H3ac and H4ac, this does not prove that EBNA 3C is modifying histones by recruiting other gene silencing proteins. It is also not possible from these data to determine that the changes in H3K27me3, H3ac and H4ac precedes or proceeds EBNA 3C association with the human genome. However, given that EBNA 3C is a potent repressor of transcription when tethered to DNA and that it can associate with chromatin remodellers such as CTBP proteins; it is very likely that this is indeed the first documented evidence of EBNA 3C binding to the human genome and directly repressing nearby genes through the recruitment of cellular factors (Bain et al. 1996; Touitou et al. 2001).

Whilst studying the effect of EBNA 3C alone is hugely informative it is clear that *in vivo* the story is far more complex. The genes targeted by EBNA 3C in this chapter have also been shown to be regulated by a variety of other EBNA proteins in other gene expression microarray data sets and cell lines. Whilst *CXCL11* has only been shown to be downregulated by EBNA 3C in the BJAB background EBNA 3C expressing cells compared to control cells, *CXCL10* has been

shown to be repressed by EBNA 3A in LCLs, upregulated by EBNA 3B in LCLs and BL cell lines, upregulated by EBNA 3C in LCLs and upregulated by EBNA 2. A similar mosaic of up and down regulation by the EBNA 3 family is observable in the other genes examined in this chapter (Table 4-3). *ADAM28* and *ADAMDEC1* are exceptions to this as the expression of both been shown to be repressed by EBNA 3C and EBNA 3A and activated by EBNA2, suggesting that cross talk between EBNA 2 and 3 proteins is context specific. It is also probably that the wide variety of cell backgrounds in which these data are obtained from play a role in gene regulation by EBNA 2 and 3 proteins. This likely represents the requirement of EBV to fine tune host gene expression with a remarkably limited set of latent genes. How the EBNA 2 and 3 proteins perform these functions, if they act synergistically, antagonistically or redundantly and if all share similar mechanisms of regulation remains poorly understood.

Gene Symbol	EBNA 3 or EBNA 2 array and cell background												
3C (BJAB)	3A (LCL)	3A (BL31)	3B (LCL)	3B (BL31)	3B/3C (LCL)	3C (LCL)	3C (BL31)	3A/3B/3C (BL31)	ER/E2 (BJAB)	ER/E2 (BL41)	ER/E2 (LCL)	ER/E2 Type 1 (AK31)	ER/E2 Type 2 (AK31)
ADAM28													
CXCL11													
ITGB1													
ITGA4													
ADAMDEC1													
CXCL10													
CCL3													
CCL3L1													
CCL4													

Table 4-3 Comparison of published gene expression microarray data Genes from the top 20 most downregulated in a BJAB EBNA 3C expressing cell compared to control cells which have also been reported to be regulated by EBNA 2 or 3 proteins are displayed (3C BJAB column). The cell line and investigated EBNA gene for each array is displayed under EBNA 3 or EBNA2 array and cell background. Green indicates that in this cell background the EBNA 2 or 3 proteins was reported to repress the corresponding gene. Red indicates up regulation by the EBNA 2 or 3 proteins. References: 3A (LCL),(Johannsen et al. 1995; Hertle et al. 2009);3A (BL31), 3B (BL31), 3B (LCL), 3C (BL31) and 3A/3B/3C (BL31), (White et al. 2010); 3B/3C (LCL) (Chen et al. 2006); 3C (LCL) (Zhao et al. 2006a); ER/E2 (BL41) and ER/E2 (BJAB) (Maier et al. 2006); ER/E2 (LCL) (Spender et al. 2006); ER/E2 type 1 (AK31) and ER/E2 type 2 (AK31), (Lucchesi et al. 2008). Data from references (Spender et al. 2002) and (Zhao et al. 2006b) were also compared but not displayed as no matches were found.

5. Investigating EBNA 3A, 3B and 3C binding to key target genes and the role of shared EBNA 2 and 3 binding sites in gene regulation

Many cellular genes are transcriptionally regulated by more than one of the EBNA 3 family in cooperative and antagonistic ways (Paschos et al. 2009; White et al. 2010). Furthermore, it is well documented that EBNA 3 family proteins can both antagonise and cooperate with EBNA 2 to regulate transcription and ChIP sequencing data indicates that EBNA 2 and 3 coincident binding sites are common in the human genome. The mechanism EBNA 2 and 3 protein transcriptional cross talk is largely unknown.

EBNA 2 and EBNA 3 binding to RBPJκ is mutually exclusive but it is not known if whether EBNA 2 and EBNA 3C bind PU.1 simultaneously (Johannsen et al. 1995; Zhao and Sample 2000) (Grossman et al. 1994; Henkel et al. 1994; Le Roux et al. 1994; Waltzer et al. 1994; Robertson et al. 1995). It is also unknown if EBNA 3A, 3B and 3C can bind RBPJκ or PU.1 simultaneously. Since our ChIP-sequencing analysis had used a pan specific antibody against EBNA 3A, 3B and 3C we did not know whether these proteins associated with the same cellular DNA elements.

We set out to investigate whether sites at specific target genes were bound by EBNA 3A, EBNA 3B, EBNA 3C and the role of EBNA 2 and EBNA 3 shared sites

5.1 Validation of antibodies specific to EBNA 3A, EBNA 3B and EBNA 3C and differential EBNA 3 binding at *ADAM* and *Bim*

We first set out to validate specific antibodies against EBNA 3A, EBNA 3B and EBNA 3C. Immunoprecipitations using ChIP conditions and buffers were carried out on non-crosslinked chromatin from the BL31 BAC cell series. We used antibodies against EBNA 3A (Exalpha, sheep polyclonal), EBNA 3B (Exalpha) sheep polyclonal and EBNA 3C (E3CD8 mouse monoclonal (Maunders et al. 1994)) and analysed precipitated proteins by western blotting. BL31 cells infected with BAC derived virus with a normal EBV genome is referred to as referred to as wild type (BL31 wtBAC) and expresses a full complement of latency III gene products. BL31 cells infected with a virus which lacks *EBNA 3A*, *EBNA 3B* or *EBNA 3C* are described as BL31 3AKO, BL31 3BKO and BL31 3CKO cell lines respectively. Cell lines also exist in which BL31 are infected with BAC derived virus that has *EBNA 3C* deleted and then restored and are described as BL31 3CREV (revertant).

A no antibody control was used in the immunoprecipitation as a negative control. Western blots from whole cell lysates revealed that EBNA 3A, EBNA 3B and EBNA 3C could not be

detected in the EBV negative BL31 line (Fig 5-1 A, D, G first lane). Analysis of whole cell lysates using antibodies against EBNA 3A, 3B and 3C confirm that EBNA 3A, EBNA 3B and EBNA 3C are knocked out in the appropriate cell lines but present in the BL31 wtBAC (Fig 5-1 A,D,G); furthermore EBNA 3B and 3C were present in BL31 3AKO (Fig 5-1 A), EBNA 3A and 3C were present in BL31 3BKO (Fig 5-1 D) and EBNA 3A and 3B were present in BL31 3CKO (Fig 5-1 G). Western blots probed with an anti-EBNA 3A antibody shows that EBNA 3A was observed in the BL31 wtBAC cell line but not in BL31 3AKO (Fig 5-1 B). Similar results were obtained from the same immunoprecipition samples probed with anti-EBNA 3B (Fig 5-1 E) or anti-EBNA 3C (Fig 5-1 H) antibodies. If the EBNA 3A antibody used in the immunoprecipitation had cross reacted with either EBNA 3B or EBNA 3C then a band would be detected in these blots. This confirms that the EBNA 3A polyclonal antibody does not cross react with either EBNA 3B or EBNA 3C. Analysis using anti EBNA 3B and 3C antibodies shows that both the EBNA 3B polyclonal and EBNA 3C E3CD8 monoclonal antibodies are also specific to EBNA 3B and EBNA 3C respectively in immunoprecipitation experiments (Fig 5-1 E, F, H, I). These data demonstrates that all three antibodies do not cross react with other members of the EBNA 3 family and should specifically precipitate DNA bound by each EBNA 3.

We sought to validate EBNA 3A, EBNA 3B or EBNA 3C specific DNA precipitation and elucidate potential differential binding of EBNA 3 proteins by examining a known EBNA 3 target sites from our ChIP sequencing data. For this we analysed the *ADAM* locus. It has been reported that regulation of genes by EBNA 3 proteins is cell type specific. For example in gene expression microarrays in EBNA 3B knock out LCL and BL31 cell lines different sets of genes have been reported to be regulated compared to wild type cell lines (White et al. 2010). This prompted us to conduct ChIP experiments in the BL line Mutu III and in PER253 LCLs, both of which express a full panel of latency III gene products. PER253 LCLs are derived from a healthy donor (code PER253) which were *ex vivo* infected with EBV resulting in LCL outgrowth and immortalisation. ChIP using the specific EBNA 3A, EBNA 3B and EBNA 3C antibodies was conducted and analysed using QPCR primer sets at the peak of binding and flanking regions as described previously. We also analysed ChIP DNA using primer sets for *PPIA* as a negative control and at *CTBP2* as a positive control. We had subsequently validated that the *CTBP2* intragenic binding site was bound by all three EBNA 3 proteins and served as a good positive control for EBNA 3A, EBNA3B and EBNA 3C binding. ChIP QPCR data at the *ADAM* intragenic peak revealed that EBNA 3A (Fig 5-2 A) and EBNA 3C (Fig 5-2 C) were capable of significantly binding to this site in Mutu III cells but that EBNA 3B was not (Fig 5-2 B).

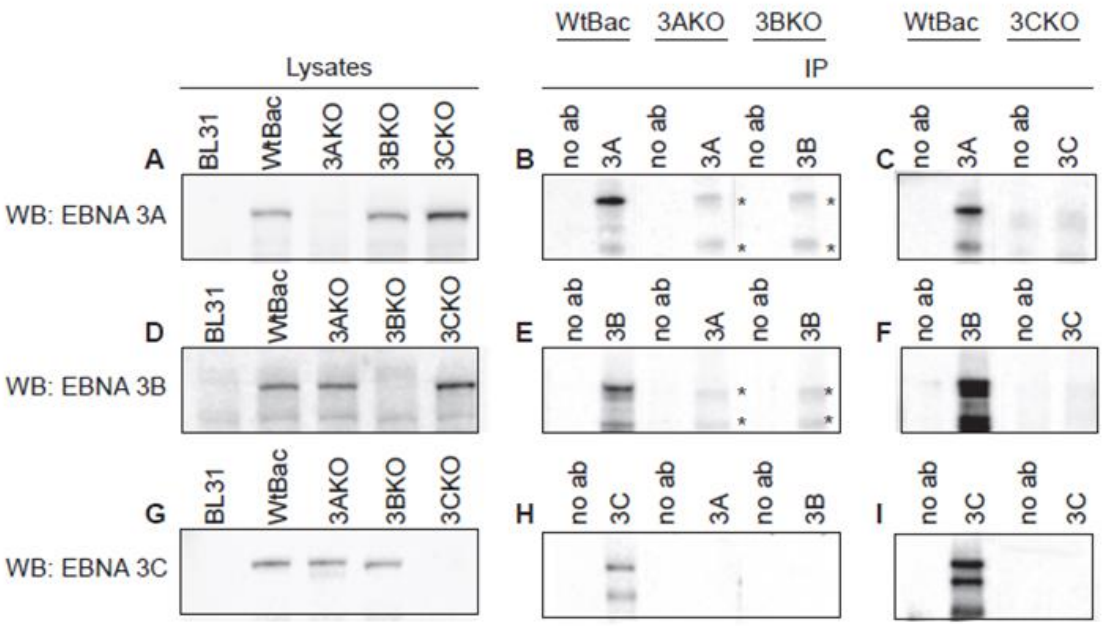


Figure 5-1 Validation of antibodies specificity for anti-EBNA 3A, anti-EBNA 3B and anti-EBNA 3C. Western blot analysis of whole cell lysates from BL31, BL31 wtBAC, BL31 3AKO, BL31 3B KO and BL31 3CKO probed with anti 3A (A), 3B (D) and 3C (G) antibodies. Immunoprecipitations conducted under ChIP conditions using anti EBNA 3A (exalpha, sheep polyclonal), anti EBNA 3B (exalpha, sheep polyclonal), anti EBNA 3C (E3CD8, mouse monoclonal and no antibody negative control were subjected to western blot analysis. These were then probed for EBNA 3A (B,C), EBNA 3B (E,F) and EBNA 3C (H, I). Asterisks denote blots where signal obtained is similar to background binding found in the corresponding knock out cell line.

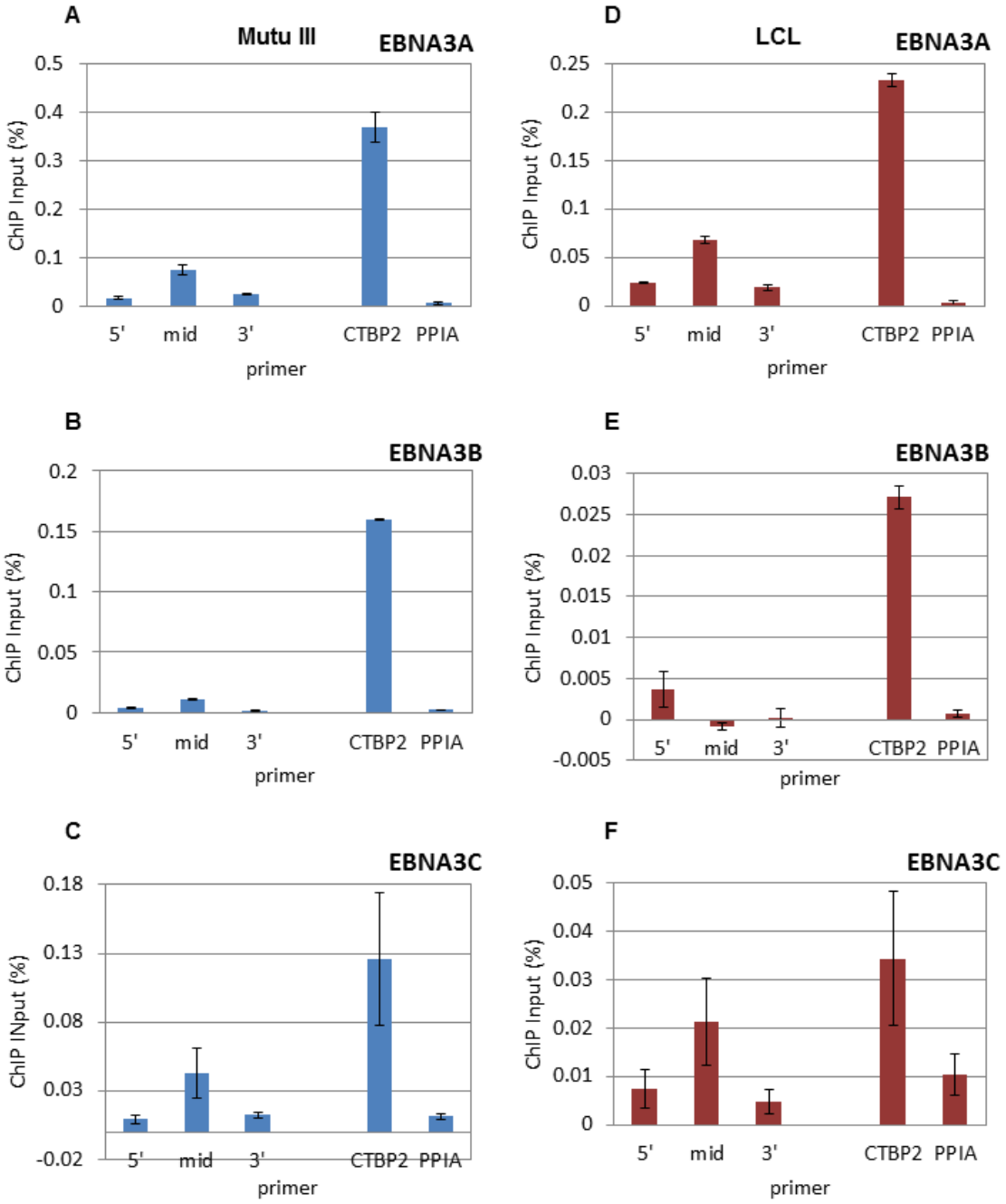


Fig 5-2 EBNA 3A and 3C bind to an intragenic peak at the *ADAM* locus in Mutu III cells and PER253 LCL. ChIP experiments conducted in Mutu III cells and PER153 LCLs. Primer sets for the peak of binding (mid) and adjacent regions (5',3') have been previously described. Control primers spanned the TSS of *PPIA*. Signal at the CTBP2 peak is displayed as a positive control. ChIP data is displayed as percentage input signals, after subtraction of signals of controls with no antibody. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 3A (A), EBNA 3B (B) and EBNA 3C (C) in Mutu III cell lines. ChIP-QPCR data for EBNA 3A (D), EBNA 3B (E) and EBNA 3C (F) in PER253 LCL.

Association with EBNA 3A (Fig 5-2 D) and EBNA 3C (Fig 5-2 F) was also detectable in the LCL cell line whilst EBNA 3B signal was below that of the negative control *PPIA* (Fig 5-2 E). It is not possible to examine the percentage of input DNA precipitated between different antibodies as they will have different affinities for their antigen. Interestingly we noted that *ADAM28* and *ADAMDEC1* have both been reported to be repressed by EBNA 3A and EBNA 3C but not by EBNA 3B (Table 4-3).

We also wished to analyse differential binding of EBNA 3 proteins at *Bim*, one of the best characterised EBNA 3A and EBNA 3C targets for transcriptional regulation. Both EBNA 3A and EBNA 3C have been shown to promote epigenetic silencing of *Bim* via unknown mechanisms. We analysed ChIP DNA precipitated by EBNA 3A, EBNA 3B and EBNA 3C by QPCR at primer sets in a small promoter proximal EBNA 3 peak at *Bim* (also known as *BCL2L11*). We also used primer sets either side of the peak as indicated and included positive and negative controls for EBNA 3 binding. This revealed that EBNA 3A and EBNA 3C also significantly associated with this binding site in both a Mutu III (Fig 5-3 B, D) and LCL (Fig 5-3 E, G) background. EBNA 3B was found to not bind to this site (Fig 5-3 C, F). To date there is no evidence of regulation of *Bim* by EBNA 3B suggesting that differential binding of EBNA 3 proteins to target genes is important to their function in co-regulating those genes.

5.2 Identifying key target genes for analysis

As previously discussed in Chapter 3, 25% of binding sites identified in our ChIP sequencing study were bound by both EBNA 2 and EBNA 3 proteins. We wished to investigate how the interplay between EBNA 2, EBNA 3A, EBNA 3B and EBNA 3C at a specific gene modulated transcription. To select targets for further analysis we analysed the 300 most significant peaks of binding in the EBNA 3 data set. The UCSC genome browser was used to investigate each peak and record the nearest gene irrespective of distance or direction of the TSS, providing a list of 300 most significantly bound genes. We also recorded all other peaks which were most proximal to the 300 genes including the height of the largest peak for EBNA 2 and if EBNA 2 and EBNA 3 peaks were found promoter proximally or at a distal site. To make the list more manageable we arbitrarily removed all genes that were not bound by an EBNA 3 peak with a height less than 10 leaving 63 genes (Table 5-1 and Table 5-2). This also removed the less significantly bound genes from further analysis.

DAVID analysis revealed that the most enriched gene ontology (GO) terms in this list were regulation of transcription, phosphorylation and leukocyte activation (Table 5-3). The GO

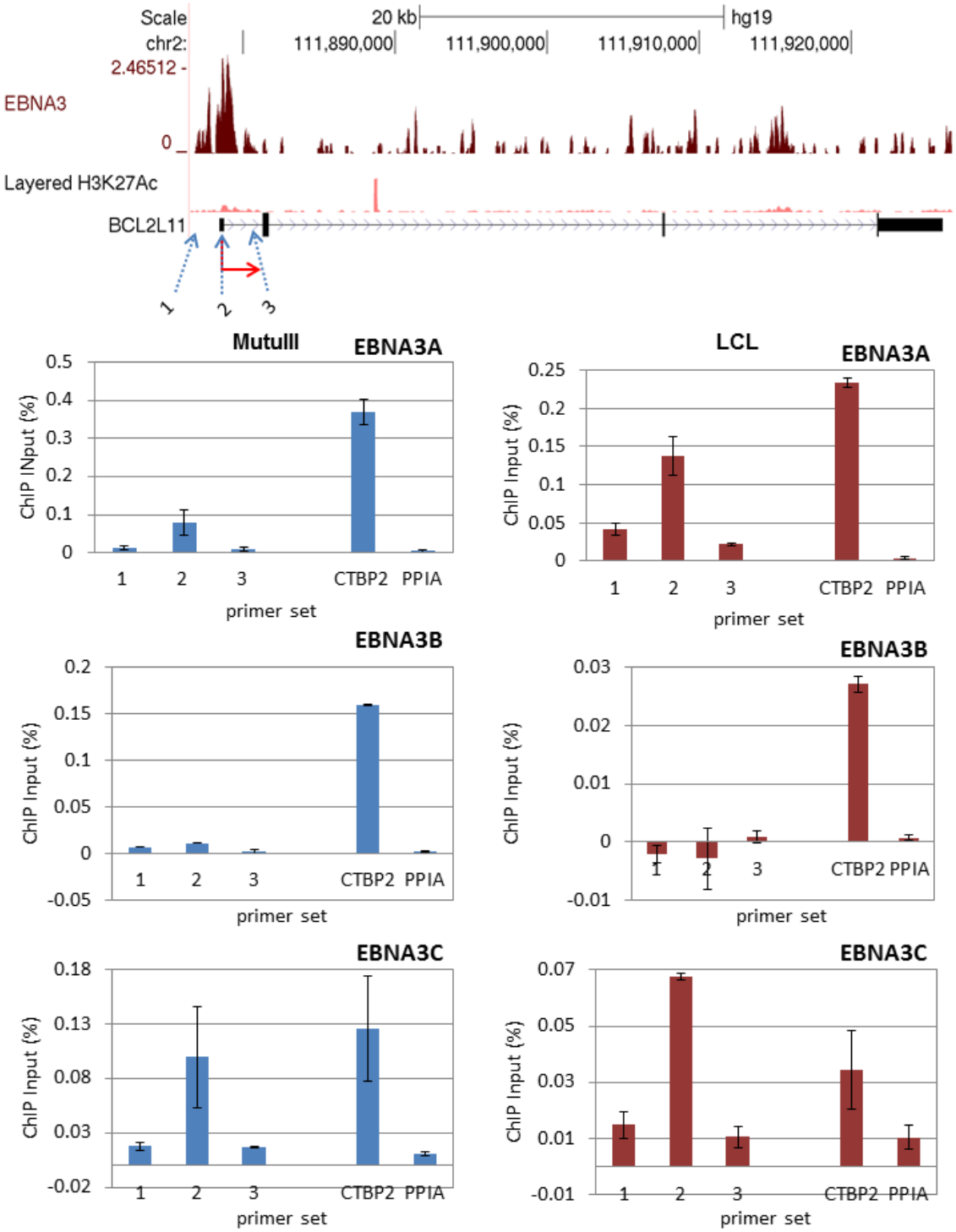


Fig 5-3 EBNA 3A and EBNA 3C bind to a promoter proximal site at *Bim* in Mutu III and PER253 LCL. ChIP experiments conducted in Mutu III cells and PER153 LCLs. (A) EBNA 3 binding at the Bim locus. Top panel; scale, EBNA 2 and EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Location of primer sets for the peak of binding (2) and adjacent regions (1,3) are annotated. Control primers spanned the TSS of *PPIA*. Signal at the CTBP2 peak is displayed as a positive control. ChIP data is displayed as percentage input signals, after subtraction of signals of controls with no antibody. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 3A (B), EBNA 3B (C) and EBNA 3C (D) in Mutu III cell lines. ChIP-QPCR data for EBNA 3A (E), EBNA 3B (F) and EBNA 3C (G) in PER253 LCL.

genes	EBNA3 prox	EBNA2 prox	EBNA3 dist	E3 count	EBNA2 dist	E2 count
AFF3	33	5	11	6	22	3
ARID1B		6	15	9	22	16
ARID5B			10	3	3	3
ATP6V1A			11	1	2	2
BCL6			14	11	22	10
BCL9L			31	7	5	10
BMP3			15	1	3	2
CDK5R1			24	3	17	5
CREB5	3	2	19	9	21	14
CRTC3			15	6	13	8
CSF2			13	3	2	2
CTBP2			24	4	11	3
CXCR4		7	28	15	12	22
CXCR5	4	2	31	7	5	10
CXXC1			22	5	1	1
DRD3			10	3	7	2
ECE1	2	11	15	4	17	12
ENPP3	8	1	13	6	3	4
ERC1			12	1		
FOXN3	2		16	4	1	1
FOXP1	10	6	8	9	21	21
GDF6			20	1		
HDAC4		2	13	8	33	21
HHEX			27	4	4	5
HNRPLL			19	5	9	3
IRF2	3	3	19	14	22	16
ITGAL	23	21		3		1
ITK			20	3	12	4
KIF13B			12	4	4	6
KLF12	2	2	14	3	13	8
KLF2		2	32	6	4	7

Table 5-1 List of genes bound by an EBNA 3 peak with a height of 10 or over. MACS peak data for EBNA 3 was sorted by p-value in Microsoft EXCEL and the 300 most significantly bound genes were manually investigated on the UCSC genome browser. Peak height was rounded to the nearest integer. Peak height of the most significant distal (>2kb from TSS) peak as well as any proximal peaks for both EBNA 3 and EBNA 2 were recorded. The height of any promoter proximal peak at a gene is displayed for EBNA 3 and EBNA 2 in the leftmost two columns. The height of the most significantly bound EBNA 3 distal peak and the number of other EBNA 3 peaks at this gene are recorded in the middle two columns. The height of the most significantly bound EBNA 2 distal peak and the number of other EBNA 2 peaks at this gene are recorded in the rightmost two columns. Cells are conditionally formatted such that that darker the shade of red or green the greater the peak height or number of other peaks at this gene for EBNA 3 and EBNA 2 respectively. Table is sorted alphabetically and split into two due to the number of genes (see Table 5-2)

genes	EBNA3 prox	EBNA2 prox	EBNA3 dist	E3 count	EBNA2 dist	E2 count
LMTK2			15	5	8	4
MAML3	17	2	3	3	7	6
MAP3K14			14	2	5	5
MAST4			10	5	5	7
MED13L	16		1	2	8	7
MET			21	2		
MYCN			15	11	14	6
NDUFS4			23	1		
NTRK2			12	1	2	2
PHF21A			25	1	29	7
PITX1	2		23	3		
RBM9			12	4	6	4
RUNX2			13	7	7	7
SIK3		1	21	2	4	4
SMAD3		3	17	7	28	16
SPTBN1			20	6	27	9
SSH2	10	22	4	5	22	9
STAT4		4	12	4	4	4
STK39			14	1		
SYK		2	10	7	5	15
TBL1XR1	7	2	22	8	3	3
TGFA			16	3		
TGFBR2			15	5	3	1
TWIST2			26	3		
ULBP3			13	2	2	3
USP22			16	2	2	1
WEE1			27	6	28	2
WWP2			10	2		
ZMIZ1			18	5	6	12
ZNF580			16	2	3	1
ZNF581			16	2	3	1
ZZZ3	2	1	19	3	1	1

Table 5-2 List of genes bound by an EBNA 3 peak with a height of 10 or over. MACS peak data for EBNA 3 was sorted by p-value in Microsoft EXCEL and the 300 most significantly bound genes were manually investigated on the UCSC genome browser. Peak height was rounded to the nearest integer. Peak height of the most significant distal (>2kb from TSS) peak as well as any proximal peaks for both EBNA 3 and EBNA 2 were recorded. The height of any promoter proximal peak at a gene is displayed for EBNA 3 and EBNA 2 in the leftmost two columns. The height of the most significantly bound EBNA 3 distal peak and the number of other EBNA 3 peaks at this gene are recorded in the middle two columns. The height of the most significantly bound EBNA 2 distal peak and the number of other EBNA 2 peaks at this gene are recorded in the rightmost two columns. Cells are conditionally formatted such that that darker the shade of red or green the greater the peak height or number of other peaks at this gene for EBNA 3 and EBNA 2 respectively. Table is sorted alphabetically and split into two due to the number of genes (see Table 5-1)

enriched biological process	number of genes	% of total genes	P-value	Benjamini
regulation of transcription	36	60	1.2E-11	1.3E-08
phosphorus metabolic process	22	36.7	3.0E-10	1.7E-07
phosphate metabolic process	22	36.7	3.0E-10	1.7E-07
phosphorylation	20	33.3	5.5E-10	2.1E-07
negative regulation of transcription	16	26.7	7.3E-10	2.1E-07
protein amino acid phosphorylation	18	30	2.0E-09	4.5E-07
negative regualtion of gene expression	16	26.7	2.6E-09	5.1E-07
negative regulation of transcription, DNA-dependent	14	23.3	3.2E-09	5.3E-07
negative regulation of nuceotide metabolic process	16	26.7	3.3E-09	4.8E-07
negative regulation of RNA metabolic process	14	23.3	3.9E-09	5.0E-07
negative regulation of nitrogen compounf metabolic process	16	26.7	3.6E-09	4.6E-07
leukocyte activation	12	20	6.3E-09	6.6E-07

Table 5-3 DAVID analysis of most significantly EBNA 3 bound genes All 63 genes from table 5-1 and 5-2 were entered into DAVID as an official gene symbol list. GOTERM_BP_FAT (biological process) enriched are shown and the count of the number of genes from the input in each pathway, the percentage of total genes this represents and measures of statistical significance are displayed. The p-value is a modified Fisher Exact P-Value, for gene-enrichment analysis. It ranges from 0 to 1. A Fisher Exact P-Value = 0 represents perfect enrichment. Usually a P-Value equal or smaller than 0.05 is considered strongly enriched in the annotation categories.

category for phosphorylation included many genes involved in cell cycle processes and another less significantly enriched GO term for the list was cell proliferation which 16/63 genes were annotated for (p-value $9.3x10^{-7}$, benjamini $3.6x10^{-5}$). To select genes for further analysis we discounted genes that were not bound by both EBNA 3 and EBNA 2 peaks with a heights of 10 or more. From this list we selected genes that were annotated in one of the GO categories of regulation of transcription, phosphorylation/cell proliferation and leukocyte activation. We then examined the distribution of peaks at these genes, whether a gene was bound by promoter proximal, distal or intragenic peaks. We then chose the only set of three genes that allowed us to investigate each GO term category and each peak distribution; *CTBP2*, *WEE1* and *ITGAL*.

5.3 Coincident binding of EBNA 2, 3A, 3B and 3C at an intragenic peak at the *CTBP2* gene

C-terminal binding proteins CTBP1 and CTBP2 (collectively known as CTBP) play an important role in gene repression and are implicated in tumourigenesis due to their ability to repress a number of tumour suppressor genes (Chinnadurai 2009). Interestingly EBNA 3A and EBNA 3C-mediated epigenetic repression of $p16^{INK4A}$ has been shown to be mediated by CTBP binding (Skalska et al. 2010). The repression of $p16^{INK4A}$ was found to be the main barrier to EBV induced proliferation as LCLs lacking $p16^{INK4A}$ could still proliferate with a conditional EBNA 3C switched off (Skalska et al. 2013). EBV mediated regulation of *CTBP2* transcription however has not been previously reported. ChIP sequencing identified a single, large intragenic peak of EBNA 2 and EBNA 3 binding between the second and third exons of the longer isoforms of *CTBP2* (Fig 5-4 A). To validate EBNA 2 binding to this site and determine which of the EBNA 3 proteins were bound, ChIP experiments were carried out using specific antibodies for EBNA 2, 3A, 3B and 3C in the BL cell line Mutu III and the EBV immortalised LCL PER253. We used the promoter of the housekeeping gene peptidylprolyl isomerase A (*PPIA*) as a control site where no EBNA proteins were bound. Primers were designed to amplify DNA at the peak of binding of both EBNA 2 and EBNA 3 (primer set 2) and flanking primers (1 and 3) were designed ~1kb away from this the region of most significant binding.

QPCR data obtained using these primers showed that all three EBNA 3 proteins associated with this site in Mutu III (Fig 5-4 C,D,E) and PER253 LCL (Fig 5-5 C,D,E). Binding of EBNA 2 to this site was validated in the Mutu III cell line (Fig 5-4 B) and we also detected binding in PER253 LCL (Fig 5-5 B). Binding signals were significantly above both the local control (flanking primers)

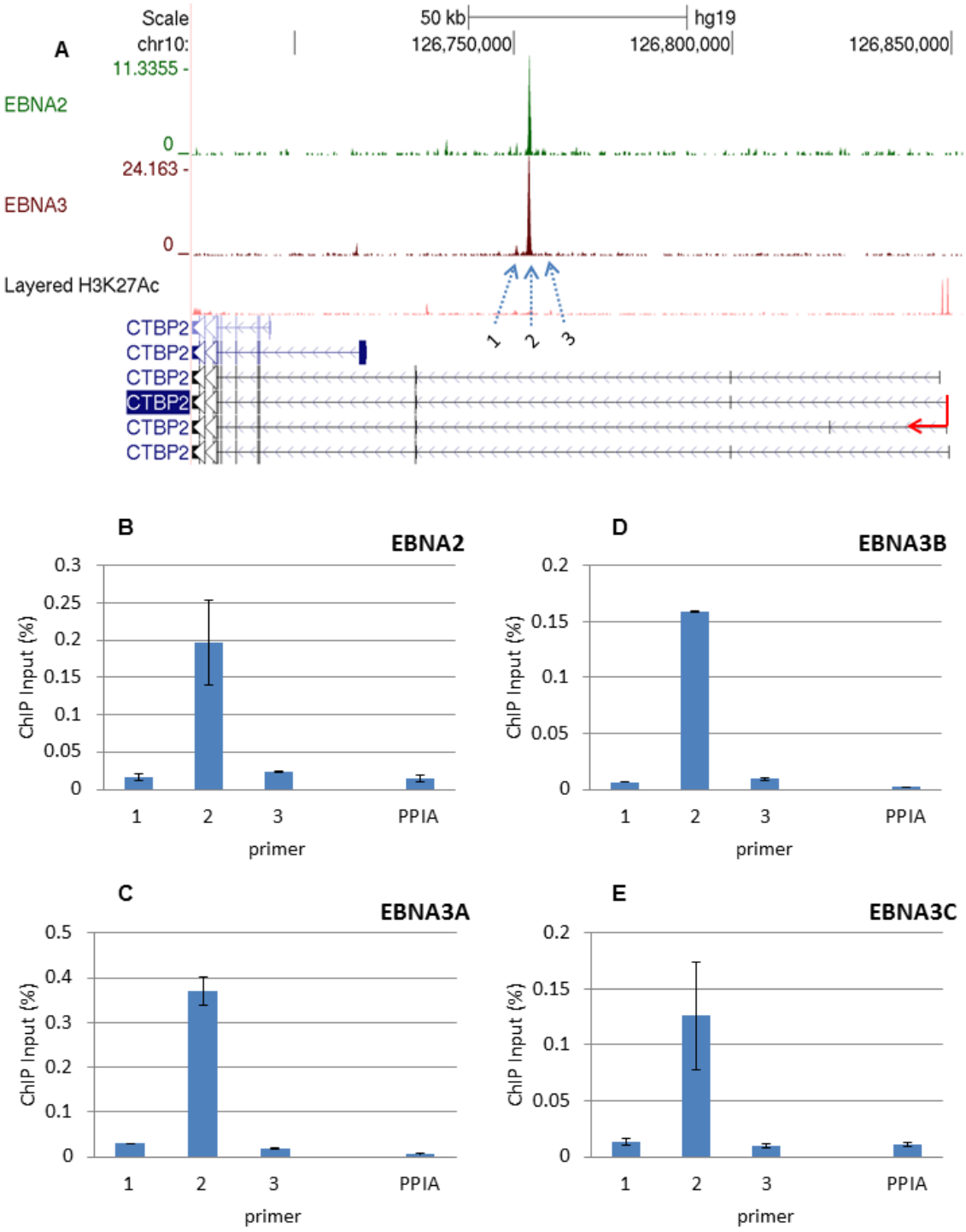


Figure 5-4 EBNA 2, EBNA 3A, EBNA 3B and EBNA 3C all bind to an intragenic peak at *CTBP2* in Mutu III cells. ChIP experiments conducted in Mutu III cells.(A) EBNA 2 (green) and EBNA 3 (red) binding at the *CTBP2* locus. Top panel; scale, EBNA 2 and EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Location of primer sets for the peak of binding (2) and adjacent regions (1,3) are annotated. Control primers spanned the TSS of PPIA. ChIP data is displayed as percentage input signals, after subtraction of signals of controls with no antibody. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 2 (B), EBNA 3A (C), EBNA 3B (D) and EBNA 3C (E).

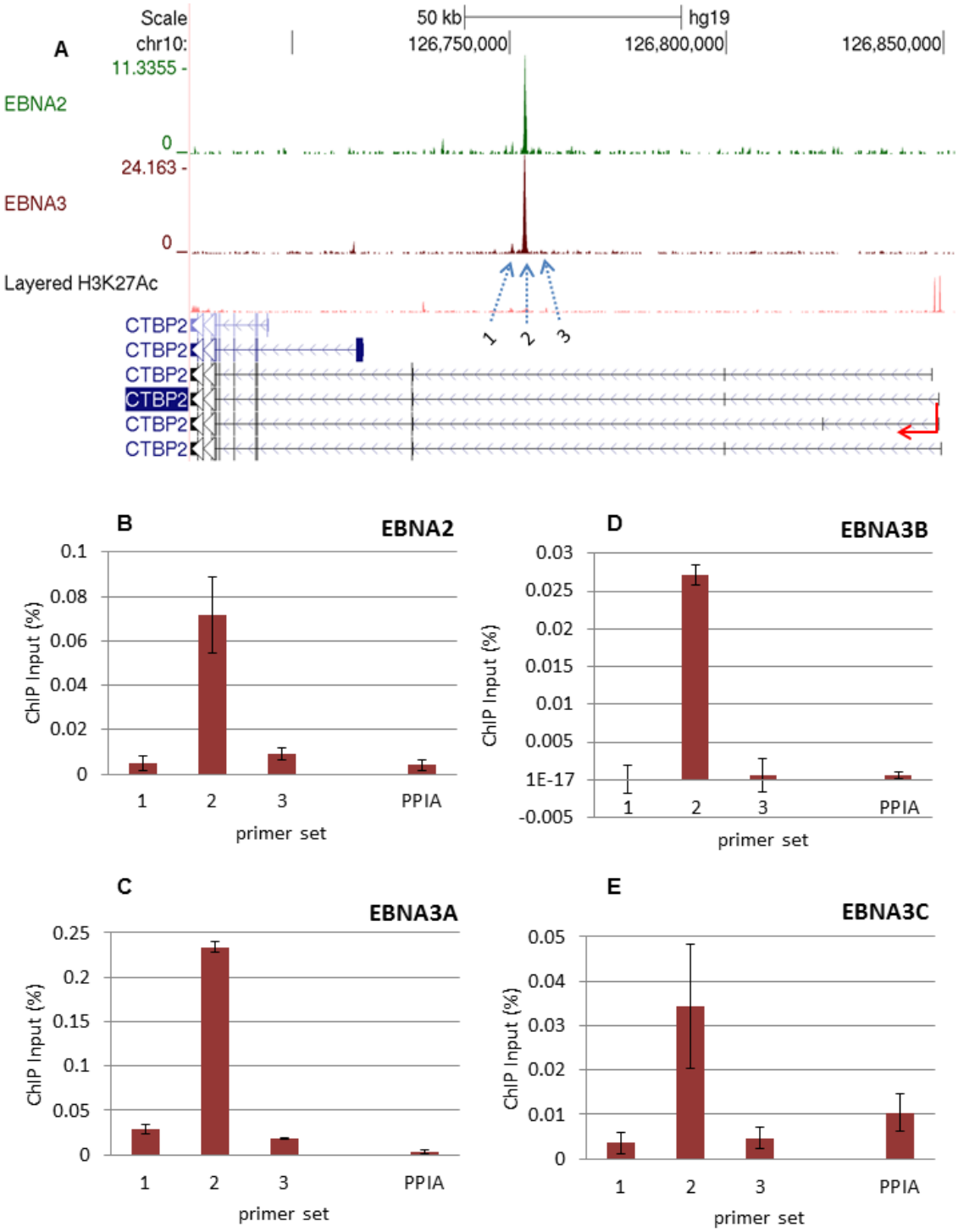


Figure 5-5 EBNA 2, EBNA 3A, EBNA 3B and EBNA 3C all bind to an intragenic peak at *CTBP2* in PER253 LCL. ChIP experiments conducted in PER253 LCL.(A) EBNA 2 (green) and EBNA 3 (red) binding at the *CTBP2* locus. Top panel; scale, EBNA 2 and EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Location of primer sets for the peak of binding (2) and adjacent regions (1,3) are annotated. Control primers spanned the TSS of PPIA. ChIP data is displayed as percentage input signals, after subtraction of signals of controls with no antibody. Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 2 (B), EBNA 3A (C), EBNA 3B (D) and EBNA 3C (E).

and at *PPIA* in all QPCR experiments. Since *CTBP2* has not previously been shown to be regulated by EBNA 2 or EBNA 3s we next investigated if binding correlated with changes in gene expression. Initial analysis was performed to determine which isoform is transcribed in B cells. In order to find a cell line in which *CTBP2* was expressed we used (http://www.epstein-barrvirus.org.uk/), a resource compiled by Dr Rob White (Imperial college) that stores gene expression microarray data for EBV proteins. Interestingly, we found evidence of regulation of *CTBP2* by EBNA 3A. A previous analysis of EBNA 3A regulated genes using wild type and EBNA 3AKO LCLs did not originally identify *CTBP2* as a regulated gene (Hertle et al. 2009). The original study obtained data from a full EBNA 3A mutant (EBV-E3AmtB) and a truncated mutant (EBV-E3AmtA), which were combined for the final analysis. These two data sets were subsequently analysed separately by Dr Rob White which revealed that *CTBP2* is de-repressed in the full EBNA 3A deletion mutant but not the truncated mutant (Fig 5-6 A). Cell lines used in this gene expression microarray were used to confirm repression of *CTBP2* by EBNA 3A and to determine isoform usage. There are six proposed isoforms of *CTBP2* as shown in Figure 5-4 A, from the top of the graphic to the bottom we termed these short (S) 1, S2 and long (L) 1, L2, L3 and L4. Primers were designed to distinguish between each of these mRNA variants. These data indicate that the L2 isoform is the primary transcript in these LCLs (D3wt1) and confirmed that EBNA 3A was able to repress *CTBP2* as LCLs generated using EBNA 3A deleted virus do not repress *CTBP2* (D3mtB1) (Fig 5-6 B).

We then confirmed de-repression of *CTBP2* in other LCLs generated from EBNA 3A knock out virus infection of cells from two donors by examining transcript levels (Fig 5-7 A). A reduction of *CTBP2* mRNA in the latency III Mutu III compared to the latency I Mutu I was also observed suggesting the combined effect of latency III proteins represses *CTBP2* (Fig 5-7 A). To investigate the effects of EBNA 3B on *CTBP2* transcription we used LCLs generated by infection of B cells from two different donors (PER142 and PER253) with wild type or EBNA 3B knock out virus (kindly provided by Dr Heather Long). In these cell lines we found that *CTBP2* transcript levels were increased in 3BKO lines relative to wild type infected cells, indicating gene repression and providing the first evidence of transcriptional regulation of *CTBP2* by EBNA 3B (Fig 5-7 B). *CTBP2* was not expressed in uninfected BL31 cells so we could not examine the BL31 infected with BAC derived EBNA 3C knock out virus. *CTBP2* is expressed in EREB 2.5 estrogen-tagged EBNA 2 system but we could not detect any regulation by EBNA 2 with or without β-estradiol (Fig 5-7 C). However subsequent ChIP-QPCR analysis revealed that EBNA 2 does not bind to the *CTBP2* intragenic peak in these cells. Together these data reveal that the

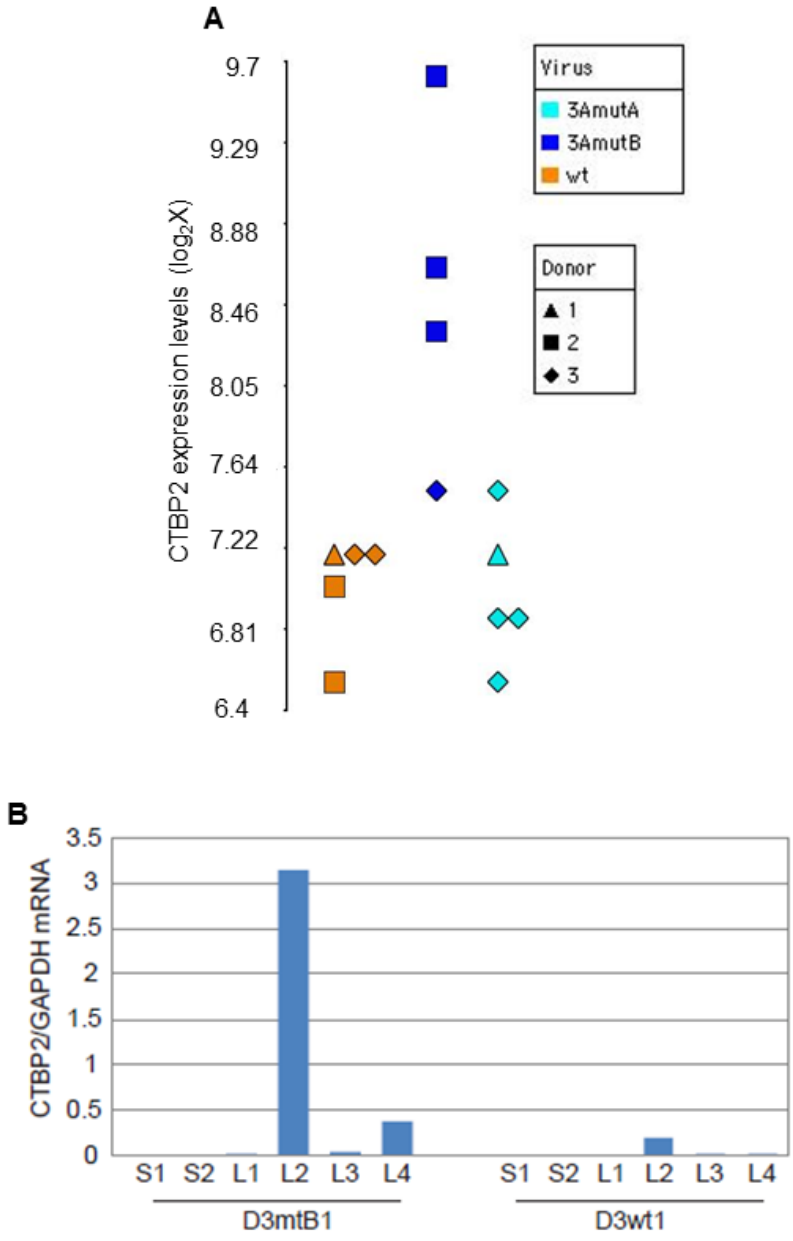


Figure 5-6 *CTBP2* isoform validation. (A) gene expression microarray data displayed as a dot plot from http://www.epstein-barrvirus.org.uk compiled by Dr Rob White. Plot shows raw expression data of *CTBP2* from donor B cells (triangles, squares, diamonds) infected with wild type (wt, orange), EBNA 3A knock out virus (3AmtB, blue) and EBNA 3A truncation mutant (3AmtA, cyan). Y axis is log2 values with values less than 3 indicating no expression and values over 8 indicates robust expression. (B) Analysis of CTBP2 cDNA by PCR using primers specific to each of the *CTBP2* isoforms (S1, S2, L1, L2, L3, L4, L5, L6) in D3mtB1 and D3wt1 cell lines in which CTBP2 is expressed and repressed respectively.

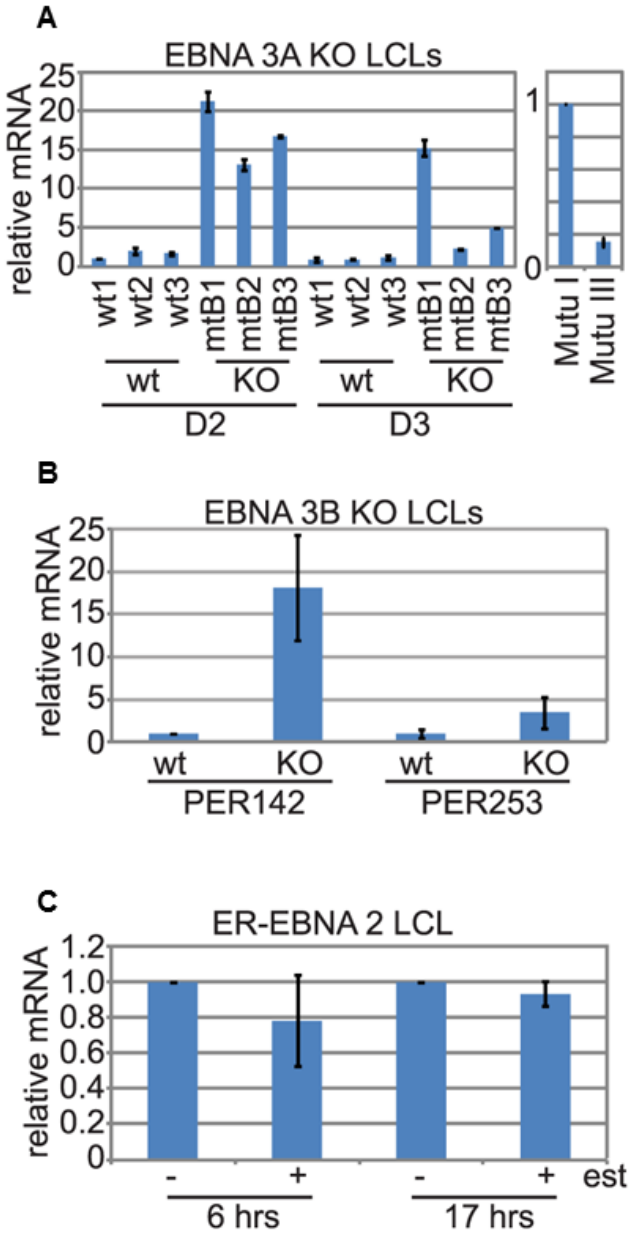


Figure 5-7 *CTBP2* is regulated by EBNA 3A and EBNA 3B in LCL backgrounds. (A) QPCR analysis of *CTBP2* transcript levels using wild type (wt1, wt2, wt3) and EBNA 3A KO (mtB1, mtB2, mtB3) virus immortalised LCLs from two different donor (D2, D3) backgrounds. Data is normalised to GAPDH transcript levels and expressed as relative to D2wt1. (B) QPCR analysis of *CTBP2* transcript levels in cDNA extracted from wild type LCLs immortalised by infection with B95.8 virus and EBNA 3B KO virus immortalised LCLs from two donors (PER142 and PER253). Data was normalised to GAPDH transcript levels and shown as relative to wild type levels of *CTBP2* from each donor. (C) QPCR analysis of *CTBP2* transcript levels using cDNA from EREB 2.5 which expresses a conditionally active EBNA 2 estrogen fusion mutant. Cells express EBNA 2 that is only transcriptionally active and nuclear localised in the presence of β-estradiol. Cells were stimulated with β-estradiol or DMSO (as a negative control) for 6 or 17 hours having been incubated in β-estradiol free media for 4 days. Transcript levels were normalised to GAPDH levels and expressed as relative to DMSO stimulated negative control. All cDNA data show the mean of two independent QPCR reaction each performed in duplicate. Error bars represent the range between two experiments.

intragenic peak at *CTBP2* is bound by EBNA 2, EBNA 3A, EBNA 3B and EBNA 3C and at least EBNA 3A and EBNA 3B are capable of repressing *CTBP2* transcription.

5.4 Regulation of *WEE1* by downstream distal EBNA 2 and EBNA 3C coincident binding sites

One of the enriched GO terms of genes closest to the most significant EBNA 3 binding sites was phosphorylation. This included many genes involved in cell proliferation due the abundance of kinases and related proteins involved in this process. Cell proliferation itself was also enriched but not as significantly. One gene in both these categories was the G2 checkpoint kinase *WEE1*. WEE1 was first characterised in *Schizosacchoromyces pombe* (*wee1*) as a cell division cycle mutant (Russell and Nurse 1987). WEE1 is active in the G2/S phase and can negatively regulate CDK1 by phosphorylation on Y15 to prevent entry into mitosis until DNA replication is complete (Featherstone and Russell 1991; McGowan and Russell 1995; Watanabe et al. 1995)(reviewed in (Kellogg 2003)). EBV has previously been reported to deregulate the cell cycle at multiple points including the G2/M phase, including misregulation of CDK1 activity (Krauer et al. 2004b; Schlick et al. 2011). *WEE1* therefore makes an attractive target for novel cell cycle regulation by EBNA 2 and 3 proteins.

ChIP sequencing data reveal that *WEE1* is the closest gene to two downstream, distal sets of EBNA 2 and EBNA 3 binding sites comprising a total of five distinct peaks (A,B,C,D,E) (Fig 5-8 A). These binding sites coincide with peaks of H3K27ac from ChIP sequencing data from the EBV immortalised LCL GM12878. Primers were designed to amplify each binding site (primer sets 2,4,6,8,10) and intervening regions (primer sets 1,3,5,7,9). ChIP-QPCR using antibodies specific to EBNA 2, 3A, 3B and 3C revealed that, unlike *CTBP2*, differential binding of EBNA 3 proteins occurs at this locus. As *CTBP2* was significantly bound by every EBNA examined we used this as a positive control. First, EBNA 2 binding to each of the five peaks was validated in Mutu III (Fig 5-8 B). However, using antibodies against EBNA 3A we were unable to detect significant binding at EBNA 3 sites in Mutu III (Fig 5-8 C). Low-level of EBNA 3B binding to peaks A and B was detected but was negligible compared to the signal obtained at *CTBP2* (Fig 5-8 D). In contrast significant EBNA 3C binding was detected at peak 5 (Fig 5-8 E). Similar result were observed in the PER253 LCL, EBNA 3B was shown not to bind to *WEE1* (Fig 5-9 D) and EBNA 3A showed only low level association with at peak E (Fig 5-9 C). Like Mutu III data, EBNA 2 was shown to bind to every binding site annotated from ChIP-sequencing data (Fig 5-9 B) and EBNA 3C bound primarily to peaks D and E at levels comparable to *CTBP2*. These data suggest that EBNA 3 binding is locus specific as EBNA 3A and EBNA 3B do not associate with *WEE1* binding sites in multiple cell types.

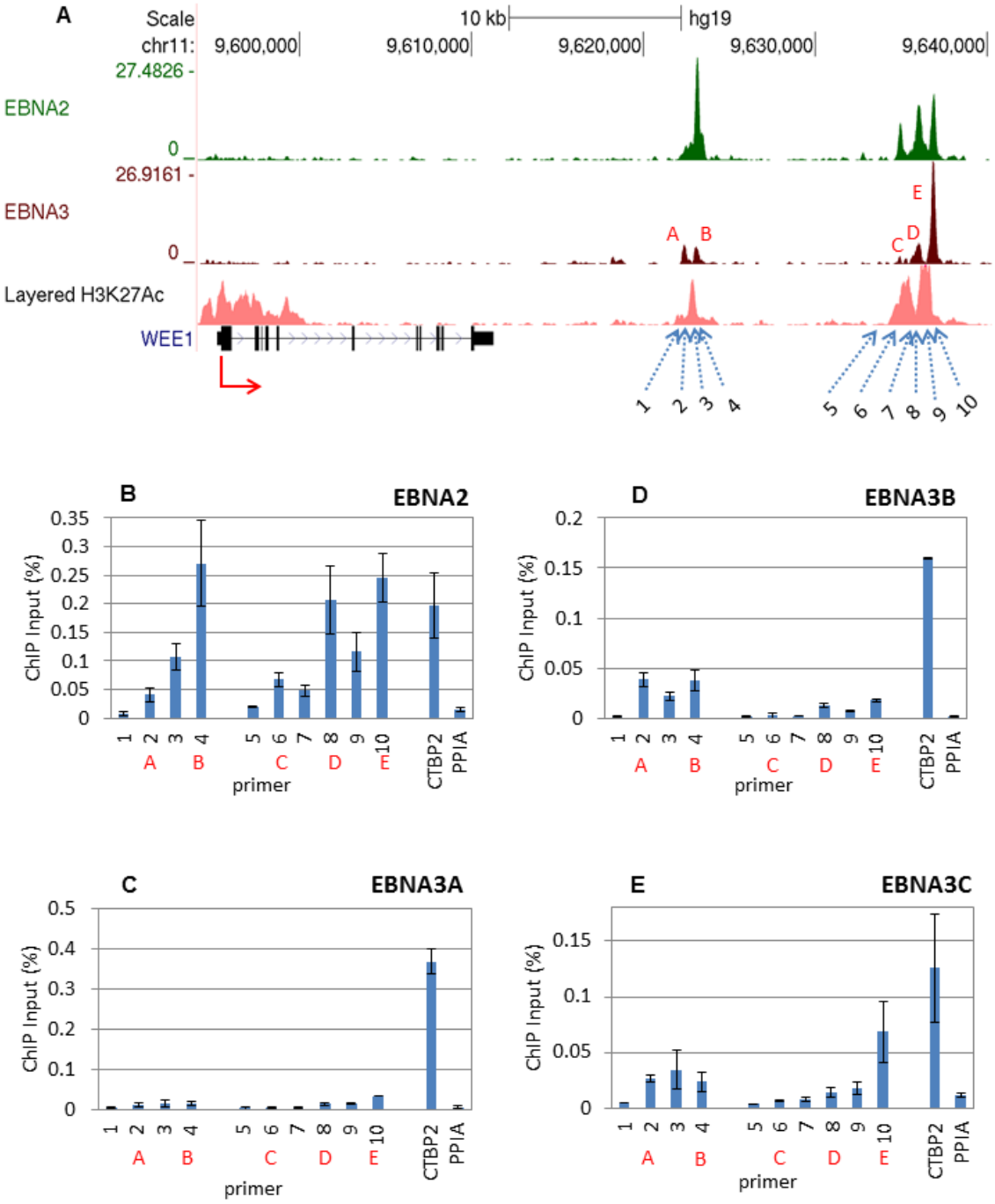


Figure 5-8 EBNA 2 and EBNA 3C bind to distal sites at *WEE1* in Mutu III. ChIP experiments conducted in Mutu III cell lines.(A) EBNA 2 (green) and EBNA 3 (red) binding at the *WEE1* locus. Top panel; scale, EBNA 2 and EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Each peak is annotated (A,B,C,D,E). Location of primer sets for the peaks of binding (2,4,6,8,10) and intervening regions (1,3,5,7,9) are annotated. Control primers spanned the TSS of *PPIA*. Signal at *CTBP2* peak is shown as a positive control. ChIP data is displayed as percentage input signals, after subtraction of signals of controls with no antibody. Corresponding peaks are displayed below primer sets in red). Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 2 (B), EBNA 3A (C), EBNA 3B (D) and EBNA 3C (E).

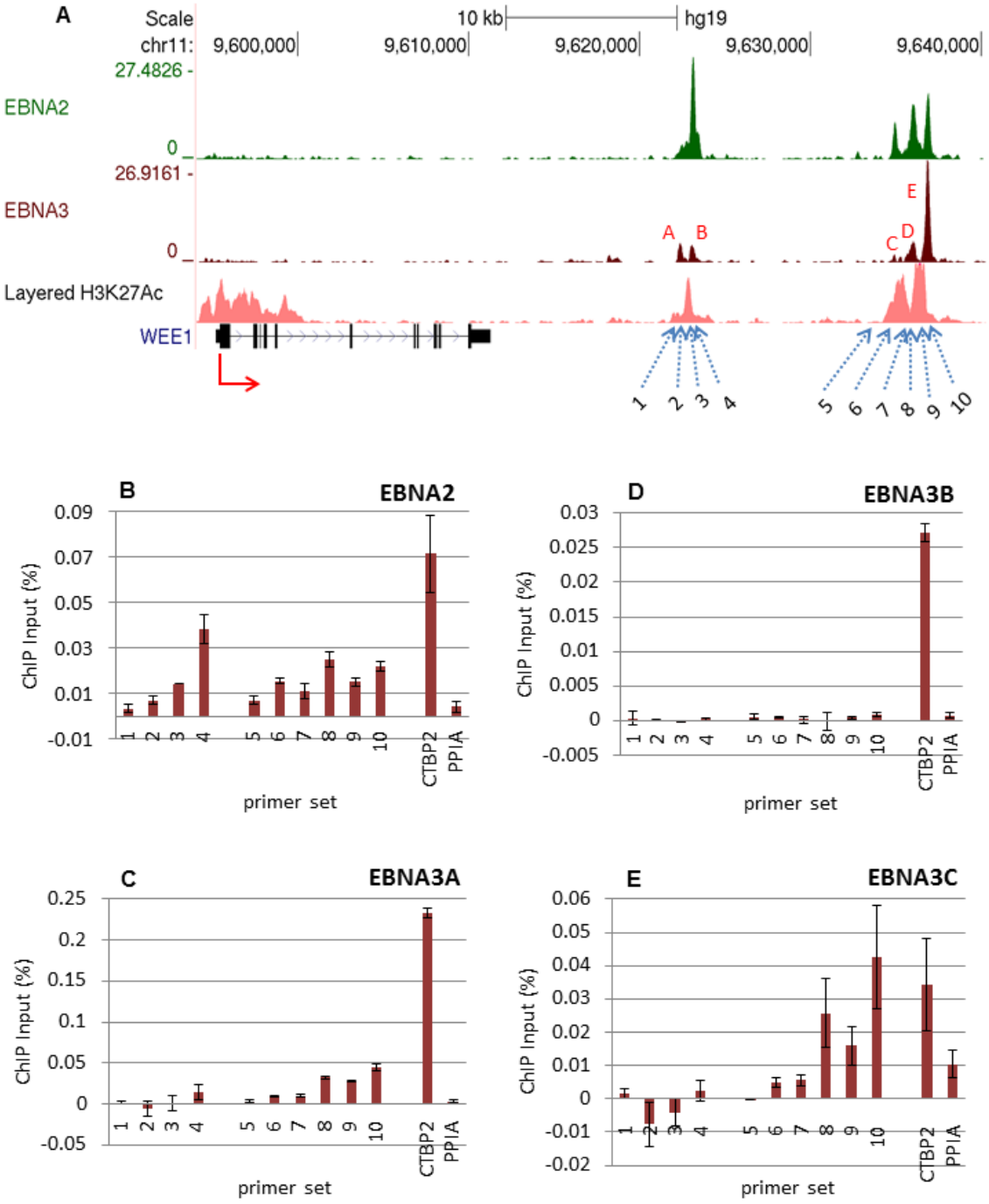


Figure 5-9 EBNA 2 and EBNA 3C bind to distal sites at *WEE1* in PER253 LCL. ChIP experiments conducted in PER253 LCL.(A) EBNA 2 (green) and EBNA 3 (red) binding at the *WEE1* locus. Top panel; scale, EBNA 2 and EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Each peak is annotated (A,B,C,D,E). Location of primer sets for the peaks of binding (2,4,6,8,10) and intervening regions (1,3,5,7,9) are annotated. Control primers spanned the TSS of *PPIA*. Signal at *CTBP2* peak is shown as a positive control. ChIP data is displayed as percentage input signals, after subtraction of signals of controls with no antibody. Corresponding peaks are displayed below primer sets in red). Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 2 (B), EBNA 3A (C), EBNA 3B (D) and EBNA 3C (E).

To determine if EBNA 3C binding at this locus had an effect on gene transcription we analysed cDNA from a cell panel of BL31, BL31 wtBAC, BL31 3CKO and BL31 3CREV. These have been previously described in 5.1. In the EBV negative BL31 cell line *WEE1* transcript could be detected and infection with wild type virus results in repression of *WEE1* (Fig 5-10 A). *WEE1* was derepressed in the 3CKO cell lines with levels of *WEE1* returning to that of BL31 consistent with 3C-mediated repression (Fig 5-10 A). Reinstating EBNA 3C in 3CREV cell lines results in the restoration of EBNA 3C mediated repression (Fig 5-10 A). These data suggest that EBNA 3C is responsible for the repression of *WEE1* in a background where all latent genes are expressed. To analyse the effects of EBNA 2 on *WEE1* mRNA expression we quantified *WEE1* cDNA levels in the EREB 2.5 cell line expressing a conditionally active EBNA 2 estrogen fusion protein. Exposure to β-estradiol and consequent EBNA 2 activation for 6 and 17 hours resulted in increased levels of *WEE1* mRNA (Fig 5-10 B). We also examined *WEE1* mRNA levels in BL31, BL31 wtBAC and BL31 infected with BAC derived virus lacking EBNA 2 (BL31 E2KO) cell lines. As in the BL31 3C KO and revertant cell panel, the net effect of latency III gene products was repression of *WEE1*, however, cells infected with a virus lacking EBNA 2 showed further repression of *WEE1* indicating that EBNA 2 acts as an activator of *WEE1* gene expression (Fig 5-10 C).

Whilst these data do not prove that the binding sites downstream of *WEE1* are directly regulating this gene, it demonstrates that EBNA 2 and EBNA 3C are capable of binding this locus and that this correlates with gene repression via EBNA 3C and increased expression in the presence of EBNA 2. Importantly we could not detect any regulatory effect of EBNA 3A or EBNA 3B, consistent with binding data at this locus. It is noteworthy that the effect of EBNA 3C on *WEE1* expression appears to be dominant over the effects of EBNA 2 although a conditional dual knock out cell line would be necessary to confirm this. This expression data supports locus specific binding being functionally relevant to EBNA 2 and 3 mediated transcriptional regulation.

5.5 *ITGAL* is differentially bound by EBNA 2 and 3 family proteins and binding is cell type specific

Integrin alpha L (ITGAL or CD11a) forms the LFA-1 heterodimer (αLβ2) with integrin beta 2 (ITGB2 or CD18). LFA-1 is widely expressed on the surface of myeloid and erythroid cells, including T and B lymphocytes. The inhibition of LFA-1 in T and B lymphocytes with a monoclonal antibody has been shown to block antigen-induced proliferation, T-cell mediated cytotoxicity, B cell aggregation and immunoglobulin production (reviewed in

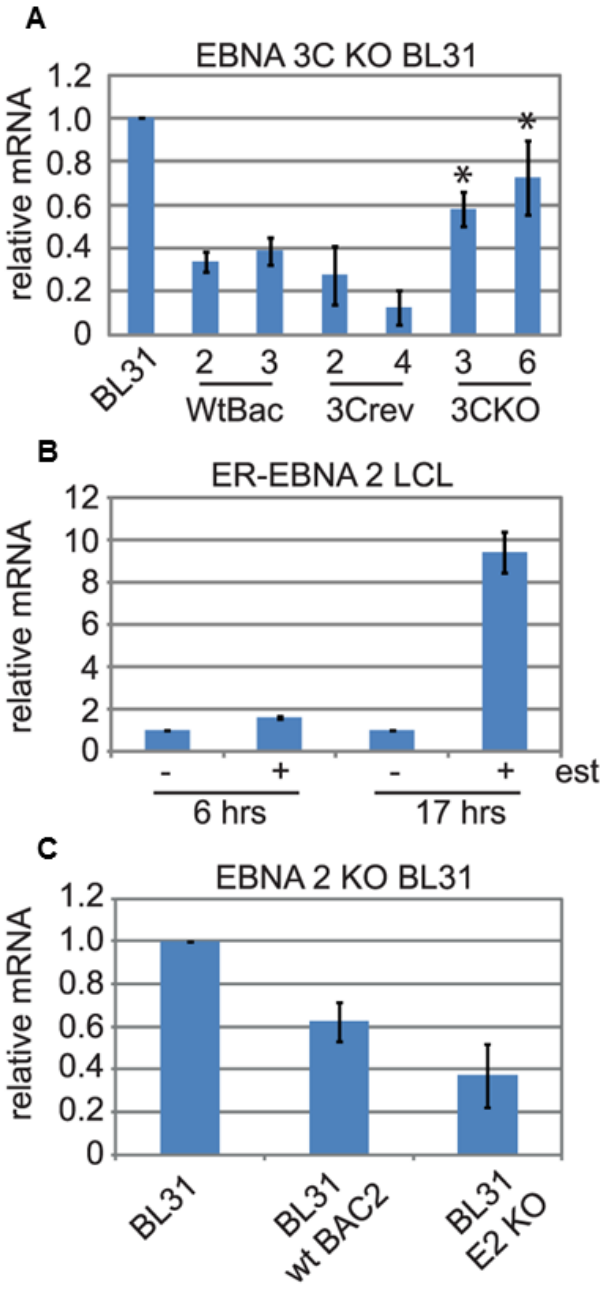


Figure 5-10 *WEE1* is regulated by EBNA 2 and EBNA 3C in a BL background and EBNA 2 in an LCL background. (A) QPCR analysis of *WEE1* transcript levels in EBV negative BL31 and BL31 cells infected with either BAC derived wild type EBV (WtBac2 and WtBac3), EBNA 3C KO EBV (3CKO3 and 3CKO6) or EBNA 3C revertant EBV (3Crev2 and 3Crev4). Data was normalised to *GAPDH* transcript levels and displayed as relative to BL31 WEE1/GAPDH mRNA. * indicates a p-value of less than 0.01 by students t-test compared to the WtBac2 cell line. (B) QPCR analysis of *WEE1* transcript levels using cDNA from EREB 2.5 LCLs expressing an EBNA 2 estrogen fusion mutant. Cells express EBNA 2 that is only transcriptionally active and nuclear localised in the presence of β-estradiol. Cells were stimulated with β-estradiol or DMSO (as a negative control) for 6 or 17 hours having been incubated in β-estradiol free media for 4 days. Transcript levels were normalised to *GAPDH* levels and expressed as relative to DMSO stimulated negative control. (C) QPCR analysis of cDNA levels in EBV negative BL31, BL31 infected with wild type BAC derived virus and BL31 cells infected with a BAC derived virus lacking EBNA 2. All cDNA data show the mean of two independent QPCR reaction each performed in duplicate. Transcript levels were normalised to GAPDH levels and expressed as relative to BL31. Error bars represent the range between two experiments.

(Mazzone and Ricevuti 1995). LFA-1 was first described as a receptor that aided adhesion of cytotoxic T cells to their targets (Harlan 1985) but can also strengthen the adhesion of T lymphocytes to dendritic cells (Springer 1990). Rapid and transient stimulation by cytokines presented on the cell surface can activate LFA-1 and promote firm adhesion to ICAM-1 which can subsequently cause transmigration through epithelial cells (reviewed in (Denucci et al. 2009). Importantly LFA-1 has also been shown to prevent B-cell apoptosis in the GC though adhesion to antigen presenting follicular dendritic cells (Lindhout et al. 1993).

ChIP sequencing data revealed that *ITGAL* is bound by three distinct, coincident promoter proximal peaks (A, B and C) (Fig 5-11 A). This binding site was also coincident with H3K27ac ChIP sequencing data from GM12878 LCLs. Using primers designed to amplify each peak (2,4,6) and intervening regions (1,3,5) we conducted ChIP-QPCR at this locus using antibodies specific to EBNA 2, 3A, 3B and 3C in the BL cell line Mutu III and the PER253 LCL. These data were plotted alongside *CTBP2* and *PPIA* as positive and negative controls respectively. We found that EBNA 2 bound primarily to peak C (primer set 6) in Mutu III, confirming ChIP sequencing data (Fig 5-11 B). Using antibodies against EBNA 3A revealed that the *ITGAL* promoter is not significantly bound by this factor as signals were comparable to that of the negative control (Fig 5-11 C). EBNA 3B and EBNA 3C on the other hand bound primarily to peak C and the signal obtained for EBNA 3B was 2 fold more than that of the positive control *CTBP2* (Fig 5-11 D, E). In the PER253 LCL we found that EBNA 2 and EBNA 3B also associated with peak C (Fig 5-12 B, D) and, similarly to the Mutu III cell line, EBNA 3A did not bind at any of the investigated sites (Fig 5-12 C). Intriguingly, unlike Mutu III, EBNA 3C did not bind above background levels observed at *PPIA* in PER253 LCL at this locus (Fig 5-12 E). These data suggest that EBNA 3 binding to particular cellular sites is cell type specific.

ITGAL has previously been reported to be regulated by EBNA 3B and EBNA 3C in data from gene expression microarray experiments in the BL31 BAC system. *ITGAL* is expressed in BL31 cells and is repressed upon infection by BAC derived wild type virus (Fig 5-13 A). Cells infected with virus lacking EBNA 3B or EBNA 3C show increased expression levels of *ITGAL* mRNA implicating them in transcriptional repression of *ITGAL* (Fig 5-13 A). This was not observed for BL31 cells infected with virus lacking EBNA 3A (Fig 5-13 A). Infection with virus lacking all three EBNA 3 proteins also results in *ITGAL* derepression but this does not pass significance thresholds set (Fig 5-13 A). We also examined *ITGAL* mRNA levels in LCLs generated from two donors (PER142 and PER253) infected with wild type or EBNA 3BKO virus, revealing that levels of ITGAL mRNA were increased in cells lacking EBNA 3B relative to wild type, implicating EBNA

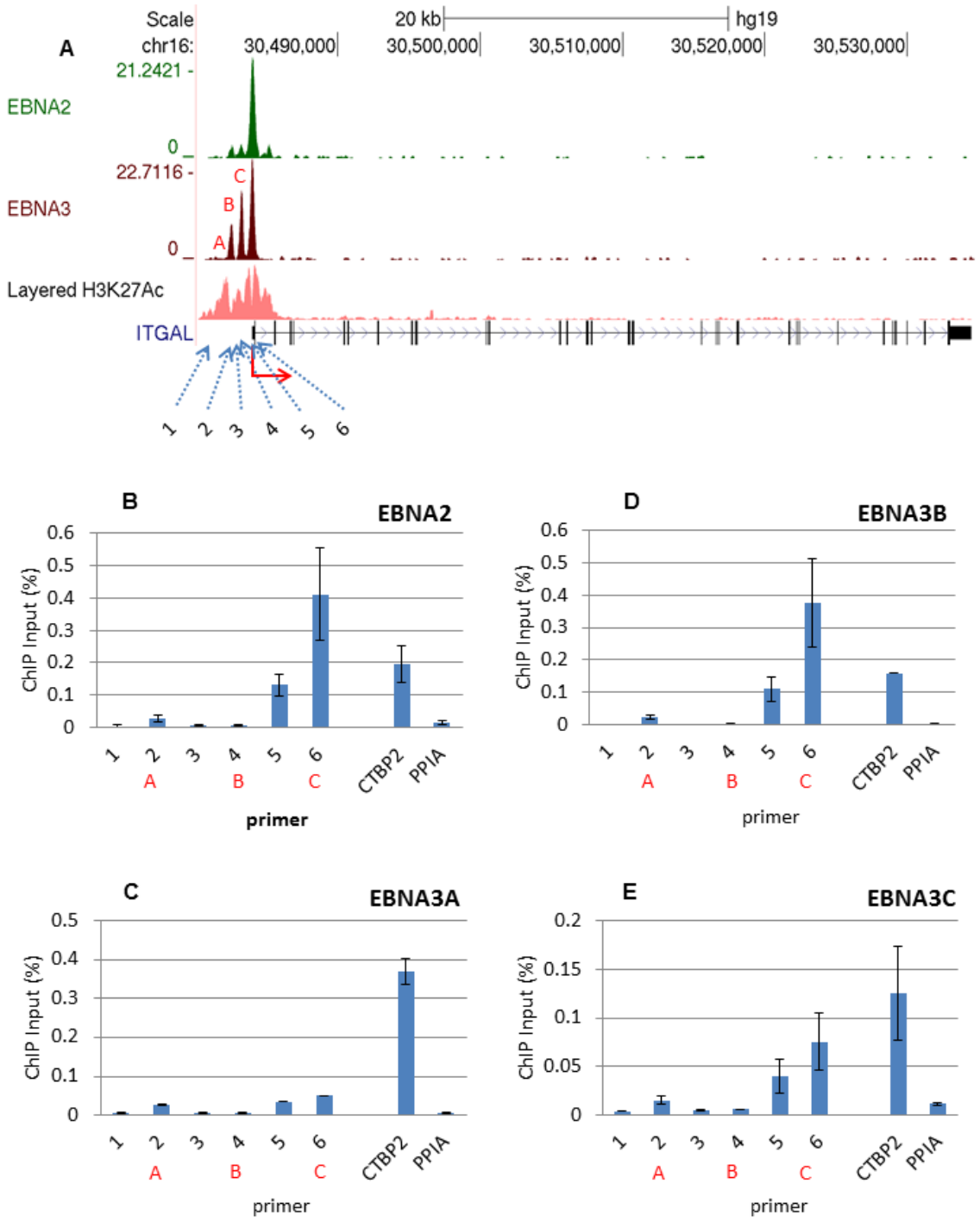


Figure 5-11 EBNA 2, EBNA 3B and EBNA 3C bind to promoter proximal sites at *ITGAL* in Mutu III. ChIP experiments conducted in Mutu III cell lines. (A) EBNA 2 (green) and EBNA 3 (red) binding at the *ITGAL* locus. Top panel; scale, EBNA 2 and EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Each peak is annotated (A,B,C). Location of primer sets for the peaks of binding (2,4,6) and intervening regions (1,3,5) are annotated. Control primers spanned the TSS of *PPIA*. Signal at *CTBP2* peak is shown as a positive control. ChIP data is displayed as percentage input signals, after subtraction of signals of controls with no antibody. Corresponding peaks are displayed below primer sets in red). Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 2 (B), EBNA 3A (C), EBNA 3B (D) and EBNA 3C (E).

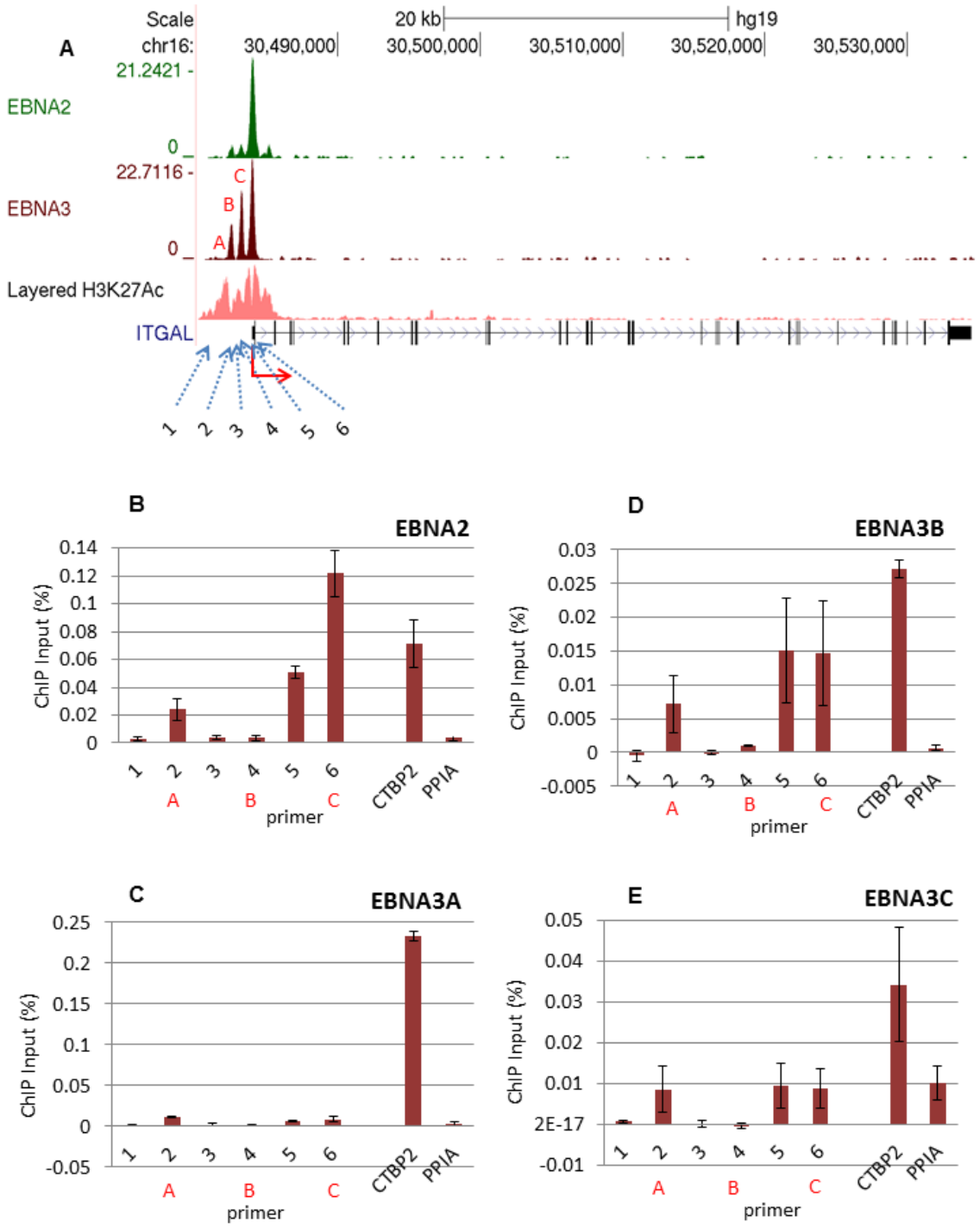


Figure 5-12 EBNA 2 and EBNA 3B bind to promoter proximal sites at *ITGAL* in PER253 LCL. ChIP experiments conducted in PER253 LCL. (A) EBNA 2 (green) and EBNA 3 (red) binding at the *ITGAL* locus. Top panel; scale, EBNA 2 and EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of gene and direction of transcription. Each peak is annotated (A,B,C). Location of primer sets for the peaks of binding (2,4,6) and intervening regions (1,3,5) are annotated. Control primers spanned the TSS of *PPIA*. Signal at *CTBP2* peak is shown as a positive control. ChIP data is displayed as percentage input signals, after subtraction of signals of controls with no antibody. Corresponding peaks are displayed below primer sets in red). Results are shown as ± standard deviations of two independent experiments. ChIP-QPCR data for EBNA 2 (B), EBNA 3A (C), EBNA 3B (D) and EBNA 3C (E).

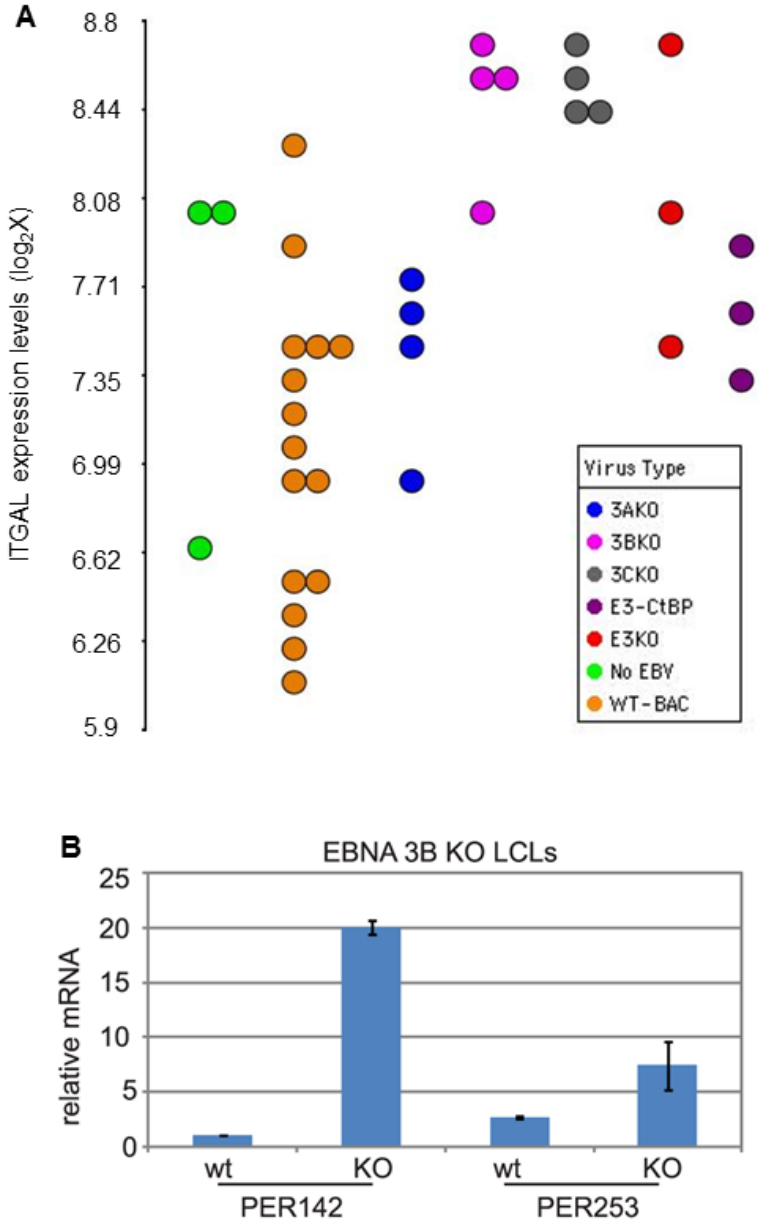


Figure 5-13 *ITGAL* is repressed by EBNA 3B and EBNA 3C in a BL background and repressed by EBNA 3B in a LCL background (A) gene expression microarray data displayed as a dot plot from http://www.epstein-barrvirus.org.uk compiled by Dr Rob White. Plot shows raw expression data of *ITGAL* from BL31 cells uninfected (green), infected with wild type BAC virus (wt-BAC, orange), EBNA 3A knock out BAC virus (3AKO, blue), EBNA 3B knock out BAC virus (3BKO, pink), EBNA 3C knock out BAC virus (3CKO, grey), EBNA 3A/B/C knock out BAC virus (E3KO, red) or EBNA 3A and EBNA3C CTBP binding defective BAC virus (E3-CtBP, purple). Y axis is log2 values with values less than 3 indicating no expression and values over 8 indicates robust expression. (B) QPCR analysis of *ITGAL* transcript levels in cDNA extracted from wild type LCLs immortalised by infection with B95.8 virus and EBNA 3B knock out virus immortalised LCLs from two donors (PER142 and PER253). Data was normalised to *GAPDH* transcript levels and shown as relative to wild type levels of *ITGAL* from each donor. Data shows the mean of two independent QPCR reaction each performed in duplicate. Error bars indicates the range between two experiments.

3B in *ITGAL* transcriptional repression in LCLs. Interestingly when we examined cDNA in other EBNA 3 positive and negative cell lines we could only detect regulation of *ITGAL* in LCLs by EBNA 3B (Fig 5-13 B). Together the ChIP and gene expression analysis reveal that in BL cell lines EBNA 3B and EBNA 3C both bind and repress this locus, whereas in the LCL studied only EBNA 3B bound and repressed *ITGAL*. These data support the role of cell-type specific binding being functionally relevant to EBNA 2 and 3 mediated transcriptional regulation.

5.6 EBNA 2 and 3 proteins compete for binding at coincident sites

The data described here suggests that coincident binding of EBNA 2 and 3 proteins is functionally relevant to gene regulation. Furthermore, differential binding of EBNA 3 factors is predictive of regulation and is cell type and locus specific. A key remaining question to address is do the EBNA 2 and 3 proteins bind simultaneously to coincident sites and exert their influence on transcription as multi-EBNA complexes or do they bind separately. To address this question we used sequential ChIP experiments (reChIP) in a Mutu III cell line.

Briefly, a ChIP experiment is performed but immunoprecipitated DNA-protein complexes are eluted from the primary antibody and then a second round of ChIP is performed with the same, or different antibodies. Using the same antibody acts as a positive control as complexes precipitates in the first ChIP will contain the protein for the second ChIP. For a negative control no antibody is added in the secondary ChIP. Using EBNA 2 as our antibody in the primary ChIP followed by the same antibody in the second round ChIP validated the reChIP experiment as DNA was enriched above the negative control of adding no antibody to the secondary ChIP (Fig 5-14 A). Similar signals to the no antibody were obtained when using EBNA 3A, EBNA 3B or EBNA 3C specific antibodies in the secondary ChIP suggesting that EBNA 3 proteins do not co-associate with EBNA 2 at the *CTBP2* peak (Fig 5-14 A). To check that EBNA 3A, EBNA 3B and EBNA 3C antibodies were capable of precipitating DNA in the secondary ChIP, reChIP experiments were performed using EBNA 3A, EBNA 3B and EBNA 3C as the primary antibody. These data reveal that when the same EBNA 3 antibody is used in the primary and secondary ChIPs a signal over background is obtained showing that the EBNA 3A (Fig 5-14 B), EBNA 3B (Fig 5-14 C)and EBNA 3C (Fig 5-14 D) antibodies are able to precipitate DNA after two rounds of ChIP. We also used EBNA 2 specific antibodies in the secondary ChIP after a primary ChIP using and EBNA 3 antibody. For EBNA 3A, EBNA 3B and EBNA 3C primary ChIP EBNA 2 secondary ChIP we detected signals comparable to that of the no antibody secondary ChIP control. This supports the finding from Figure 5-14 A that EBNA 3 and EBNA 2 proteins do not associate with the *CTBP2* intragenic peak at the same time.

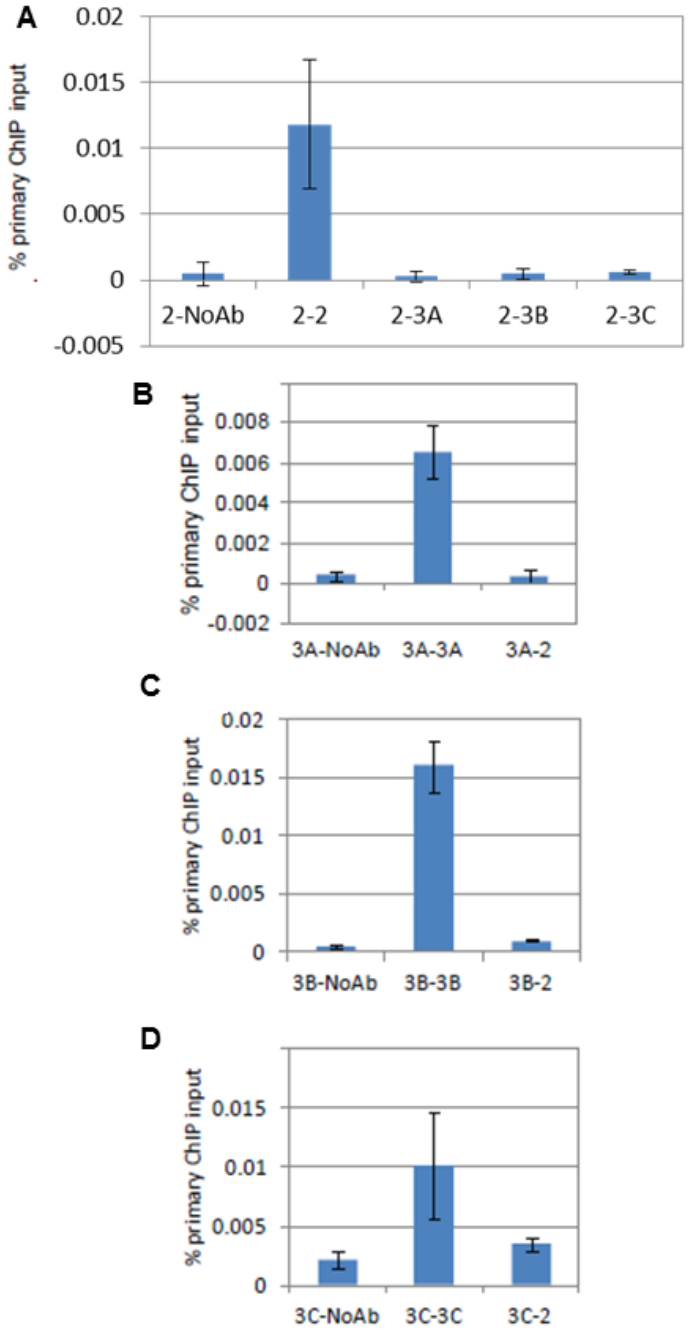


Figure 5-14 ReChIP analysis at the intragenic *CTBP2* peak. ChIP experiments were performed with an antibody (primary), complexes were eluted and subjected to a second ChIP using a second antibody (secondary). DNA was analysed by QPCR using *CTBP2* peak primers. (A) Re-ChIP analysis using anti EBNA 2 as a primary antibody and anti EBNA 2 (positive control)(2-2), EBNA 3A (2-3A), EBNA 3B (2-3B) and EBNA 3C 2-3C) as secondary antibodies. A no antibody secondary was also conducted (negative control)(2-NoAb). (B) Re-ChIP analysis using anti EBNA 3A as a primary antibody and anti EBNA 3A (positive control)(3A-3A) and EBNA 2 (3A-2) secondary antibodies. A no antibody secondary was also conducted (negative control)(3A-NoAb). (C) Re-ChIP analysis using anti EBNA 3B as a primary antibody and anti EBNA 3B (positive control)(3B-3B) and EBNA 2 (3B-2) secondary antibodies. A no antibody secondary was also conducted (negative control)(3B-NoAb). (D) Re-ChIP analysis using anti EBNA 3C as a primary antibody and anti EBNA 3C (positive control)(3C-3C) and EBNA 2 (3C-2) secondary antibodies. A no antibody secondary was also conducted (negative control)(3C-NoAb). Data is shown as a percentage of the primary ChIP input. Error bars represent ± standard deviations of two independent experiments.

We also analysed ReChIP DNA using primers specific for *WEE1* peak E, a site bound only by EBNA 2 and EBNA 3C in Mutu III cell lines (see Fig 5-8). Using an EBNA 2 specific antibody in the primary ChIP we found that using an anti-EBNA 2 antibody in the secondary ChIP produced a signal above that of the negative control (Fig 5-15 A). We also found that anti-EBNA 3C antibody added to an EBNA 2 antibody primary ChIP and analysed by QPCR produced a signal comparible to the negative control, and EBNA 3A and EBNA 3B antibodies, which do not bind to this site (Fig 5-15 A). We then conducted QPCR on DNA precipitated from a ChIP using EBNA 3C followed by a secondary ChIP using the same antibody and found that a signal above background was detectable, whereas using an EBNA 2 specific antibody produced a signal comparable to the negative control (Fig 5-15 B). QPCR analysis of reChIP DNA was performed using primers for *ITGAL* peak C. In Mutu III cells this site is bound by EBNA 2, EBNA 3B and EBNA 3C (see Fig 5-11). Data here shows that EBNA 3B and EBNA 3C antibodies do not enrich DNA at this site when added to a primary ChIP conducted with an anti-EBNA 2 antibody (Fig 5-16 A) despite EBNA 3B (Fig 5-16 B) and EBNA 3C (Fig 5-16 C) antibodies being able to enrich DNA from this site after two rounds of ChIP. This data shows that EBNA 2 and EBNA 3 proteins do not associate with *WEE1* peak E or *ITGAL* peak C at the same time.

ReChIP experiments described here implicate competition between EBNA 2 and EBNA 3 factors for binding sites. At all the binding sites investigated EBNA 2 coincidentally bound with at least one EBNA 3 protein but in no QPCR experiments could we detect evidence that both were in a complex at the same site. ChIP data is a population average of millions of cells, therefore signals obtained by QPCR or sequencing showing coincident binding between EBNA 2 and EBNA 3 proteins likely represents the site being bound in subpopulations of cells by either EBNA 2, EBNA 3A, EBNA 3B or EBNA 3C.

5.7 Discussion

Prior to this investigation there were no known human DNA elements targeted by EBNA 2 and EBNA 3 proteins. Here we describe over 5300 coincident binding sites for EBNA 2 and 3 proteins in the human genome which target over 3100 genes. We chose to investigate three genes targeted by coincident sites involved in important pathways for EBV pathogenicity and immortalisation.

5.7.1 CTBP2

We show here that the C-terminal binding protein *CTBP2* is bound at an intragenic site by EBNA 2, 3A, 3B and 3C in the BL cell line MutuIII and the PER253 LCL (Fig 5-4 and Fig 5-5).

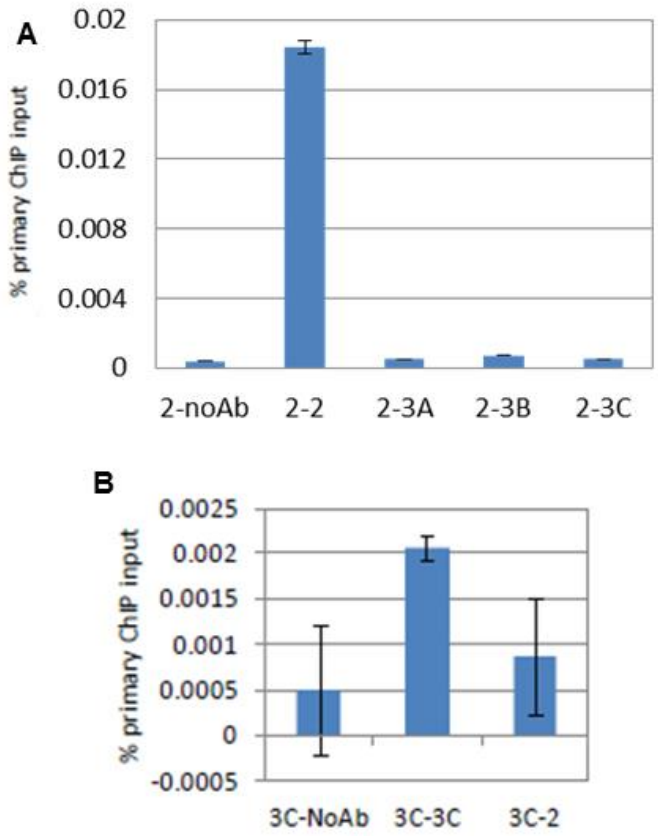


Figure 5-15 ReChIP analysis at the distal *WEE1* peak E. ChIP experiments were performed with an antibody (primary), complexes were eluted and subjected to a second ChIP using a second antibody (seconday). DNA was analysed by QPCR using *WEE1* peak E primers. (A) Re-ChIP analysis using anti EBNA 2 as a primary antibody and anti EBNA 2 (positive control)(2-2), EBNA 3A (2-3A), EBNA 3B (2-3B) and EBNA 3C 2-3C) as secondary antibodies. A no antibody secondary was also conducted (negative control)(2-NoAb). (B) Re-ChIP analysis using anti EBNA 3C as a primary antibody and anti EBNA 3C (positive control)(3C-3C) and EBNA 2 (3C-2) secondary antibodies. A no antibody secondary was also conducted (negative control)(3C-NoAb). Data is shown as a percentage of the primary ChIP input. Error bars represent ± standard deviations of two independent experiments.

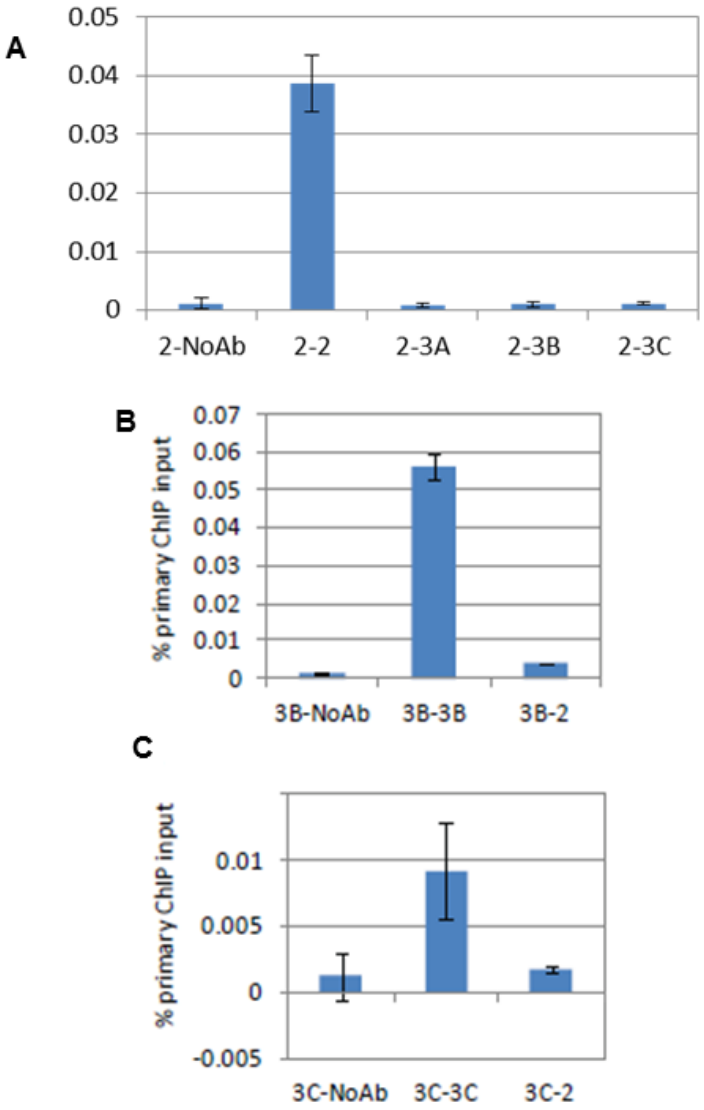


Figure 5-16 ReChIP analysis at the promoter proximal *ITGAL* peak C. ChIP experiments were performed with an antibody (primary), complexes were eluted and subjected to a second ChIP using a second antibody (seconday). DNA was analysed by QPCR using *ITGAL* peak C primers. (A) Re-ChIP analysis using anti EBNA 2 as a primary antibody and anti EBNA 2 (positive control)(2-2), EBNA 3A (2-3A), EBNA 3B (2-3B) and EBNA 3C 2-3C) as secondary antibodies. A no antibody secondary was also conducted (negative control)(2-NoAb). (B)Re-ChIP analysis using anti EBNA 3B as a primary antibody and anti EBNA 3B (positive control)(3B-3B) and EBNA 2 (3B-2) secondary antibodies. A no antibody secondary was also conducted (negative control)(3B-NoAb). (C) Re-ChIP analysis using anti EBNA 3C as a primary antibody and anti EBNA 3C (positive control)(3C-3C) and EBNA 2 (3C-2) secondary antibodies. A no antibody secondary was also conducted (negative control)(3C-NoAb). Data is shown as a percentage of the primary ChIP input. Error bars represent ± standard deviations of two independent experiments.

Furthermore this gene is regulated by EBNA 3A and 3B in LCLs (Fig 5-7). Currently *CTBP2* has not been reported as regulated in any gene expression microarray investigating EBNA 2 or 3 proteins. However re-analysis of data from LCLs generated from wild type virus or EBNA 3A deficient virus (Hertle et al. 2009) by Dr Rob White (Imperial College, London) revealed that in one EBNA 3A mutant virus studied regulation is detectable. *CTBP2* was not included previously as a regulated gene lists because data from two different EBNA 3A mutant viruses was amalgamated. In one mutant virus EBNA 3A is deleted, and in the other mutant EBNA 3A is truncated and no protein is expressed (Hertle et al. 2009). *CTBP2* mRNA levels are elevated in the full deletion mutant compared to wild type infected cells, but not in the truncation mutant (Fig 5-6 A). The mechanism of this remains unclear but it is possible that a low level and undetectable quantity of EBNA 3A protein is expressed in EBNA 3A truncation mutants and that this is capable of repressing *CTBP2*.

In follow up experiments from this initial characterisation of binding sites, a colleague in our lab (Dr David Wood, university of Sussex) performed chromosome conformation capture (CCC) on this locus. CCC allows detection of a specific promoter-enhancer looping event. We know that many long range enhancer functions are mediated via this mechanism as discussed in the introduction. First, chromatin is cross linked with formaldehyde, creating a snapshot of the global 3D chromatin architecture by covalently binding proteins and DNA together. Then DNA is digested with a restriction enzyme chosen to isolate promoter and enhancer fragments of DNA. Complexes are then ligated and cross-links reversed. Proteins are digested with enzymes and DNA is purified. In this way a promoter enhancer DNA chimera will be created if the two sites were brought to close proximity via protein interactions. Primers can then be used that span restriction sites between the enhancer and promoter and the presence of enhancer promoter looping can be detected by QPCR.

These experiments were carried out in a EBNA 3A KO LCL in which *CTBP2* was expressed, and in a wild type infected LCL in which *CTBP2* was repressed. In the wild type infected LCLs no promoter and intragenic enhancer looping was detected. However, in cells infected with a virus lacking EBNA 3A a loop between the enhancer and promoter was observed. This suggests that EBNA 3A is critical for the prevention of an activating loop from forming between the enhancer and promoter at *CTBP2* (Fig 5-17). Unfortunately we could not investigate the effects of EBNA 3C on either regulation or looping at the *CTBP2* locus as no such cell series we examined expressed *CTBP2*. We were able to detect both binding and repression of *CTBP2* by EBNA 3B, however we do not know if repression is mediated by the prevention of activating

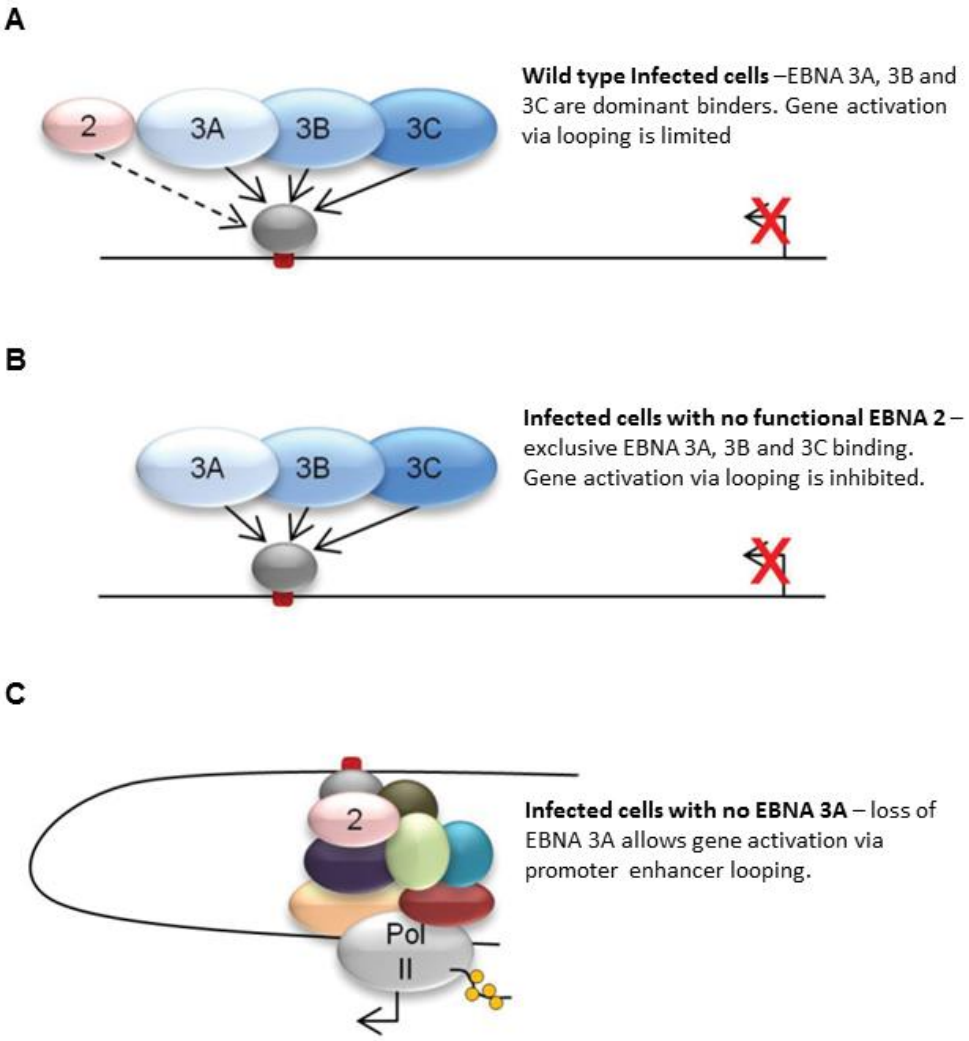


Figure 5-17 Model for regulation of *CTBP2* via enhancer looping mediated by EBNA 2 and 3 binding. (A) model for EBV modulation of *CTBP2* promoter enhancer looping in wild type infected cells. (B) model for EBV modulation of *CTBP2* promoter enhancer looping in infected cells lacking functional EBNA 2. (C) model for EBV modulation of *CTBP2* promoter enhancer looping in infected cells lacking EBNA 3A. Intragenic enhancer is displayed as a red box. Promoter is displayed as an arrow with a cross through it if it is repressed. Unknown cellular binding partner of EBNA 2, 3A, 3B or 3C is displayed as a grey oval. EBNA 2 is shown as a pink oval with 2 in and EBNA 3A, EBNA 3B and EBNA 3C as blue ovals with 3A, 3B or 3C in them respectively. Components of transcriptional machinery and looped DNA are annotated in situations where *CTBP2* expression is de-repressed.

enhancer promoter looping. EBNA 3B is expressed in the EBNA 3A KO cell line in which the described CCC experiment was conducted in, therefore EBNA 3B does not prevent an activating loop by itself but may be required for EBNA 3A dependant repression of *CTBP2*. Another possibility is that activating loop formation may be partially inhibited by EBNA 3B in cells lacking EBNA 3A but as the assay performed was not quantitative we could observe no difference. We have been unable to address the role of EBNA 2 on the regulation of *CTBP2* because it is not expressed in cell series studied. However analysis of *CTBP2* transcript levels in Mutu I compared to Mutu III reveals that the overall effect of latency III protein expression is repression of *CTBP2* (Fig 5-7 A). This suggests that EBNA 3A (and potentially 3B and 3C) have a dominant role in the repression of this gene.

C-terminal binding proteins were first characterised in 1993 as a binding partner of the C-terminus of the adenovirus E1A protein (Boyd et al. 1993). E1A can activate and repress transcription and it appears that E1A mediated repression requires its association with cellular CTBP (Sollerbrant et al. 1996). The consensus binding motif for CTBP has been described and mutants in the E1A binding region for CTBP are defective for transformation and immortalisation (Schaeper et al. 1995). More recently the interaction of the C terminal region of E1A and CTBP has been shown to repress the transforming properties of E1A and play a role in enhancing viral replication during productive infection (Cohen et al. 2013; Subramanian et al. 2013).

The CTBP family consists of two highly related but unique proteins CTBP1 and CTBP2. CTBP1 and CTBP2 share a 72% amino acid identity and are encoded by genes on chromosome 4 and 21 respectively. This distinction was only revealed in 1998 (Katsanis and Fisher 1998) and to this day CTBP1 and 2 are often collectively termed as CTBP, making the differentiation of the function of these two proteins a more difficult task. It is noteworthy that *CTBP1* has not been identified as a regulated target of EBNA 2 and 3 proteins, nor was it identified as a gene closest to a peak. However, a single significant coincident intragenic peak is found at this gene (Fig 5-18). *CTBP1* may not have been identified as a gene closest to a EBNA 2 or EBNA 3 binding site as each peak is assigned to a single gene, in this case this may have been determined to be *HV535469* or *HV535487*. These are completely unannotated human cDNAs with no known function or known protein expression. It would be informative to investigate *CTBP1* and this intragenic binding site.

Both CTBP1 and 2 appear to bind to the same consensus motif described in the C terminus of E1A and recent work has shown that CTBP1 and CTBP2 perform redundant and unique

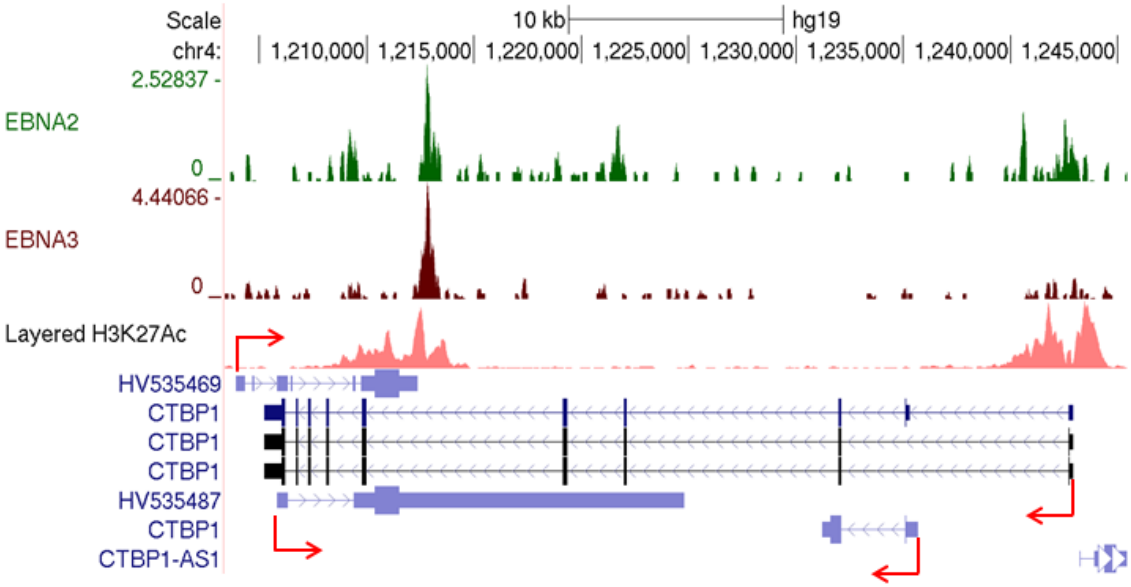


Figure 5-18 EBNA 2 and EBNA 3 proteins bind to an intragenic site at *CTBP1* in Mutu III. EBNA 2 (green) and EBNA 3 (red) binding at the *CTBP1* locus. Top panel; scale, EBNA 2 and EBNA 3 panel; our ChIP sequencing data displayed as reads per million background subtracted total reads and aligned with the human genome, H3K27ac panel; H3K27ac ChIP sequencing data from GM12878 cells in the ENCODE project, bottom panel; annotation of genes and direction of transcription.

functions; for example, during productive adenovirus infection the E1A-CTBP2 complex restricts viral replication more than E1A-CTBP1 (Subramanian et al. 2013). Crucially whilst CTBP1 is both nuclear and cytosolic, CTBP2 appears to be nuclear only due to the presence of a nuclear localisation domain in its N terminus (Zhao et al. 2006c). Furthermore mice deficient in CTBP1 are viable whereas CTBP2 null mice are not (Hildebrand and Soriano 2002). Proteomics analysis in Hela cells of the binding partners of CTBP1 and CTBP2 revealed both common and unique binding partners of both proteins. Unique binding partners of CTBP2 include the transcription factor E2F7 and components of the NuRD complex which have been shown to play a critical role in cell cycle regulation and histone modifications (Zhao et al. 2014). Importantly, here we report that EBV preferentially targets *CTBP2* for transcriptional repression. These findings are important to the study of EBV as transcriptional regulation by the key EBNAs has not distinguished between CTBP1 and CTBP2. For example, EBNA 3A and 3C have been shown to repress some of its targets, including $p16^{INK4A}$, (Skalska et al. 2010) in a CTBP dependant manner, but it is unknown whether CTBP1 or CTBP2 or both are involved. As we show here that EBNA 3A is a potent repressor of *CTBP2* transcription this may infer that EBV preferentially requires CTBP1 as a cellular binding partner. It is also possible that repression of *CTBP2* by EBNA 3A is part of an auto-regulatory loop whereby EBNA 3A controls its own repressive activity by regulating *CTBP2* transcription.

The role of CTBP proteins in the cell appears to be very diverse. It has been proposed that CTBP proteins exert their transcriptional influence by acting as a scaffold to recruit chromatin modifying enzymes such as HDACs, polycomb group proteins and histone methylatransferases (Shi et al. 2003; Kuppuswamy et al. 2008). siRNA knock down of CTBP1 and CTBP2 results in a decrease in mitotic events, reduction of proliferation and increase in apoptosis in breast cancer cells caused by abnormalities in chromosome alignment (Bergman et al. 2009). Genome-wide analysis of CTBP binding sites in the human genome reveals that CTBP proteins primarily bind to distal elements (1823/6607 sites are promoter proximal) and that genes targeted are enriched for GO terms including: cellular response to DNA damage, cell cycle, cell proliferation, cell death, cell adhesion and chromatin modifications (Di et al. 2013). Notably, consensus binding motifs for cellular TFs centred under CTBP peaks show an enrichment for ETS (PU.1) transcription factors amongst others (Di et al. 2013). CTBP2 specifically has also been reported to influence cancer cell migration by induction of the T-cell lymphoma invasion and metastasis I (Tiam1) protein. Upregulation of Tiam1 was discovered through reporter assays in which KLF8 was shown to recruit CTBP2 to three elements up to 1.8kb from the Tiam1 promoter (Paliwal et al. 2012). This highlights the fact the CTBP2 can act as a co-

repressor and activator of transcription. CTBP proteins are also bound by $p14^{ARF}$ which targets them for proteasomal degradation, modulating their repressive role in the cell death pathway. CTBP2 was shown to bind to the Bik promoter and repress its transcription. Transcriptional repression found to be alleviated in the presence of $p14^{ARF}$ (Kovi et al. 2010). Intriguingly depletion of CTBP2 has also been shown to decrease the level of MYC and some of its downstream targets identifying further unique activities of CTBP2 that may be relevant to EBV induced transformation (Zhang et al. 2014).

Our analysis has therefore highlighted a potential role for the regulation of *CTBP2* in EBV immortalised cells, and additional experiments are now required to investigate this further. It is important to determine if EBNA 3A and EBNA 3C bind to CTBP2 or CTBP1 and explore the effect of CTBP2 on growth. Importantly, like *Bim*, we have found that *CTBP2* is not expressed in many of the cell lines available to us. The repression of Bim is well characterised and involves initial recruitment of PRC components and H3K27me3 by EBNA 3A and EBNA 3C. This is followed by PRC2 protein recruitment and CpG methylation and persistent repression. It will therefore be informative to investigate H3K27me3 and CpG methylation at *CTBP2* upon expression of EBNA 3A and EBNA 3C.

5.7.2 WEE1

Whilst EBV nuclear antigens have not been previously shown to regulate *WEE1*, it is upregulated following initial infection. It has been reported that 4-7 days post infection EBV infected cells undergo a period of hyperproliferation which is marked by a heightened DNA damage response (DDR) and cell proliferation gene expression in the host cell (Nikitin et al. 2010). One of the upregulated cell proliferation genes at this time after infection is *WEE1*, the expression of which is then reduced during the process of outgrowth to an LCL. Furthermore the DDR kinases ATM and Chk2, which are induced in the hyperproliferation period have been shown to supress EBV-mediated transformation and initial B-cell proliferation (Nikitin et al. 2010). EBNA 3C is required to attenuate the DDR during this period as cells infected with EBNA 3CKO virus rapidly accumulated DDR genes and marks, crippling them for long term outgrowth. Analysis of raw data from gene expression microarray data from the BL31 BAC cell series (White et al. 2010) reveals that EBNA 3C may regulate *WEE1* although this did not pass significance thresholds set in their analysis (Fig 5-19). The large spread of expression levels of *WEE1* in wild type BAC and 3C revertant cell lines is likely responsible for this. This suggests that even in established EBV immortalised cells *WEE1* is still partially repressed by EBNA 3C.

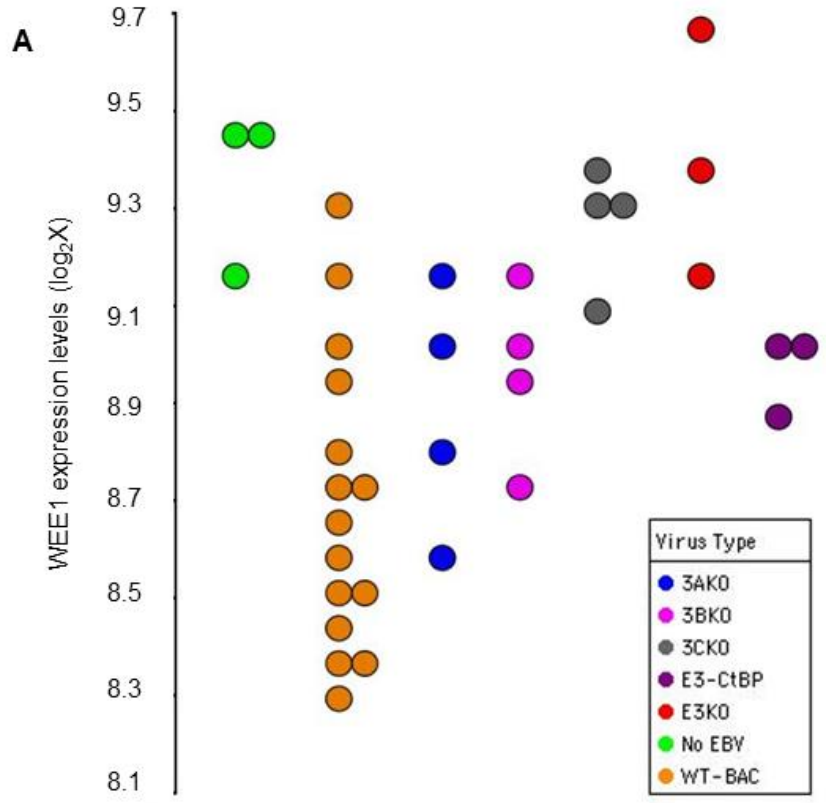


Figure 5-19 *WEE1* is regulated by EBNA 3C in BL background. Gene expression microarray data displayed as a dot plot from http://www.epstein-barrvirus.org.uk compiled by Dr Rob White. Plot shows raw expression data of *WEE1* from BL31 cells uninfected (green), infected with wild type BAC virus (wt-BAC, orange), EBNA 3A knock out BAC virus (3AKO, blue), EBNA 3B knock out BAC virus (3BKO, pink), EBNA 3C knock out BAC virus (3CKO, grey), EBNA 3A/B/C knock out BAC virus (E3KO, red) or EBNA 3A and EBNA3C CTBP binding defective BAC virus (E3-CtBP, purple). Y axis is log2 values with values less than 3 indicating no expression and values over 8 indicates robust expression.

We detected five distinct peaks of binding for EBNA 2 and 3 proteins by ChIP sequencing and showed that these sites were bound by EBNA 2 and EBNA 3C by ChIP QPCR. We further discovered that *WEE1* is upregulated by EBNA 2 and downregulated by EBNA 3C. Subsequent chromosome confirmation capture experiments performed by Dr David Wood showed that this regulation is mediated through the modulation of enhancer-promoter looping (McClellan et al. 2013). In EBV negative BL31 cell lines where *WEE1* is expressed (Fig 5-10 A), no looping between the promoter and enhancers was observed. Cells infected with wild type EBV BAC expressing all latent genes, enhancer-promoter looping was detectable from each enhancer to the promoter. In these cells expression of *WEE1* is repressed (Fig 5-10 A). Infection with an EBV BAC with *EBNA 3C* deleted did not promote enhancer-promoter looping, in these cells *WEE1* is expressed despite the presence of all other latent proteins (Fig 5-10 A). BL31 cells infected with a EBV BAC that has had *EBNA 3C* removed and subsequently restored also showed looping from both enhancers to the *WEE1* promoter. These cells exhibit re-repression of *WEE1* (Fig 5-10 A). We therefore propose a model where in wild type infected cells EBNA 3C is able to bind predominantly to peak 5, promoting looping between both enhancers and the promoter and gene repression (Fig 5-20 A). Wild type cells lacking functional EBNA 2 showed increased repression indicating that EBNA 2 may be able to compete with EBNA 3C and partially inhibit the formation of a repressive loop (Fig 5-20 B). In cells with no functional EBNA 3C, WEE1 is derepressed and no looping between the promoter and enhancers is detected indicating that EBNA 3C is primarily responsible for loop formation (Fig 5-20 C).

WEE1 was first discovered through characterisation of *Saccharomyces cerevisiae* mutants that produced cells half the length of a wild type and were partly impaired for cell viability and growth rate (Thuriaux et al. 1978). It was further found that of the mutants bearing the wee phenotype, 51 mapped to *WEE1* and one mapped to *cdc2* (*CDK1*) implicating interplay between these proteins. From this data *WEE1* was proposed to be involved in the rate limiting control determining when mitosis occurs and to be a negative regulator of the cell cycle (Nurse and Thuriaux 1980). Elevated levels of *WEE1* transcription during the hyperproliferative stage of infection may be a response to DNA damage induced by aberrant proliferation to delay entry into mitosis. EBNA 3C repression of *WEE1* may be counteracting WEE1 induced cell cycle arrest following *WEE1* activation by the DNA damage response. However, this does not explain the role of EBNA 2 transcriptional activation of *WEE1*. Furthermore, EBNA 2 protein is detectable 12 hours post infection and is at a level similar to that of LCLs by 32 hours (Alfieri et al. 1991). EBNA 3C is then transcribed after EBNA 2 mediated promoter switching from Wp to

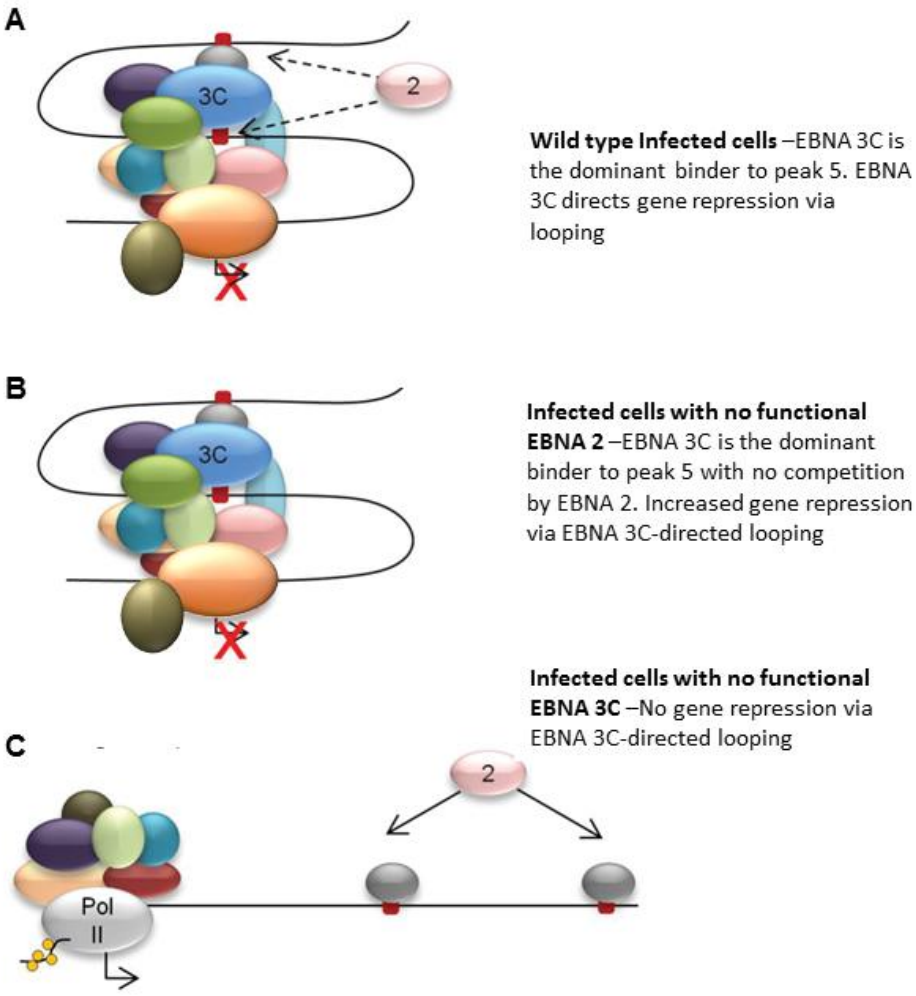


Figure 5-20 Model for regulation of WEE1 via enhancer looping mediated by EBNA 2 and 3C binding. (A) model for EBV modulation of *WEE1* promoter enhancer looping in wild type infected cells. (B) model for EBV modulation of *WEE1* promoter enhancer looping in infected cells lacking functional EBNA 2. (C) model for EBV modulation of *WEE1* promoter enhancer looping in infected cells lacking EBNA 3C. Distal enhancers are displayed as a red box. Promoter is displayed as an arrow with a cross through it if it is repressed. Unknown cellular binding partner of EBNA 2 or 3C is displayed as a grey oval. EBNA 2 is shown as a pink oval with 2 in and EBNA 3C as a blue oval with 3C in. Components of transcriptional machinery and looped DNA are annotated in situations where *WEE1* expression is de-repressed. Unknown repressive complex proteins and promoter enhancer looping are displayed in situations where *WEE1* transcription is repressed.

Cp at around 48 hours (Woisetschlaeger et al. 1991) so precedes the events described by Nikitin *et al* (Nikitin et al. 2010). Given this information *WEE1* is likely to be regulated through a change in pattern of cellular transcription factor expression or chromatin landscape. It would be informative to examine epigenetic modifications at the *WEE1* gene and enhancers during the time course of the initial stages of infection.

Together this data and published work suggests that EBNA 3C directly represses *WEE1* via a repressive promoter enhancer loop shortly after initial infection. A period of hyperproliferation after initial infection induces a DNA damage response and over expression of *WEE1* which is subsequently attenuated by EBNA 3C. It may be that part of this attenuation is due to *WEE1* transcriptional repression by EBNA 3C. It is not clear exactly when during the period of hyperproliferation EBNA 3C binds to and represses *WEE1* and should be investigated further.

5.7.3 ITGAL

ITGAL plays a vital role in directing T and B lymphocytes to sites of infection or damage and mediates pathways such as apoptosis through cell-cell contacts. Secreted chemokines can attract immune cells by increasing their adhesion to local endothelial cells via integrin signalling molecules including ITGAL and integrin β2 (collectively known as LFA-1) interaction with ICAM1 (reviewed in (Newton and Dixit 2012)). LFA-1 is essential to lymphocyte migration through endothelial cells and is involved in rolling, weak adhesion, firm adhesion and transmigration. Furthermore LFA-1 has been previously described to be upregulated by LMP1, mimicking B-cell activation (Wang et al. 1987; Wang et al. 1990a).

At *ITGAL* we found three distinct promoter proximal binding sites for EBNA 2 and 3 proteins by ChIP seq data. EBNA 2 and EBNA 3B were confirmed to bind to peak C in both Mutu III and PER253 LCLs. Interestingly, EBNA 3C was found to only significantly bind in the BL cell line Mutu III and not in PER253 LCLs. The mechanism of this is unknown but it is possible that a particular TF required by EBNA 3C to bind this site is absent in the PER253 LCL, or that the binding of another factor specifically occludes EBNA 3C binding in this cell line. These data were confirmed by analysis of cDNA data in which we could only detect regulation of *ITGAL* by EBNA 3B in LCLs whereas both EBNA 3B and EBNA 3C were shown to regulate it in a BL line (Fig 5-13). This is consistent with a model where gene regulation is determined by specific binding of EBNA 2 and EBNA 3 proteins.

Unlike *CTBP2* and *WEE1*, *ITGAL* was only bound by promoter proximal sites so there was no need to perform CCC experiments. However, subsequent luciferase reporter assays performed

in our lab by Opeoluwa Ojeniyi, in which the *ITGAL* promoter and binding sites were cloned upstream of the luciferase gene, revealed EBNA 2 mediated activation of *ITGAL* (McClellan et al. 2013). EBNA 2 dependent activation was inhibited in the presence of any of the EBNA 3 family of proteins. In contrast, ChIP assays performed in chromatin context revealed only EBNA 3B and EBNA 3C were able to bind and repress this locus. Luciferase assays were conducted out of chromatin context, and this data may suggest that EBNA 3A binding and repression of *ITGAL* may require certain histone modifications. Similarly to *CTBP2* and *WEE1*, the repressive effect of EBNA 3 proteins appears to be dominant at *ITGAL* shared EBNA 2 and EBNA 3 binding sites.

Given the prominent role of LFA-1 in leukocyte migration, EBNA 2 and EBNA 3 regulation of *ITGAL* could facilitate infected B-cell entry into the GC. LMP1 and EBNA 2 mediated activation of *ITGAL* transcription could be responsible for producing an activated B-cell phenotype which begins the process of differentiation and migration to the GC. As EBNA 2 expression precedes that of EBNA 3 proteins in vivo (Woisetschlaeger et al. 1991), it is possible that repression of *ITGAL* is required at a later time point, perhaps evading the immune response by inhibiting interactions with CD8+ and CD4+ T cells. EBNA 3B and EBNA 3C mediated repression may also involve recruitment of PcG proteins and persistent repression of *ITGAL*. This may then aid survival in the GC after EBV loses latency III gene expression.

Here we provide important mechanistic information into how EBV deregulates the host transcriptome. We show that EBNA 2 and EBNA 3 proteins bind to enhancer elements and others in the lab have shown that this can direct regulation of target genes by stimulating enhancer-promoter looping. We also demonstrate that binding of EBNA 2 and 3 proteins is cell type specific, suggesting that the balance of cellular TFs and cofactors within a cell modulates which sites are targeted. Binding is also context specific, and EBNA 3 proteins do not all bind to the same sites within a cell, which indicates that each EBNA 3 has either differential binding partners, or can respond to different histone modifications at a site. ReChIP experiments have revealed that despite EBNA 2 and EBNA 3 proteins targeting many shared sites they do not appear to bind to them at the same time, suggesting that part of the transcriptional regulation by these factors may be through competitive binding. Finally, we have used our ChIP sequencing data to uncover some of the mechanisms employed by EBV to regulate *Bim*, *ADAM*, *CTBP2*, *WEE1* and *ITGAL*, all of which have relevance to EBV induced transformation and immortalisation.

6. Discussion

The EBV EBNA 2 and EBNA 3 family are potent regulators of transcription; and EBV mediated infection and immortalisation is dependent on the activity of EBNA 2, EBNA 3A and EBNA 3C (Cohen et al. 1989; Tomkinson and Kieff 1992; Tomkinson et al. 1993; Young and Murray 2003). Upon infection of a resting B cell EBNA2 and the EBNA 3 family alter the cellular transcriptome facilitating B-cell activation like changes causing entry of the infected cell into the GC; eventually resulting in lifelong persistent infection. In this study we characterise EBNA 2 and EBNA 3 binding sites in the human genome and investigate their functional implications on transcription.

6.1 EBNA 2 and EBNA 3 binding sites in the human genome

ChIP sequencing was carried out in the Mutu III cell line using antibodies against EBNA 2 and a pan specific antibody which precipitates EBNA 3A, 3B and 3C revealed over 21000 EBNA 2 and over 7000 EBNA 3 binding sites in the human genome. These sites are predominantly distal to gene TSSs and 25% of all binding sites are shared between EBNA 2 and EBNA 3 proteins. Recently ChIP-sequencing experiments have shown that enhancer elements are more common than originally thought. Enhancers can modulate transcription regardless of the orientation of the element and can act over distances of several hundred to several thousand base pairs (Banerji et al. 1983). It is now believed that enhancers alone may account for 10% of the human genome (Bulger and Groudine 2011). Modulation of enhancer function plays a role in cancer development as a number of single nucleotide polymorphisms at cancer risk loci map to known enhancer elements (Pomerantz et al. 2009; Ahmadiyeh et al. 2010). Enhancers commonly consist of a cluster of TF binding sites and interplay between TFs at enhancers is functionally relevant (Thanos and Maniatis 1995). It appears protein bound enhancers are functionally similar to promoter proximal elements and can physically contact the promoter in question via DNA looping. In fact promoter enhancer looping may be the common mechanism by which enhancers exert their transcriptional influence (Vilar and Saiz 2005). The enhancer sites that EBNA 2 and 3 proteins target also co-localise with ChIP sequencing data from the ENCODE project for epigentetic marks associated with enhancers and B-cell transcription factors. These include RBPJκ and PU.1, which are known DNA targeting cellular TFs for the EBNA 2 and EBNA 3 family (Grossman et al., 1994)(Johannsen et al. 1995; Sjoblom et al. 1995; Sjoblom et al. 1998; Zhao and Sample 2000). Other B-Cell transcription factors enriched at binding sites included EBF1 which has also been implicated in targeting EBNA 2 to cellular sites (Zhao et al. 2011).

Many other B-cell transcription factors co-localised with our ChIP-sequencing data sets including PAX5, IRF4 and BATF. However, we have no direct evidence that these TFs were aiding the recruitment of EBNA 2 and 3 proteins to binding sites. As EBV induces an activated B-cell phenotype upon infection it is highly likely that many of the enhancers it occupies are involved in B-cell differentiation and development. Such enhancers will also be targeted by B-cell transcription factors is the absence of EBV infection. Indeed, the question of what exactly EBNA proteins are doing differently to normal B-cell transcription factors is a fundamental one. The best characterised TF binding partner of EBNA 2 and 3 proteins is RBPJκ. RBPJκ is bound by intracellular NOTCH which is cleaved at the membrane following activation (Radtke et al. 2010). NOTCH/RBPJκ signalling in B-cells is critical to B-cell development. Strong NOTCH1 signalling drives cells towards a T-cell fate (Dallas et al. 2005) whereas NOTCH2 signalling preferentially selects specific B-cell fates (Saito et al. 2003; Gibb et al. 2010). Furthermore, NOTCH signalling in GC B cells protects B-cells from apoptosis, increasing the chances of the cells survival and subsequent maturation into a circulating memory B-cell (Yoon et al. 2009). Importantly, EBNA 2 can mimic the effects of intracellular NOTCH and perform similar functions at particular genes (Sakai et al. 1998). So are EBNA 2 and 3 proteins simply copying the function, or 'pretending' to be numerous B-cell transcription factors involved in driving B-cells towards the memory compartment? This appears unlikely as EBV is nearly 100% associated with endemic BL suggesting that EBV proteins have a unique regulatory effect compared to any human proteins (Young and Rickinson, 2004).

EBNA 2 and 3 may also play a role in targeting of cellular TFs to aberrant sites. PU.1 can bind its cognate site in condensed chromatin and recruit histone modifying factors to create a poised chromatin state, facilitating binding of other factors such as EBF1, E2A and HATs leading to gene transcription (Zaret and Carroll 2011). However at other loci PU.1 is unable to bind its cognate sequence and a requirement for additional factors, histone marks or non-coding RNA molecules, has been shown (Choukrallah and Matthias 2014). EBNA 2 and 3 activity could therefore modulate which sites can be bound by cellular TFs. This may in fact prove a great problem for the study of human DNA elements as the majority of ChIP sequencing experiments conducted in the ENCODE project, including the data set to which our ChIP sequencing data is aligned to, was conducted in EBV immortalised LCLs expressing a full latency III pattern of gene expression. If EBV is capable of targeting cellular TFs to aberrant sites many known enhancer sites described in ENCODE could in fact be EBV specific.

These mechanisms may play a part in EBNA 2 and 3 interactions with cellular TFs and enhancers but the most likely mechanism of transcriptional regulation is by altering bound TF function, and aberrant recruitment of regulatory cofactors to TF binding sites by EBNA 2 and EBNA 3. EBNA 2 is known to bind to the repression domain of RBPJκ without interfering with the RBPJκ DNA binding domain thereby activating repressed RBPJκ targets (Hsieh and Hayward 1995). Our ChIP sequencing data also revealed an association of EBNA 2 and EBNA 3 bound sites with epigenetic marks associated with active and poised enhancers. Given that the repressive function of EBNA 3A and 3C is best characterised by the activity of PcG group proteins it was surprising that H3K27me3 and EZH2 were not shown to be enriched at EBNA 3 binding sites. This may be due to the fact that PcG protein activity results in broad areas of H3K27me3 in the genome rather than defined peaks. However, whether EBNA 2 and 3 proteins cause or are recruited by poised and active marks remains unknown. EBNA 3C can interact with the histone acetylatransferase p300, mediating gene activation by histone acetylation (Subramanian et al. 2002) and EBNA 2 can recruit HATs including p300, CBP and KAT2B (Wang et al. 2000). This suggests that EBNA 2 and 3 proteins are capable of altering the chromatin at which they are bound to create poised and activating enhancer marks. It is therefore likely that EBNA2 and 3 proteins are recruited to and create active and poised enhancers.

To determine the significance of EBNA 2 and 3 genomic associations we compared our ChIP sequencing data with published gene expression microarray data on EBNA 2 and EBNA 3 proteins. A gene was determined as bound by EBNA 2 or EBNA 3 proteins if it was the most proximal gene to a peak. Our analysis revealed that 8386 genes were bound by EBNA 2 and 3937 were bound by EBNA 3 and that in some cases there was remarkable overlap between bound and regulated genes. This was especially true in EBNA 2 gene expression microarray data sets, each of which had over half of reported regulated genes bound by EBNA 2. In one data set 22/23 (95.7%) genes regulated were bound and in another 56/64 (87.5%). EBNA 3 bound and regulated genes also showed a large degree of overlap. In an EBNA 3C regulated data set 195/430 (45.3%) genes regulated were bound by EBNA 3 proteins. This is especially remarkable given that there are less than half the number of EBNA 3 bound genes than EBNA 2 bound genes. The determination of a gene being bound because it is most proximal to a peak is imperfect as it is documented that enhancers can act over significant genomic distances, interact with multiple promoters and even interact with genes on other chromosomes (Spilianakis et al. 2005; Lomvardas et al. 2006). Furthermore enhancers ‘skip’ genes, contacting and regulating a gene specifically over another that is closer. Also a gene that is bound in our

ChIP sequencing data and regulated in gene expression microarray data sets does not prove that it is the activity of EBNA 2 and 3 proteins at that site coordinating its regulation. Despite this EBNA 2 and 3 binding seems to correlate well with gene regulation.

Therefore we can conclude that EBNA 2 and 3 proteins associate extensively with cellular DNA elements, especially enhancers, in order to regulate host gene transcription.

6.2 Histone modifications and EBNA 3C DNA binding

Analysis of gene expression microarray data set performed in our lab revealed that EBNA 3C specifically represses chemokine signalling and integrin receptor signalling pathways (Table 4-2). We further determined that many of the most repressed genes in these pathways were bound by EBNA 3 proteins from our ChIP sequencing data. It has previously been reported that genes repressed by EBNA 3C show changes in histone modification associated with repression. For example histones at $p16^{INK4A}$ and *Bim* show increased H3K27me3 in the presence of EBNA 3A and EBNA 3C (Paschos et al. 2009; Skalska et al. 2010). This mark is deposited by the PcG protein EZH2 and is proposed to result in recruitment of other chromatin modifying enzymes including PRC1 which can lead to CpG methylation and persistent gene repression (Simon and Kingston 2009). Recruitment of PcG proteins by EBNA 3A and 3C is dependent on their interactions with co-repressors such as CTBP proteins (Touitou et al. 2001; Skalska et al. 2010). Furthermore EBNA 3 target genes have been shown to undergo changes in histone acetylation associated with repression including *NOTCH2* and *TOX* (White et al. 2010).

The gene expression microarray experiment performed in our lab compared mRNA from EBV negative B-cell lymphoma (BJAB) E3C4 stably expressing EBNA 3C with that of the BJAB PZ2 which is transfected with an empty plasmid. We therefore conducted ChIP experiments using antibodies for histone modifications at repressed genes in the related cell lines E3C3, E3C7, PZ1 and PZ3. We also used antibodies against all EBNA 3 proteins as E3C3 and E3C7 only express EBNA 3C and we could be certain that this would be the only EBNA 3 protein we would precipitate. In agreement with current knowledge on EBNA 3C mediated repression we detected changes in histone modifications associated with repression at *ITGB1*, *ITGA4*, *ADAM28*, *ADAMDEC1*, *CXCL10* and *CXCL11* (Fig 4-2 – Fig 4-5). We also detected EBNA 3C binding at the *ADAM* locus, *ITGB1* and *ITGA4* (Fig 4.2 – Fig 4-5). At the time of publication this was the first evidence of EBNA 3C interaction with the human genome.

We could not confirm EBNA 3C binding to two low level EBNA 3 binding sites at the *CXCL* locus (Fig 4-7), furthermore, we could not validate these sites in Mutu III cells using the same

antibody that was used for EBNA 3 proteins in the ChIP sequencing experiment. It may therefore be that these sites are an artefact of ChIP sequencing analysis. This would suggest that regulation of *CXCL10* and *CXCL11* by EBNA 3C is via a long range interactions or an indirect effect of the presence of EBNA 3C. Conversely it could be that the ChIP sequencing experiment is more sensitive than our conventional ChIP method and this site is indeed responsible for the direct repression of *CXCL10* and *CXCL11*.

Increases in levels of H3K27me3 and decreases in H3ac and H4ac were almost ubiquitous across repressed loci. An interesting observation was that in EBNA 3C expressing cells levels of H3K27me3 were increased more downstream of the TSS rather than at the promoter (see Fig 4.4 C, Fig 4-5 C and Fig 4-8 A for examples). Furthermore, at the two distal peaks studied there was a limited increase in H3K27me3 in the presence of EBNA 3C compared to control cells (Fig 4-3 C and Fig 4-5 C). This supports our finding that EBNA 3 peaks are not enriched for H3K27me3 (Fig 3-10). These observations are however difficult to explain. If EBNA 3C is associating with a TF to bind a distinct site, then recruiting co-regulators like CTBP2 which then recruits PcG proteins it would be expected that high levels of H3K27me3 at EBNA 3C binding sites would be observed. It may be that PcG proteins are indirectly recruited to EBNA 3C sites and then loop to intragenic sites to deposit H3K27me3. It may also be that the effect of EBNA 3C is to create poised chromatin at silencers facilitating repressive complex recruitment. The delineation between enhancers and silencers is not a strong one and whether a DNA element acts as an enhancer or silencer is likely dependent on what is bound to it at the time (Murayama et al. 2004). Acetylation of histones H3 and H4 was found to be most abundant at promoters and at the *ADAM* locus distal site EBNA 3C binding site (Fig 4-3 – Fig 4-6). H3ac and H4ac was almost always reduced in cells expressing EBNA 3C at sites that were acetylated in PZ1 and PZ3. This suggests that EBNA 3C promotes chromatin condensing excluding binding of activating proteins to promoters and enhancers.

It is important to note that we have not proved that EBNA 3C binding directly recruits proteins that repress these loci. It is possible that EBNA 3C binding to investigated sites is coincidental and the presence of repressive histone modifications and gene repression in EBNA 3C expressing cells is a downstream effect of a different target of EBNA 3C regulation. However, as discussed previously, binding does appear to strongly correlate with regulation from analysis of published microarray data (Table 3-2). Furthermore we know that EBNA 3C can associate with histone deacetylases and repressive cofactors that can in turn recruit PcG proteins (reviewed in (Allday 2013)). CCC experiments also showed that at the *ADAM* locus,

EBNA 3C alone was able to direct enhancer promoter loops to both *ADAM28* and *ADAMDEC1*. This data, along with confirmed EBNA 3C binding (Fig 4-5 B), transcriptional repression and associated changes in histone modifications (Fig 4-5 C and Fig 4-6) allows a near complete mechanism for repression of *ADAM28* and *ADAMDEC1* by distal EBNA 3C binding sites. Future experiments at this locus should attempt to determine which cellular TF targets EBNA 3C to the enhancer and what proteins recruited by EBNA 3C are essential for epigenetic repression.

The relevance of this data *in vivo* is more difficult to determine. Chemotaxis assays performed in our lab by Sarika Khasnis showed that EBNA 3C expressing BJAB cells caused reduced attraction of CXCR3 (the receptor for CXCL10 and CXCL11) expressing cells compared to control cells (McClellan et al. 2012). Integrin receptor signalling may also play a role in immune evasion as lymphocyte infiltration into infected areas is facilitated by interactions between α4β1 integrin, VCAM-1 and ADAM28 (Bridges et al. 2002; Bridges et al. 2005). Furthermore, integrin signalling is important for cell adhesion and migration so its regulation may allow EBV infected B-cells entry into the GC. Once in the GC integrin receptor signalling is crucial in providing adequate survival signals for B-cells via T-helper cells (Taylor et al. 1999; Yoon et al. 2014). It is important to note that by the time an EBV infected B-cell gets to the GC it no longer expresses EBNA 3C. However, at *Bim* epigenetic marks and gene repression persist after conditional EBNA 3C is removed, and recently epitope-tagged EBNA 3C has been shown to bind to the *Bim* promoter (Paschos et al. 2009; Paschos et al. 2012). We further know that EBNA 3C can physically associate with CTBP proteins and that the CTBP protein binding domain of EBNA 3C is required for repression of *Bim* (Touitou et al. 2001). Future experiments at these loci should focus on the timing of the EBNA 3C mediated repression via H3K27me3 and look for evidence for subsequent CpG methylation. Furthermore, analysis of this gene subset in other cell lines expressing a full panel of latent gene expression will be necessary to describe the function of EBNA 2, which has been shown to activate most genes investigated in this chapter (Table 4-3). In conclusion, this data suggests that association of EBNA 3C with DNA elements results in the recruitment of cofactors which alter the chromatin landscape and repress transcription.

6.3 The influence of shared sites and DNA looping

Previous studies at the viral *LMP1* promoter have shown co-activation by EBNA 2 and EBNA 3C which is entirely dependent on PU.1 binding sites at the gene promoter (Zhao and Sample 2000). Initial analysis of ChIP sequencing data revealed that 25% of all binding sites for EBNA 2 and EBNA 3 proteins were shared sites (Fig 3-8 C). This was perhaps not surprising as EBNA 2

and 3 proteins are known to share cellular TF binding partners. This data provided us with the first opportunity to study how EBNA 2 and 3 binding to shared genomic sites modulates transcription. To do this we required specific EBNA 3A, EBNA 3B and EBNA 3C antibodies. We confirmed antibody specificity by immunoprecipitation experiments using ChIP conditions and buffers in the BL31 cell series which showed that antibodies used did not co-precipitate other EBNA 3 proteins (Fig 5-1). We first validated these antibodies in ChIP experiments at EBNA 3 only binding sites at the *ADAM* intragenic peak and the *Bim* promoter. Only EBNA 3A and EBNA 3C were shown to bind to these sites (Fig 5-2 and Fig 5-3). Importantly, *ADAM28*, *ADAMDEC1* and *Bim* have only been reported to be regulated by EBNA 3A and EBNA 3C and not EBNA 3B (Hertle et al. 2009; White et al. 2010; McClellan et al. 2012).

We then investigated shared binding sites at *CTBP2*, *WEE1* and *ITGAL* and found that EBNA 2 bound to every binding site discovered by ChIP sequencing in both Mutu III and PER253 LCL (Fig 5-4, 5-5, 5-8, 5-9, 5-11 and 5-12). EBNA 3 binding was more varied and showed locus specific and cell type dependent binding which correlated with gene repression. Only EBNA 3C was shown to bind significantly and regulate *WEE1* (Fig 5-8 and Fig 5-9). At *ITGAL* EBNA 3B and EBNA 3C were shown to bind in Mutu III cell lines, correlating with gene repression (Fig 5-11 and Fig 5-13). In PER 253 LCL lines only EBNA 3B showed significant binding at *ITGAL*, which was only found to be regulated in LCL lines by EBNA 3B (Fig 5-12 and Fig 5-13). This striking correlation between binding and regulation provides a mechanism for differential gene regulation by EBNA 3 proteins.

Extensive overlap in genes regulated by EBNA 3 proteins has been reported (White et al. 2010). Interestingly, in this study genes only regulated by a single EBNA 3 protein or any combination of pairs of EBNA 3 proteins were also described. EBNA 3C appears to be especially important in EBNA 3 mediated regulation as only 9/1201 genes were reported to be regulated by EBNA 3A and EBNA 3B but not EBNA 3C. Furthermore, EBNA 3C was shown to regulate the most genes without another EBNA 3 protein regulating the same gene (see (White et al. 2010) Fig 1.). Here we provide a simple mechanism to explain this finding, that genes regulated by EBNA 3B are targeted by an EBNA 3B binding site. Genes regulated by EBNA 3A and EBNA 3C are targeted by an EBNA 3A and EBNA 3C binding site and so on.

Unfortunately, we have very little evidence about how differential binding of EBNA 2 and 3 proteins occurs. The obvious answer is that they associate with different cellular TFs. However, to date most research has been conducted on known TF binding partners, particularly RBPJκ. This study and others indicate a number of other B-cell transcription factor consensus motifs

and binding by published ChIP sequencing being enriched at EBNA 2 and 3 binding sites. This however is not sufficient to prove these cellular TFs are targeting EBNA 2 and 3 proteins to DNA. One candidate TF for differential targeting of EBNA 2 and 3 proteins is EBF1 which has been implicated in targeting EBNA 2 to genomic sites (Zhao et al. 2011) but to date shows no evidence of EBNA 3 proteins association. Furthermore the most significant TF consensus motif for the top 300 most significant EBNA 2 binding sites in our ChIP sequencing data set was EBF1 (Fig 3-15). This suggests that differential EBNA 2 and 3 binding and correlated regulation may depend on which TF they are able to associate with, the prediction being that at each has at least one TF that only it can associate with. EBF1 should be investigated further as a potential EBNA 2 only targeting TF. It is also possible that the binding of EBNA 2, EBNA 3A, EBNA 3B and EBNA 3C to the same cellular TF alters which sites that particular TF binds to. Both mechanisms for differential gene targeting are likely true. The different EBNA 2 and 3 proteins almost certainly have different affinities for different cellular TFs and are also likely to be able to individually specify the same TF to particular DNA elements. ChIP sequencing studies using specific EBNA 3A, EBNA 3B and EBNA 3C antibodies would allow detection of EBNA 3A, EBNA 3B and EBNA 3C only sites. These could then be compared to existing ChIP sequencing data on cellular TFs and enriched TF consensus motifs at EBNA 3A, EBNA 3B or EBNA 3C only sites could be compared to identify candidate TFs.

One fundamental question we wished to address was do the EBNA 2 and 3 proteins bind to the same site at the same time as part of a multi-EBNA2/3 complex. Using ReChIP exeriments and QPCR at *CTBP2*, *WEE1* and *ITGAL* we show here that EBNA 2 does not appear to bind to shared sites at these genes at the same time as EBNA 3 proteins (Fig 5-14 – Fig 5-16). This suggests that mutually exclusive binding of EBNA 2 and EBNA 3 proteins has functional relevance. After publication of this data an EBNA 3A binding site was identified in an intergenic region between *CXCL9* and *CXCL10* (Harth-Hertle et al. 2013). This site was also bound by EBNA 2 and RBPJκ in ChIP sequencing data published from another group (Zhao et al. 2011). By using a HA-tagged Dox inducible EBNA 3A they showed that EBNA 2 occupancy of this enhancer was reduced 24 hours after Dox induction with EBNA 3A, with a corresponding increase in HA-tagged EBNA 3A binding to this site. This data further suggests that EBNA 2 and EBNA 3 proteins cannot bind a single site at the same time. This may be a site specific phenomena and evidence that this is true at every shared site will be difficult to obtain.

We could not perform reChIP experiments to observe if EBNA 3A, EBNA 3B and EBNA 3C were able to bind the same site at the same time. This is partly because the EBNA 3 antibodies

generally precipitate less DNA than the EBNA 2 antibody, either due to reduced affinity for the epitopes, less EBNA 3 protein or less EBNA 3 genomic associations. This is problematic in reChIP experiments as DNA is lost at each precipitation, wash and elution step in a conventional ChIP, and a reChIP has double the amount of these steps to a conventional ChIP. To get an estimate of how much DNA is lost it is possible to compare percentage ChIP input signals in Fig 5-4 with Fig 5-14, Fig 5-8 with Fig 5-15 and Fig 5-11 with Fig 5-16. This is in no way quantitative as these data are from two different ChIP experiments. This reveals that signal at binding sites is fairly consistently reduced by 10 fold or more by performing a secondary ChIP. Recent data has reported that EBNA 3A is co-precipitated by antibodies against a TAP-tagged EBNA 3C in IP experiments (Paschos et al. 2012). Also, EBNA 3A and 3C have been reported to physically interact in a yeast two hybrid system and colocalise in immunofluorescence studies (Krauer et al. 2004a; Calderwood et al. 2007). However, this does not show that EBNA 3A and EBNA 3C target DNA at the same time. Future experiments should focus on improving the ReChIP technique. The creation of an oligo library which is fused to two different EBNA proteins would help as this would act as a perfect positive control. Optimisation of this experiment or using a scaled up ChIP method may also help to validate EBNA 3 differential binding via reChIP.

Follow up CCC experiments performed in our lab by Dr. David Wood (University of Sussex) revealed that promoter enhancer looping was mediated by EBNA 3A and EBNA 3C at *CTBP2* and *WEE1* respectively. In wild type infected LCLs in which *CTBP2* is repressed, no loop between the intragenic enhancer and promoter was detected (Fig 5-7 A and (McClellan et al. 2013)). In cells infected with EBNA 3AKO virus in which *CTBP2* is derepressed, a loop between the promoter and enhancer was detected. This directly implicates EBNA 3A in the repression of *CTBP2* by EBNA 3A by the prevention of activating loop formation. CCC experiments carried out in EREB 2.5 expressing a conditionally active EBNA 2 did not detect promoter enhancer looping with or without EBNA 2 but this is perhaps not surprising as EBNA 2 does not bind to this site in EREB 2.5 cells (McClellan et al. 2013). We therefore do not know anything about EBNA 2 function at *CTBP2* as we could not detect *CTBP2* expression in any cell panel with conditional or knock out EBNA 2 (Fig 5-7 C). We also do not know the role of EBNA 3B or EBNA 3C in this interaction although we present evidence here that EBNA 3B can repress *CTBP2* transcription (Fig 5-7 B). A few possibilities exist as to the role of differential binding at *CTBP2*. Given that EBNA 3B can repress *CTBP2* it is possible that EBNA 3B is able to partially prevent activating loop formation but as the CCC experiments performed were not quantitative we could not detect this. It is also possible that EBNA 3B represses *CTBP2* via mechanisms

completely unrelated to EBNA 3A mediated looping, or that EBNA 3B repression requires EBNA 3A mediated looping and vice versa. Furthermore, as repression of *CTBP2* by EBNA 3B was detected in a different cell line it is possible that which EBNA 3 performs repression via promoter enhancer looping is cell type specific. To improve our understanding of differential binding and promoter enhancer looping it will be necessary to analyse CCC experiments using quantitative techniques such as QPCR. In the case of *CTBP2* it would also be informative to explore cell lines with conditional EBNA 3C and EBNA 2 which express *CTBP2*, but to our knowledge, no such cell lines exist.

Promoter-enhancer looping at *WEE1* was shown to be entirely dependent on the presence of EBNA 3C. Here, EBNA 3C mediated promoter enhancer looping was associated with gene repression (McClellan et al. 2013). This is, to our knowledge, the first demonstration of repressive promoter enhancer DNA looping. Data presented here shows that EBNA 2 has an activating role at this locus although the mechanism of this is unknown. It may be that EBNA 2 partially prevents EBNA 3C mediated promoter enhancer looping by directly competing for binding sites (Fig 5-10 B, C).

At *ITGAL* we observed a cell type specific difference in binding correlated with a cell type specific difference in regulation (Fig 5-11 E and Fig 5-12 E). Cell type specific differences in regulation have been reported in gene expression microarray data for EBNA 3B (White et al. 2010) and EBNA 2 (Maier et al. 2006). This data offers a simple explanation for these findings; in one cell line it binds, in the other it does not bind. This is made more complex by the high proportion of shared sites between EBNA 2 and EBNA 3 such that one protein not binding may increase binding of other factors, as demonstrated at *CXCL9* and *CXCL10* (Harth-Hertle et al. 2013). The fundamental question that needs to be addressed now is why is binding different between cell lines. An obvious answer is epigenetic silencing of a site in one cell line and not another, and histone modifications undoubtedly play a role, but at ITGAL EBNA 3B was shown to bind to the promoter in both cell types, whereas EBNA 3C could not. This suggests that the chromatin environment was specifically permissive to EBNA 3B at this site. If it is true that each EBNA has its own set of TF binding partners then relative abundances of TFs between cell lines could also account for these data. It is likely that cell type specific binding will only start to be understood once ChIP sequencing experiments have been carried out on multiple, well characterised, EBV infected cell lines. One obvious candidate is GM12878 and other EBV infected cells from which the ENCODE project is constructed. It is possible to hypothesise that

cell type specific binding will be a mixture of: expression levels of individual EBNAs, expression levels of their TF binding partners and histone modifications at target sites.

There are many *in vivo* implications of this work. Apart from showing the mechanisms of EBNA 2 and EBNA 3 mediated transcriptional regulation we have identified novel targets of regulation from this ChIP sequencing data. *CTBP2* has never been reported to be regulated by EBNA 2 or EBNA 3 proteins, indeed previous work in the EBV field has failed to distinguish between *CTBP1* and *CTBP2*. Our findings strongly suggests that we should as EBNA 3A seems to preferentially repress *CTBP2* by preventing activating loop formation. Furthermore this work allows us to understand how genes are regulated. *WEE1* expression was reported to be induced in the hyperproliferation phase shortly after infection and is then repressed. The data presented here shows this may be at least partly due to EBNA 3C mediated repressive loop formation. *ITGAL* is expressed upon EBV infection resulting in increased surface LFA-1 and an activated B-cell phenotype, our data suggests that at some point in EBV life cycle this is required to be repressed, primarily by EBNA 3B via binding to promoter proximal sites. *ADAM28* and *ADAMDEC1* show similar patterns of regulation in array data and both are consistently repressed by EBNA 3A and EBNA 3C. Here we find that EBNA 3A and EBNA 3C bind to an intragenic site and EBNA 3C is able to coordinate repressive loops to the promoter of each gene and stimulate epigenetic silencing of the whole locus.

The data presented here is of crucial importance to understanding how EBV regulates the host transcriptome. We show that EBNA 2 and 3 proteins associate extensively with the human genome, primarily targeting enhancer elements. We provide evidence for gene-specific and cell background specific binding of EBNA 3 proteins being functionally relevant to the transcriptional regulation of target genes, and that genes targeted have extensive *in vivo* importance to EBV induced transformation and biology. The transcriptional control processes described here provide novel mechanisms of host-cell reprogramming by EBV and further the understanding of transcription in general. The ChIP sequencing data is now publically available and we hope it will be a useful resource for others in the field.

7. Bibliography

2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489:57-74.

Adhikary, D., U. Behrends, A. Moosmann, K. Witter, G. W. Bornkamm, and J. Mautner. 2006. Control of Epstein-Barr virus infection in vitro by T helper cells specific for virion glycoproteins. The Journal of experimental medicine 203:995-1006.

Adolfsson, J., O. J. Borge, D. Bryder, K. Theilgaard-Monch, I. Astrand-Grundstrom, E. Sitnicka, Y. Sasaki, and S. E. Jacobsen. 2001. Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. Immunity 15:659-669.

Adolfsson, J., R. Mansson, N. Buza-Vidas, A. Hultquist, K. Liuba, C. T. Jensen, D. Bryder, L. Yang, O. J. Borge, L. A. Thoren, K. Anderson, E. Sitnicka, Y. Sasaki, M. Sigvardsson, and S. E. Jacobsen. 2005. Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. Cell 121:295-306.

Aghdassi, A., M. Sendler, A. Guenther, J. Mayerle, C. O. Behn, C. D. Heidecke, H. Friess, M. Buchler, M. Evert, M. M. Lerch, and F. U. Weiss. 2012. Recruitment of histone deacetylases HDAC1 and HDAC2 by the transcriptional repressor ZEB1 downregulates E-cadherin expression in pancreatic cancer. Gut 61:439-448.

Ahmadiyeh, N., M. M. Pomerantz, C. Grisanzio, P. Herman, L. Jia, V. Almendro, H. H. He, M. Brown, X. S. Liu, M. Davis, J. L. Caswell, C. A. Beckwith, A. Hills, L. Macconaill, G. A. Coetzee, M. M. Regan, and M. L. Freedman. 2010. 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. Proceedings of the National Academy of Sciences of the United States of America 107:9742-9746.

Akashi, K., D. Traver, T. Miyamoto, and I. L. Weissman. 2000. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. Nature 404:193-197.

Alfieri, C., M. Birkenbach, and E. Kieff. 1991. Early events in Epstein-Barr virus infection of human B lymphocytes. Virology 181:595-608.

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. 2013. EBV finds a polycomb-mediated, epigenetic solution to the problem of oncogenic stress responses triggered by infection. Frontiers in genetics 4:212.

Allis, C. D., S. L. Berger, J. Cote, S. Dent, T. Jenuwien, T. Kouzarides, L. Pillus, D. Reinberg, Y. Shi, R. Shiekhattar, A. Shilatifard, J. Workman, and Y. Zhang. 2007. New nomenclature for chromatin-modifying enzymes. Cell 131:633-636.

Altmann, M., and W. Hammerschmidt. 2005. Epstein-Barr virus provides a new paradigm: a requirement for the immediate inhibition of apoptosis. PLoS Biol 3:e404.

Altmann, M., D. Pich, R. Ruiss, J. Wang, B. Sugden, and W. Hammerschmidt. 2006. Transcriptional activation by EBV nuclear antigen 1 is essential for the expression of EBV's transforming genes. Proceedings of the National Academy of Sciences of the United States of America 103:14188-14193.

Andersson, J. 2000. An Overview of Epstein-Barr Virus: from Discovery to Future Directions for Treatment and Prevention. Herpes 7:76-82.

Anderton, E., J. Yee, P. Smith, T. Crook, R. E. White, and M. J. Allday. 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.

Atchison, M. L. 1988. Enhancers: mechanisms of action and cell specificity. Annual review of cell biology 4:127-153.

Back, J., A. Dierich, C. Bronn, P. Kastner, and S. Chan. 2004. PU.1 determines the self-renewal capacity of erythroid progenitor cells. Blood 103:3615-3623.

Bain, G., E. C. Robanus Maandag, H. P. te Riele, A. J. Feeney, A. Sheehy, M. Schlissel, S. A. Shinton, R. R. Hardy, and C. Murre. 1997. Both E12 and E47 allow commitment to the B cell lineage. Immunity 6:145-154.

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

Banerji, J., L. Olson, and W. Schaffner. 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. Cell 33:729-740.

Barski, A., S. Cuddapah, K. Cui, T. Y. Roh, D. E. Schones, Z. Wang, G. Wei, I. Chepelev, and K. Zhao. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129:823-837.

Basu, A., and M. L. Atchison. 2010. CtBP levels control intergenic transcripts, PHO/YY1 DNA binding, and PcG recruitment to DNA. Journal of cellular biochemistry 110:62-69.

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

Bell, A. C., A. G. West, and G. Felsenfeld. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell 98:387-396.

Ben-Porath, I., M. W. Thomson, V. J. Carey, R. Ge, G. W. Bell, A. Regev, and R. A. Weinberg. 2008. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nature genetics 40:499-507.

Benninger-Doring, G., S. Pepperl, L. Deml, S. Modrow, H. Wolf, and W. Jilg. 1999. Frequency of CD8(+) T lymphocytes specific for lytic and latent antigens of Epstein-Barr virus in healthy virus carriers. Virology 264:289-297.

Benveniste, D., H. J. Sonntag, G. Sanguinetti, and D. Sproul. 2014. Transcription factor binding predicts histone modifications in human cell lines. Proceedings of the National Academy of Sciences of the United States of America.

Bergman, L. M., C. N. Birts, M. Darley, B. Gabrielli, and J. P. Blaydes. 2009. CtBPs promote cell survival through the maintenance of mitotic fidelity. Molecular and cellular biology 29:4539-4551.

Bernasconi, N. L., E. Traggiai, and A. Lanzavecchia. 2002. Maintenance of serological memory by polyclonal activation of human memory B cells. Science 298:2199-2202.

Bickham, K., C. Munz, M. L. Tsang, M. Larsson, J. F. Fonteneau, N. Bhardwaj, and R. Steinman. 2001. EBNA1-specific CD4+ T cells in healthy carriers of Epstein-Barr virus are primarily Th1 in function. The Journal of clinical investigation 107:121-130.

Blake, N., T. Haigh, G. Shaka'a, D. Croom-Carter, and A. Rickinson. 2000. The importance of exogenous antigen in priming the human CD8+ T cell response: lessons from the EBV nuclear antigen EBNA1. Journal of immunology 165:7078-7087.

Boccellato, F., E. Anastasiadou, P. Rosato, B. Kempkes, L. Frati, A. Faggioni, and P. Trivedi. 2007. EBNA2 interferes with the germinal center phenotype by downregulating BCL6 and TCL1 in non-Hodgkin's lymphoma cells. J Virol 81:2274-2282.

Boeger, H., J. Griesenbeck, J. S. Strattan, and R. D. Kornberg. 2003. Nucleosomes unfold completely at a transcriptionally active promoter. Molecular cell 11:1587-1598.

Borghesi, L., J. Aites, S. Nelson, P. Lefterov, P. James, and R. Gerstein. 2005. E47 is required for V(D)J recombinase activity in common lymphoid progenitors. The Journal of experimental medicine 202:1669-1677.

Bornkamm, G. W. 2009. Epstein-Barr virus and the pathogenesis of Burkitt's lymphoma: more questions than answers. Int J Cancer 124:1745-1755.

Bourillot, P. Y., L. Waltzer, A. Sergeant, and E. Manet. 1998. Transcriptional repression by the Epstein-Barr virus EBNA3A protein tethered to DNA does not require RBP-Jkappa. The Journal of general virology 79 (Pt 2):363-370.

Boyd, J. M., T. Subramanian, U. Schaeper, M. La Regina, S. Bayley, and G. Chinnadurai. 1993. A region in the C-terminus of adenovirus 2/5 E1a protein is required for association with a cellular phosphoprotein and important for the negative modulation of T24-ras mediated transformation, tumorigenesis and metastasis. The EMBO journal 12:469-478.

Bray, S. J. 2006. Notch signalling: a simple pathway becomes complex. Nature reviews. Molecular cell biology 7:678-689.

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

Bridges, L. C., P. H. Tani, K. R. Hanson, C. M. Roberts, M. B. Judkins, and R. D. Bowditch. 2002. The lymphocyte metalloprotease MDC-L (ADAM 28) is a ligand for the integrin alpha4beta1. The Journal of biological chemistry 277:3784-3792.

Brink, A. A., D. F. Dukers, A. J. van den Brule, J. J. Oudejans, J. M. Middeldorp, C. J. Meijer, and M. Jiwa. 1997. Presence of Epstein-Barr virus latency type III at the single cell level in post-transplantation lymphoproliferative disorders and AIDS related lymphomas. Journal of clinical pathology 50:911-918.

Bulger, M., and M. Groudine. 2011. Functional and mechanistic diversity of distal transcription enhancers. Cell 144:327-339.

Bulger, M., D. Schubeler, M. A. Bender, J. Hamilton, C. M. Farrell, R. C. Hardison, and M. Groudine. 2003. A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus. Molecular and cellular biology 23:5234-5244.

Burkitt, D. 1958. A sarcoma involving the jaws in African children. Br J Surg 46:218-223.

Burkitt, D. 1962. A children's cancer dependent on climatic factors. Nature 194:232-234.

Calderwood, M. A., K. Venkatesan, L. Xing, M. R. Chase, A. Vazquez, A. M. Holthaus, A. E. Ewence, N. Li, T. Hirozane-Kishikawa, D. E. Hill, M. Vidal, E. Kieff, and E. Johannsen. 2007. Epstein-Barr virus and virus human protein interaction maps. Proceedings of the National Academy of Sciences of the United States of America 104:7606-7611.

Calender, A., M. Billaud, J. P. Aubry, J. Banchereau, M. Vuillaume, and G. M. Lenoir. 1987. Epstein-Barr virus (EBV) induces expression of B-cell activation markers on in vitro infection of EBV-negative B-lymphoma cells. Proceedings of the National Academy of Sciences of the United States of America 84:8060-8064.

Callan, M. F., C. Fazou, H. Yang, T. Rostron, K. Poon, C. Hatton, and A. J. McMichael. 2000. CD8(+) T-cell selection, function, and death in the primary immune response in vivo. The Journal of clinical investigation 106:1251-1261.

Cao, R., Y. Tsukada, and Y. Zhang. 2005. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Molecular cell 20:845-854.

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

Cao, R., and Y. Zhang. 2004. SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. Molecular cell 15:57-67.

Cattoretti G., C. Angelin-Duclos, R. Shaknovich, H. Zhou, D. Wang, B. Alobeid. 2005. PRDM1/Blimp-1 is expressed in human B-lymphocytes committed to the plasma cell lineage. Journal of Pathology 206(1):76-86.

Chan, C. W., A. K. Chiang, K. H. Chan, and A. S. Lau. 2003. Epstein-Barr virus-associated infectious mononucleosis in Chinese children. The Pediatric infectious disease journal 22:974-978.

Chang, C. M., K. J. Yu, S. M. Mbulaiteye, A. Hildesheim, and K. Bhatia. 2009. The extent of genetic diversity of Epstein-Barr virus and its geographic and disease patterns: a need for reappraisal. Virus Res 143:209-221.

Chang, K. H., A. Sanchez-Aguilera, S. Shen, A. Sengupta, M. N. Madhu, A. M. Ficker, S. K. Dunn, A. M. Kuenzi, J. L. Arnett, R. A. Santho, X. Agirre, J. P. Perentesis, M. W. Deininger, Y. Zheng, X. R. Bustelo, D. A. Williams, and J. A. Cancelas. 2012. Vav3 collaborates with p190-BCR-ABL in lymphoid progenitor leukemogenesis, proliferation, and survival. Blood 120:800-811.

Chapman, A. L., A. B. Rickinson, W. A. Thomas, R. F. Jarrett, J. Crocker, and S. P. Lee. 2001. Epstein-Barr virus-specific cytotoxic T lymphocyte responses in the blood and tumor site of Hodgkin's disease patients: implications for a T-cell-based therapy. Cancer Res 61:6219-6226.

Chen, A., M. Divisconte, X. Jiang, C. Quink, and F. Wang. 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., B. Zhao, E. Kieff, J. C. Aster, and F. Wang. 2006. EBNA-3B- and EBNA-3C-regulated cellular genes in Epstein-Barr virus-immortalized lymphoblastoid cell lines. J Virol 80:10139-10150.

Chen, E. J., M. H. Shaffer, E. K. Williamson, Y. Huang, and J. K. Burkhardt. 2013. Ezrin and moesin are required for efficient T cell adhesion and homing to lymphoid organs. PLoS One 8:e52368.

Chen, S., J. Ma, F. Wu, L. J. Xiong, H. Ma, W. Xu, R. Lv, X. Li, J. Villen, S. P. Gygi, X. S. Liu, and Y. Shi. 2012. The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. Genes Dev 26:1364-1375.

Chinnadurai, G. 2009. The transcriptional corepressor CtBP: a foe of multiple tumor suppressors. Cancer research 69:731-734.

Cho, Y. W., T. Hong, S. Hong, H. Guo, H. Yu, D. Kim, T. Guszczynski, G. R. Dressler, T. D. Copeland, M. Kalkum, and K. Ge. 2007. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. The Journal of biological chemistry 282:20395-20406.

Chopra, V. S., D. A. Hendrix, L. J. Core, C. Tsui, J. T. Lis, and M. Levine. 2011. The polycomb group mutant esc leads to augmented levels of paused Pol II in the Drosophila embryo. Molecular cell 42:837-844.

Choukrallah, M. A., and P. Matthias. 2014. The Interplay between Chromatin and Transcription Factor Networks during B Cell Development: Who Pulls the Trigger First? Frontiers in immunology 5:156.

Cieniewicz, A. M., L. Moreland, A. E. Ringel, S. G. Mackintosh, A. Raman, T. M. Gilbert, C. Wolberger, A. J. Tackett, and S. D. Taverna. 2014. The bromodomain of Gcn5 regulates site-specificity of lysine acetylation on histone H3. Mol Cell Proteomics.

Claessens, F., and D. T. Gewirth. 2004. DNA recognition by nuclear receptors. Essays in biochemistry 40:59-72.

Clements, A., A. N. Poux, W. S. Lo, L. Pillus, S. L. Berger, and R. Marmorstein. 2003. Structural basis for histone and phosphohistone binding by the GCN5 histone acetyltransferase. Molecular cell 12:461-473.

Cludts, I., and P. J. Farrell. 1998. Multiple functions within the Epstein-Barr virus EBNA-3A protein. Journal of virology 72:1862-1869.

Clybouw, C., B. McHichi, S. Mouhamad, M. T. Auffredou, M. F. Bourgeade, S. Sharma, G. Leca, and A. Vazquez. 2005. EBV infection of human B lymphocytes leads to down-regulation

of Bim expression: relationship to resistance to apoptosis. Journal of immunology 175:2968-2973.

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

Cohen, J. I., F. Wang, and E. Kieff. 1991. Epstein-Barr virus nuclear protein 2 mutations define essential domains for transformation and transactivation. Journal of virology 65:2545-2554.

Cohen, J. I., F. Wang, J. Mannick, and E. Kieff. 1989. Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation. Proceedings of the National Academy of Sciences of the United States of America 86:9558-9562.

Cohen, M. J., A. F. Yousef, P. Massimi, G. J. Fonseca, B. Todorovic, P. Pelka, A. S. Turnell, L. Banks, and J. S. Mymryk. 2013. Dissection of the C-terminal region of E1A redefines the roles of CtBP and other cellular targets in oncogenic transformation. Journal of virology 87:10348-10355.

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

Crawford, D. H., K. F. Macsween, C. D. Higgins, R. Thomas, K. McAulay, H. Williams, N. Harrison, S. Reid, M. Conacher, J. Douglas, and A. J. Swerdlow. 2006. A cohort study among university students: identification of risk factors for Epstein-Barr virus seroconversion and infectious mononucleosis. Clin Infect Dis 43:276-282.

Creyghton, M. P., A. W. Cheng, G. G. Welstead, T. Kooistra, B. W. Carey, E. J. Steine, J. Hanna, M. A. Lodato, G. M. Frampton, P. A. Sharp, L. A. Boyer, R. A. Young, and R. Jaenisch. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America 107:21931-21936.

Cunningham, S. A., J. M. Rodriguez, M. P. Arrate, T. M. Tran, and T. A. Brock. 2002. JAM2 interacts with alpha4beta1. Facilitation by JAM3. The Journal of biological chemistry 277:27589-27592.

D'Souza, D. C., R. B. Gil, E. Zuzarte, L. M. MacDougall, L. Donahue, J. S. Ebersole, N. N. Boutros, T. Cooper, J. Seibyl, and J. H. Krystal. 2006. gamma-Aminobutyric acid-serotonin interactions in healthy men: implications for network models of psychosis and dissociation. Biol Psychiatry 59:128-137.

Dai, Y., Y. Tang, F. He, Y. Zhang, A. Cheng, R. Gan, and Y. Wu. 2012. Screening and functional analysis of differentially expressed genes in EBV-transformed lymphoblasts. Virology journal 9:77.

Dallas, M. H., B. Varnum-Finney, C. Delaney, K. Kato, and I. D. Bernstein. 2005. Density of the Notch ligand Delta1 determines generation of B and T cell precursors from hematopoietic stem cells. The Journal of experimental medicine 201:1361-1366.

Dang, C. V., A. O'Donnell K, and T. Juopperi. 2005. The great MYC escape in tumorigenesis. Cancer cell 8:177-178.

Davids, M. S., and J. A. Burger. 2012. Cell Trafficking in Chronic Lymphocytic Leukemia. Open J Hematol 3.

de la Serna, I. L., Y. Ohkawa, C. A. Berkes, D. A. Bergstrom, C. S. Dacwag, S. J. Tapscott, and A. N. Imbalzano. 2005. MyoD targets chromatin remodeling complexes to the myogenin locus prior to forming a stable DNA-bound complex. Molecular and cellular biology 25:3997-4009.

de Leval, L., and R. P. Hasserjian. 2009. Diffuse large B-cell lymphomas and burkitt lymphoma. Hematology/oncology clinics of North America 23:791-827.

Decker, T., M. Pasca di Magliano, S. McManus, Q. Sun, C. Bonifer, H. Tagoh, and M. Busslinger. 2009. Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. Immunity 30:508-520.

Defossez, P. A., K. F. Kelly, G. J. Filion, R. Perez-Torrado, F. Magdinier, H. Menoni, C. L. Nordgaard, J. M. Daniel, and E. Gilson. 2005. The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. The Journal of biological chemistry 280:43017-43023.

DeKoter, R. P., and H. Singh. 2000. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. Science 288:1439-1441.

Delecluse, H. J., R. Feederle, B. O'Sullivan, and P. Taniere. 2007. Epstein Barr virus-associated tumours: an update for the attention of the working pathologist. Journal of clinical pathology 60:1358-1364.

Delogu, A., A. Schebesta, Q. Sun, K. Aschenbrenner, T. Perlot, and M. Busslinger. 2006. Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. Immunity 24:269-281.

Denucci, C. C., J. S. Mitchell, and Y. Shimizu. 2009. Integrin function in T-cell homing to lymphoid and nonlymphoid sites: getting there and staying there. Crit Rev Immunol 29:87-109.

Di, L. J., J. S. Byun, M. M. Wong, C. Wakano, T. Taylor, S. Bilke, S. Baek, K. Hunter, H. Yang, M. Lee, C. Zvosec, G. Khramtsova, F. Cheng, C. M. Perou, C. R. Miller, R. Raab, O. I. Olopade, and K. Gardner. 2013. Genome-wide profiles of CtBP link metabolism with genome stability and epithelial reprogramming in breast cancer. Nat Commun 4:1449.

Dorsett, Y., D. F. Robbiani, M. Jankovic, B. Reina-San-Martin, T. R. Eisenreich, and M. C. Nussenzweig. 2007. A role for AID in chromosome translocations between c-myc and the IgH variable region. The Journal of experimental medicine 204:2225-2232.

Dresang, L. R., D. T. Vereide, and B. Sugden. 2009. Identifying sites bound by Epstein-Barr virus nuclear antigen 1 (EBNA1) in the human genome: defining a position-weighted matrix to predict sites bound by EBNA1 in viral genomes. Journal of virology 83:2930-2940.

Dukers, D. F., P. Meij, M. B. Vervoort, W. Vos, R. J. Scheper, C. J. Meijer, E. Bloemena, and J. M. Middeldorp. 2000. Direct immunosuppressive effects of EBV-encoded latent membrane protein 1. Journal of immunology 165:663-670.

Edmondson, D. G., M. M. Smith, and S. Y. Roth. 1996. Repression domain of the yeast global repressor Tup1 interacts directly with histones H3 and H4. Genes Dev 10:1247-1259.

Egle, A., A. W. Harris, P. Bouillet, and S. Cory. 2004. Bim is a suppressor of Myc-induced mouse B cell leukemia. Proceedings of the National Academy of Sciences of the United States of America 101:6164-6169.

Egloff, S., and S. Murphy. 2008. Cracking the RNA polymerase II CTD code. Trends Genet 24:280-288.

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

Enders, A., P. Bouillet, H. Puthalakath, Y. Xu, D. M. Tarlinton, and A. Strasser. 2003. Loss of the pro-apoptotic BH3-only Bcl-2 family member Bim inhibits BCR stimulation-induced apoptosis and deletion of autoreactive B cells. The Journal of experimental medicine 198:1119-1126.

Epstein, M. A., B. G. Achong, and Y. M. Barr. 1964. Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma. Lancet 1:702-703.

Ernst, T., A. J. Chase, J. Score, C. E. Hidalgo-Curtis, C. Bryant, A. V. Jones, K. Waghorn, K. Zoi, F. M. Ross, A. Reiter, A. Hochhaus, H. G. Drexler, A. Duncombe, F. Cervantes, D. Oscier, J. Boultwood, F. H. Grand, and N. C. Cross. 2010. Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nature genetics 42:722-726.

Erwin, G. D., N. Oksenberg, R. M. Truty, D. Kostka, K. K. Murphy, N. Ahituv, K. S. Pollard, and J. A. Capra. 2014. Integrating diverse datasets improves developmental enhancer prediction. PLoS computational biology 10:e1003677.

Farrell, C. M., A. G. West, and G. Felsenfeld. 2002. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. Molecular and cellular biology 22:3820-3831.

Featherstone, C., and P. Russell. 1991. Fission yeast p107wee1 mitotic inhibitor is a tyrosine/serine kinase. Nature 349:808-811.

Fejer, G., A. Koroknai, F. Banati, I. Gyory, D. Salamon, H. Wolf, H. H. Niller, and J. Minarovits. 2008. Latency type-specific distribution of epigenetic marks at the alternative promoters Cp and Qp of Epstein-Barr virus. J Gen Virol 89:1364-1370.

Filipovich, A. H., A. Mathur, D. Kamat, and R. S. Shapiro. 1992. Primary immunodeficiencies: genetic risk factors for lymphoma. Cancer research 52:5465s-5467s.

Fischer, S. F., P. Bouillet, K. O'Donnell, A. Light, D. M. Tarlinton, and A. Strasser. 2007. Proapoptotic BH3-only protein Bim is essential for developmentally programmed death of germinal center-derived memory B cells and antibody-forming cells. Blood 110:3978-3984.

Florquin, K., Y. Saeys, S. Degroeve, P. Rouze, and Y. Van de Peer. 2005. Large-scale structural analysis of the core promoter in mammalian and plant genomes. Nucleic acids research 33:4255-4264.

Francis, N. J., R. E. Kingston, and C. L. Woodcock. 2004. Chromatin compaction by a polycomb group protein complex. Science 306:1574-1577.

Fuentes-Panana, E. M., R. Peng, G. Brewer, J. Tan, and P. D. Ling. 2000. Regulation of the Epstein-Barr virus C promoter by AUF1 and the cyclic AMP/protein kinase A signaling pathway. Journal of virology 74:8166-8175.

Fujinaga, K., D. Irwin, Y. Huang, R. Taube, T. Kurosu, and B. M. Peterlin. 2004. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Molecular and cellular biology 24:787-795.

Fuxa, M., and M. Busslinger. 2007. Reporter gene insertions reveal a strictly B lymphoid-specific expression pattern of Pax5 in support of its B cell identity function. Journal of immunology 178:3031-3037.

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

Gao, H., K. Lukin, J. Ramirez, S. Fields, D. Lopez, and J. Hagman. 2009. Opposing effects of SWI/SNF and Mi-2/NuRD chromatin remodeling complexes on epigenetic reprogramming by EBF and Pax5. Proceedings of the National Academy of Sciences of the United States of America 106:11258-11263.

Garmy-Susini, B., C. J. Avraamides, J. S. Desgrosellier, M. C. Schmid, P. Foubert, L. G. Ellies, A. M. Lowy, S. L. Blair, S. R. Vandenberg, B. Datnow, H. Y. Wang, D. A. Cheresh, and J. Varner. 2013. PI3Kalpha activates integrin alpha4beta1 to establish a metastatic niche in lymph nodes. Proceedings of the National Academy of Sciences of the United States of America 110:9042-9047.

Georges, A. B., B. A. Benayoun, S. Caburet, and R. A. Veitia. 2010. Generic binding sites, generic DNA-binding domains: where does specific promoter recognition come from? FASEB journal : official publication of the Federation of American Societies for Experimental Biology 24:346-356.

Gershenzon, N. I., and I. P. Ioshikhes. 2005. Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. Bioinformatics 21:1295-1300.

Ghisletti, S., I. Barozzi, F. Mietton, S. Polletti, F. De Santa, E. Venturini, L. Gregory, L. Lonie, A. Chew, C. L. Wei, J. Ragoussis, and G. Natoli. 2010. Identification and characterization of

enhancers controlling the inflammatory gene expression program in macrophages. Immunity 32:317-328.

Gibb, D. R., M. El Shikh, D. J. Kang, W. J. Rowe, R. El Sayed, J. Cichy, H. Yagita, J. G. Tew, P. J. Dempsey, H. C. Crawford, and D. H. Conrad. 2010. ADAM10 is essential for Notch2-dependent marginal zone B cell development and CD23 cleavage in vivo. The Journal of experimental medicine 207:623-635.

Gilchrist, D. A., S. Nechaev, C. Lee, S. K. Ghosh, J. B. Collins, L. Li, D. S. Gilmour, and K. Adelman. 2008. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes & development 22:1921-1933.

Gold, M. R., L. Matsuuchi, R. B. Kelly, and A. L. DeFranco. 1991. Tyrosine phosphorylation of components of the B-cell antigen receptors following receptor crosslinking. Proceedings of the National Academy of Sciences of the United States of America 88:3436-3440.

Gombert, W. M., and A. Krumm. 2009. Targeted deletion of multiple CTCF-binding elements in the human C-MYC gene reveals a requirement for CTCF in C-MYC expression. PloS one 4:e6109.

Goodnow, C. C., C. G. Vinuesa, K.L. Randall, F. Mackay, R. Brink. 2010. Control systems and decision making for antibody production. Nature Immunology 11:681–688.

Gordon, W. R., K. L. Arnett, and S. C. Blacklow. 2008. The molecular logic of Notch signaling--a structural and biochemical perspective. Journal of cell science 121:3109-3119.

Green, M. R. 2005. Eukaryotic transcription activation: right on target. Molecular cell 18:399-402.

Greenwald, I. 2012. Notch and the awesome power of genetics. Genetics 191:655-669.

Gregory, C. D., M. Rowe, and A. B. Rickinson. 1990. Different Epstein-Barr virus-B cell interactions in phenotypically distinct clones of a Burkitt's lymphoma cell line. The Journal of general virology 71 (Pt 7):1481-1495.

Groom, J. R., and A. D. Luster. 2011. CXCR3 ligands: redundant, collaborative and antagonistic functions. Immunol Cell Biol 89:207-215.

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

Haan, K. M., S. K. Lee, and R. Longnecker. 2001. Different functional domains in the cytoplasmic tail of glycoprotein B are involved in Epstein-Barr virus-induced membrane fusion. Virology 290:106-114.

Hadinoto, V., M. Shapiro, C. C. Sun, and D. A. Thorley-Lawson. 2009. The dynamics of EBV shedding implicate a central role for epithelial cells in amplifying viral output. PLoS Pathog 5:e1000496.

Hagman, J., M. J. Gutch, H. Lin, and R. Grosschedl. 1995. EBF contains a novel zinc coordination motif and multiple dimerization and transcriptional activation domains. The EMBO journal 14:2907-2916.

Hagman, J., J. Ramirez, and K. Lukin. 2012. B lymphocyte lineage specification, commitment and epigenetic control of transcription by early B cell factor 1. Current topics in microbiology and immunology 356:17-38.

Hagman, J., A. Travis, and R. Grosschedl. 1991. A novel lineage-specific nuclear factor regulates mb-1 gene transcription at the early stages of B cell differentiation. The EMBO journal 10:3409-3417.

Halder, S., M. Murakami, S. C. Verma, P. Kumar, F. Yi, and E. S. Robertson. 2009. Early events associated with infection of Epstein-Barr virus infection of primary B-cells. PLoS One 4:e7214.

Han, H., K. Tanigaki, N. Yamamoto, K. Kuroda, M. Yoshimoto, T. Nakahata, K. Ikuta, and T. Honjo. 2002. Inducible gene knockout of transcription factor recombination signal binding protein-J reveals its essential role in T versus B lineage decision. International immunology 14:637-645.

Handoko, L., H. Xu, G. Li, C. Y. Ngan, E. Chew, M. Schnapp, C. W. Lee, C. Ye, J. L. Ping, F. Mulawadi, E. Wong, J. Sheng, Y. Zhang, T. Poh, C. S. Chan, G. Kunarso, A. Shahab, G. Bourque, V. Cacheux-Rataboul, W. K. Sung, Y. Ruan, and C. L. Wei. 2011. CTCF-mediated functional chromatin interactome in pluripotent cells. Nature genetics 43:630-638.

Hansen, K. H., A. P. Bracken, D. Pasini, N. Dietrich, S. S. Gehani, A. Monrad, J. Rappsilber, M. Lerdrup, and K. Helin. 2008. A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol 10:1291-1300.

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

Hardy, R. R., P. W. Kincade, and K. Dorshkind. 2007. The protean nature of cells in the B lymphocyte lineage. Immunity 26:703-714.

Hark, A. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. Levorse, and S. M. Tilghman. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature 405:486-489.

Harlan, J. M. 1985. Leukocyte-endothelial interactions. Blood 65:513-525.

Harp, J. M., B. L. Hanson, D. E. Timm, and G. J. Bunick. 2000. Asymmetries in the nucleosome core particle at 2.5 A resolution. Acta Crystallogr D Biol Crystallogr 56:1513-1534.

Harris, N. L., E. S. Jaffe, J. Diebold, G. Flandrin, H. K. Muller-Hermelink, J. Vardiman, T. A. Lister, and C. D. Bloomfield. 1999. The World Health Organization classification of neoplastic diseases of the hematopoietic and lymphoid tissues. Report of the Clinical Advisory Committee meeting, Airlie House, Virginia, November, 1997. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 10:1419-1432.

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

He, A., X. Shen, Q. Ma, J. Cao, A. von Gise, P. Zhou, G. Wang, V. E. Marquez, S. H. Orkin, and W. T. Pu. 2012. PRC2 directly methylates GATA4 and represses its transcriptional activity. Genes Dev 26:37-42.

Heinz, S., C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, and C. K. Glass. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell 38:576-589.

Henikoff, S., and A. Shilatifard. 2011. Histone modification: cause or cog? Trends Genet 27:389-396.

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

Henle, G., W. Henle, P. Clifford, V. Diehl, G. W. Kafuko, B. G. Kirya, G. Klein, R. H. Morrow, G. M. Munube, P. Pike, P. M. Tukei, and J. L. Ziegler. 1969. Antibodies to Epstein-Barr virus in Burkitt's lymphoma and control groups. J Natl Cancer Inst 43:1147-1157.

Hensbergen, P. J., P. G. Wijnands, M. W. Schreurs, R. J. Scheper, R. Willemze, and C. P. Tensen. 2005. The CXCR3 targeting chemokine CXCL11 has potent antitumor activity in vivo involving attraction of CD8+ T lymphocytes but not inhibition of angiogenesis. J Immunother 28:343-351.

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

Herz, H. M., and A. Shilatifard. 2010. The JARID2-PRC2 duality. Genes Dev 24:857-861.

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

Hildebrand, J. D., and P. Soriano. 2002. Overlapping and unique roles for C-terminal binding protein 1 (CtBP1) and CtBP2 during mouse development. Molecular and cellular biology 22:5296-5307.

Hislop, A. D., M. Kuo, A. B. Drake-Lee, A. N. Akbar, W. Bergler, N. Hammerschmitt, N. Khan, U. Palendira, A. M. Leese, J. M. Timms, A. I. Bell, C. D. Buckley, and A. B. Rickinson. 2005. Tonsillar homing of Epstein-Barr virus-specific CD8+ T cells and the virus-host balance. The Journal of clinical investigation 115:2546-2555.

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

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

Hofelmayr, H., L. J. Strobl, C. Stein, G. Laux, G. Marschall, G. W. Bornkamm, and U. Zimber-Strobl. 1999. Activated mouse Notch1 transactivates Epstein-Barr virus nuclear antigen 2-regulated viral promoters. Journal of virology 73:2770-2780.

Holwerda, S. J., and W. de Laat. 2013. CTCF: the protein, the binding partners, the binding sites and their chromatin loops. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 368:20120369.

Hombach, J., F. Lottspeich, and M. Reth. 1990. Identification of the genes encoding the IgM-alpha and Ig-beta components of the IgM antigen receptor complex by amino-terminal sequencing. European journal of immunology 20:2795-2799.

Hsieh, J. J., and S. D. Hayward. 1995. Masking of the CBF1/RBPJ kappa transcriptional repression domain by Epstein-Barr virus EBNA2. Science 268:560-563.

Huck, K., O. Feyen, T. Niehues, F. Ruschendorf, N. Hubner, H. J. Laws, T. Telieps, S. Knapp, H. H. Wacker, A. Meindl, H. Jumaa, and A. Borkhardt. 2009. Girls homozygous for an IL-2-inducible T cell kinase mutation that leads to protein deficiency develop fatal EBV-associated lymphoproliferation. The Journal of clinical investigation 119:1350-1358.

Hunkapiller, J., Y. Shen, A. Diaz, G. Cagney, D. McCleary, M. Ramalho-Santos, N. Krogan, B. Ren, J. S. Song, and J. F. Reiter. 2012. Polycomb-like 3 promotes polycomb repressive complex 2 binding to CpG islands and embryonic stem cell self-renewal. PLoS genetics 8:e1002576.

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

Hutchings, I. A., R. J. Tierney, G. L. Kelly, J. Stylianou, A. B. Rickinson, and A. I. Bell. 2006. Methylation status of the Epstein-Barr virus (EBV) BamHI W latent cycle promoter and promoter activity: analysis with novel EBV-positive Burkitt and lymphoblastoid cell lines. J Virol 80:10700-10711.

Hutt-Fletcher, L. M. 2007. Epstein-Barr virus entry. J Virol 81:7825-7832.

Iwasaki, H., C. Somoza, H. Shigematsu, E. A. Duprez, J. Iwasaki-Arai, S. Mizuno, Y. Arinobu, K. Geary, P. Zhang, T. Dayaram, M. L. Fenyus, S. Elf, S. Chan, P. Kastner, C. S. Huettner, R. Murray, D. G. Tenen, and K. Akashi. 2005. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. Blood 106:1590-1600.

Jackson, K. J., M. J. Kidd, Y. Wang, and A. M. Collins. 2013. The shape of the lymphocyte receptor repertoire: lessons from the B cell receptor. Frontiers in immunology 4:263.

Jacobson, R. H., A. G. Ladurner, D. S. King, and R. Tjian. 2000. Structure and function of a human TAFII250 double bromodomain module. Science 288:1422-1425.

Jaffe, E. S., J. K. Chan, I. J. Su, G. Frizzera, S. Mori, A. C. Feller, and F. C. Ho. 1996. Report of the Workshop on Nasal and Related Extranodal Angiocentric T/Natural Killer Cell Lymphomas. Definitions, differential diagnosis, and epidemiology. The American journal of surgical pathology 20:103-111.

Jang, M. K., K. Mochizuki, M. Zhou, H. S. Jeong, J. N. Brady, and K. Ozato. 2005. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Molecular cell 19:523-534.

Jang, Y.S., W. Kang, D. Y. Chang, P. S. Sung, B. C. Park, S. H. Yoo, Y.W. Park, E. C. Shin. 2013. CD27 engagement by a soluble CD70 protein enhances non-cytolytic antiviral activity of CD56bright natural killer cells by IFN-γ secretion. Clinical Immunology 149(3):379-87.

Ji, M., H. Li, H. C. Suh, K. D. Klarmann, Y. Yokota, and J. R. Keller. 2008. Id2 intrinsically regulates lymphoid and erythroid development via interaction with different target proteins. Blood 112:1068-1077.

Jiang, R., R. S. Scott, and L. M. Hutt-Fletcher. 2006. Epstein-Barr virus shed in saliva is high in B-cell-tropic glycoprotein gp42. J Virol 80:7281-7283.

Jimenez-Ramirez, C., A. J. Brooks, L. P. Forshell, K. Yakimchuk, B. Zhao, T. Z. Fulgham, and C. E. Sample. 2006. Epstein-Barr virus EBNA-3C is targeted to and regulates expression from the bidirectional LMP-1/2B promoter. Journal of virology 80:11200-11208.

Jochum, S., A. Moosmann, S. Lang, W. Hammerschmidt, and R. Zeidler. 2012a. The EBV Immunoevasins vIL-10 and BNLF2a Protect Newly Infected B Cells from Immune Recognition and Elimination. PLoS Pathog 8:e1002704.

Jochum, S., R. Ruiss, A. Moosmann, W. Hammerschmidt, and R. Zeidler. 2012b. RNAs in Epstein-Barr virions control early steps of infection. Proc Natl Acad Sci U S A 109:E1396-1404.

Johannsen, E., E. Koh, G. Mosialos, X. Tong, E. Kieff, and S. R. Grossman. 1995. Epstein-Barr virus nuclear protein 2 transactivation of the latent membrane protein 1 promoter is mediated by J kappa and PU.1. Journal of virology 69:253-262.

Jones, P. L., G. J. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis, and A. P. Wolffe. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature genetics 19:187-191.

Jones, S., X. Zhang, D. W. Parsons, J. C. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, H. Kamiyama, A. Jimeno, S. M. Hong, B. Fu, M. T. Lin, E. S. Calhoun, M. Kamiyama, K. Walter, T. Nikolskaya, Y. Nikolsky, J. Hartigan, D. R. Smith, M. Hidalgo, S. D. Leach, A. P. Klein, E. M. Jaffee, M. Goggins, A. Maitra, C. Iacobuzio-Donahue, J. R. Eshleman, S. E. Kern, R. H. Hruban, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V. E. Velculescu, and K. W. Kinzler. 2008. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science 321:1801-1806.

Kaiser, C., G. Laux, D. Eick, N. Jochner, G. W. Bornkamm, and B. Kempkes. 1999. The proto-oncogene c-myc is a direct target gene of Epstein-Barr virus nuclear antigen 2. Journal of virology 73:4481-4484.

Kalla, M., C. Gobel, and W. Hammerschmidt. 2012. The lytic phase of epstein-barr virus requires a viral genome with 5-methylcytosine residues in CpG sites. J Virol 86:447-458.

Kalla, M., A. Schmeinck, M. Bergbauer, D. Pich, and W. Hammerschmidt. 2010. AP-1 homolog BZLF1 of Epstein-Barr virus has two essential functions dependent on the epigenetic state of the viral genome. Proc Natl Acad Sci U S A 107:850-855.

Kapatai, G., and P. Murray. 2007. Contribution of the Epstein Barr virus to the molecular pathogenesis of Hodgkin lymphoma. Journal of clinical pathology 60:1342-1349.

Katsanis, N., and E. M. Fisher. 1998. A novel C-terminal binding protein (CTBP2) is closely related to CTBP1, an adenovirus E1A-binding protein, and maps to human chromosome 21q21.3. Genomics 47:294-299.

Kee, B. L., and C. Murre. 1998. Induction of early B cell factor (EBF) and multiple B lineage genes by the basic helix-loop-helix transcription factor E12. The Journal of experimental medicine 188:699-713.

Kellogg, D. R. 2003. Wee1-dependent mechanisms required for coordination of cell growth and cell division. Journal of cell science 116:4883-4890.

Kelly, G. L., A. E. Milner, R. J. Tierney, D. S. Croom-Carter, M. Altmann, W. Hammerschmidt, A. I. Bell, and A. B. Rickinson. 2005. Epstein-Barr virus nuclear antigen 2 (EBNA2) gene deletion is consistently linked with EBNA3A, -3B, and -3C expression in Burkitt's lymphoma cells and with increased resistance to apoptosis. Journal of virology 79:10709-10717.

Kempkes, B., D. Spitkovsky, P. Jansen-Durr, J. W. Ellwart, E. Kremmer, H. J. Delecluse, C. Rottenberger, G. W. Bornkamm, and W. Hammerschmidt. 1995. B-cell proliferation and induction of early G1-regulating proteins by Epstein-Barr virus mutants conditional for EBNA2. The EMBO journal 14:88-96.

Kerr, B. M., A. L. Lear, M. Rowe, D. Croom-Carter, L. S. Young, S. M. Rookes, P. H. Gallimore, and A. B. Rickinson. 1992. Three transcriptionally distinct forms of Epstein-Barr virus latency in somatic cell hybrids: cell phenotype dependence of virus promoter usage. Virology 187:189-201.

Ketel, C. S., E. F. Andersen, M. L. Vargas, J. Suh, S. Strome, and J. A. Simon. 2005. Subunit contributions to histone methyltransferase activities of fly and worm polycomb group complexes. Molecular and cellular biology 25:6857-6868.

Kim do, N., Y. J. Song, and S. K. Lee. 2011. The role of promoter methylation in Epstein-Barr virus (EBV) microRNA expression in EBV-infected B cell lines. Exp Mol Med 43:401-410.

Kim, J. L., D. B. Nikolov, and S. K. Burley. 1993. Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature 365:520-527.

Kim, T. H., Z. K. Abdullaev, A. D. Smith, K. A. Ching, D. I. Loukinov, R. D. Green, M. Q. Zhang, V. V. Lobanenkov, and B. Ren. 2007. Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128:1231-1245.

Kirby, H., A. Rickinson, and A. Bell. 2000. The activity of the Epstein-Barr virus BamHI W promoter in B cells is dependent on the binding of CREB/ATF factors. J Gen Virol 81:1057-1066.

Kitagawa, N., M. Goto, K. Kurozumi, S. Maruo, M. Fukayama, T. Naoe, M. Yasukawa, K. Hino, T. Suzuki, S. Todo, and K. Takada. 2000. Epstein-Barr virus-encoded poly(A)(-) RNA supports Burkitt's lymphoma growth through interleukin-10 induction. The EMBO journal 19:6742-6750.

Klein, U., Y. Tu, G. A. Stolovitzky, J. L. Keller , J. Jr Haddad , V. Miljkovic , G. Cattoretti , A. Califano , R. Dalla-Favera. 2003. Gene expression dynamics during germinal center transit in B cells. Annals of the New York Academy of Sciences 987:166-72.

Klemsz, M. J., S. R. McKercher, A. Celada, C. Van Beveren, and R. A. Maki. 1990. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. Cell 61:113-124.

Klenova, E. M., R. H. Nicolas, H. F. Paterson, A. F. Carne, C. M. Heath, G. H. Goodwin, P. E. Neiman, and V. V. Lobanenkov. 1993. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. Molecular and cellular biology 13:7612-7624.

Knight, J. S., K. Lan, C. Subramanian, and E. S. Robertson. 2003. Epstein-Barr virus nuclear antigen 3C recruits histone deacetylase activity and associates with the corepressors mSin3A and NCoR in human B-cell lines. J Virol 77:4261-4272.

Knutson, J. C. 1990. The level of c-fgr RNA is increased by EBNA-2, an Epstein-Barr virus gene required for B-cell immortalization. Journal of virology 64:2530-2536.

Kopan, R., and M. X. Ilagan. 2009. The canonical Notch signaling pathway: unfolding the activation mechanism. Cell 137:216-233.

Koshland, D., and A. Strunnikov. 1996. Mitotic chromosome condensation. Annu Rev Cell Dev Biol 12:305-333.

Kovi, R. C., S. Paliwal, S. Pande, and S. R. Grossman. 2010. An ARF/CtBP2 complex regulates BH3-only gene expression and p53-independent apoptosis. Cell Death Differ 17:513-521.

Krauer, K. G., M. Buck, D. K. Belzer, J. Flanagan, G. M. Chojnowski, and T. B. Sculley. 2004a. The Epstein-Barr virus nuclear antigen-6 protein co-localizes with EBNA-3 and survival of motor neurons protein. Virology 318:280-294.

Krauer, K. G., A. Burgess, M. Buck, J. Flanagan, T. B. Sculley, and B. Gabrielli. 2004b. The EBNA-3 gene family proteins disrupt the G2/M checkpoint. Oncogene 23:1342-1353.

Kumano, K., S. Chiba, A. Kunisato, M. Sata, T. Saito, E. Nakagami-Yamaguchi, T. Yamaguchi, S. Masuda, K. Shimizu, T. Takahashi, S. Ogawa, Y. Hamada, and H. Hirai. 2003. Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. Immunity 18:699-711.

Kuo, M. H., J. E. Brownell, R. E. Sobel, T. A. Ranalli, R. G. Cook, D. G. Edmondson, S. Y. Roth, and C. D. Allis. 1996. Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature 383:269-272.

Kuppuswamy, M., S. Vijayalingam, L. J. Zhao, Y. Zhou, T. Subramanian, J. Ryerse, and G. Chinnadurai. 2008. Role of the PLDLS-binding cleft region of CtBP1 in recruitment of core and auxiliary components of the corepressor complex. Molecular and cellular biology 28:269-281.

Kurdistani, S. K., and M. Grunstein. 2003. Histone acetylation and deacetylation in yeast. Nature reviews. Molecular cell biology 4:276-284.

Kutok, J. L., and F. Wang. 2006. Spectrum of Epstein-Barr virus-associated diseases. Annu Rev Pathol 1:375-404.

Kuzmichev, A., T. Jenuwein, P. Tempst, and D. Reinberg. 2004. Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Molecular cell 14:183-193.

Laichalk, L. L., and D. A. Thorley-Lawson. 2005. Terminal differentiation into plasma cells initiates the replicative cycle of Epstein-Barr virus in vivo. J Virol 79:1296-1307.

Landais, E., X. Saulquin, and E. Houssaint. 2005. The human T cell immune response to Epstein-Barr virus. Int J Dev Biol 49:285-292.

Laudencia-Chingcuanco, D., S. Ganeshan, F. You, B. Fowler, R. Chibbar, and O. Anderson. 2011. Genome-wide gene expression analysis supports a developmental model of low temperature tolerance gene regulation in wheat (Triticum aestivum L.). BMC Genomics 12:299.

Laux, G., F. Dugrillon, C. Eckert, B. Adam, U. Zimber-Strobl, and G. W. Bornkamm. 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.

Le Roux, A., B. Kerdiles, D. Walls, J. F. Dedieu, and M. Perricaudet. 1994. The Epstein-Barr virus determined nuclear antigens EBNA-3A, -3B, and -3C repress EBNA-2-mediated transactivation of the viral terminal protein 1 gene promoter. Virology 205:596-602.

Lee, S. P., J. M. Brooks, H. Al-Jarrah, W. A. Thomas, T. A. Haigh, G. S. Taylor, S. Humme, A. Schepers, W. Hammerschmidt, J. L. Yates, A. B. Rickinson, and N. W. Blake. 2004. CD8 T cell recognition of endogenously expressed epstein-barr virus nuclear antigen 1. The Journal of experimental medicine 199:1409-1420.

Lee, S. P., R. J. Tierney, W. A. Thomas, J. M. Brooks, and A. B. Rickinson. 1997. Conserved CTL epitopes within EBV latent membrane protein 2: a potential target for CTL-based tumor therapy. Journal of immunology 158:3325-3334.

Lee, T. I., and R. A. Young. 2000. Transcription of eukaryotic protein-coding genes. Annual review of genetics 34:77-137.

Leen, A., P. Meij, I. Redchenko, J. Middeldorp, E. Bloemena, A. Rickinson, and N. Blake. 2001. Differential immunogenicity of Epstein-Barr virus latent-cycle proteins for human CD4(+) T-helper 1 responses. Journal of virology 75:8649-8659.

Lefstin, J. A., and K. R. Yamamoto. 1998. Allosteric effects of DNA on transcriptional regulators. Nature 392:885-888.

Leight, E. R., and B. Sugden. 2000. EBNA-1: a protein pivotal to latent infection by Epstein-Barr virus. Rev Med Virol 10:83-100.

Lhoumaud, P., M. Hennion, A. Gamot, S. Cuddapah, S. Queille, J. Liang, G. Micas, P. Morillon, S. Urbach, O. Bouchez, D. Severac, E. Emberly, K. Zhao, and O. Cuvier. 2014. Insulators recruit histone methyltransferase dMes4 to regulate chromatin of flanking genes. The EMBO journal 33:1599-1613.

Li, Q., K. R. Peterson, X. Fang, and G. Stamatoyannopoulos. 2002. Locus control regions. Blood 100:3077-3086.

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

Li, Q., S. M. Turk, and L. M. Hutt-Fletcher. 1995. The Epstein-Barr virus (EBV) BZLF2 gene product associates with the gH and gL homologs of EBV and carries an epitope critical to infection of B cells but not of epithelial cells. Journal of virology 69:3987-3994.

Li, R., H. Pei, and D. K. Watson. 2000. Regulation of Ets function by protein - protein interactions. Oncogene 19:6514-6523.

Lieberman-Aiden, E., N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326:289-293.

Lin, H., and R. Grosschedl. 1995. Failure of B-cell differentiation in mice lacking the transcription factor EBF. Nature 376:263-267.

Lin, J., E. Johannsen, E. Robertson, and E. Kieff. 2002. Epstein-Barr virus nuclear antigen 3C putative repression domain mediates coactivation of the LMP1 promoter with EBNA-2. J Virol 76:232-242.

Lin, Y. C., S. Jhunjhunwala, C. Benner, S. Heinz, E. Welinder, R. Mansson, M. Sigvardsson, J. Hagman, C. A. Espinoza, J. Dutkowski, T. Ideker, C. K. Glass, and C. Murre. 2010. A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. Nature immunology 11:635-643.

Lindhout, E., M. L. Mevissen, J. Kwekkeboom, J. M. Tager, and C. de Groot. 1993. Direct evidence that human follicular dendritic cells (FDC) rescue germinal centre B cells from death by apoptosis. Clinical and experimental immunology 91:330-336.

Ling, P. D., D. R. Rawlins, and S. D. Hayward. 1993. The Epstein-Barr virus immortalizing protein EBNA-2 is targeted to DNA by a cellular enhancer-binding protein. Proceedings of the National Academy of Sciences of the United States of America 90:9237-9241.

Linka, R. M., S. L. Risse, K. Bienemann, M. Werner, Y. Linka, F. Krux, C. Synaeve, R. Deenen, S. Ginzel, R. Dvorsky, M. Gombert, A. Halenius, R. Hartig, M. Helminen, A. Fischer, P. Stepensky, K. Vettenranta, K. Kohrer, M. R. Ahmadian, H. J. Laws, B. Fleckenstein, H. Jumaa, S. Latour, B. Schraven, and A. Borkhardt. 2012. Loss-of-function mutations

within the IL-2 inducible kinase ITK in patients with EBV-associated lymphoproliferative diseases. Leukemia 26:963-971.

Liu, Z., D. R. Scannell, M. B. Eisen, and R. Tjian. 2011. Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. Cell 146:720-731.

Lobanenkov, V. V., R. H. Nicolas, V. V. Adler, H. Paterson, E. M. Klenova, A. V. Polotskaja, and G. H. Goodwin. 1990. A novel sequence-specific DNA binding protein which interacts with three regularly spaced direct repeats of the CCCTC-motif in the 5'-flanking sequence of the chicken c-myc gene. Oncogene 5:1743-1753.

Lomvardas, S., G. Barnea, D. J. Pisapia, M. Mendelsohn, J. Kirkland, and R. Axel. 2006. Interchromosomal interactions and olfactory receptor choice. Cell 126:403-413.

Lonard, D. M., and B. W. O'Malley. 2005. Expanding functional diversity of the coactivators. Trends in biochemical sciences 30:126-132.

Lotz, M., C. D. Tsoukas, S. Fong, D. A. Carson, and J. H. Vaughan. 1985. Regulation of Epstein-Barr virus infection by recombinant interferons. Selected sensitivity to interferon-gamma. Eur J Immunol 15:520-525.

Lu, F., P. Wikramasinghe, J. Norseen, K. Tsai, P. Wang, L. Showe, R. V. Davuluri, and P. M. Lieberman. 2010. Genome-wide analysis of host-chromosome binding sites for Epstein-Barr Virus Nuclear Antigen 1 (EBNA1). Virology journal 7:262.

Lucchesi, W., G. Brady, O. Dittrich-Breiholz, M. Kracht, R. Russ, and P. J. Farrell. 2008. Differential gene regulation by Epstein-Barr virus type 1 and type 2 EBNA2. J Virol 82:7456-7466.

Luger, K., A. W. Mader, R. K. Richmond, D. F. Sargent, and T. J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature 389:251-260.

Luger, K., and T. J. Richmond. 1998. The histone tails of the nucleosome. Curr Opin Genet Dev 8:140-146.

Lund, J., O. H. Olsen, E. S. Sorensen, H. R. Stennicke, H. H. Petersen, and M. T. Overgaard. 2013. ADAMDEC1 is a metzincin metalloprotease with dampened proteolytic activity. The Journal of biological chemistry 288:21367-21375.

MacLennan, I. C. 1994. Germinal centers. Annual review of immunology 12:117-139.

Maeda, T., T. Merghoub, R. M. Hobbs, L. Dong, M. Maeda, J. Zakrzewski, M. R. van den Brink, A. Zelent, H. Shigematsu, K. Akashi, J. Teruya-Feldstein, G. Cattoretti, and P. P. Pandolfi. 2007. Regulation of B versus T lymphoid lineage fate decision by the proto-oncogene LRF. Science 316:860-866.

Mahadevan, L. C., A. C. Willis, and M. J. Barratt. 1991. Rapid histone H3 phosphorylation in response to growth factors, phorbol esters, okadaic acid, and protein synthesis inhibitors. Cell 65:775-783.

Maier, H., R. Ostraat, H. Gao, S. Fields, S. A. Shinton, K. L. Medina, T. Ikawa, C. Murre, H. Singh, R. R. Hardy, and J. Hagman. 2004. Early B cell factor cooperates with Runx1 and mediates epigenetic changes associated with mb-1 transcription. Nature immunology 5:1069-1077.

Maier, S., G. Staffler, A. Hartmann, J. Hock, K. Henning, K. Grabusic, R. Mailhammer, R. Hoffmann, M. Wilmanns, R. Lang, J. Mages, and B. Kempkes. 2006. Cellular target genes of Epstein-Barr virus nuclear antigen 2. J Virol 80:9761-9771.

Maini, M. K., N. Gudgeon, L. R. Wedderburn, A. B. Rickinson, and P. C. Beverley. 2000. Clonal expansions in acute EBV infection are detectable in the CD8 and not the CD4 subset and persist with a variable CD45 phenotype. Journal of immunology 165:5729-5737.

Majumder, P., and J. M. Boss. 2011. Cohesin regulates MHC class II genes through interactions with MHC class II insulators. Journal of immunology 187:4236-4244.

Mannick, J. B., J. I. Cohen, M. Birkenbach, A. Marchini, and E. Kieff. 1991. The Epstein-Barr virus nuclear protein encoded by the leader of the EBNA RNAs is important in B-lymphocyte transformation. Journal of virology 65:6826-6837.

Mansson, R., S. Zandi, E. Welinder, P. Tsapogas, N. Sakaguchi, D. Bryder, and M. Sigvardsson. 2010. Single-cell analysis of the common lymphoid progenitor compartment reveals functional and molecular heterogeneity. Blood 115:2601-2609.

Margueron, R., N. Justin, K. Ohno, M. L. Sharpe, J. Son, W. J. Drury, 3rd, P. Voigt, S. R. Martin, W. R. Taylor, V. De Marco, V. Pirrotta, D. Reinberg, and S. J. Gamblin. 2009. Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461:762-767.

Marmorstein, R. 2001. Structure of histone deacetylases: insights into substrate recognition and catalysis. Structure 9:1127-1133.

Marquitz, A. R., and N. Raab-Traub. 2012. The role of miRNAs and EBV BARTs in NPC. Seminars in cancer biology 22:166-172.

Marshall, D., and C. Sample. 1995. Epstein-Barr virus nuclear antigen 3C is a transcriptional regulator. Journal of virology 69:3624-3630.

Marshall, N. F., and D. H. Price. 1995. Purification of P-TEFb, a transcription factor required for the transition into productive elongation. The Journal of biological chemistry 270:12335-12338.

Maston, G. A., S. K. Evans, and M. R. Green. 2006. Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet 7:29-59.

Mathas, S., M. Janz, F. Hummel, M. Hummel, B. Wollert-Wulf, S. Lusatis, I. Anagnostopoulos, A. Lietz, M. Sigvardsson, F. Jundt, K. Johrens, K. Bommert, H. Stein, and B. Dorken. 2006. Intrinsic inhibition of transcription factor E2A by HLH proteins ABF-1 and Id2 mediates reprogramming of neoplastic B cells in Hodgkin lymphoma. Nat Immunol 7:207-215.

Maunders, M. J., L. Petti, and M. Rowe. 1994. Precipitation of the Epstein-Barr virus protein EBNA 2 by an EBNA 3c-specific monoclonal antibody. The Journal of general virology 75 (Pt 4):769-778.

Mazzone, A., and G. Ricevuti. 1995. Leukocyte CD11/CD18 integrins: biological and clinical relevance. Haematologica 80:161-175.

McCabe, M. T., A. P. Graves, G. Ganji, E. Diaz, W. S. Halsey, Y. Jiang, K. N. Smitheman, H. M. Ott, M. B. Pappalardi, K. E. Allen, S. B. Chen, A. Della Pietra, 3rd, E. Dul, A. M. Hughes, S. A. Gilbert, S. H. Thrall, P. J. Tummino, R. G. Kruger, M. Brandt, B. Schwartz, and C. L. Creasy. 2012. Mutation of A677 in histone methyltransferase EZH2 in human B-cell lymphoma promotes hypertrimethylation of histone H3 on lysine 27 (H3K27). Proceedings of the National Academy of Sciences of the United States of America 109:2989-2994.

McClellan, M. J., S. Khasnis, C. D. Wood, R. D. Palermo, S. N. Schlick, A. S. Kanhere, R. G. Jenner, and M. J. West. 2012. Downregulation of integrin receptor-signaling genes by Epstein-Barr virus EBNA 3C via promoter-proximal and -distal binding elements. Journal of virology 86:5165-5178.

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

McGowan, C. H., and P. Russell. 1995. Cell cycle regulation of human WEE1. The EMBO journal 14:2166-2175.

McKercher, S. R., B. E. Torbett, K. L. Anderson, G. W. Henkel, D. J. Vestal, H. Baribault, M. Klemsz, A. J. Feeney, G. E. Wu, C. J. Paige, and R. A. Maki. 1996. Targeted disruption of the PU.1 gene results in multiple hematopoietic abnormalities. The EMBO journal 15:5647-5658.

McManus, S., A. Ebert, G. Salvagiotto, J. Medvedovic, Q. Sun, I. Tamir, M. Jaritz, H. Tagoh, and M. Busslinger. 2011. The transcription factor Pax5 regulates its target genes by

recruiting chromatin-modifying proteins in committed B cells. The EMBO journal 30:2388-2404.

Medina, K. L., J. M. Pongubala, K. L. Reddy, D. W. Lancki, R. Dekoter, M. Kieslinger, R. Grosschedl, and H. Singh. 2004. Assembling a gene regulatory network for specification of the B cell fate. Developmental cell 7:607-617.

Mendenhall, E. M., R. P. Koche, T. Truong, V. W. Zhou, B. Issac, A. S. Chi, M. Ku, and B. E. Bernstein. 2010. GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS genetics 6:e1001244.

Mercer, E. M., Y. C. Lin, C. Benner, S. Jhunjhunwala, J. Dutkowski, M. Flores, M. Sigvardsson, T. Ideker, C. K. Glass, and C. Murre. 2011. Multilineage priming of enhancer repertoires precedes commitment to the B and myeloid cell lineages in hematopoietic progenitors. Immunity 35:413-425.

Middeldorp, J. M., and D. M. Pegtel. 2008. Multiple roles of LMP1 in Epstein-Barr virus induced immune escape. Semin Cancer Biol 18:388-396.

Miller, C. L., A. L. Burkhardt, J. H. Lee, B. Stealey, R. Longnecker, J. B. Bolen, and E. Kieff. 1995. Integral membrane protein 2 of Epstein-Barr virus regulates reactivation from latency through dominant negative effects on protein-tyrosine kinases. Immunity 2:155-166.

Misteli, T., A. Gunjan, R. Hock, M. Bustin, and D. T. Brown. 2000. Dynamic binding of histone H1 to chromatin in living cells. Nature 408:877-881.

Morris, J. R., D. A. Petrov, A. M. Lee, and C. T. Wu. 2004. Enhancer choice in cis and in trans in Drosophila melanogaster: role of the promoter. Genetics 167:1739-1747.

Mullighan, C. G., S. Goorha, I. Radtke, C. B. Miller, E. Coustan-Smith, J. D. Dalton, K. Girtman, S. Mathew, J. Ma, S. B. Pounds, X. Su, C. H. Pui, M. V. Relling, W. E. Evans, S. A. Shurtleff, and J. R. Downing. 2007. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. Nature 446:758-764.

Murat, P., J. Zhong, L. Lekieffre, N. P. Cowieson, J. L. Clancy, T. Preiss, S. Balasubramanian, R. Khanna, and J. Tellam. 2014. G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. Nature chemical biology 10:358-364.

Murata, T., Y. Kondo, A. Sugimoto, D. Kawashima, S. Saito, H. Isomura, T. Kanda, and T. Tsurumi. 2012. Epigenetic histone modification of Epstein-Barr virus BZLF1 promoter during latency and reactivation in Raji cells. J Virol 86:4752-4761.

Murayama, A., M. S. Kim, J. Yanagisawa, K. Takeyama, and S. Kato. 2004. Transrepression by a liganded nuclear receptor via a bHLH activator through co-regulator switching. The EMBO journal 23:1598-1608.

Murray, P. G., and L. S. Young. 2001. Epstein-Barr virus infection: basis of malignancy and potential for therapy. Expert Rev Mol Med 3:1-20.

Muylaert, I., and P. Elias. 2010. Contributions of nucleotide excision repair, DNA polymerase eta, and homologous recombination to replication of UV-irradiated herpes simplex virus type 1. J Biol Chem 285:13761-13768.

Neri, F., D. Incarnato, A. Krepelova, S. Rapelli, A. Pagnani, R. Zecchina, C. Parlato, and S. Oliviero. 2013. Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. Genome Biol 14:R91.

Neve, R. M., K. Chin, J. Fridlyand, J. Yeh, F. L. Baehner, T. Fevr, L. Clark, N. Bayani, J. P. Coppe, F. Tong, T. Speed, P. T. Spellman, S. DeVries, A. Lapuk, N. J. Wang, W. L. Kuo, J. L. Stilwell, D. Pinkel, D. G. Albertson, F. M. Waldman, F. McCormick, R. B. Dickson, M. D. Johnson, M. Lippman, S. Ethier, A. Gazdar, and J. W. Gray. 2006. A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer cell 10:515-527.

Newton, K., and V. M. Dixit. 2012. Signaling in innate immunity and inflammation. Cold Spring Harb Perspect Biol 4.

Niedobitek, G. 2000. Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma. Molecular pathology : MP 53:248-254.

Nikitin, P. A., C. M. Yan, E. Forte, A. Bocedi, J. P. Tourigny, R. E. White, M. J. Allday, A. Patel, S. S. Dave, W. Kim, K. Hu, J. Guo, D. Tainter, E. Rusyn, and M. A. Luftig. 2010. An ATM/Chk2-mediated DNA damage-responsive signaling pathway suppresses Epstein-Barr virus transformation of primary human B cells. Cell Host Microbe 8:510-522.

Nossal, G. J. V. 1983. Cellular mechanisms of immunological tolerance. Annual Review of Immunology 1:33–62.

Ntziachristos, P., A. Tsirigos, P. Van Vlierberghe, J. Nedjic, T. Trimarchi, M. S. Flaherty, D. Ferres-Marco, V. da Ros, Z. Tang, J. Siegle, P. Asp, M. Hadler, I. Rigo, K. De Keersmaecker, J. Patel, T. Huynh, F. Utro, S. Poglio, J. B. Samon, E. Paietta, J. Racevskis, J. M. Rowe, R. Rabadan, R. L. Levine, S. Brown, F. Pflumio, M. Dominguez, A. Ferrando, and I. Aifantis. 2012. Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. Nat Med 18:298-301.

Nurse, P., and P. Thuriaux. 1980. Regulatory genes controlling mitosis in the fission yeast Schizosaccharomyces pombe. Genetics 96:627-637.

Nutt, S. L., B. Heavey, A. G. Rolink, and M. Busslinger. 1999. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature 401:556-562.

O'Meara, M. M., and J. A. Simon. 2012. Inner workings and regulatory inputs that control Polycomb repressive complex 2. Chromosoma 121:221-234.

O'Reilly, R. J., T. N. Small, E. Papadopoulos, K. Lucas, J. Lacerda, and L. Koulova. 1997. Biology and adoptive cell therapy of Epstein-Barr virus-associated lymphoproliferative disorders in recipients of marrow allografts. Immunol Rev 157:195-216.

O'Riordan, M., and R. Grosschedl. 1999. Coordinate regulation of B cell differentiation by the transcription factors EBF and E2A. Immunity 11:21-31.

Odumade, O. A., K. A. Hogquist, and H. H. Balfour, Jr. 2011. Progress and problems in understanding and managing primary Epstein-Barr virus infections. Clin Microbiol Rev 24:193-209.

Oktaba, K., L. Gutierrez, J. Gagneur, C. Girardot, A. K. Sengupta, E. E. Furlong, and J. Muller. 2008. Dynamic regulation by polycomb group protein complexes controls pattern formation and the cell cycle in Drosophila. Dev Cell 15:877-889.

Orphanides, G., T. Lagrange, and D. Reinberg. 1996. The general transcription factors of RNA polymerase II. Genes & development 10:2657-2683.

Osmond, D. Proliferation kinetics and the lifespan of B cells in central and peripheral lymphoid organs. (1991). Current Opinion in Immunology 3:179–185.

Oudet, P., M. Gross-Bellard, and P. Chambon. 1975. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell 4:281-300.

Ouyang, Q., W. M. Wagner, S. Walter, C. A. Muller, A. Wikby, G. Aubert, T. Klatt, S. Stevanovic, T. Dodi, and G. Pawelec. 2003. An age-related increase in the number of CD8+ T cells carrying receptors for an immunodominant Epstein-Barr virus (EBV) epitope is counteracted by a decreased frequency of their antigen-specific responsiveness. Mech Ageing Dev 124:477-485.

Pabo, C. O., and R. T. Sauer. 1992. Transcription factors: structural families and principles of DNA recognition. Annual review of biochemistry 61:1053-1095.

Palaga, T., L. Miele, T. E. Golde, and B. A. Osborne. 2003. TCR-mediated Notch signaling regulates proliferation and IFN-gamma production in peripheral T cells. Journal of immunology 171:3019-3024.

Palermo, R. D., H. M. Webb, and M. J. West. 2011. RNA polymerase II stalling promotes nucleosome occlusion and pTEFb recruitment to drive immortalization by Epstein-Barr virus. PLoS pathogens 7:e1002334.

Paliwal, S., N. Ho, D. Parker, and S. R. Grossman. 2012. CtBP2 Promotes Human Cancer Cell Migration by Transcriptional Activation of Tiam1. Genes Cancer 3:481-490.

Palstra, R. J., B. Tolhuis, E. Splinter, R. Nijmeijer, F. Grosveld, and W. de Laat. 2003. The beta-globin nuclear compartment in development and erythroid differentiation. Nature genetics 35:190-194.

Paschos, K., G. A. Parker, E. Watanatanasup, R. E. White, and M. J. Allday. 2012. BIM promoter directly targeted by EBNA3C in polycomb-mediated repression by EBV. Nucleic acids research 40:7233-7246.

Paschos, K., P. Smith, E. Anderton, J. M. Middeldorp, R. E. White, and M. J. Allday. 2009. Epstein-barr virus latency in B cells leads to epigenetic repression and CpG methylation of the tumour suppressor gene Bim. PLoS pathogens 5:e1000492.

Pasini, D., M. Malatesta, H. R. Jung, J. Walfridsson, A. Willer, L. Olsson, J. Skotte, A. Wutz, B. Porse, O. N. Jensen, and K. Helin. 2010. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic acids research 38:4958-4969.

Pasqualucci, L., G. Bhagat, M. Jankovic, M. Compagno, P. Smith, M. Muramatsu, T. Honjo, H. C. Morse, 3rd, M. C. Nussenzweig, and R. Dalla-Favera. 2008. AID is required for germinal center-derived lymphomagenesis. Nature genetics 40:108-112.

Pathmanathan, R., U. Prasad, R. Sadler, K. Flynn, and N. Raab-Traub. 1995. Clonal proliferations of cells infected with Epstein-Barr virus in preinvasive lesions related to nasopharyngeal carcinoma. The New England journal of medicine 333:693-698.

Patte, C., A. Auperin, J. Michon, H. Behrendt, G. Leverger, D. Frappaz, P. Lutz, C. Coze, Y. Perel, M. Raphael, M. J. Terrier-Lacombe, and P. Societe Francaise d'Oncologie. 2001. The Societe Francaise d'Oncologie Pediatrique LMB89 protocol: highly effective multiagent chemotherapy tailored to the tumor burden and initial response in 561 unselected children with B-cell lymphomas and L3 leukemia. Blood 97:3370-3379.

Pegtel, D. M., A. Subramanian, T. S. Sheen, C. H. Tsai, T. R. Golub, and D. A. Thorley-Lawson. 2005. Epstein-Barr-virus-encoded LMP2A induces primary epithelial cell migration and invasion: possible role in nasopharyngeal carcinoma metastasis. J Virol 79:15430-15442.

Peng, C. W., B. Zhao, H. C. Chen, M. L. Chou, C. Y. Lai, S. Z. Lin, H. Y. Hsu, and E. Kieff. 2007. Hsp72 up-regulates Epstein-Barr virus EBNALP coactivation with EBNA2. Blood 109:5447-5454.

Perez-Andres, M., B. Paiva, W. G. Nieto, A. Caraux, A. Schmitz, J. Almeida, R. F. Vogt, Jr., G. E. Marti, A. C. Rawstron, M. C. Van Zelm, J. J. Van Dongen, H. E. Johnsen, B. Klein, A. Orfao, and M. B. L. Primary Health Care Group of Salamanca for the Study of. 2010. Human peripheral blood B-cell compartments: a crossroad in B-cell traffic. Cytometry. Part B, Clinical cytometry 78 Suppl 1:S47-60.

Peterson, C. L., and M. A. Laniel. 2004. Histones and histone modifications. Curr Biol 14:R546-551.

Plasschaert, R. N., S. Vigneau, I. Tempera, R. Gupta, J. Maksimoska, L. Everett, R. Davuluri, R. Mamorstein, P. M. Lieberman, D. Schultz, S. Hannenhalli, and M. S. Bartolomei. 2014. CTCF binding site sequence differences are associated with unique regulatory and functional trends during embryonic stem cell differentiation. Nucleic acids research 42:774-789.

Pomerantz, M. M., N. Ahmadiyeh, L. Jia, P. Herman, M. P. Verzi, H. Doddapaneni, C. A. Beckwith, J. A. Chan, A. Hills, M. Davis, K. Yao, S. M. Kehoe, H. J. Lenz, C. A. Haiman, C. Yan, B. E. Henderson, B. Frenkel, J. Barretina, A. Bass, J. Tabernero, J. Baselga, M. M. Regan, J. R. Manak, R. Shivdasani, G. A. Coetzee, and M. L. Freedman. 2009. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nature genetics 41:882-884.

Pongubala, J. M., D. L. Northrup, D. W. Lancki, K. L. Medina, T. Treiber, E. Bertolino, M. Thomas, R. Grosschedl, D. Allman, and H. Singh. 2008. Transcription factor EBF restricts

alternative lineage options and promotes B cell fate commitment independently of Pax5. Nature immunology 9:203-215.

Portal, D., B. Zhao, M. A. Calderwood, T. Sommermann, E. Johannsen, and E. Kieff. 2011. EBV nuclear antigen EBNALP dismisses transcription repressors NCoR and RBPJ from enhancers and EBNA2 increases NCoR-deficient RBPJ DNA binding. Proceedings of the National Academy of Sciences of the United States of America 108:7808-7813.

Postigo, A. A., and D. C. Dean. 1999. ZEB represses transcription through interaction with the corepressor CtBP. Proceedings of the National Academy of Sciences of the United States of America 96:6683-6688.

Precopio, M. L., J. L. Sullivan, C. Willard, M. Somasundaran, and K. Luzuriaga. 2003. Differential kinetics and specificity of EBV-specific CD4+ and CD8+ T cells during primary infection. Journal of immunology 170:2590-2598.

Ptashne, M., and A. Gann. 1997. Transcriptional activation by recruitment. Nature 386:569-577.

Pui, J. C., D. Allman, L. Xu, S. DeRocco, F. G. Karnell, S. Bakkour, J. Y. Lee, T. Kadesch, R. R. Hardy, J. C. Aster, and W. S. Pear. 1999. Notch1 expression in early lymphopoiesis influences B versus T lineage determination. Immunity 11:299-308.

Raab-Traub, N., K. Flynn, G. Pearson, A. Huang, P. Levine, A. Lanier, and J. Pagano. 1987. The differentiated form of nasopharyngeal carcinoma contains Epstein-Barr virus DNA. International journal of cancer. Journal international du cancer 39:25-29.

Radkov, S. A., R. Touitou, A. Brehm, M. Rowe, M. West, T. Kouzarides, and M. J. Allday. 1999. Epstein-Barr virus nuclear antigen 3C interacts with histone deacetylase to repress transcription. J Virol 73:5688-5697.

Radtke, F., N. Fasnacht, and H. R. Macdonald. 2010. Notch signaling in the immune system. Immunity 32:14-27.

Radtke, F., A. Wilson, G. Stark, M. Bauer, J. van Meerwijk, H. R. MacDonald, and M. Aguet. 1999. Deficient T cell fate specification in mice with an induced inactivation of Notch1. Immunity 10:547-558.

Recillas-Targa, F., M. J. Pikaart, B. Burgess-Beusse, A. C. Bell, M. D. Litt, A. G. West, M. Gaszner, and G. Felsenfeld. 2002. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. Proceedings of the National Academy of Sciences of the United States of America 99:6883-6888.

Renne, C., J. I. Martin-Subero, M. Eickernjager, M. L. Hansmann, R. Kuppers, R. Siebert, and A. Brauninger. 2006. Aberrant expression of ID2, a suppressor of B-cell-specific gene expression, in Hodgkin's lymphoma. The American journal of pathology 169:655-664.

Ringrose, L., and R. Paro. 2007. Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development 134:223-232.

Robertson, E. S., S. Grossman, E. Johannsen, C. Miller, J. Lin, B. Tomkinson, and E. Kieff. 1995. Epstein-Barr virus nuclear protein 3C modulates transcription through interaction with the sequence-specific DNA-binding protein J kappa. Journal of virology 69:3108-3116.

Robertson, E. S., J. Lin, and E. Kieff. 1996. The amino-terminal domains of Epstein-Barr virus nuclear proteins 3A, 3B, and 3C interact with RBPJ(kappa). Journal of virology 70:3068-3074.

Robertson, M. J. 2002. Role of chemokines in the biology of natural killer cells. J Leukoc Biol 71:173-183.

Roche, J., P. Nasarre, R. Gemmill, A. Baldys, J. Pontis, C. Korch, J. Guilhot, S. Ait-Si-Ali, and H. Drabkin. 2013. Global Decrease of Histone H3K27 Acetylation in ZEB1-Induced Epithelial to Mesenchymal Transition in Lung Cancer Cells. Cancers (Basel) 5:334-356.

Rojas, J. R., R. C. Trievel, J. Zhou, Y. Mo, X. Li, S. L. Berger, C. D. Allis, and R. Marmorstein. 1999. Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. Nature 401:93-98.

Rolink, A. G., S. L. Nutt, F. Melchers, and M. Busslinger. 1999. Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. Nature 401:603-606.

Romanow, W. J., A. W. Langerak, P. Goebel, I. L. Wolvers-Tettero, J. J. van Dongen, A. J. Feeney, and C. Murre. 2000. E2A and EBF act in synergy with the V(D)J recombinase to generate a diverse immunoglobulin repertoire in nonlymphoid cells. Molecular cell 5:343-353.

Roth, S. Y., and C. D. Allis. 1992. Chromatin condensation: does histone H1 dephosphorylation play a role? Trends Biochem Sci 17:93-98.

Roughan, J. E., and D. A. Thorley-Lawson. 2009. The intersection of Epstein-Barr virus with the germinal center. J Virol 83:3968-3976.

Roughan, J. E., C. Torgbor, and D. A. Thorley-Lawson. 2010. Germinal center B cells latently infected with Epstein-Barr virus proliferate extensively but do not increase in number. J Virol 84:1158-1168.

Rowland S. L., K. Tuttle, R. M. Torres, R. Pelanda. 2013. Antigen and cytokine receptor signals guide the development of the naïve mature B cell repertoire. Immunologic Research 55(1-3):231-40.

Rubio, E. D., D. J. Reiss, P. L. Welcsh, C. M. Disteche, G. N. Filippova, N. S. Baliga, R. Aebersold, J. A. Ranish, and A. Krumm. 2008. CTCF physically links cohesin to chromatin. Proceedings of the National Academy of Sciences of the United States of America 105:8309-8314.

Russell, P., and P. Nurse. 1987. Negative regulation of mitosis by wee1+, a gene encoding a protein kinase homolog. Cell 49:559-567.

Saito, T., S. Chiba, M. Ichikawa, A. Kunisato, T. Asai, K. Shimizu, T. Yamaguchi, G. Yamamoto, S. Seo, K. Kumano, E. Nakagami-Yamaguchi, Y. Hamada, S. Aizawa, and H. Hirai. 2003. Notch2 is preferentially expressed in mature B cells and indispensable for marginal zone B lineage development. Immunity 18:675-685.

Sakai, T., Y. Taniguchi, K. Tamura, S. Minoguchi, T. Fukuhara, L. J. Strobl, U. Zimber-Strobl, G. W. Bornkamm, and T. Honjo. 1998. Functional replacement of the intracellular region of the Notch1 receptor by Epstein-Barr virus nuclear antigen 2. Journal of virology 72:6034-6039.

Sample, J., E. B. Henson, and C. Sample. 1992. The Epstein-Barr virus nuclear protein 1 promoter active in type I latency is autoregulated. Journal of virology 66:4654-4661.

Santos, M. A., L. M. Sarmento, M. Rebelo, A. A. Doce, I. Maillard, A. Dumortier, H. Neves, F. Radtke, W. S. Pear, L. Parreira, and J. Demengeot. 2007. Notch1 engagement by Delta-like-1 promotes differentiation of B lymphocytes to antibody-secreting cells. Proceedings of the National Academy of Sciences of the United States of America 104:15454-15459.

Sarma, K., R. Margueron, A. Ivanov, V. Pirrotta, and D. Reinberg. 2008. Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Molecular and cellular biology 28:2718-2731.

Schaeper, U., J. M. Boyd, S. Verma, E. Uhlmann, T. Subramanian, and G. Chinnadurai. 1995. Molecular cloning and characterization of a cellular phosphoprotein that interacts with a conserved C-terminal domain of adenovirus E1A involved in negative modulation of oncogenic transformation. Proceedings of the National Academy of Sciences of the United States of America 92:10467-10471.

Schlick, S. N., C. D. Wood, A. Gunnell, H. M. Webb, S. Khasnis, A. Schepers, and M. J. West. 2011. Upregulation of the cell-cycle regulator RGC-32 in Epstein-Barr virus-immortalized cells. PLoS One 6:e28638.

Schmidlin, H., S. A. Diehl , B. Blom. 2009. New insights into the regulation of human B-cell differentiation. Trends in Immunology 30(6):277-85.

Scholz, J.L., M.P. Cancro. 2012. Resolve, revise, and relax: the 3 Rs of B cell repertoire adjustment. Immunology Letters 143(1):2-8.

Schram B.R., L. E. Tze, L. B. Ramsey, J. Liu, L. Najera, A. L. Vegoe, R. R. Hardy, K. L. Hippen, M. A. Farrar, T. W. Behrens. 2008. B cell receptor basal signaling regulates antigen-induced Ig light chain rearrangements. Journal of Immunology 180(7):4728–41.

Schuettengruber, B., and G. Cavalli. 2009. Recruitment of polycomb group complexes and their role in the dynamic regulation of cell fate choice. Development 136:3531-3542.

Seet, C. S., R. L. Brumbaugh, and B. L. Kee. 2004. Early B cell factor promotes B lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. The Journal of experimental medicine 199:1689-1700.

Sgadari, C., A. L. Angiolillo, B. W. Cherney, S. E. Pike, J. M. Farber, L. G. Koniaris, P. Vanguri, P. R. Burd, N. Sheikh, G. Gupta, J. Teruya-Feldstein, and G. Tosato. 1996. Interferon-inducible protein-10 identified as a mediator of tumor necrosis in vivo. Proceedings of the National Academy of Sciences of the United States of America 93:13791-13796.

Shaffer A. L., K. I. Lin, T. C. Kuo, X. Yu, E. M. Hurt, A. Rosenwald, J. M. Giltnane, L. Yang, H. Zhao, K. Calame, L. M. Staudt. 2002. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. Immunity 17(1):51-62.

Shi, Y., J. Sawada, G. Sui, B. Affar el, J. R. Whetstine, F. Lan, H. Ogawa, M. P. Luke, and Y. Nakatani. 2003. Coordinated histone modifications mediated by a CtBP co-repressor complex. Nature 422:735-738.

Shukla, S., E. Kavak, M. Gregory, M. Imashimizu, B. Shutinoski, M. Kashlev, P. Oberdoerffer, R. Sandberg, and S. Oberdoerffer. 2011. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. Nature 479:74-79.

Siemer, D., J. Kurth, S. Lang, G. Lehnerdt, J. Stanelle, and R. Kuppers. 2008. EBV transformation overrides gene expression patterns of B cell differentiation stages. Mol Immunol 45:3133-3141.

Simon, J. A., and R. E. Kingston. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. Nature reviews. Molecular cell biology 10:697-708.

Simon, J. A., and R. E. Kingston. 2013. Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. Molecular cell 49:808-824.

Sinclair, A. J., I. Palmero, G. Peters, and P. J. Farrell. 1994. EBNA-2 and EBNA-LP cooperate to cause G0 to G1 transition during immortalization of resting human B lymphocytes by Epstein-Barr virus. The EMBO journal 13:3321-3328.

Siponen, M. I., M. Wisniewska, L. Lehtio, I. Johansson, L. Svensson, G. Raszewski, L. Nilsson, M. Sigvardsson, and H. Berglund. 2010. Structural determination of functional domains in early B-cell factor (EBF) family of transcription factors reveals similarities to Rel DNA-binding proteins and a novel dimerization motif. The Journal of biological chemistry 285:25875-25879.

Sive, J. I., and B. Gottgens. 2014. Transcriptional network control of normal and leukaemic haematopoiesis. Experimental cell research.

Sjoblom, A., A. Jansson, W. Yang, S. Lain, T. Nilsson, and L. Rymo. 1995. PU box-binding transcription factors and a POU domain protein cooperate in the Epstein-Barr virus (EBV) nuclear antigen 2-induced transactivation of the EBV latent membrane protein 1 promoter. The Journal of general virology 76 (Pt 11):2679-2692.

Sjoblom, A., W. Yang, L. Palmqvist, A. Jansson, and L. Rymo. 1998. An ATF/CRE element mediates both EBNA2-dependent and EBNA2-independent activation of the Epstein-Barr virus LMP1 gene promoter. Journal of virology 72:1365-1376.

Skalska, L., R. E. White, M. Franz, M. Ruhmann, and M. J. Allday. 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., R. E. White, G. A. Parker, E. Turro, A. J. Sinclair, K. Paschos, and M. J. Allday. 2013. Induction of p16(INK4a) is the major barrier to proliferation when Epstein-Barr virus (EBV) transforms primary B cells into lymphoblastoid cell lines. PLoS pathogens 9:e1003187.

Smale, S. T., and J. T. Kadonaga. 2003. The RNA polymerase II core promoter. Annual review of biochemistry 72:449-479.

Smith, E. M., R. Gisler, and M. Sigvardsson. 2002. Cloning and characterization of a promoter flanking the early B cell factor (EBF) gene indicates roles for E-proteins and autoregulation in the control of EBF expression. Journal of immunology 169:261-270.

Sollerbrant, K., G. Chinnadurai, and C. Svensson. 1996. The CtBP binding domain in the adenovirus E1A protein controls CR1-dependent transactivation. Nucleic acids research 24:2578-2584.

Souza, T. A., B. D. Stollar, J. L. Sullivan, K. Luzuriaga, and D. A. Thorley-Lawson. 2005. Peripheral B cells latently infected with Epstein-Barr virus display molecular hallmarks of classical antigen-selected memory B cells. Proc Natl Acad Sci U S A 102:18093-18098.

Spender, L. C., G. H. Cornish, A. Sullivan, and P. J. Farrell. 2002. Expression of transcription factor AML-2 (RUNX3, CBF(alpha)-3) is induced by Epstein-Barr virus EBNA-2 and correlates with the B-cell activation phenotype. Journal of virology 76:4919-4927.

Spender, L. C., and G. J. Inman. 2011. Inhibition of germinal centre apoptotic programmes by epstein-barr virus. Adv Hematol 2011:829525.

Spender, L. C., W. Lucchesi, G. Bodelon, A. Bilancio, C. E. Karstegl, T. Asano, O. Dittrich-Breiholz, M. Kracht, B. Vanhaesebroeck, and P. J. Farrell. 2006. Cell target genes of Epstein-Barr virus transcription factor EBNA-2: induction of the p55alpha regulatory subunit of PI3-kinase and its role in survival of EREB2.5 cells. J Gen Virol 87:2859-2867.

Spilianakis, C. G., M. D. Lalioti, T. Town, G. R. Lee, and R. A. Flavell. 2005. Interchromosomal associations between alternatively expressed loci. Nature 435:637-645.

Splinter, E., H. Heath, J. Kooren, R. J. Palstra, P. Klous, F. Grosveld, N. Galjart, and W. de Laat. 2006. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes & development 20:2349-2354.

Springer, T. A. 1990. Adhesion receptors of the immune system. Nature 346:425-434.

Srinivasan, L., and M. L. Atchison. 2004. YY1 DNA binding and PcG recruitment requires CtBP. Genes & development 18:2596-2601.

Stadanlick, J.E., M. Kaileh, F. G. Karnell, J. L. Scholz, J. P. Miller, W.J. Quinn, 3rd, R. J. Brezski, L. S. Treml, K. A. Jordan, J.G. Monroe, R. Sen, M.P. Cancro. 2008. Tonic B cell antigen receptor signals supply an NF-kappaB substrate for prosurvival BLyS signaling. Nature Immunology 9:1379–1387.

Strahl, B. D., and C. D. Allis. 2000. The language of covalent histone modifications. Nature 403:41-45.

Stratowa, C., G. Loffler, P. Lichter, S. Stilgenbauer, P. Haberl, N. Schweifer, H. Dohner, and K. K. Wilgenbus. 2001. CDNA microarray gene expression analysis of B-cell chronic lymphocytic leukemia proposes potential new prognostic markers involved in lymphocyte trafficking. International journal of cancer. Journal international du cancer 91:474-480.

Strobl, L. J., H. Hofelmayr, G. Marschall, M. Brielmeier, G. W. Bornkamm, and U. Zimber-Strobl. 2000. Activated Notch1 modulates gene expression in B cells similarly to Epstein-Barr viral nuclear antigen 2. Journal of virology 74:1727-1735.

Su G. H., H. M. Chen, N. Muthusamy, L. A. Garrett-Sinha, D. Baunoch, D.G. Tenen, M.C. Simon. 1997. Defective B cell receptor-mediated responses in mice lacking the Ets protein, Spi-B. The EMBO Journal 16(23):7118-29.

Subramanian, C., S. Hasan, M. Rowe, M. Hottiger, R. Orre, and E. S. Robertson. 2002. Epstein-Barr virus nuclear antigen 3C and prothymosin alpha interact with the p300

transcriptional coactivator at the CH1 and CH3/HAT domains and cooperate in regulation of transcription and histone acetylation. J Virol 76:4699-4708.

Subramanian, T., L. J. Zhao, and G. Chinnadurai. 2013. Interaction of CtBP with adenovirus E1A suppresses immortalization of primary epithelial cells and enhances virus replication during productive infection. Virology 443:313-320.

Swedlow, J. R., and T. Hirano. 2003. The making of the mitotic chromosome: modern insights into classical questions. Molecular cell 11:557-569.

Taby, R., and J. P. Issa. 2010. Cancer epigenetics. CA Cancer J Clin 60:376-392.

Tamaru, Y., T. Miyawaki, K. Iwai, T. Tsuji, R. Nibu, A. Yachie, S. Koizumi, and N. Taniguchi. 1993. Absence of bcl-2 expression by activated CD45RO+ T lymphocytes in acute infectious mononucleosis supporting their susceptibility to programmed cell death. Blood 82:521-527.

Tan, L., and Y. G. Shi. 2012. Tet family proteins and 5-hydroxymethylcytosine in development and disease. Development 139:1895-1902.

Tanimoto, K., Q. Liu, J. Bungert, and J. D. Engel. 1999. Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. Nature 398:344-348.

Tanner, J., Y. Whang, J. Sample, A. Sears, and E. Kieff. 1988. Soluble gp350/220 and deletion mutant glycoproteins block Epstein-Barr virus adsorption to lymphocytes. J Virol 62:4452-4464.

Taylor, S. T., J. A. Hickman, and C. Dive. 1999. Survival signals within the tumour microenvironment suppress drug-induced apoptosis: lessons learned from B lymphomas. Endocr Relat Cancer 6:21-23.

Tempera, I., M. Klichinsky, and P. M. Lieberman. 2011. EBV latency types adopt alternative chromatin conformations. PLoS pathogens 7:e1002180.

Thal, M. A., T. L. Carvalho, T. He, H. G. Kim, H. Gao, J. Hagman, and C. A. Klug. 2009. Ebf1-mediated down-regulation of Id2 and Id3 is essential for specification of the B cell lineage. Proceedings of the National Academy of Sciences of the United States of America 106:552-557.

Thanos, D., and T. Maniatis. 1995. Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. Cell 83:1091-1100.

Thomas, M., M. Calamito, B. Srivastava, I. Maillard, W. S. Pear, and D. Allman. 2007. Notch activity synergizes with B-cell-receptor and CD40 signaling to enhance B-cell activation. Blood 109:3342-3350.

Thomson, S., L. C. Mahadevan, and A. L. Clayton. 1999. MAP kinase-mediated signalling to nucleosomes and immediate-early gene induction. Semin Cell Dev Biol 10:205-214.

Thuriaux, P., P. Nurse, and B. Carter. 1978. Mutants altered in the control co-ordinating cell division with cell growth in the fission yeast Schizosaccharomyces pombe. Mol Gen Genet 161:215-220.

Tie, F., R. Banerjee, P. A. Conrad, P. C. Scacheri, and P. J. Harte. 2012. Histone demethylase UTX and chromatin remodeler BRM bind directly to CBP and modulate acetylation of histone H3 lysine 27. Molecular and cellular biology 32:2323-2334.

Tie, F., R. Banerjee, C. A. Stratton, J. Prasad-Sinha, V. Stepanik, A. Zlobin, M. O. Diaz, P. C. Scacheri, and P. J. Harte. 2009. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development 136:3131-3141.

Tierney, R. J., H. E. Kirby, J. K. Nagra, J. Desmond, A. I. Bell, and A. B. Rickinson. 2000. Methylation of transcription factor binding sites in the Epstein-Barr virus latent cycle promoter Wp coincides with promoter down-regulation during virus-induced B-cell transformation. J Virol 74:10468-10479.

Tobollik, S., L. Meyer, M. Buettner, S. Klemmer, B. Kempkes, E. Kremmer, G. Niedobitek, and B. Jungnickel. 2006. Epstein-Barr virus nuclear antigen 2 inhibits AID expression during EBV-driven B-cell growth. Blood 108:3859-3864.

Tomkinson, B., and E. Kieff. 1992. Use of second-site homologous recombination to demonstrate that Epstein-Barr virus nuclear protein 3B is not important for lymphocyte infection or growth transformation in vitro. Journal of virology 66:2893-2903.

Tomkinson, B., E. Robertson, and E. Kieff. 1993. Epstein-Barr virus nuclear proteins EBNA-3A and EBNA-3C are essential for B-lymphocyte growth transformation. Journal of virology 67:2014-2025.

Tong, X., R. Drapkin, D. Reinberg, and E. Kieff. 1995a. The 62- and 80-kDa subunits of transcription factor IIH mediate the interaction with Epstein-Barr virus nuclear protein 2. Proceedings of the National Academy of Sciences of the United States of America 92:3259-3263.

Tong, X., R. Drapkin, R. Yalamanchili, G. Mosialos, and E. Kieff. 1995b. The Epstein-Barr virus nuclear protein 2 acidic domain forms a complex with a novel cellular coactivator that can interact with TFIIE. Molecular and cellular biology 15:4735-4744.

Tong, X., F. Wang, C. J. Thut, and E. Kieff. 1995c. The Epstein-Barr virus nuclear protein 2 acidic domain can interact with TFIIB, TAF40, and RPA70 but not with TATA-binding protein. Journal of virology 69:585-588.

Torlakovic, E., A. Tierens, H. D. Dang, and J. Delabie. 2001. The transcription factor PU.1, necessary for B-cell development is expressed in lymphocyte predominance, but not classical Hodgkin's disease. The American journal of pathology 159:1807-1814.

Touitou, R., M. Hickabottom, G. Parker, T. Crook, and M. J. Allday. 2001. Physical and functional interactions between the corepressor CtBP and the Epstein-Barr virus nuclear antigen EBNA3C. Journal of virology 75:7749-7755.

Touitou, R., J. O'Nions, J. Heaney, and M. J. Allday. 2005. Epstein-Barr virus EBNA3 proteins bind to the C8/alpha7 subunit of the 20S proteasome and are degraded by 20S proteasomes in vitro, but are very stable in latently infected B cells. The Journal of general virology 86:1269-1277.

Treiber, T., E. M. Mandel, S. Pott, I. Gyory, S. Firner, E. T. Liu, and R. Grosschedl. 2010. Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poising of chromatin. Immunity 32:714-725.

Tsurumi, T., M. Fujita, and A. Kudoh. 2005. Latent and lytic Epstein-Barr virus replication strategies. Rev Med Virol 15:3-15.

Tzellos, S., P. B. Correia, C. E. Karstegl, L. Cancian, J. Cano-Flanagan, M. J. McClellan, M. J. West, and P. J. Farrell. 2014. A Single Amino Acid in EBNA-2 Determines Superior B Lymphoblastoid Cell Line Growth Maintenance by Epstein-Barr Virus Type 1 EBNA-2. Journal of virology 88:8743-8753.

Ullius, A., J. Luscher-Firzlaff, I. G. Costa, G. Walsemann, A. H. Forst, E. G. Gusmao, K. Kapelle, H. Kleine, E. Kremmer, J. Vervoorts, and B. Luscher. 2014. The interaction of MYC with the trithorax protein ASH2L promotes gene transcription by regulating H3K27 modification. Nucleic acids research 42:6901-6920.

Van Bortle, K., E. Ramos, N. Takenaka, J. Yang, J. E. Wahi, and V. G. Corces. 2012. Drosophila CTCF tandemly aligns with other insulator proteins at the borders of H3K27me3 domains. Genome research 22:2176-2187.

van den Borne, P., P. H. Quax, I. E. Hoefer, and G. Pasterkamp. 2014. The multifaceted functions of CXCL10 in cardiovascular disease. Biomed Res Int 2014:893106.

van den Bosch, C. A. 2004. Is endemic Burkitt's lymphoma an alliance between three infections and a tumour promoter? The Lancet. Oncology 5:738-746.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. Zhang, G. L. Gabor Miklos, C. Nelson, S. Broder, A. G. Clark, J. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. Ge, F. Gong, Z. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. Ke, K. A. Ketchum, Z. Lai, Y. Lei, Z. Li, J. Li, Y. Liang, X. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W. Shao, B. Shue, J. Sun, Z. Wang, A. Wang, X. Wang, J. Wang, M. Wei, R. Wides, C. Xiao, C. Yan, A. Yao, J. Ye, M. Zhan, W. Zhang, H. Zhang, Q. Zhao, L. Zheng, F. Zhong, W. Zhong, S. Zhu, S. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M. L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y. H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N. N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J. F. Abril, R. Guigo, M. J. Campbell, K. V. Sjolander, B. Karlak, A. Kejariwal, H. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y. H. Chiang, M. Coyne, C. Dahlke, A. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S. Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh, and X. Zhu. 2001. The sequence of the human genome. Science 291:1304-1351.

Verkoczy L., B. Duong, P. Skog, D. Ait-Azzouzene, K. Puri, J. L. Vela, D. Nemazee. 2007. Basal B cell receptor-directed phosphatidylinositol 3-kinase signaling turns off RAGs and promotes B cell-positive selection. Journal of Immunology 178(10):6332–41.

Vilar, J. M., and L. Saiz. 2005. DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. Current opinion in genetics & development 15:136-144.

Voigt, P., W. W. Tee, and D. Reinberg. 2013. A double take on bivalent promoters. Genes Dev 27:1318-1338.

Wada, T., T. Takagi, Y. Yamaguchi, D. Watanabe, and H. Handa. 1998. Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. The EMBO journal 17:7395-7403.

Waltzer, L., F. Logeat, C. Brou, A. Israel, A. Sergeant, and E. Manet. 1994. The human J kappa recombination signal sequence binding protein (RBP-J kappa) targets the Epstein-Barr virus EBNA2 protein to its DNA responsive elements. The EMBO journal 13:5633-5638.

Waltzer, L., M. Perricaudet, A. Sergeant, and E. Manet. 1996. Epstein-Barr virus EBNA3A and EBNA3C proteins both repress RBP-J kappa-EBNA2-activated transcription by inhibiting the binding of RBP-J kappa to DNA. Journal of virology 70:5909-5915.

Wang, F., C. Gregory, C. Sample, M. Rowe, D. Liebowitz, R. Murray, A. Rickinson, and E. Kieff. 1990a. Epstein-Barr virus latent membrane protein (LMP1) and nuclear proteins 2 and 3C are effectors of phenotypic changes in B lymphocytes: EBNA-2 and LMP1 cooperatively induce CD23. Journal of virology 64:2309-2318.

Wang, F., C. D. Gregory, M. Rowe, A. B. Rickinson, D. Wang, M. Birkenbach, H. Kikutani, T. Kishimoto, and E. Kieff. 1987. Epstein-Barr virus nuclear antigen 2 specifically induces expression of the B-cell activation antigen CD23. Proceedings of the National Academy of Sciences of the United States of America 84:3452-3456.

Wang, F., S. F. Tsang, M. G. Kurilla, J. I. Cohen, and E. Kieff. 1990b. Epstein-Barr virus nuclear antigen 2 transactivates latent membrane protein LMP1. Journal of virology 64:3407-3416.

Wang, H., M. T. Maurano, H. Qu, K. E. Varley, J. Gertz, F. Pauli, K. Lee, T. Canfield, M. Weaver, R. Sandstrom, R. E. Thurman, R. Kaul, R. M. Myers, and J. A. Stamatoyannopoulos. 2012. Widespread plasticity in CTCF occupancy linked to DNA methylation. Genome research 22:1680-1688.

Wang, L., S. R. Grossman, and E. Kieff. 2000. Epstein-Barr virus nuclear protein 2 interacts with p300, CBP, and PCAF histone acetyltransferases in activation of the LMP1 promoter. Proceedings of the National Academy of Sciences of the United States of America 97:430-435.

Wang, S., M. H. Linde, M. Munde, V. D. Carvalho, W. D. Wilson, and G. M. Poon. 2014. Mechanistic Heterogeneity in Site Recognition by the Structurally Homologous DNA-binding Domains of the ETS Family Transcription Factors Ets-1 and PU.1. The Journal of biological chemistry 289:21605-21616.

Wang, Z., J. Song, T. A. Milne, G. G. Wang, H. Li, C. D. Allis, and D. J. Patel. 2010. Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. Cell 141:1183-1194.

Wardemann H., S. Yurasov, A. Schaefer, J. W. Young, E. Meffre, M. C. Nussenzweig. 2003. Predominant autoantibody production by early human B cell precursors. Science 301(5638):1374–7.

Watanabe, N., M. Broome, and T. Hunter. 1995. Regulation of the human WEE1Hu CDK tyrosine 15-kinase during the cell cycle. The EMBO journal 14:1878-1891.

Weng, A. P., J. M. Millholland, Y. Yashiro-Ohtani, M. L. Arcangeli, A. Lau, C. Wai, C. Del Bianco, C. G. Rodriguez, H. Sai, J. Tobias, Y. Li, M. S. Wolfe, C. Shachaf, D. Felsher, S. C. Blacklow, W. S. Pear, and J. C. Aster. 2006. c-Myc is an important direct target of Notch1 in T-cell acute lymphoblastic leukemia/lymphoma. Genes & development 20:2096-2109.

Wensing, B., and P. J. Farrell. 2000. Regulation of cell growth and death by Epstein-Barr virus. Microbes and infection / Institut Pasteur 2:77-84.

Westhoff Smith, D., and B. Sugden. 2013. Potential cellular functions of Epstein-Barr Nuclear Antigen 1 (EBNA1) of Epstein-Barr Virus. Viruses 5:226-240.

White, R. E., M. A. Calderwood, and A. Whitehouse. 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. The Journal of general virology 84:3393-3403.

White, R. E., I. J. Groves, E. Turro, J. Yee, E. Kremmer, and M. J. Allday. 2010. Extensive co-operation between the Epstein-Barr virus EBNA3 proteins in the manipulation of host gene expression and epigenetic chromatin modification. PLoS One 5:e13979.

Williams, H., and D. H. Crawford. 2006. Epstein-Barr virus: the impact of scientific advances on clinical practice. Blood 107:862-869.

Woisetschlaeger, M., X. W. Jin, C. N. Yandava, L. A. Furmanski, J. L. Strominger, and S. H. Speck. 1991. Role for the Epstein-Barr virus nuclear antigen 2 in viral promoter switching during initial stages of infection. Proceedings of the National Academy of Sciences of the United States of America 88:3942-3946.

Wolf, H., H. Zur Hausen, G. Klein, V. Becker, G. Henle, and W. Henle. 1975. Attempts to detect virus-specific DNA sequences in human tumors. III. Epstein-Barr viral DNA in non-lymphoid nasopharyngeal carcinoma cells. Medical microbiology and immunology 161:15-21.

Wolffe, A. P., and J. J. Hayes. 1999. Chromatin disruption and modification. Nucleic acids research 27:711-720.

Woo, C. J., P. V. Kharchenko, L. Daheron, P. J. Park, and R. E. Kingston. 2010. A region of the human HOXD cluster that confers polycomb-group responsiveness. Cell 140:99-110.

Wood, V. H., J. D. O'Neil, W. Wei, S. E. Stewart, C. W. Dawson, and L. S. Young. 2007. Epstein-Barr virus-encoded EBNA1 regulates cellular gene transcription and modulates the STAT1 and TGFbeta signaling pathways. Oncogene 26:4135-4147.

Woodcock, C. L., and S. Dimitrov. 2001. Higher-order structure of chromatin and chromosomes. Curr Opin Genet Dev 11:130-135.

Wright, D. H. 1999. What is Burkitt's lymphoma and when is it endemic? Blood 93:758.

Wu, D. Y., A. Krumm, and W. H. Schubach. 2000. Promoter-specific targeting of human SWI-SNF complex by Epstein-Barr virus nuclear protein 2. Journal of virology 74:8893-8903.

Wu, H., A. C. D'Alessio, S. Ito, K. Xia, Z. Wang, K. Cui, K. Zhao, Y. E. Sun, and Y. Zhang. 2011. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature 473:389-393.

Wu, T. C. 2007. The role of vascular cell adhesion molecule-1 in tumor immune evasion. Cancer Res 67:6003-6006.

Wu, X., J. V. Johansen, and K. Helin. 2013. Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation. Molecular cell 49:1134-1146.

Xu, C., C. Bian, W. Yang, M. Galka, H. Ouyang, C. Chen, W. Qiu, H. Liu, A. E. Jones, F. MacKenzie, P. Pan, S. S. Li, H. Wang, and J. Min. 2010. Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proceedings of the National Academy of Sciences of the United States of America 107:19266-19271.

Xu, K., Z. J. Wu, A. C. Groner, H. H. He, C. Cai, R. T. Lis, X. Wu, E. C. Stack, M. Loda, T. Liu, H. Xu, L. Cato, J. E. Thornton, R. I. Gregory, C. Morrissey, R. L. Vessella, R. Montironi, C. Magi-Galluzzi, P. W. Kantoff, S. P. Balk, X. S. Liu, and M. Brown. 2012. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. Science 338:1465-1469.

Yang, J. J., D. Bhojwani, W. Yang, X. Cai, G. Stocco, K. Crews, J. Wang, D. Morrison, M. Devidas, S. P. Hunger, C. L. Willman, E. A. Raetz, C. H. Pui, W. E. Evans, M. V. Relling, and W. L. Carroll. 2008. Genome-wide copy number profiling reveals molecular evolution from diagnosis to relapse in childhood acute lymphoblastic leukemia. Blood 112:4178-4183.

Yang, Z., J. H. Yik, R. Chen, N. He, M. K. Jang, K. Ozato, and Q. Zhou. 2005. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. Molecular cell 19:535-545.

Yang, Z. H., C. L. Zhou, H. Zhu, J. H. Li, and C. D. He. 2014. A functional SNP in the MDM2 promoter mediates E2F1 affinity to modulate cyclin D1 expression in tumor cell proliferation. Asian Pacific journal of cancer prevention : APJCP 15:3817-3823.

Yoon, S. O., I. Y. Lee, X. Zhang, M. C. Zapata, and Y. S. Choi. 2014. CD9 may contribute to the survival of human germinal center B cells by facilitating the interaction with follicular dendritic cells. FEBS Open Bio 4:370-376.

Yoon, S. O., X. Zhang, P. Berner, B. Blom, and Y. S. Choi. 2009. Notch ligands expressed by follicular dendritic cells protect germinal center B cells from apoptosis. Journal of immunology 183:352-358.

Young, L. S., C. W. Dawson, and A. G. Eliopoulos. 2000. The expression and function of Epstein-Barr virus encoded latent genes. Molecular pathology : MP 53:238-247.

Young, L. S., and P. G. Murray. 2003. Epstein-Barr virus and oncogenesis: from latent genes to tumours. Oncogene 22:5108-5121.

Young, L. S., and A. B. Rickinson. 2004. Epstein-Barr virus: 40 years on. Nature reviews. Cancer 4:757-768.

Yu, J., D. R. Rhodes, S. A. Tomlins, X. Cao, G. Chen, R. Mehra, X. Wang, D. Ghosh, R. B. Shah, S. Varambally, K. J. Pienta, and A. M. Chinnaiyan. 2007. A polycomb repression signature in metastatic prostate cancer predicts cancer outcome. Cancer Res 67:10657-10663.

Yu, M., T. Mazor, H. Huang, H. T. Huang, K. L. Kathrein, A. J. Woo, C. R. Chouinard, A. Labadorf, T. E. Akie, T. B. Moran, H. Xie, S. Zacharek, I. Taniuchi, R. G. Roeder, C. F. Kim, L. I. Zon, E. Fraenkel, and A. B. Cantor. 2012. Direct recruitment of polycomb repressive complex 1 to chromatin by core binding transcription factors. Molecular cell 45:330-343.

Yuan, W., T. Wu, H. Fu, C. Dai, H. Wu, N. Liu, X. Li, M. Xu, Z. Zhang, T. Niu, Z. Han, J. Chai, X. J. Zhou, S. Gao, and B. Zhu. 2012. Dense chromatin activates Polycomb repressive complex 2 to regulate H3 lysine 27 methylation. Science 337:971-975.

Yuan, W., M. Xu, C. Huang, N. Liu, S. Chen, and B. Zhu. 2011. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. The Journal of biological chemistry 286:7983-7989.

Zampieri, M., T. Guastafierro, R. Calabrese, F. Ciccarone, M. G. Bacalini, A. Reale, M. Perilli, C. Passananti, and P. Caiafa. 2012. ADP-ribose polymers localized on Ctcf-Parp1-Dnmt1 complex prevent methylation of Ctcf target sites. The Biochemical journal 441:645-652.

Zaret, K. S., and J. S. Carroll. 2011. Pioneer transcription factors: establishing competence for gene expression. Genes & development 25:2227-2241.

Zentner, G. E., and S. Henikoff. 2013. Regulation of nucleosome dynamics by histone modifications. Nat Struct Mol Biol 20:259-266.

Zetterberg, H., C. Borestrom, T. Nilsson, and L. Rymo. 2004. Multiple EBNA1-binding sites within oriPI are required for EBNA1-dependent transactivation of the Epstein-Barr virus C promoter. Int J Oncol 25:693-696.

Zhang, C., C. Gao, Y. Xu, and Z. Zhang. 2014. CtBP2 could promote prostate cancer cell proliferation through c-Myc signaling. Gene 546:73-79.

Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li, and X. S. Liu. 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biol 9:R137.

Zhang, Z., C. V. Cotta, R. P. Stephan, C. G. deGuzman, and C. A. Klug. 2003. Enforced expression of EBF in hematopoietic stem cells restricts lymphopoiesis to the B cell lineage. The EMBO journal 22:4759-4769.

Zhao, B., J. C. Mar, S. Maruo, S. Lee, B. E. Gewurz, E. Johannsen, K. Holton, R. Rubio, K. Takada, J. Quackenbush, and E. Kieff. 2006a. Epstein-Barr virus nuclear antigen 3C regulated genes in lymphoblastoid cell lines. Proc Natl Acad Sci U S A 108:337-342.

Zhao, B., S. Maruo, A. Cooper, R. C. M, E. Johannsen, E. Kieff, and E. Cahir-McFarland. 2006b. RNAs induced by Epstein-Barr virus nuclear antigen 2 in lymphoblastoid cell lines. Proc Natl Acad Sci U S A 103:1900-1905.

Zhao, B., and C. E. Sample. 2000. Epstein-barr virus nuclear antigen 3C activates the latent membrane protein 1 promoter in the presence of Epstein-Barr virus nuclear antigen 2 through sequences encompassing an spi-1/Spi-B binding site. Journal of virology 74:5151-5160.

Zhao, B., J. Zou, H. Wang, E. Johannsen, C. W. Peng, J. Quackenbush, J. C. Mar, C. C. Morton, M. L. Freedman, S. C. Blacklow, J. C. Aster, B. E. Bernstein, and E. Kieff. 2011. 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, L. J., T. Subramanian, S. Vijayalingam, and G. Chinnadurai. 2014. CtBP2 proteome: Role of CtBP in E2F7-mediated repression and cell proliferation. Genes Cancer 5:31-40.

Zhao, L. J., T. Subramanian, Y. Zhou, and G. Chinnadurai. 2006c. Acetylation by p300 regulates nuclear localization and function of the transcriptional corepressor CtBP2. The Journal of biological chemistry 281:4183-4189.

Zheng, H., L. L. Li, D. S. Hu, X. Y. Deng, and Y. Cao. 2007. Role of Epstein-Barr virus encoded latent membrane protein 1 in the carcinogenesis of nasopharyngeal carcinoma. Cellular & molecular immunology 4:185-196.

Zimber-Strobl, U., E. Kremmer, F. Grasser, G. Marschall, G. Laux, and G. W. Bornkamm. 1993. The Epstein-Barr virus nuclear antigen 2 interacts with an EBNA2 responsive cis-element of the terminal protein 1 gene promoter. The EMBO journal 12:167-175.

Appendix A

antibodies ChIP and immunoprecipitation antibodies							
protein	antibody	antibody species	quantity used (µg)	company/ generated by	secondary antibody	quantity used (µg)	company/ generated by
EBNA 3	EBNA 3C (ab16128)	Sheep polyclonal	10	Abcam			
EBNA 3C	E3CD8	Mouse monoclonal	8	Martin Rowe[1]	rabbit anti-mouse-IgG	13.5	Dako
EBNA 2	PE2	Mouse monoclonal	8	Martin Rowe[1]	rabbit anti-mouse-IgG	13.5	Dako
EBNA 3A	EBNA 3A (F115P)	Sheep polyclonal	8	exalpha			
EBNA 3B	EBNA 3B (F120P)	Sheep polyclonal	8	exalpha			
Histone H3	acetyl-Histone H3 (06-599)	rabbit polyclonal	5	millipore			
Histone H4	acetyl-Histone H4 (06-598)	rabbit polyclonal	5	millipore			
Histone H3	Tri-methylated K27 (17-622)	rabbit polyclonal	5	millipore			
PU.1	PU.1 (H-135) X (sc-22805)	rabbit polyclonal	8	santa cruz			

[1]: Kindly provided by Prof. Martin Rowe: University of Birmingham

Appendix B

antibodies Western blotting							
protein	antibody	antibody species	antibody dilution	company/ generated by	secondary antibody	antibody dilution	company/ generated by
EBNA 3A	T2.78	mouse monoclonal	1/100	Martin Rowe[1]	anti-mouse-HRP	1/1000	cell signalling
EBNA 3B	EBNA 3B (F120P)	Sheep polyclonal	1/500	exalpha	Rabbit anti-goat-HRP	1/1000	sigma
EBNA 3C	A10	mouse monoclonal	1:1	Martin Rowe[1]	anti-mouse-HRP	1/1000	cell signalling
TBP	TFIID (TBP) 58C9 (sc-421)	mouse monoclonal	1/200	santa cruz	anti-mouse-HRP	1/1000	cell signalling

[1]: Kindly provided by Prof. Martin Rowe: University of Birmingham

Appendix C

primers Real time PCR primers for QPCR			
primer set	primer	sequence 5' to 3'	notes
viral genome			
LMP1			
(-)310	Fw (MW 311)	GGCCAAGTGCAACAGGAA	
	Rv (MW 310)	GCAGATTACACTGCCGCTTC	
(+)34	Fw (MW 140)	AGAGGAGGAGAAGGAGAGCAA	
	Rv (MW 139)	CCTGAGGATGGAACACGAC	
(+)574	Fw (MW 148)	TGAGCAGGATGAGGTCTAGGA	
	Rv (MW 147)	GGAGATTCTCTGGCGACTTG	
(+)1564	Fw (MW 150)	GTGTTGTGCAGAGGTCTGATG	
	Rv (MW 149)	CCTTCACACACCACACAGGT	
(+)2238	Fw (MW 144)	GGACACGCTCCTTCTTGG	
	Rv (MW 143)	ACTGGCTGGATTCTACGCTACT	
salmon genome			
insulin			
salmon insulin	Fw (MW424)	TGAGACAGACACCACCAAAGA	
	Rv (MW425)	CGAACCCAGCCTTCACAA	
Human genome			
ADAM28			
(-)587	Fw (MW506)	TCCTCATTGCCAGTTATGACA	
	Rv (MW507)	CTCACACAAGTGTACTCTCAGGAAG	
(-)87	Fw (MW508)	AGGACCACAGCTTCGAGGT	
	Rv (MW509)	CCTCCTCTCCAGTGAGACAGA	
(+)1523	Fw (MW615)	GTGATCTTGGCTCACTGCAA	
	Rv (MW616)	AATTAGCCGAGCGTGGTG	
(+)13274	Fw (MW613)	CAACCAGAGAACTGAGGCACT	
	Rv (MW614)	AGCAATGTGGCATGTGTGAT	
(+)19387	Fw (MW510)	ACAGGAGCATGCACTCTTCA	
	Rv (MW511)	GGCAATGTTCTGCTGCAA	
ADAMDEC1			
(-)1281	Fw (MW514)	AGAGTTCACCATTCCAGGATTC	
	Rv (MW515)	TCCAGAGATAGGCATCCTTCTTA	
(-)79	Fw (MW516)	GTAGATATGCACGCGACCAC	
	Rv (MW517)	TTGAGGACTCACATCTGGACA	
(+)228	Fw (MW518)	GTGGGATCTCCCAGCTACC	
	Rv (MW519)	AATGCTGGTGTGATGGTACG	
ADAM peak			
5'	Fw (MW544)	ACCATAAAGCCATGGGTCAG	
	Rv (MW545)	GGAGTACAGTTTGGAGCTCATATTT	
mid	Fw (MW540)	CTTCATGGCTACAGACTCTTGG	peak primer
	Rv (MW541)	CCTATGTCTCGCTTCCTGCT	peak primer

3'	Fw (MW542)	TCCGGTGACTAAAGTATATGTAGGG	
	Rv (MW543)	CTCCATGCTGGATTGATTACAA	
Bim (BCL2L11)			
1	Fw (MW870)	CAGAGGGAGGAGAGCTCAAA	
	Rv (MW871)	GAGTTTCTAAGCCGCTCTGG	
2	Fw (MW876)	CTGGTGAAGGGTCGTAGGTC	peak primer
	Rv (MW877)	CCCGATACTACGAGCAGGTC	peak primer
3	Fw (MW878)	AGGACAAGTGGCGAGGACT	
	Rv (MW879)	GCCAGTCACCTGGAGACAA	
CTBP2			
1	Fw (MW834)	TGGCTATGTCACAGCGATTC	
	Rv (MW835)	ACCGTGCCTGATGGAGTT	
2	Fw (MW836)	TTGCATATTTGGGATTTCAGTTT	peak primer
	Rv (MW837)	TTCTGTGTGAAACAGTTGTGGTT	peak primer
3	Fw (MW838)	CCAGGACAACGTCTGAGTGA	
	Rv (MW839)	CCATCGCCATGTTAACAGAA	
CXCL10			
(-)893	Fw (MW476)	ACAGTGTCTTGGAGCTGAACC	
	Rv (MW477)	GCACGCATAGAGACAGACCTT	
(-)177	Fw (MW474)	TCAAAGCAGGCCAGTCCTAT	
	Rv (MW475)	AACAGCAAGTTGTGCCTGATT	
(+)130	Fw (MW488)	AAGGAACATCAAAGGATACTTAATTTG	
	Rv (MW489)	TCCCTCTTATTTATAAGCATGCAGT	
(+)541	Fw (MW490)	TGATGTGATCCATTCCTCCTC	
	Rv (MW491)	AGAGAAGGTTAGCACAGTGTTCATT	
(+)1426	Fw (MW472)	ATGAATGCATAGCAGCAGGA	
	Rv (MW473)	CTGACTCTGGCTTCAGATTGG	
(+)1975	Fw (MW599)	GAATGCTCTTTACTTCATGGACTTC	
	Rv (MW600)	TCTACCTTCCTGTATGTGTTTGGA	
CXCL11			
(-)893	Fw (MW482)	CGTGGCATGGAGTTACTGAA	
	Rv (MW4830	AGCTCCTCCTGCCTTCTCTT	
(-)295	Fw (MW480)	CAGTCTTCCTGAATGAATGACAA	
	Rv (MW481)	CCAGATGGTAACCAGCCTTC	
(+)46	Fw (MW492)	AAGCTGAAGTAGCAGCAGCA	
	Rv (MW493)	GCACACAATATCACAGCCAAG	
(+)594	Fw (MW494)	CCAGTAACAGTTGTGTTAAGTGCTATT	
	Rv (MW495)	TAGCTGGTGGTGAACGTGAG	
(+)1660	Fw (MW478)	TTGATGCTTACAACTATTCTGTTGTG	
	Rv (MW479)	CAATGTCTCCACCGTAACCA	
(+)2147	Fw (MW595)	GCAATATCTTGGACACATCTGAA	
	Rv (MW596)	AAAGTGATTGCTAGGTATACATTTGCT	
CXCL11 upstream peak (our ChIP sequencing data)			
1	Fw (MW611)	GTGGTGGTAACTGGATGAGTGA	peak primer
	Rv (MW612)	ACACTATGTACATTACGTGCCACA	peak primer

2	Fw (MW609)	ACCAACGCTCATTTCCTGTT	
	Rv (MW610)	ACCTTTATTTCCTTCACTCCTCTTC	
3	Fw (MW607)	TGGTTGGCTGTAAAGGAAGAATG	peak primer (not MACS)
	Rv (MW608)	AATCAGCAGTGATTGATTCCTAGA	peak primer (not MACS)
CXCL10 downstream EBNA 3A binding site (Harth-Hertle et al 2013)			
R1	Fw (MW1237)	CAGGGACGGTAAGAGCCTTC	described as peak see ref
	Rv (MW1238)	AAATTCAAACAGGCCTGGAG	described as peak see ref
R2	Fw (MW1239)	TTAGCAAGGGTGGACGGTAG	described as peak see ref
	Rv (MW1240)	TCACAAGGCACTTCATCGTC	described as peak see ref
R3	Fw (MW1241)	GTGTTTGCTCAAGGCCCTAC	described as peak see ref
	Rv (MW1242)	TGCTTGCAGGGAAGGATATAAG	described as peak see ref
ITGA4			
(-)445	Fw (MW548)	GCTGTCTCTCTGGTTGCTGA	
	Rv (MW549)	AACGCAACACACCTGAACTG	
(+)77	Fw (MW550)	GGGCTGCAGAGGAAGTGT	
	Rv (MW551)	TGACAAAGACGTTATGGCTATTC	
(+)657	Fw (MW552)	CTGCGCCTCATCTCTTGG	
	Rv (MW553)	AACACTAAACGGCCACTACCC	
(+)1145	Fw (MW554)	GTGCGTTATGGTGCAAGGT	
	Rv (MW555)	CGCCACAGATGTGCAGTC	
ITGAL			
1	Fw MW820	CAGGCTGGAGTACAGTGGTG	
	Rv MW821	GAGGCTGAGGCAGGAGAAT	
2	Fw MW822	TGTAAAGGGTATCTCACTGTGGTT	peak primer A
	Rv MW823	ATAAATGGCCAACATGCACA	peak primer A
3	Fw MW824	AGGCTGGTCTCGAACTCTTG	
	Rv MW825	CGGTGGCTCACACCTCTAA	
4	Fw MW826	TGCACCTGTGGTTTCAGCTA	peak primer B
	Rv MW827	CGATCACAGCTCAATGCAAC	peak primer B
5	Fw MW828	ACCCAGCCTCCAATTCTTTAG	
	Rv MW829	TTTCTCTGGACCTTGAAAGATGT	
6	Fw MW830	TGCTTACACTTCCTCCCTGAA	peak primer C
	Rv MW831	TTTCTCACAGAGGCAACAGG	peak primer C
ITGB1			
(-)13540	Fw (MW627)	TGCACCTTGGCAGATTGA	
	Rv (MW628)	GGTTGAGAGCTGGTTGCTG	
(-)885	Fw (MW558)	CAGGATAGCAGCTTGCCAGT	
	Rv (MW559)	GCATGTCGATTCGATGATTG	
(-)220	Fw (MW560)	CTTCGCAGAGGAGGAAACTG	
	Rv (MW561)	GACCGGGACAAAGGAACC	
(+)1019	Fw (MW562)	CGCGTGTCATTTCTTAGGC	
	Rv (MW563)	TGATCATACCAATGAGTGCAAA	
PPIA			
PPIA	Fw (MW200)	GGGCCGAACGTGGTATAA	Spans PPIA TATA box
	Rv (MW201)	CCATGGCTAATAGTACACGGTTT	Spans PPIA TATA box

WEE1		
1	Fw (MW798)	CAAGCCATCTTCACACCTCA
	Rv (MW799)	TTGAGACCAGCCTAGACAACAA
2	Fw (MW800)	TGCTAAGTGGATGAGGCAGA
	Rv (MW801)	TGCACAGCTGCATTCCTC
3	Fw (MW802)	TTGCTGCAACACCAGCTC
	Rv (MW803)	CACCAGGAACACTGGAGACA
4	Fw (MW804)	TCCAGACAGCACCAGGAAG
	Rv (MW805)	CTGTGGGAAAGTCCTCCAGT
5	Fw (MW806)	TACAGGCGTGAGCCACTG
	Rv (MW807)	GTCTGCACCAAGTGACAAGC
6	Fw (MW808)	GGGACTCCTGGTGTTAGTCCT
	Rv (MW809)	CATAGAAGGAGGCCTGTCAAA
7	Fw (MW810)	ACCACTGCAGGAAGCGTTAT
	Rv (MW811)	GATTCGTGCACATCTGTTGAA
8	Fw (MW812)	ACCCGGTCCTAACTGGAGA
	Rv (MW813)	GAGCCTCGCTCCAGAGACT
9	Fw (MW814)	GGACATTGTATTTATTCAGGTCCA
	Rv (MW815)	GCTCCCTGAATCCTATCCACT
10	Fw (MW816)	CCCTGGGTTAGTCATGCAA
	Rv (MW817)	GACTCCCATGTGTGGTTGG

Appendix D

primers cDNA amplifying primers			
primer set	primer	sequence 5' to 3'	notes
Human genome			
CTBP2	Fw (MW840)	GAGTGATCGTGCGGATAGG	CTBP2 exon 5
	Rv (MW841)	GAGTCCGCTGTCTCTTCCAC	CTBP2 exon 6
S1	Fw (MW853)	TACAGGTCGTGGGCTGGTAT	CTBP2 S1 exon 1/2
	Rv (MW854)	CTTCAGGATGGGCATCTCC	CTBP2 S1 exon 2
S2	Fw (MW855)	ACCATGCTTGCACCAGAAC	CTBP2 S2 exon 1
	Rv (MW856)	TTCAGGATGGGCATCTCC	CTBP2 S2 exon 2
L1	Fw (MW857)	ATGGCCTACGTCGTCATACC	CTBP2 L1 exon 1
	Rv (MW858)	GGAATACACCTTCATTGGTTCAA	CTBP2 L1 exon 2
L2	Fw (MW859)	CGGGACGAGGGTTTCATC	CTBP2 L2 exon 1/2
	Rv (MW860)	CTATAGTTCACACGATGGGCTGT	CTBP2 L2 exon 2
L3	Fw (MW861)	CGGGACGAGATGGAGAAA	CTBP2 L3 exon1/2
	Rv (MW862)	GTACTCTGCAGCCAGGAGGA	CTBP2 L3 exon 2
L4	Fw (MW863)	CCAAGTTGGAATTAGCATCCTC	CTBP2 L4 exon 1
	Rv (MW864)	TAGTTCACACGATGGGCTGT	CTBP2 L4 exon 2
ITGAL	Fw (MW832)	CCAAGTCAAGCACATGTACCA	ITGAL exon 25
	Rv (MW833)	ACGCTCCACTGGTGTGTG	ITGAL exon 26
WEE1	Fw (MW818)	TGAAGAGGGCGATAGTCGTT	WEE1 exon 8
	Rv (MW819)	CACCAGCAGCACATACCACT	WEE1 exon 9