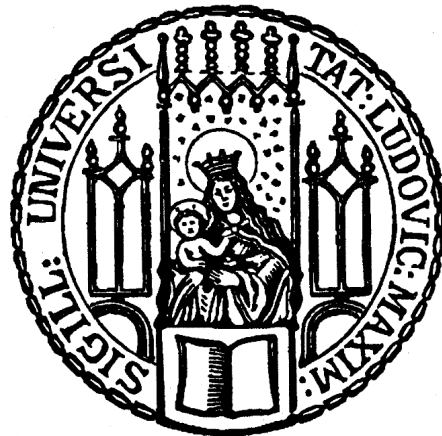


Dissertation zur Erlangung des Doktorgrades  
der Fakultät für Biologie  
der Ludwig-Maximilians-Universität München

# Time-course analysis of human primary B lymphocytes infected with EBV



Paulina Mrozek-Górska

Dissertation submitted on 20.03.2018

oral examination on 11.10.2018

Erstgutachter:	Prof. Dr. Dirk Eick
Zweitgutachter:	Prof. Dr. Angelika Böttger

## Erklärung

Hiermit erkläre ich an Eides statt, dass die vorliegende Arbeit mit dem Titel

„Time-course analysis of human primary B cells infected with EBV“

von mir selbständig und ohne unerlaubte Hilfsmittel angefertigt wurde, und ich mich dabei nur der ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe. Die Arbeit wurde weder in der jetzigen noch in einer abgewandelten Form einer anderen Prüfungskommission vorgelegt.

München, den 20.03.2018

---

Paulina Mrozek-Górska

# Table of Contents

<b>1. ABSTRACT</b>	<b>1</b>
<b>2. INTRODUCTION</b>	<b>2</b>
<b>2.1. EBV - ITS LIFE CYCLE AND THE REPROGRAMMING OF ITS HOST CELL</b>	<b>2</b>
2.1.1. EBV AND ITS INGENIOUS LIFE CYCLE	2
2.1.2. B CELL REPROGRAMING DURING EBV INFECTION	4
<b>2.2. EPIGENETIC CHANGES DURING VIRAL INFECTION</b>	<b>7</b>
2.2.1. HISTONE ACQUISITION AND THE HISTONE “CODE”	7
2.2.2. TRANSCRIPTOME REGULATION – FUNCTIONS OF THE CHROMATIN ARCHITECTURE AND LNCRNAs	11
2.2.3. THE EPIGENETIC MECHANISMS DURING $\gamma$ HERPESVIRUS INFECTION	14
<b>2.3. POTENTIAL METABOLIC CHANGES IN THE COURSE OF VIRAL INFECTION</b>	<b>16</b>
2.3.1. THE GLUCOSE METABOLISM	16
2.3.2. THE MITOCHONDRIAL ACTIVITY	18
2.3.3. VIRAL INFECTION AND METABOLISM	20
<b>2.4. SCOPE OF MY THESIS</b>	<b>21</b>
<b>3. RESULTS</b>	<b>23</b>
<b>3.1. MODEL OF EARLY EBV INFECTION – THE EXPERIMENTAL DESIGN</b>	<b>23</b>
<b>3.2. PHENOTYPIC CHANGES DURING THE FIRST EIGHT DAYS OF EBV INFECTION</b>	<b>25</b>
<b>3.3. REGULATION OF THE CELLULAR METABOLISM IN THE COURSE OF EBV INFECTION</b>	<b>27</b>
<b>3.4. HISTONES AND HISTONE CHAPERONES DYNAMICS IN THE EARLY PHASE OF EBV INFECTION</b>	<b>32</b>
<b>3.5. GLOBAL TRANSCRIPTIONAL REGULATION IN THE HOST CELL AFTER VIRAL INFECTION</b>	<b>34</b>
3.5.1. PREPARATION, DESIGN AND QUALITY CONTROL OF THE TIME-RESOLVED RNA-SEQ EXPERIMENT	34
3.5.2. DYNAMIC REGULATION OF SPECIFIC BIOLOGICAL PROCESSES IN THE COURSE OF EBV INFECTION	43
3.5.1. nCRNAs AND TARGET GENES INVOLVED IN CHROMATIN PROCESSING OR HOST CELL METABOLISM DURING EARLY EBV INFECTION	47
<b>3.6. DYNAMIC NUCLEOSOME ASSEMBLY DURING THE FIRST DAYS OF EBV INFECTION</b>	<b>55</b>
3.6.1. MNASE-SEQ EXPERIMENTS IDENTIFY MONONUCLEOSOMES ON VIRAL DNA VERY EARLY AFTER INFECTION	56
3.6.2. KINETICS OF NUCLEOSOME POSITIONING ON VIRAL AND CELLULAR DNAs	59
<b>4. DISCUSSION</b>	<b>67</b>
<b>4.1. SCOPE AND AIMS OF MY THESIS</b>	<b>67</b>
<b>4.2. PHENOTYPIC AND METABOLIC CHANGES DURING THE FIRST DAYS OF VIRAL INFECTION</b>	<b>68</b>
<b>4.3. GLOBAL CHANGES IN THE HOST TRANSCRIPTOME AS A CONSEQUENCE OF VIRAL INFECTION</b>	<b>70</b>
4.3.1. THE DYNAMICS OF THE CELLULAR PATHWAYS	71
4.3.2. BIOLOGICAL PROCESSES IDENTIFIED DURING B CELL TRANSFORMATION	74



<b>4.4. CHROMATIN ALTERATIONS IN THE COURSE OF VIRAL INFECTION - NUCLEOSOME OCCUPANCY AND POSITIONING</b>	<b>78</b>
4.4.1. VIRAL CHROMATIN ASSEMBLY	78
4.4.2. CELLULAR CHROMATIN REARRANGEMENTS AS A RESPONSE TO VIRAL INFECTION	82
<b>4.5. OPEN QUESTIONS AND OUTLOOK</b>	<b>83</b>
4.5.1. WHICH SPECIFIC METABOLIC PROCESSES ARE ALTERED DURING EBV INFECTION?	84
4.5.2. WHICH CELLULAR AND VIRAL GENES IDENTIFIED IN MY RNA-SEQ EXPERIMENTS ARE CRUCIAL FOR THE EARLY PHASE OF EBV INFECTION?	84
4.5.3. WHEN AND WHERE DO HISTONE MODIFICATIONS OCCUR ON VIRAL DNA DURING THE EARLY PHASE OF INFECTION?	85
<b>5. MATERIAL</b>	<b>86</b>
<hr/>	
<b>5.1. OLIGONUCLEOTIDES</b>	<b>86</b>
<b>5.2. PLASMIDS</b>	<b>86</b>
<b>5.3. VIRUS SUPERNATANTS</b>	<b>86</b>
<b>5.4. ANTIBODIES</b>	<b>87</b>
<b>5.5. EUKARYOTIC CELL LINES</b>	<b>87</b>
<b>5.6. CELL CULTURE MEDIA AND ADDITIVES</b>	<b>88</b>
<b>5.7. ENZYMES AND OTHER CHEMICALS</b>	<b>88</b>
<b>5.8. BUFFERS</b>	<b>90</b>
<b>5.9. COMMERCIAL KITS</b>	<b>91</b>
<b>5.10. SOFTWARE AND PACKAGES</b>	<b>92</b>
<b>5.11. DEVICES AND CONSUMABLES</b>	<b>94</b>
<b>6. METHODS</b>	<b>95</b>
<hr/>	
<b>6.1. EUKARYOTIC CELL CULTURE METHODS</b>	<b>95</b>
6.1.1. CELL CULTURE CONDITIONS	95
6.1.2. FREEZING AND THAWING OF EUKARYOTIC CELLS	95
6.1.3. EXTRACTION OF B CELLS FROM THE ADENOID TISSUE	95
6.1.4. COLLECTING VIRUS SUPERNATANTS FROM THE HEK293 2089 CELL LINE	96
6.1.5. COLLECTING VIRUS SUPERNATANT FROM THE B95-8 CELL LINE	96
6.1.6. DETERMINING THE VIRUS TITER	97
6.1.7. PURIFICATION OF B CELLS ON MACS COLUMNS	97
6.1.8. PRIMARY B CELLS INFECTION WITH EBV	98
<b>6.2. FLOW CYTOMETRY MEASUREMENTS</b>	<b>98</b>
6.2.1. FACS SORTING FOR PRIMARY B CELLS	98
6.2.2. STAINING WITH TMRE AND ANNEXIN V	99
6.2.3. GLUCOSE ANALOGUE UPTAKE ANALYSIS USING 2-NBDG	99
<b>6.3. NUCLEIC ACID METHODS</b>	<b>99</b>
6.3.1. GENOMIC DNA PURIFICATION FROM EUKARYOTIC CELLS	99
6.3.2. DNA PURIFICATION FROM AGAROSE GEL	100
<b>6.4. TIME-COURSE RNA-SEQ EXPERIMENTS</b>	<b>100</b>

6.4.1. SAMPLE COLLECTION	100
6.4.2. RNA ISOLATION, ERCC SPIKE-IN ADDITION	100
6.4.3. LIBRARY PREPARATION AND SEQUENCING	101
6.4.4. BIOINFORMATIC ANALYSIS	102
<b>6.5. MNASE-SEQ EXPERIMENT</b>	<b>105</b>
6.5.1. SAMPLE PREPARATION, MNASE DIGESTION	105
6.5.2. LIBRARY PREPARATION AND SEQUENCING	106
6.5.3. BIOINFORMATIC ANALYSIS	106
<b>6.6. PROTEIN ANALYSIS METHODS</b>	<b>109</b>
6.6.1. WHOLE CELL LYSATE PREPARATION	109
6.6.2. SODIUM DODECYL SULFATE POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)	109
6.6.3. WESTERN BLOT	109
<b><u>7. ABBREVIATIONS</u></b>	<b><u>110</u></b>
<b><u>8. LITERATURE</u></b>	<b><u>115</u></b>
<b><u>9. APPENDIX</u></b>	<b><u>128</u></b>
9.1. OUTLIER SAMPLE FROM RNA-SEQ DATA	128
9.2. QUALITY CONTROLS OF rRNA DEPLETED SEQUENCING LIBRARIES	129
9.3. GENES CONTRIBUTING TO PC1 AND PC2 IN THE POLYA-ENRICHED SEQUENCING LIBRARIES	130
9.4. TIME-COURSE EXPRESSION OF EBV GENES DURING THE PRE-LATENT PHASE OF INFECTION	131
9.5. CELL CYCLE PHASE PREDICTION FROM RNA-SEQ DATA	132
<b><u>10. CURRICULUM VITAE</u></b>	<b><u>133</u></b>
<b><u>11. ACKNOWLEDGMENT</u></b>	<b><u>134</u></b>

## 1. Abstract

*In vitro* infection of human quiescent B cells with Epstein-Barr Virus (EBV) activates them and leads to infinitely proliferating lymphoblastoid B cell lines (LCL), in which the virus establishes a stable latent infection. In this process, the virus triggers dramatic changes in the host cell biology to allow its long-term persistence and guarantee the survival of its host cell. During latency, EBV encoded proteins mimic activating physiological cellular pathways and processes promoting viral success. In my PhD thesis, I wished to investigate the very early steps of EBV infection focusing on changes in the host cell biology. I investigated the early phenotypes of the infected cells, which grow in cell volume and massively induce RNA and proteins synthesis, initially. Within the first two days, the cells show stable metabolic activity and do not proliferate. On day three post infection (p.i), the cells initiate DNA replication and divide followed by a phase of intense cellular proliferation. Starting at day four p.i., the cells show very high metabolic activities characterized by an increased uptake of glucose and enhanced mitochondrial activity. To monitor the obvious alterations in the biology of the EBV-infected B cells, I devised well-controlled and time-resolved RNA expression profiling (RNA-seq) experiments. The analyses of these experiments identified seven different clusters of genes with very specific gene expression patterns in EBV-infected cells compared with uninfected cells. My results document that the virus governs all important cellular processes including proliferation, cell metabolism, various epigenetic mechanisms, and ncRNAs biosynthesis in a very strictly time-controlled manner during infection.

The virus delivers its epigenetically naïve genomic DNA to the host cell upon infection, but in latently infected cells the viral DNA is extrachromosomally maintained and organized identical to cellular chromatin with nucleosomes including mostly repressive histone marks. In my PhD work I wanted to investigate the kinetics of nucleosome assembly on viral DNA as well as the specific positioning of nucleosomes. Within 24 hours after infection, EBV DNA had acquired nucleosomal structures, which accumulated and became more prevalent until day three p.i.. Nucleosome acquisition did not appear to be random, but I found pre-defined locations occupied with nucleosomes very early after infection. In my MNase-seq experiments nucleosome occupancy at certain cellular loci also changes dramatically indicating that dynamic alterations in nucleosome positioning might cause the profound changes in the transcriptome of the virally infected B cells.

## 2. Introduction

### 2.1. EBV - its life cycle and the reprogramming of its host cell

#### 2.1.1. EBV and its ingenious life cycle

Epstein-Barr virus (EBV) is a ubiquitous human pathogen and a member of the  $\gamma$  herpesviruses family. EBV has a very strong tropism towards human B cells probably because the major viral glycoprotein gp350/220 binds CD21 on the surface of B cells to facilitate virus attachment. EBV's genome is composed of a single double-stranded DNA molecule contained in an icosahedral nucleocapsid, which is surrounded by a tegument and contained in a membranous envelope. The viral DNA is approximately 170 kb length in size and codes for about 80 proteins (Speck et al., 2000; Young and Rickinson, 2004).

EBV was described as the first human oncovirus and is now classified as a WHO group I carcinogen, because EBV is causally associated with several, diverse human tumours (Henle, 1968; Esau, 2017; Pagano, 1999). EBV is known to be involved in a number of malignancies including Burkitt's lymphoma (BL), nasopharyngeal carcinoma (NPC), Hodgkin's Lymphoma (HL), gastric cancer, and certain T/NK cell lymphomas (Carbone et al., 2008). In immunocompromised patients, EBV can cause non-Hodgkin lymphomas such as post-transplant lymphoproliferative diseases (PTLD).

In healthy individuals, the virus persists as an asymptomatic, lifelong infection. This is because the virus establishes a latent infection in mostly resting, quiescent B cells, but it can also replicate its genome in concert with the cellular genome in proliferating cells. More than 90 % of the population is seropositive for EBV, but the infection mostly goes unnoticed. *In vitro*, the virus infects mature resting B lymphocytes, activates them, induces cell division and their infinite proliferation as lymphoblastoid cell lines (LCL).

EBV is a very successful pathogen. The virus has evolved an inventive life cycle that consists of three distinct phases of infection: the pre-latent, latent, and lytic phase. Upon infection, EBV's genome enters the cell as a linear, naked, and epigenetically naïve DNA molecule, which rapidly circularizes in the nucleus of the infected cell. During the initial pre-latent phase, certain latent genes are already expressed, but many viral lytic genes are also temporary transcribed. Certain lytic genes seem to have an essential role in the early survival

of the virus regulating important host cell processes that will support viral long-term persistence (Altmann and Hammerschmidt, 2005). The naked, histone-free DNA supports the permissive viral gene expression, but lytic viral genes need to become silenced to escape from the host immune responses as well as to establish the stable, latent phase. To support gene silencing the viral genome acquires a very high nucleosome occupancy with specific, mostly repressive histone marks and DNA CpG methylation (Woellmer and Hammerschmidt, 2013; Woellmer et al., 2012). These mechanisms lead to the establishment of the latent phase of EBV infection, when several copies of the viral genome are maintained in the cell as plasmid-like chromosomes, called episomes. During this stable latent state, the virus expresses only a small subset of viral genes that is characteristic for the distinct type of latency induced during EBV infection. In an EBV-infected individual, four types of latency programs are known: latency 0, I, II, and latency III. Latency 0 is found in quiescent B cells, when no viral proteins but only viral noncoding RNAs are expressed. The remaining latency programs share the expression of Epstein-Barr virus Nuclear Antigen 1 (*EBNA1*), but differ in the expression of latent membrane proteins (*LMPs*) and additional *EBNA* genes. The latency programs I and II are characteristic for certain malignancies such as NPC or BL, respectively. *In vitro* established LCLs show the latency III program, which includes the expression of six *EBNAs*, two *LMPs*, two *EBERs*, and up to 44 miRNAs. This discrete program is the consequence of the well-regulated silencing of the EBV genome during the latent phase of infection.

Eventually, latently infected cells can switch to the productive lytic phase of EBV after induction of the signalling cascades upon B cell receptor activation. The viral lytic phase requires the expression of the *BZLF1* gene (Countryman and Miller, 1985; Takada et al., 1986), which encodes a viral sequence-specific transcription factor. It preferentially binds to an AP-1-like sequence and to a second class of sequences containing CpG methylated motifs, commonly called BZLF1 responsive elements (ZREs) (Kalla et al., 2012). Together with the second viral factor BRLF1, BZLF1 induces the very rapid expression of early lytic genes, which are mostly involved in lytic EBV DNA replication from the origin of lytic replication (*oriLyt*) element. Finally, 48 hours after reactivation, all viral lytic genes are expressed, which are essential for the synthesis of structural proteins, packaging of the new EBV DNA into viral capsids, and the release of virus progeny.

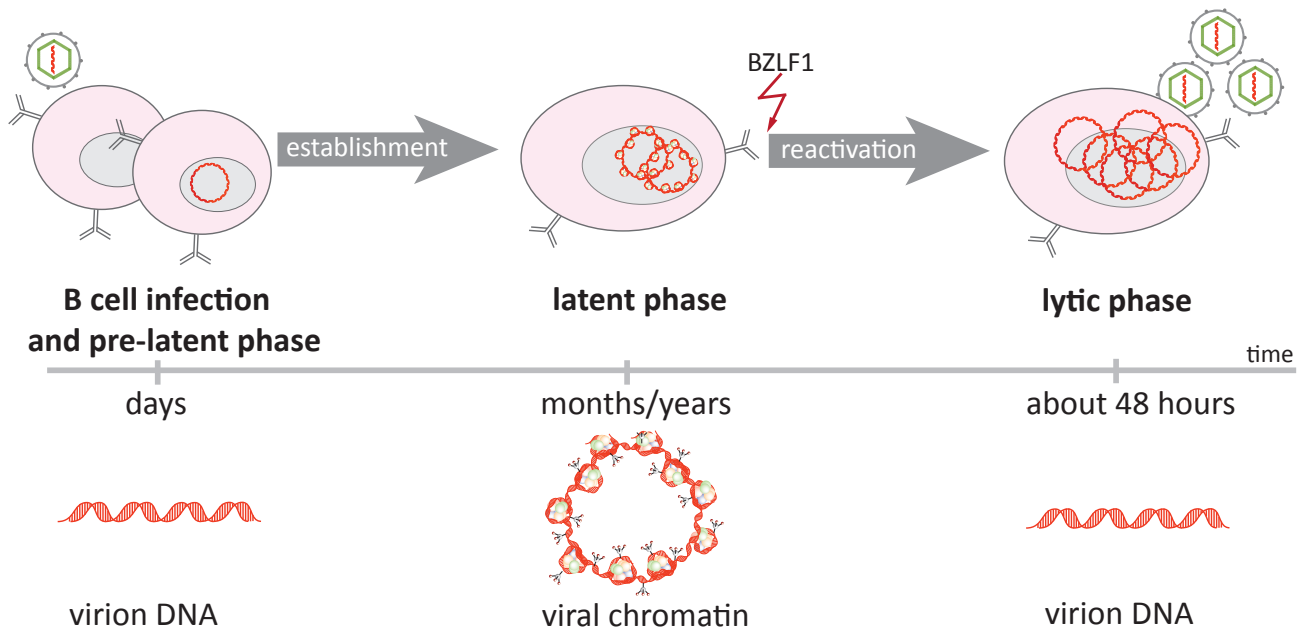


Fig. 1. **Schematic overview of EBV's life cycle.**

The scheme depicts the three main phases of EBV infection. B cells infected with EBV enter the pre-latent phase, which leads to establishment of the stable latent phase after about two weeks p.i. Upon antigen encounter and the subsequent expression of viral transactivator BZLF1 latently infected cells can enter the lytic phase. In the lytic phase, the viral genome replicates independently of the cell cycle phase, is encapsidated and released to allow infection of the neighbouring cells. The lower panel of the figure indicates the changes of EBV DNA. The virus delivers its linear, naked DNA to the cell, where it rapidly circularizes and acquires a fully chromatinized and methylated episome, which is a prerequisite for establishing the virus latent phase. During asynchronous viral DNA replication in the lytic phase the viral DNA loses all its epigenetic modifications and virus progeny is released that contain linear, unmethylated viral DNA free of histones.

### 2.1.2. B cell reprogramming during EBV infection

The success of EBV to establish and maintain a lifelong infection in B cells lies in the complex viral strategy to reprogram the cellular host. Initially, the virus has to find an efficient way to escape from the host immune responses that encompasses many defence lines such as the innate and adaptive immunity or induction of apoptosis. On the other hand, the virus has to manipulate host cell processes to create a favourable environment that allows the virus to persist in the infected cell and to replicate its genome in synchrony with the dividing cell, but also enables virus to effectively switch to the productive lytic phase if needed. Several reports describe viral factors and mechanisms that control immune responses against EBV or foster the process of B cell growth transformation and the regulation of host gene expression.

### *Immune responses against EBV*

EBV can persist in the host B cell despite strong and efficient immune responses against EBV antigens and viral T cell epitopes presented on the surface of the infected cell. Especially during the early phase of EBV infection, when viral latency is not fully established, the virus has to evade from the rapid cellular defence against the invading pathogen. Several viral genes are expressed immediately after infection to fulfil this task. The viral gene *BCRF1* mimics the host cytokine IL-10, which has an anti-inflammatory function as it controls the expression of the transporter proteins associated with antigen presentation (TAP) (Zeidler et al., 1997). Additionally, the *BNLF2a* gene reduces the recognition of the infected cell by CD8<sup>+</sup> T cells in the pre-latent phase of infection, because BNLF2a diminishes the surface presentation of EBV antigens (Jochum et al., 2012a). EBV also expresses two homologs of the cellular anti-apoptotic B cell lymphoma 2 protein (BCL-2): *BALF1* and *BHRF1*. The expression of these genes allows the virus to evade immediate apoptosis and to establish the subsequent latent phase (Altmann and Hammerschmidt, 2005). Yet another process, which is used by EBV during the early phase of EBV infection is RNA interference mediated by EBV's 44 miRNAs. Viral RNA interference reduces protein levels by neutralizing the encoding transcripts. Certain viral miRNAs have a role in inhibiting the recognition of EBV-infected cells by NK cells or limit pro-apoptotic interleukins of the host cell (Seto et al., 2010; Jochum et al., 2012b). In addition, several EBV miRNAs control the recognition of EBV infected B cells by CD8<sup>+</sup> and CD4<sup>+</sup> T cells. Directly or indirectly, viral miRNAs target HLA molecules, the pro-inflammatory IL-12 cytokine, lysosomal enzymes, and TAP transporters to efficiently prevent EBV antigen presentation (Albanese et al., 2017; Tagawa et al., 2016; Albanese et al., 2016). Clearly, EBV has evolved a variety of ingenious mechanisms to promote its escape from complex host cell immune responses.

### *B cell growth transformation and host gene regulation*

B cells infected with EBV are immediately activated and enter a highly proliferative state before the latent phase is established, eventually. Certain viral genes trigger and control this extraordinary process. A well-known example of one of the cell transforming viral factors is the latent membrane protein 1, LMP1. It has constitutive activating functions and mimics the activated ligand-bound cellular CD40 receptor, which stimulates and activates B cells. LMP1 is constitutively active and thus drives B cell transformation supporting their survival, activation,

and continuous proliferation (Dent et al., 1997). LMP1 was also found to have functions in different processes such as cytokine and chemokine release, immune modulation or gene regulation (Kieser and Sterz, 2015). The LMP2 viral protein is also involved in B cell signalling but in contrast to LMP1, it induces a basic level of activation similar to the tonic signalling of a non-engaged B cell receptor (BCR) (Mancao and Hammerschmidt, 2007).

EBNA proteins have additional and very important roles in activating B cells and regulating host genes. EBNA2 is a viral transcription factor, which together with its co-activator EBNA-LP are transcriptional activators of many cellular genes and are responsible for the establishment of a stable latent phase. Among the EBNA-LP and EBNA2 regulated genes are the proto-oncogene *c-MYC* and the cellular transcription factor PU.1, which promote the proliferation of EBV-infected genes. EBNA2 acts by binding to enhancer regions of its targets inducing subsequent alterations of the chromatin architecture (Zhao et al., 2011; Kempkes and Ling, 2015). EBV's EBNA3 proteins consist of EBNA3A, EBNA3B, and EBNA3C. EBNA3A and 3C are believed to function as oncoproteins by affecting cellular proliferation, survival, and immune responses. They are found to be essential for the regulation of cellular proliferation by targeting important tumour suppressor pathways constituents such as the pro-apoptotic BIM or cyclin-dependent kinase inhibitor (CDKI) p16<sup>INK4a</sup> (Paschos et al., 2012). EBNA3A and 3C are co-activators together with EBNA2 and interact with many cellular factors to facilitate gene expression and chromatin organization (Allday et al., 2015). The EBNA3B protein is dispensable for B cell transformation by EBV and functions as a tumour suppressor promoting for instance chemokines expression (White et al., 2012). EBNA1 protein is a viral factor that has a crucial function in maintaining the EBV episome, thus promoting cell survival during the stable latent phase (Frappier, 2012).

EBV uses a very complex network of viral strategies, which together reprogram the infected B cell and permit virus persistence in the host cell during latency. As mentioned in the previous paragraph, there are many reports about viral factors and their functions, which modulate cellular mechanisms, but our knowledge is still incomplete. In my PhD thesis, I focused on the phenotypes of B cell reprogramming at very early time-points after EBV infection, but I also investigated the dynamic changes of the cellular transcriptome induced by the virus. I aimed at verifying already known principles of EBV infection and at identifying new time-resolved processes that are essential for the establishment of EBV's latent phase. A key component of EBV's latent phase is the function of viral DNA, which includes acquisition of



nucleosomes and other chromatin constituents and occurs early in the pre-latent phase after infection.

## 2.2. Epigenetic changes during viral infection

Epigenetics is a very dynamic and fast developing field of research, which investigates the molecular basics of epigenetic control in different types of cells and epigenetic regulation of various diseases. Initially, the term “epigenetics” described apparent differences in phenotypes that do not stem from changes of the underlying DNA sequence, i.e. the genotype (Weinhold, 2006). Epigenetics implements heritable changes in gene expression that result from how cells “read” their genes. Several epigenetic mechanisms regulate the entire cellular transcriptome and they govern the identities of differentiated cells and tissues. Moreover, epigenetic mechanisms are critically involved in the dynamic response to changes in the cellular environment. The most important levels of epigenetic controls include nucleosome occupancy, positioning and their mobility, histone modifications, DNA CpG methylation, chromatin architecture, and lncRNAs. Epigenetics also plays an important role in the interaction between the host and its pathogens during different phases of infection with viruses, bacteria, or parasites (Allis and Jenuwein, 2016).

### 2.2.1. Histone acquisition and the histone “code”

#### *Nucleosomes and their acquisition to DNA*

Histone proteins are the basic “bricks” that together with double-stranded DNA form an octameric complex called the nucleosome. It consists of two copies each of the histone proteins H2A, H2B, H3, and H4. Nucleosomes package genomic DNA into a bead-like structure that consists of units of exactly 147 bp of DNA wrapped around the histone octamer in 1.65 turns of a left-handed superhelix (Luger et al., 1997; Thomas and Kornberg, 1975). Nucleosomes are formed in mostly regularly-spaced arrays, in which individual nucleosomes are separated by a DNA linker region, which is occupied by H1, a special linker histone. From a wider perspective, chromatin is a complex of DNA and nucleosomes, which together form chromatin fibres, which are coiled and condensed to form a chromosome. In principle, nucleosomes compact the DNA genome, but they also act as a physical barrier preventing the binding of cellular transcription factors to DNA and therefore regulate gene expression indirectly.

Nucleosome acquisition on DNA is a very tightly controlled process, which is essential for transcription, differentiation, DNA replication, and genome stability. The most important

cellular protein complexes involved in the mechanisms of nucleosome acquisition are histone chaperones. This class of enzymes deposits one tetramer of H3-H4 histone proteins onto DNA forming a tetrasome, which assembles almost one turn of the DNA helix. In the following step, two dimers of H2A-H2B histones are loaded on the tetrasome to complete the wrapping of the remaining DNA on the nucleosomal octamer (Richmond and Davey, 2003; Burgess and Zhang, 2010). Dedicated histone chaperones act on specific histone dimers, tetramers or even distinct histone variants, and they have two modes of action: replication-dependent or replication-independent assembly of histones. CAF-1 and ASF1 belong to the group of replication-dependent, S phase related histone chaperones. Canonical histones are newly synthesized in the S phase of the cell cycle, they are transported into the nucleus and bound by ASF1. This chaperone transfers the H3-H4 tetramers onto CAF-1, which interacts with proliferating cell nuclear antigen (PCNA) directly behind the replication forks and assembles histones onto newly synthesized DNA (Krude, 1999; Burgess and Zhang, 2013). NAP-1, another S phase chaperone, is presumably involved in the transport and deposition of H2A-H2B dimers on DNA during replication.

Histone loading can also take place independent of the cell cycle by a special class of replication-independent histone chaperones (Henikoff and Smith, 2015). They mostly act on variants of histone proteins such as CenH3 (centromere variant histone), H3.3, H2A.Z, H2A.X, H2A.Bbd (Barr body deficient histone), or MacroH2A. HIRA and the Daxx-ATRX complex are examples of chaperones specific for histone H3.3 variants and are associated with transcriptionally active sites but also silenced regions such as telomeres. Histone variants and their chaperones are involved in histone turnover, which is important for activity-dependent but cell cycle-independent gene expression during differentiation or in age-related signalling processes in neurons (Voon and Wong, 2016; Bano et al., 2017; Deaton et al., 2016). Histone variants presumably play a role in the stability of histone octamers and their strength associating with DNA (Jin and Felsenfeld, 2007; Henikoff, 2009).

### *Nucleosome dynamics*

The term “nucleosome dynamics” combines features of nucleosomes such as positioning and occupancy on the genome, post-translational modifications of histones and their mobility. Nucleosome dynamics results from a combination of different parameters and players that include DNA sequence preferences, the dynamic actions of chromatin remodelling

enzymes, cellular transcription factors (TFs), and the RNA polymerase II (PolII) transcription machinery. Transcriptionally active promoters or enhancer regions are often characterised by constitutive nucleosome free regions (NFR) or conditionally regulated nucleosome depleted regions (NDR); both of which ensure DNA accessibility for PolII, TFs, and the DNA replication machineries. Throughout the entire genome, nucleosome binding and spacing is dynamic and can affect different chromatin functions by nucleosomes sliding (their fuzziness), dissociation from DNA and histone nucleosome compositions, which can also alter their stability (Lai and Pugh, 2017; Struhl and Segal, 2013). Nucleosome occupancy is operationally defined as the average amount of nucleosomes on a specific region in the genome, i.e. it describes the probability that nucleosomes are present or absent at a certain location. In contrast, nucleosome positioning describes the probability that a nucleosome is placed at a given locus with discrete genomic coordinates; this term defines the exact point of nucleosome binding. Nucleosome positioning is a feature that is strongly dependent on the action of chromatin remodelers. Both, nucleosome occupancy and positioning allow to properly regulate the accessibility of DNA regions and thus influence gene expression. This chromatin plasticity is a mechanism, with which the genome of a cell can quickly respond to external stimuli or which can regulate a recovery from internal transcription or replication defects (Allis and Jenuwein, 2016).

In addition to nucleosome dynamics, histone proteins carry a vast variety of post translational modifications (PTMs), which play a role in processes involved in transcription, replication, or DNA damage response (DDR). Most post-translational modifications of histones appear on their highly basic N-terminal tails, which stick out from the core of the nucleosome and thus are presented to specific, histone-modifying enzymes. The best known and described histone PTMs are acetylation, methylation, and phosphorylation (Bannister and Kouzarides, 2011; Zentner and Henikoff, 2013). In addition, recent studies have revealed a number of novel modifications such as propionylation, butyrylation, crotonylation, 2-hydroxyisobutyrylation, malonylation, and succinylation (Kebede et al., 2015; Lawrence et al., 2016). The majority of these recently identified modification does not yet have assigned functions in chromatin biology and gene expression regulation; therefore, I will concentrate on histone modifications with well-known functions.

Histone acetylation was identified as the first histone modification by Phillips in 1963 (Phillips, 1963). Further studies found the correlation of histone acetylation and active gene transcription, which initiated the search for new histone PTMs and their roles in modulating

genes expression (Verdone et al., 2005; Allfrey et al., 1964). Acetylation is based on the enzymatic addition of an acetyl group ( $\text{COCH}_3$ ) derived from acetyl coenzyme A and promoted by histone acetyltransferases (HATs). Lysine acetylation neutralizes the positive charge of lysine residues, which weakens DNA-nucleosome interactions and increases chromatin accessibility for different cellular factors involved in transcription or replication. Histone acetylation is a very dynamic modification and can be removed by enzymes called histone deacetylases (HDACs), which were also found to have a role in cancer progression (Zentner and Henikoff, 2013; Verdone et al., 2005).

Histone methylation catalyses the addition of mono-, di- or tri- methyl groups from S-adenosyl-L-methionine to lysine or arginine residues, a chemical reaction, which is mediated by a group of enzymes called histone methyltransferases (HMTs) (Greer and Shi, 2012). Histone trimethylation at H3K4me3 and H3K36me3 are marks linked to gene activation in regulatory modules of a certain genomic context. H3K4me3 has an antirepressor-like function as this modification counteracts the repressive H3K27me3 modification introduced by Polycomb repressive complex (PRC) proteins. Yet another activation mark, H3K36me3, might have a role in DNA Mismatch Repair and DDR. H3K9me3 is a histone mark indicative of constitutively repressed chromatin and has an important role in gene silencing together with H3K27me3 introduced by PRC proteins. These methylation marks at the N-terminus of histones increase the affinity to certain protein modules such as HP1 chromodomain, which results in the destabilization of nucleosomes (Zentner and Henikoff, 2013; Li and Wang, 2017).

Similar to histone acetylation, histone phosphorylation interferes with the negatively charged phosphates on the DNA backbone, which reduces the interactions between DNA and nucleosome and thus increases DNA accessibility. Phosphorylation marks play an important role in many biological processes such as DDR or chromatin remodelling (Rossetto et al., 2012).

In more general terms, nucleosome occupancy and histone modifications define two major states of chromatin: euchromatin, which is characterized by decondensed and transcriptionally accessible regions, and heterochromatin with condensed regions and transcriptionally silenced or repressed genes (Allis and Jenuwein, 2016).

Specific combinatorial patterns of histone modifications create the so-called “histone code” (Peterson and Laniel, 2004). This additional information is used to recruit numbers of functionally different histone modifying factors, which are chromatin “readers” (e.g. chromodomain, PWWP domain proteins), chromatin “writers” (e.g. HAT, HMT), and chromatin

“erasers” (HDACs). The histone code provides an epigenetic marking that stores the instructions for regulating fundamental mechanisms such as transcription, replication, DNA repair and imprinting, which regulates cell differentiation and the cell fate (Jenuwein and Allis, 2001).

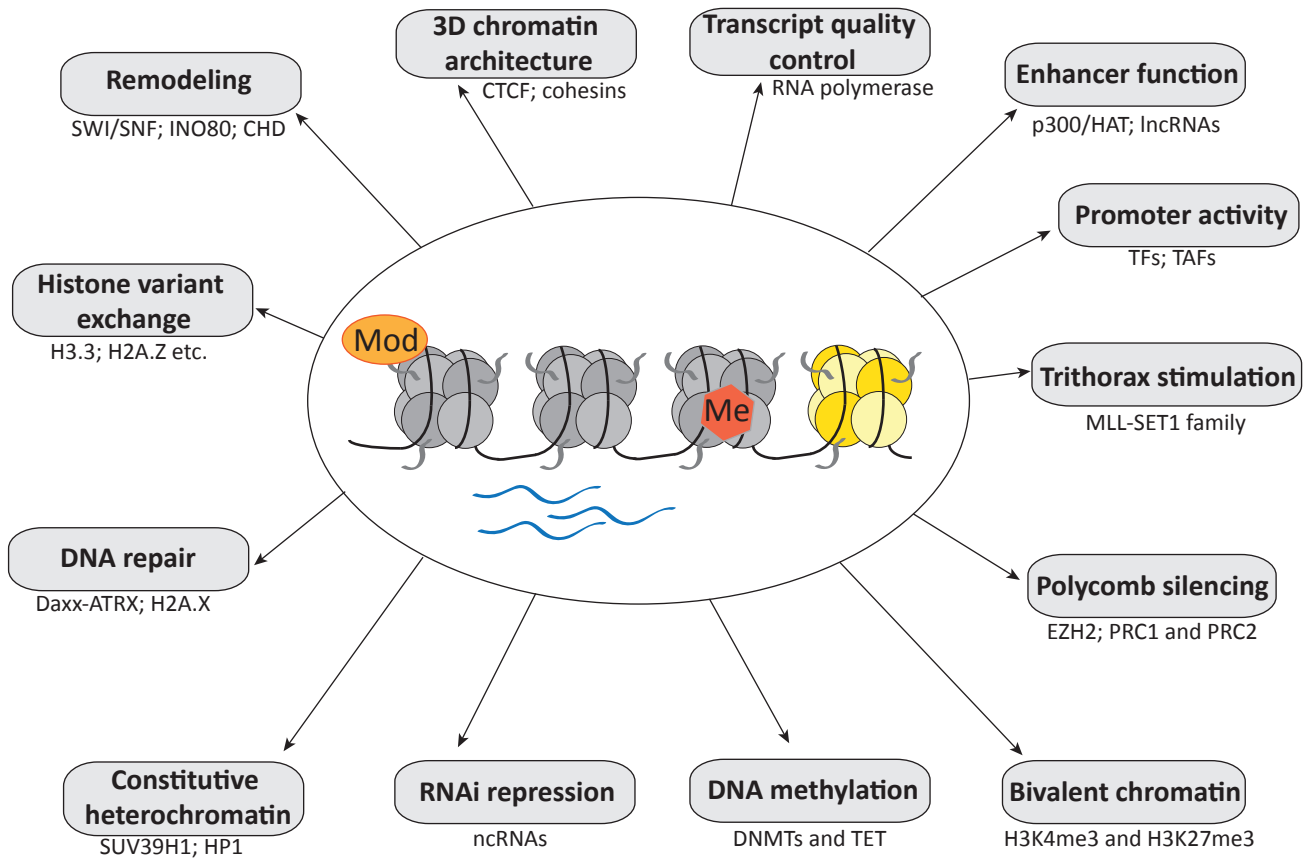


Fig. 2. **Examples of epigenetic modifications and their functions.**

The graph depicts four nucleosomes in its centre, which include different epigenetic features such as histone modifications (Mod), DNA methylation (Me), chromatin remodelling and histone turnover (yellow nucleosome), and lncRNAs (blue lines). These modifications are associated with their known and presumable functions on the chromatin landscape (grey boxes). Examples of factors involved in these mechanisms are provided (adapted from, Allis and Jenuwein, 2016).

### 2.2.2. Transcriptome regulation – functions of the chromatin architecture and lncRNAs

#### *Chromatin architecture*

Combined and complex interactions between DNA and chromatin control the transcription of all metazoan cells. The communication between promoter and enhancer

regions is another layer of complexity that affects gene expression and is mediated by the dynamic chromatin structure. Gene transcription initiates at promoter elements, where TFs commonly bind upstream of transcription start sites (TSS) and promote binding of the RNA Polymerase II (PolII) preinitiation complex again supported by TFs and their co-regulators. Efficient transcription elongation needs additional signals, which will help RNA Polymerase II to escape from pausing and to continue transcription. This complex regulation depends also on distal cis-regulatory elements such as enhancers or insulators (Hu and Tee, 2017; Kolovos et al., 2012).

Enhancers are important regulatory regions of the genome, which are able to interact with their promoters bridging long distances. Enhancers shape the gene expression program in a specific cell type and biological context. The program depends on the exact chromatin state, i.e. accessible local chromatin structure for TFs binding and higher order chromatin structure (chromatin architecture). The latter facilitates physical interactions between enhancers and their cognate promoters. Epigenetic marks define different types of enhancers: active enhancers are decorated with H3K27ac and H3K4me1, while silent enhancers are marked with H3K27me3, the typical repression mark. Poised enhancers are characterized by both types of histone modifications, repressive and activating marks such as H3K27me3 and H3K4me1, respectively (Shlyueva et al., 2014). Gene expression during development is often characterized by poised enhancers that modulate gene expression in embryonic stem cells and at various stages of differentiation.

Chromatin accessibility is a very important feature for enhancer mediated regulation and thus enhancers recruit several chromatin remodelers and transcription factors, in particular pioneer transcription factors (Calo and Wysocka, 2013). Together with these properties, enhancers also promote the three-dimensional shaping (looping) of chromatin to make physical contacts with their cognate promoter regions. Specific interactions are facilitated by compartmentalizing and insulating the genome into regulatory domains called “topologically associated domains” (TADs). TADs are able to arrange and bring together specific activating regions including promoter-enhancer loops via intra-TAD chromatin interactions preventing interactions between different TADs (inter-TADs) and separating these interaction domains. TAD borders are bound by architectural proteins such as CTCF and cohesins (Hu and Tee, 2017). Within TADs enhancers can also form groups of enhancers (super enhancers), which can massively amplify the expression of genes depending on the activating signal. On the contrary, a disruption or changes in certain enhancers are known to cause disease such as

sporadic Parkinson's disease or pancreatic agenesis, an extremely rare anomaly (Dekker et al., 2013; Hu and Tee, 2017).

### *lncRNAs and epigenetic regulation*

During the last ten years, the interest in ncRNAs and their functions in cellular biology has increased dramatically. lncRNA are defined as transcripts that do not encode proteins and are longer than 200 nt. Initially, lncRNAs were described as junk or artefactual RNAs, but they have been recognized to provide a new level of complexity in the regulation of gene expression. lncRNAs were found to act in cis targeting the transcripts of genes from which they originate. lncRNAs can also work in trans on transcripts, which are unrelated. By now, lncRNAs are described to be cell specific and to have roles in cell growth, apoptosis, development, and cell differentiation (Rinn and Chang, 2012). The molecular mechanisms that support lncRNAs in epigenetic regulation of the gene expression are diverse. For example, lncRNAs can efficiently interact with chromatin modifiers, serve as guide RNAs, locate to complementary region and change the local chromatin structure. Xist was described as the first functional lncRNA, which is essential for the inactivation of the X chromosome and gene dosage compensation. Silencing of the X chromosome is promoted by binding of Xist to the PRC2 complex via a repetitive motif in 5' end of the lncRNA. Xist was also found to directly bind to SHARP, a transcriptional repressor, and SMRT, which activates HDAC3 and mediates histone deacetylation of the X chromosome (McHugh et al., 2015).

Interactions of lncRNAs and proteins can guide the complexes to specific chromatin regions (as in case of Xist) or act as decoy for chromatin modifying enzymes preventing repressive modifications of a specific genomic region. lncRNAs can provide a scaffold to tether more than one protein complex and synchronize their activity. A prominent example is HOTAIR lncRNA, which guides the PRC2 complex to *HOX* genes and subsequently interacts with the histone demethylase LSD1, which removes activation mark from this locus (Marchese and Huarte, 2014; Flynn and Chang, 2014). Besides chromatin modifications, the lncRNA SRA is able to bind to and enhance the insulator function of CTCF, which together with cohesins are involved in chromatin looping and long range interactions (Yao et al., 2010). The known mechanisms of lncRNA molecules and their functions are still incomplete and gene regulation mediated by non-coding transcripts might be even more diverse and complex.

### 2.2.3. The epigenetic mechanisms during $\gamma$ herpesvirus infection

As discussed in the previous chapter, epigenetics has an important role in regulating cellular mechanisms including transcription. Therefore, it is not surprising that pathogens and in particular viruses use this layer of regulation to promote their infection and/or alter cellular processes. Viruses can indirectly control epigenetic regulation of the infected cell by modulating host cell signalling or transcription of chromatin constituents with enzymatic functions. Viruses can also directly impact on epigenetic mechanisms including histone modifications, DNA methylation, and chromatin structure. A recent report suggested that viruses are able to control each and every step of the virus life cycle by epigenetic processes (Milavetz and Balakrishnan, 2015). As different viruses clearly have diverse life styles that necessitate distinct mechanisms, I would like to concentrate here on known principles of epigenetic regulation in the group of  $\gamma$  herpesviruses focusing on KSHV and EBV.

The first step of infection is the entry of viral DNA into the nucleus of the host cell. Herpesviruses deliver their genomes to the nucleus as linear, epigenetically naïve DNAs free of CpG methylations and lack of histones. The viral genome circularizes immediately, which can be a very important step to prevent an antiviral DDR response that might emerge upon detection of free DNA ends in the infected cell. During infection with certain  $\alpha$  herpesviruses, the viral DNA rapidly acquires the histone variant H3.3, which seems to have an important role in the ensuing lytic phase of infection (Oh et al., 2012). The process of chromatinization of viral DNA of  $\gamma$  herpesviruses might be different since they commonly establish a stable latent phase upon infection and persist in the host cell. Tegument proteins, which are well-conserved among herpesviruses, might have a functional role in this process. Certain tegument proteins bind to cellular factors in promyelocytic leukemia nuclear bodies (PML NB), which contain chromatin factors such as Daxx, ATRX, or SP100. It was described that EBV's tegument protein BHRF1 can interact with Daxx and prevent its binding to ATRX. These two cellular proteins are known as the histone variant H3.3 chaperone complex. Thus, inhibiting their binding interferes with the loading of H3.3 histone onto viral DNA (Tsai et al., 2014; Huang et al., 2016). This mechanism was interpreted to prevent the silencing of latent genes, which the histone variant H3.3 might induce, but this report does not explain how and when the nucleosomes are actually assembled on viral DNA. The exact timing and the mechanisms of the chromatinization of EBV's genomic DNA remain to be explored.



The most critical step in  $\gamma$  herpesvirus infection is the establishment of a stable latent phase, which allows the pathogen to persist in the host cell long-term. The latent state has to be strictly controlled, because only a small and distinct group of viral genes is expressed, while all lytic viral genes must be repressed to escape from immune surveillance. This state of infection was investigated in more detail in both viruses EBV and KSHV. Mechanisms that tightly control epigenetic gene regulation during this phase encompass DNA methylation, nucleosome occupancy, histone modifications, and chromatin architecture. During latency, the EBV genome is fully methylated at CpG sites, densely occupied by nucleosomes with the histone marks H3K27me3 and, at a lower level, H3K9me3. During reactivation of the lytic phase the repressive modification H3K27me3 seems to be removed as compared with H3K9me3, which remains unchanged (Woellmer et al., 2012). KSHV, similar to EBV, acquires CpG methylation, but the post-translational modifications of histones differ during the latent phase of infection. Latent KSHV genes, which are expressed in this phase acquire active histone marks such as H3K9ac or H3K4me3; late lytic genes, which have to be silenced are associated with H3K9me3 and H3K27me3 (Toth et al., 2013). Viral genes, which are important for lytic reactivation and escape from latency contain bivalent epigenetic marks (H3K4me3 plus H3K27me3). This strategy, which differs from EBV, allows KSHV to enter the lytic phase in a very short time, but this strategy can promote a certain level of leakiness of viral genes and provoke antiviral immune responses *in vivo*.

Recently, the three-dimensional structure of the viral episome was causally linked with viral gene regulation during latency. Proteins involved in chromatin architecture such as CTCF and cohesins were also found on the viral genome. CTCF might play also a role in disrupting nucleosome positioning and preventing the spread of specific histone marks (Tempera et al., 2010), but CTCF also flanks TADs and might contribute to the three-dimensional structure of EBV DNA. During latency, chromatin loops between transcriptional regulatory elements were found in EBV as well as KSHV. Interactions between the oriP enhancer element with the latent Cp and LMP1 promoters enhance the expression of these EBV genes (Arvey et al., 2012). In the KSHV genome, CTCF-mediated loops were identified between *LANA* and the 3' end of the *K12* gene and between lytic genes encoding the transcripts of ORF50 and ORF45. These loops presumably repress lytic gene expression (Kang et al., 2011).

Escape from latency and reactivation of the productive, lytic phase allows herpesviruses to replicate and synthesize viral progeny to infect neighbouring cells. This process is very fast and has to be extremely efficient. All repressive modifications have to be erased or replaced by

activation marks, but also the entire viral chromatin has to be rearranged to promote rapid viral lytic gene expression. These mechanisms involve global epigenetic regulators such as HDACs, histone demethylases, or chromatin remodelers (Lieberman, 2013). Clearly, there are several studies that describe various epigenetic mechanisms employed by  $\gamma$  herpesviruses, but there are many crucial steps such as histone acquisition and *de novo* nucleosome formation that are largely unknown.

### **2.3. Potential metabolic changes in the course of viral infection**

Metabolic pathways play very important roles during all processes in the cell, because they provide a constant source of energy and basic components used for nucleic acid biosynthesis, protein synthesis or yield precursors for other specific biochemical processes. Cells adapt to various environmental stresses by modulating cellular metabolic pathways, which support the physiological development and differentiation of the cell, or result from responses to pathological situations. There are many reports that describe changes in the metabolism of glucose or energy consumption during the development of cancer cells or the activation and differentiation of T cells. There is an increasing interest to understand these cellular processes and their alterations in more detail during acute or latent viral infections.

#### **2.3.1. The glucose metabolism**

Glucose is the ubiquitous fuel in the biology of many organisms. It is used as a basic source of energy for cellular respiration. This simple carbohydrate is catabolized in a process called glycolysis. Briefly, this pathway converts one molecule of glucose, which is taken up by a transporter mediated pathway and initially converted to glucose-6-phosphate (G6P). It further breaks down into two molecules of pyruvic acid, which is metabolised to finally release the energy in the form of ATP and reduced NADH molecules. The conjugate base of pyruvic acid is pyruvate, which is a key intermediate in several metabolic pathways. The most important pyruvate pathway is the import into mitochondria, the conversion to acetyl-CoA, and entry into the tricarboxylic acid (TCA) cycle. As a product of this reaction cascade, NADH and FADH<sub>2</sub> are produced, which can be immediately used for ATP production in the electron transport chain reaction (ETC). During glycolysis, the TCA yields several precursors, which can be used in diverse biosynthesis pathways such as serine biosynthesis, the pentose phosphate pathway (PPP), or fatty acid synthesis (FAS) (Berg et al., 2002). The alternative pathway for pyruvate is its reduction to lactate in the cytosol mediated by lactate dehydrogenase (LDH) especially

under limited oxygen conditions. During this process, NAD<sup>+</sup> can be produced independently from mitochondria, which is a very fast but less efficient mechanism. This process can be also used under normal oxygen (aerobic) conditions, called the Warburg effect, which was first identified in cancer cells (DeBerardinis et al., 2008; Warburg, 1925).

Interestingly, the modulation of the glucose metabolism pathways plays a decisive role during the activation of naïve T cells upon antigen encounter and activation of the T cell receptor. Naïve T cells are in the quiescent state, where they do not undergo clonal divisions, do not secrete significant amounts of cytokines, and hence have limited requirements for inducing catabolic and anabolic pathways. As a consequence, they use available nutrients and produce energy using standard glycolysis supported by mitochondrial respiration (Finlay, 2012; Almeida et al., 2016). T cell activation mediated by antigen recognition induces their intense proliferation and their differentiation to T effector cells. T cell receptor (TCR) stimulation significantly increases the uptake of glucose and glycolysis in these cells. Additionally, glycolysis related pathways are also activated such as the pentose phosphate and the hexosamine pathways, which are important for the biogenesis of nucleic acids, NADPH, and UDP-N-acetylglucosamine (UDP-GlcNac) (Fig. 6). The latter is essential for the post-translational glycosylation of proteins and is involved in intracellular signalling. Activated T cells make use of both aerobic and anaerobic glycolysis, which is not fully understood, but supposedly contributes to the synthesis of different biomolecules. Another regulated mechanism is fatty acid synthesis (FAS) providing the cell with phospholipids, which are used during T cell proliferation and cell divisions (Fig. 3) (Almeida et al., 2016; Lunt and Vander Heiden, 2011; Buck et al., 2015).

There are extensive studies on the metabolic pathways induced upon T cells activation and differentiation, but details about B cells and their bioenergetic metabolism are incomplete. It was shown that BCR crosslinking rapidly increases glucose uptake, aerobic glycolysis, and upregulates specific glycolytic enzymes. Regulation of glucose utilization is driven by PI3 kinase signalling and *c-Myc* dependent transcription of glycolytic genes, which together are involved in B cell growth (Doughty et al., 2006). Additionally, B cell activation mediated by IL-4 and toll-like receptor (TLR) also leads to an induction of similar metabolic changes as described for BCR cross-linking (Dufort et al., 2007).

Cancer cells display an abnormally high proliferation rate and an increased cell volume, which stem from alterations in cancer cell metabolism supporting the cell with essential nutrients and biochemical energy. Cancer cells change all metabolic programs and show an

increase in glucose uptake, enhanced glutaminolysis or fatty acid synthesis. Characteristic for many malignant cells is the switch from glycolysis to the Warburg effect, which produces less energy but is faster than mitochondrial respiration. This process might be also beneficial to cancer cells, because it provides intermediates for many biosynthetic pathways such as amino acid and nucleic acid precursors. The Warburg effect is supposedly triggered by the *c-Myc* oncogene, which also drives cellular proliferation (Fadaka et al., 2017). The described processes illustrate that glucose metabolism and its regulation are crucial for cell survival and determination of a cell fate. Thus, glucose metabolism is also likely to play an important role during infection with EBV, which is able to activate and to transform the host cell.

### 2.3.2. The mitochondrial activity

Mitochondria, which are essential organelles in all metazoan cells, contain a specific double membrane (outer OMM- and inner mitochondrial membrane – IMM). The inner membrane of the mitochondrion is more complex and comprises a number of enzymes and proteins that are involved in carbohydrate and fat processing to finally release energy via the ATP synthetase complex. This membrane is surrounded by the mitochondrial matrix, where the TCA cycle generates electrons that can be transported to the cascade of IMM protein complexes. This process is called electron transport chain (ETC), which produces water as the last acceptor of electrons and energy in the form of ATP. Alternatively, this process is also called oxidative phosphorylation (OXPHOS) (Kühlbrandt, 2015). The number of mitochondria in a cell differs and depends on the cell state, because mitochondria biosynthesis is stimulated by a greater energy expenditure of the cell. Mitochondria also play an important role in maintaining the proper concentration of calcium ions within cellular compartments, which is crucial for the efficient communication with the cytosol.

Recently, it was found that T cells activation triggers an increase in mitochondrial respiration (Sena et al., 2013; Liu and Ho, 2017). Clearly, activated T cells need more energy and molecular precursors during activation and proliferation, additionally the generation of reactive oxygen species (ROS) also play a role during this activation. ROS are by-products of mitochondrial respiration and are essential for the optimal induction of the transcription factor NFAT (nuclear factor of activated T cells) as well as the release of cytokines (Sena et al., 2013; Weinberg et al., 2015). Another important mitochondrial function is calcium flux, which is supportive during naïve T cells stimulation. Upon TCR engagement the mitochondria increase the uptake of cytosolic calcium to stabilize its concentration. Thus, mitochondria act

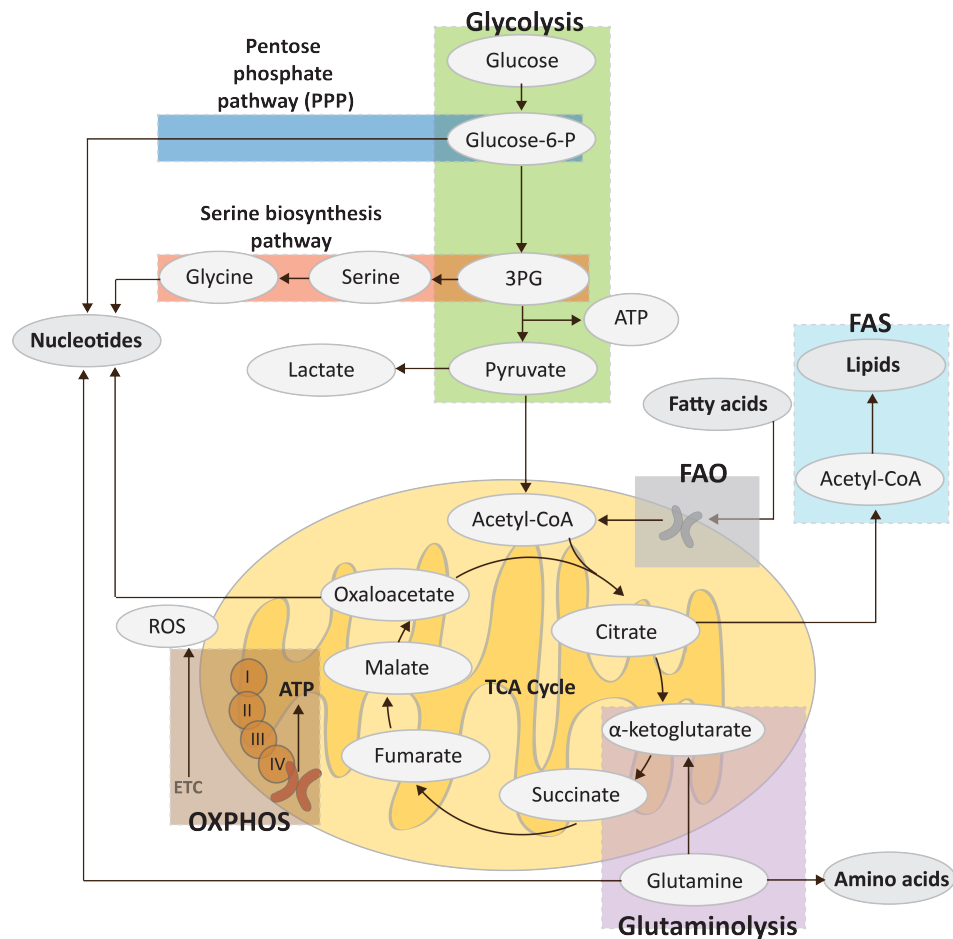


Fig. 3. **The scheme of metabolic pathways during T cell activation** (adapted from Buck, O'Sullivan and Pearce, 2015).

The glycolysis pathway in green, intersects with other biosynthesis pathways such as the pentose phosphate pathway (PPP, blue) and serine biosynthesis (red). Pyruvate, the end product of the catabolic glucose breakdown, can be used directly for lactate synthesis or is transported to mitochondria (in yellow) where it enters the TCA cycle followed by OXPHOS production of ATP (brown). Other substrates can be also metabolized via the TCA cycle such as glutamine (purple) or fatty acid synthesis (FAS, light blue), which yields acetyl-CoA via fatty acid  $\beta$ -oxidation (FAO, grey). All these pathways are tightly connected providing the efficient cellular metabolic machinery.

as a local buffering system. It supports the cell to sustain the continuous calcium influx and express calcineurin, a protein that initiates the migration of NFAT from the cytoplasm to the nucleus. Mitochondria can also actively support the import of calcium via purinergic receptors that work as ATP-gated calcium channels. This will lead to the increase of calcium uptake to the cell and therefore promote the activation of the T cell.

There are also reports that link the highly-synchronized program of mitochondria biogenesis and proteome remodelling, which can generate a specific type of mitochondria (Ron-Harel et al., 2016). The distinct mitochondria population is supporting the one-carbon (1C) metabolism, a pathway responsible for the synthesis of precursors for purine, NADPH and glutathione. This mechanism is actively responsible for cell growth and proliferation during T

cell activation. Cycling of NADH and NAD<sup>+</sup> might also be important in downstream pathways, since NAD<sup>+</sup> is a necessary co-factor in many enzymatic reactions (Almeida et al., 2016).

Recent reports claim that mitochondrial activation has a distinct function during B cell activation to allow the cells to accommodate the phenotypic and environmental changes encountered over their lifespan (Sandoval et al., 2017). Class-switch recombination of activated B cells correlates with an increase of the mitochondria mass and their membrane potential, while the differentiation to plasma cells is in line with a decrease in mitochondria biomass. These changes result from increased ROS levels, which act as additional signalling molecules and suppress the synthesis of haeme, promoting plasma cell differentiation (Jang et al., 2015; Sandoval et al., 2017). The increased mitochondrial activity in activated B cells is a consequence of an elevated glycolysis and OXPHOS activity indicating that mitochondrial homeostasis is crucial during B cell activation and differentiation.

### **2.3.3. Viral infection and metabolism**

Viruses rely on the metabolic machinery of their host cells, because even large viruses, e.g. herpes and poxviruses lack the capacity to encode many enzymes involved in metabolic processes in the infected cell. Recently, several reports highlighted the importance of studying and describing the impact of viral infection on the host cell metabolism (Freyberg and Harvill, 2017; Miyake-Stoner and O'Shea, 2014; Sanchez and Lagunoff, 2015). Many viruses induce glycolysis, fatty acid synthesis and glutaminolysis, which are clearly universal mechanisms of the cell, but virally induced metabolic processes are also cell type specific as viruses have evolved to adapt to the very specialized and differentiated host cells. It is common knowledge that closely related viruses can promote different metabolic alterations in the infected cells, but very different viruses can also cause similar metabolic perturbations. It is very likely that individual viruses follow unique strategies to promote their successful infection and *de novo* synthesis (Sanchez and Lagunoff, 2015).

Enveloped viruses, which turn the infected cell into a virus factory, regulate cellular lipogenesis required for the biosynthesis of lipid membranes of the virus. Human cytomegalovirus (HCMV) requires also aerobic glycolysis, which leads to lactate production, glutaminolysis as well as fatty acid synthesis (FAS) that is supported by metabolites derived from the TCA cycle. The expression of early viral genes promotes these metabolic changes. An increased glucose uptake is necessary to support the replication of the HCMV genome and is mediated by a specific glucose transporter (GLUT4) (Yu et al., 2011a). In contrast, herpes

simplex virus 1 (HSV-1) does not induce aerobic glycolysis, but uses the TCA cycle to promote pyrimidine synthesis rather than FAS (Vastag et al., 2011). Finally, vaccinia virus uses glutaminolysis to support its synthesis, thus the glucose metabolism is not critical for this virus and its production (Fontaine et al., 2014).

There are very few reports that focus on latent viral infection and the consequences of metabolic alterations. One example is the latent infection of endothelial cells with KSHV (Lagunoff, 2016; Sanchez and Lagunoff, 2015). In this report changes in three major pathways of central carbon metabolism, glycolysis, fatty acid synthesis, and glutaminolysis were described similar to metabolic alterations found in cancer cells. KSHV induces the production and release of lactate at the expense of the OXPHOS pathway, but this virus can also induce many anabolic pathways. Most importantly, inhibition of each of the three major pathways of carbon metabolism leads to the induction of cell death in infected cells compared with uninfected cells indicating that metabolic perturbations are critical for the survival of latently infected cells. Interestingly, it was also found that KSHV's latent miRNAs are involved in reducing OXPHOS and activation of aerobic glycolysis (Yogev et al., 2014). In EBV-infected Nasopharyngeal carcinoma (NPC) cell lines, high levels of glycolysis were reported to result from LMP1 expression. LMP1 induces the hexokinase 2 (HK2) enzyme and its knock-down in these cells leads to cell death confirming that glycolysis is critical for the survival of the infected cells (Xiao et al., 2014).

The metabolic state of the cell has an important role in each of the many cellular processes with a strong impact on cell fate, differentiation and response to environmental factors. Viruses are able to modulate these mechanisms to their advantage providing optimal conditions for their replication, spreading, or persistence in the infected host cell. We know little about EBV-induced metabolic changes in newly infected B lymphocytes, when the cells undergo massive phenotypic changes. Thus, it was interesting to study the metabolic parameters in this phase of EBV's life cycle.

## **2.4. Scope of my thesis**

Viral infection has a strong impact on the biology of the host cell. The pathogen induces changes that will support its persistence in the host cell, its genome replication, and prepare for the productive lytic phase reactivation. It is clear that the virus has to evade from the immune responses of the infected cell and organism as EBV does using its viral proteins and miRNAs, which can support cell survival, diminish the viral antigen presentation, or antiviral

processes (Jochum et al., 2012a; Albanese et al., 2016; Tagawa et al., 2016). Recently, several reports identified changes in the cell metabolism or epigenetic processes upon viral infection (Milavetz and Balakrishnan, 2015; Sanchez and Lagunoff, 2015). The knowledge about changes in the host cell during the early steps of EBV's infection is still poor and incomplete. The main aim of my PhD thesis was to examine the kinetics of alterations in the host cell biology that are triggered by EBV during the pre-latent phase of infection.

First, I focused on the phenotypes of the uninfected primary B cells and cells infected with EBV at different days post infection looking at the cell size and cell content, but I also considered already existing data from our lab regarding cell cycle distribution during infection. Since the cells seemed to change dramatically upon EBV infection, I wanted to elucidate the mechanisms, which are responsible for this effect. Second, I investigated the basic metabolic activities in EBV infected B cells and compared these phenotypes with described processes and documented models such as T cells activation or cancer cells. I studied the kinetics of glucose uptake and mitochondrial activity to decipher when exactly the virus redirects cellular metabolic mechanisms. Third, I designed and performed a comprehensive time-course transcriptome analysis (RNA-seq) to gain a broader overview of cellular processes regulated at different days after EBV infection. I found seven specific clusters of genes, which comprise a plethora of cellular mechanisms altered by the infection with EBV. The expression kinetics of these clusters suggested that cellular genes are very strictly controlled in time to support the viral pre-latent phase. The virus not only modulates its host cell, but also explores basic cellular mechanisms such as epigenetics to establish and to maintain latency and to efficiently reactivate during the productive lytic phase of infection (Woellmer et al., 2012; Woellmer and Hammerschmidt, 2013). Fourth, I investigated the kinetics of nucleosome assembly on EBV DNA and their exact positioning at viral loci using an MNase-seq approach. I wanted to define the exact timing of nucleosome acquisition and viral genome chromatinization in the course of the pre-latent phase of infection with EBV.



## 3. Results

### 3.1. Model of early EBV infection – the experimental design

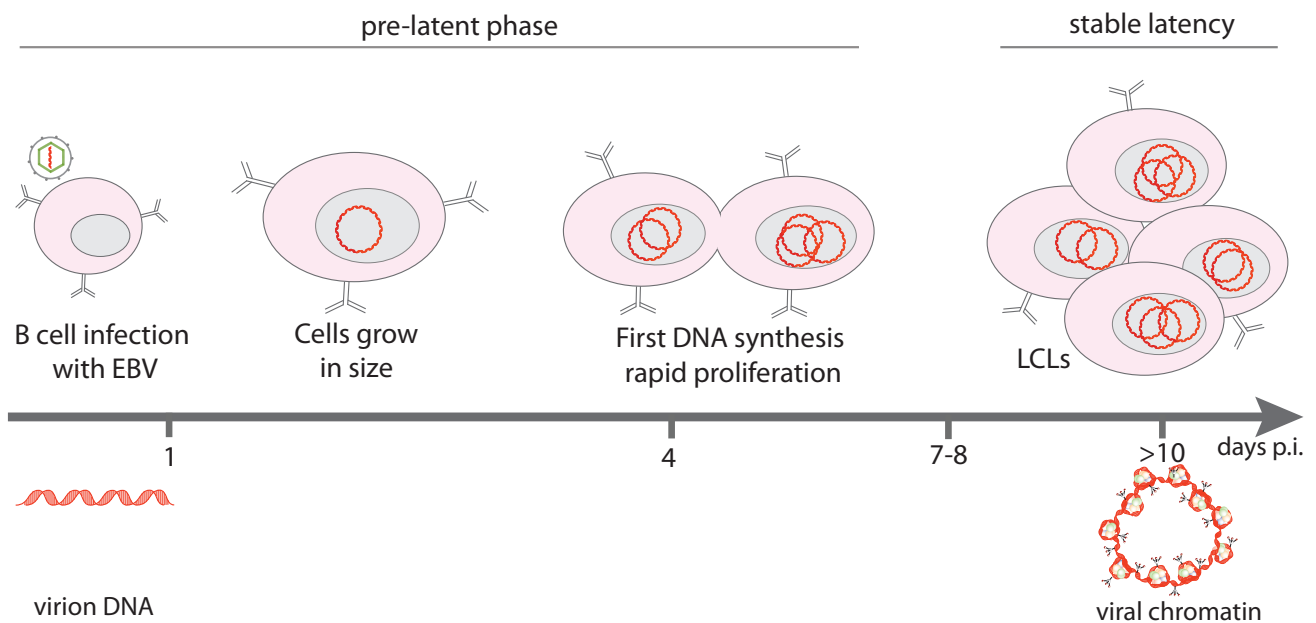
EBV infects human B cells and establishes a life-long but generally asymptomatic infection in us. Most B cells are latently infected and the virally infected cells are well controlled by the immune system of the host. EBV is also associated with several malignancies such as Burkitt's lymphoma (BL), nasopharyngeal carcinoma (NPC), Hodgkin lymphoma (HL), and non-Hodgkin lymphomas (NHL) (Young and Rickinson, 2004; Young et al., 2016). In all infected cells EBV initiates its latent phase, during which several copies of the genomic viral DNA persist as episomes that replicate in synchrony with cellular DNA. Latently EBV infected cells express a small subset of viral genes, only, which are essential for the success of viral infection. Latently infected B cells can initiate the lytic phase after expression of BZLF1, a viral transcription factor, which is induced when the infected cell encounters and binds its cognate antigen via its surface immunoglobulin. Antigen binding induces the B cell receptor signalling pathway and readily starts the productive phase, during which all viral lytic genes are expressed that are essential for replicating and packaging of viral DNA and release of viral progeny (Countryman and Miller, 1985; Kalla et al., 2010). EBV's life-cycle is regulated in a very strict and ingenious way that involves many fundamental host cell processes such as epigenetics, transcription, metabolism, and immune responses.

*In vivo*, only a small subset of human memory B cells are infected with EBV, which makes it impossible to investigate the infection process and perform a high-throughput molecular analysis in the small and rare population of cells. Suitable animal *in vivo* models also do not exist that mimic EBV infection in humans. A close relative is the murine gamma herpes virus MHV68, which shares many features with EBV but lacks homologs of critical oncoproteins such as LMP1, LMP2A/B or the EBNA encoding genes (Olivadoti *et al.*, 2007; Aligo *et al.*, 2015). In contrast, the *in vitro* model of EBV infection is very well established, which allows to study the early phase of EBV infection in great detail in mature human B lymphocytes isolated *ex vivo*.

Adenoids surgically removed from children's nasopharynx contain large numbers of primary B cells, which can be easily isolated. B cells can be separated from other cells types such as T cells, granulocytes, or erythrocytes using Ficoll-Hypaque density gradient

centrifugations. The cells can be further purified and fractionated to enrich non-activated, naïve B cells by FACS sorting using specific antibodies against CD38 and IgD surface proteins or with magnetic beads to remove activated or memory B cells and other cell types. Wild-type (wt) EBV (2089 or B95-8) readily infect these cells, which can be tracked to record phenotypic changes during the very early stages of viral infection.

It was known from previous experiments in our lab that *in vitro* infected cells first grow in size. Around day four post infection (p.i.) they enter a phase of very rapid proliferation, which leads to established lymphoblastoid cell line (LCL) after about two weeks p.i. (Fig. 4). During the first days after infection, in the so-called pre-latent phase, the host cells undergo very dramatic changes, which the virus induces to establish a latent phase in them, eventually. It is also obvious that the virus itself has to go through rapid changes to adapt its DNA genome to the environment of the cellular nucleus and to escape from immune responses of the host cell. With this *in vitro* infection model, I focussed on the daily changes of the infected cells and compared them to non-infected cells with respect to metabolic, transcriptomic, as well as epigenetic criteria. My experiments demonstrate that dramatic metabolic, transcriptional, and epigenetic changes occur during the pre-latent phase of B cell infection with EBV.



**Fig. 4. Schematic overview of primary B cells during the early phase of EBV infection.**

B cells infected with EBV directly enter the pre-latent phase, during which the cells become reprogrammed to proliferating, activated lymphoblastic cells. Initially, the cells grow in size and enter a phase of rapid proliferation on day four post infection (p.i.). After about eight to ten days p.i. cell proliferation decelerates and constantly proliferating lymphoblastoid cell lines (LCL) emerge. The linear genomic DNA of EBV virions is epigenetically naïve and free of histones (left), but it circularizes immediately after infection, acquires cellular histones, becomes fully chromatinized and undergoes extensive DNA CpG methylation (right).

### 3.2. Phenotypic changes during the first eight days of EBV infection

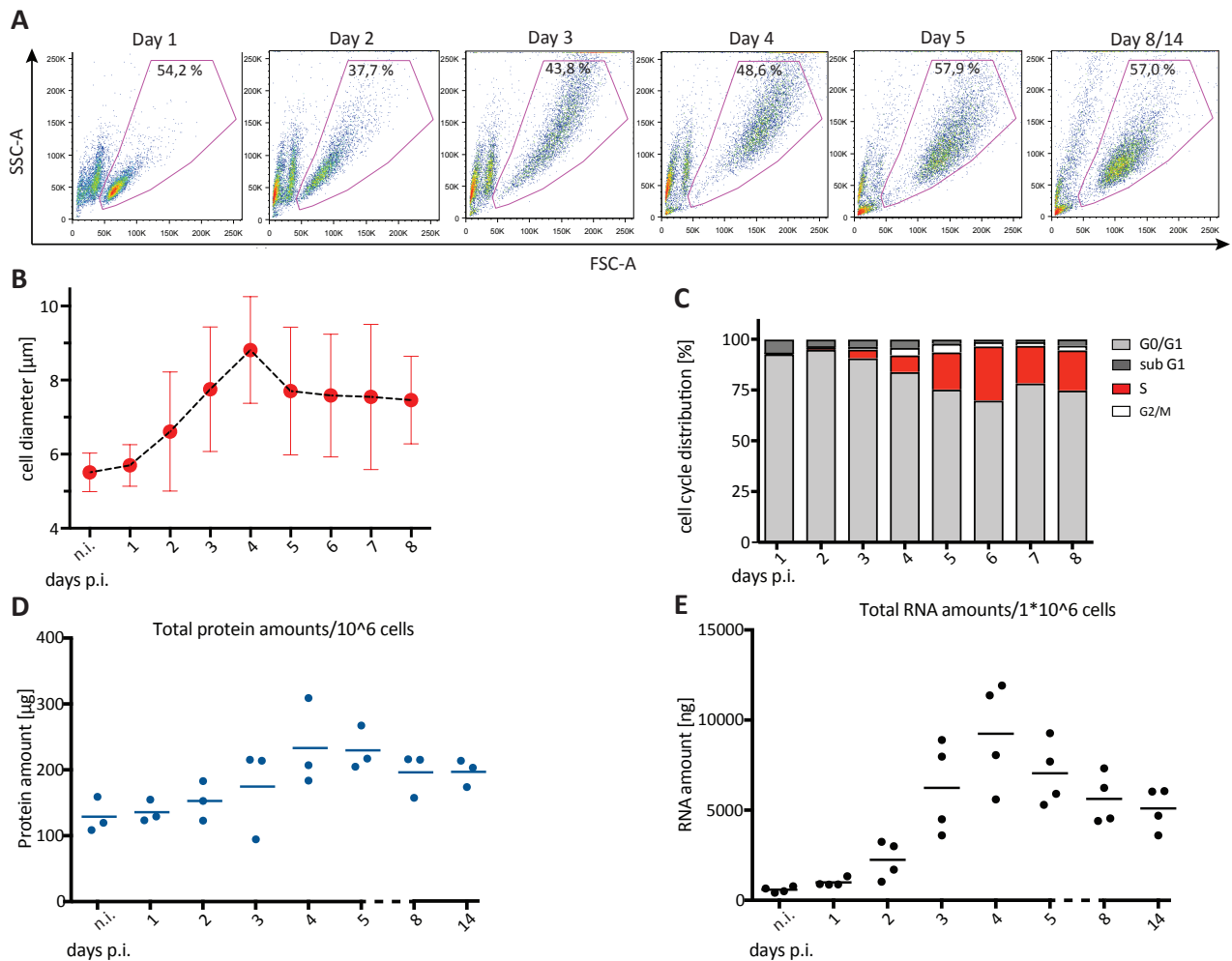
The phenotypic changes that occur after EBV infection were tracked on a daily basis employing mainly fluorescence-activated cell sorting (FACS) based methods together with other techniques to assess several biophysical parameters of the cells. First, I focused on the size of the cells, their cell cycle distribution as well as the content of total cellular RNA and protein. Freshly prepared donor B cells and infected B cells were checked every day and viable cells were physically enriched using a FACS Aria IIIu sorter with established gating strategies, which allowed to exclude cell debris and disintegrating cells from the subsequent analyses.

For infecting naïve, quiescent B cells I used the wt EBV (2089) strain at an optimal multiplicity of infection (MOI). The infected cells grew in size until day four p.i. (Fig. 5 A) with an increase in cellular granularity according to forward scatter (FSC-A) and sideward scatter criteria (SSC-A), respectively. Five days p.i., the cells started to form a more homogenous population and from day eight onwards the cells located to the gate that is characteristic for established lymphoblastoid cell lines. The absolute numbers of gated living cells changed accordingly. Until day three, the numbers of viable cells decreased considerably, but from day four onwards the cells started to divide and proliferated such that the cell numbers increased (data not shown). The exact diameters of uninfected and infected cells were measured by Dagmar Pich. Uninfected B cells were small with a diameter of 5.5  $\mu\text{m}$ , whereas infected cells grew gradually to reach a maximum diameter of 9  $\mu\text{m}$  at day four p.i., but shrank a bit in the following days to reach a cell size in the range of 7.5-8  $\mu\text{m}$  (Fig. 5 B).

EBV infected cells did not only change in size, but the infection also strongly affected their resting state. The cell cycle distribution of the cells was measured by Dagmar Pich (Pich et al., to be published) using short-term BrdU incorporation, which was recorded by FACS. B cells infected for one or two days stayed arrested in G<sub>0</sub>/G<sub>1</sub> compared with non-infected cells (Fig. 5 C). Starting on day three, a small fraction of cells entered the S phase of the cell cycle and this fraction increased further on day four. From day five to seven p.i. the cells underwent a very rapid phase of DNA synthesis and proliferation by dividing every 10-12 hours. Thereafter, they decelerated and adopted a generation time of about 24 to 36 hours.

The dynamic phenotypic changes and the entry of the activated cells into the cell cycle led to impressive alterations in the content of cellular macromolecules such as RNAs and proteins. I measured the amount of total protein contained in one million cells and compared it with established LCLs infected for 14 days (Fig. 5 D). The protein amounts increased up to 2.5

times from day two until day four and reached a stable level after about one week post infection. Additionally, the total RNA content showed a similar kinetics, but the differences were much more prominent. Uninfected B cells contained several hundred ng of RNA per one million cells. The RNA content increased up to 10,000 ng at day four p.i., a ten fold increase compared to non-infected cells (Fig. 5 E). As observed in the previous experiments, the amounts of total RNA appeared to be stable from day eight p.i. onwards.



**Fig. 5. Phenotypic changes of the host cell during the first two weeks of EBV infection.**

B cells were purified from adenoid tissue and infected with wt EBV (2089) with an MOI 0.1. Every day p.i. the cells were collected and analysed for multiple parameters.

- Time-resolved phenotypic changes of EBV-infected cells analysed using FACS and FSC-A (x-axis) and SSC-A (y-axis) criteria. Newly infected cells on day one p.i. formed a homogenous population of small cells with low granularity. The cells increased in size and reached maximal size and granularity four days p.i. From then on, the cells adopted a more discrete population. On day eight and 14 post infection the cells were indistinguishable (data not shown).
- The panel shows the changes of the cellular diameter during the first eight days after EBV infection. Mean and standard deviation were calculated from the images of about thousand cells at the specified time points p.i.. Uninfected cells (n.i.) were small, whereas infected cells grew constantly until they reached their maximum at day four p.i.. At day five, the cells became smaller and from this time point the cell diameter was stable.
- B cells after wt EBV (2089) infection were incubated with BrdU for one hour followed by intracellular staining with a BrdU specific antibody. The cells were measured by FACS to determine the cell cycle distribution. Until day three the cells did not proliferate and stayed arrested in G0/G1 cell cycle phase.

From day three p.i. the cells started to synthesize cellular DNA indicated by the fraction of the cells in the S phase of the cell cycle. (Experiments in panels B and C were performed by Dagmar Pich; panel B is based on three independent replicates)

- D. Uninfected cells and cells infected with wt EBV (2089) were collected daily and total protein lysates were obtained. The samples were measured using the Bradford assay. The graph shows the total amount of proteins ( $\mu\text{g}$ ) obtained from one million cells. Measurements were performed every day after infection as indicated. The protein content increased slowly, peaked at day four p.i. and decreased again in the following days. The solid blue lines show the mean average from cells prepared from three different donors, individual samples are represented as single dots.
- E. Amounts of total RNA (ng) obtained from one million cells were measured using the Shimadzu MultiNA Electrophoresis System for DNA/RNA analysis. The RNA content increased up to ten times until day four p.i., but on day five and onwards it decreased over time. Solid black lines represent the mean average from four different donors. Each donor is displayed as a single dot.

My data documented that the infection with EBV triggers many profound changes in the host cell, which are highly dynamic during the first eight days of infection. The most dramatic changes with respect to cell size, RNA and protein content took place within the first four days after infection, and the cells adopted a less dynamic, established state from day eight p.i. onwards.

The cell cycle distribution was also very dynamic during the first eight days of infection. Initially, the infected cells did not synthesize cellular DNA, but starting from day three they underwent very rapid rounds of DNA replication and cell divisions to slow down later and enter a stable state equivalent to that of established lymphoblastoid cell lines.

### 3.3. Regulation of the cellular metabolism in the course of EBV infection

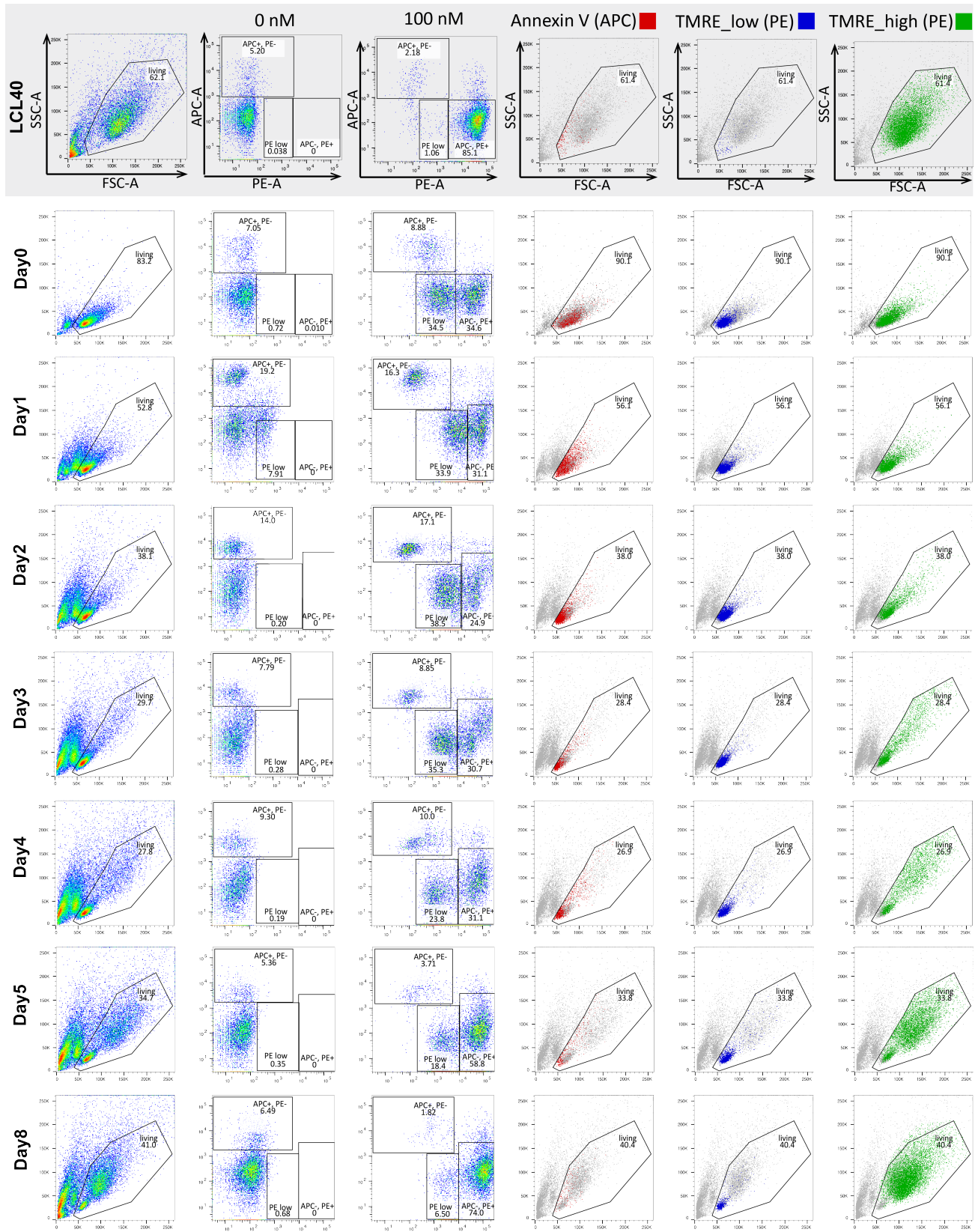
Alterations in the cellular metabolism affect many biological processes such as cancerogenesis or activation and differentiation of immune cells. Interestingly, many viruses are also able to trigger changes in the cellular metabolism to promote their productive, lytic phase or to support their replication and long-term infection. This is the reason why I investigated certain selected metabolic processes and their kinetics during the early stages of EBV infection.

B cells from adenoid tissue were isolated and purified using a commercial B cell isolation kit that yields “untouched”, non-manipulated B cells, but removes all other cell types present in the preparation. The B cells were infected with wt EBV (B95-8), which lacks a fluorescence marker gene that would have interfered with the subsequent FACS analyses. Uninfected B cells as well as cells collected at certain time points post infection were incubated with TMRE and Annexin V and analysed by FACS. TMRE (tetramethylrhodamine ethyl ester) is a red-

fluorescent, positive-charged dye, which is sequestered and accumulated by active mitochondria due to their negative charge and polarized membranes. Apoptotic cells with depolarized mitochondrial membranes will fail to incorporate TMRE and will not emit its red colour. Annexin V staining allows to detect early stages of cellular apoptosis. It binds to exposed phosphatidylserine residues (PS), which translocate from the plasma membrane to the surface of the cell in the early phase of cellular apoptosis. Established LCLs (LCL40 in the first row, Fig. 6) contained a minor fraction of Annexin V positive cells (2-5 %), the majority of the cells (about 85 %) stained intensively with TMRE indicating a very high mitochondrial activity; very few cells (about 1 %) showed a low level of TMRE staining.

Cells collected at different time points after EBV infection revealed very different characteristics. Fewer than 10 % of the non-infected cells were Annexin V positive immediately after preparation, but this fraction increased to almost 20 % one day post infection (Fig. 6). Annexin-positive cells decreased to about 5 to 7 % on day five and day eight. The majority of the infected and living cells incorporated TMRE, but two subpopulations, TMRE\_high (PE) and TMRE\_low (PE), were obvious (third column; Fig. 6). The TMRE\_low population was rather consistent in infected cells compared with the non-infected cells until day three post infection, but decreased dramatically starting from day four to become a minor fraction later in well-established cells. This subpopulation consisted of small cells with a low granularity, indicative of non-activated, probably quiescent and resting cells (fifth columns in Fig. 6). The summary of my four independent experiments using B cells from different donors is shown in Fig. 7 A. The graph indicates that the fraction of the cells with a strong mitochondria activity (TMRE\_high) increased starting on day four p.i. to eventually reach a level of more than 80 %. On the contrary, Annexin V-positive cells decreased over time to reach a level lower than 5 %.

The mitochondrial activity depends on a constant supply of the cells with nutrients and the subsequent catabolic processes. One of the basic compounds is glucose, which the cells readily take up from their environment. I wanted to examine if EBV-infected B cells show glucose uptake kinetics that differ from uninfected B cells or established LCLs. To scrutinize the dynamics of glucose uptake in the EBV infection model, I used fluorescently labelled (FITC) deoxyglucose analogue 2-[N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl) amino]-2-deoxy-D-glucose (2-NBDG), which is transported into the cell via the glucose transporter and is phosphorylated by the cellular hexokinase 1 (HK1) mimicking glucose uptake and the very first step of glycolysis,



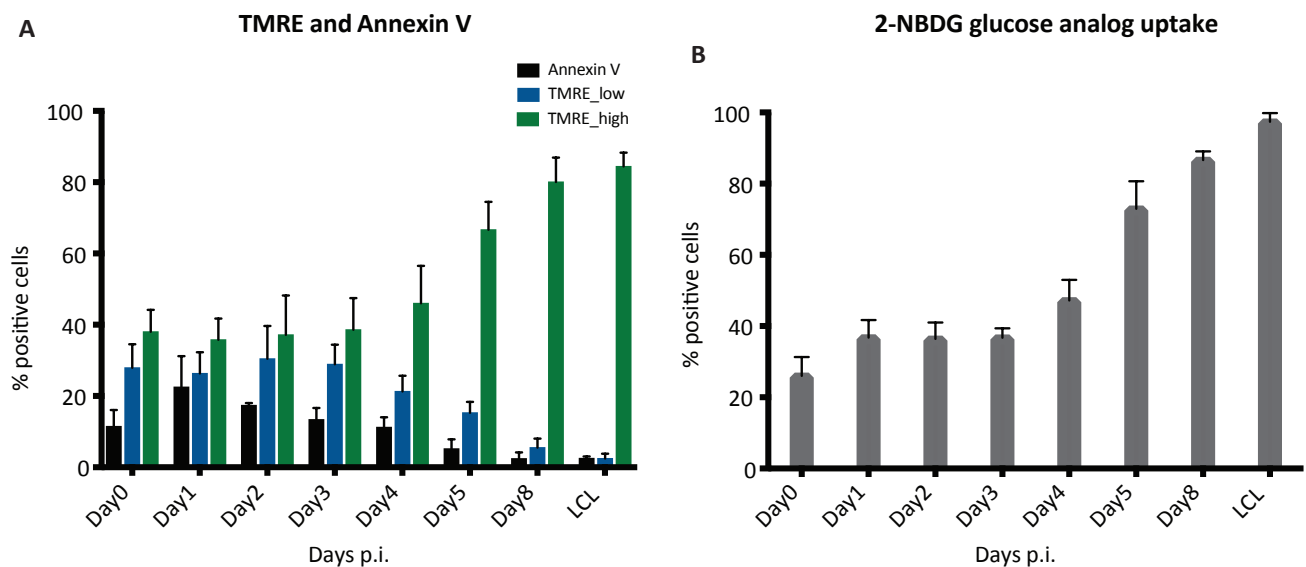
**Fig. 6. Tracking of apoptosis and mitochondrial activity during the first eight days after EBV infection.**

The FACS plots show the apoptosis rate (Annexin V staining) and mitochondrial activity (TMRE) of the cells at early time points after EBV infection.  $5 \times 10^5$  uninfected or EBV infected B cells collected at different time points after infection were incubated with 100 nM TMRE for 30 min followed by incubation with APC-coupled Annexin V for another 10 min. The percentages of positive cells measured by FACS are indicated.



The first row shows established LCLs as a control, the remaining rows represent uninfected primary B cells or cells on different days p.i.. In the first column, I compared the forward and sideward scatter of the different cell samples. The next two columns compare untreated cells (0 nM) with cells incubated with 100 nM TMRE for 30 min. Columns four to six show the results after back-gating of the cell populations indicating the fraction of cells that were Annexin V positive (red), TMRE\_low positive (blue) or TMRE\_high positive (green). During the first three days, 10-20 % of the cells were Annexin V positive and from day five this fraction decreased. TMRE incorporation revealed two clearly positive populations in uninfected B cells (Day0). The population of TMRE\_high cells became dominant from day four p.i. indicating a high mitochondrial activity of stably infected cells. The figure shows the results of one representative experiment out of four biological replicates.

respectively. Glucose uptake and its catabolic breakdown to pyruvate is one pathway to regulate the metabolic cellular rate.

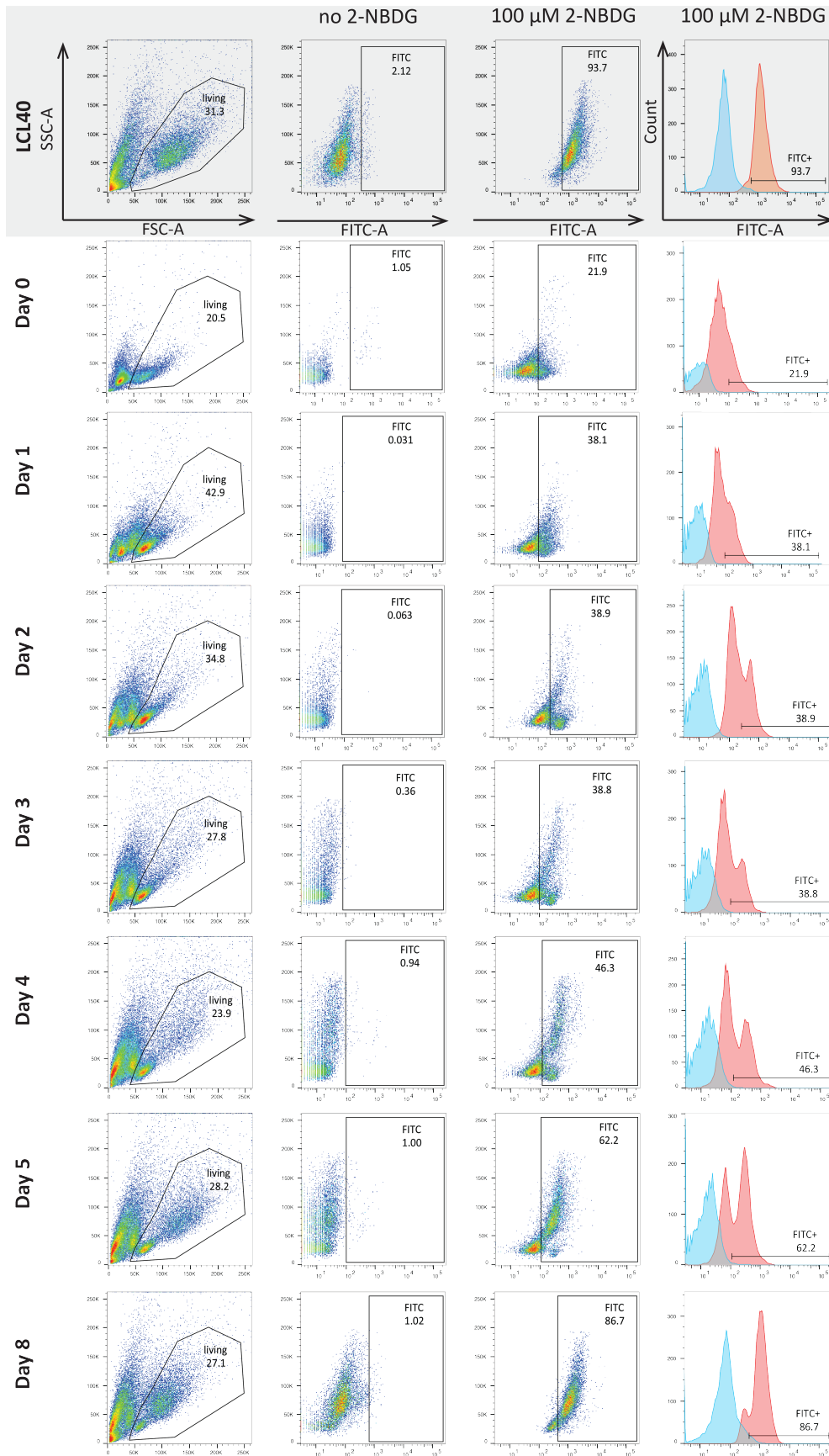


**Fig. 7. Summary of the metabolic activities of B cells early after EBV infection.**

- A.** The fractions of Annexin V binding cells and cells that take up TMRE as a measure of their mitochondrial activity. The fraction of Annexin V-positive and TMRE\_low cells decreased in the course of EBV infection. The fraction of TMRE\_high cells stayed constant until day three p.i., but increased starting on day four to constitute more than 80 % one week p.i.. Mean and standard deviations from four independent biological replicates are shown.
- B.** Uptake of the glucose analogue (2-NBDG) in the course of early EBV infection. During the first three days p.i. the fraction of cells that took up glucose was constant. Starting on day four the fraction of glucose analogue-positive cells increased up to 100 % indicating a very high rate of glucose uptake. Bars in the graph display mean and standard deviation from four independent experiments.

Identical numbers of freshly purified uninfected B cells and cells infected with EBV were plated in a low-glucose media and incubated with 2-NBDG for one hour. After incubation, the cells were washed with ice-cold PBS and kept on ice to inhibit the dephosphorylation and export of 2-NBDG from the cells. Samples were analysed at the indicated time-points by FACS using the FITC channel to record the uptake of the glucose analogue. Established LCLs showed a very efficient glucose analogue uptake (more than 90 % positive cells) in a distinct positive population compared with cells that were not incubated with the glucose analogue (2-NBDG)





**Fig. 8. Uptake of the glucose analogue 2-NBDG during B cells infection with EBV.**

FACS analysis of the glucose uptake in B cells at different days after EBV infection.  $5 \times 10^5$  of uninfected and EBV infected B cells were incubated with  $100 \mu\text{M}$  2-NBDG at the indicated time points p.i. for one hour at  $37^\circ\text{C}$ . The cells were washed with PBS and analysed by FACS including the untreated samples (no 2-NBDG). The first row

shows an established LCL as a reference, the remaining seven rows represent uninfected B cells and cells infected for different time periods as indicated. The first column compares forward and sideward scatter of cells incubated with 2-NBDG. The next three columns show the analysis of untreated cells (no 2-NBDG) and cells incubated with 2-NBDG. During the first three days after viral infection the cells did not show significant changes in glucose uptake. Starting from day four p.i., the glucose uptake activity increased substantially to reach more than 85 % cells with increased glucose uptake. The figure depicts one experiment out of four performed with B cells from different donors.

(first row in Fig. 8). The populations of the uninfected B cells and cells infected with EBV for three days contained about 40 % 2-NBDG positive cells. Starting from day four, the fraction of FITC positive cells constantly increased. After eight days p.i., more than 80 % of the cells (Fig. 8) were 2-NBDG-positive indicating an efficient glucose uptake and an increased energy consumption. Fig. 7 B summarizes the results from four independent donors.

My data revealed that during the first three days after EBV infection the population of cells did not seem to undergo dramatic changes in their uptake of glucose. Similarly, mitochondrial activity was undistinguishable from uninfected cells during this infection period. However, from day four onwards there was a detectable gain in cell activity as revealed by a larger fraction of cells with active mitochondria and a rise in the percentage of the cells that efficiently took up glucose. When the cells reached the stable stage of LCLs most of them had active mitochondria and took up glucose in a short period of time (Fig. 7). These results need to be more intensively investigated in the future focusing on other metabolic pathways such as the Warburg effect or glutamate metabolism, which likely are also involved in the reprogramming of primary B cells and their dynamic metabolic program after EBV infection.

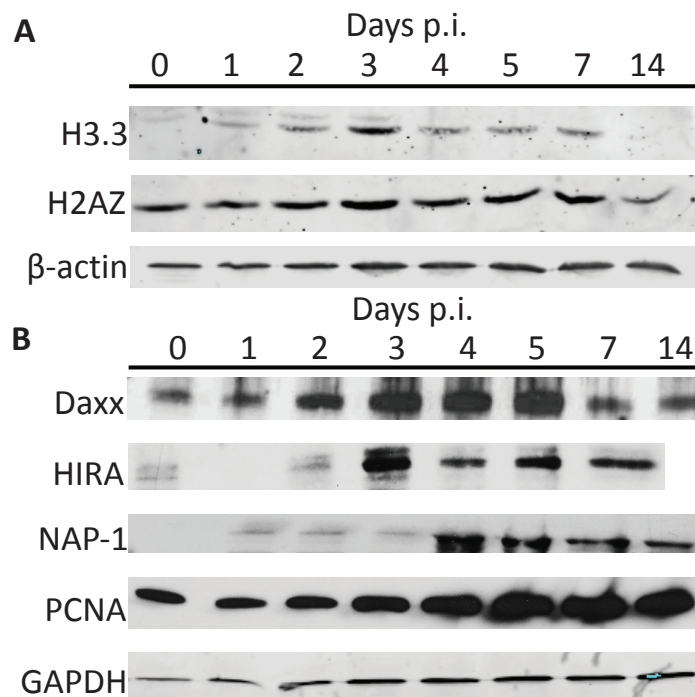
#### **3.4. Histones and histone chaperones dynamics in the early phase of EBV infection**

The EBV genome enters the cell nucleus as an epigenetically naïve, linear, double-stranded DNA molecule, which circularizes immediately. From previous work by Anne Wöllmer in our lab we learnt that EBV DNA methylation occurs late after infection, reaching a stable, high level about three weeks p.i.. CpG methylation of viral DNA plays a very important role during lytic phase reactivation, but does not seem to contribute to transcriptional regulation in the first week of infection (Woellmer and Hammerschmidt, 2013; Woellmer et al., 2012). Nucleosome occupancy appears to be a much earlier event that is completed already at the end of the first week p.i.. Nucleosome occupancy could be important for the

establishment of the latent phase and the downregulation of the initial, promiscuous viral gene expression to escape from antiviral immune responses of the human host.

It is unknown if nucleosome acquisition of viral DNA takes place prior to the first round of DNA replication and cell division (prior to day three p.i.) or later during infection. Histone variant proteins (e.g. H3.3, H2A.Z), which can be synthesized independently of the cell cycle might play a role prior to day three p.i. together with histones that stem from the small pool of freely available histones. Alternatively, nucleosome transcription is likely linked with the first S phase and cell division (after day three p.i.), when all canonical histone proteins such as H2A, H2B, H3 and H4 are synthesized at the beginning of the S phase.

Initially, I analysed the protein levels of the histone variants H3.3, H2A.Z and selected histone chaperones in the course of EBV infection (Fig. 9 A and B). The protein levels of histone H3.3 peaked at day three p.i., prior to the first cell division and decreased at later time points.



**Fig. 9. Immunoblots of chromatin constituents in the course of EBV infection.**

One million uninfected B cells or B cells infected with wt EBV (2089) were collected at the indicated time points and prepared for total protein lysates. Lysates obtained from  $5 \times 10^5$  B cells were loaded per lane.

- A.** Histone variant H3.3 protein levels increased early after infection, peaked at day three and were reduced later during infection. Levels of the histone variant H2A.Z protein levels seemed to be constant over time.  $\beta$ -actin served as a protein loading control. One representative experiment out of two is shown.
- B.** Levels of histone chaperone proteins varied during the early phase of EBV infection. Daxx and HIRA are two chaperones of histone variant H3.3, whereas NAP-1 is a chaperone of the canonical H2A and H2B histones. Levels of Daxx and HIRA chaperone proteins peaked at day three and then decreased slowly. NAP-1 showed low levels during the first three days p.i. but increased noticeably on day four and was stable later after infection. Levels of the PCNA protein, which is a hallmark of DNA replication increased from day four p.i. and stayed high on days five and seven, when cells proliferate intensively. GAPDH protein was used as a protein loading control. One representative experiments out of two is shown.

In contrast, expression of the histone variant H2A.Z protein was constant (Fig. 9 A). Furthermore, the histone H3.3 specific chaperones Daxx and HIRA showed a similar expression over time as the histone H3.3 variant protein, but with a stronger expression on day three and reduced protein levels later during infection. The Daxx histone chaperone retained higher protein levels from day three till day five p.i. The canonical histone chaperone NAP-1, which acts in the S phase of the cell cycle, showed elevated protein levels from day four to five in a period when the cells proliferated rapidly. PCNA is an essential DNA replication-associated factor, that had higher steady state levels during the intense proliferation phase that started on day four until day seven after EBV infection (Fig. 9 B).

These results may suggest that histone variants and their specific chaperones play a role in the process of nucleosome acquisition of the EBV genome in the early period of viral infection. Immunodetection of proteins is not a very robust and comprehensive approach, because it is highly dependent on the specificity and quality of the antibodies. Thus, I decided to investigate changes triggered by EBV infection further applying genome-wide, high-throughput experiments such as ribonucleic acid sequencing (RNA-seq) and micrococcal nuclease sequencing (MNase-seq), which are described in following chapters of my thesis.

### **3.5. Global transcriptional regulation in the host cell after viral infection**

Every viral infection triggers a number of host cell responses and biological processes that viral factors modify, repress, or control. The encounter between host and pathogen results in very complicated interactions, which can cause diverse alterations in the transcriptome of the infected host cell. In this PhD work I focused on the pre-latent phase of EBV infection, when the cells are on their way to become reprogrammed to investigate the dynamics of these changes in the cellular transcriptome.

#### **3.5.1. Preparation, design and quality control of the time-resolved RNA-seq experiment**

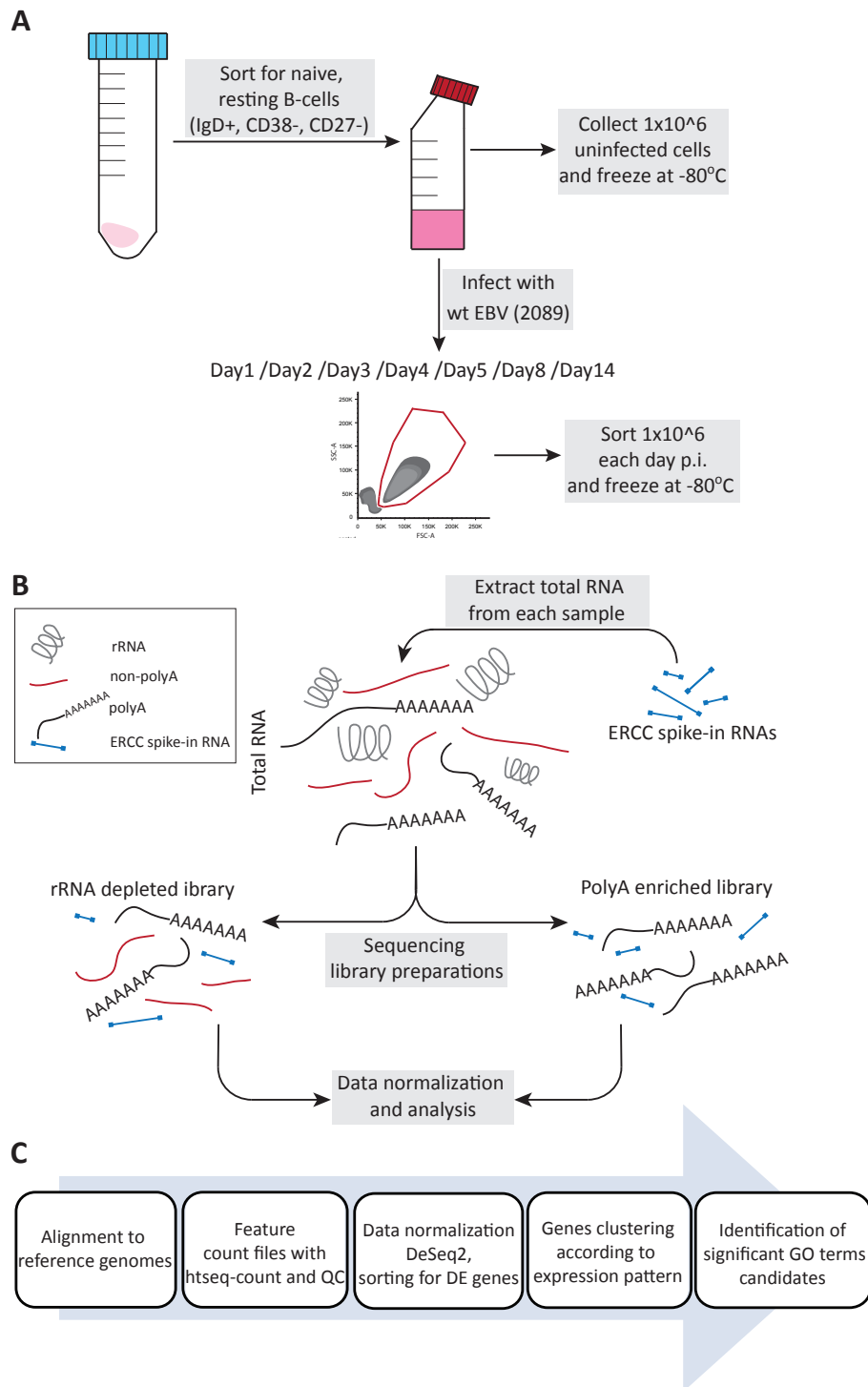
##### **3.5.1.1. Sample collection and library preparations for NGS**

The development of the high-throughput technologies based on the Next Generation Sequencing (NGS) such as RNA-seq opens a wide range of possibilities analysing cellular transcriptomes. This method combines a high resolution, a wide dynamic range and superior precision to identify nearly all existing transcripts and quantify them accurately (Wang et al., 2009; Nagalakshmi et al., 2010). I wanted to employ this powerful technique to investigate the

alterations of the cellular transcriptomes of B cell infected with EBV and identify the biological processes that lead to the successful implementation of EBV's latent phase, eventually.

Primary B cells isolated from adenoid tissue obtained from surgically removed nasal polyps of young children were stained with CD38 and IgD antibodies to allow the sorting of quiescent, primary and naïve B cells. As uninfected control samples, one million of the sorted cells from each donor were collected, washed, immediately frozen in TRIZOL reagent, and stored at -80 °C. The remaining cells were infected with wt EBV (2089) with an MOI 0.1 and incubated overnight at 37 °C. The next day, cells were washed to remove unbound virus and resuspended in fresh medium with cyclosporine to inhibit contaminating antiviral T cells from being activated. Each day after infection the cells were sorted according to the established gating strategy to collect living cells, only, and remove cellular debris as schematically shown in Fig. 10 A. One million sorted cells were washed, resuspended in TRIZOL reagent and stored at -80 °C. The daily samples from three different donors were processed together to extract total cell RNA. To control this and subsequent steps, certain molar amounts of commercial External RNA Controls Consortium (ERCC) spike-in controls were added to each sample. They allowed tracking every step of library preparation as well as monitoring the sequencing depth. Extracted total RNA samples were then split into two aliquots for preparing two complementary types of cDNA libraries: polyA enriched libraries and rRNA depleted random hexamer enriched libraries (Fig. 10 B). With these two alternative approaches I was aiming to identify and quantify all RNAs in EBV infected cells including also the subset of lncRNAs (Zhang et al., 2014) or histone gene RNAs that commonly are not polyadenylated (Kari et al., 2013; Marzluff et al., 2008).

The next steps in library preparation included cDNA amplification and ligation of specific adaptors to allow multiplex sequencing. The finished libraries were sequenced in a single flow cell using the Illumina HiSeq4000 platform. From all samples 100 base - long paired-end sequence information was obtained. A thorough bioinformatic analysis was the next step that included mapping to the human and EBV genomes, counting reads for each gene combined with data normalization (Fig. 10 C). The main aim of this experiment was to identify time-resolved clusters of co-expressed genes and to reveal their biological functions using standard gene ontology (GO) enrichment analyses. My results confirmed obvious biological processes such as metabolic pathways and DNA replication but also identified unexpected



**Fig. 10. Schematic overview of the workflow and the analytical steps of my time-resolved RNA-seq experiments.**

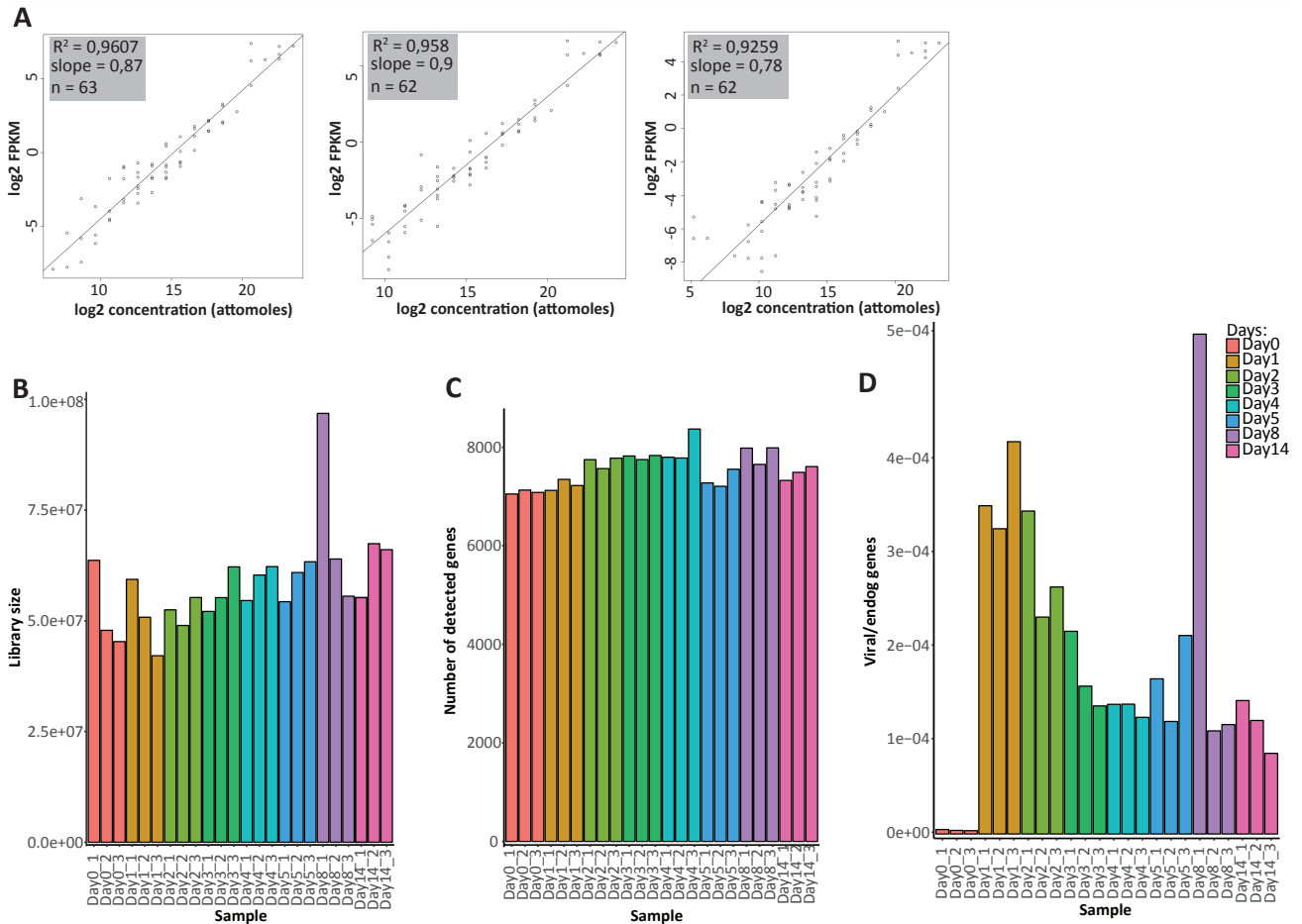
- A.** Sample collection and preparation for RNA-seq experiments. B cells isolated from adenoids tissue were sorted to obtain naïve, resting B-cell populations. One million non-infected cells were collected, lysed in TRIZOL and immediately frozen at  $-80^{\circ}\text{C}$ . The remaining cells were infected with wt EBV (2089) virus with MOI 0.1. At the indicated time points cells were sorted according to the forward and sideward scatter criteria of physically intact cells. One million alive cells were pelleted, snap frozen in TRIZOL reagent and stored at  $-80^{\circ}\text{C}$ .
- B.** Total RNA was extracted from each sample that had been supplemented with ERCC spike-in control RNAs. The prepared RNA samples were used to prepare two types of NGS libraries: polyA enriched libraries and rRNA depleted libraries. All libraries were sequenced together in a single flow cell by paired-end sequencing of 100 nucleotides followed by bioinformatic analyses.
- C.** The scheme illustrates the main steps of the bioinformatic analyses of the RNA-seq data (QC, quality control; DE, differentially expressed; GO, gene ontology).

pathways such as ncRNAs processing and chromatin silencing including novel, interesting targets for an analysis in the future.

### 3.5.1.2. Quality controls of the sequenced libraries

In general, RNA-seq experiments provide an enormous amount of sequencing data, when designed and performed carefully. RNA-seq experiments are demanding as they require many individual and sophisticated steps such as RNA isolation, reverse transcription, adaptor ligation, limited rounds of PCR amplification, etc.. To minimize possible biases in the downstream RNA-seq data, comprehensive quality controls have to be applied from the very beginning and after each step of data analysis. Starting from the raw sequencing data, it is highly recommended to perform a FastQC analysis (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> or FastQC on the Galaxy platform). This analysis provides a rapid control of the sequencing output including quality scores per base, GC content, etc.. In the next steps of my experiments, I relied on externally added trace amounts of artificial RNA sets (ERCC spike-in RNAs) purchased from ThermoScientific.

The RNA-seq data, obtained from the many samples of my time-course infection experiments were initially checked for raw sequencing reads quality, as mentioned above, followed by the analysis of the ERCC spike-in control samples and post-mapping quality controls. The set of ERCC spike-in RNAs consists of 92 polyadenylated, artificial RNAs of various lengths and different concentrations mimicking the complexity of eukaryotic transcripts. Figure 8 A presents dose-response plots with  $\log_2$  concentrations of ERCC spike-ins versus  $\log_2$  of their FPKM values to analyse their linear regression equations. My sequencing approach detected about 70% of all spike-in RNAs in the samples missing those that are present in very low concentrations. These results indicated the lower limit of detection of the RNA-seq experiments and their very wide dynamic range covering about 12-14  $\log_2$  of their initial concentrations. After mapping to the human and EBV genomes all samples were also analysed for certain additional parameters to check the overall quality of the libraries. The library size graph (Fig. 11 B) shows that different samples were sequenced with similar depth about 60 million reads per library with the exception of sample day8\_1 that had many more reads in contrast to all other samples.



**Fig. 11. Quality control of the sequenced libraries.**

- A.** Detection of the ERCC spike-in RNAs to assess the lower limit of detection and the dynamic range of the sequenced libraries. Artificial ERCC spike-in control RNAs were added according to the manufacturer's recommendations to each RNA sample prior to library preparation. The three graphs are representative examples of regression curves. The x-axis indicates the  $\log_2$  concentrations of the spike-in RNAs added to the samples, the y-axis depicts the  $\log_2$  of Fragments Per Kilobase of transcript per Million mapped reads (FPKM) per individual ERCC spike-in RNA. The coefficient of determination ( $R^2$ ), the slope of the regression line and the numbers of spike-in RNAs (n) identified out of 92 added to the samples are indicated.
- B.** The quality control graph provides information about the size of the sequenced libraries prepared from the different samples from three independent donors. The majority of the samples shows comparable sequencing depth except sample Day8\_1. This sample was further investigated for a possible bias and excluded, finally; see Supplementary Figure 1 for additional information in the Appendix.
- C.** The bar plot depicts the numbers of detected genes with a threshold of > 10 reads per million (RPM) in each sequenced sample. The numbers of detected genes were comparable.
- D.** Ratios of the numbers of viral genes versus the numbers of cellular genes at each day for three replicates are shown. Viral genes were absent in uninfected cells (Day0) as expected, in contrast to day one and two, when the ratio of detected viral genes was very high. From day three on, the numbers of detected EBV genes decreased to reach a stable level at day four p.i.. Sample Day8\_1 appears to be an outlier because of its high ratio of viral versus cellular genes fraction (see Supplementary Figure 1 for more information).

The numbers of mapped and identified genes (Fig. 11 C) seemed to be very similar in all samples encompassing around 7000 genes. The ratios of the detected viral versus cellular genes were dynamic during the course of infection (Fig. 11 D). In the samples from uninfected cells



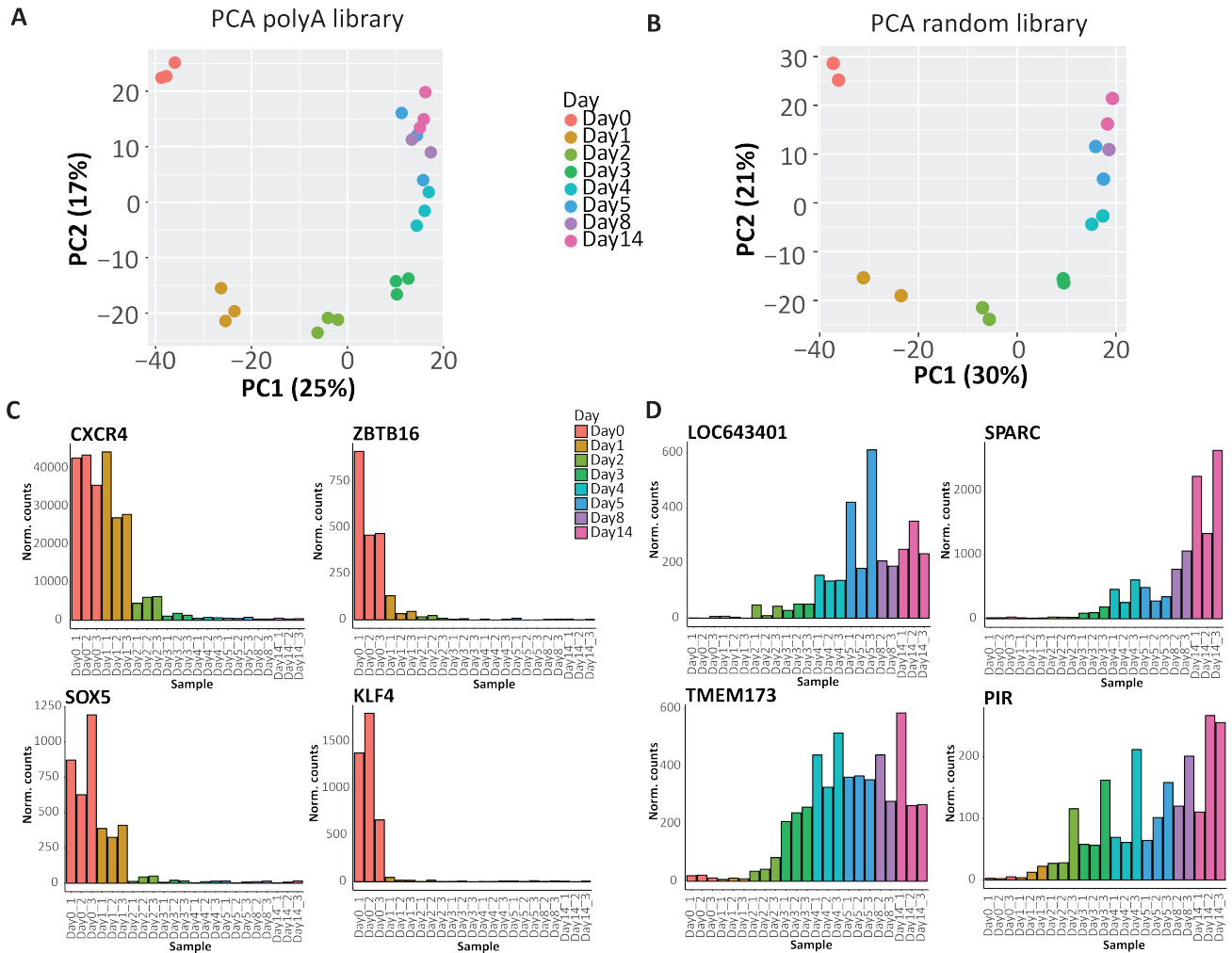
(Day0), no viral transcripts were found as expected. In contrast, the samples from cells infected for one and two days reached level of  $3-3.5 \times 10^{-4}$  viral/cellular genes. Starting at day three p.i., the ratios of viral genes decreased to reach an almost stable level. Again, sample Day8\_1 looked different and revealed almost five-times more viral versus cellular genes compared to the other two replicates at this day p.i.. A Principal Component Analysis (PCA) as described in the following chapter documented that this sample differed dramatically from all other samples (Supplementary Figure 1 in Appendix). I decided to exclude it from my further analyses to eliminate this outlier to prevent a bias in the final results. The data shown in Fig. 11 were derived from polyA enriched libraries. The quality controls of the rRNA depleted libraries were very comparable and are shown in Supplementary Figure 2 in the Appendix.

### 3.5.1.3. Principal Component Analysis (PCA)

The comparison of RNA-seq data collected at different time-points after EBV infection should reveal the differences and dynamics in the host cell transcriptome in the course of infection. PCA is a statistical method that can reduce the large number of variables that are present within multivariate data sets such as sequencing samples, by identifying the orthogonal directions (called Principal Components - PC) that explain most of the variations in the data set. The PCA analysis was performed on the basis of the log-transformed expression matrix, and in Fig. 12 A and B the first two Principal Components (those that express the highest fraction of variance) are plotted. The PCA plot reveals that the largest contribution to the variation in the data is the infection status of the cells (PC2) followed by the progression of the viral infection over time (PC1). Samples obtained from day five, eight and 14 are closer to each other than the earlier time-points, indicating reduced variances present in samples from these later time points.

PCA from my two different types of RNA-seq libraries (polyA enriched and rRNA depleted) showed almost identical characteristics in the PCA plots (Fig. 12 A). The well-defined and clear changes between different days after EBV infection suggested that discrete sets of genes contributed to specific PC1 or PC2 variances. From the ranked PC1 or PC2 matrix I investigated the top entries of the two lists to look at these potentially interesting candidates (Fig. 12 C, D). Four examples of the gene candidates (the full lists of the top 30 candidates are provided in Supplementary Table 1 of the Appendix), which were highly expressed in the

uninfected cells is a chemokine receptor and three transcription factors that may control the quiescent state of the non-activated B cells (Fig. 12 C).



**Fig. 12. Principle Component Analysis (PCA) of the time-resolved RNA-seq data.**

- A.** The PCA plot shows samples from all polyA enriched libraries at the given time points p.i.. With the exception of day 8 p.i. each day is represented by three replicates. The samples are shown as a function of the first (PC1) and the second (PC2) principal components. X and y axes show the percentages of variance explained by PC1 and PC2.
- B.** The PCA plot indicates PC1 and PC2 from RNA-seq data obtained from rRNA depleted, random hexamer enriched libraries. The results were very similar to the data in panel A. Again, PC2 defined changes between uninfected and EBV-infected cells and PC1 depicted differences between different days after infection.
- C.** The bar plots show the time-resolved expression of selected genes, which primarily contribute to sample variances as defined by PC2. The x-axis displays samples from different days and three different donors, whereas y-axis depicts normalized reads counts. All genes are highly expressed in uninfected cells but are strongly repressed shortly after EBV infection. *CXC4R4* and *SOX5* showed also significantly lower expression level at day one compared with the uninfected cells.
- D.** Examples of four genes that stand for variances according to PC1. The indicated genes were barely expressed in uninfected cells but became activated at different time points post EBV infection.

*CXC4R4* is a chemokine receptor that is involved in haematopoiesis, migration of resting leukocytes, and in maintaining the undifferentiated state of hematopoietic stem cell (Zou et al.,

1998; Sugiyama et al., 2006). Another candidate was *SOX5*, a transcription factor (TF) mainly associated with embryonic development and cell fate determination, but it is apparently also involved in B cell differentiation reducing their proliferative capacity (Rakhmanov et al., 2014). *KLF4* encodes a transcription factor of the larger family of zinc-finger SP-1 like TFs. It is mainly known as one of the factors that can lead to induced pluripotent stem cells (iPS) (Takahashi and Yamanaka, 2006), but it also has many important functions in processes such as proliferation, apoptosis, or cell reprogramming and is highly expressed in non-dividing, cell cycle-arrested cells (Chen et al., 2001). Another example, the *ZBTB16* gene, also known as *PLZF*, is a member of yet another subfamily of zinc-finger TFs, which is involved in cell cycle processes as well as B cell malignancies (Parrado et al., 2000). It also interacts with histone deacetylases, which together with other factors were reported to induce innate immune responses (Xu et al., 2009; Sadler et al., 2015). In contrast to transcripts that contributed to PC2 of the uninfected samples, the PC1 coordinates of the transcripts from samples infected for two or three days increased gradually to become rather invariant on day four and later (Fig. 12 A, B). One of the interesting genes in samples infected for eight days was *SPARC*, which encodes a glycoprotein and has emerged to correlate with progression of many cancers such as ovarian or prostate cancer, for example (Chen et al., 2012). In addition, there are reports that claim a role of *SPARC* in B cell malignancies (Sangaletti et al., 2014). *TMEM173*, also known as *STING* encodes a transmembrane protein that promotes innate immune response by stimulating the secretion of interferon type I as a response to DNA-containing pathogens such as DNA viruses or bacteria (Ishikawa et al., 2009). On the contrary, the *TMEM173* protein might support the success of HSV-1 in certain cell and animal models of herpesvirus infection (Kalamvoki and Roizman, 2014). The *PIR* gene that encodes pirin is also expressed late during EBV infection. It is a nuclear protein, which presumably acts as a transcriptional cofactor. It was found to interact with B cell lymphoma 3-encoded oncoprotein (BCL-3) and to promote gene transcription (Dechend et al., 1999). Another candidate is *LOC643401*, a lncRNA also known as *PURPL*. Recently, it was found that *PURPL* suppresses basal p53 levels in colorectal cancer cells by preventing the interaction of p53 with MYBBP1A, which can activate expression of p53 in the nucleoplasm (Li et al., 2017).

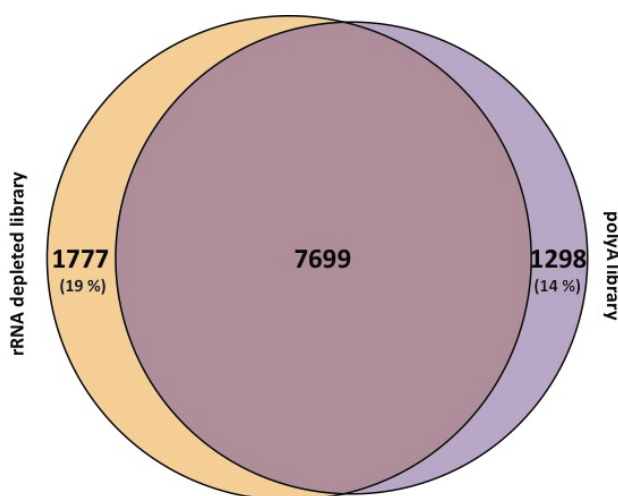
The PCA analysis identified linear combinations of different genes that explain most of the variations in the data (PC1 and PC2). The component scores along PC1 and PC2 of the top 30 candidates suggested that these transcripts are critically involved in the steps of cellular reprogramming (Supplementary Table 1 in the Appendix). Their functions as a group of genes

are not immediately obvious, but my analysis revealed that they act in a time-dependent manner during infection of primary human B cells with EBV.

#### 3.5.1.4. Comparison of the two types of NGS libraries

I analysed two different types of sequencing libraries in my RNA-seq experiments. One set of samples with three replicates where each was polyA enriched prior to cDNA synthesis, the second set with two replicates each was based on rRNA depleted random hexamer enriched RNA samples to generate the NGS libraries. These two techniques together should maximize the coverage of cellular transcripts. This is because certain cellular transcripts are non-polyadenylated including histone RNAs and certain non-coding RNAs (ncRNAs). RNA-seq samples obtained from both libraries entered the same bioinformatic analysis pipeline. After processing with DEseq2, I looked for transcripts detected in both libraries and ones that were uniquely contained in one of the two libraries, only.

The comparison of significant differentially expressed (DE) genes from polyA enriched libraries with rRNA depleted libraries is summarized in the Venn diagram (Fig. 13). Only in the rRNA depleted libraries I detected 29 viral transcripts and a number of different types of cellular ncRNA (miRNAs, lncRNAs, snoRNAs etc.) that were absent in the polyA enriched libraries. More than 400 transcripts from this list could not be classified to a known gene. Interestingly, in this type of library I exclusively discovered eleven additional histone genes, which were unique in the rRNA depleted libraries.



**Fig.13. Comparison between the polyA enriched and the rRNA depleted random hexamer primed RNA-seq libraries displayed as a Venn diagram.**

The two sets of transcripts show the intersection of 7699 genes found in both types of libraries. The rRNA depleted samples differed from the polyA enriched samples by 1777 genes, about 19 % of all entries of the list. The polyA enriched libraries contained 1298 unique and differentially expressed transcripts, which were 14 % of all entries.

In the group of transcripts identified in the polyA enriched libraries, only, I could not find any EBV specific mRNAs but certain ncRNAs although fewer than in the rRNA depleted

libraries. In the set of genes identified in the rRNA depleted libraries more than 250 genes did not have a specific gene name but an alphanumerical code. The entries from these lists do not point to obvious functions or molecular mechanisms, which might be specifically enriched. This comparison and additional, previously performed quality controls led me to choose the RNA-seq results from the polyA enriched libraries for my further analyses. These libraries rely on three independent donors and technical replicates and apparently contain sufficient information including chromatin components such as histone genes, some of which are not polyadenylated, which was my greatest concern. My interest was on daily changes of the cellular transcriptomes, thus the polyA enriched libraries contained all the information I wanted to collect. The rRNA depleted libraries can be used in the future to look more specifically for changes of ncRNAs levels during viral infection.

### **3.5.2. Dynamic regulation of specific biological processes in the course of EBV infection**

As mentioned previously, the main goal of my RNA-seq experiments was to identify significant differentially regulated genes and classify them into functional groups that share a similar timely-resolved pattern of gene expression in the early phase of EBV infection. Time-course RNA-seq experiments are very valuable in understanding dynamic biological processes, but on the other hand, they also introduce additional experimental and computational challenges during the different steps of data analysis (Spies and Ciaudo, 2015). In a conventional RNA-seq experiment that, for example, compares two different states of a given cell, the identification of differentially expressed genes is the main task of data analysis. In an experiment in which time-resolved RNA-seq data are key to understand the underlying processes differentially expressed genes have to be identified across different time points. My RNA-seq experiments consisted of eight time points collected at specific days after EBV infection as indicated in Fig. 10.

In the initial step of my data analysis, genes that are expressed at very low levels were excluded from the subsequent analyses. Filtering of genes was followed by the identification of genes that were differentially expressed (DE) between at least one pair of time points. The applied criteria were a false discovery rate (FDR) of  $< 0.1$  and an estimated fold change (FC) greater than 2 or smaller than 0.5, as estimated by the DEseq2 analysis package for R. This approach led to the identification of 8997 DE genes, which were used to create a distance matrix between them followed by a hierarchical clustering. The number of clusters was

estimated with a dynamic hybrid cut algorithm (Langfelder et al., 2008; van Dongen and Enright, 2012). This analysis yielded seven different clusters named according to an arbitrary colour code. The clusters demonstrated peculiar patterns of genes expression during the course of EBV infection (Fig. 14 and Fig 15). Upon first visual inspection, three clusters appeared to be related (Fig. 14 A, B, C), but in each of these three clusters gene expression peaked at different days post infection. Two other clusters showed high gene expression in uninfected cells (Day 0) followed by a gradual decrease of gene expression during the first four days after EBV infection (Fig. 15 A and B). The two remaining clusters depicted very individual gene expression patterns. The yellow cluster encompassed genes with expression levels that constantly increased over time (Fig. 15 D). Genes in the brown cluster were highly expressed in uninfected cells but dropped on the first day p.i. to rise thereafter (Fig. 15 C). The clustering of the identified differentially expressed genes allowed to reduce the high complexity of the sequencing data and to find a specific expression pattern of host transcripts altered by EBV infection. Next, the defined clusters were analysed if they are enriched for certain biological processes using a common gene ontology (GO) enrichment algorithm. GO studies can identify groups of genes sharing the same biological function and/or pathway in complicated data sets.

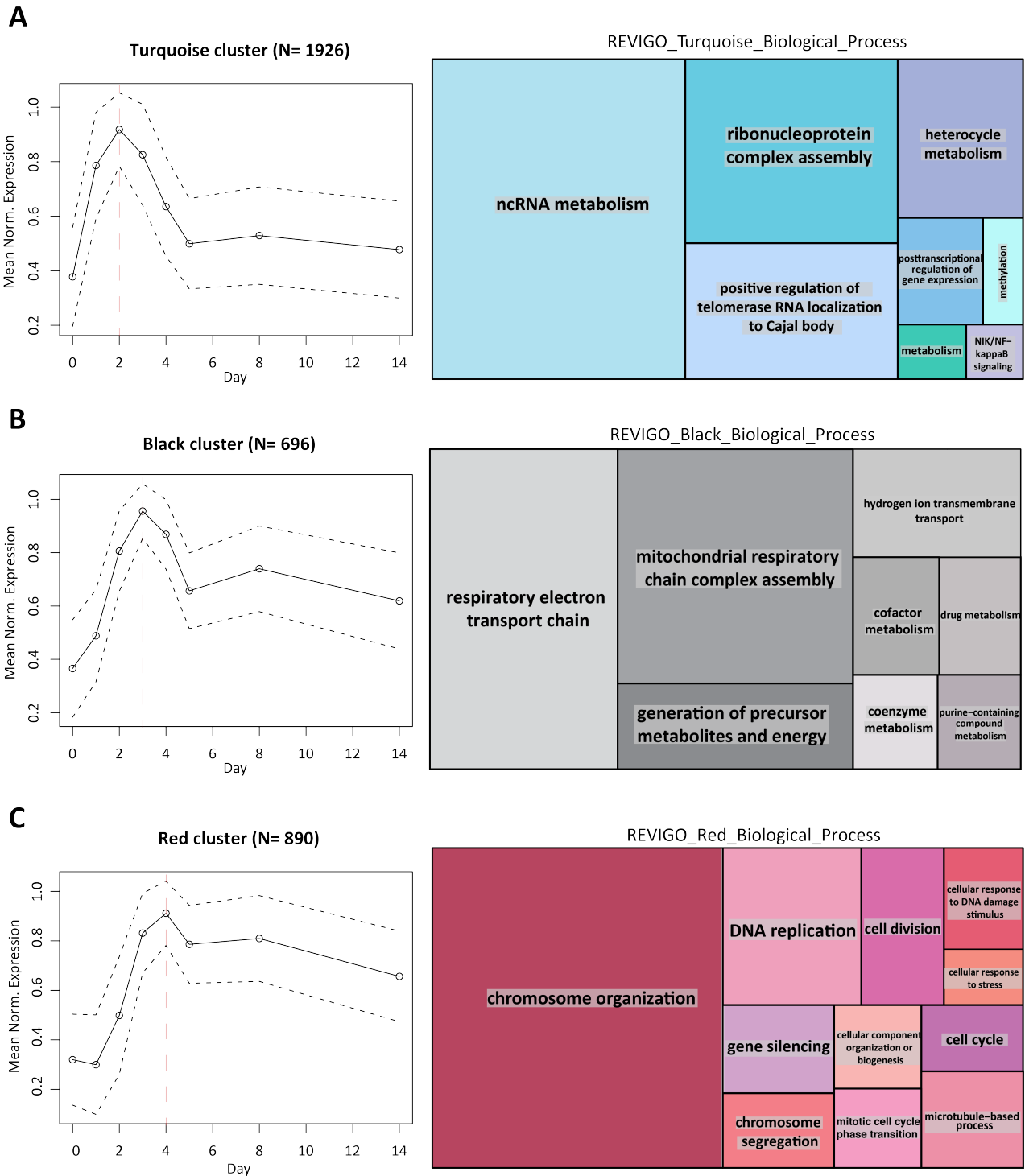
GO analyses can be performed with several freely available online tools and databases, which in principle translate gene lists into biological functions. To characterize specific biological processes, which occur within each of the seven clusters, I analysed the lists of genes from a single cluster in comparison to a set of background genes using the GOrilla online platform (Eden et al., 2009). The union of genes that were tested by DEseq2 across the pairwise comparisons were implemented as the “background genes” set. The GOrilla tool employs rigorous statistics using the minimum hypergeometric (mHG) model to calculate enrichment p-values for each identified GO term. In addition, I set the p-value threshold to  $p \leq 10^{-7}$ . To increase the precision and specificity of the identified GO terms they were further analysed using the “reduce and visualize gene ontology” (REViGO) online tool. This program is designed to work with the long lists of GO terms to remove redundant ones. It also offers different visualization options that can help to display the data in a more transparent and attractive way (Supek et al., 2011). The RNA-seq data enriched GO terms characteristic for a specific gene expression cluster were displayed as REViGO tree maps (Figs. 14 and 15 right panels).

The plots display graphs with the identified enriched GO terms with background colour codes reflecting the name of the cluster. Each diagram consists of tiles representing “super

groups” of related GO terms. The sizes of the tiles reflect the defined p-values as absolute log<sub>10</sub> terms. Each cluster was indicative of very specific biological pathways, which differed dramatically among all seven identified clusters. Three clusters were combined in Fig. 14, because they revealed very similar pattern, but differed with regard to the timing of the peak values. In contrast to this almost congruent pattern, the clusters belonged to very different individual sets of GO terms. Interestingly, the turquoise cluster was enriched in pathways involved in ncRNA metabolism and processing (Fig. 14 A). In line with the metabolic experiments shown in Fig. 7, the black cluster highlighted catabolic processes and biochemical energy metabolism in particular the mitochondrial respiratory chain complex (Fig. 14 B), which peaked at day three p.i.. Genes from the red cluster peaked at day four p.i., when DNA synthesis is way up and the cells divide rapidly (Fig. 5). Very accurately, this cluster encompassed genes involved in chromosome organization, cell cycle processes, and DNA replication (Fig. 14 C).

Four additional, more heterogeneous clusters represent diverse expression patterns that lack clear expression peaks in contrast to the three clusters shown in Fig. 14. The green cluster (Fig. 15 A) is characterized by a rapid and later constant loss of gene expression. The identified processes are typical of uninfected resting cells as the enriched pathways were indicative of cell differentiation and development processes (Fig. 15 A). On the contrary, in the blue cluster, terms characteristic of DNA-templated transcription dominated (Fig. 15 B). The two last clusters summarized biological pathways that are commonly involved in antiviral responses: the brown cluster encompassed leukocyte activation and other systemic immune processes (Fig. 15 C). The yellow cluster showed an increase in pathways regulating type I interferon signalling and responses to common external stimuli (Fig. 15 D).

In summary, my RNA-seq data and their bioinformatic analyses nicely reflected the biological phenotypes of B cells infected with EBV. The data indicated that EBV triggers a plethora of substantial changes in the biology of the entire cell. In very defined and short time intervals, the virus drastically overturns the transcriptional program of resting and quiescent B cells and induces many different biological processes and pathways. Some were anticipated according to the biological phenotypes observed in the course of viral infection such as cell activation, DNA replication, mitosis, etc.. My RNA-seq data confirmed all these processes. In addition, the cluster analyses pointed to additional, unexpected and therefore very interesting processes (and genes involved therein) that warrant further investigation.



**Fig. 14. Cluster analyses of differentially expressed (DE) genes in the course of EBV infection define peak expression pathways and a group of three clusters of biological processes.**

A cluster analysis was performed to identify gene groups with similar expression pattern during infection. In a pair-wise comparison of all cluster sets, a group of significant differentially expressed genes between every possible pair of the days p.i. was found using DESeq2. Across all samples, the Spearman’s correlation coefficient between pairs of genes was calculated followed by a hierarchical clustering analysis in R. The program identified seven clusters that differed in their specific expression pathways in the course of infection. Three cluster pathway schemes are shown on the left side with their arbitrarily assigned colour name codes. The y-axis represents Mean Normalized Expression as a function of time (days p.i.) on the x-axis. The numbers on top of the plots (N) indicate the numbers of the genes that form the cluster. The circles indicate average expression levels, while the dashed lines delineate the standard deviations, the red, dashed lines point to the expression peak of certain cluster. All genes within a given specific cluster were used to perform a gene ontology (GO) analysis using the online



platform ("GORilla", <http://cbl-gorilla.cs.technion.ac.il/>). The GO analysis is summarized in a TreeMap plot (right side) created using the REVIGO online tool that reduces the redundancy in the GO terms. The plots contain rectangles of different sizes indicating the absolute log<sub>10</sub> of the calculated p-values of the shown GO terms. The colour background of the tree maps is related to the colour name of the clusters.

- A. The graph depicts the expression pathway of the turquoise cluster. This group contains genes that peaked on day two p.i., decreased until day five and showed stable levels thereafter. The GO analysis identified the most significant biological processes involved in ncRNAs metabolism, RNA processing and ribonucleoprotein complex assembly.
- B. The black cluster is similar to the previous cluster, but the genes peaked on day three p.i.. Within this group, the most frequently found GO terms were cell metabolism and mitochondrial processes linked to energy production such as respiratory electron transport chain or generation of precursor metabolites.
- C. In the red cluster of expression pathways, genes peaked on day four after EBV infection and decreased slightly thereafter. The analysis pointed to the biological processes in the categories of chromosome organization, DNA replication, cell cycle, and mitotic processes.

The time-controlled expression of EBV genes is another very interesting aspect, which was already described and reviewed in the literature (Price and Luftig, 2014). My RNA-seq data confirmed most of the known expression kinetics of EBV genes (Supplementary Figure 3 in Appendix), but has to be improved since a self-made annotation file of EBV genes was used, which lacks certain viral transcripts. My main interest was to explore an unknown aspect of the early phase of EBV infection, which describes dynamic time-related changes in the host transcriptome upon the EBV infection.

### **3.5.1. ncRNAs and target genes involved in chromatin processing or host cell metabolism during early EBV infection**

As described in the previous chapters, my data analyses identified previously unknown aspects and mechanisms that act in the early phase of EBV infection including changes in metabolic pathways or chromatin dynamics. My transcriptome data revealed the kinetics of specific pathways and allowed the identification of individual targets, which might be key to the success of EBV infection and the establishment of its latent phase later. For example, the RNA-seq data in the red cluster (Fig. 14 C) confirmed phenotypic observations regarding cell cycle entry and cell division on day four (Fig. 5). GO terms on this day highlighted key processes such as the cell cycle, cell division and DNA replication. This example confirmed that NGS data indeed reflected the processes that take place in the cell after EBV infection. The next very important step is to look more closely at interesting clusters to try to define target genes that are essential for the regulation of specific biological pathways. In the next subchapters, I would like to focus on three broad aspects of transcriptomic changes after EBV infection: energy metabolism, chromatin regulation and ncRNA processing.

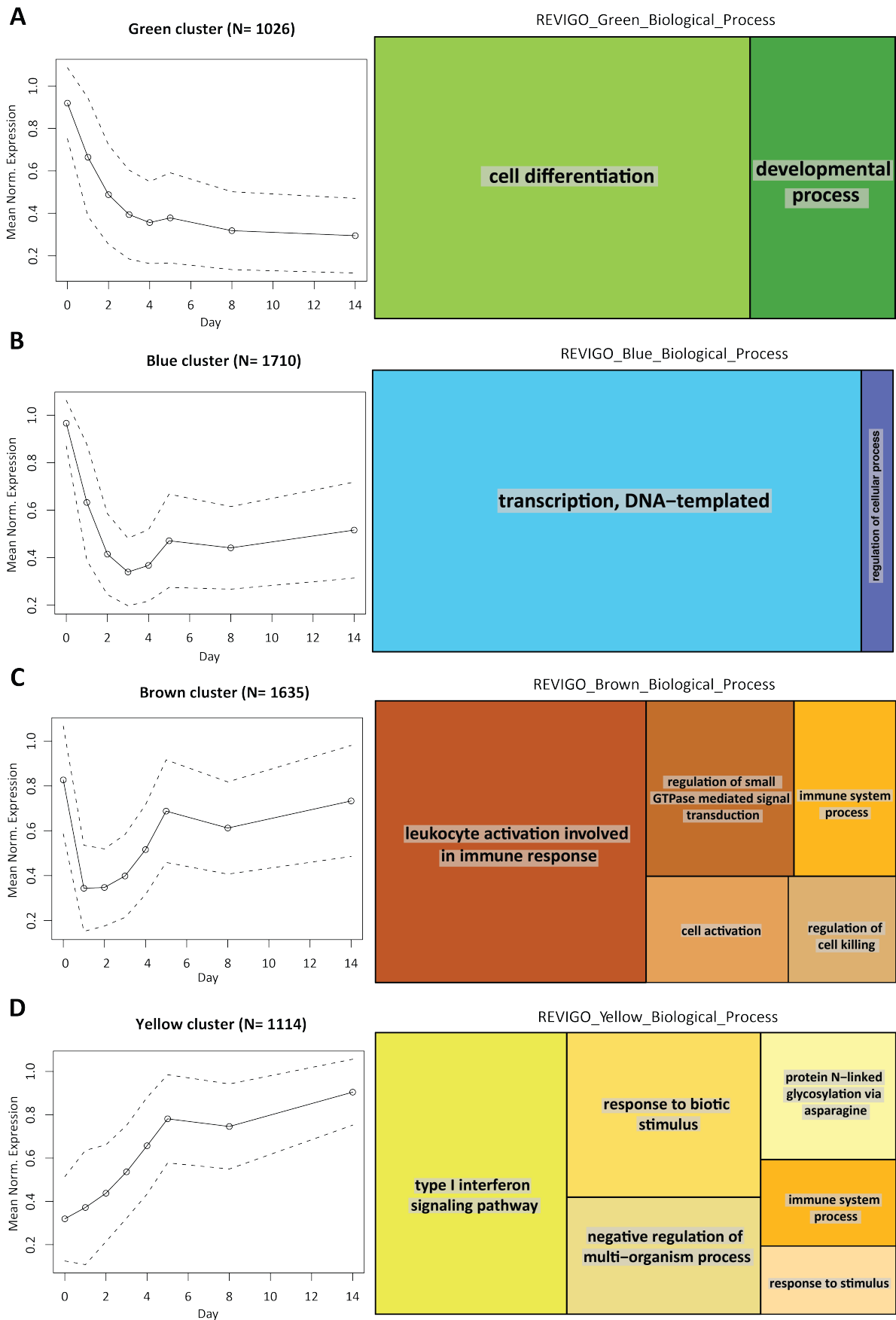
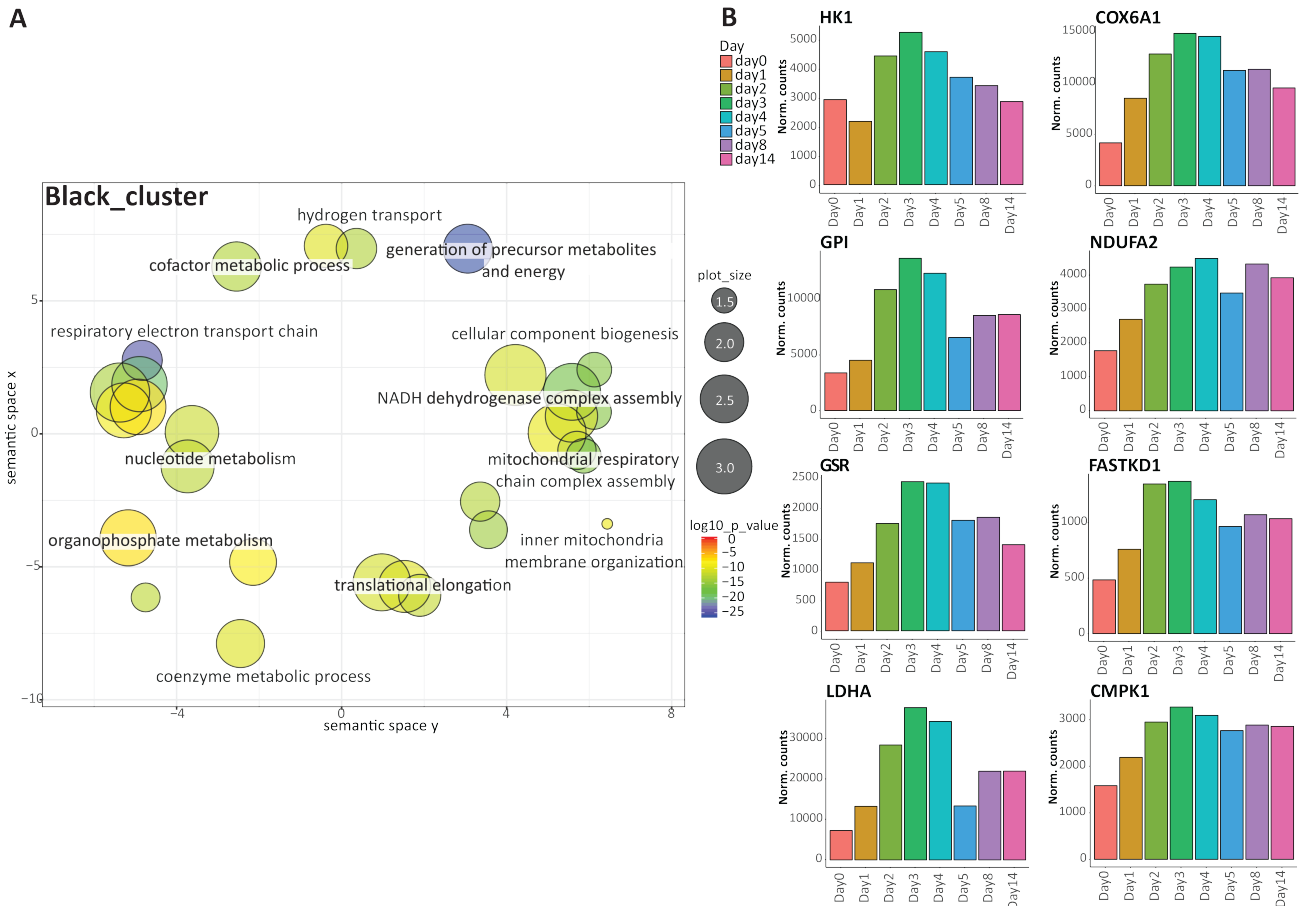


Fig. 15. Four clusters of differentially expressed genes during the early phase of EBV infection with individual expression dynamics. The specific expression clusters were identified as described in Fig. 14.

- A. The graph displays pathway common to the green cluster. It encompasses genes that are highly expressed in uninfected cells but downregulated in infected cells until day four p.i. when the genes adapted a reduced but stable expression level. GO terms enriched in this group of genes indicated processes of cell differentiation and development.
- B. Compared to non-infected (Day 0) cells, a precipitous loss of the gene expression until day three p.i. was obvious in the blue cluster pathway followed by a recovery on day five p.i.. In this group, very broad GO terms were found, which almost all belong to the category of DNA-templated transcription and their regulation.
- C. The group of genes in the brown cluster is also characterized by an initial loss of expression on the first days after infection and a rapid recovery until day five. The GO analysis of this cluster highlighted biological processes related to leukocyte activation, systemic immune processes or regulation of small GTPase mediated signal transduction.
- D. The yellow cluster summarized genes that constantly but slowly increased during infection. Most of these identified genes belong to the type I interferon signalling pathway or are genes that react to various stimuli or are involved in general immunological processes.

### 3.5.1.1. The early regulation of energy metabolism in EBV-infected B cells

The cellular metabolism consists of a very tightly controlled network with numerous pathways. It allows the cell to generate sufficient chemical energy and biochemical precursors that are needed for the different steps in resting or cycling cells. Pathogens also depend on the host cell metabolism as a source for basic precursors, components, or energy that are all required to support viral genome replication and protein synthesis, but also virion assembly inside the cell. Recently, many reports have described different mechanisms that viruses use to manipulate the host cell metabolism for their purposes. Some of them follow similar alteration in the metabolic pathways, which were observed in cancer cells such as the Warburg effect. Many viruses apply also their own, specific mechanisms that differ depending on the type of the viral infection such as latent or lytic infection. Certain aspects of energy metabolism during EBV infection became apparent in this thesis work (Fig. 7). EBV induced the activation of mitochondrial pathways and led to a substantial increase in the uptake of glucose from day four post infection onwards. The transcriptome analysis identified one cluster of similarly regulated genes (black cluster) with GO terms that pointed to respiratory electron transport chain, nucleotide metabolism, NADH dehydrogenase complex assembly, and the generation of precursor for metabolites and energy (Fig. 14 B and Fig. 16 A). These processes form a large group of selected GO terms indicating their abundance. I selected eight different genes, which these GO terms contained (Fig. 16 B). All eight genes followed very similar expression kinetics during the course of EBV infection. They were expressed at low levels in uninfected cells, but increased dramatically until day three p.i.. Thereafter, their expression levels declined and seemed to stabilize later. Four depicted examples relate directly to the very early steps of glucose metabolism. These are hexokinase 1 (HK1), which phosphorylates glucose to



**Fig. 16. Detailed GO analysis of genes in the black cluster and selected examples of regulated genes.**

The set of genes characteristic for the black cluster underwent a GO analysis using the GOrilla tool followed by a reduction of redundant GO terms by REViGO. (A) The graph visualizes the GO enrichment of certain terms using a scatterplot. (B) Selection and visualization of the expression levels (normalized counts) of selected genes.

- A.** The scatterplot displays GO terms as circles in a semantic space, where similar or related GO terms will be arranged closer to each other. Specific colours indicate the log<sub>10</sub> p-values calculated by the GOrilla tool as indicated in the legend on the right. The different sizes of the circles indicate the frequency of the GO term in the underlying database (smaller circles represent terms that are more specific). Three main groups of GO terms that consist of sets of partially overlapping or adjacent circles denote respiratory electron transport chains, NADH dehydrogenase complex assembly, and translational elongation.
- B.** The bar plots show mean expression levels of normalized counts (y-axis) of individual genes in the course of EBV infection (days p.i., x-axis). Each colour is specific for a certain time-point as indicated.

glucose-6-phosphate (G6P) in the very first step of glycolysis or glucose-6-phosphate isomerase (GPI), which converts G6P to fructose-6-phosphate (F6P) (Berg et al., 2002). GSR is a glutathione reductase, which is important in preventing oxidative stress in different types of human cells (Garrett and Grisham, 2005). Lactate dehydrogenase A (LDHA) catalyses the reduction of pyruvate to L-lactate and was reported to play an important role in cancer cells metabolism or during embryonic development (Valvona et al., 2016). The second selected group of four genes contained important components involved in the mitochondrial respiratory activities. The encoded enzymes are key players in the electron transport chain

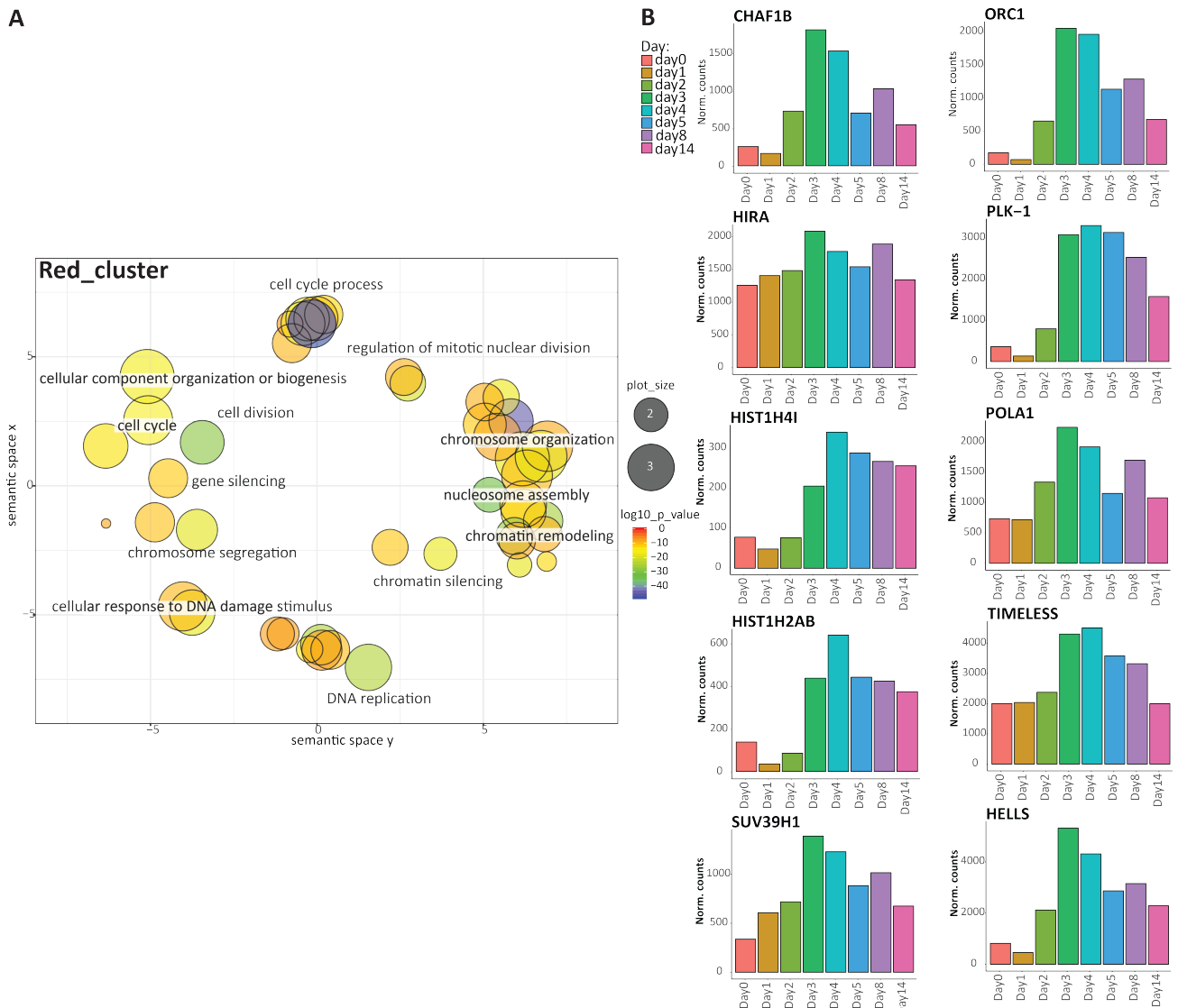
such as cytochrome c oxidase subunit 6A1 (COX6A1) or NADH dehydrogenase 1 alpha subcomplex (NDUFA2). FASTKD1 is the AST kinase domain-containing protein 1, which can modulate the energy balance of stressed mitochondria (Friedman and Nunnari, 2014). On the other hand, cytidine monophosphate kinase (CMPK1) is an enzyme involved in the cellular nucleic acid biosynthesis (Liou et al., 2002). Together, these examples indicate that the EBV infected cells consume a lot of energy and need basic cellular components especially at day three p.i.. Which gene products that are absolutely necessary for the activation and survival of the cell or for the establishment of the latent phase of infection are currently unknown.

### 3.5.1.2. Chromatin components and their regulation during early EBV infection

Chromatin does not only function to package the long cellular DNA molecules into the nuclear space of chromosomes, but chromatin also works as an additional and very important layer regulating global gene expression. This adjustment can be achieved by nucleosomes that occupy and block the access to DNA but also by post-translationally modified histones that mark active or repressed genes. Many processes related to chromatin remodeling depend on certain phases of the cell cycle. Newly synthesized DNA acquires canonical histones during the S phase of the cell cycle. EBV infects resting, quiescent cells, which are in G<sub>0</sub> phase, activates them and induces their DNA synthesis and mitosis.

The red cluster identified on the basis of my RNA-seq data consists of genes that are involved mainly in chromosome organization, DNA replication, cell cycle processes, and gene silencing. Genes found in this cluster showed a very high expression on day three and/or on day four p.i. (Fig. 14 B). The group of GO terms, which are generally involved in chromosome organization (Fig. 17 A) encompasses many different histone proteins such as HIST1H4I or HISTH2AB, which were highly upregulated starting on day four after EBV infection. Histone chaperones were also present in this group for instance chromatin assembly factor 1 (CHAF1A), which is responsible for assembling histone proteins onto newly synthesised DNA, a prerequisite for efficient DNA replication during S phase. Another example of histone chaperones is HIRA, which has a role in placing the histone H3.3 variant onto DNA independent of the cell cycle (Burgess and Zhang, 2010). Histone-lysine N-methyltransferase (SUV39H1) is a very interesting candidate that acts in chromatin silencing and heterochromatin formation. It was also found that SUV39H1 might have a special role in

transcriptional silencing of BZLF in the EBV genome during established latency (Imai et al., 2014). Two smaller groups of GO terms depicted in the scatter plot of the red cluster (Fig. 17 A) are involved in many important steps during DNA replication. The three selected genes act during the early stage of DNA replication: the origin replication complex subunit 1 (ORC1), DNA polymerase alpha (POLA1), and polo-like kinase 1 (PLK-1), which triggers the G2/M transition. PLK1 also has a potential oncogenic function and correlates with cell cycle



**Fig. 17. GO analysis of the red cluster genes and examples of interesting candidates.** The plots were generated as described in Fig. 16.

- The scatterplot displays four large groups of GO terms. A first group of terms indicates functions related to chromosome organization, nucleosome assembly and chromatin remodelling. The second more heterogeneous group encompasses terms with diverse functions in the cell cycle including chromosome segregation and DNA damage responses. A third group relates to terms with a distinct role in DNA replication. The fourth very homogenous group deals with the global term “cell cycle process”.
- The bar plots show the expression levels as the mean of normalized counts of ten genes in non-infected B cells and cells at different days after EBV infection. With the exception of HIRA, the majority of the genes are rapidly activated on the second day after EBV infection. Their levels increased further until day four p.i. and decreased thereafter.

progression (Bell and Dutta, 2002). The TIMELESS gene is mostly related to the circadian clock regulation, but also plays an important role in maintaining many other biological processes such as cell cycle functions, DNA replication and DNA damage response (DDR) (Engelen et al., 2013). The last example is a lymphoid-specific helicase (HELLS), which plays an essential role in regulating the expansion and survival of lymphoid cells during leukemogenesis (Geiman et al., 1998).

Additionally, my RNA-seq data were used by my collaboration partners to predict the cell cycle phases from the RNA-seq data according to a modified algorithm established for single cell sequencing samples (Scialdone et al., 2015). Briefly, this algorithm identifies pairs of genes whose relative expression changes with the cell cycle phase. The obtained data confirmed our biochemical analysis, in which uninfected cells and cells at day one and two p.i. are arrested in the G0/G1, whereas, starting from day three onwards S phase related transcripts were identified, which is in line with our findings (Fig. 5 C; Supplementary Figure 4 in the Appendix).

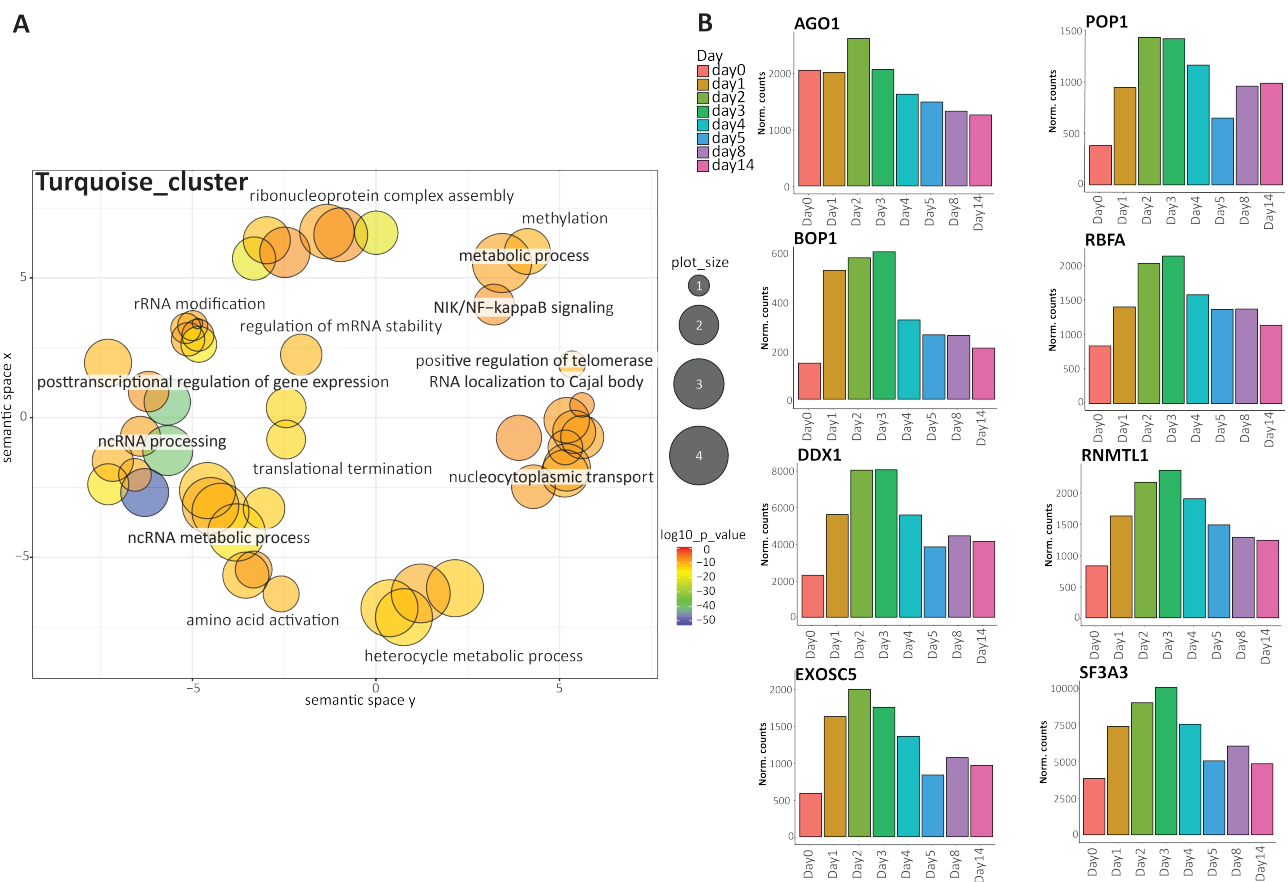
The ten genes exemplified in Fig. 17 B emphasize the important connections between epigenetic regulation and cell cycle progression. Clearly, it is very important for the newly introduced EBV DNA to replicate and to acquire cellular histones with the proper modifications to establish the latent phase in the infected and proliferating host cell. Histone deposition and nucleosome formation on viral DNA will be discussed in more details in the next chapter of this thesis using a different technique – MNase-seq.

### **3.5.1.3. The presumable functions of non-coding RNAs (ncRNAs) during the first days after EBV infection**

My data documented that EBV alters the expression of cellular ncRNAs, which are thought to play essential roles in gene regulation at various levels and to control the cellular machineries. ncRNAs make up a large group of transcripts that include transfer RNAs (tRNAs), ribosomal RNAs (rRNAs), microRNA (miRNAs), small nuclear RNA (snRNAs), small nucleolar (snoRNAs), and long non-coding RNAs (lncRNAs). EBV expresses its own set of ncRNAs, which consists of 44 miRNAs, two lncRNAs (EBER1 and EBER2), and a single snoRNA. Viral miRNAs have a prominent role in viral immunoevasion and the regulation of host genes expression.

Interestingly, RNA-seq data that defined the turquoise cluster of genes (Fig. 14 A) revealed GO terms, which are predominantly involved in ncRNA processing, ncRNA metabolic

processes, rRNA modifications, or ribonucleoprotein complex assembly (Fig. 18 A). The majority of these genes was activated very early after EBV infection already at day one or two p.i. indicating that they might have a crucial role shortly after virus entry. The group of eight genes displayed in Fig. 18 panel B comprises potential viral targets that are involved in important processes mediated by cellular ncRNAs. The first example, argonaute 1 protein (AGO1), is a component of the RNA-induced silencing complex (RISC), which has a function in gene silencing mediated by miRNAs (Okamura et al., 2004; Winter and Diederichs, 2011). The other depicted genes are also involved in general RNA processing such as Exosome Component 5 (EXOSC5), which is a non-catalytic component of the RNA exosome complex participating in



**Fig. 18. The turquoise GO cluster and examples of regulated genes.**

The plots were generated with the workflow described in Fig. 16.

- A.** The scatterplot displays many groups of GO terms dedicated to the regulation of ncRNA processing, rRNA modifications and translation besides NIK/NF-kappaB signalling and metabolic processes involved in heterocycle synthesis and methylation.
- B.** The bar plots show examples of dynamically regulated genes in the turquoise cluster during different days after EBV infection. The majority of these genes peak at day two to three p.i. and decrease thereafter.

maturation of many ncRNA types and degradation of defective transcripts (Vanacova and Stefl, 2007). ATP-dependent RNA helicase (DDX1) has a role in processes regulating RNA secondary



structures to facilitate translation initiation (Cordin et al., 2006; Bléoo et al., 2001). Genes such as BOP1 or RBFA are implicated in the biogenesis of eukaryotic and mitochondrial rRNAs or in ribosome-binding and therefore have a role in the cell cycle process (Hölzel et al., 2005; Rozanska et al., 2017). POP1 is a Ribonucleases P/MRP protein subunit localized and activated in the nucleus during pre-RNA processing (Lygerou et al., 1994). Mitochondrial RNAs are also altered by EBV infection. RNA methyltransferase-like protein 1 (RNMTL1, known also as a MRM3) is a mitochondrial rRNA methyltransferase, which is crucial in the biogenesis and function of the large subunit of the mitochondrial ribosome (Rorbach et al., 2014). The MRM3 protein was recently found to be manipulated by HCMV to support its replication (Karniely et al., 2016). EBV not only impairs transcription or modification of ncRNAs but also splicing events. One of the example is SF3A3, which is the Splicing factor subunit 3A related to the broad aspect of mRNA splicing (Tanackovic and Krämer, 2005).

Obviously, changes in ncRNAs processing are very important for EBV to control the very early phase of viral infection. The virus may use this initial mechanism to alter downstream biological processes. Recently, many reports have indicated that lncRNAs might play important roles in regulating the virus-host cell relationship including epigenetic and metabolic processes. These phenomena could also be very important steps in establishing the latent phase of infection, eventually.

### **3.6. Dynamic nucleosome assembly during the first days of EBV infection**

The nucleosome is a nucleoprotein with a core structure, which consist of an octamer with two molecules from each of the four histone proteins H2A, H2B, H3, and H4. 147 bp of DNA is wrapped in approximately 1.65 turns around it to form the nucleosome. Nucleosomes are separated by linker fragments with a length of approximately 80 bp, which are bound by linker histone 1 (H1). Nucleosomes, DNA and also certain RNAs create a highly ordered structure called chromatin (Mondal et al., 2010; Luger et al., 1997). The main function of chromatin is to organize and condense the long genome segments of the individual chromosomes to pack them into the small volume of the nucleus. An additional, important role of this complex is the tight regulation of gene expression. Nucleosomes are physical barriers for many transcription factors and other proteins involved in DNA transcription. Consequently, chromatin can modulate gene transcription through the regional accessibility of the nucleosome-bound DNA. Yet another important mechanism is histone protein modifications

such as acetylation, methylation, phosphorylation, and others. They act in combination creating the so-called “histone code”, which together with numerous enzyme complexes define the status of gene expression depending on cell differentiation and the biological context (Jenuwein and Allis, 2001). In general, the acquisition of canonical histones takes place in a replication-dependent process during the S phase of the cell cycle. In contrast, histone variants can also be assembled independently of the cell cycle (Henikoff and Smith, 2015). Histone variants also play an important role in the alteration of the nucleosomal structure, its stability, dynamics and consequently gene expression (Weber and Henikoff, 2014) When viruses with double-stranded DNA genomes infect metazoan cells, histone variants might also play a role in the acquisition of chromatin of the incoming viral DNA during latent and lytic infection (Albright and Kalejta, 2016; Placek et al., 2009).

Little is known about nucleosomes on EBV DNA. In the established latent phase, EBV DNA is very densely occupied with histone octamers comprising mostly repressive modifications (Woellmer et al., 2012). During the induction of the lytic phase, the viral chromatin is very dynamically remodelled to ensure the rapid and efficient expression of all viral lytic genes (Woellmer and Hammerschmidt, 2013). A single report found a possible role of the viral tegument protein BNRF1, which can hijack a component of the histone H3.3 variant chaperone Daxx and use it to prevent loading of H3.3 histone variant onto viral DNA and therefore avoid silencing of viral latent genes (Tsai et al., 2014).

One of the goals of my thesis was to study the acquisition of chromatin constituents on EBV DNA and its kinetics during the early phase of viral infection. I employed the MNase-seq approach and could identify and localize mononucleosomes at fixed positions on viral DNA at certain genomic loci very early after infection. I also observed dynamic rearrangements of nucleosomes on certain host genes, which might reflect alterations in gene expression during the pre-latent phase of EBV infection.

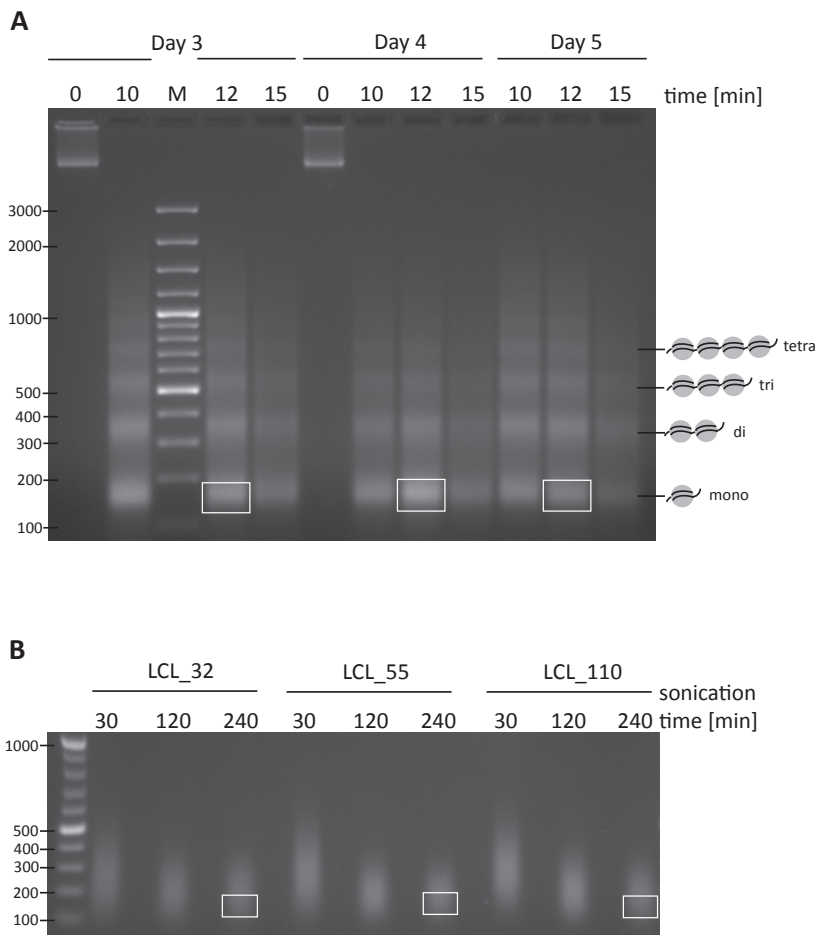
### **3.6.1. MNase-seq experiments identify mononucleosomes on viral DNA very early after infection**

One of the most common and reliable experiments to determine nucleosome occupancy and positioning is the limited digestion of chromatin with micrococcal nuclease followed by a high-throughput DNA sequencing technique. Micrococcal nuclease (MNase) is an endo-exonuclease generating double-strand breaks preferentially in accessible, hence nucleosome-free DNA and inter-nucleosomal linker DNA. After the double-strand break, the enzyme

“nibbles” on the DNA molecule until it reaches a physical obstruction, such as a nucleosomal structure. Therefore, it is a reliable tool to analyse nucleosome occupancy, because MNase cuts only linker fragment DNA, but spares nucleosome-bound DNA. DNA wrapped around the histone octamer is inaccessible and protected from digestion. In the next steps, nucleosomal DNA can be isolated and identified using for instance a next generation sequencing approach. This technique, termed MNase-seq allows to look for nucleosome occupancy or even mononucleosomes positioning using single-base resolution (Wei et al., 2012; Cui and Zhao, 2012). I used my infection model to examine the dynamic chromatin acquisition during the first eight days p.i. with EBV. In the first step, I briefly fixed the cells with formaldehyde to preserve or freeze the protein-DNA interactions by introducing covalent bounds. Chromatin was prepared from the cells infected for the indicated time periods and digested with pre-adjusted concentration of MNase and increasing time. The reaction was stopped by adding EGTA to the samples. Digested chromatin was then de-crosslinked, protein was removed and the DNA was isolated using the phenol-chloroform extraction method. Samples were loaded onto an agarose gel to analyse the partial digestion of chromatin with MNase and reveal mononucleosomal DNA with a length of approximately 147 bp as well as multimers of nucleosomes. Mononucleosomal DNA fragments were cut out and purified from agarose gels (Fig. 19 A). As a control, I used protein-free genomic DNA from established lymphoblastoid cell lines from different donors. Extracted DNA was sonicated to obtain randomly sheared DNA fragments with a length of approximately 150 bp. Sonicated DNA fragments were then isolated after size fractionation on agarose gels (Fig. 19 B). MNase digested and randomly shared DNA fragments were used for library preparation and subsequently sequenced on an Illumina platform. The paired-end reads of 50 bases were mapped to the human and EBV genomes and mapping data were further processed to identify nucleosomal fragments using the DANPOS Python package. At early time-points post EBV infection, detection of the incoming EBV DNA was expected to be challenging, because only a small fraction of the B cells was thought to be successfully infected.

Latently infected established B cell lines harbour several genomic copies of EBV DNA per cell, but the copy number of EBV DNA in newly infected cells is probably in the order of one per cell. Limited MNase digestions were expected to decrease the molar concentration of viral DNA further, because we hypothesized that it is predominantly free of nucleosomes and thus prone to become degraded. This setting makes it difficult to ensure a sufficient read coverage of viral DNA by NGS to obtain sufficient number of reads for mapping and the subsequent

bioinformatics analysis. In well-established LCLs, the fraction of sequencing reads that mapped to the EBV genome was about 0.1 %, in newly infected B-lymphocytes it was in the range of 0.01-0.03 % of the total numbers of reads (Fig. 20 A).



**Fig. 19. Preparation of MNase-seq samples.**

**A.** Chromatin obtained from EBV-infected cells was digested with 1.5 U of MNase for different time periods (10, 12, and 15 min) at 37 °C. After protein removal and DNA purification, samples were separated on an agarose gel and visualized by UV light after ethidium bromide staining. Partial digestion of chromatin on the agarose gel results in a nucleosomal ladder. Mononucleosomes with a length of about 150 bp and multimers (di, tri and tetra) are indicated on the right side of the panel. White rectangles depict samples, which were excised from the gel, purified and used for NGS libraries preparation.

**B.** Protein-free genomic DNA from three LCL samples (LCL\_32, LCL\_55 and LCL\_110) were prepared with a commercial kit followed by two rounds of phenol-chloroform DNA extraction. DNA samples were then sonicated for 30, 120, and 240 min in 10 min intervals using a Bioruptor instrument. Samples were loaded on agarose gels and sheared DNA fragments of approximately 150 bp were excised (white rectangles). DNA fragments were purified and used for NGS library preparation as described in panel A.

Despite the low viral coverage, it was possible to identify reads that indicated single nucleosomes on EBV DNA. The results from a single set of MNase experiments and three independent sonicated LCL control DNAs are shown in the Integrative Genomics Viewer (IGV). Fig. 20 B visualizes specific positions of fragments of nucleosomal length on EBV and cellular DNA. EBV DNA contains several regions with repetitive sequences of various lengths, where reads cannot be unequivocally allocated and which appear to be free of nucleosomes. Hence, these gaps are of technical nature, only.

The global view of the DANPOS results revealed massive changes in nucleosome occupancy on EBV DNA in the course of viral infection (Fig. 20 A, B). On day 0, the non-infected cells, which served as negative controls had no reads on EBV DNA as expected. Starting already on day one after infection, single nucleosomes were present on EBV DNA, but with very small

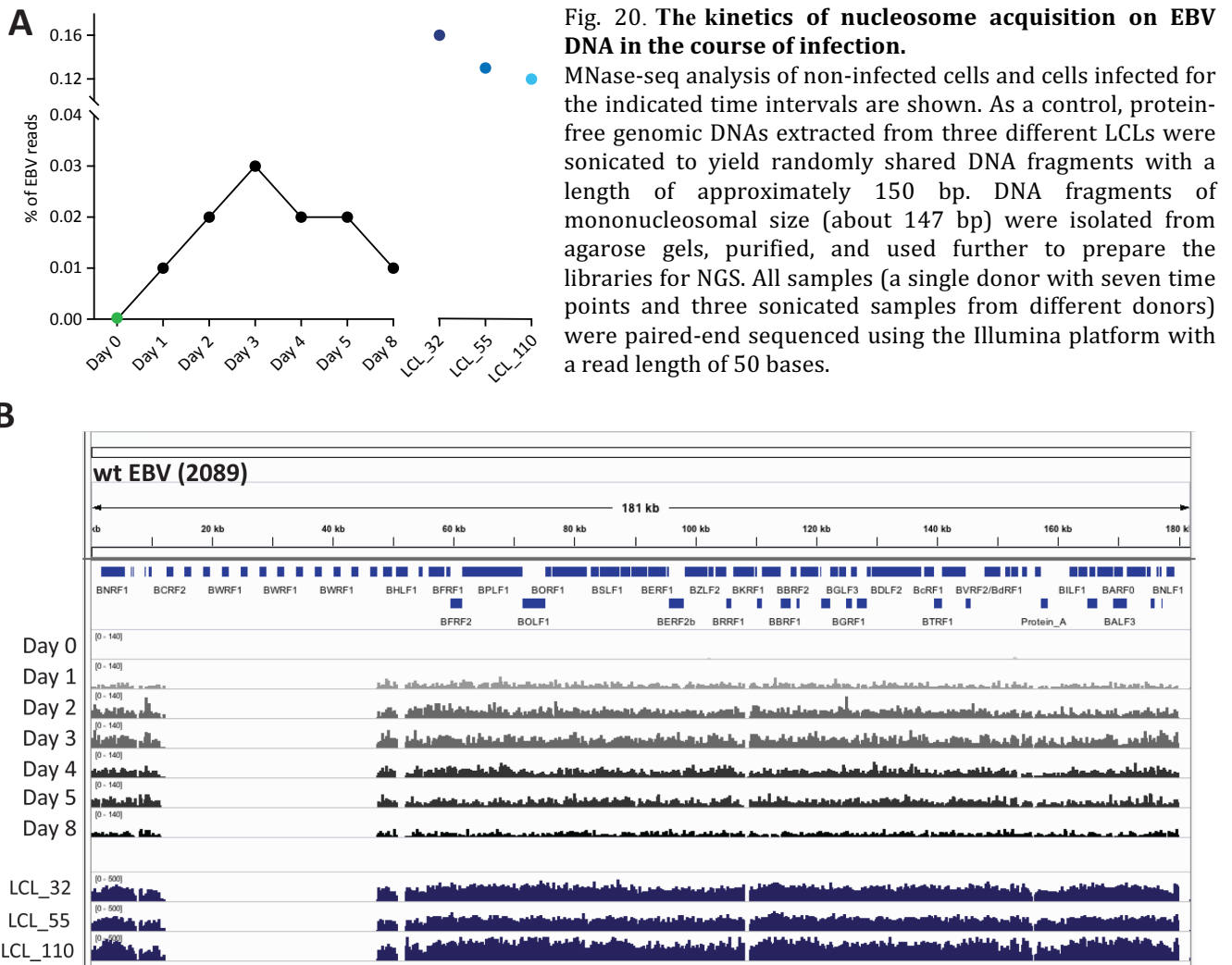
numbers of reads (Fig. 20 A). From day two onwards, mapped reads indicative of nucleosomes increased steadily to reach a stable level at day five p.i. On day eight p.i. fewer nucleosomes were identified as compared with the previous time points of analysis. Controls with sonicated DNA prepared from LCLs showed higher numbers of reads that mapped to EBV DNA.

The length distribution of mononucleosomal fragments were determined and plotted (Fig. 21) to obtain quantitative and qualitative data. The graphs confirmed my previous observations that nucleosomes on the EBV genome were already detectable one day p.i.. They increased considerably on day two p.i. to accumulate later during infection until day five. The analysis of the frequency of length distributions of the fragments clearly indicated a maximum at exactly 147 bp even at early time points of infection confirming the presence and identity of nucleosomal DNA fragments. In contrast, the sonicated samples had wider read distributions with a maximum frequency of length about 125 bp, which reflects a size-selected but rather random population of DNA fragments isolated from the agarose gels (Fig. 19 B). According to these observations, I concluded that the sonicated samples contained randomly sheared fragments, whereas the MNase-digested DNA contained fragments indicative of mononucleosomes with an expected length of 147 bp.

### **3.6.2. Kinetics of nucleosome positioning on viral and cellular DNAs**

My MNase experiment revealed that the assembly of mononucleosomes on EBV DNA was apparent already 24 hours after viral infection suggesting a very rapid process. The nucleosome density on EBV DNA of the infected B cells increased further during the next three days. From the PhD work of Anne Wöllmer in our lab we knew that the acquisition of the nucleosomes at the oriP locus of the dyad symmetry (DS) element in EBV's DNA (Zhou et al., 2005) took place very early after infection. In her work, which was based on a qPCR approach, three nucleosomes accumulated gradually over time to reach a stable level about one week after infection. The well described oriP element served as a positive control in my MNase-seq experiments to identify exactly positioned nucleosomes.

The mapped and normalized MNase-seq results were further used to accurately calculate fragments that can be considered as nucleosomes. DANPOS is a comprehensive, widely used bioinformatic pipeline (Python) that allows to achieve a significant enhancement of nucleosomes occupancy and their dynamics of single-nucleotide resolution (Chen et al., 2013). Evaluated nucleosome peaks after DANPOS processing were displayed in the IGV Browser to



- A.** Percentage of reads mapped on the EBV genome as a function of total read counts. At early time-points, the percentages of mapped viral reads were low but they increased during EBV infection. Reads from randomly shared LCL DNAs amounted to 0.12 to 0.16 % of EBV mapped reads of the total sequencing reads.
- B.** A snapshot of the IGV browser window with the entire wt EBV (2089) genome map. The first track displays an annotation of selected viral genes. The next seven tracks represent results from mapped MNase-seq samples obtained from B cells from a single donor at the indicated day p.i.. The last three tracks show mapped sequencing reads from randomly sonicated LCL DNAs.

visualize and inspect certain selected loci. Within the EBV genome, my MNase-seq analysis confirmed that the *oriP* element was occupied at the known positions and that this binding was clearly detectable within 24 hours p.i.. The three clearly and accurately positioned nucleosomes accumulated further from day two p.i. onwards (Fig. 22 A and B).

I used this approach to investigate whether the acquisition of viral chromatin is a random process or if nucleosomes preferentially assemble at fixed, specific loci, creating regions with nucleosome “hot spots”. It seemed as if nucleosome positioning was not an entirely random

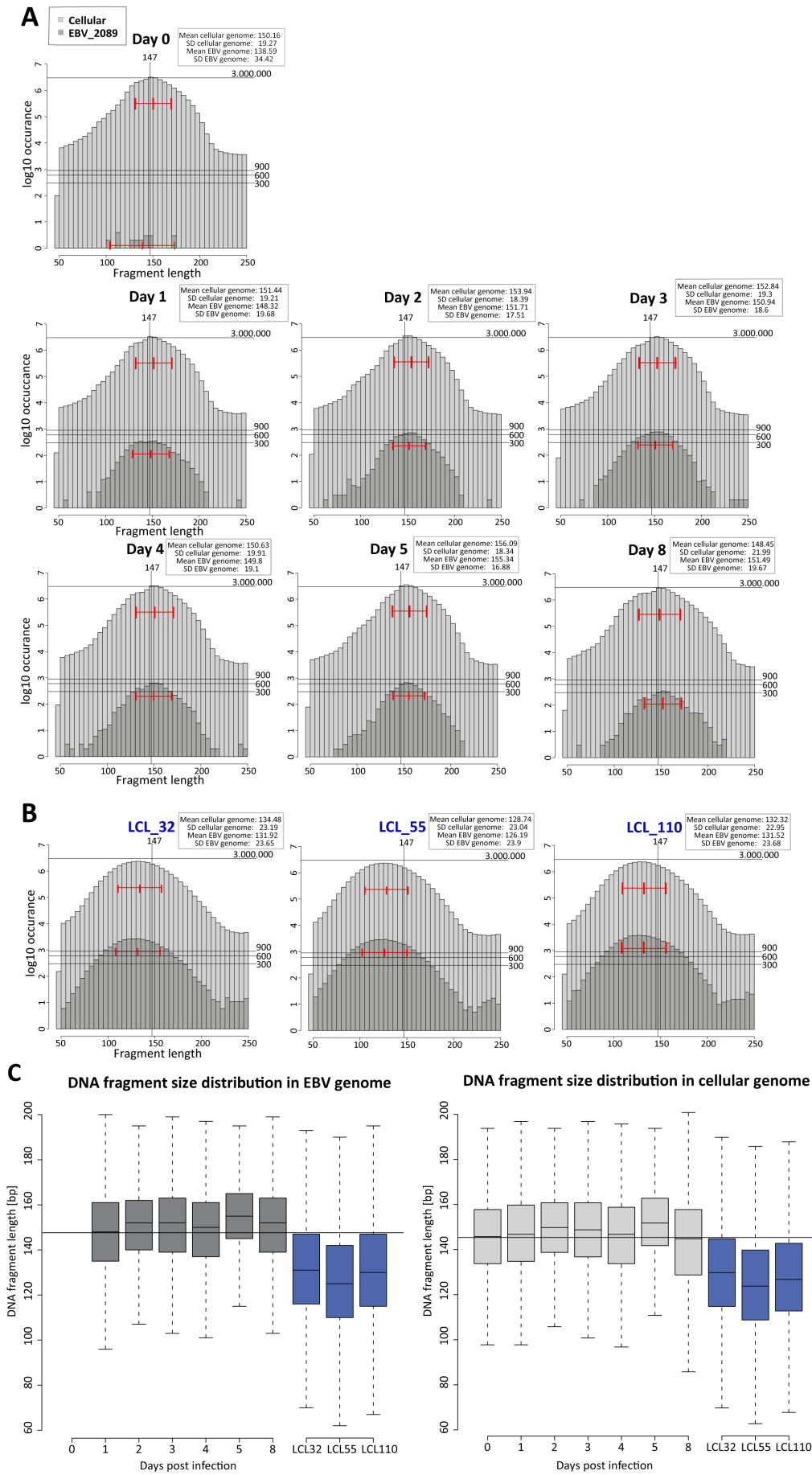


Fig. 21. Distribution of mononucleosomal fragments on EBV and cellular genomes.

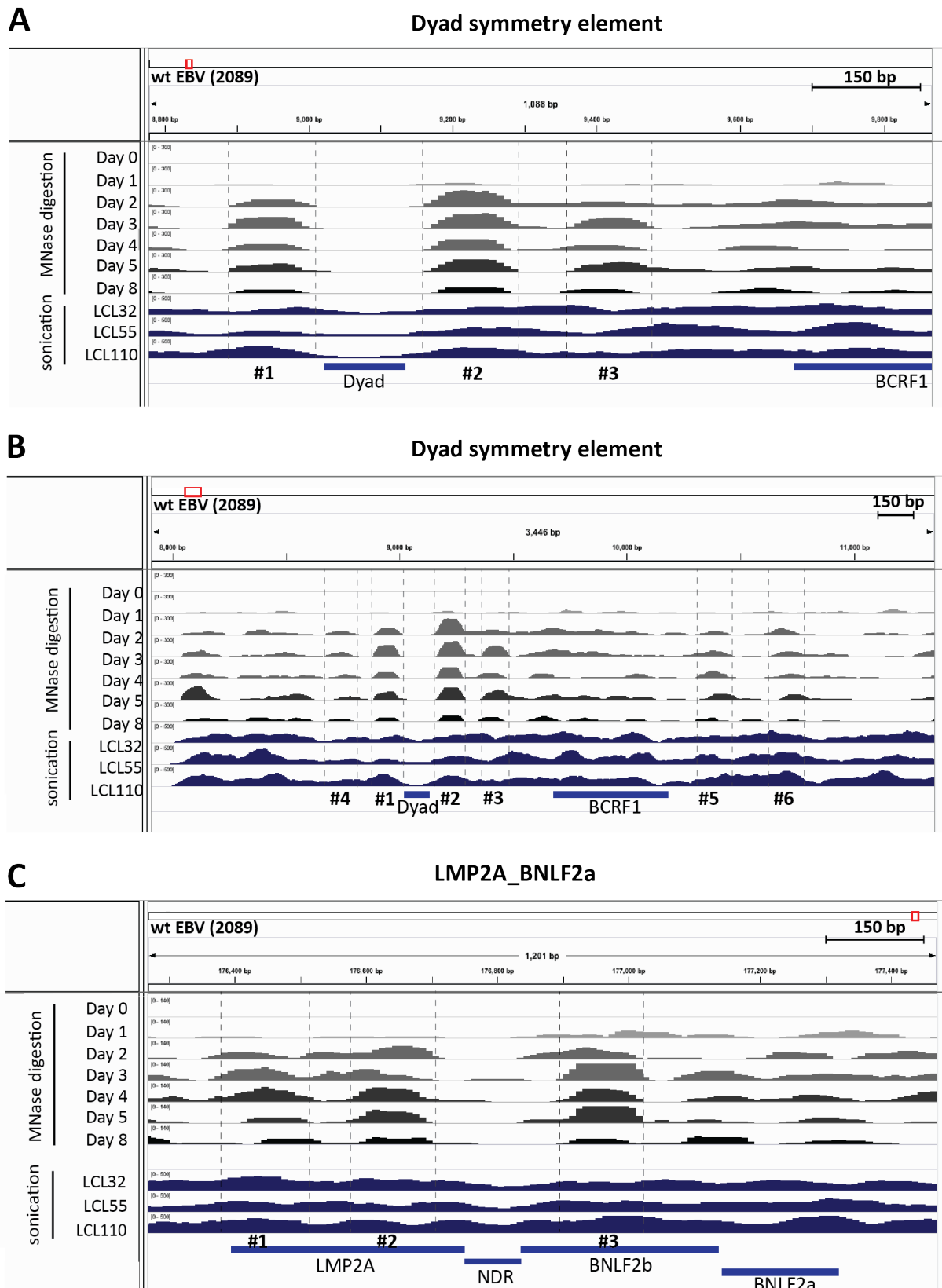
- A. The plots represent the read length distributions in a window of 50-250 bps (x-axis) indicating reads mapped to viral (dark grey) and cellular genomic DNA (light grey) calculated as the log<sub>10</sub> occurrence of mapped fragments (y-axis) within samples collected from different days after EBV infection as indicated. The MNase-digested samples clearly peaked exactly at the expected length of mononucleosomal DNA fragments (about 147 bp), which start to appear already at day one p.i. Red lines indicate mean average and standard deviation of length distribution, the grey boxes contain basic statistic data on the length of cellular and viral DNA fragments.
- B. The three graphs indicated LCL\_32, LCL\_55 and LCL\_110 show the fragment distributions of the control samples after sonication and size fractionation. Randomly shared LCL DNAs exhibited a wider distribution of fragments within the window of 50 to 250 bps with a peak frequency in the range of 125-135 bps.
- C. Summary of the DNA fragment length distribution shown as boxplots. The x-axis displays samples from different days post infection or sonicated LCL samples. On the y-axis, the lengths of DNA fragments are indicated with mononucleosomal fragment of 147 bp depicted as a horizontal line. Each box represents Inter Quartile Range (IQR), which contains 50 % of the data, the top of the box is defined as an upper quartile and bottom of the box is described as a lower quartile. The line within a box is a median of IQR. The upper whiskers define the largest value within the area described as (upper quartile) + 1.5 x IQR and lower whiskers indicate the area with the lowest values, which are (lower quartile) - 1.5 x IQR. Outliers were excluded from the plots. The left and right panels represent data from the EBV and the cellular genomes, respectively.

process. Certain loci acquired single nucleosomal DNA structures rapidly and the initially defined positions were maintained throughout the observation period.

Another example shows an IGV snapshot of the EBV genome with the gene loci LMP2A, BNLF2a and BNLF2b. We knew from our unpublished data that a prominent nucleosome depleted region (NDR) is present between LMP2A and BNLF2b in a latently infected lymphoma B cell line. This NDR at about nucleotide coordinate 176,800 bp is obvious also in my data (Fig. 22 C). Flanking the NDR, well-pointed nucleosomes become visible, which were gradually acquired starting from day two p.i. (Fig. 22 C, #1, 2 and 3). The sonicated LCL DNAs revealed mostly chaotic or inconsistent random fragments distribution covering the entire EBV genome, which supports the correctness and specificity of my MNase data and their interpretation.

The nucleosome positioning on the cellular genome in response to EBV infection was also an intriguing aspect. It was not clear if the virus is capable of altering the host nucleosome occupancy or even their positioning. Nucleosomes could become more dynamic as a consequence of viral infection and the noticed transcriptomic changes or certain viral factors might provoke or induce a reorganization of cellular chromatin. Regions with positioned nucleosomes in the host chromatin were selected and shown in Fig. 23. The *CRIP1* locus had very sharp and clear positions of two nucleosomes upstream from the TSS of the gene, but the nucleosome location at the indicated position #1 was inconsistent during infection (Fig. 23 A). The analysis of the region between two histone genes (Fig. 23 B) showed few, better organized





**Fig. 22. Visualization of positioned nucleosomes on EBV DNA at different days after infection.**

The normalized and mapped DNA sequencing data were analysed with DANPOS to define the nucleosomes and their inspected positions (dashed grey line) on the viral genome. The results are shown as snapshots of the IGV browser looking at three discrete regions of the wt EBV (2089) genome. The bars indicate the scale of a single nucleosome (150 bp).

- A.** The three positions of individual, well-described mononucleosomes (#1, #2, #3) at the DS element of EBV were occupied at very early time-points p.i (day one and two). The density of these nucleosomes increased

further during infection. The sonicated samples (LCL\_32, LCL\_55 and LCL\_110) did not show these clearly defined, MNase-protected regions as expected.

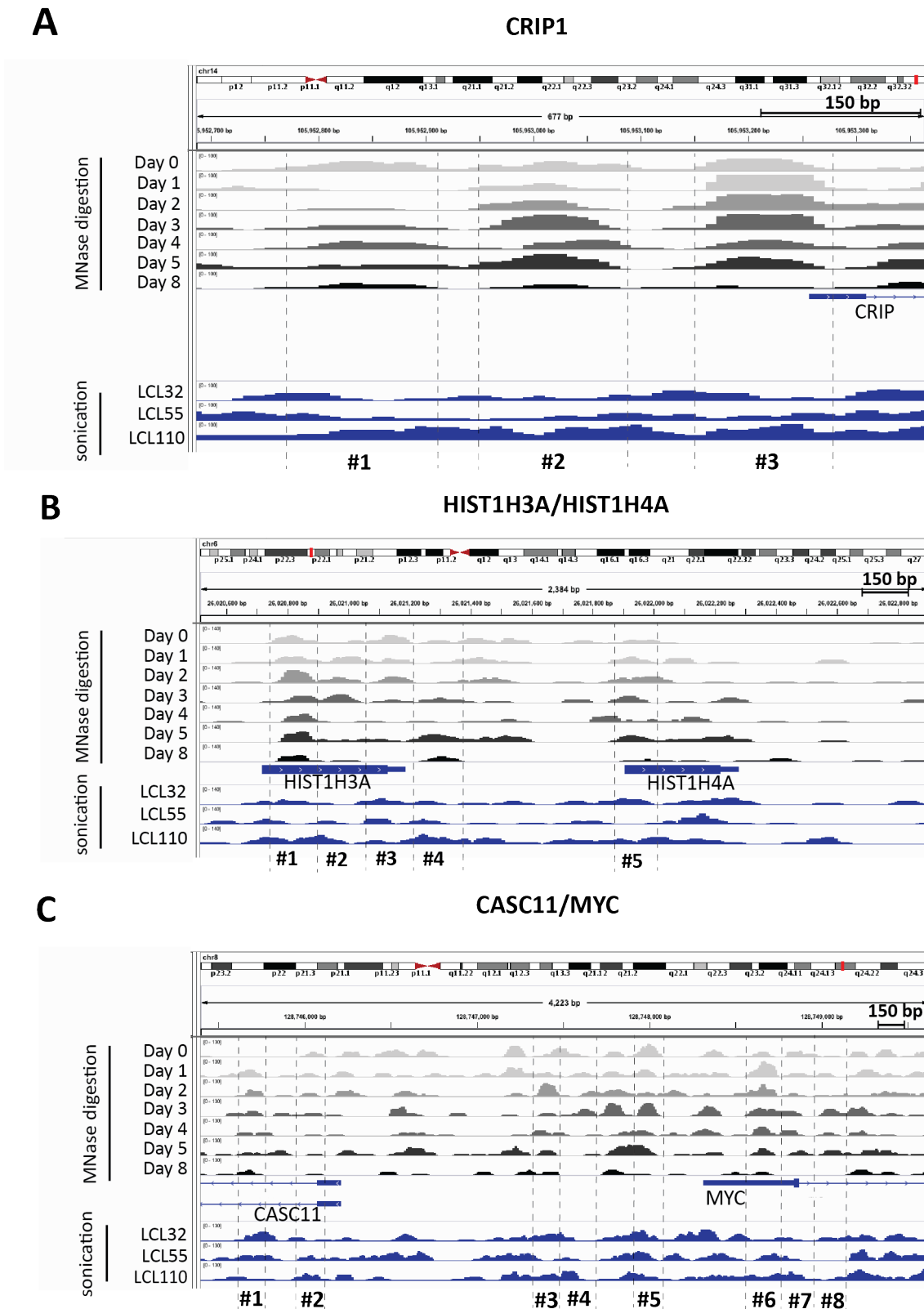
- B. A distant view of the dyad symmetry element and its flanks to compare nucleosome positioning in the vicinity of the DS element. The nucleosomes flanking *oriP* (#1, #2 and #3) have the highest and sharpest defined patterns within a region of around 3500 bp of the EBV genome compared to other indicated but putative nucleosomes (#4, #5 and #6).
- C. A second example of well-defined and positioned nucleosomes (#1, #2, #3) at the LMP2A and BNLF2a loci. Chromatin acquisition increased gradually within the first four to five days p.i..

nucleosomes (Fig. 23 B, position #1, 5) as compared with other positions (Fig. 23 B, position #2, 3 and 4), which seemed to experience a loss or a gain of nucleosomes during late steps of EBV infection. The analysis of the *MYC* locus suggested that there might be dynamic changes of nucleosome positions (Fig. 23 C), but it also appeared as if two depicted nucleosomes of the *MYC* locus were mostly stable during EBV infection (Fig. 23 C, position #2, 6).

Many stable nucleosomes in uninfected B cells appeared to become mobile and seemed to change their position over time. (Fig. 23 C, positions: #1, 3, 4, 5, 7 or 8). From the three selected loci in the host genome, all of them displayed certain changes in the nucleosome positioning. The *MYC* gene seemed to undergo many and various nucleosomes rearrangements, which might be critical for its expression regulation in the EBV-infected B cell. These findings suggest that EBV is able to influence different levels of the cellular processes, having also substantial role in the epigenetic landscape of the host cell.

Summing up, this experiment helped me to study the dynamics of nucleosome occupancy on viral DNA in the early time points after EBV infection. It appeared that the positions of certain nucleosomes on EBV DNA are pre-defined at certain viral loci and that there is a gain of signal strength in the course of infection. Acquisition of viral chromatin seemed to start already on the very first day p.i.. The mechanism that might trigger this process is enigmatic and needs to be investigated in further experiments, because nucleosome formation precedes cellular DNA replication and hence the *de novo* synthesis of most canonical histones.

It is necessary to improve the significance of my results with additional biological replicates, since the MNase-digested data stem from a single donor, only. Biological replicates will help to judge the inconsistent behaviour of the sample collected at day eight after infection. Another aspect, which should be considered is to include additional time points within the first 24 hours after infection to define the exact timing of the onset of nucleosome acquisition on EBV DNA and chromatin rearrangements in the cell.



**Fig. 23. Visualization of nucleosome positions at selected cellular loci during EBV infection.**

The normalized DNA sequencing data were analysed with DANPOS to define the nucleosomes and their positions on the cellular genome.

- A.** The snapshot from the IGV browser, which displays three well-positioned nucleosomes upstream of the *CRIP1* gene. It appears that there are certain changes of nucleosome positioning at location #1 on day one and two p.i., but nucleosomes at positions #2 and 3 were stable after EBV infection. The last three tracks indicate randomly distributed peaks derived from sonicated DNA of LCLs samples.

- B.** A more distant view of possible nucleosome positions in the genomic region between two histone genes (*HIST1H3A* and *HIST1H4A*). Most of the nucleosomes altered their positions during the course of EBV infection, only nucleosomes #1 and 5 appeared to be stable.
- C.** The graph depicts possible nucleosomes that are present in the region of 4500 bp encompassing the *CASC11/MYC* genes. Only two nucleosomes positions (#2 and 6) appeared to be stable, in contrast to the other indicated nucleosomes (#1, 3, 4, 5, 7 and 8).

## 4. Discussion

### 4.1. Scope and aims of my thesis

EBV establishes a stable latent infection in cells that forms a long-lasting virus-host cell relationship. *In vitro*, B cell infection with EBV transforms the B cell and gives rise to infinitely proliferating lymphoblastoid cell lines (LCLs). Many years of research identified multiple viral factors, which mimic cellular signalling processes supporting EBV's long-term persistence in the host cell (Young and Rickinson, 2004; Fields et al., 2007). EBV is able to efficiently evade the host immune responses with immunoevasive cytokine-like viral proteins such as BCRF1 or BNLF2a and by reducing antigen presentation by a set of viral miRNAs (Jochum et al., 2012a; Albanese et al., 2017). Infection with EBV promotes cell proliferation, because most of the viral EBNA proteins bind several cellular transcription factors and induce cellular proto-oncogenes. In addition, at least two cellular receptor pathways of B cells, i.e. CD40 and BCR, are activated by LPM1 and LMP2A, respectively (Münz, 2015; Young, Arrand and Murray, 2007). Much is known about the situation in stable and latently EBV-infected cells, but the initial events that lead to B cell activation and transformation are less well studied and incomplete.

In my thesis, I focused on the alterations in the host cell during the early time-points after EBV infection. I examined the kinetics of the cellular changes, which result in latently infected and immortalized B cells, eventually. My interest was also on epigenetic processes, which regulate viral and cellular chromatin such as nucleosome acquisition and positioning over time. I used a well-established *in vitro* infection model, in which primary B cells obtained from children's adenoid biopsies were infected with wt EBV. Uninfected samples as well as seven samples of infected cells were collected daily on day one to five, eight and two weeks p.i., which provided a broad view of the changes in the biology of the infected cell. My prime interest focused on: (i) phenotypic changes, (ii) metabolic processes, (iii) global transcriptome alterations, and (iv) nucleosome positioning during the first days after EBV infection. High-throughput techniques used for the analyses of gene expression and nucleosome acquisition yielded comprehensive results, which were confirmed by additional biochemical methods. I found very specific and compelling phenotypes of the infected B cells and identified key parameters of their biology, which were mediated by the viral infection in the pre-latent and latent phases of EBV infection.

## 4.2. Phenotypic and metabolic changes during the first days of viral infection

It is known that EBV infection of primary B cells yields bulk or even clonal lymphoblastoid cell lines, but it was not clear when and how this process initiates after EBV infection. In recent years, our lab has investigated the early phenotypes of primary B cells infected with EBV, studying the contributions of the known latent viral genes with various virus mutants. From this study and our published work (Woellmer et al., 2012) we learnt that the viral genes that are essential in LCLs are dispensable in the early phase of infection, which we termed “pre-latent phase” (Pich et al., to be published). Naïve B cells isolated from adenoid tissues were very small, quiescent (Fig. 5 A, B), and arrested in the G<sub>0</sub>/G<sub>1</sub> phase of the cell cycle (Fig. 5 C). In the course of EBV infection, the cells gradually grew in size to become substantially larger (Fig. 5 B) and more granular (Fig. 5 A). The infected cells reached the maximum of their size on day four p.i. and then shrank slightly to a stable level at day five and onwards. These changes strictly correlated with differences in the cell content at different time points post infection. Protein as well as total RNA amounts per cell peaked at day four p.i. (Fig. 5 D and E). At this time point the infected cells entered the S phase of the cell cycle and initiated a short period of cellular hyperproliferation starting on day five (Fig. 5 C). These observations are very comparable to the cascade of events that take place after antigen encounter or artificial activation of the CD3 and CD28 surface molecules of naïve T cells. Initially, inactive T cells are quiescent and non-dividing, but require the cytokines IL-7 and IL-15 for their survival. After activation the cells undergo coordinated changes in their cell size and rates of macromolecules synthesis such as RNA, DNA, or proteins and showed alterations in the patterns of surface immune receptors and adhesion molecules (Teague et al., 1993; Berard and Tough, 2002). T cells respond to the engagement of their T cell receptor (TCR) within minutes, but the process continues for hours afterwards. Activation of naïve T cells induces not only their intensive proliferation but also their subsequent differentiation to i.e. effector or helper T cells. When antigen is no longer available most of the activated T cells undergo apoptosis, but a small fraction of the cells survives and persists long-term as memory T cells (Pennock et al., 2013).

Changes in the cellular biomass such as cell size and cell content result from an altered metabolism of the host cell. Many reports define the metabolic changes, which occur during T cells activation and support the bioenergetically demanding processes of their growth and proliferation. Clearly, metabolic conversions determine the T cell fate. Naïve T cells use only available nutrients to maximize energy production in catabolic pathways such as oxidative phosphorylation (OXPHOS), whereas activated T cells use both OXPHOS and the Warburg effect.

T cell activation induces fatty acid synthesis (FAS) and supports high mitochondrial activity and their biogenesis (Ron-Harel et al., 2016; Almeida et al., 2016). It is also well-known that cancer cells activate many metabolic pathways to obtain rapid access to energy and molecular precursors, which are essential for the synthesis of biomolecules in highly proliferating cells. Hallmarks of the metabolism of cancer cells include the increased transport and consumption of glucose, which is metabolized to lactate under aerobic conditions, a process, which is well known as the Warburg effect, which is far less efficient in ATP production than OXPHOS. Cancer cells have a high glucose uptake and rate of glycolysis, but they also manipulate FAS and engage glutaminolysis (DeBerardinis et al., 2008; Fadaka et al., 2017).

Similarly, many viruses reshape the metabolism of the infected cells during productive lytic infection to gain energy and materials for the replication of the viral genomes, their packaging and release of virus progeny. There are also viruses such as Kaposi Sarcoma-associated Herpesvirus (KSHV) that manipulates the cellular metabolism during the establishment of its latent phase of infection by increasing glucose uptake, FAS and glutaminolysis (Delgado et al., 2010; Lagunoff, 2016). In EBV-associated nasopharyngeal carcinoma (NPC), which is an example of type I EBV latency, LMP1 controls the cellular metabolism, in particular glycolysis (Xiao et al., 2014). My experimental model leads to a different type of EBV latency (type III), which depends on the expression of several latent viral genes including *LMP1*. Together, they govern important host cell processes, which presumably induce also cell metabolism. Interestingly, the expression of LMP1 was reported to be dispensable in the early phase of infection prior to the establishment of the latent phase (Price and Luftig, 2014). Therefore, it is not fully understood what metabolic host cell functions EBV induces during latency as compared with the pre-latent, initial phase of infection. Three scenarios in the latently infected cells could argue for a regulation of the cellular metabolic pathways. (i) Latent viral infection can cause cellular stress, to which the cells respond by alterations in the cellular metabolic pathways. (ii) Latently infected cells can switch to the lytic productive phase, which requires synthesis of many metabolites. (iii) To support long-term persistence in the host cell, the virus needs to modulate cellular metabolic activity. EBV seems to pursue third strategy by altering several metabolic pathways.

In my PhD work I tracked changes in certain basic metabolic processes during the course of EBV infection. I used samples from uninfected cells and daily collected samples of infected B cells to investigate the uptake of the glucose analogue (2-NBDG) and mitochondrial activation using TMRE, which is sequestered by active mitochondria. Similar to the uninfected state, the

cells did not seem to be highly metabolically active during the first three days after EBV infection. The uptake of glucose analogue 2-NBDG did not differ (Fig. 7 B and 8), nor did the cells show an increased mitochondrial activity (Fig. 6 and 7 A). In uninfected cells and in cells infected for up to three days TMRE incorporation revealed two distinct populations, TMRE\_high and TMRE\_low (Fig. 6). The TMRE\_low population was always represented by small, probably quiescent cells, which might suggest that they were not infected with EBV and were prone to die. Alternatively, these cells could be delayed in their response to viral infection and became fully active later. After day four and onwards, the cells grew and entered the phase of rapid proliferation, which was linked to the strong activation of all analysed metabolic processes. Around day eight the TMRE\_low population was significantly reduced, whereas the TMRE\_high cell population increased rapidly to reach a level at which the majority of the cells had fully active mitochondria. A similar kinetics was observed analysing the uptake of glucose analogue. After day four the cells increased the uptake of glucose and probably other nutrients necessary for the intensively proliferating cells (Fig. 8 and 7 B). My observations revealed that B cell infection with EBV promotes substantial phenotypic changes, which are comparable to those observed during the activation of T cells.

In principle, the metabolic phenomena in activated B and T cells do not differ, but the timing does: antigen encounter of T cells induces the rapid proliferation of T cells within hours, but B cells infected with EBV take four days to enter their proliferating mode. The molecular differences are unknown, but seem to be physiological. Naïve B lymphocytes activated by CD40 receptor engagement in combination with IL-4 start to proliferate also about three to four days post induction, although both stimuli are genuine signals in B cells and mediated by T helper cells during a germinal centre reaction (Pich et al., unpublished data).

#### **4.3. Global changes in the host transcriptome as a consequence of viral infection**

I wanted to look at global changes in central cellular processes at early time points after EBV infection. The kinetics and alterations in gene expression prior to and after EBV infection were investigated using time-resolved RNA-seq experiments. I used my infection model and collected the identical numbers of physically intact, uninfected, and quiescent B cells as well as cells infected with EBV at the indicated time points (Fig. 10) from day one to five, day eight, and day 14. RNA was extracted from the cell samples and used to prepare suitable libraries followed by NGS sequencing. The sequencing data were very carefully controlled and entered detailed



bioinformatics analyses. My emphasis in analysing the RNA-seq data was on the dynamics of the changes triggered by EBV and the active processes during the early time points of B cell transformation.

During stable latent infection, the cells are reprogrammed by the pathogen to retain the viral genome and successfully establish a long-term infection. There are many examples in the literature that describe differences that occur after viral entry altering the expression and regulation of protein-coding genes as well as genes encoding ncRNAs. THP-1 cells latently infected with Human Cytomegalovirus (HCMV) display a significant number of genes that are assigned to pathways involved in apoptosis, inflammatory response, and the cell cycle, which seem essential during infection. Additionally, many known and newly identified non-coding RNAs were found up- or downregulated in the HCMV infected cells indicating their distinct roles during the latent phase establishment (Zhang et al., 2016; Cheng et al., 2017). The closest relative of EBV, KSHV induces changes in the host cell transcriptome during the latent infection, which are very much related to changes observed in conditions of hypoxia including glucose metabolism and ncRNAs including human miRNAs (Viollet et al., 2017, 2015).

There are numerous reports, which characterize the changes in the viral transcriptome during infection or in the already established latent phase (Purushothaman et al., 2015; Sijmons et al., 2014). The studies describe the kinetics and the role of EBV genes during the process of B cell immortalization with a special focus on the genes encoding the EBNA, LMPs and ncRNAs (Price and Luftig, 2014; Young and Rickinson, 2004; Humme et al., 2003). However, it is still not fully understood how EBV reprograms the B cell and which/when cellular pathways are induced, modified, or regulated by viral infection.

#### **4.3.1. The dynamics of the cellular pathways**

Studying the kinetics of transcriptomic changes is important to understand the time-related interactions that take place between the pathogen and the host cell. Interpreting the alterations in gene expression of the infected cell are crucial to identify the molecular components and biological processes that coordinate the initial mechanisms supporting viral persistence in the infected B cell, eventually. There are no reports to my knowledge that describe the kinetics of the host transcriptome as a response to EBV infection.

The PCA analysis of my time-resolved RNA-seq experiments suggested that the most dramatic variances exist between resting, uninfected B cells and cells infected with EBV (Fig. 12 A, B). The distant position of the non-infected compared with the infected samples indicated

that there is a well-defined transcriptome characteristic of quiescent B cells, which is immediately lost upon EBV infection probably as a direct response to viral infection and the activating signals. Lymphocytes in a quiescent state are non-dividing cells of small sizes with reduced macromolecular synthesis and energy requirements but elevated stress resistance (Yusuf and Fruman, 2003). My data also suggest that the quiescence state of the cells is clearly not an arbitrary transcription program, but is actively and well controlled. There are several reports describing T and NK cells and the regulation of their resting states, but the exact factors and mechanisms that determine the state of quiescent B cells are not comprehensively investigated.

My PCA analysis identified a number of genes, which were highly expressed in the uninfected cells and probably define the resting state of naïve B cells (Fig. 12 C). Certain genes have been already described to have a role in maintaining the dormant state of B cells such as *ZBTB16* or *KLF4* (Fig. 12 C), which were initially identified in murine and human B cells to be strongly downregulated upon activation. Additionally, the ectopic expression of *KLF4* and *ZBTB16* in proliferating or memory B cells lead to their cell cycle arrest in the G1 phase or their delayed entry into the S phase and cell division (Yusuf et al., 2008; Good and Tangye, 2007). The role of *CXCR4* was suggested to be an universal control switch regulating the quiescence in various somatic stem cells (Nie et al., 2008). In B cells, *CXCR4* was found to be downregulated upon different B cell stimulation factors such as IL-21 or CD63 (Yoshida et al., 2011). *SOX5* was recently defined as a transcription factor (TF), which limits the proliferation of B cells but promotes plasma blasts differentiation (Rakhmanov et al., 2014). Thus, it seems that EBV downregulates a number of cellular genes such as *SOX5* to support B cell proliferation and/or differentiation. EBV must have a special interest to downregulate certain factors that drive B cell differentiation to maintain the balance of the virus and its host cell and to prevent the onset of the lytic phase of EBV or antibodies synthesis. Both processes, B cells differentiation to plasma cells and antibodies synthesis, would be incompatible with EBV's latent phase of infection. It is speculative but most genes, which were found in the green cluster (Fig. 15 A) and are highly expressed in uninfected B cells but downregulated immediately after EBV expression, might predominantly act as "quiescence factors" to maintain the resting state of the B cells.

Samples collected from EBV-infected B cells at different time points after infection contributed to variances defined by PC1 (Fig. 12 A, B). Clearly, samples from day one to three differed from each other as compared with samples collected on days four to 14, which appeared as an almost homogenous population. Genes that were defined by PC1 variances

encompass various genes with more diverse biological functions. For example, *PIR* and *SPARC* have a role in cancer progression, whereas *TMEM173* modulates innate immune responses (Fig. 12 D). There is also a prominent representative of lncRNAs, *LOC643401*, which was recently identified to suppress p53 expression (Li et al., 2017) that might have an obvious role in the activation and induced proliferation of B cells.

The comparison of the transcriptomic profiles at different days after infection was an intriguing point of my research. I wanted to investigate how the infection promotes specific alterations in the host gene expression and when the most dramatic changes in the transcriptomes occur. I applied a visualization tool to identify the intersecting sets using the UpSet package for R (Fig. 24). This tool is based on the “set theory” and replaces the better known, conventional Venn diagrams to visualize numerous data sets (Lex et al., 2014; Conway et al., 2017). The UpSet tool employs a matrix-based layout to display differences and intersections from multiple data sets. I identified the genes differentially expressed between different days post infection using the DESeq2 package (Love et al., 2014). Genes are called if their differential expression was found to be statistically significant according to a model used in DESeq2 comparing any combination of two different time points (days p.i.). I observed very specific changes in the gene expression at different days after EBV infection (Fig. 24). The largest group with more than 8000 DE genes was identified comparing uninfected cells with cells infected for one day (Day0\_1). The pairwise comparison of the following days p.i. (Day1\_2, Day2\_3 etc.) identified a rapid decrease in the numbers of DE genes over time. This finding suggested that the most profound changes in the host transcriptomes occur within 24 hours post infection.

I analysed the pairwise sets of DE genes in more detail using the UpSet tool, to visualize their intersections. Briefly, each row of the UpSet plot shows the number of DE genes from the pairwise comparison. The columns in the matrix visualize the indicated intersections and the bars in the graph depict the cardinality, which is the number of elements in a given set of intersections (Fig. 24).

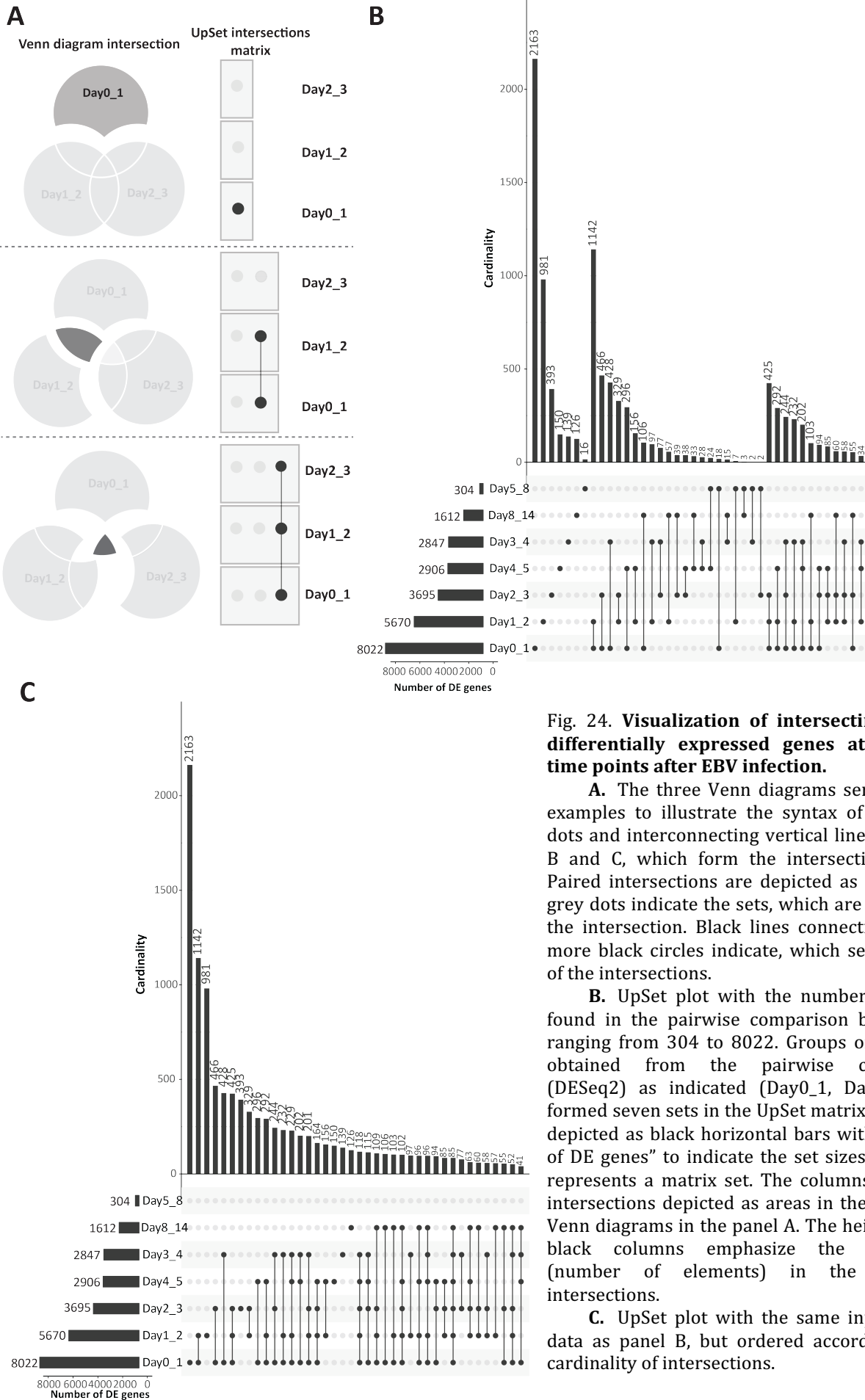
The viral infection causes a profound and immediate shift in host gene expression and further modulates it to strictly control the biological processes in the infected cell. This interaction is supported by the analysis shown in Figure 24. Intersections between all data sets are virtually not existing, but the UpSet matrix points to the early time points after infection that differ substantially from later time points. The Day0\_1 set of genes contained 2163 genes, which were unique for this pairwise comparison and did not intersect with any set of later time points.

This unique group of genes might govern the switch from the quiescent to the activated B cell state. The intersection between the two comparisons Day0\_1 and Day1\_2 of DE genes displayed the second highest cardinality (n=1142) indicating that the data sets are still related in terms of their DE genes (Fig. 24 B), although there are many unique genes within each of these two comparisons. Day1\_2 and Day2\_3 shared a considerable number of DE genes (n=329), but the pairwise intersection of later time points identified gradually lower numbers of intersections. These results suggested that the infection with EBV triggers dynamic changes in the host gene expression within the early days p.i. (day one to four), but later modulations of the host transcriptome were less extensive. It seems as if already at day five p.i. the latent phase of infection is fully implemented.

#### **4.3.2. Biological processes identified during B cell transformation**

Dynamic changes in gene expression indicated that various biological processes are altered during the early phase of EBV infection. To decrease the complexity of the RNA-seq data, DE genes were grouped into sets, which displayed similar expression profiles during the course of EBV infection. I observed seven different clusters encompassing very individual, time-dependent patterns of gene expression (Fig. 14 and 15). Three of the seven identified clusters showed very similar expression pathways in principle, but differed with respect to the individual peaks of maximal gene expression (Fig. 14). Very specific biological processes were enriched in the three individual clusters. The clusters correlated perfectly with the phenotypes I identified with my different techniques confirming the recent literature (Price et al., 2017; Hatton et al., 2014).

Cell proliferation and division were found to be activated starting on day three to four p.i. (Fig. 5 C) along with an activation of cellular metabolic pathways (Fig. 5). Genes identified in the red cluster referred to chromosome organization, DNA replication, and cell division. Moreover, this cluster represented genes, which are involved in the chromatin acquisition such as histone genes, histone chaperones, and chromatin modifiers, which might indicate important epigenetic alterations including chromatin silencing in the cellular and perhaps also viral chromatin at day four p.i. (Fig. 17). The black cluster peaked on day three post EBV infection and encompassed a large group of genes indicative of common mitochondrial activities, respiratory chain complex assembly and genes involved in the glucose metabolism (Fig. 14 B and 13). These processes likely are the consequences of the high energy requirements of the cell and correlate with the



**Fig. 24. Visualization of intersecting sets of differentially expressed genes at different time points after EBV infection.**

**A.** The three Venn diagrams serve only as examples to illustrate the syntax of the single dots and interconnecting vertical lines in panels B and C, which form the intersection matrix. Paired intersections are depicted as black dots, grey dots indicate the sets, which are not part of the intersection. Black lines connecting two or more black circles indicate, which sets are part of the intersections.

**B.** UpSet plot with the numbers of genes found in the pairwise comparison by DESeq2, ranging from 304 to 8022. Groups of DE genes obtained from the pairwise comparison (DESeq2) as indicated (Day0\_1, Day1\_2, etc.) formed seven sets in the UpSet matrix, which are depicted as black horizontal bars with “Number of DE genes” to indicate the set sizes. Each row represents a matrix set. The columns show the intersections depicted as areas in the schematic Venn diagrams in the panel A. The heights of the black columns emphasize the cardinality (number of elements) in the indicated intersections.

**C.** UpSet plot with the same input sets of data as panel B, but ordered according to the cardinality of intersections.

metabolic assays (Fig. 7), which showed significantly increased glucose uptake and mitochondrial activity from day four p.i onwards. Activation of cellular metabolic pathways is important during the cell growth preparing for the subsequent intensive cellular proliferation. Many reports identified alterations in the metabolic transcriptome to be an important step during viral infection. For example, HCMV regulates the expression of glucose transporter (*GLUT4*) during infection, induces mitochondrial biogenesis and respiration presumably mediated by several viral factors (Yu et al., 2011b; Karniely et al., 2016; Sanchez and Lagunoff, 2015).

The turquoise cluster, which showed a peak of gene expression at day two p.i., identified pathways in ncRNA processing and ribonucleoprotein complex assembly (Fig. 14 C and 15). Genes in this cluster were involved in a variety RNA-coupled processes such as pre-RNA processing, RNA helicases, RNA methyltransferases, splicing factors, and exosome components needed for RNA degradation (Fig. 18). In the last years, the research focus on non-coding RNAs increased substantially and a large number of ncRNAs, especially long ncRNAs (lncRNA) have been identified to encompass very diverse and essential regulatory functions in almost all biological processes (Fortes and Morris, 2016; Liu and Ding, 2017). Recently, several lncRNAs were found to play a role during viral infection introducing repressive epigenetic marks via the PRC complex (Flynn and Chang, 2014) or DNA methylation (Zhao et al., 2016), which might both stabilize the latent phase of viral infection. lncRNAs supposedly play a role in regulating viral replication (Ge and Lin, 2014), which is advantageous, because RNAs are normally not recognized by the antiviral mechanisms and immune responses of the host cell much in contrast to viral proteins. My data support the view that EBV induces cellular ncRNAs and RNA processing factors very early after infection, which might be crucial for preparing the cells for the subsequent steps such as hyperproliferation, metabolic alterations, or epigenome changes. Many EBV-induced ncRNAs, which I found, lack appropriate annotations, which make their identification and functional analysis uncertain. More work is needed to unravel their importance during the pre-latent, but also during the latent phase of EBV's life cycle.

My RNA-seq analysis identified two clusters with high gene expression in non-infected cells, which were found to be downregulated during EBV infection (Fig. 15 A and B). In the green cluster, genes with discrete functions in cell differentiation were prominent, whereas the blue cluster identified a group of genes referring to DNA-templated transcription. EBV suppresses the differentiation of latently infected B cells into plasma cells by epigenetic mechanisms triggered by EBNA3A and 3C (Styles et al., 2017). In this recent report, EBNA3A and EBNA3C

can mediate the removal of activation-associated histone marks and promote gene repression inducing H3K27me3 marks via the EZH2 component of the PRC2 complex. Thus, during latency, EBV can prevent B cell differentiation to favour long-term latency and enable virus persistence. In my study, I observed the very rapid and early downregulation of several cellular genes involved in differentiation, which might indicate the presence of additional mechanisms, which EBV induces in the newly infected B cells.

The downregulation of cellular genes in the blue cluster point to a viral function that might repress certain antiviral factors, which could counteract B cell transformation. A similar function was reported in productively HSV-1 infected cells. HSV-1 can negatively regulate the host RNA PolII transcription in the majority of cellular genes, which requires several immediate early genes of HSV-1 (Spencer et al., 1997). It is currently unclear if EBV factors are directly involved in transcriptional repression of selected cellular genes and how this mechanism might work.

Two gene clusters found in my RNA-seq data showed a rather constant gradual increase in gene expression as a response to EBV infection. The yellow cluster included functions with respect to type I interferon signalling and responses to external stimuli, the brown cluster included pathways with genes primary involved in leukocyte activation (Fig. 15 C, D). EBV initially suppresses an interferon-mediated immunity by the tegument protein LF2, which interacts with the central domain of interferon regulated factor 7 (IRF7) (Wu et al., 2009). Unpublished genetic analysis from my lab challenge this view but point to viral miRNAs, which regulate the interferon response in EBV-infected B cells. On the contrary, it seems as if LMP-1, a latent EBV protein, may promote the interferon production later during infection, which can have a role in the establishment of the latent phase of infection (Xu et al., 2006). These processes nicely correlate with my RNA-seq data and are in line with the reported kinetics of the type I interferon signalling pathways. Infection of naïve B cells with EBV leads to cell activation identified in the brown cluster (Fig. 15 D), which includes genes that promote cellular proliferation and establishment of the latent phase (Calender et al., 1987; Hatton et al., 2014).

A recent report covers alterations in the host transcriptome and proteome during KSHV infection and identified global changes in the cellular networks that are essential for virus latency (Sychev et al., 2017). The authors identified alterations in approximately 17 % of the host transcriptome and 13 % of the cellular proteome that includes also dramatic changes in the phosphoproteome 48 hours after KSHV infection of endothelial cells. This analysis identified multiple cellular pathways highlighting the complexity of the host-virus interactions.

Cell activation mediated by BCR or TCR signalling promotes a large set of transcription responses such as the recruitment of RNA PolII to promoters, the acquisition of H3K4me3 activation mark, or the binding of mutually exclusive transcription factors to boost cellular gene expression (Garruss and Fowler, 2015; Fowler et al., 2015). The virus might mimic the general processes in B cells that occur during activation of the B cell receptor (BCR) upon antigen encounter and modulate the different steps during the activation of the infected B cell according to its requirements. This can be achieved by the time-dependent expression of certain viral latent genes (Price and Luftig, 2014), which might have a role in regulating cellular gene expression.

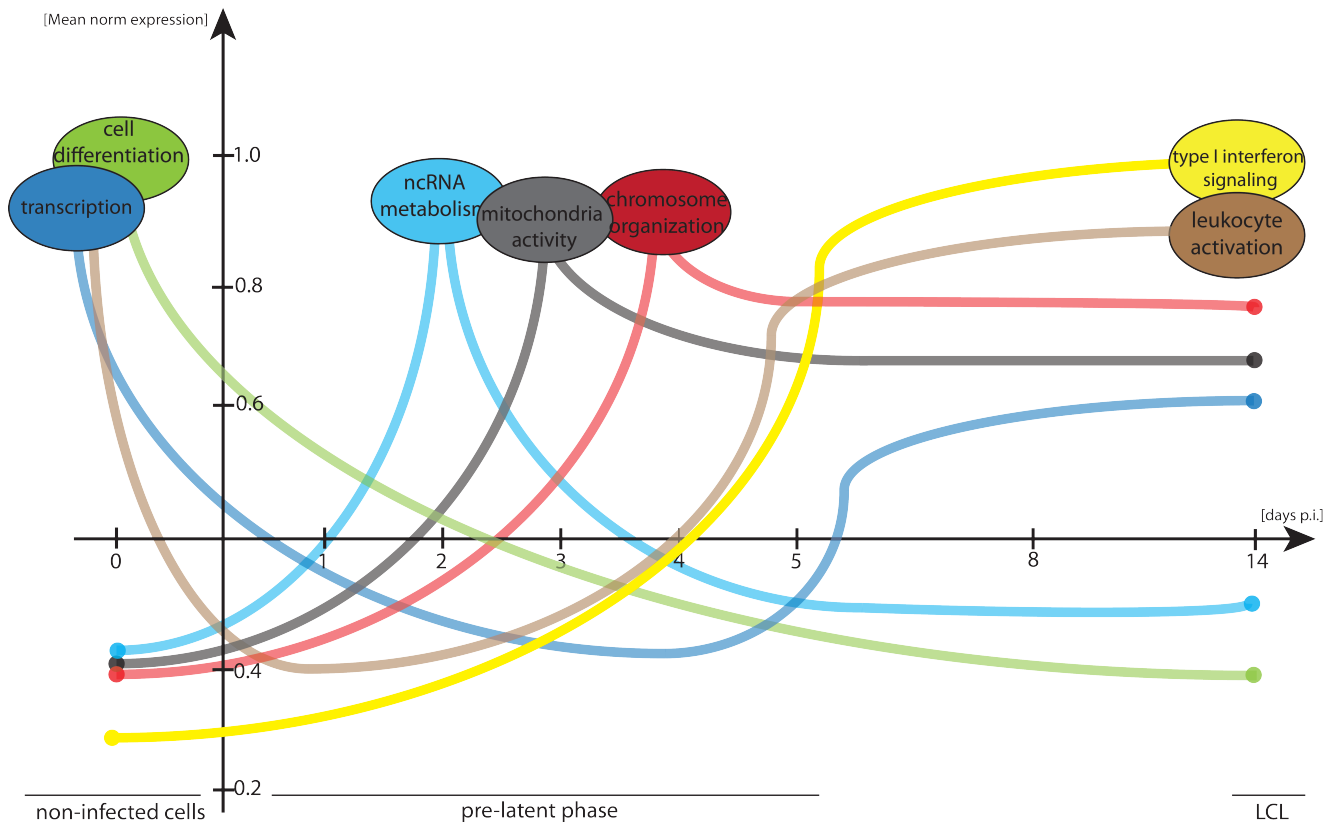
My comprehensive RNA-seq experiment allowed me to examine the very dynamic changes in the biological processes, which EBV regulates during the early phase of infection. The results are summarized schematically in Fig. 25 indicating that the virus regulates the majority of cellular mechanisms in a very strictly time-controlled mode to promote the establishment of the latent phase of infection. My data provided a deep insight into the transcriptomic profiles of the host cell that reflect the complex interplay between EBV and the infected B cell.

#### **4.4. Chromatin alterations in the course of viral infection - nucleosome occupancy and positioning**

##### **4.4.1. Viral chromatin assembly**

EBV enters the B cell with a naked, epigenetically naïve double-stranded DNA, which in the course of establishing the latent phase acquires nucleosomes with mostly repressive modifications and gains a high level of CpG methylation (Woellmer and Hammerschmidt, 2013; Woellmer et al., 2012). These events are thought to be essential for the virus to successfully persist in the host cell and escape from latency to later induce the lytic phase. How and when exactly EBV acquires nucleosomes is neither known nor understood. In my PhD work I was interested in the kinetics of nucleosome acquisition on viral DNA at early time points post infection. Three scenarios of viral DNA chromatinization in the course of infection appear plausible: (i) The virus can enter the cell, circularizes its DNA and waits until the cell will start DNA replication and synthesis of the new histone proteins, which can then be used and assembled on the viral genome. (ii) The virus directly acquires histone variants, which can be synthesized independent of the cell cycle (Weber and Henikoff, 2014) and uses cellular chaperons to deposit histones on viral DNA, or (iii) the virus can make use of the relatively small





**Fig. 25. Schematic overview of seven discrete biological processes at early time-points after EBV infection.**

The figure schematically summarizes the results obtained from RNA-seq experiments and their bioinformatic analyses. The x-axis indicates the time points (days p.i.), whereas the y-axis displays mean normalized expression values, which were used in the clusters created in this thesis and shown in Fig. 14 and 15. The coloured lines represent the expression pathways of the seven individual clusters with their arbitrarily assigned cluster names. The ellipses point to the peaks of individual cluster and provide the most prominent GO terms of the depicted clusters.

pool of free and available nucleosomes, which stem from nucleosome eviction and the continuous exchange of histones during cellular transcription or (iv) histones that might be directly mobilized by the viral infection (Adam et al., 2014; Venkatesh and Workman, 2015). It is evident, that cells infected with EBV start their DNA replication at the end of day three p.i. (Fig. 5 C), which would concomitantly foster the synthesis of new canonical histones (Lipford and Bell, 2001). It is not clear how the virus copes with this situation. Anne Wöllmer from our lab tried to address this question in her PhD work using a qPCR approach and three well-defined nucleosomes flanking EBV's dyad symmetry (DS) element (Zhou et al., 2005). She found that the three highly ordered and exactly positioned nucleosomes are gradually assembled at around day two to three days post infection and that this process is completed within one week.

I used her data as a positive control for my high-throughput MNase-seq analysis with freshly infected B cells collected at different days after infection. I wanted to study not only the kinetics of nucleosomes assembly on viral DNA, but also to understand the mechanism of the

initial chromatinization of viral genes. In my experiments, nucleosomal occupancy of the EBV genome seemed to be a very rapid process, which started already on the first day and gradually increased until day four p.i. (Fig. 20 B, C). High nucleosome occupancy correlates with gene silencing (Arvey et al., 2012), which is required during latency, when the majority of viral genes has to be repressed.

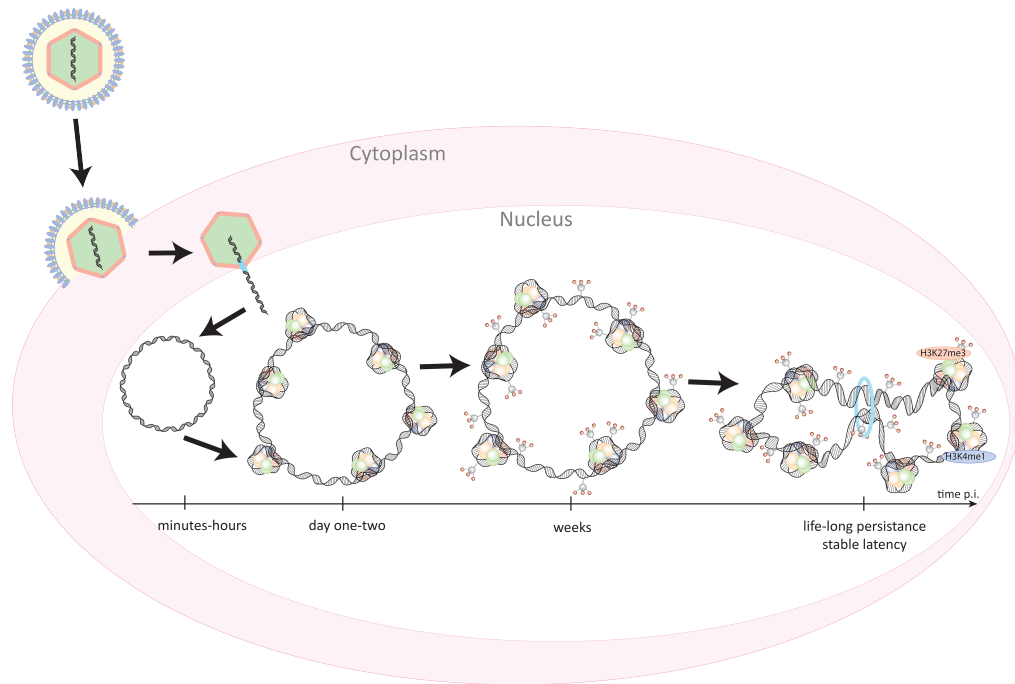
My results are in line with published work on nucleosome acquisition during herpesviral infection. During its lytic phase of infection, HSV-1 acquires histones as fast as 3 hours p.i., a process, which is mediated by the Asf1 histone chaperone (Oh et al., 2012). Later reports using this acute model of lytic infection also identified differences in the mobilization of the canonical H3.1 histone and the histone variant H3.3. Initially, the histone variant H3.3 binds to the HSV-1 genome, but histone H3.1 molecules are subsequently acquired after the initiation of HSV-1 DNA replication (Conn et al., 2013). This finding suggests an important role of H3.3 histone in the very early phase of HSV-1 infection. The rapid association of herpesviral DNA with nucleosomes seems to play an important role during the lytic as well as latent infection with HCMV. The HCMV genome successively acquires replication-independent and replication-dependent histone proteins. Additionally, both H3.1 and H3.3 are deposited on HCMV DNA in both types of infection indicating possibly similar functions of histone variants in regulating the expression of the viral genome (Albright and Kalejta, 2016; Lee et al., 2011).

Nucleosome assembly is a general mechanism of epigenetic regulation, which also depends on the exact nucleosome positioning on the DNA molecule. It is not clear, which mechanisms EBV employs during the rapid chromatinization of its genome: (i) nucleosome acquisition can be a random process, which modulates or even controls silencing of the viral genome independent of the exact position. Alternatively, (ii) nucleosomes might be preferentially and actively positioned on pre-defined viral loci creating so-called nucleosome “hot spots”. Later during infection, the “hot spots” may serve as platforms from which the residual genome is organized to acquire chromatin and to complete the process. Using my MNase-seq approach, I was able to identify several exact nucleosome positions on EBV’s genomic DNA (Fig. 22). Upon EBV infection, I observed nucleosomal footprints on certain loci already on day one, which became more pronounced and stronger in the course of infection. The three nucleosomes flanking the control locus with the DS element confirmed that the DANPOS program can authentically and reliably identify nucleosomal DNA. These results suggested that nucleosome positioning on viral DNA at certain loci take place at specific and well-defined positions. It seems as if the localization of nucleosomes is important during the early phase of

infection, which warrants functional studies in the future with better resolution. Additionally, I found that histone H3.3 protein levels are not stable during different days p.i.. The levels peaked around day three together with the histone variant chaperones Daxx and HIRA, which displayed the highest protein expression between day two and four p.i. (Fig. 9). These results support a functional role of the histone H3.3 variant in EBV's pre-latent phase, but did not explain how complete nucleosomes form and become deposited after the first day of infection. The HIRA and Daxx chaperones were identified as important players in the acquisition of histones onto viral DNA during HCMV, HSV-1 but also EBV infection (Rai et al., 2017; Tsai et al., 2014). Tsai et al. identified EBV's tegument protein BNRF1 to replace ATRX and bind to the Daxx histone variant H3.3 chaperone disrupting the Daxx-ATRAX complex. As a consequence, BNRF1 is expected to interfere with silencing of viral latent genes, which is mediated by loading H3.3 onto the viral genome in the early phase of infection. BNRF1 thus seems to ensure the expression of important viral latent genes. This mechanism appears to be counter-intuitive in the pre-latent phase of EBV infection, when most latent viral genes are well expressed (Supplementary Figure 4 in the Appendix) but it might be important for the virus to escape from latency later. More effort is needed to investigate the function of BNRF1 and its effect on H3.3 binding on viral DNA in the early days of infection.

Nucleosome assembly on DNA is the central epigenetic mechanism, but post-translational histone modifications also play an important role in the regulation of the viral gene expression (Bannister and Kouzarides, 2011). Repressive histone modifications are characteristic of latent herpesvirus infections. For example, HCMV maintains a low-average histone occupancy in the early stages of infection, but it increases further during infection and is associated with the process of viral DNA replication (Nitzsche et al., 2008). Nothing is known about histone modifications in the pre-latent phase of EBV infection, indicating a neglected aspect.

In my RNase-seq analysis, the red cluster enriched for chromatin organization processes peaking at day four p.i., which does not correlate with my MNase-seq results. The initial acquisition of nucleosomes to viral DNA is a much faster process suggesting that the viral chromatin is established prior to cellular DNA replication, which begins on day three post infection. Timing is an essential component of biological processes. Viral genome circularization, chromatinization and methylation are time-controlled mechanisms, which are very important for the viral persistence in the host cell (Fig. 26).



**Fig. 26. Model of the epigenetic modifications of the EBV genome during the pre-latent phase of infection.** EBV enters the B cell with a linear, epigenetically naïve DNA, which immediately circularizes. It seems that during the first days p.i., the virus acquires nucleosomes and probably also histone modifications. Cytosine methylation of viral DNA is a slower process, which is completed after about three weeks p.i.. Additionally, in the course of infection, viral chromatin acquires a three-dimensional conformation, which together with a dense nucleosome occupancy and specific histone modifications ensure the life-long persistence of the virus in the host cell (adapted from Buschle, 2017).

#### 4.4.2. Cellular chromatin rearrangements as a response to viral infection

Chromatin plays an important role during the different phases of viral infection, because gene expression of both the virus and the host cell are controlled by the chromatin structure, accessibility, and local chromatin constituents. Pathogens employ factors and mechanisms that alter viral chromatin (Knipe et al., 2013). It is still intriguing whether viruses can also modulate the host cell chromatin, which might result in the activation or silencing of specific cellular genes or affect cellular genes globally.

Certain nucleosomes at selected cellular loci showed early dynamic changes upon EBV infection (Fig. 23). There are very stable nucleosomes at specific positions in the course of infection, but certain nucleosomes seemed to be evicted, shifted or even gained at different time points p.i.. These observations suggest that cellular nucleosomes are regulated in response to viral infection. A very dynamic nucleosome mobilization took place at histone genes (Fig. 23 B) and around the *c-Myc* locus (Fig. 23 C). Changes in the chromatin organization might be a prerequisite to allow the dynamic activation of cellular gene expression following the infection with EBV. My MNase-seq experiments have to be repeated with more biological replicates to

confirm my finding, but they also indicate that additional intermediate time points should be analysed very early after EBV infection at 6 and 12h p.i..

Viruses have developed mechanisms to regulate cellular chromatin by hijacking cellular enzymes and molecular machineries such as histone acetyltransferases, histone deacetylases, and chromatin remodelling components (Wei and Zhou, 2010). HSV-1 is able to induce chromatin remodelling via the viral transactivators ICP0 and VP16, which recruit critical chromatin modifying factors such as SNF2h, BRG/BRM, or CHD3. They are involved in the high dynamics of viral and cellular chromatin in HSV-1 infected cells, thus regulating their gene expression and DNA replication (Conn and Schang, 2013; Arbuckle and Kristie, 2014). Hernando et al. identified global changes in chromatin accessibility and a decrease of three repressive marks (H3K27me3, H3K9me3, and H4K20me3) upon transformation of B cells by EBV (Hernando et al., 2014). These processes are specific for EBV-induced B cell transformation, because the stimulation of resting B cells with CD40L and IL-4, which promote their proliferation does not cause a substantial decrease of heterochromatin marks. This group also identified dynamic histone modifications and altered histone protein levels in the course of EBV infection as well as local changes in chromatin structure at selected loci using limited DNaseI and MNase digestions. In general, this earlier report supports my findings, but did not investigate nucleosome positions and the dynamics of chromatin remodelling (Fig. 23), which is more sensitive, but requires more and comprehensive experiments for completing this task.

B cell activation itself was also identified to trigger global changes in cell chromatin. It was shown recently that stimulated B lymphocytes can decondensate chromatin via different mechanisms, which involve induced global acetylation and ATP-dependent de-compaction. It was found that some of these mechanisms strictly dependent on *c-Myc* activation and a continuous energy supply (Kieffer-Kwon et al., 2017).

My MNase-seq experiments suggest that EBV triggers also changes in the global cellular chromatin structure. Mobilization of cellular nucleosomes might be used by the virus to assemble chromatin on its own genomic DNA or is used as yet another level of regulating cellular gene expression to support B cell transformation and the establishment of the latent phase of EBV infection.

#### 4.5. Open questions and outlook

This thesis contributed to a better understanding of the global changes in the host B cell, which take place and are triggered by EBV infection. My experiments indicated that the virus

alters a plethora of cellular processes including host cell metabolism, the cellular transcriptome, and epigenome. These results gave rise to another set of questions, which have to be investigated and answered in more detail in the future.

#### **4.5.1. Which specific metabolic processes are altered during EBV infection?**

The results from this PhD thesis suggested that the metabolism of glucose and the mitochondrial activity are modulated in the course of early infection with EBV. It will be intriguing to examine whether EBV alters other metabolic processes such as FAS, Warburg effect, and glutamine metabolism similar to other viruses such as KSHV (Delgado et al., 2010; Sanchez et al., 2017; Lagunoff, 2016) or whether EBV-infected and activated B cells induce pathways that are characteristic of T cell activation (Almeida et al., 2016) or cancerogenesis (DeBerardinis et al., 2008; Biswal et al., 2017). EBV might trigger very individual metabolic alterations, which are essential for its persistence in the host cell. Additionally, it is not known whether specific viral factors such as EBNA2 or LMP1 are involved in the changes of the host cell metabolism or whether the cell activation mediated by viral infection is sufficient to modulate biomolecules synthesis and energy production.

#### **4.5.2. Which cellular and viral genes identified in my RNA-seq experiments are crucial for the early phase of EBV infection?**

My RNA-seq experiments identified very specific biological pathways, which were regulated during the indicated, early time-points after EBV infection. The aim of this thesis was to look at changes, which occur in the host cell transcriptome, but it is also interesting to look at the kinetics of viral genes in the course of the pre-latent phase of infection. The RNA-seq data have to be re-mapped to a newly created wt EBV (2089) genome, which will provide more reliable and precise information about time-related viral gene expression. This will help to confirm existing data (Price and Luftig, 2014) and to identify viral genes, which might be crucial at certain time-points during viral infection. In addition, experiments with different EBV mutants can identify the viral factors that are essential and/or contribute to the regulation of specific host genes. The recent development of single-cell transcriptome analyses will likely increase our knowledge about the exact changes and cell fate pathways at the level of single cells (Grimes et al., 2015; Athanasiadis et al., 2017). To confirm that the global changes on the cellular transcriptome are specific effects triggered only by EBV infection, the next experiments should be conducted with additional controls. *Ex vivo* isolated B cells do not survive under the

cell culture conditions for more than 48 hours. My experiments included the transcriptomic profiles of these uninfected B cells, but did not consider the possibility that a fraction of the B cells was uninfected on day one or two p.i., which is an uncertainty and needs to be addressed. The use of samples comprising the transcriptomes of the non-infected primary B cells cultivated for up to two days should definitely increase the specificity of my results and might provide valuable new information about the gene regulation of these prone to die cells.

#### **4.5.3. When and where do histone modifications occur on viral DNA during the early phase of infection?**

The kinetics of nucleosome acquisition and assembly on EBV genome has to be confirmed and refined with more biological replicates using my MNase-seq approach. Additional time points within the first 24 hours of infection at 6 and 12 hours p.i. will allow to define the exact start when EBV DNA becomes chromatinized. To exclude a possible sequence-specific MNase bias, an additional control with naked DNA digested with MNase should be included in the next set of experiments. The role of histone variants has to be studied in more detail to confirm or refute published data and the role of BNRF1 viral protein in the pre-latent phase (Tsai et al., 2014). The reported deposition of H3.1 and H3.3 histones as described for HSV-1 genomic DNA, needs to be revisited to understand the modulation and function of early EBV chromatin (Albright and Kalejta, 2016). Another interesting aspect in my infection model is how and when viral lytic genes acquire their repressed state, which is a prerequisite to establish the latent phase and to persist as a life-long, asymptomatic infection. Close to nothing is known about this process, which likely results from a dynamic interplay of the cell-mediated deposition of repressive chromatin marks (histone H3K9 and H3K27 methylation) and the virally implemented modulation of histone modifications and nucleosome remodelling.

## 5. Material

### 5.1. Oligonucleotides

**NEBNext Multiplex Oligos kit for Illumina (primers set 2)** used for library preparation containing:

- NEBNext USER enzyme and adaptor for Illumina

5' -/Phos/GATCGGAAGAGCACACGTCTGAACTCCAGTC /ideoxyU/ ACACTCTTTCCCTACACGACGCTCTTCCGATCT - 3'

- Universal PCR Primer

5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'

- 12 oligos with different barcodes

5' - CAAGCAGAAGACGGCATACGAGATTG /barcode/ GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC-s-T - 3'

In this PhD work the following oligos were used: 13, 14, 15, 16, 18, 19, 20, 22, 23, 25.

### 5.2. Plasmids

The two plasmids, p509 and p2670, were used to induce EBV's lytic cycle in HEK293 2089 cells. The p509 plasmid expresses the *BZLF1* gene under the control of the CMV promoter (Hammerschmidt and Sugden, 1988), which is a trigger for the lytic phase reactivation. The plasmid p2670 contains the *BALF4* gene under the CMV promoter to improve the virus progeny production and to increase the virus titer.

### 5.3. Virus supernatants

Table 2. Virus supernatants used to infect B cells during this work.

EBV SUPERNATANT	DESCRIPTION
<b>2089</b>	Wild type maxi EBV: F plasmid (pMB0131) with eGFP and hygromycin resistance gene in B95-8 DNA. Virus supernatant was collected from the stably transfected HEK293 cells.
<b>B95-8</b>	Wild type EBV derived from the marmoset B-lymphoblastoid cell line (B95-8) (Miller et al., 1972)



## 5.4. Antibodies

Table 3. The antibodies from different suppliers used in this work are listed in the table below.

TARGET	APPLICATION	SPECIES	SUPPLIER	DILUTION/AMOUNT USED
<b>ANTI-MOUSE IgG, HRP</b>	WB	goat	CellSignaling (#7076S)	1:10,000
<b>ANTI-RABBIT IgG, HRP</b>	WB	rabbit	CellSignaling (#7074S)	1:10,000
<b>CD38</b>	FACS sorting	mouse	eBioscience (#25-0389-42)	70 µl 1:10 dilution/1x10 <sup>8</sup> cells
<b>DAXX</b>	WB	rabbit	SantaCruz (#sc-7152)	1:1,000
<b>GAPDH</b>	WB	mouse	GeneTex (#GT239)	1:10,000
<b>HIRA</b>	WB	rabbit	Abcam (#ab20655)	1:2,000
<b>HISTONE H2A.Z</b>	WB	rabbit	Abcam (#ab4174)	1:1,000
<b>HISTONE H3.3</b>	WB	rabbit	Abcam (#ab62642)	1:1,000
<b>IgD</b>	FACS sorting	mouse	BD Pharmingen (#555778)	70 µl/1x10 <sup>8</sup> cells
<b>IRDYE ANTI-MOUSE 680RD</b>	WB	goat	Licor (#925-68070)	1:15,000
<b>IRDYE ANTI-RABBIT 800CW</b>	WB	donkey	Licor (#925-32210)	1:15,000
<b>NAP-1</b>	WB	rabbit	Abcam (#ab79490)	1:1,000
<b>PCNA</b>	WB	mouse	SantaCruz (#sc-56)	1:1,000
<b>β-ACTIN</b>	WB	mouse	Sigma (#A5316)	1:5,000

## 5.5. Eukaryotic cell lines

Table 4. All stable cell lines used during this PhD work stem from the collection of the Research Unit Vectors group or were derived from adenoid tissue as described below.

NAME	DESCRIPTION
<b>B95-8</b>	Marmoset lymphoblastoid cell line immortalised by EBV infection (Miller et al., 1972)
<b>HEK293 2089</b>	Human embryonic kidney cell line (HEK293) stably transfected with the maxi-EBV plasmid (p2089) (Delecluse et al., 1998)
<b>LCL</b>	Lymphoblastoid cell line derived from B cell transformed with EBV
<b>PRIMARY B CELLS</b>	B cells purified from human adenoid tissue and used for infection with EBV, different anonymous donors
<b>RAJI</b>	Human Burkitt-lymphoma cell line, EBV-positive
<b>ELIJAH</b>	EBV- negative Burkitt lymphoma cell line

## 5.6. Cell culture media and additives

Table 5. Media and media supplementation used for cultivating eukaryotic cell lines.

<b>MEDIUM</b>	<b>APPLICATION</b>	<b>SUPPLIER</b>
<b>ALPHA-THIOGLYCEROLE</b>	Antioxidant reagent as an additive to the media	Sigma-Aldrich, Munich, Germany
<b>FETAL BOVINE SERUM (FBS)</b>	Nutritive additive to the cell media (8%)	BioSell, Feuch/Nürnberg, Germany
<b>HYGROMYCIN</b>	Antibiotics used for cell culture selection (for HEK293 2089 cell line)	AppliChem, Darmstadt, Germany
<b>PENICILLIN/STREPTOMYCIN</b>	Antibiotics used as an additive to the media	Life technologies Corporation (Gibco), USA
<b>RPMI-1640</b>	Cell culture medium used for cultivation of all cell lines used in this PhD work	Life technologies Corporation (Gibco), USA
<b>SODIUM PYRUVATE</b>	Antioxidant reagent as an additive to the media	Life technologies Corporation (Gibco), USA
<b>SODIUM SELENITE</b>	Antioxidant reagent as an additive to the media	Sigma-Aldrich, Munich, Germany
<b>TRYPsin-EDTA</b>	0.05 % solution used for cell dissociation from the cell culture dish	Life technologies Corporation (Gibco), USA

## 5.7. Enzymes and other chemicals

Table 6. The list of enzymes and other chemicals used in this PhD thesis.

<b>ENZYME/CHEMICAL</b>	<b>SUPPLIER</b>
<b>2-NBDG</b>	Thermo Scientific, USA
<b>AGAROSE</b>	Invitrogen, Karlsruhe, Germany
<b>AMERSHAM ECL WESTERN BLOTTING DETECTION REAGENT</b>	GE Healthcare, USA
<b>AMMONIUM PERSULFATE (APS)</b>	SERVA, Heidelberg, Germany
<b>BETA-MERCAPTOETHANOL</b>	Sigma-Aldrich, Munich, Germany
<b>BOVINE SERUM ALBUMIN (BSA)</b>	Sigma-Aldrich, Munich, Germany
<b>BRADFORD SOLUTION</b>	BioRad, Munich

<b>CIPROBAY 200</b>	Bayer Vital, Leverkusen, Germany
<b>CYCLOSPORIN A</b>	Sigma-Aldrich, Munich, Germany
<b>DNA LADDERS (GENE RULER 1KB AND 100 BP PLUS)</b>	Thermo Scientific, USA
<b>ETHANOL (ABSOLUTE, PURE)</b>	Merck, Darmstadt, Germany
<b>ETHIDIUM BROMIDE</b>	Roth, Karlsruhe, Germany
<b>FICOLL-HYPAQUE</b>	PAN Biotech, Aidenbach, Germany
<b>FORMALDEHYDE</b>	Sigma-Aldrich, Munich, Germany
<b>LIGHTCYCLER 480 SYBR GREEN I MASTER MIX</b>	Roche Diagnostics, Unterhaching, Germany
<b>MICROCOCAL NUCLEASE (MNASE)</b>	Sigma-Aldrich, Munich, Germany
<b>MINI PROTEASE INHIBITIOR COCKTAIL (COMPLETE)</b>	Roche Diagnostics, Unterhaching, Germany
<b>NONFAT DRIED MILK POWDER</b>	AppliChem, Darmstadt, Germany
<b>NP-40 (IGEPAL)</b>	Sigma-Aldrich, Munich, Germany
<b>PAGE RULER PLUS PRESTAINED PROTEIN LADDER</b>	Thermo Scientific, USA
<b>PBS</b>	Biochrome AG, Berlin
<b>POLYETHYLENIMINE MAX (PEI MAX)</b>	Sigma-Aldrich, Munich, Germany
<b>PROTEINASE K</b>	Roth, Karlsruhe, Germany
<b>RNASE A</b>	AppliChem, Darmstadt, Germany
<b>ROTIPHORESE GEL A POLYACRYLAMIDE (30 %)</b>	Roth, Karlsruhe, Germany
<b>SODIUM DODECYL SULFATE (SDS, 10 %)</b>	SERVA, Heidelberg, Germany
<b>SPERMIDINE</b>	Sigma-Aldrich, Munich, Germany
<b>SPERMINE</b>	Sigma-Aldrich, Munich, Germany
<b>TEMED</b>	Roth, Karlsruhe, Germany
<b>TMRE (MITOSTATUS)</b>	BD Pharmingen, USA

<b>TRITON X-100</b>	Sigma, Steinheim, Germany
<b>TRIZOL REAGENT</b>	Thermo Scientific, USA

## 5.8. Buffers

Table 7. The list of buffers and solutions prepared and used in this work.

<b>BUFFER</b>	<b>COMPOSITION</b>	<b>APPLICATION</b>
<b>BLOCKING BUFFER</b>	5 % nonfat dried milk PBS-T	Western Blot
<b>BLOTTING BUFFER (1 X)</b>	25 mM Tris, 190 mM glycine, 20 % methanol in H <sub>2</sub> O	Western Blot
<b>DIGESTION BUFFER</b>	10 mM Tris pH 7.5, 15 mM NaCl, 60 mM KCl, 0.5 % NP-40, 1 mM CaCl <sub>2</sub> in ddH <sub>2</sub> O add fresh: 0.15 mM Spermine, 0.5 mM Spermidine and 1 x proteinase inhibitor cocktail (Roche)	MNase digestion
<b>FACS STAINING BUFFER</b>	0.5 % BSA, 2mM EDTA in PBS	FACS based applications, staining
<b>FIXATION BUFFER</b>	1 % Formaldehyde, 10 % FBS in PBS	MNase experiment
<b>LAEMMLI BUFFER</b>	2 % SDS, 100 mM DTT, 10 mM EDTA, 10 % glycerol, 60 mM Tris/HCl (pH 6.8), 0.01 % bromophenol blue, 1 mM PMSF in H <sub>2</sub> O	Western Blot
<b>LOADING BUFFER</b>	10 mM Tris-HCl (pH 7.5), 60 % glycerol, 60 mM EDTA, Orange G	Agarose gel electrophoresis
<b>NP-40 LYSIS BUFFER</b>	10 mM Tris pH 7.5, 10 mM NaCl, 3 mM MgCl <sub>2</sub> , 0.5 % NP-40 in ddH <sub>2</sub> O add fresh: 0.15 mM Spermine, 0.5 mM Spermidine and 1 x proteinase inhibitor cocktail (Roche)	MNase digestion
<b>PBS-T</b>	0.1 % Tween-20 in PBS	Western Blot
<b>RIPA BUFFER</b>	50 mM Tris-HCl pH8.0, 150 mM NaCl, 0.1 % SDS, 0.5 % sodium deoxycholate, 1 % NP-40, supplemented with 1x "Complete" proteinase inhibitor cocktail (Roche)	Western Blot
<b>RUNNING BUFFER (10 X)</b>	25 mM Tris, 190 mM glycine, 0.1 % SDS in H <sub>2</sub> O	Western Blot
<b>SEPARATION GEL (10 % OR 12 %)</b>	Polyacrylamide (30 %) 3.3 ml/4.0 ml, 1.66 ml 2M Tris-HCl pH 8.9, 66.7 µl 0.5 M EDTA, 100 µl 10 % SDS, 70 µl 10 % APS, 10 µl TEMED, in 10 ml ddH <sub>2</sub> O	Western Blot
<b>STACKING GEL (5 %)</b>	Polyacrylamide (30 %) 1.34 ml, 1.5 ml 2M Tris-HCl pH 6.6, 80 µl 10 % SDS, 70 µl 10 % APS, 10 µl	Western Blot

	TEMED, in 10 ml ddH <sub>2</sub> O	
<b>STOP BUFFER</b>	50 % (v/v) 0.5 M EDTA, 50 % (v/v) 10 % SDS	MNase digestion
<b>TAE</b>	40 mM Tris-acetate, 1 mM EDTA in H <sub>2</sub> O	Agarose gel electrophoresis
<b>TBE</b>	89 mM Tris, 89 mM boric acid, 2 mM EDTA, in H <sub>2</sub> O	Agarose gel electrophoresis
<b>WASH BUFFER</b>	10 mM Tris pH 7.5, 15 mM NaCl, 60 mM KCl, 0.5% NP-40, in ddH <sub>2</sub> O add fresh: 0.15 mM Spermine, 0.5 mM Spermidine and 1 x proteinase inhibitor cocktail (Roche)	MNase digestion

### 5.9. Commercial kits

Table 8. Commercial kits used in this PhD work.

<b>KIT</b>	<b>SUPPLIER</b>
<b>AGENCOURT AMPURE XP, PCR CLEAN-UP</b>	Beckman Coulter, Munich, Germany
<b>AMERSHAM ECL SELECT WESTERN BLOTTING DETECTION REAGENT</b>	GE Healthcare, USA
<b>B CELL ISOLATION KIT II, HUMAN</b>	Miltenyi Biotec GmbH, Bergisch Gladbach, Germany
<b>ERCC RNA SPIKE-IN MIX</b>	Thermo Fischer Scientific, USA
<b>NEBNEXT MULTIPLEX OLIGOS FOR ILLUMINA (INDEX PRIMERS SET 2)</b>	New England Biolabs, USA
<b>NEBNEXT ULTRA II LIBRARY PREP KIT FOR ILLUMINA</b>	New England Biolabs, USA
<b>NUCLEOSPIN EXTRACT II KIT PCR/GEL PURIFICATION KIT</b>	Machery-Nagel, Düren, Germany
<b>QIAMP DNA MINI KIT</b>	Qiagen, Hilden
<b>QUBIT dsDNA HS ASSAY KIT</b>	Thermo Fischer Scientific, USA
<b>RIBOZERO rRNA REMOVAL KIT</b>	Illumina, USA
<b>RNAEASY MINI KIT</b>	Qiagen, Hilden
<b>TRUSEQ DUAL INDEX SEQUENCING PRIMER BOX, PAIR-END</b>	Illumina, USA

## 5.10. Software and packages

Table 9. Software and packages used for data analysis.

<b>SOFTWARE</b>	<b>APPLICATION</b>	<b>SOURCE</b>
<b>ANACONDA COMMAND LINE CLINET 1.6.6</b>	Open source of data science packages	Anaconda Inc., USA <a href="https://anaconda.org/">https://anaconda.org/</a>
<b>BD FACSDIVA SOFTWARE</b>	FACS analysis, FACS sorting	BD Biosciences, USA
<b>BEDTOOLS 2.25.0</b>	NGS data processing	(Quinlan and Hall, 2010), <a href="http://bedtools.readthedocs.io/en/latest/">http://bedtools.readthedocs.io/en/latest/</a>
<b>BOWTIE 1.1.0</b>	NGS data processing	(Langmead et al., 2009), <a href="http://bowtie-bio.sourceforge.net/index.shtml">http://bowtie-bio.sourceforge.net/index.shtml</a>
<b>BOWTIE 2.3.3</b>	NGS data processing	(Langmead and Salzberg, 2012) <a href="http://bowtie-bio.sourceforge.net/bowtie2/index.shtml">http://bowtie-bio.sourceforge.net/bowtie2/index.shtml</a>
<b>DANPOS 2.2.2</b>	A tool for Dynamic Analysis of Nucleosome and Protein Occupancy by Sequencing, Python package	(Chen et al., 2013) <a href="https://sites.google.com/site/danposdoc/news">https://sites.google.com/site/danposdoc/news</a>
<b>DESEQ2</b>	NGS data processing and normalization	(Love et al., 2014), <a href="https://bioconductor.org/packages/release/bioc/html/DESeq2.html">https://bioconductor.org/packages/release/bioc/html/DESeq2.html</a>
<b>dynamicTreeCut 1.6.3</b>	R package for clusters visualization	(Langfelder et al., 2008) <a href="https://cran.r-project.org/web/packages/dynamicTreeCut/index.html">https://cran.r-project.org/web/packages/dynamicTreeCut/index.html</a>
<b>FASTCLUSTER 1.1.24</b>	R package for clustering	Daniel Müllner, <a href="https://cran.r-project.org/web/packages/fastcluster/index">https://cran.r-project.org/web/packages/fastcluster/index</a>
<b>FLOWJO 10.4</b>	FACS data analysis	FlowJo, USA
<b>GALAXY 15.10</b>	NGS data processing	(Afgan et al., 2016), <a href="https://galaxyproject.org/">https://galaxyproject.org/</a>
<b>GGPLOT2 2.2.1</b>	R package for graphs construction	Handley Wickham, <a href="http://ggplot2.tidyverse.org/reference/">http://ggplot2.tidyverse.org/reference/</a>
<b>HOMEBREW</b>	Installation helper	Max Howell, <a href="https://brew.sh/">https://brew.sh/</a>
<b>HTSEQ (V 0.6.0)</b>	Python program to count mapped sequencing reads	(Anders et al., 2015) <a href="https://htseq.readthedocs.io/en/release_0.9.1/">https://htseq.readthedocs.io/en/release_0.9.1/</a>

<b>ILLUSTRATOR CS5 PHOTOSHOP CS5</b>	Graphics design	Adobe Systems Inc., USA
<b>INTEGRATIVE GENOMICS VIEWER (IGV) 2.5.9</b>	Visualization of NGS data results	(Thorvaldsdottir et al., 2013; Robinson et al., 2011) <a href="http://software.broadinstitute.org/software/igv/">http://software.broadinstitute.org/software/igv/</a>
<b>MACVECTOR 15.5</b>	Nucleic acid visualization, cloning, primer design	MacVector, USA
<b>MENDELEY DESKTOP 1.17.12</b>	Literature citation	Mendeley Inc., USA, <a href="https://www.mendeley.com/newsfeed/">https://www.mendeley.com/newsfeed/</a>
<b>MICROSOFT OFFICE 2017</b>	Text, tables and presentation processing	Microsoft, USA
<b>PRISM 7</b>	Basic statistics, graphs processing	Graphpad Software, USA
<b>PYTHON 2.7.14</b>	Programing language	Python, <a href="https://www.python.org/">https://www.python.org/</a>
<b>R VERSION 3.3.3</b>	Programming language for statistical computing and graphics	The R Foundation, Austria <a href="https://www.r-project.org/">https://www.r-project.org/</a>
<b>RSTUDIO VERSION 1.1.383</b>	Integrated Development Environment for R	RStudio Inc., USA, <a href="https://www.rstudio.com/">https://www.rstudio.com/</a>
<b>SAMTOOLS 0.1.7</b>	NGS data processing	(Li et al., 2009) <a href="http://samtools.sourceforge.net/">http://samtools.sourceforge.net/</a>
<b>STAR ALIGNER (V 2.4.2A)</b>	Sequence reads alignment	(Dobin et al., 2013) <a href="https://github.com/alexdobin/STAR/">https://github.com/alexdobin/STAR/</a>
<b>UPSETR</b>	R package	(Conway et al., 2017; Lex et al., 2014) <a href="https://cran.rproject.org/web/packages/UpSetR/README.html">https://cran.rproject.org/web/packages/UpSetR/README.html</a>
<b>VENNDIAGRAM 1.6.18</b>	R package	Hanbo Chen, <a href="https://cran.r-project.org/web/packages/VennDiagram/index.html">https://cran.r-project.org/web/packages/VennDiagram/index.html</a>

### 5.11. Devices and consumables

Table 10. Special devices used in this work

DEVICE	SUPPLIER
<b>BIORUPTOR STANDARD SONICATOR</b>	Diagenode, Belgium
<b>CP1000 DEVELOPER MACHINE</b>	AGFA HealthCare NV, Belgium HealthCare NV, Belgium
<b>FACS CANTO II, FACS FORTESSA, FACS ARIA IIIU</b>	BD Bioscience, USA
<b>X-RAY FILMS</b>	AGFA HealthCare NV, Belgium
<b>ILLUMINA HISEQ1500, ILLUMINA HISEQ4000</b>	Illumina, USA
<b>LI-COR ODYSSEY 9120 INFRARED IMAGING SYSTEM</b>	LI-COR, USA
<b>MACS LD COLUMNS AND SEPARATOR</b>	Miltenyi Biotec GmbH, Bergisch Gladbach, Germany
<b>MINI PROTEAN II TETRA GEL ELECTROPHORESIS UNIT</b>	Bio-Rad, Munich, Germany
<b>NITROCELLULOSE, 0.45 <math>\mu</math>M</b>	Bio-Rad, Munich, Germany
<b>QUBIT FLUOROMETRIC QUANTITATION</b>	Thermo Fischer Scientific, USA
<b>SHIMADZU MICROCHIP ELECTROPHORESIS SYSTEM FOR DNA/RNA ANALYSIS (MCE-202 MULTINA)</b>	Shimadzu Group, Japan



## 6. Methods

### 6.1. Eukaryotic cell culture methods

#### 6.1.1. Cell culture conditions

Eukaryotic cell manipulations were performed in class II biosafety laminar flow hoods. Cells were cultured in incubators at 37 °C with 95 % humidity and 5 % CO<sub>2</sub>. Centrifugation of the cells was performed at 300 g for 10 min. Cells were cultivated in RPMI-1640 medium supplemented with 8 % FBS, 100 µg/ml streptomycin, 100 U/ml penicillin, 1 mM sodium pyruvate, 100 nM sodium selenite, and 0.43 % α-thioglycerols. Cells were counted using a haemocytometer (Neubauer cell chamber) according to the manufacturer's instructions.

The suspension cells were cultivated and kept at an average density of 5x10<sup>5</sup> cells/ml. HEK293 2089 cells were cultured using 100 µg/ml of hygromycin selection on the 13 cm dishes until the cells reached around 80 % confluency. In the next step, the cells were dissociated from the dish using a Trypsin-EDTA 0.05 % solution and the detached cells were resuspended in fresh medium. A certain fraction of the cells was transferred to a new dish.

#### 6.1.2. Freezing and thawing of eukaryotic cells

Thawing of eukaryotic cells was performed by warming a frozen vial to 37 °C and immediately transferring it to 50 ml of pre-warmed medium. Cells were pelleted, resuspended in 5 ml fresh medium, and cultured overnight in the incubator. The next day the cells were transferred to fresh cell culture medium containing appropriate antibiotic for selection.

To freeze eukaryotic cells, cold medium containing 20 % FBS and 10 % DMSO was prepared. Around 5x10<sup>6</sup> cells were collected for one freezing vial, the cells were centrifuged and resuspended in 1 ml of ice-cold freezing medium. The cell suspension was transferred to the pre-cooled vial and stored in a freezing-box at -80 °C for approximately one week. Afterwards, the vials were transferred to a liquid nitrogen tank for long term storage.

#### 6.1.3. Extraction of B cells from the adenoid tissue

The adenoid tissues were first cleaned with PBS and then transferred to a sterile petri dish. In the next steps, the adenoids were cut extensively using two sterile scalpels and washed

with 2-3 ml of PBS. The cell suspension was taken up with a Pasteur pipette and filtered through a 100 µm cell strainer into a new falcon tube. This procedure was repeated multiple times until the adenoid was fully disintegrated to isolate as many cells as possible. The 50 ml falcon tube with the collected cell suspension was filled with PBS and centrifuged at 300 g. Afterwards, the cell pellet was resuspended in 30 ml fresh PBS and supplemented with 0.5 ml defibrinated sheep blood to deplete most of the T cells from the preparation by rosetting. The prepared cell suspension was loaded very carefully on top of 15 ml Ficol Hypaque to obtain two clearly separated phases. Next, the samples were centrifuged at 500 g for 30 min with minimum break and acceleration. The cells were carefully collected from the turbid interphase and transferred to a new falcon tube, which was filled with PBS. The cells were washed with PBS three times, using decreasing centrifugation speed (450, 400 and 300 g for 10 min). After the final wash, the cell pellet was resuspended in fresh, pre-warmed medium. The cells were counted and immediately used for further experiment such as cell sorting or MACS depletion.

#### **6.1.4. Collecting virus supernatants from the HEK293 2089 cell line**

One day prior to transfection, HEK293 2089 cells were seeded on a 13 cm dish and cultivated until they reached 70-80 % confluency. The transfection mix containing 6 µg of p509, carrying the *BZLF1* gene, and 6 µg of p2670 with the *BALF4* gene in RPMI medium was mixed with PEI Max at a 1:6 ratio. The mixed solution was incubated for 15 min at room temperature (RT) and added dropwise onto the HEK293 2089 cells. Prior this step, the cell culture medium was replaced to fresh medium without hygromycin. Dishes with medium were carefully mixed and incubated for another three days in the incubator. Afterwards, the supernatants from the transfected plates were collected in falcon tubes and centrifuged for 8 min at 300 g to remove cells from the medium. Then, the supernatant was transferred to a new falcon tube, and centrifuged once more at 1200 g for 8 min. Collected virus supernatants were stored at 4 °C and used to determine the virus concentration with Raji cells (Material 5.5).

#### **6.1.5. Collecting virus supernatant from the B95-8 cell line**

For collecting EBV virus stocks, the B95-8 cells were cultivated at higher density. A fraction of the cells spontaneously reactivates the lytic phase and releases virus to the medium. Medium was collected from the cells, centrifuged two times at 300 and 1200 g for 10 min and stored at 4 °C.

### 6.1.6. Determining the virus titer

The virus titer of wt EBV (2089) was determined by infecting Raji cells with different amounts of the virus stock, and GFP-positive Raji cells were measured after three days.  $1 \times 10^5$  Raji cells were seeded onto a 24-well plate in 100  $\mu$ l medium. Different amounts of virus supernatant were added to the Raji cells and the wells were filled up to a total volume of 2 ml. On the next day, 80 % of medium was removed from the cells and replaced by fresh RPMI. Cells were incubated for another two days and analysed with a FACS Cano for GFP expression. The Green Raji Units (GRU) were calculated by considering the volume of the added supernatant, the percentage of the GFP-positive cells, and the total initial number of Raji cells ( $GRU = \left( \frac{\% \text{ GFP positive cells}}{100 \%} \right) \times \left( \frac{1000 \mu\text{l}}{x \mu\text{l of supernatant}} \right) \times 10^5$ ). GRU per ml indicates the concentration of virus in a given sample. This titer will be used to obtain a certain multiplicity of infection (MOI).

The titers of wt EBV (B95-8) were obtained by a different method performed by Dagmar Pich. Briefly, Elijah cells were infected with different amounts of virus supernatant for 3 hours at 4 °C under rotation and later the cells were stained with a directly coupled antibody directed against the viral envelop protein gp350, which binds to viral particles attached to the cells. The fractions of gp350-positive cells infected with B95-8 virus stock were compared to the cells infected with 2089 virus supernatant with the known titer, which reveal as a calibrated reference.

### 6.1.7. Purification of B cells on MACS columns

All the steps were performed on ice and with pre-cooled solutions. Cells obtained from adenoid tissue were centrifuged (300 g for 10 min) and washed once with ice-cold FACS staining buffer. Samples were resuspended in FACS staining buffer in 400  $\mu$ l/ $10^8$  cells and 100  $\mu$ l of the Biotin-Antibody Cocktail, provided by the B cell isolation kit from Miltenyi, was added. Non-B cells were labelled with a mix of biotinylated antibodies directed against: CD2, CD14, CD16, CD36, CD43, and CD235a surface markers. The cell suspension was mixed and incubated for 10 min on ice. In the next step, 300  $\mu$ l FACS buffer was added to the samples followed by supplementation with 200  $\mu$ l of anti-Biotin MicroBeads. The samples were incubated for another 15 min on ice. Meanwhile, the MACS Separator and MACS LD Columns were prepared according to the manufacturer's protocol. Briefly, columns were placed on the magnetic separator and pre-incubated with 5 ml FACS staining buffer and afterwards they were loaded with the cell-MicroBeads suspensions. The flow through containing unlabelled B cells was

collected and stored on ice. Further, the columns were washed two times with 5 ml of FACS staining buffer and the flow through was collected into the pre-cooled falcon tubes. Optionally, columns can be removed from the magnetic separator and washed with FACS Staining Buffer using a column plunger. This fraction contained the labelled non-B cells. Collected cell fractions were used for staining with an appropriate marker (e.g. CD19 for B cells) to control the purity of the B cell purification. Highly pure B cells were obtained by this depletion strategy were used for my infection experiments.

### **6.1.8. Primary B cells infection with EBV**

Infection of primary B cells was performed using wt EBV (2089 or B95-8) with a multiplicity of infection (MOI) 0.1, which refers to the number of virions that are added per cell. The next day the cells were centrifuged (300g at RT for 10 min) to remove unbound virus and resuspended in the fresh medium.

## **6.2. Flow cytometry measurements**

### **6.2.1. FACS sorting for primary B cells**

Primary, naïve B cells were sorted using a FACS Aria IIIu device according to the instructions of the manufacturer to get pure populations of cells with our staining strategy. Briefly, the cells were washed once and resuspended in FACS staining buffer (chapter 5.8) to a final concentration of  $1 \times 10^8$  cells/ml. The cells were stained using fluorophore-coupled antibodies against CD38 and IgD (chapter 5.8) for 60 min in the dark. The samples were washed once with ice-cold FACS Staining buffer and resuspended in the same buffer followed by filtration through 35  $\mu\text{m}$  cell strainer to obtain single cell suspensions. Sorting was performed with a 70  $\mu\text{m}$  nozzle (for primary B cells) and a velocity of about 8,000 events/second with a sorting mask of “4-way purity”. The gating criteria included: (i) living cells, (ii) single cells, and (iii) CD38-negative and IgD-positive cells. The cells were sorted to obtain 20 to 30 million primary, naïve B cells, which were used for the following experiments. Daily sorting for infected B cells was performed using the 100  $\mu\text{m}$  nozzle and a velocity of about 5,000 events/second with the sorting mask of “4-way purity”.

### 6.2.2. Staining with TMRE and Annexin V

Primary B cells were purified using the MACS B cell isolation kit.  $5 \times 10^5$  uninfected B cells were resuspended in 1 ml of fresh pre-warmed medium and transferred to polypropylene FACS tubes. Prepared cells were directly mixed with four different concentrations of TMRE (0, 20, 50 and 100 nM). The samples were incubated for 30 min at 37°C protected from light. In the next steps, the cells were washed two times with FACS buffer to remove the uncoupled dye. Finally, the cells were resuspended in 500 µl Annexin V staining buffer and incubated with APC-coupled Annexin V for another 10 min in the dark. Co-stainings were measured on a FACS Fortessa machine looking at TMRE incorporated cells in the PE channel and the Annexin V positive population in the APC channel. Measurements were performed daily for the first eight days after wt EBV (B95-8) infection as described above (chapter 6.1.8). The experiment was performed in the same way also with LCLs, which were cultured for two weeks after infection.

### 6.2.3. Glucose analogue uptake analysis using 2-NBDG

The experiments were performed with primary B cells purified with the MACS B cell isolation kit (uninfected cells and cells infected with wt EBV (B95-8) collected on daily basis: day zero to day five and day eight). About  $5 \times 10^5$  B cells were seeded into a 24-well plate and resuspended in 500 µl of glucose-free medium supplemented with 1 % FBS. Three different concentrations (0, 20 and 50 µM) of the glucose analogue 2-NBDG (2-[N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl) amino]-2-deoxy-D-glucose) were added to the prepared samples, which were incubated for one hour. Afterwards, the cells were washed two times with ice-cold PBS to remove the remaining reagent from the medium. The treated samples were resuspended in 300 µl ice-cold PBS and measured directly on the FACS Fortessa machine using the FITC channel.

## 6.3. Nucleic acid methods

General nucleic acid methods such as phenol/chloroform extraction and DNA precipitation with ethanol or DNA electrophoresis were performed using standard laboratory protocols (Sambrook and Russell, 2001).

### 6.3.1. Genomic DNA purification from eukaryotic cells

Genomic DNA was isolated from eukaryotic cells by using the QIAamp DNA Mini Kit (Qiagen) following the manufacturer's instructions.

### **6.3.2. DNA purification from agarose gel**

Purification of DNA from agarose gels was performed using the NucleoSpin Extract II Kit (Machery-Nagel) according to the instructions in the manual.

## **6.4. Time-course RNA-seq experiments**

### **6.4.1. Sample collection**

B cells were isolated from adenoid tissues and sorted to enrich for the naïve B cells population as described above (6.1.3 and 6.2.1). The “Day 0” samples with one million of uninfected naïve B cells were washed two times with PBS and resuspended in 1 ml TRIZOL reagent followed by snap freezing in liquid nitrogen. The cells were stored for further analysis at -80 °C. The remaining sorted cells were bulk infected with wt EBV (2089) using MOI 0.1 (as described in 6.1.7). The cells were kept overnight in the incubator at a final concentration of around  $1 \times 10^6$  cells/ml supplemented with 10 µg/ml Ciprobay and 0.5g/ml Cyclosporine A. On the next day, cells were spun down to remove unbound viral particles and resuspended in fresh medium with Ciprobay and Cyclosporine A. Cells were split to separated bottles and kept for different days after infection. On a daily basis, the cells were collected and resuspended in FACS buffer. Living cells were sorted according to the established gating strategy to obtain one million cells. Samples were washed two times with PBS, resuspended in TRIZOL reagent and snap frozen in liquid nitrogen. Frozen cells were further used for RNA extraction.

### **6.4.2. RNA isolation, ERCC spike-in addition**

Total RNA extractions from sorted and frozen samples (day zero to five, day eight, and day 14) were performed by the company Vertis Biotechnologie AG (Lise-Meintner-Straße 30, 85354 Freising). Before isolation of total RNA, all samples were supplemented with the specified amounts of ERCC spike-ins according to the provided manual. Briefly, after the addition of ERCC spike-in controls, the samples were thawed and mixed by vortexing. 200 µl chloroform was added to the homogenate, mixed and centrifuged with maximal speed for 20 min at 4 °C. The aqueous phase was carefully transferred to a new sterile RNase-free tube and combined with an equal volume of 100 % RNA-free ethanol. Samples were then loaded onto the RNeasy columns and total RNA extraction was performed using the RNeasy Kit from Qiagen including DNase treatment. Total RNA quality and concentrations were measured by capillary electrophoresis using the Shimadzu MultiNA Electrophoresis System for DNA/RNA analysis.

### 6.4.3. Library preparation and sequencing

Preparation of sequencing libraries was performed by the company Vertis Biotechnologie AG (Lise-Meintner-Straße 30, 85354 Freising). Extracted total RNA samples were split in two to prepare two types of libraries: i) samples from three donors were used for polyA enriched libraries, which were primed with an oligo(dT) primer ii) with RNA from two of three donors rRNA depleted, random hexamer enriched libraries were established.

- (i) For the preparation of polyA enriched libraries, indicated amounts of RNA were used (30 ng from Day0, 50 ng from Day1, 100 ng from Day2 and 200ng from samples collected at Day3-5, 8 and 14). The first-strand cDNA was synthesized using oligo(dT) primer. PolyA enriched cDNA samples were fragmented and further used for the ligation of Illumina TruSeq sequencing adapters to the 5' and 3' ends of the cDNA fragments. Barcoded cDNA samples were finally amplified by PCR using a proof-reading enzyme with numbers of cycles ranging from 15-18. The primers used for PCR amplification were designed for TruSeq Dual-Index sequencing according to the instructions by Illumina. The following adapter sequences flank the DNA inserts:

TruSeq\_Sense\_primer i5 Index:

5'- AATGATACGGCGACCACCGAGATCTACAC-NNNNNNNN-ACACTCTTCCCTACACGACGCTCTTCCGATCT-3'

TruSeq\_Antisense\_primer i7 Index:

5'-CAAGCAGAAGACGGCATACGAGAT-NNNNNNNN-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT-3'

- (ii) The remaining RNAs samples were used for preparing the rRNA depleted, random hexamer enriched libraries. rRNAs were depleted using the Ribo-Zero rRNA Removal Kit GOLD from Illumina according to the manufacturer's instructions. The recovery rates after rRNA depletion were checked on a capillary electrophoresis system and calculated to be on average 2 % of the initial rRNA. From these rRNA depleted samples, the first strand cDNA synthesis was performed using a N6 randomized primer. Afterwards the samples were fragmented and TruSeq adapters were ligated to the 5' and 3' ends of cDNA fragments. Finally, PCR amplification using a proof reading enzyme was performed with 15-18 cycles. Primer sets for TruSeq Dual-Index sequencing were used for amplification similar to the ones used for polyA enriched libraries.

All amplified cDNA samples were later size fractionated to obtain a range of 250-600 bp fragments using a clean-up system from Agencourt AMPure XP kit. The aliquots of size selected

cDNAs were measured by capillary electrophoresis to confirm the correct size and to determine the concentration of each sample. The prepared libraries were sequenced (paired-end, 100 bases) on two flow cells of the Illumina HiSeq4000 platform (Institute of Human Genetics, Helmholtz Zentrum München).

#### 6.4.4. Bioinformatic analysis

Mapping and counting of the sequence files were performed with the help of Thomas Schwarzmayr (Institute of Human Genetics, Helmholtz Zentrum München).

##### Mapping of the sequencing reads

Sequenced libraries on average yielded about 15.7 Gb of sequences per sample. STAR aligner (v 2.4.2a) with modified parameter settings (`--twopassMode=Basic`) was used for split-read alignment against the human genome assembly Hg19 (GRCh37) combined with the human herpes virus v4 wild type genome sequence (AJ507799 from GenBank) and the ERCC spike-in transcript sequences (Thermo Fischer Scientific). The UCSC knownGene annotation was used together with the human herpes virus v4 annotation (manually created by Dr. Takanobu Tagawa) to provide the known exon boundaries to STAR.

```
/usr/local/STAR \--runThreadN4\--  
genomeDir/genomes/hg19_ebv7/chromosome/STARindex/ \--  
readFilesIn<FASTQFILES_R1><FASTQFILES_R2>\-  
outFileNamePrefix<OUTPUTFOLDER_PREFIX>\--readFilesCommand zcat \--outSAMtype BAM  
Unsorted \--chimSegmentMin 20 \--twopassMode Basic \--limitOutSJcollapsed  
10000000 \--limitIObufferSize 500000000 \--outSAMattrRGline READGROUPS4BAMFILE \-  
-outSAMattributes All
```

##### Gene counting

In order to quantify the number of reads mapped to the annotated genes the HTseq-count (v0.6.0) program was used with the default settings.

```
/usr/local/bin/htseq-count \-f bam \-t exon \-i gene_id \-m intersection-nonempty  
\-r name \--stranded=no <PATHTOBAMFILE> \ucsc_ebv_spikeins.gtf \>  
OUTPUTFILE.htseqcount
```

Before testing for the expression changes across the eight time points, gene expression values were normalized. Normalization and clustering analysis were performed with the help and expertise of Dr. Antonio Scialdone (Institute of Epigenetics and Stem Cells, Institute of



Computational Biology, Helmholtz Zentrum München). The R script for this analysis with around 800 lines of code can be found on the CD under “RNA\_seq/RNaseq\_clusters\_script.R”.

### **Quality control and data normalization**

As a first assessment of sample quality, the total number of mapped reads, the numbers of detected genes, and the ratio between the reads mapped to viral genes and those mapped to endogenous genes were computed. The data were normalized for sequencing depth using size factors (Anders and Huber, 2010), which will normalize the data according to the amount of RNA in each sample, which correlates with cell size. After normalization, principal component analysis (PCA) was performed on the log-transformed expression matrix.

### **Differentially expressed genes**

The DESeq2 R package was used to identify differentially expressed genes between pairs of samples from different days post infection. Before running DESeq2, low expressed genes were removed by excluding all of those that had an average expression of lower than 50 normalized counts. Next, all genes that were statistically significant at a false discovery rate of lower than 0.1 and had an estimated fold-change greater than 2 or lower than 0.5 were selected.

### **Gene clustering**

Identification and classification of genes with a dynamic expression pattern during the time-course of infection were calculated using a clustering method.

In the first step, differentially expressed genes between all possible combination of pairs of days were identified by DESeq2. From this analysis, all genes (N=8997) that were significantly different in at least one of the comparisons were selected for further analytical steps.

To classify these 8997 genes into clusters, a distance matrix between genes was identified as:  $\sqrt{(1 - \rho)/2}$ , where  $\rho$  is the Spearman's correlation coefficient between pairs of genes across all samples (van Dongen and Enright, 2012). Hierarchical clustering was performed on this distance matrix (“hclust” function in R, with an average method), followed by the dynamic hybrid cut algorithm (dynamicTreeCut package, with minimum cluster size of 500 and default value of “deepSplit” parameter). The analysis resulted in seven clusters with specific expression patterns of cellular genes according to the time-course of EBV infection.

### **GO analysis**

The list of genes identified in each of the seven gene clusters were used to look for the enrichment of specific Gene Ontology (GO) terms using the GOrilla online tool (Eden et al., 2009). Additionally, a list of defined background genes was used to increase the specificity of the enrichment analysis. The background genes were defined as union of differentially expressed (DE) genes, which were defined in all pairwise analyses by DESeq2. For the search of specific GO terms the p-value threshold was set to  $10^{-7}$ . The defined biological process terms were used for further analysis with the REViGO online tool (Supek et al., 2011). The program accepts long lists of Gene Ontology terms and summarizes them by removing redundant GO terms. I used settings, which defined medium similarity, database used for Homo sapiens and SimRel as a default setting for calculating the semantic similarity of the GO terms. I used R scripts created by the REViGO tool to create tree plots and scatter plots from my enriched GO terms.

### **Venn diagram of two types of libraries**

The Venn diagrams of the sets of differentially expressed (DE) genes from polyA enriched library and rRNA depleted library were generated using R, the script can be found on the CD under “RNA\_seq/VennDiagramm\_script.R”.

### **Cell cycle phases prediction from RNA-seq data**

The analysis was performed by Ron Fechtner, a bachelor student from Dr. Antonio Scialdone using a published algorithm generated originally for single cell sequencing data.

To predict the cell cycle phase of each RNA-seq sample, the Pairs algorithm was used as described in (Scialdone et al., 2015) and optimized for the data of bulk human cells data. In brief, the method identifies pairs of genes whose relative expression change between the phases of the cell cycle. The proportion of marker pairs for a given cell cycle phase within the data was used to assign scores for each phase, which provided a prediction on the most likely phase of the sample. The three phase scores (G1, S and G2M cell cycle phase) in the range of 0 to 1 describe how likely a certain proportion can be expected by chance. By shuffling gene counts the actual proportion was compared to the distribution of random proportions. Samples with G1 or G2M scores above 0.5 were assigned to the greater of both. If both were below 0.5, S phase was assigned. Originally, the algorithm was designed and used on single-cell RNA-seq data generated

from human embryonic stem cells (Leng et al., 2015), genes associated with the cell cycle GO term (GO:0007049) were considered for this analysis.

The pairs algorithm can be found and installed from the official Python Package Index (PyPI) using pip installer: `pip install pypairs`

### **UpSet analysis of the DE pairwise comparisons**

The plots were generated using R with the data from sets of differentially expressed genes from pairwise comparisons from DESeq2. The adequate R script can be found on the CD under “RNA\_se/UpSet\_script.R”.

## **6.5. MNase-seq experiment**

### **6.5.1. Sample preparation, MNase digestion**

Uninfected and EBV-infected B cells collected at certain days p.i. (day one to five and day eight) were briefly fixed in 1 % fixation mix and rotated for 5 minutes at RT followed by centrifugation and washing with FACS staining buffer. Prepared and fixed cells were sorted to obtain  $1 \times 10^6$  cells within the living cell gate at the day of harvest. In the next step, cells were washed with cold PBS. The fixed cell pellet was snap frozen and stored at  $-80$  °C. After collecting all samples, the following steps were performed. Cell pellets were thawed and lysed for 5 min on ice using 600  $\mu$ l NP-40 lysis buffer. The samples were centrifuged for 7 min at 300 g and 4 °C. The pellets were washed with 500  $\mu$ l MNase wash buffer, spun down, and finally resuspended in 650  $\mu$ l MNase digestion buffer. 50  $\mu$ l of each sample were taken as an undigested control and 1.5 U of MNase enzyme was added to the remaining sample and incubated at 25 °C for certain time points. From this bulk reaction, 200  $\mu$ l were collected at 10 min, 12 min, and 15 min directly into prepared tubes with 25  $\mu$ l of pre-warmed Stop buffer and incubated for another 5 min to quench the MNase digestion reaction. After collecting all digested and undigested samples, 80  $\mu$ g Proteinase K was added to each sample and incubated over night at 65 °C. The next day, 40  $\mu$ g of RNase A was added to the samples and incubated for 30 min at 37 °C. Afterwards, DNA was purified with phenol/chloroform, precipitated with ethanol and sodium acetate and resuspended in TE. The samples were loaded quantitatively on a 2 % TAE agarose gel to purify the mononucleosomal DNA. Fragments with the length of 150 bp were excised from the gels, purified using the NucleoSpin Extract II Kit from Machery-Nagel, and resuspended in 20  $\mu$ l

elution buffer.

For the preparation of the sonicated samples, one million of well-established LCLs from three different donors (LCL32, LCL55, and LCL110) were used to extract total DNA using the QIAamp DNA Mini Kit. To ensure removal of the proteins, phenol/chloroform extraction was employed on the samples two times followed by ethanol precipitation, purification, and resuspension in TE. Two  $\mu\text{g}$  of the prepared DNA were sheared for either 30, 120, and 240 min with a Bioruptor device (Diagenode) to obtain DNA fragment in similar size to MNase digested DNA. 150 bp fragments after 240 min sonication were excised from the gel and purified with the same kit as MNase digested samples. DNA purified from the gels was measured to determine the concentrations using the Qubit System with HS DNA kit and stored at  $-20\text{ }^{\circ}\text{C}$  until the library preparation step was performed.

### **6.5.2. Library preparation and sequencing**

I prepared the sequencing libraries in the lab of Dr. Helmut Blum in the Gene Center (Feodor-Lynen-Straße 25, 81375 Munich) using the NEBNext Ultra II DNA Library Prep Kit and adapters from the NEBNext Multiplex Oligos II kit for Illumina. 50 ng of MNase digested or sonicated DNA samples were filled up to 50  $\mu\text{l}$  with 1 x TE (pH 8.0) and used for libraries preparations. All steps were performed according to the manufacturer's protocol (version 1.1, 11/15). At step 1.2 of protocol the NEBNext Adaptor for Illumina was diluted 1:10 according to the suggestion from the manual. AMPure XP Beads were used for size-selection clean-up of adaptor ligated DNA. Samples were amplified using a universal primer and certain index primers (including different barcodes for each sample) with 11 to 12 cycles. After the PCR reactions, the samples were cleaned up using AMPure XP Beads and resuspended in 0.1 x TE. The concentrations of the prepared libraries were checked on the Qubit System with the HS DNA Kit and frozen at  $-20\text{ }^{\circ}\text{C}$ . Before loading the samples onto the flow cell, they were checked by Dr. Stefan Krebs from the Gene Centre for their quality on the capillary electrophoresis system, BioAnalyzer 2100 from Agilent. Sequencing was performed with 50 bases (paired-end) on an Illumina HiSeq 1500 platform in the Dr. Blum's lab in the Gene Center.

### **6.5.3. Bioinformatic analysis**

Mapping, normalization and further bioinformatic analyses were performed with the help and advice from Dr. Alexander Buschle. The data from the Illumina sequencer were obtained from the Galaxy server of Dr. Blum's laboratory. The processing of the raw sequencing data was

performed on this server. Initially, the data were split according to their unique indices (demultiplexing) used during the libraries preparation with the Illumina Demultiplex tool. Prior to further downstream analyses, the quality of the sequencing reads was checked using the FastQC tool. Since all of the samples passed the quality control test, they could be used for further bioinformatic steps. The commands described in the methods are examples, which were adapted to the certain input file.

### Mapping and counting of sequencing reads

Demultiplexed FASTQ data were downloaded from the Galaxy server. Mapping of the sequencing reads to the human Hg19 (GRCh37) and 2089 EBV (AGV database) genomes was performed with the help of the command line program Bowtie2.

```
bowtie_index="-x path/Hg19_EBV2089.ebwt"
bowtie_opts="-p 2 -X 250 --no-discordant --no-mixed --no-unal"
bowtie2 $bowtie_opts -x $ bowtie_index -1 Input1.fastq -2 Input2.fastq >
Output.sam
```

In the next step, the total numbers of reads were counted in each individual SAM file (day p.i.).

```
for rows in `ls *.sam`; do
wc -l $rows >> 0_sam_rows.txt;
done
```

Data were normalized according to the libraries sizes, which was done by reducing all SAM files to 53960846 reads, which corresponds to reads of the smallest library. Since SAM files are ordered randomly, no bias was introduced by this step.

```
for rowstocut in `ls *.sam`;do
sed -n 1,53960846p $rowstocut > 1_Sams_Cut/cutTo_53960846_${rowstocut};
done
```

Columns 3 and 9 of the normalized data, which represent chromosome name and insert size, respectively, were extracted from the file. For further analyses reads with an insert length greater than 147- 30 bp, or smaller than 147+30 bp were selected and used for further analyses.

```
for run in `ls *.sam`; do
placeholder=`basename $run .sam`
cut -f3,9 $run > /col_3_9_${placeholder}.txt
awk '( ($9 >= 117) && ( $9 <= 177 ) )' $run > plus_${run}
awk '( ($9 <= -117) && ( $9 >= -177 ) )' $run > minus_${run}
paste -d'\n' plus_${run} minus_${run} > stripped_${run}
samtools view -H $run > header_${run}
cat header_${run} stripped_${run} > stripped_H_${run}
```

```
rm plus_$run
rm minus_$run
rm header_$run
rm stripped_$run
```

The processed files were converted to BAM files using an additional filter (filter.orphans.py). In this step all reads, which were not paired and did not map uniquely to the genomes were excluded. The BAM files were further sorted by chromosome and nucleotide coordinates.

```
chmod +x filter_orphans.py
```

```
samtools view -hf 0x2 Input.sam | grep -v "XS:i:" | filter_orphans.py | samtools
view -b -o Output.bam
samtools sort Input.bam -o Output_sort.bam
```

### **Nucleosome peak calling using the DANPOS Python program**

Sorted BAM files were used as an input file for the DANPOS program. The dpos function, which is a peak-calling algorithm to accurately define occupancy and positioning of nucleosomes. The function was used with the default settings for paired-end sequencing reads. The output WIG files were converted to TDF files for faster data display in the IGV Browser. The called nucleosomes were visualized in the IGV Browser selecting certain viral or cellular loci.

```
python danpos.py dpos -m 1 Input_sort.bam -o /Results
igvtools toTDF Input.wig Output.tdf /IGVTools/genomes/hg19_2089.chrom.sizes
```

### **Analyses of DNA fragment length distribution**

For this analysis, normalized SAM files were used to create a text file containing only information about chromosome number (column 3) and insert size (column 9). These data were further used to calculate and visualize DNA fragment length distribution.

```
for col9 in `ls cutTo_*.sam`; do
cut -f3,9 $col9 > /col_3_9_$col9.txt;
done
```

R was used to create histograms and boxplots with DNA fragment length distribution and their statistics on the cellular and EBV genomes to compare MNase digested samples with sonicated LCLs controls. The R script used for these analyses can be found on the CD under "MNase\_seq/MNaseseq\_plots\_script.R".

## **6.6. Protein analysis methods**

### **6.6.1. Whole cell lysate preparation**

Preparation of the whole cell extracts was performed with exactly  $2 \times 10^6$  of uninfected cells and cells collected at different time points after EBV infection. The cells were washed once with PBS (6 min at 300 g and 4 °C) and lysed in a volume of 200  $\mu$ l RIPA buffer (Chapter 5.8) for 30 minutes on ice. The lysates were sonicated with a Bioruptor Sonicator (30 sec on/30 sec off, HIGH setting, 4 x 5 min, on ice). Afterwards, the samples were centrifuged to remove cellular debris (20 min at maximal speed at 4 °C) and the supernatants were transferred to new tubes. The protein concentrations of the lysates were determined by Bradford's protein assay. The absorbance was measured at 595 nm on a photometer and protein concentrations were determined using a BSA standard curve.

### **6.6.2. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)**

The SDS-PAGE gels were prepared beforehand according to chapter 5.8. The gels were run using the Mini Protean Tetra Cell System (Bio-Rad). The samples were prepared as described above, diluted with Laemmli buffer (Chapter 5.8), and heat denatured for seven minutes at 95 °C in a thermomixer. The proteins were separated in the sampling gel (Chapter 5.8) at 90 V until the samples entered the separating gel. Afterwards, the voltage was increased to 110-120 V and run until the bromophenol blue dye migrated out of the gel.

### **6.6.3. Western Blot**

The separated gels were used for transferring proteins to nitrocellulose membranes. The blotting was performed in blotting buffer (chapter 5.8) at 100 V for 80 min. Afterwards, the membrane was blocked for 45 min with blocking solution (chapter 5.8) at room temperature under agitation. The blocked membrane was incubated with selected primary antibodies at certain dilutions (chapter 5.4) overnight at 4 °C. The next day, the membrane was extensively washed five times for 5 min with PBS-T buffer (chapter 5.8) and subsequently incubated with the secondary antibody (chapter 5.4) for 45 min at room temperature. Thereafter, the membrane was washed three times with PBS-T for 5 min and used for the downstream analysis. Protein signals were visualized by chemiluminescence using the ECL reagent (chapter 5.9) and either exposed to X-ray film (chapter 5.11) or the LI-COR Odyssey Infrared System (chapter 5.11).

## 7. Abbreviations

%	percent
°C	degree Celsius
2-NBDG	2-[N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl) amino]-2-deoxy-D-glucose
A	adenine
AP-1	activator protein 1
APS	ammonium persulfate
ATP	adenosine-5'-triphosphate
ATRX	alpha Thalassemia/Mental Retardation Syndrome X-Linked
BAM	Binary Alignment Map
BCL-2	B cell lymphoma 2 protein
BCR	B cell receptor
BED	Browser Extensible Data
BL	Burkitt's lymphoma
bp	base pair
BSA	bovine serum albumine
C	cytosine
C-terminus	carboxy-terminal
CAF-1	chromatin assembly factor-1
CD	cluster of differentiation
CDKI	cyclin-dependent kinase inhibitor
cDNA	complementary DNA
CENPA	centromer protein A
CMPK1	cytidine monophosphate kinase 1
CO <sub>2</sub>	carbon dioxide
COX6A1	cytochrome C Oxidase Subunit 6A1
CpG	cytosine-phosphatidyl-guanosine
CTCF	CCCTC-binding factor
CXCR4	C-X-C motif chemokine receptor 4
cytC	cytochrome C
Daxx	death-associated protein 6
DDR	DNA damage response
DE	differentially expressed
DMSO	dimethylsulfoxid
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DNMT	DNA methyltransferase
DS	dyad symmetry
dsDNA	double stranded DNA
DTT	dithiothreitol
e.g.	exempli gratia
EBER	EBV-encoded RNA



---

EBNA	EBV-encoded nuclear antigen
EBV	Epstein-Barr virus
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic
EGTA	ethyleneglycoltetraacetic
ERCC	external RNA controls consortium
et al.	et alii
ETC	electron transport chain
etc.	et cetera
EtOH	ethanol
EZH2	human enhancer of zeste
FACS	fluorescence-activated cell sorting
FAS	fatty acid synthesis
FBS	fetal bovine serum
FC	fold change
FDR	false discovery rate
Fig.	figure
FPKM	Fragments Per Kilobase of transcript per Million mapped reads
FSC-A	forward scatter area
G	guanine
g	gram
G6P	glucose-6-phosphate
GFP	green fluorescent protein
GO	Gene Ontology
gp	glycoprotein
GPI	glucose-6-phosphate isomerase
GSR	gluthathione-disulfide reductase
H <sub>2</sub> O	dihydrogen monoxide; water
H3	histone H3
H3K27ac	acetylated histone H3 at lysine 27
H3K27me3	trimethylated histone H3 at lysine 27
H3K36me3	trimethylated histone H3 at lysine 36
H3K4me1	monomethylated histone H3 at lysine 4
H3K4me3	trimethylated histone H3 at lysine 4
H3K9me3	trimethylated histone H3 at lysine 9
HAT	histone acetyltransferase
HCMV	Human Cytomegalovirus
HDAC	histone deacetylase
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
Hg19	human genome 19
HK	hexokinase
HL	Hodgkin's Lymphoma
HLA	human leukocyte antigen

HMT	histone methyltransferase
HOX	homeobox
HP1	heterochromatin protein 1
HRP	horseradish peroxidase
HSV-1	Herpes Simplex Virus 1
i.e.	id est
IFN	interferon
IgG	immunoglobulin G
IGV	integrated genome viewer
IL	interleukin
IMM	inner mitochondrial membrane
IQR	interquartile range
kb	kilo base
kDa	kilo Dalton
KLF4	Kruppel-like factor 4
KSHV	Kaposi Sarcoma associated herpesvirus
l	liter
LCL	lymphoblastoid cell line
LDH	lactate dehydrogenase
LMP	latent membrane protein
lncRNA	long non-coding RNA
LSD1	lysine-specific histone demethylase 1
M	molar
MACS	magnetic activated cell sorting
MHC	major histocompatibility complex
mHG	minimum hypergeometric
min	minute
miRNA	microRNA
ml	mililiter
mM	micromolar
MNase	micrococcal nuclease
MNase-seq	micrococcal nuclease sequencing
MOI	multiplicity of infection
mRNA	messenger RNA
N-terminus	amino-terminus
NaCl	sodium chloride
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NAP1	nucleosome assembly protein 1
ncRNA	non-coding RNA
NDR	nucleosome depleted region
NDUFA2	NADH dehydrogenase 1 alpha subcomplex subunit 2
NEB	New England Biolabs

---

NFkB	nuclear factor "kappa-light-chain-enhancer" of activated B cells
NFR	nucleosome free regions
ng	nanogram
NGS	next generation sequencing
NK	natural killer
nmol	nanomole
NPC	nasopharyngeal carcinoma
nt	nucleotide
OMM	outer mitochondrial membrane
ORC	origin recognition complex
oriLyt	origin of lytic replication
oriP	origin of plasmid replication
OXPHOS	oxydative phosphorylation
p.i.	post infection
PAA	polyacrylamide
padj	adjusted p-value
PBMCs	peripheral blood mononuclear cells
PBS	phosphate buffered saline
PBS-T	phosphate buffered saline with tween
PCA	principal component analysis
PcG	polycomb group
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PEI	polyethylenimine
pH	potential of hydrogen
PI3K	phosphoinositide 3-kinase
PIC	proteinase inhibitor cocktail
PTLD	post-transplant lymphoproliferative diseases
PML NB	promyelocytic leukemia nuclear bodies
Pol II	RNA polymerase II
PPP	pentose phosphate pathway
PRC	polycomb repressive complex
PS	phosphatidylserine residues
PTM	posttranscriptional modifications
QC	quality control
RNA	ribonucleic acid
RNA-seq	ribonucleic acid sequencing
ROS	reactive oxygen species
rpm	rounds per minute
RPM	Reads Per Million
RPMI	Roswell Park Memorial Institute
rRNA	ribosomal RNA
RT	room temperature

s	second
SAM	Sequence Alignment Map
SDS	sodium dodecyl sulphate
SDS page	SDS polyacrylamide gel electrophoresis
sec	second
SSC-A	sideward scatter area
SWI/SNF	switch/sucrose nonfermentable
T	thymine
TAD	topologically associating domains
TAP	transporter associated with antigen processing
TBE	buffer containing Tris base, boric acid and EDTA
TCA	tricarboxylic acid
TCR	T cell receptor
TE	tris-EDTA
TEMED	tetramethylethylenediamine
TF	transcription factor
TLR	toll-like receptor
TMRE	tetramethylrhodamine ethyl ester
Tris	2-amino-2-hydroxymethyl-propane-1,3-diol
tRNA	transfer RNA
TSS	transcriptional start sites
U	untis
U	uracil
USA	United States of America
V	Volt
WHO	World Health Organization
wt	wild type
ZRE	BZLF1-responsive element
µg	microgram
µl	microliter
µM	micromolar

## 8. Literature

- Adam, S., S.E. Polo, and G. Almouzni. 2014. How to restore chromatin structure and function in response to DNA damage - let the chaperones play. *FEBS J.* 281:2315–2323. doi:10.1111/febs.12793.
- Afgan, E., D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, B. Grüning, A. Guerler, J. Hillman-Jackson, G. Von Kuster, E. Rasche, N. Soranzo, N. Turaga, J. Taylor, A. Nekrutenko, and J. Goecks. 2016. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* 44:W3–W10. doi:10.1093/nar/gkw343.
- Albanese, M., T. Tagawa, M. Bouvet, L. Maliqi, D. Lutter, J. Hoser, M. Hastreiter, M. Hayes, B. Sugden, L. Martin, A. Moosmann, and W. Hammerschmidt. 2016. Epstein-Barr virus microRNAs reduce immune surveillance by virus-specific CD8+ T cells. *Proc. Natl. Acad. Sci. U. S. A.* 113:E6467–E6475. doi:10.1073/pnas.1605884113.
- Albanese, M., T. Tagawa, A. Buschle, and W. Hammerschmidt. 2017. MicroRNAs of Epstein-Barr Virus control innate and adaptive antiviral immunity. *J. Virol.* 91:e01667-16. doi:10.1128/JVI.01667-16.
- Albright, E.R., and R.F. Kalejta. 2016. Canonical and variant forms of histone H3 are deposited onto the human cytomegalovirus genome during lytic and latent infections. *J. Virol.* 90:10309. doi:10.1128/JVI.01220-16.
- Aligo, J., M. Walker, P. Bugelski, and D. Weinstock. 2015. Is murine gammaherpesvirus-68 (MHV-68) a suitable immunotoxicological model for examining immunomodulatory drug-associated viral recrudescence? *J. Immunotoxicol.* 12:1–15. doi:10.3109/1547691X.2014.882996.
- Allday, M.J., Q. Bazot, and R.E. White. 2015. The EBNA3 family: two oncoproteins and a tumour suppressor that are central to the biology of EBV in B cells. 61–117.
- Allfrey, V.G., R. Faulkner, and A.E. Mirsky. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U. S. A.* 51:786–94.
- Allis, C.D., and T. Jenuwein. 2016. The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 2016 178. 17:487. doi:10.1038/nrg.2016.59.
- Almeida, L., M. Lochner, L. Berod, and T. Sparwasser. 2016. Metabolic pathways in T cell activation and lineage differentiation. *Semin. Immunol.* 28. doi:10.1016/j.smim.2016.10.009.
- 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. doi:10.1371/journal.pbio.0030404.
- Anders, S., and W. Huber. 2010. Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. doi:10.1186/gb-2010-11-10-r106.
- Anders, S., P.T. Pyl, and W. Huber. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics.* 31:166–9. doi:10.1093/bioinformatics/btu638.
- Arbuckle, J.H., and T.M. Kristie. 2014. Epigenetic repression of herpes simplex virus infection by the nucleosome remodeler CHD3. *MBio.* 5:e01027-13. doi:10.1128/mBio.01027-13.
- Arvey, A., I. Tempera, K. Tsai, H.-S. Chen, N. Tikhmyanova, M. Klichinsky, C. Leslie, and P.M. Lieberman. 2012. An atlas of the Epstein-Barr Virus transcriptome and epigenome reveals host-virus regulatory interactions. *Cell Host Microbe.* 12:233–245.

doi:10.1016/J.CHOM.2012.06.008.

- Athanasiadis, E.I., J.G. Botthof, H. Andres, L. Ferreira, P. Lio, and A. Cvejic. 2017. Single-cell RNA-sequencing uncovers transcriptional states and fate decisions in haematopoiesis. *Nat. Commun.* 8:2045. doi:10.1038/s41467-017-02305-6.
- Bannister, A.J., and T. Kouzarides. 2011. Regulation of chromatin by histone modifications. *Cell Res.* 21:381–95. doi:10.1038/cr.2011.22.
- Bano, D., A. Piazzesi, P. Salomoni, and P. Nicotera. 2017. The histone variant H3.3 claims its place in the crowded scene of epigenetics. *Aging (Albany, NY)*. 9:602–614. doi:10.18632/aging.101194.
- Bell, S.P., and A. Dutta. 2002. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* 71:333–374. doi:10.1146/annurev.biochem.71.110601.135425.
- Berard, M., and D.F. Tough. 2002. Qualitative differences between naïve and memory T cells. *Immunology*. 106:127–38. doi:10.1046/J.1365-2567.2002.01447.X.
- Berg, J.M. (Jeremy M., J.L. Tymoczko, L. Stryer, and L. Stryer. 2002. *Biochemistry*. W.H. Freeman.
- Biswal, B.N., S.N. Das, B.K. Das, and R. Rath. 2017. Alteration of cellular metabolism in cancer cells and its therapeutic prospects. *J. Oral Maxillofac. Pathol.* 21:244–251. doi:10.4103/jomfp.JOMFP\_60\_17.
- Bléoo, S., X. Sun, M.J. Hendzel, J.M. Rowe, M. Packer, and R. Godbout. 2001. Association of human DEAD box protein DDX1 with a cleavage stimulation factor involved in 3'-end processing of pre-mRNA. *Mol. Biol. Cell.* 12:3046–59.
- Buck, M.D., D. O'Sullivan, and E.L. Pearce. 2015. T cell metabolism drives immunity. *J. Exp. Med.* 212:1345–60. doi:10.1084/jem.20151159.
- Burgess, R.J., and Z. Zhang. 2010. Histones, histone chaperones and nucleosome assembly. *Protein Cell.* 1:607–612. doi:10.1007/s13238-010-0086-y.
- Burgess, R.J., and Z. Zhang. 2013. Histone chaperones in nucleosome assembly and human disease. *Nat. Struct. Mol. Biol.* 20:14–22. doi:10.1038/nsmb.2461.
- Calender, A., M. Billaud, J.P. Aubry, J. Banchemereau, 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. *Proc. Natl. Acad. Sci. U. S. A.* 84:8060–4.
- Calo, E., and J. Wysocka. 2013. Modification of enhancer chromatin: what, how, and why? *Mol. Cell.* 49:825–837. doi:10.1016/j.molcel.2013.01.038.
- Carbone, A., A. Gloghini, and G. Dotti. 2008. EBV-associated lymphoproliferative disorders: classification and treatment. *Oncologist.* 13:577–85. doi:10.1634/theoncologist.2008-0036.
- Chen, J., M. Wang, B. Xi, J. Xue, D. He, J. Zhang, and Y. Zhao. 2012. SPARC is a key regulator of proliferation, apoptosis and invasion in human ovarian cancer. *PLoS One.* 7:e42413. doi:10.1371/journal.pone.0042413.
- Chen, K., Y. Xi, X. Pan, Z. Li, K. Kaestner, J. Tyler, S. Dent, X. He, and W. Li. 2013. DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res.* 23:341–51. doi:10.1101/gr.142067.112.
- Chen, X., D.C. Johns, D.E. Geiman, E. Marban, D.T. Dang, G. Hamlin, R. Sun, and V.W. Yang. 2001. Krüppel-like factor 4 (gut-enriched Krüppel-like factor) inhibits cell proliferation by blocking G1/S progression of the cell cycle. *J. Biol. Chem.* 276:30423–8. doi:10.1074/jbc.M101194200.
- Cheng, S., K. Caviness, J. Buehler, M. Smithey, J. Nikolich-Žugich, and F. Goodrum. 2017. Transcriptome-wide characterization of human cytomegalovirus in natural infection and experimental latency. *Proc. Natl. Acad. Sci. U. S. A.* 114:E10586–E10595. doi:10.1073/pnas.1710522114.
- Conn, K.L., M.J. Hendzel, and L.M. Schang. 2013. The differential mobilization of histones H3.1 and H3.3 by Herpes Simplex Virus 1 relates histone dynamics to the assembly of viral

- chromatin. *PLoS Pathog.* 9:e1003695. doi:10.1371/journal.ppat.1003695.
- Conn, K.L., and L.M. Schang. 2013. Chromatin dynamics during lytic infection with herpes simplex virus 1. *Viruses.* 5:1758–86. doi:10.3390/v5071758.
- Conway, J.R., A. Lex, and N. Gehlenborg. 2017. UpSetR: an R package for the visualization of intersecting sets and their properties. *Bioinformatics.* 33:2938–2940. doi:10.1093/bioinformatics/btx364.
- Cordin, O., J. Banroques, N.K. Tanner, and P. Linder. 2006. The DEAD-box protein family of RNA helicases. *Gene.* 367:17–37. doi:10.1016/J.GENE.2005.10.019.
- Countryman, J., and G. Miller. 1985. Activation of expression of latent Epstein-Barr herpesvirus after gene transfer with a small cloned subfragment of heterogeneous viral DNA. *Proc. Natl. Acad. Sci. U. S. A.* 82:4085–9.
- Cui, K., and K. Zhao. 2012. Genome-wide approaches to determining nucleosome occupancy in metazoans using MNase-Seq. *Methods Mol. Biol.* 833:413–419. doi:10.1007/978-1-61779-477-3\_24.
- Deaton, A.M., M. Gómez-Rodríguez, J. Mieczkowski, M.Y. Tolstorukov, S. Kundu, R.I. Sadreyev, L.E. Jansen, and R.E. Kingston. 2016. Enhancer regions show high histone H3.3 turnover that changes during differentiation. *Elife.* 5:e15316. doi:10.7554/eLife.15316.
- DeBerardinis, R.J., J.J. Lum, G. Hatzivassiliou, and C.B. Thompson. 2008. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab.* 7:11–20. doi:10.1016/j.cmet.2007.10.002.
- Dechend, R., F. Hirano, K. Lehmann, V. Heissmeyer, S. Ansieau, F.G. Wulczyn, C. Scheidereit, and A. Leutz. 1999. The Bcl-3 oncoprotein acts as a bridging factor between NF- $\kappa$ B/Rel and nuclear co-regulators. *Oncogene.* 18:3316–3323. doi:10.1038/sj.onc.1202717.
- Dekker, J., M.A. Marti-Renom, and L.A. Mirny. 2013. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat. Rev. Genet.* 14:390–403. doi:10.1038/nrg3454.
- Delecluse, H.J., T. Hilsendegen, D. Pich, R. Zeidler, and W. Hammerschmidt. 1998. Propagation and recovery of intact, infectious Epstein-Barr virus from prokaryotic to human cells. *Proc. Natl. Acad. Sci. U. S. A.* 95:8245–50.
- Delgado, T., P.A. Carroll, A.S. Punjabi, D. Margineantu, D.M. Hockenbery, and M. Lagunoff. 2010. Induction of the Warburg effect by Kaposi's sarcoma herpesvirus is required for the maintenance of latently infected endothelial cells. *Proc. Natl. Acad. Sci. U. S. A.* 107:10696–701. doi:10.1073/pnas.1004882107.
- Dent, A.L., A.L. Shaffer, X. Yu, D. Allman, L.M. Staudt, N. Raab-Traub, and H. Kikutani. 1997. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science (80- ).* 276:589–592. doi:10.1126/science.276.5312.589.
- Dobin, A., C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T.R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 29:15–21. doi:10.1093/bioinformatics/bts635.
- van Dongen, S., and A.J. Enright. 2012. Metric distances derived from cosine similarity and Pearson and Spearman correlations.
- Doughty, C.A., B.F. Bleiman, D.J. Wagner, F.J. Dufort, J.M. Mataraza, M.F. Roberts, and T.C. Chiles. 2006. Antigen receptor-mediated changes in glucose metabolism in B lymphocytes: role of phosphatidylinositol 3-kinase signaling in the glycolytic control of growth. *Blood.* 107:4458–65. doi:10.1182/blood-2005-12-4788.
- Dufort, F.J., B.F. Bleiman, M.R. Gumina, D. Blair, D.J. Wagner, M.F. Roberts, Y. Abu-Amer, and T.C. Chiles. 2007. Cutting edge: IL-4-mediated protection of primary B lymphocytes from apoptosis via Stat6-dependent regulation of glycolytic metabolism. *J. Immunol.* 179:4953–7.
- Eden, E., R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini. 2009. GOrilla: a tool for discovery and

- visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*. 10:48. doi:10.1186/1471-2105-10-48.
- Engelen, E., R.C. Janssens, K. Yagita, V.A.J. Smits, G.T.J. van der Horst, and F. Tamanini. 2013. Mammalian TIMELESS is involved in period determination and DNA damage-dependent phase advancing of the circadian clock. *PLoS One*. 8:e56623. doi:10.1371/journal.pone.0056623.
- Esau, D. 2017. Viral causes of lymphoma: the history of Epstein-Barr Virus and Human T-Lymphotropic Virus 1. *Virology (Auckl)*. 8:1178122X17731772. doi:10.1177/1178122X17731772.
- Fadaka, A., B. Ajiboye, O. Ojo, O. Adewale, I. Olayide, and R. Emuowhochere. 2017. Biology of glucose metabolism in cancer cells. *J. Oncol. Sci.* 3:45–51. doi:10.1016/J.JONS.2017.06.002.
- Fields, B.N., D.M. Knipe, and P.M. Howley. 2007. Fields virology. Wolters Kluwer Health/Lippincott Williams & Wilkins. 86 pp.
- Finlay, D.K. 2012. Regulation of glucose metabolism in T cells: new insight into the role of Phosphoinositide 3-kinases. *Front. Immunol.* 3:247. doi:10.3389/fimmu.2012.00247.
- Flynn, R.A., and H.Y. Chang. 2014. Long noncoding RNAs in cell-fate programming and reprogramming. *Cell Stem Cell*. 14:752–61. doi:10.1016/j.stem.2014.05.014.
- Fontaine, K.A., R. Camarda, and M. Lagunoff. 2014. Vaccinia virus requires glutamine but not glucose for efficient replication. *J. Virol.* 88:4366–74. doi:10.1128/JVI.03134-13.
- Fortes, P., and K. V Morris. 2016. Long noncoding RNAs in viral infections. *Virus Res.* 212:1–11. doi:10.1016/j.virusres.2015.10.002.
- Fowler, T., A.S. Garruss, A. Ghosh, S. De, K.G. Becker, W.H. Wood, M.T. Weirauch, S.T. Smale, B. Aronow, R. Sen, and A.L. Roy. 2015. Divergence of transcriptional landscape occurs early in B cell activation. *Epigenetics Chromatin*. 8:20. doi:10.1186/s13072-015-0012-x.
- Frappier, L. 2012. Contributions of Epstein-Barr nuclear antigen 1 (EBNA1) to cell immortalization and survival. *Viruses*. 4:1537–47. doi:10.3390/v4091537.
- Freyberg, Z., and E.T. Harvill. 2017. Pathogen manipulation of host metabolism: A common strategy for immune evasion. *PLoS Pathog.* 13:e1006669. doi:10.1371/journal.ppat.1006669.
- Friedman, J.R., and J. Nunnari. 2014. Mitochondrial form and function. *Nat.* 2014 5057483. 505:335. doi:10.1038/nature12985.
- Garrett, R.H., and C.M. Grisham. 2005. Biochemistry. Thomson Brooks/Cole.
- Garruss, A.S., and T. Fowler. 2015. Dataset of transcriptional landscape of B cell early activation. *Genomics Data*. 5:238–240. doi:10.1016/J.GDATA.2015.06.007.
- Ge, X.Q., and H. Lin. 2014. Noncoding RNAs in the regulation of DNA replication. *Trends Biochem. Sci.* 39:341–3. doi:10.1016/j.tibs.2014.06.003.
- Geiman, T.M., S.K. Durum, and K. Muegge. 1998. Characterization of gene expression, genomic structure, and chromosomal localization of Hells(Lsh). *Genomics*. 54:477–483. doi:10.1006/GENO.1998.5557.
- Good, K.L., and S.G. Tangye. 2007. Decreased expression of Kruppel-like factors in memory B cells induces the rapid response typical of secondary antibody responses. *Proc. Natl. Acad. Sci. U. S. A.* 104:13420–5. doi:10.1073/pnas.0703872104.
- Greer, E.L., and Y. Shi. 2012. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* 13:343–357. doi:10.1038/nrg3173.
- Grimes, H.L., S. Harinder, A. Olsson, N. Salomonis, B.J. Aronow, and V. Chaudhry. 2015. Single cell RNA seq for analysis of cell fate decisions. *Blood*. 126.
- Hammerschmidt, W., and B. Sugden. 1988. Identification and characterization of oriLyt, a lytic origin of DNA replication of Epstein-Barr virus. *Cell*. 55:427–33. doi:10.1016/0092-



8674(88)90028-1.

- Hatton, O.L., A. Harris-Arnold, S. Schaffert, S.M. Krams, and O.M. Martinez. 2014. The interplay between Epstein-Barr virus and B lymphocytes: implications for infection, immunity, and disease. *Immunol. Res.* 58:268–76. doi:10.1007/s12026-014-8496-1.
- Henikoff, S. 2009. Labile H3.3+H2A.Z nucleosomes mark “nucleosome-free regions.” *Nat. Genet.* 41:865–866. doi:10.1038/ng0809-865.
- Henikoff, S., and M.M. Smith. 2015. Histone variants and epigenetics. *Cold Spring Harb. Perspect. Biol.* 7:a019364. doi:10.1101/cshperspect.a019364.
- Henle, W. 1968. Evidence for viruses in acute leukemia and Burkitt’s tumor. *Cancer.* 21:580–6.
- Hernando, H., A.B.M.M.K. Islam, J. Rodriguez-Ubrevia, I. Forne, L. Ciudad, A. Imhof, C. Shannon-Lowe, and E. Ballestar. 2014. Epstein-Barr virus-mediated transformation of B cells induces global chromatin changes independent to the acquisition of proliferation. *Nucleic Acids Res.* 42:249–263. doi:10.1093/nar/gkt886.
- Hölzel, M., M. Rohmoser, M. Schlee, T. Grimm, T. Harasim, A. Malamoussi, A. Gruber-Eber, E. Kremmer, W. Hiddemann, G.W. Bornkamm, and D. Eick. 2005. Mammalian WDR12 is a novel member of the Pes1-Bop1 complex and is required for ribosome biogenesis and cell proliferation. *J. Cell Biol.* 170:367–78. doi:10.1083/jcb.200501141.
- Hu, Z., and W.-W. Tee. 2017. Enhancers and chromatin structures: regulatory hubs in gene expression and diseases. *Biosci. Rep.* 37:BSR20160183. doi:10.1042/BSR20160183.
- Huang, H., Z. Deng, O. Vladimirova, A. Wiedmer, F. Lu, P.M. Lieberman, and D.J. Patel. 2016. Structural basis underlying viral hijacking of a histone chaperone complex. *Nat. Commun.* 7:12707. doi:10.1038/ncomms12707.
- Humme, S., G. Reisbach, R. Feederle, H.-J. Delecluse, K. Bousset, W. Hammerschmidt, and A. Schepers. 2003. The EBV nuclear antigen 1 (EBNA1) enhances B cell immortalization several thousandfold. *Proc. Natl. Acad. Sci. U. S. A.* 100:10989–94. doi:10.1073/pnas.1832776100.
- Imai, K., N. Kamio, M.E. Cueno, Y. Saito, H. Inoue, I. Saito, and K. Ochiai. 2014. Role of the histone H3 lysine 9 methyltransferase Suv39 h1 in maintaining Epstein-Barr virus latency in B95-8 cells. *FEBS J.* 281:2148–2158. doi:10.1111/febs.12768.
- Ishikawa, H., Z. Ma, and G.N. Barber. 2009. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature.* 461:788–92. doi:10.1038/nature08476.
- Jang, K.-J., H. Mano, K. Aoki, T. Hayashi, A. Muto, Y. Nambu, K. Takahashi, K. Itoh, S. Taketani, S.L. Nutt, K. Igarashi, A. Shimizu, and M. Sugai. 2015. Mitochondrial function provides instructive signals for activation-induced B-cell fates. *Nat. Commun.* 6:6750. doi:10.1038/ncomms7750.
- Jenuwein, T., and C.D. Allis. 2001. Translating the histone code. *Science.* 293:1074–80. doi:10.1126/science.1063127.
- Jin, C., and G. Felsenfeld. 2007. Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev.* 21:1519–29. doi:10.1101/gad.1547707.
- 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. doi:10.1371/journal.ppat.1002704.
- 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:E1396-404. doi:10.1073/pnas.1115906109.
- Kalamvoki, M., and B. Roizman. 2014. HSV-1 degrades, stabilizes, requires, or is stung by STING depending on ICP0, the US3 protein kinase, and cell derivation. *Proc. Natl. Acad. Sci. U. S. A.* 111:E611-7. doi:10.1073/pnas.1323414111.
- 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. doi:10.1128/JVI.06314-11.
- Kalla, M., A. Schmeink, 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. doi:10.1073/pnas.0911948107.
- Kang, H., A. Wiedmer, Y. Yuan, E. Robertson, and P.M. Lieberman. 2011. Coordination of KSHV latent and lytic gene control by CTCF-cohesin mediated chromosome conformation. *PLoS Pathog.* 7:e1002140. doi:10.1371/journal.ppat.1002140.
- Kari, V., O. Karpiuk, B. Tieg, M. Kriegs, E. Dikomey, H. Krebber, Y. Begus-Nahrman, and S.A. Johnsen. 2013. A subset of histone H2B genes produces polyadenylated mRNAs under a variety of cellular conditions. *PLoS One.* 8:e63745. doi:10.1371/journal.pone.0063745.
- Karniely, S., M.P. Weekes, R. Antrobus, J. Rorbach, L. van Haute, Y. Umrana, D.L. Smith, R.J. Stanton, M. Minczuk, P.J. Lehner, and J.H. Sinclair. 2016. Human Cytomegalovirus infection upregulates the mitochondrial transcription and translation machineries. *MBio.* 7:e00029. doi:10.1128/mBio.00029-16.
- Kebede, A.F., R. Schneider, and S. Daujat. 2015. Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. *FEBS J.* 282:1658–1674. doi:10.1111/febs.13047.
- Kempkes, B., and P.D. Ling. 2015. EBNA2 and its coactivator EBNA-LP. Springer, Cham. 35–59.
- Kieffer-Kwon, K.-R., K. Nimura, S.S.P. Rao, J. Xu, S. Jung, A. Pekowska, M. Dose, E. Stevens, E. Mathe, P. Dong, S.-C. Huang, M.A. Ricci, L. Baranello, Y. Zheng, F. Tomassoni Ardori, W. Resch, D. Stavreva, S. Nelson, M. McAndrew, A. Casellas, E. Finn, C. Gregory, B.G. St Hilaire, S.M. Johnson, W. Dubois, M.P. Cosma, E. Batchelor, D. Levens, R.D. Phair, T. Misteli, L. Tessarollo, G. Hager, M. Lakadamyali, Z. Liu, M. Floer, H. Shroff, E.L. Aiden, and R. Casellas. 2017. Myc regulates chromatin decompaction and nuclear architecture during B Cell activation. *Mol. Cell.* 67:566–578.e10. doi:10.1016/j.molcel.2017.07.013.
- Kieser, A., and K.R. Sterz. 2015. The Latent Membrane Protein 1 (LMP1). Springer, Cham. 119–149.
- Knipe, D.M., P.M. Lieberman, J.U. Jung, A.A. McBride, K. V. Morris, M. Ott, D. Margolis, A. Nieto, M. Nevels, R.J. Parks, and T.M. Kristie. 2013. Snapshots: Chromatin control of viral infection. *Virology.* 435:141–156. doi:10.1016/J.VIROL.2012.09.023.
- Kolovos, P., T.A. Knoch, F.G. Grosveld, P.R. Cook, and A. Papantonis. 2012. Enhancers and silencers: an integrated and simple model for their function. *Epigenetics Chromatin.* 5:1. doi:10.1186/1756-8935-5-1.
- Krude, T. 1999. Chromatin assembly during DNA replication in somatic cells. *Eur. J. Biochem.* 263:1–5. doi:10.1046/j.1432-1327.1999.00508.x.
- Kühlbrandt, W. 2015. Structure and function of mitochondrial membrane protein complexes. *BMC Biol.* 13:89. doi:10.1186/s12915-015-0201-x.
- Lagunoff, M. 2016. Activation of cellular metabolism during latent Kaposi's Sarcoma herpesvirus infection. *Curr. Opin. Virol.* 19:45–9. doi:10.1016/j.coviro.2016.06.012.
- Lai, W.K.M., and B.F. Pugh. 2017. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat. Rev. Mol. Cell Biol.* 18:548–562. doi:10.1038/nrm.2017.47.
- Langfelder, P., B. Zhang, and S. Horvath. 2008. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics.* 24:719–720. doi:10.1093/bioinformatics/btm563.
- Langmead, B., and S.L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* 9:357–359. doi:10.1038/nmeth.1923.

- Langmead, B., C. Trapnell, M. Pop, and S.L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10:R25. doi:10.1186/gb-2009-10-3-r25.
- Lawrence, M., S. Daujat, and R. Schneider. 2016. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet.* 32:42–56. doi:10.1016/J.TIG.2015.10.007.
- Lee, S.-B., C.-F. Lee, D.S.-C. Ou, K. Dulal, L.-H. Chang, C.-H. Ma, C.-F. Huang, H. Zhu, Y.-S. Lin, and L.-J. Juan. 2011. Host-viral effects of chromatin assembly factor 1 interaction with HCMV IE2. *Cell Res.* 21:1230–47. doi:10.1038/cr.2011.53.
- Leng, N., L.-F. Chu, C. Barry, Y. Li, J. Choi, X. Li, P. Jiang, R.M. Stewart, J.A. Thomson, and C. Kendzierski. 2015. Oscope identifies oscillatory genes in unsynchronized single-cell RNA-seq experiments. *Nat. Methods.* 12:947–950. doi:10.1038/nmeth.3549.
- Lex, A., N. Gehlenborg, H. Strobel, R. Vuillemot, and H. Pfister. 2014. UpSet: Visualization of intersecting sets. *IEEE Trans. Vis. Comput. Graph.* 20:1983–92. doi:10.1109/TVCG.2014.2346248.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 25:2078–2079. doi:10.1093/bioinformatics/btp352.
- Li, L., and Y. Wang. 2017. Cross-talk between the H3K36me3 and H4K16ac histone epigenetic marks in DNA double-strand break repair. *J. Biol. Chem.* 292:11951–11959. doi:10.1074/jbc.M117.788224.
- Li, X.L., M. Subramanian, M.F. Jones, R. Chaudhary, D.K. Singh, X. Zong, B. Gryder, S. Sindri, M. Mo, A. Schetter, X. Wen, S. Parvathaneni, D. Kazandjian, L.M. Jenkins, W. Tang, F. Elloumi, J.L. Martindale, M. Huarte, Y. Zhu, A.I. Robles, S.M. Frier, F. Rigo, M. Cam, S. Ambs, S. Sharma, C.C. Harris, M. Dasso, K. V Prasanth, and A. Lal. 2017. Long noncoding RNA PURPL suppresses basal p53 levels and promotes tumorigenicity in colorectal cancer. *Cell Rep.* 20:2408–2423. doi:10.1016/j.celrep.2017.08.041.
- Lieberman, P.M. 2013. Keeping it quiet: chromatin control of gammaherpesvirus latency. *Nat. Rev. Microbiol.* 11:863–75. doi:10.1038/nrmicro3135.
- Liou, J.-Y., G.E. Dutschman, W. Lam, Z. Jiang, and Y.-C. Cheng. 2002. Characterization of human UMP/CMP kinase and its phosphorylation of D- and L-form deoxycytidine analogue monophosphates. *Cancer Res.* 62:1624–31.
- Lipford, J.R., and S.P. Bell. 2001. Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell.* 7:21–30. doi:10.1016/S1097-2765(01)00151-4.
- Liu, P.-S., and P.-C. Ho. 2017. Mitochondria: A master regulator in macrophage and T cell immunity. *Mitochondrion.* doi:10.1016/J.MITO.2017.11.002.
- Liu, W., and C. Ding. 2017. Roles of lncRNAs in viral infections. *Front. Cell. Infect. Microbiol.* 7:205. doi:10.3389/fcimb.2017.00205.
- Love, M.I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi:10.1186/s13059-014-0550-8.
- Luger, K., A.W. Mäder, R.K. Richmond, D.F. Sargent, and T.J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature.* 389:251–60. doi:10.1038/38444.
- Lunt, S.Y., and M.G. Vander Heiden. 2011. Aerobic glycolysis: Meeting the metabolic requirements of cell proliferation. *Annu. Rev. Cell Dev. Biol.* 27:441–464. doi:10.1146/annurev-cellbio-092910-154237.
- Lygerou, Z., P. Mitchell, E. Petfalski, B. Séraphin, and D. Tollervy. 1994. The POP1 gene encodes a protein component common to the RNase MRP and RNase P ribonucleoproteins. *Genes Dev.* 8:1423–33.
- Mancao, C., and W. Hammerschmidt. 2007. Epstein-Barr virus latent membrane protein 2A is a B-cell receptor mimic and essential for B-cell survival. *Blood.* 110:3715–21.

- doi:10.1182/blood-2007-05-090142.
- Marchese, F.P., and M. Huarte. 2014. Long non-coding RNAs and chromatin modifiers: their place in the epigenetic code. *Epigenetics*. 9:21–6. doi:10.4161/epi.27472.
- Marzluff, W.F., E.J. Wagner, and R.J. Duronio. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat. Rev. Genet.* 9:843–54. doi:10.1038/nrg2438.
- McHugh, C.A., C.-K. Chen, A. Chow, C.F. Surka, C. Tran, P. McDonel, A. Pandya-Jones, M. Blanco, C. Burghard, A. Moradian, M.J. Sweredoski, A.A. Shishkin, J. Su, E.S. Lander, S. Hess, K. Plath, and M. Guttman. 2015. The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*. 521:232–6. doi:10.1038/nature14443.
- Milavetz, B.I., and L. Balakrishnan. 2015. Viral epigenetics. *Methods Mol. Biol.* 1238:569–96. doi:10.1007/978-1-4939-1804-1\_30.
- Miller, G., T. Shope, H. Lisco, D. Stitt, and M. Lipman. 1972. Epstein-Barr virus: transformation, cytopathic changes, and viral antigens in squirrel monkey and marmoset leukocytes. *Proc. Natl. Acad. Sci. U. S. A.* 69:383–7.
- Miyake-Stoner, S.J., and C.C. O’Shea. 2014. Metabolism goes viral. *Cell Metab.* 19:549–50. doi:10.1016/j.cmet.2014.03.022.
- Mondal, T., M. Rasmussen, G.K. Pandey, A. Isaksson, and C. Kanduri. 2010. Characterization of the RNA content of chromatin. *Genome Res.* 20:899–907. doi:10.1101/gr.103473.109.
- Münz, C. 2015. Epstein Barr virus. Volume 1 : one herpes virus : many diseases. 391 pp.
- Nagalakshmi, U., K. Waern, M. Snyder, U. Nagalakshmi, K. Waern, and M. Snyder. 2010. RNA-seq: A method for comprehensive transcriptome analysis. *In Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., Hoboken, NJ, USA. 4.11.1-4.11.13.
- Nie, Y., Y.-C. Han, and Y.-R. Zou. 2008. CXCR4 is required for the quiescence of primitive hematopoietic cells. *J. Exp. Med.* 205:777–83. doi:10.1084/jem.20072513.
- Nitzsche, A., C. Paulus, and M. Nevels. 2008. Temporal dynamics of cytomegalovirus chromatin assembly in productively infected human cells. *J. Virol.* 82:11167–80. doi:10.1128/JVI.01218-08.
- Oh, J., N. Ruskoski, and N.W. Fraser. 2012. Chromatin assembly on herpes simplex virus 1 DNA early during a lytic infection is Asf1a dependent. *J. Virol.* 86:12313–21. doi:10.1128/JVI.01570-12.
- Okamura, K., A. Ishizuka, H. Siomi, and M.C. Siomi. 2004. Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. *Genes Dev.* 18:1655–66. doi:10.1101/gad.1210204.
- Olivadoti, M., L.A. Toth, J. Weinberg, and M.R. Opp. Murine Gammaherpesvirus 68: A model for the study of Epstein-Barr Virus infections and eelated diseases.
- Pagano, J.S. 1999. Epstein-Barr Virus: The first human tumor virus and its role in cancer. *Proc. Assoc. Am. Physicians.* 111:573–580. doi:10.1046/j.1525-1381.1999.t01-1-99220.x.
- Parrado, A., M.E. Noguera, A. Delmer, S. McKenna, J. Davies, I. Le Gall, P. Bentley, J.A. Whittaker, F. Sigaux, C. Chomienne, and R.A. Padua. 2000. Dereglated expression of promyelocytic leukemia zinc finger protein in B-cell chronic lymphocytic leukemias does not affect cyclin A expression. *Hematol. J. Off. J. Eur. Haematol. Assoc.* 1:15–27. doi:10.1038/sj/thj/6200012.
- 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 Res.* 40:7233–46. doi:10.1093/nar/gks391.
- Pennock, N.D., J.T. White, E.W. Cross, E.E. Cheney, B.A. Tamburini, and R.M. Kedl. 2013. T cell responses: naive to memory and everything in between. *Adv. Physiol. Educ.* 37:273–83. doi:10.1152/advan.00066.2013.
- Peterson, C.L., and M.-A. Laniel. 2004. Histones and histone modifications. *Curr. Biol.* 14:R546-51. doi:10.1016/j.cub.2004.07.007.

- Phillips, D.M. 1963. The presence of acetyl groups of histones. *Biochem. J.* 87:258–63.
- Placek, B.J., J. Huang, J.R. Kent, J. Dorsey, L. Rice, N.W. Fraser, and S.L. Berger. 2009. The histone variant H3.3 regulates gene expression during lytic infection with herpes simplex virus type 1. *J. Virol.* 83:1416–21. doi:10.1128/JVI.01276-08.
- Price, A.M., J. Dai, Q. Bazot, L. Patel, P.A. Nikitin, R. Djavadian, P.S. Winter, C.A. Salinas, A.P. Barry, K.C. Wood, E.C. Johannsen, A. Letai, M.J. Allday, and M.A. Luftig. 2017. Epstein-Barr virus ensures B cell survival by uniquely modulating apoptosis at early and late times after infection. *Elife.* 6:e22509. doi:10.7554/eLife.22509.
- Price, A.M., and M.A. Luftig. 2014. Dynamic Epstein-Barr virus gene expression on the path to B-cell transformation. *Adv. Virus Res.* 88:279–313. doi:10.1016/B978-0-12-800098-4.00006-4.
- Purushothaman, P., S. Thakker, and S.C. Verma. 2015. Transcriptome analysis of Kaposi's sarcoma-associated herpesvirus during de novo primary infection of human B and endothelial cells. *J. Virol.* 89:3093–111. doi:10.1128/JVI.02507-14.
- Quinlan, A.R., and I.M. Hall. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 26:841–2. doi:10.1093/bioinformatics/btq033.
- Rai, T.S., M. Glass, J.J. Cole, M.I. Rather, M. Marsden, M. Neilson, C. Brock, I.R. Humphreys, R.D. Everett, and P.D. Adams. 2017. Histone chaperone HIRA deposits histone H3.3 onto foreign viral DNA and contributes to anti-viral intrinsic immunity. *Nucleic Acids Res.* 45:11673–11683. doi:10.1093/nar/gkx771.
- Rakhmanov, M., H. Sic, A.-K. Kienzler, B. Fischer, M. Rizzi, M. Seidl, K. Melkaoui, S. Unger, L. Moehle, N.E. Schmit, S.D. Deshmukh, C.K. Ayata, W. Schuh, Z. Zhang, F.-L. Cosset, E. Verhoeyen, H.-H. Peter, R.E. Voll, U. Salzer, H. Eibel, and K. Warnatz. 2014. High levels of SOX5 decrease proliferative capacity of human B cells, but permit plasmablast differentiation. *PLoS One.* 9:e100328. doi:10.1371/journal.pone.0100328.
- Richmond, T.J., and C.A. Davey. 2003. The structure of DNA in the nucleosome core. *Nature.* 423:145–150. doi:10.1038/nature01595.
- Rinn, J.L., and H.Y. Chang. 2012. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81:145–66. doi:10.1146/annurev-biochem-051410-092902.
- Robinson, J.T., H. Thorvaldsdóttir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, and J.P. Mesirov. 2011. Integrative genomics viewer. *Nat. Biotechnol.* 29:24–26. doi:10.1038/nbt.1754.
- Ron-Harel, N., D. Santos, J.M. Ghergurovich, P.T. Sage, A. Reddy, S.B. Lovitch, N. Dephoure, F.K. Satterstrom, M. Sheffer, J.B. Spinelli, S. Gygi, J.D. Rabinowitz, A.H. Sharpe, and M.C. Haigis. 2016. Mitochondrial biogenesis and proteome remodeling promote one-carbon metabolism for T cell activation. *Cell Metab.* 24:104–17. doi:10.1016/j.cmet.2016.06.007.
- Rorbach, J., P. Boesch, P.A. Gammage, T.J.J. Nicholls, S.F. Pearce, D. Patel, A. Hauser, F. Perocchi, and M. Minczuk. 2014. MRM2 and MRM3 are involved in biogenesis of the large subunit of the mitochondrial ribosome. *Mol. Biol. Cell.* 25:2542–55. doi:10.1091/mbc.E14-01-0014.
- Rossetto, D., N. Avvakumov, and J. Côté. 2012. Histone phosphorylation. *Epigenetics.* 7:1098–1108. doi:10.4161/epi.21975.
- Rozanska, A., R. Richter-Dennerlein, J. Rorbach, F. Gao, R.J. Lewis, Z.M. Chrzanowska-Lightowlers, and R.N. Lightowlers. 2017. The human RNA-binding protein RBFA promotes the maturation of the mitochondrial ribosome. *Biochem. J.* 474:2145–2158. doi:10.1042/BCJ20170256.
- Sadler, A.J., F.J. Rossello, L. Yu, J.A. Deane, X. Yuan, D. Wang, A.T. Irving, M. Kaparakis-Liaskos, M.P. Gantier, H. Ying, H.C.H. Yim, E.L. Hartland, A.J. Notini, S. de Boer, S.J. White, A. Mansell, J.-P. Liu, D.N. Watkins, S. Gerondakis, B.R.G. Williams, and D. Xu. 2015. BTB-ZF transcriptional regulator PLZF modifies chromatin to restrain inflammatory signaling

- programs. *Proc. Natl. Acad. Sci. U. S. A.* 112:1535–40. doi:10.1073/pnas.1409728112.
- Sambrook, J., and D.W. (David W. Russell. 2001. *Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory Press.
- Sanchez, E.L., and M. Lagunoff. 2015. Viral activation of cellular metabolism. *Virology*. 479–480:609–18. doi:10.1016/j.virol.2015.02.038.
- Sanchez, E.L., T.H. Pulliam, T.A. Dimairo, A.B. Thalhofer, T. Delgado, and M. Lagunoff. 2017. Glycolysis, glutaminolysis, and fatty acid synthesis are required for distinct stages of Kaposi's Sarcoma-Associated Herpesvirus lytic replication. *J. Virol.* 91. doi:10.1128/JVI.02237-16.
- Sandoval, H., S. Kodali, and J. Wang. 2017. Regulation of B cell fate, survival, and function by mitochondria and autophagy. *Mitochondrion*. doi:10.1016/J.MITO.2017.11.005.
- Sangaletti, S., C. Tripodo, P. Portararo, M. Dugo, C. Vitali, L. Botti, C. Guarnotta, B. Cappetti, A. Gulino, I. Torselli, P. Casalini, C. Chiodoni, and M.P. Colombo. 2014. Stromal niche communalities underscore the contribution of the matricellular protein SPARC to B-cell development and lymphoid malignancies. *Oncoimmunology*. 3:e28989. doi:10.4161/onci.28989.
- Scialdone, A., K.N. Natarajan, L.R. Saraiva, V. Proserpio, S.A. Teichmann, O. Stegle, J.C. Marioni, and F. Buettner. 2015. Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods*. 85:54–61. doi:10.1016/J.YMETH.2015.06.021.
- Sena, L.A., S. Li, A. Jairaman, M. Prakriya, T. Ezponda, D.A. Hildeman, C.-R. Wang, P.T. Schumacker, J.D. Licht, H. Perlman, P.J. Bryce, and N.S. Chandel. 2013. Mitochondria are required for antigen-specific T cell activation through Reactive Oxygen Species signaling. *Immunity*. 38:225–236.
- Seto, E., A. Moosmann, S. Grömminger, N. Walz, A. Grundhoff, and W. Hammerschmidt. 2010. Micro RNAs of Epstein-Barr Virus promote cell cycle progression and prevent apoptosis of primary human B cells. *PLoS Pathog.* 6:e1001063. doi:10.1371/journal.ppat.1001063.
- Shlyueva, D., G. Stampfel, and A. Stark. 2014. Transcriptional enhancers: from properties to genome-wide predictions. *Nat. Rev. Genet.* 15:272–286. doi:10.1038/nrg3682.
- Sijmons, S., M. Van Ranst, and P. Maes. 2014. Genomic and functional characteristics of human cytomegalovirus revealed by next-generation sequencing. *Viruses*. 6:1049–72. doi:10.3390/v6031049.
- Speck, P., K.M. Haan, and R. Longnecker. 2000. Epstein-Barr Virus entry into cells. *Virology*. 277:1–5. doi:10.1006/VIRO.2000.0624.
- Spencer, C.A., M.E. Dahmus, and S.A. Rice. 1997. Repression of host RNA polymerase II transcription by herpes simplex virus type 1. *J. Virol.* 71:2031–40.
- Spies, D., and C. Ciaudo. 2015. Dynamics in transcriptomics: Advancements in RNA-seq time course and downstream analysis. *Comput. Struct. Biotechnol. J.* 13:469–77. doi:10.1016/j.csbj.2015.08.004.
- Struhl, K., and E. Segal. 2013. Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* 20:267–73. doi:10.1038/nsmb.2506.
- Styles, C.T., Q. Bazot, G.A. Parker, R.E. White, K. Paschos, and M.J. Allday. 2017. EBV epigenetically suppresses the B cell-to-plasma cell differentiation pathway while establishing long-term latency. *PLoS Biol.* 15:e2001992. doi:10.1371/journal.pbio.2001992.
- Sugiyama, T., H. Kohara, M. Noda, and T. Nagasawa. 2006. Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity*. 25:977–88. doi:10.1016/j.immuni.2006.10.016.
- Supek, F., M. Bošnjak, N. Škunca, and T. Šmuc. 2011. REVIGO summarizes and visualizes long lists of Gene Ontology terms. *PLoS One*. 6:e21800. doi:10.1371/journal.pone.0021800.

- Sychev, Z.E., A. Hu, T.A. DiMaio, A. Gitter, N.D. Camp, W.S. Noble, A. Wolf-Yadlin, and M. Lagunoff. 2017. Integrated systems biology analysis of KSHV latent infection reveals viral induction and reliance on peroxisome mediated lipid metabolism. *PLoS Pathog.* 13:e1006256. doi:10.1371/journal.ppat.1006256.
- Tagawa, T., M. Albanese, M. Bouvet, A. Moosmann, J. Mautner, V. Heissmeyer, C. Zielinski, D. Lutter, J. Hoser, M. Hastreiter, M. Hayes, B. Sugden, and W. Hammerschmidt. 2016. Epstein-Barr viral miRNAs inhibit antiviral CD4+ T cell responses targeting IL-12 and peptide processing. *J. Exp. Med.* 213:2065–80. doi:10.1084/jem.20160248.
- Takada, K., N. Shimizu, S. Sakuma, and Y. Ono. 1986. trans activation of the latent Epstein-Barr virus (EBV) genome after transfection of the EBV DNA fragment. *J. Virol.* 57:1016–22.
- Takahashi, K., and S. Yamanaka. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* 126:663–76. doi:10.1016/j.cell.2006.07.024.
- Tanackovic, G., and A. Krämer. 2005. Human splicing factor SF3a, but not SF1, is essential for pre-mRNA splicing in vivo. *Mol. Biol. Cell.* 16:1366–77. doi:10.1091/mbc.E04-11-1034.
- Teague, T.K., L. Munn, K. Zygourakis, and B.W. McIntyre. 1993. Analysis of lymphocyte activation and proliferation by video microscopy and digital imaging. *Cytometry.* 14:772–782. doi:10.1002/cyto.990140710.
- Tempera, I., A. Wiedmer, J. Dheekollu, and P.M. Lieberman. 2010. CTCF prevents the epigenetic drift of EBV latency promoter Qp. *PLoS Pathog.* 6:e1001048. doi:10.1371/journal.ppat.1001048.
- Thomas, J.O., and R.D. Kornberg. 1975. An octamer of histones in chromatin and free in solution. *Proc. Natl. Acad. Sci. U. S. A.* 72:2626–30.
- Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14:178–192. doi:10.1093/bib/bbs017.
- Toth, Z., K. Brulois, and J.U. Jung. 2013. The chromatin landscape of Kaposi's sarcoma-associated herpesvirus. *Viruses.* 5:1346–73. doi:10.3390/v5051346.
- Tsai, K., L. Chan, R. Gibeault, K. Conn, J. Dheekollu, J. Domsic, R. Marmorstein, L.M. Schang, and P.M. Lieberman. 2014. Viral reprogramming of the Daxx histone H3.3 chaperone during early Epstein-Barr virus infection. *J. Virol.* 88:14350–63. doi:10.1128/JVI.01895-14.
- Valvona, C.J., H.L. Fillmore, P.B. Nunn, and G.J. Pilkington. 2016. The regulation and function of lactate dehydrogenase A: Therapeutic potential in brain tumor. *Brain Pathol.* 26:3–17. doi:10.1111/bpa.12299.
- Vanacova, S., and R. Stefl. 2007. The exosome and RNA quality control in the nucleus. *EMBO Rep.* 8:651–7. doi:10.1038/sj.embor.7401005.
- Vastag, L., E. Koyuncu, S.L. Grady, T.E. Shenk, and J.D. Rabinowitz. 2011. Divergent effects of Human Cytomegalovirus and Herpes Simplex Virus-1 on cellular metabolism. *PLoS Pathog.* 7:e1002124. doi:10.1371/journal.ppat.1002124.
- Venkatesh, S., and J.L. Workman. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* 16:178–189. doi:10.1038/nrm3941.
- Verdone, L., M. Caserta, and E. Di Mauro. 2005. Role of histone acetylation in the control of gene expression. *Biochem. Cell Biol.* 83:344–353. doi:10.1139/o05-041.
- Viollet, C., D.A. Davis, M. Reczko, J.M. Ziegelbauer, F. Pezzella, J. Ragoussis, and R. Yarchoan. 2015. Next-Generation Sequencing analysis reveals differential expression profiles of miRNA-mRNA target pairs in KSHV-infected cells. *PLoS One.* 10:e0126439. doi:10.1371/journal.pone.0126439.
- Viollet, C., D.A. Davis, S.S. Tekeste, M. Reczko, J.M. Ziegelbauer, F. Pezzella, J. Ragoussis, and R. Yarchoan. 2017. RNA sequencing reveals that Kaposi Sarcoma-associated herpesvirus

- infection mimics hypoxia gene expression signature. *PLoS Pathog.* 13:e1006143. doi:10.1371/journal.ppat.1006143.
- Voon, H.P.J., and L.H. Wong. 2016. New players in heterochromatin silencing: histone variant H3.3 and the ATRX/DAXX chaperone. *Nucleic Acids Res.* 44:1496–501. doi:10.1093/nar/gkw012.
- Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 10:57–63. doi:10.1038/nrg2484.
- Warburg, O. 1925. The metabolism of carcinoma cells. *J. Cancer Res.* 9:148–163. doi:10.1158/jcr.1925.148.
- Weber, C.M., and S. Henikoff. 2014. Histone variants: dynamic punctuation in transcription. *Genes Dev.* 28:672–82. doi:10.1101/gad.238873.114.
- Wei, G., G. Hu, K. Cui, and K. Zhao. 2012. Genome-wide mapping of nucleosome occupancy, histone modifications, and gene expression using next-generation sequencing technology. *Methods Enzymol.* 513:297–313. doi:10.1016/B978-0-12-391938-0.00013-6.
- Wei, H., and M.-M. Zhou. 2010. Viral-encoded enzymes that target host chromatin functions. *Biochim. Biophys. Acta - Gene Regul. Mech.* 1799:296–301. doi:10.1016/J.BBAGRM.2009.08.007.
- Weinberg, S.E., L.A. Sena, and N.S. Chandel. 2015. Mitochondria in the regulation of innate and adaptive immunity. *Immunity.* 42:406–17. doi:10.1016/j.immuni.2015.02.002.
- Weinhold, B. 2006. Epigenetics: the science of change. *Environ. Health Perspect.* 114:A160-7.
- White, R.E., P.C. Rämer, K.N. Naresh, S. Meixlsperger, L. Pinaud, C. Rooney, B. Savoldo, R. Coutinho, C. Bödör, J. Gribben, H.A. Ibrahim, M. Bower, J.P. Nourse, M.K. Gandhi, J. Middeldorp, F.Z. Cader, P. Murray, C. Münz, and M.J. Allday. 2012. EBNA3B-deficient EBV promotes B cell lymphomagenesis in humanized mice and is found in human tumors. *J. Clin. Invest.* 122:1487–502. doi:10.1172/JCI58092.
- Winter, J., and S. Diederichs. 2011. Argonaute proteins regulate microRNA stability: Increased microRNA abundance by Argonaute proteins is due to microRNA stabilization. *RNA Biol.* 8:1149–1157. doi:10.4161/rna.8.6.17665.
- Woellmer, A., J.M. Arteaga-Salas, and W. Hammerschmidt. 2012. BZLF1 governs CpG-methylated chromatin of Epstein-Barr Virus reversing epigenetic repression. *PLoS Pathog.* 8:e1002902. doi:10.1371/journal.ppat.1002902.
- Woellmer, A., and W. Hammerschmidt. 2013. Epstein-Barr virus and host cell methylation: regulation of latency, replication and virus reactivation. *Curr. Opin. Virol.* 3:260–5. doi:10.1016/j.coviro.2013.03.005.
- Wu, L., E. Fossum, C.H. Joo, K.-S. Inn, Y.C. Shin, E. Johannsen, L.M. Hutt-Fletcher, J. Hass, and J.U. Jung. 2009. Epstein-Barr virus LF2: an antagonist to type I interferon. *J. Virol.* 83:1140–6. doi:10.1128/JVI.00602-08.
- Xiao, L., Z.-Y. Hu, X. Dong, Z. Tan, W. Li, M. Tang, L. Chen, L. Yang, Y. Tao, Y. Jiang, J. Li, B. Yi, B. Li, S. Fan, S. You, X. Deng, F. Hu, L. Feng, A.M. Bode, Z. Dong, L.-Q. Sun, and Y. Cao. 2014. Targeting Epstein-Barr virus oncoprotein LMP1-mediated glycolysis sensitizes nasopharyngeal carcinoma to radiation therapy. *Oncogene.* 33:4568–78. doi:10.1038/onc.2014.32.
- Xu, D., K. Brumm, and L. Zhang. 2006. The latent membrane protein 1 of Epstein-Barr virus (EBV) primes EBV latency cells for type I interferon production. *J. Biol. Chem.* 281:9163–9. doi:10.1074/jbc.M511884200.
- Xu, D., M. Holko, A.J. Sadler, B. Scott, S. Higashiyama, W. Berkofsky-Fessler, M.J. McConnell, P.P. Pandolfi, J.D. Licht, and B.R.G. Williams. Promyelocytic leukemia zinc finger protein regulates interferon-mediated innate immunity. *Immunity.* 30:802–816. doi:10.1016/j.immuni.2009.04.013.

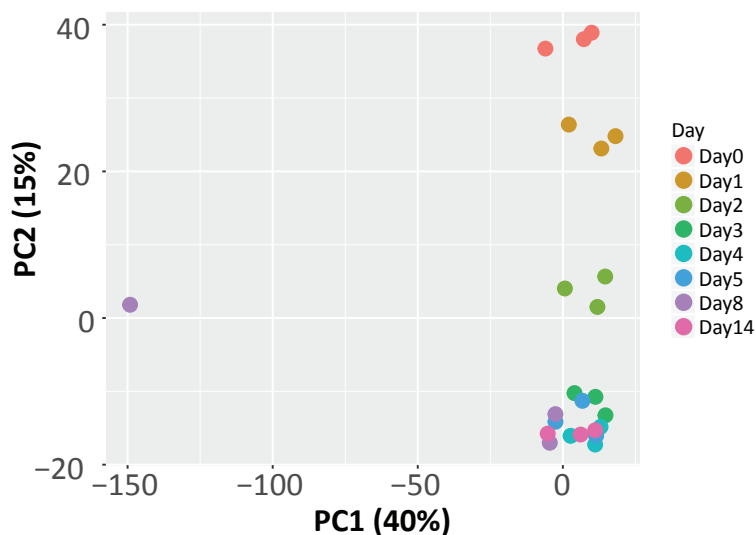


- Yao, H., K. Brick, Y. Evrard, T. Xiao, R.D. Camerini-Otero, and G. Felsenfeld. 2010. Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev.* 24:2543–2555. doi:10.1101/gad.1967810.
- Yogev, O., D. Lagos, T. Enver, and C. Boshoff. 2014. Kaposi's Sarcoma Herpesvirus microRNAs induce metabolic transformation of infected cells. *PLoS Pathog.* 10:e1004400. doi:10.1371/journal.ppat.1004400.
- Yoshida, N., D. Kitayama, M. Arima, A. Sakamoto, A. Inamine, H. Watanabe-Takano, M. Hatano, T. Koike, and T. Tokuhisa. 2011. CXCR4 expression on activated B cells is downregulated by CD63 and IL-21. *J. Immunol.* 186:2800–8. doi:10.4049/jimmunol.1003401.
- Young, L.S., J.R. Arrand, and P.G. Murray. 2007. EBV gene expression and regulation. Cambridge University Press.
- Young, L.S., and A.B. Rickinson. 2004. Epstein–Barr virus: 40 years on. *Nat. Rev. Cancer.* 4:757–768. doi:10.1038/nrc1452.
- Young, L.S., L.F. Yap, and P.G. Murray. 2016. Epstein–Barr virus: more than 50 years old and still providing surprises. *Nat. Rev. Cancer.* 16:789–802. doi:10.1038/nrc.2016.92.
- Yu, Y., A.J. Clippinger, and J.C. Alwine. 2011a. Viral effects on metabolism: changes in glucose and glutamine utilization during human cytomegalovirus infection. *Trends Microbiol.* 19:360–7. doi:10.1016/j.tim.2011.04.002.
- Yu, Y., T.G. Maguire, and J.C. Alwine. 2011b. Human cytomegalovirus activates glucose transporter 4 expression to increase glucose uptake during infection. *J. Virol.* 85:1573–80. doi:10.1128/JVI.01967-10.
- Yusuf, I., and D.A. Fruman. 2003. Regulation of quiescence in lymphocytes. *Trends Immunol.* 24:380–6. doi:10.1016/S1471-4906(03)00141-8.
- Yusuf, I., M.G. Kharas, J. Chen, R.Q. Peralta, A. Maruniak, P. Sareen, V.W. Yang, K.H. Kaestner, and D.A. Fruman. 2008. KLF4 is a FOXO target gene that suppresses B cell proliferation. *Int. Immunol.* 20:671–681. doi:10.1093/intimm/dxn024.
- Zeidler, R., G. Eissner, P. Meissner, S. Uebel, R. Tampé, S. Lazis, and W. Hammerschmidt. 1997. Downregulation of TAP1 in B lymphocytes by cellular and Epstein-Barr Virus–encoded Interleukin-10. *Blood.* 90.
- Zentner, G.E., and S. Henikoff. 2013. Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.* 20:259–266. doi:10.1038/nsmb.2470.
- Zhang, Q., M.-M. Lai, Y.-Y. Lou, B.-H. Guo, H.-Y. Wang, and X.-Q. Zheng. 2016. Transcriptome altered by latent human cytomegalovirus infection on THP-1 cells using RNA-seq. *Gene.* 594:144–150. doi:10.1016/J.GENE.2016.09.014.
- Zhang, Y., L. Yang, and L.-L. Chen. 2014. Life without A tail: New formats of long noncoding RNAs. *Int. J. Biochem. Cell Biol.* 54:338–349. doi:10.1016/j.biocel.2013.10.009.
- Zhao, B., J. Zou, H. Wang, E. Johannsen, C. 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. *Proc. Natl. Acad. Sci. U. S. A.* 108:14902–7. doi:10.1073/pnas.1108892108.
- Zhao, Y., H. Sun, and H. Wang. 2016. Long noncoding RNAs in DNA methylation: new players stepping into the old game. *Cell Biosci.* 6:45. doi:10.1186/s13578-016-0109-3.
- Zhou, J., C.M. Chau, Z. Deng, R. Shiekhata, M.-P. Spindler, A. Schepers, and P.M. Lieberman. 2005. Cell cycle regulation of chromatin at an origin of DNA replication. *EMBO J.* 24:1406–17. doi:10.1038/sj.emboj.7600609.
- Zou, Y.-R., A.H. Kottmann, M. Kuroda, I. Taniuchi, and D.R. Littman. 1998. Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature.* 393:595–599. doi:10.1038/31269.

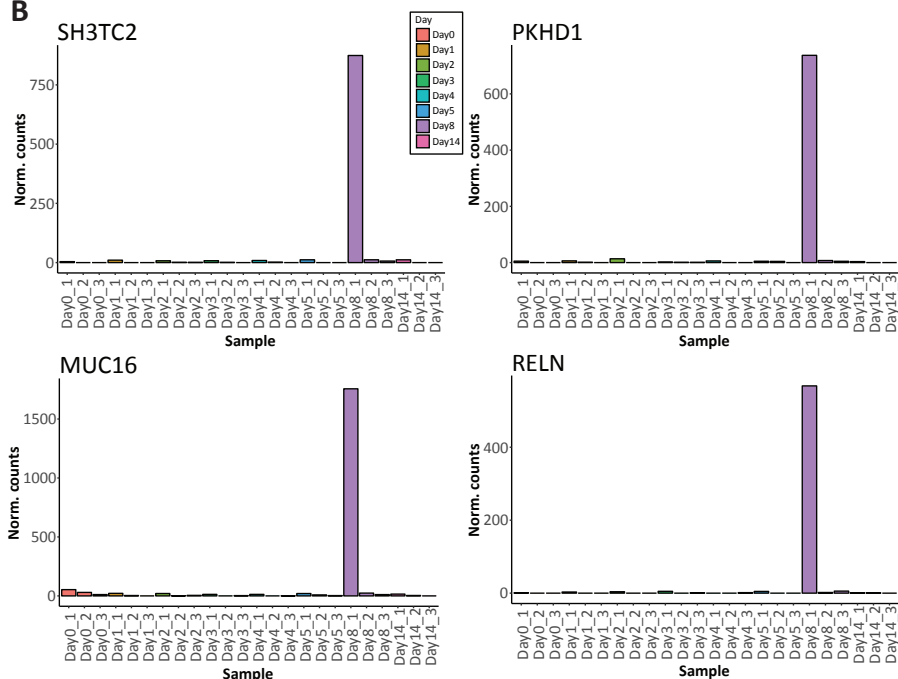
## 9. Appendix

### 9.1. Outlier sample from RNA-seq data

A



B

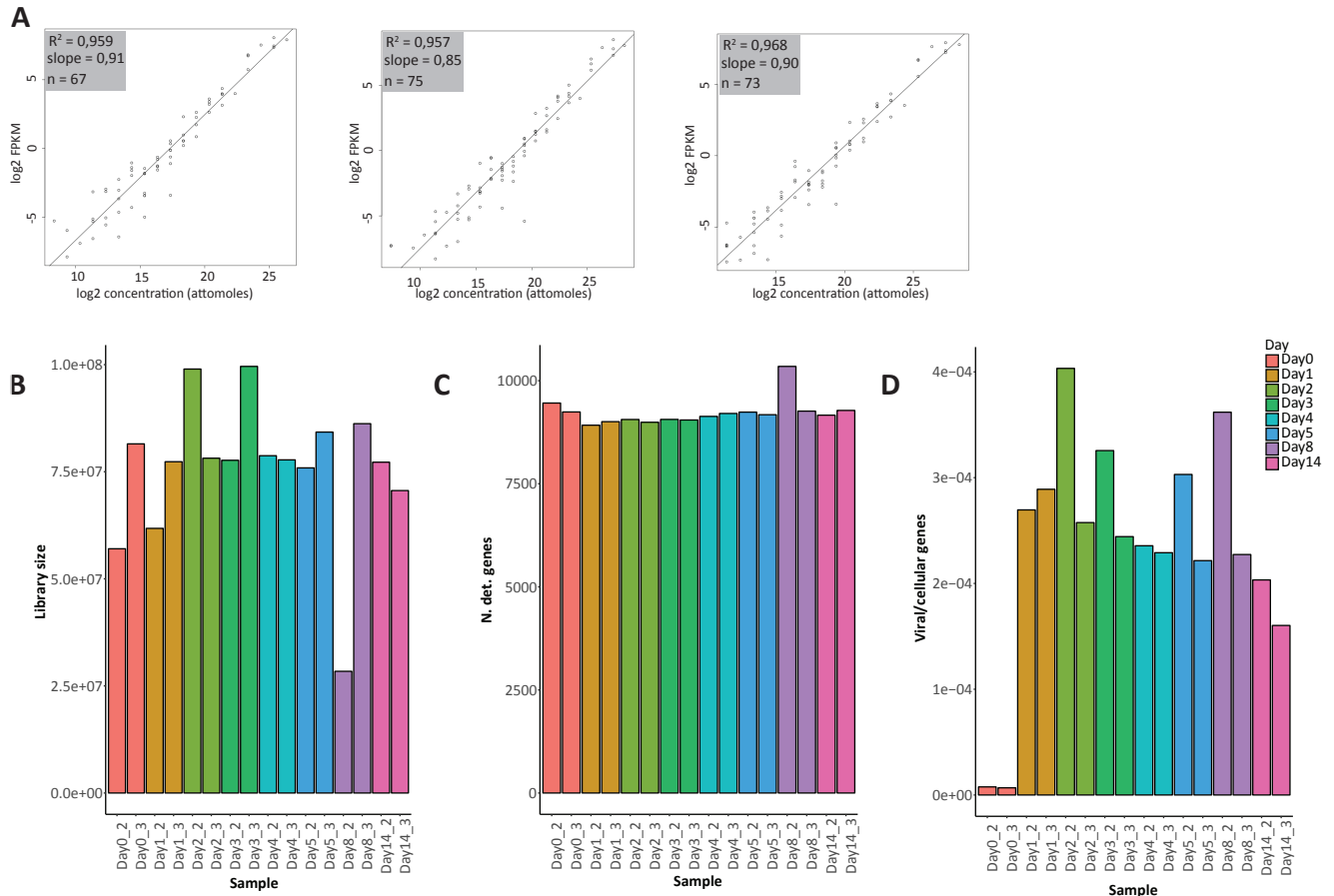


Supplementary figure 1. **RNA-seq data analysis and identification of an outlier sample (Day8\_1) in the polyA enriched libraries.**

- Principal component analysis (PCA) of all RNA-seq samples. The percentage of variance explained by principal component 1 and 2 (PC1, PC2) are indicated. Sample at day 8 (Day8\_1) appears to be an outlier and determines all variances between sequencing samples indicated by PC1. Further RNA-seq analysis was performed without Day8\_1 sample to minimize bias in the final results.
- Expression levels of the four genes with the most negative loadings on PC1 of the PCA plot (panel A). The depicted genes show very high expression levels only in the Day8\_1 sample and were found to be involved

in different, biologically unrelated processes from various types of the cells such as MUC16 in epithelia cells or RELN and SH3TC2 in neuronal cells. These results confirm that sample Day8\_1 had to be excluded from further RNA-seq analysis.

## 9.2. Quality controls of rRNA depleted sequencing libraries



Supplementary figure 2. **Quality controls of rRNA depleted, hexamer enriched RNA-seq libraries.**

The statistics for quality controls of the sequencing libraries were performed as described in Fig. 11.

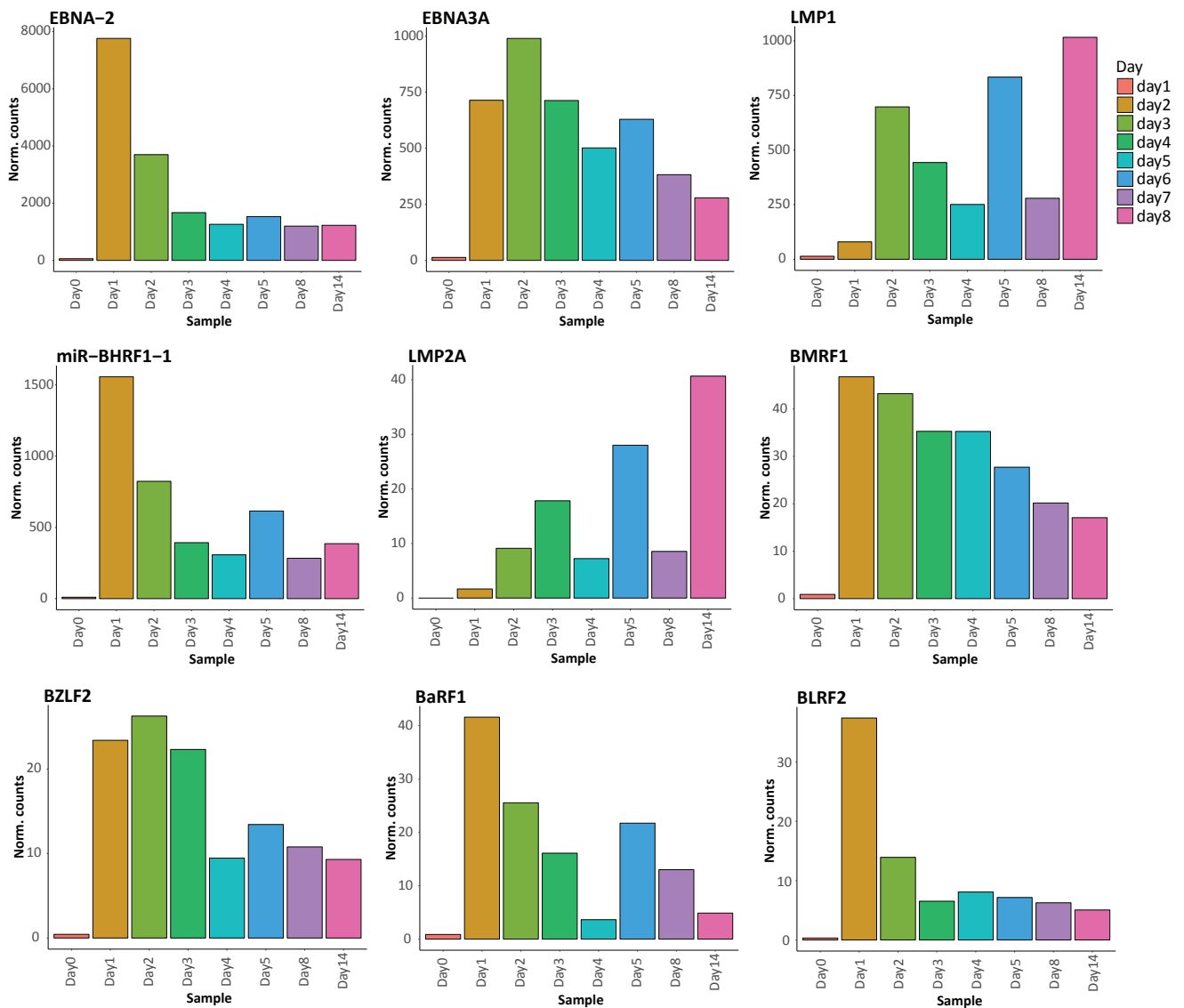
- Detection of the ERCC spike-in RNAs in rRNA depleted sequencing libraries. The depicted three graphs are representative examples of regression curves. The x-axis indicates the  $\log_2$  concentrations of the spike-in RNAs added to the samples, the y-axis depicts the  $\log_2$  of Fragments Per Kilobase of transcript per Million mapped reads (FPKM) per individual ERCC spike-in RNA. The square coefficient of determination ( $R^2$ ), the slope of the regression line, and the numbers of spike-in RNAs ( $n$ ) identified out of 92 added to the samples are indicated.
- The graph shows the sizes of the sequenced rRNA depleted libraries prepared from the different samples from two independent donors. The majority of the samples shows comparable sequencing depth, but sample Day8\_2 obtained fewer reads. This sample was excluded from further bioinformatics analysis because of its possible bias in the final results.
- The bar plot illustrates the numbers of detected genes in each sequenced sample with reads per million (RPM) >10. The numbers of detected genes are very comparable between samples.
- The graph displays the ratio of the numbers of reads from viral genes versus the numbers of reads from cellular genes at each day p.i. Viral genes are absent in the uninfected cells (Day 0) as expected. Initially after infection, the cells display higher viral genes ratios, especially at day one and two, which decrease gradually at the later time points.

### 9.3. Genes contributing to PC1 and PC2 in the polyA-enriched sequencing libraries

Supplementary table 1. The list of the top 30 genes with the highest loadings of the principal component (PC1 or 2).

Genes contributing to PC1		Genes contributing to PC2	
Gene	PC1 loading score	Gene	PC2 loading score
LOC643401	0.03461165	ZBTB16	-0.04016592
TIMD4	0.03461143	VAV3-AS1	-0.03945671
HRASLS2	0.03205240	CXCR4	-0.03768658
SPARC	0.03136722	HTR3A	-0.03694487
BC005081	0.03116985	SOX5	-0.03678185
CCL25	0.03008670	COBLL1	-0.03532114
GPR15	0.02972569	FOSB	-0.03410085
TMEM173	0.02961373	SDK2	-0.03405426
PLA2G4A	0.02949812	VAV3	-0.03327525
LAMP5	0.02942003	KLF4	-0.03300732
BCAR1	0.02920159	CLIC3	-0.03177428
LIPH	0.02884333	NR4A2	-0.03131990
COL16A1	0.02874806	FOS	-0.03125702
QPRT	0.02860624	SNORA70F	-0.03040951
BCAN	0.02834911	SERPINA9	-0.02900152
CCDC74A	0.02833998	GPRIN3	-0.02899787
CD276	0.02831831	GRASP	-0.02892397
PYGL	0.02816162	KIAA1683	-0.02891723
APOBEC3B	0.02811052	FAM49A	-0.02863372
IGF1	0.02791771	CXCR5	-0.02861571
IFNA2	0.02760686	GYLTL1B	-0.02846314
PLEKHG6	0.02737123	JUN	-0.02833765
IPCEF1	0.02702188	FCRL1	-0.02792099
LINC00029	0.02699818	DUSP1	-0.02791355
PIR	0.02681159	AK026379	-0.02757292
SAMD7	0.02670240	ADAM28	-0.02743359
SPC24	0.02661545	RGS2	-0.02734632
MIR148A	0.02657479	LINC00304	-0.02724341
EMID1	0.02656506	FCRL2	-0.02721448
CMBL	0.02626588	EHD3	-0.02718472

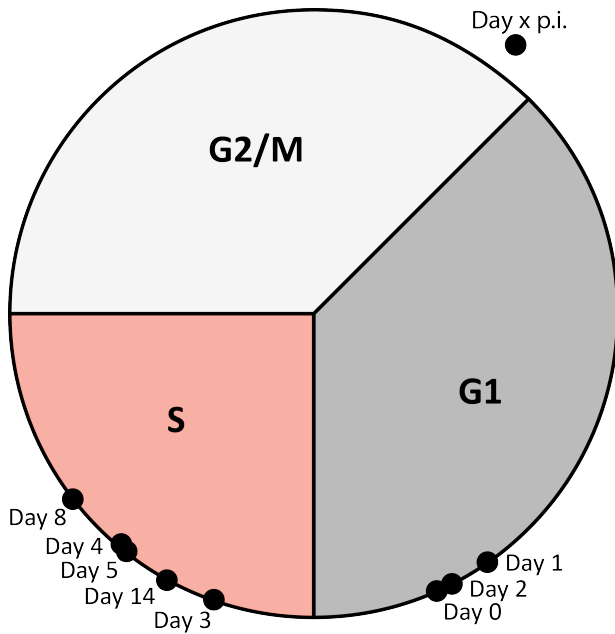
## 9.4. Time-course expression of EBV genes during the pre-latent phase of infection



### Supplementary figure 3. Expression of selected viral genes in the course of EBV infection.

The plots display the expression of nine selected EBV genes at different days p.i. The x-axis depicts the days p.i. and the y-axis shows the numbers of normalized counts (mean from three independent donors). The first five plots display known EBV latent genes with relatively high expression levels in particular EBNA-2. The latent genes were either highly expressed at the very early time points and then their expression decreased such as EBNA2, EBNA3A, miR-BHRF1-1, or the expression of viral genes gradually increased over time such as LMP1 and LMP2A. BMRF1, BZLF2, BaRF1 and BLRF2 are described as lytic genes and displayed overall lower expression levels compared to the latent genes. EBV lytic genes showed higher expression levels very early after infection, which were repressed at the later time points after infection.

### 9.5. Cell cycle phase prediction from RNA-seq data



Supplementary figure 4. **Assignment of cell cycle phases within samples from different days after EBV infection according to the RNA-seq data.**

The pie chart illustrates the average scores and predictions across the three replicates for each day p.i.. Samples resting in the G1 phase (dark grey) were detected in the uninfected cells and cells infected for one or two days. Samples collected from day three, four, five, eight and 14 were predicted to be in the S phase (light red) of the cell cycle as actively replicating cells. The assignment is based on the analysis using algorithm described in chapter 6.4.4.

---

## 10. Curriculum vitae

---

### Personal information:

Name: **Paulina Mrozek-Górska**  
Nationality: **polish**  
Email: **mrozek.paulina@gmail.com**

---

### Education:

Since 02/2014 **Doctoral studies on the topic: “Time-course analysis of human primary B cells infected with EBV”**  
**Research Unit Gene Vectors, Helmholtz Center Munich, Germany**  
**Prof. Dr. Wolfgang Hammerschmidt**  
**Prof. Dr. Dirk Eick**

---

## 11. Acknowledgment

I would like to thank all people, who helped and supported me on the way to complete my PhD thesis!

**Wolfgang:** Thank you for your enormous patience and great mentoring! I really appreciate your priceless help in correcting my thesis and solving THE problems, which occur on the way. Thank you for an opportunity to discuss basically everything, what bothered me during my PhD thesis, for your always “open doors” and for the valuable advice.

**Dirk:** Thank you very much for your constant support during my TAC meetings, your nice suggestions and advice for writing my thesis.

**Sandra Hake:** Thank you for taking part in my TAC meeting, for your valuable suggestions as well as nice atmosphere during these meetings☺

Big thank you to all smart and skilled people, who helped me a lot with bioinformatic analyses!

**Thomas Schwarzmayr:** Thank you very much for your help and dedication to my “exotic” RNA-seq project. I really appreciate your patience in answering all my emails and for trying to go even deeper into the analysis of my data.

**Dr. Antonio Scialdone:** I would like to express my gratitude for this wonderful opportunity to discuss with you and analyse my RNA-seq data. Thank you for sharing with me your great bioinformatic knowledge, your open mind, and understanding my important, but difficult biological tasks! Thank you also for your advice and suggestions to my thesis!

**Stefan Krebs, LAFUGA Team:** Thank you for all help during preparation of my MNase-seq samples, for giving new and nice ideas for solving some technical problems. Thank you also for an opportunity to work in your lab, while preparing the libraries and having access to all necessary lab equipment.



My beloved group of “*Hammerschmidts*” I would like to thank all of you together and everyone one by one for great and inspiring working atmosphere, for being something more than just colleagues from work! **Dagi** - thank you for amazing help with some experiments e.g. sorting or cell culture problems ☺, **Christine** - thank you for motivating me and for help with long and laborious sonication ☺, **Mika** - thank you for sharing a bench with me, it is always a pleasure, whatever I say ☺ and also for hugs and sweets, when I am mad ☺, **Adam** and **Manuel** - thank you for nice chatting time, but also for very important help with some experiments and scientific problems ☺, **Marisa** - thank you for introducing me to the lab, for your patience in the beginning and nice cooperation later!!!

And **Alex**: I told you I will make a special 30 pages to thank you... I really should ☺. Thank you for your enormous help with data analysis, advice, patience in solving my small or big R problems. Thank you for always being there with an open mind and will to cooperate! I really appreciate our formal and informal meeting times, thank you for trying to explain me some German “things” and for extreme fun from some nerdy jokes! I know that I can always rely on you in science, and outside it ☺. Just thank you, Alex!

Special thanks go to people from the 2<sup>nd</sup> floor (AG Zeidler, AG Moosmann, AG Mautner and AG Strobl) for nice conversations during lunch time, but also for kindly sharing with me with your experience and skills ☺

Dear **Kathrin**, thank you for your exceptional support during my PhD with sweets, coffee, and your always positive attitude! Thank you for great time that we always spend together!

Thanks to my beloved family and husband for your everlasting support!

*Kochani rodzice, kochana siostró, kochana rodzinko Górskich: dziękuję, że wspieraliście, wspieracie i jestem pewna, że będziecie mnie wspierać w moich marzeniach i dążeniach naukowych, ale nie tylko. Dziękuję Wam, za każdą duchową formę wsparcia i wiarę we mnie!*

*Kochany mężu, dziękuję Ci za Twoją nieocenioną cierpliwość i wsparcie w ważnych i trudnych dla mnie momentach. Z Tobą wszystko staje się możliwe!*

*Dziękuję wszystkim, którzy wspierali mnie dobrym słowem, żartem i jedzeniem:)*