PLOS Pathogens

EBF1 binds to EBNA2 and promotes the assembly of EBNA2 chromatin complexes in B cells

--Manuscript Draft--

Manuscript Number:	
Full Title:	EBF1 binds to EBNA2 and promotes the assembly of EBNA2 chromatin complexes in B cells
Short Title:	EBF1, a cofactor of EBNA2
Article Type:	Research Article
Section/Category:	Virology
Keywords:	EBNA2 target genes CBF1 RBPJ CSL EBF1 Notch B cell chromatin transcription factor
Corresponding Author:	Bettina Kempkes Helmholtz Zentrum München, German Research Center for Environmental Helth Munich, GERMANY
Corresponding Author's Institution:	Helmholtz Zentrum München, German Research Center for Environmental Helth
First Author:	Laura V Glaser
Order of Authors:	Laura V Glaser
	Simone Rieger
	Sybille Thumann
	Cornelia Kuklik-Roos
	Dietmar E Martin
	Kerstin C. Maier
	Marie L. Harth-Hertle
	Rolf Backofen
	Björn Grüning
	Florian Zimmer
	Bettina Kempkes
Abstract:	Epstein-Barr virus (EBV) infection converts resting human B cells into permanently growing lymphoblastoid cell lines (LCLs). The viral Epstein-Barr virus nuclear antigen 2 (EBNA2) plays a key role in this process. It preferentially binds to B cell enhancers and establishes a specific viral and cellular gene expression program in LCLs. The cellular DNA binding factor CBF1/CSL serves as a sequence specific chromatin anchor for EBNA2. The ubiquitous expression of this highly conserved protein raises the question whether additional cellular factors might determine EBNA2 chromatin binding selectively in B cells. Here we used CBF1 deficient B cells to identify cellular genes up or downregulated by EBNA2 as well as CBF1 independent EBNA2 chromatin binding sites. Apparently, CBF1 independent EBNA2 target genes and chromatin binding sites can be identified but both are less frequent than CBF1 dependent EBNA2 functions. CBF1 independent EBNA2 binding sites are highly enriched for EBF1 binding motifs. We show that EBNA2 binds to EBF1 in CBF1 proficient and deficient B cells and requires EBF1 to bind to CBF1 independent binding sites. Our results identify EBF1 as

Additional Information:	
Question	**Response**
Financial Disclosure Please describe all sources of funding that have supported your work. **This information is required for submission and will be published with your article, should it be accepted.** A complete funding statement should do the following: Include **grant numbers and the URLs** of any funder's website. Use the full name, not acronyms, of funding institutions, and	Deutsches Zentrum für Infektionsforschung (DZIF) -Vorhaben TI 07.003 Deutsche Krebshilfe (Project: 109258) The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript

use initials to identify authors who received the funding. **Describe the role** of any sponsors or funders in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. If the funders had **no role** in any of the above, include this sentence at the end of your statement: "*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript*." However, if the study was **unfunded**, please provide a statement that clearly indicates this, for example: "*The author(s) received no specific funding for this work*." * typeset	
Competing Interests You are responsible for recognizing and disclosing on behalf of all authors any competing interest that could be perceived to bias their work, acknowledging all financial support and any other relevant financial or non-financial competing interests. Do any authors of this manuscript have competing interests (as described in the PLOS Policy on Declaration and Evaluation of Competing Interests)? **If yes**, please provide details about any and all competing interests in the box below. Your response should begin with this statement: *I have read the journal's policy and the authors of this manuscript have the following competing interests:* **If no** authors have any competing interests to declare, please enter this statement in the box: "*The authors have declared that no competing interests exist*." * typeset	The authors have declared that no competing interests exist
Data Availability PLOS journals require authors to make all	Yes - all data are fully available without restriction

data underlying the findings described in their manuscript fully available, without restriction and from the time of publication, with only rare exceptions to address legal and ethical concerns (see the PLOS Data Policy and FAQ for further details). When submitting a manuscript, authors must provide a Data Availability Statement that describes where the data underlying their manuscript can be found. Your answers to the following constitute your statement about data availability and will be included with the article in the event of publication. **Please note that simply stating 'data available on request from the author' is not acceptable.** ***If, however, your data are only available upon request from the author(s), you must answer "No" to the first question below, and explain your exceptional situation in the text box provided.*** Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?	
Please describe where your data may be found, writing in full sentences. **Your answers should be entered into the box below and will be published in the form you provide them, if your manuscript is accepted.** If you are copying our sample text below, please ensure you replace any instances of XXX with the appropriate details. If your data are all contained within the paper and/or Supporting Information files, please state this in your answer below. For example, "All relevant data are within the paper and its Supporting Information files." If your data are held or will be held in a public repository, include URLs, accession numbers or DOIs. For example, "All XXX files are available from the XXX database (accession number(s) XXX, XXX)." If this information will only be available after acceptance, please indicate this by ticking the box below. If neither of these applies but you are able to provide details of access elsewhere, with or without limitations, please do so in the box below. For example: "Data are available from the XXX Institutional Data Access / Ethics Committee for researchers who meet the criteria for access to confidential data."	Affymetrix gene array hybridization results have been deposited at GEO Gene Expression Omnibus (GSE96762) https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=ohmrsmywlfkrzul&acc=GSE96762 ChIP Seq will be available from Sequence Read Archive (SRA Identifier SRP101966). The Galaxy workflow used to analyze ChIP-Seq data is avaiable at: https://github.com/bgruening/galaxytools/tree/master/workflows/peak_calling

“Data are from the XXX study whose authors may be contacted at XXX.” * typeset	
Additional data availability information:	Tick here if the URLs/accession numbers/DOIs will be available only after acceptance of the manuscript for publication so that we can ensure their inclusion before publication.

Dear Editors,

please find enclosed our manuscript "EBF1 binds to EBNA2 and promotes the assembly of EBNA2 chromatin complexes in B cells".

CBF1/CSL (C promoter binding factor, Suppressor of Hairless, and Lag1 also called RBPJ or RBPJκ) is a cellular DNA binding protein, ubiquitously expressed in all mammalian tissues. CBF1 serves as a DNA sequence specific adaptor molecule that anchors either repressors or activators to transcriptional control elements. EBNA2 is one of the first proteins expressed in Epstein-Barr virus infected B cells and plays a key role during the immortalization process which generates permanently growing B cell cultures (LCLs). Since EBNA2 cannot bind to DNA directly, it uses cellular DNA binding factors as anchors to bind to regulatory regions in the genome. CBF1 is a well characterized anchor protein of EBNA2. Surprisingly, despite the ubiquitous expression of CBF1, EBNA2 is found preferentially at B cell super enhancers which are occupied with multiple B cell specific transcription factors.

In brief, our studies suggest, that EBF1 serves as a co-factor that assists EBNA2 chromatin binding for a subpopulation of EBNA2 binding sites in B cells. Our manuscript supports this finding by combining several separate lines of evidence based on genetic, biochemical and bioinformatics methodologies.

- We used genetically engineered CBF1 deficient B cell lines to identify EBNA2 target genes and EBNA2 binding sites to show that EBNA2 can exert some of its functions in the absence of CBF1.
- EBNA2 chromatin binding sites fall into two subclasses: CBF1 dependent and independent binding sites. CBF1 independent EBNA2 binding sites also strongly bind EBNA2 in LCLs.
- CBF1 dependent and independent EBNA2 binding sites carry similar chromatin signatures, suggesting that the chromatin activation state is not the distinguishing feature of the two subclasses.
- In silico de novo DNA motif discovery and prediction of transcription factor binding sites for the two subclasses was performed. The results suggested that the DNA sequence motifs at CBF1 independent sites are unlikely to bind CBF1. Instead the results predict significant binding probability for EBF1. Thus, CBF1 independent binding sites carry low affinity binding motifs indicating that additional factors contribute to CBF1 binding that can be detected in chromatin immunoprecipitation studies.
- Quantitative correlation studies of signal intensities for transcription factor binding in LCLs were performed to identify candidate proteins which are co-enriched with EBNA2. These quantitative analyses were based on publicly available information resources. Cluster analysis based on correlation coefficients revealed a strong co-binding of EBNA2 and CBF1 as expected, but also a strong co-binding of EBNA2 and EBF1.
- Importantly, our co-immunoprecipitation studies demonstrated that EBF1 and EBNA2 form complexes in CBF1 proficient and deficient B cells and thus identified EBNA2 as the first viral protein bound to EBF1.
- In order to test whether EBNA2 requires EBF1 to bind to chromatin, EBF1 expression was silenced by EBF1 specific siRNAs. Chromatin immunoprecipitations studies of CBF1 independent and dependent EBNA2 binding sites revealed that EBF1 is critical for EBNA2

complex formation specifically at CBF1 independent binding sites. Since we can show that EBNA2 and EBF1 physically interact in cells, our studies identify EBF1 as novel EBNA2 chromatin anchor.

EBNA2 expression is a hallmark of B cell lymphomas arising in immunocompromised patients and considered to drive the proliferation of these cells. Until today, there is no established therapeutic strategy to target latent EBV infection. We are convinced that our studies are of general importance since a detailed understanding of the mechanics that underlie EBNA2 functions in the host B cell are required to establish novel targeted therapies.

We hope that our manuscript will be accepted for publication by PLOS Pathogens.

Yours sincerely,

Bettina Kempkes.

EBF1 binds to EBNA2 and promotes the assembly of EBNA2 chromatin complexes in B cells

Laura V. Glaser[1*], Simone Rieger[1*], Sybille Thumann[1], Cornelia Kuklik-Roos[1], Dietmar E. Martin[2], Kerstin C. Maier[2] Marie L. Harth-Hertle[1], Björn Grüning[3], Rolf Backofen[3], Stefan Krebs[2], Helmut Blum[2], Ralf Zimmer[4], Florian Erhard[4], Bettina Kempkes[1**]

Department of Gene Vectors, Helmholtz Center Munich, Munich, Germany[1], Gene Center, Ludwig-Maximilians-University, Munich, Germany[2], Bioinformatics, Institute for Informatics, Albert-Ludwigs-University, Freiburg, Germany[3], Teaching and Research Unit Bioinformatics, Institute of Informatics, Ludwig-Maximilians-University Munich, Germany[4]

* These authors contributed equally to this work.
** Corresponding author

Short title: EBF1, a cofactor of EBNA2

Current address LG: Max Planck Institute for Molecular Genetics, Department of Computational Molecular Biology, Berlin, Germany
Current address FE: Institute for Virology and Immunobiology, Julius-Maximilians University, Würzburg, Germany
Current address KCM: Max Planck Institute for Biophysical Chemistry Göttingen, Department of Molecular Biology, Germany

Abstract

Epstein-Barr virus (EBV) infection converts resting human B cells into permanently growing lymphoblastoid cell lines (LCLs). The viral Epstein-Barr virus nuclear antigen 2 (EBNA2) plays a key role in this process. It preferentially binds to B cell enhancers and establishes a specific viral and cellular gene expression program in LCLs. The cellular DNA binding factor CBF1/CSL serves as a sequence specific chromatin anchor for EBNA2. The ubiquitous expression of this highly conserved protein raises the question whether additional cellular factors might determine EBNA2 chromatin binding selectively in B cells. Here we used CBF1 deficient B cells to identify cellular genes up or downregulated by EBNA2 as well as CBF1 independent EBNA2 chromatin binding sites. Apparently, CBF1 independent EBNA2 target genes and chromatin binding sites can be identified but both are less frequent than CBF1 dependent EBNA2 functions. CBF1 independent EBNA2 binding sites are highly enriched for EBF1 binding motifs. We show that EBNA2 binds to EBF1 in CBF1 proficient and deficient B cells and requires EBF1 to bind to CBF1 independent binding sites. Our results identify EBF1 as a co-factor of EBNA2 which conveys B cell specificity to EBNA2.

Introduction

CBF1/CSL (C promoter binding factor, Suppressor of Hairless, and lag1 also called RBPJ or RBPJκ) is a cellular DNA binding protein, ubiquitously expressed in all mammalian tissues. CBF1 serves as a DNA adaptor molecule that recruits either repressors or activators to transcriptional control elements like enhancers and transcription start sites of genes and is described as the major downstream effector of the cellular Notch signal transduction pathway (1). Notch signaling controls the development and differentiation of diverse organs and tissues. Despite the ubiquitous expression of its chromatin anchor CBF1, target gene control by Notch is context dependent and requires tissue and lineage specific cooperating transcription factors (2). In B cells, latently infected with Epstein-Barr virus (EBV), CBF1 anchors the viral transactivator protein EBV nuclear antigen 2 (EBNA2) to chromatin and thereby initiates a cascade of signaling events that coordinate B cell activation and proliferation of infected cells (3-6). Thus, EBNA2 is considered to mimic Notch signaling (7). In contrast to the universal expression and pleiotropic activities of Notch, the expression and the biological activity of EBNA2 is strictly confined to EBV infected B cells, characterized by a transcription program that phenocopies antigen activated B cell blasts (8, 9).

CBF1 and EBNA2 frequently co-occupy cellular enhancer and super-enhancer regions reinforcing the concept that CBF1 is the major adaptor for EBNA2 to chromatin (10). In addition, EBNA2 bound regions are co-occupied with multiple additional transcription factors including IRF4, BATF, NFκB, Runx, and ETS family members as well as the B cell lineage defining and pioneer factors PU.1/SPI1 and EBF1 (10). While the adaptor function of CBF1 is well defined, a potential functional contribution of these co-occurring factors to EBNA2 function has not been studied thoroughly. These proteins are active transcription factors which carry transactivation domains and can actively promote or impair transcription of target genes. PU.1/SPI1 promotes B cell development and is expressed throughout B cell differentiation, but also controls T cell, myeloid and dendritic cell differentiation (11). PU.1/SPI1 DNA binding sites are critical for LMP1 promoter luciferase activation (3, 12-14). However, its contribution to LMP1 expression in the context of the entire viral genome is surprisingly weak (15). Most recently it has

been shown, that EBNA2 enhances the binding of CBF1 and EBF1 to chromatin and EBF1 is critical for expression of the EBNA2 viral target gene LMP1 (15, 16). Importantly, within the hematopoietic compartment EBF1 is exclusively expressed in B cells and their lymphocytic precursors. The other EBF gene family members EBF2, 3, and 4 are expressed at very low or undetectable levels in B cells. EBF1 initiates B cell lineage commitment, development and differentiation as a pioneer factor that promotes chromatin accessibility and DNA demethylation in lymphocyte precursors (17, 18).

Strong EBNA2 binding correlates with extended regions of extraordinarily high histone 3 lysine 27 acetylation (H3K27ac) and H3K4 mono-methylation (H3K4me1) marks which are characteristic features of activated super enhancers (19). In addition, EBNA2 modulates the formation of chromatin loops to connect enhancer and promoters of its target genes (20). In theory, EBNA2 co-occurring factors, like PU.1/SPI1 and EBF1 could function as pioneer factors for EBNA2 by modulating the chromatin state and thereby promoting access of EBNA2 to chromatin, indirectly. Alternatively, EBNA2 co-occurring factors might serve as alternate adaptors that promote DNA binding of EBNA2.

CBF1 is ubiquitously expressed in all mammalian cells including primary human B cells and EBV infected and non-infected human B cell lines. For this study we used a CBF1 deficient human B cell line, which had been generated by homologous recombination in the somatic B cell line DG75, to screen for CBF1 independent functions of EBNA2. The parental DG75 B cell line is an EBV negative Burkitt's lymphoma cell line that, in contrast to EBV immortalized B cells, tolerates inactivation of the CBF1 gene without loss of viability (21, 22). We compared EBNA2 induced cellular genes in CBF1 proficient and deficient DG75 cells and found the majority of EBNA2 target genes to be CBF1 dependent. A minor fraction of EBNA2 target genes is regulated CBF1 independently. By chromatin immunoprecipitation and genome wide sequencing of EBNA2 bound DNA fragments (ChIP-Seq) we identified a subpopulation of CBF1 independent EBNA2 binding sites that was significantly enriched for EBF1 binding motifs. We show that CBF1 independent EBNA2 binding to chromatin is dependent on EBF1 protein expression. Importantly, we demonstrate that EBNA2 and EBF1 can form protein

complexes in CBF1 positive and negative cells, indicating that EBF1 serves as B cell specific DNA anchor for EBNA2.

Results

Genome wide expression profiling identifies cellular transcripts regulated by EBNA2 in CBF1 deficient B cells.

In order to rigorously test, if EBNA2 can exert any functions in the absence of its DNA adaptor CBF1, a microarray based genome wide screen for EBNA2 target genes in DG75 B cells that are either proficient (wt) or deficient (ko) for CBF1 was performed. Both cell lines constitutively express an estrogen receptor (ER) hormone binding domain EBNA2 fusion protein (ER/EBNA2). ER/EBNA2 is retained in the cytoplasm of the cell but is rapidly activated and translocated to the nucleus in response to estrogen (21, 22). For expression profiling, DG75$^{ER/EBNA2}$ CBF1 wt and CBF1 ko cells were cultured in estrogen supplemented media for 24 h, total cellular RNAs were harvested and processed for the hybridization of gene arrays that detect 30645 coding transcripts, 11086 lincRNAs (long intergenic non-coding RNA transcripts) and 148 miRNAs (micro RNAs; Fig. 1A, B and F). Cell cultures of the parental DG75 CBF1 wt and CBF1 ko cell lines, which do not express ER/EBNA2, were treated with estrogen and processed for the microarray analysis as specificity controls. Neither in DG75 CBF1 wt nor in DG75 CBF1 ko cells statistically significant changes ($p \leq 0.05$) of cellular transcript abundance in response to estrogen treatment were observed, proving that target gene activation is strictly dependent on ER/EBNA2 (S1 Fig. A). In addition, estrogen responsive target genes described in the literature did not change expression levels proving that the estrogen receptor response is not functional in DG75 B cells (S1 Fig. B) (23-27).

Multiple previously characterized EBNA2 target genes were significantly upregulated in DG75$^{ER/EBNA2}$ CBF1 wt cells (S1 Fig. C). In total, 99 cellular transcripts were up- and 37 cellular transcripts were downregulated ≥ 4-fold ($p \leq 0.001$) (Fig. 1A). Importantly, 15 transcripts were upregulated and 6 transcripts were downregulated in CBF1 deficient DG75$^{ER/EBNA2}$ ≥ 4-fold (Fig. 1B). Thus, as expected, the number of differentially expressed EBNA2 target genes was markedly higher in CBF1 proficient cells.

Unexpectedly, a robust response to EBNA2 was also seen in CBF1 deficient cells (Fig.1B and C and S2 Fig.). The dynamic range of gene regulation is illustrated for genes regulated ≥ 2-fold ($p \leq 0.05$) for both cell lines (Fig. 1C). Many CBF1 independent target genes are also regulated in CBF1 proficient cells (Fig. 1D and E), while a small group of targets is regulated by EBNA2 in CBF1 deficient cells, only. In order to verify the microarray results, a panel of 12 CBF1 dependent and independent targets was selected for re-testing. RT-qPCR experiments confirmed that most CBF1 independent targets also responded to EBNA2 in CBF1 proficient cells. As already seen in the microarray experiment, the degree to which individual targets responded in CBF1 proficient cells varied considerably, but was faithfully reproduced by RT-qPCR (S3 Fig.). Interestingly, the CBF1 dependent target genes included a substantial number of miRNAs that are up- or downregulated by EBNA2 (S4 Fig.).

Figure 1. Comparative transcript profiling of EBNA2 target gene expression in CBF1 proficient and deficient DG75 cells. DG75 cells expressing ER/EBNA2 were cultivated in estrogen supplemented medium for 24 h or were left untreated. Total cellular RNA was isolated and submitted to gene expression analysis using the Human Gene 2.0 ST array. All probe sets represent single transcripts trxs). For each condition, 3 biological replicates were examined. Each vertical column represents the results obtained after hybridizing a single microarray. Horizontal rows represent data obtained for a particular probe set across all cell lines and conditions adjusted to a scale ranging from -2.0 to + 2.0. The relative high, medium and low expression values are represented by red, white and blue color, respectively. Vertical columns are ranked according to fold changes from highest induction on top to highest repression levels at the bottom. (A) Expression levels of 136 transcripts which change expression levels at least 4-fold ($p \leq 0.001$) in response to EBNA2 in CBF1 proficient DG75 ($DG75^{ER/EBNA2}$ CBF1 wt) cells are displayed. The transcript cluster ID and the assigned genes/transcripts and long non-coding RNAs are annotated. (B) 21 transcripts regulated at least 4-fold ($p \leq 0.001$) in CBF1 deficient DG75 ($DG75^{ER/EBNA2}$ CBF1 ko). (C) Boxplots depicting the fold change distribution of EBNA2 induced and repressed transcripts at least 2-fold ($p \leq 0.05$) in CBF1 wt and ko cells, respectively. EBNA2 induced (D) and repressed (D) transcripts are shown to illustrate the dynamic range of each system. Boxplot whiskers extend to 1.5x interquartile range. (F) Expression levels of EBNA2 (prior to and after estrogen treatment) and CBF1 proteins were monitored by western blot analysis. Equal amounts of total protein lysates were applied and GAPDH served as an internal loading control. One representative experiment (n=3) is shown.

CBF1 independent EBNA2 repressed target genes are enriched for genes involved in B cell signaling

To functionally characterize EBNA2 target genes, biological processes associated with individual subsets of genes were analyzed. The subsets considered here consisted of genes that were on average induced or repressed in the CBF1 proficient and deficient cell lines, or genes where induction or repression was dependent or independent of CBF1 (S5 Fig.). Only genes significantly ($q < 0.01$) regulated in at least one of the two cell lines were considered. Thresholds on fold changes were chosen by the online tool GOrilla in a data dependent manner to identify subsets enriched in GO terms in the "Biological Process" category (S5 Fig.).

Neither genes repressed in CBF1 proficient cells only (repressed/CBF1 dependent) nor induced in CBF1 deficient cells (induced/CBF1 independent) were significantly ($q \leq 10^{-4}$) enriched for any biological process. Genes induced in CBF1 proficient cells only (induced/CBF1 dependent) were strongly and most significantly enriched for immunoglobulin receptor binding and moderately enriched for biological processes involving several enzymatic activities (Table 1).

Table 1: Gene Ontology Enrichment Analysis of CBF1 dependent EBNA2 induced target genes

Term ID*	Term	Genes in term	Target genes in term**	Enrichment Score***	q-value
GO:0034987	immunoglobulin receptor binding	19	10	58,08	5,59E-12
GO:0004252	serine-type endopeptidase activity	30	8	29,43	1,13E-06
GO:0008236	serine-type peptidase activity	33	8	26,75	2,16E-06
GO:0017171	serine hydrolase activity	33	8	26,75	1,80E-06
GO:0003823	antigen binding	49	11	24,77	5,03E-09
GO:0070011	peptidase activity, acting on L-amino acid peptides	137	8	15,39	1,14E-04
GO:0008233	peptidase activity	142	8	14,85	1,35E-04
GO:0004175	endopeptidase activity	84	12	12,79	1,25E-06
GO:0004872	receptor activity	226	33	2,61	6,65E-04
GO:0060089	molecular transducer activity	226	33	2,61	5,98E-04

*****The top 10 GO terms in the "Biological Process" category are depicted. Note that a given gene can be annotated to multiple terms.
**number of genes in the top of the EBNA2 target gene list (chosen by GOrilla)
***Enrichment of a given GO term among differentially regulated genes with respect to the total number of genes assayed and annotated to them, calculated by GOrilla, see Material and Methods

Target genes repressed by EBNA2 in the absence of CBF1 (repressed/ CBF1 independent) showed a remarkable profile (Table 2). They map to several GO terms that cover diverse immune responses. Since the study had been performed in B cells, the enrichment for genes involved in immune responses and B cell receptor biology could have been expected. However, our study indicates that EBNA2 also represses immune response genes and this feature of EBNA2 is CBF1 independent.

Table 2: Gene Ontology Enrichment Analysis for CBF1 independent EBNA2 repressed target genes

Term ID*	Term	Genes in term	Target genes in term**	Enrichment Score***	q-value
GO:0002768	immune response-reg. cell surface receptor signaling pathway	46	27	3,52	1,09E-06
GO:0002757	immune response-activating signal transduction	52	30	3,46	3,87E-07
GO:0002764	immune response-reg. signaling pathway	54	31	3,44	3,08E-07
GO:0050778	positive regulation of immune response	68	36	3,18	2,96E-07
GO:0002429	immune response-act. cell surface receptor signaling pathway	44	31	2,87	1,29E-06
GO:0050776	regulation of immune response	84	39	2,79	1,25E-06
GO:0002253	activation of immune response	54	36	2,72	8,67E-07
GO:0002376	immune system process	149	56	2,26	1,16E-06
GO:0007166	cell surface receptor signaling pathway	152	73	1,9	1,90E-06
GO:0007165	signal transduction	311	128	1,59	6,01E-07

*****The top 10 GO terms in the "Biological Process" category are depicted. Note that a given gene can be annotated to multiple terms.
**number of genes in the top of the EBNA2 target gene list (chosen by GOrilla)
***Enrichment of a given GO term among differentially regulated genes with respect to the total number of genes assayed and annotated to them, calculated by GOrilla, see Material and Methods

EBNA2 is recruited to chromatin in CBF1 deficient B cells

In summary our differential expression analysis of EBNA2 target genes shows that EBNA2 can regulate a small fraction of its target genes without using CBF1 as DNA anchor. In order to identify alternate strategies of EBNA2 to bind to chromatin we performed chromatin immunoprecipitation (ChIP) studies to identify genomic loci that are bound by EBNA2 in CBF1 negative cells. In ER/EBNA2 expressing cells, EBNA2 shuttles from the cytoplasm to the nucleus in response to estrogen. In order to avoid a potential impact of cytoplasmic ER/EBNA2 contamination on our biochemical studies,

we switched to a doxycycline inducible HA-EBNA2 expression system (doxHA-E2) in DG75 (S6 Fig. A). In the absence of doxycycline EBNA2 is not expressed and cannot interfere with the immunoprecipitation procedure in DG75$^{doxHA\text{-}E2}$/CBF1 wt and DG75$^{doxHA\text{-}E2}$/CBF1 ko cells. Up to 90% of the cells express EBNA2 when treated with doxycycline (S6 Fig. B). EBNA2 protein signal detected by immunostaining was 5-10-fold stronger than EBNA2 in LCLs (data not shown). In comparison to LCLs, some of the EBNA2 co-occurring transcription factors like BATF and IRF4 were expressed at very low levels (data not shown) while EBF1 and PU.1/SPI1 were robustly expressed (S6 Fig. C). ChIP followed by high throughput sequencing (ChIP-seq) was performed to determine EBNA2 genome occupancy. 1,789 EBNA2 binding sites were identified in CBF1 proficient DG75$^{dox\text{-}HA2}$ (Fig. 2A), while 22,500 EBNA2 peaks were identified in LCLs, which had been performed in parallel. 1,325 of the EBNA2 peaks in DG75$^{doxHA\text{-}E2}$ cells were also present in LCLs (shared peaks), while 464 binding sites occurred exclusively in DG75$^{doxHA\text{-}E2}$ cells. EBNA2 signal intensities was most prominent at LCL/DG75$^{doxHA\text{-}E2}$ shared EBNA2 binding (Fig. 2B).

In LCLs, EBNA2 is preferentially recruited to enhancer elements which pre-exist in peripheral CD19 positive B cells before they are infected by EBV to generate LCLs (10). Chromatin marks characteristic for activated enhancer elements are H3K27ac in combination with H3K4me1 signals that are stronger than H3K4me3. We speculated that DG75 specific chromatin signatures in the absence of EBV infection might influence EBNA2 binding. We thus compared H3K4me1, H3K4me3, and H3K27ac signal intensities at EBNA2 binding sites i) shared by LCLs and DG75$^{doxHA\text{-}E2}$, ii) unique for LCLs and iii) unique for DG75 in naïve CD19 positive B cells with those in non-transfected DG75.

EBNA2 binding sites, shared by LCLs and DG75$^{doxHA\text{-}E2}$, stand out as the subset with the most prominent enrichment for all three investigated histone modifications associated with the chromatin state of active enhancers (Fig. 2C). In contrast, DG75$^{doxHA\text{-}E2}$ unique EBNA2 binding sites were highly enriched for active chromatin marks in the DG75 precursor only, while LCL unique EBNA2 peaks showed significantly lower signal intensities in DG75. These data indicate that a set of enhancers, which are pre-activated in DG75 cells, but not in the CD19 positive LCL precursors, might allow

the formation of "DG75 unique" EBNA2 binding sites. DG75 lack pre-formed enhancer signatures at "LCL unique" binding sites. In addition, the absence or low abundance of IRF4 and BATF proteins or other co-occurring transcription factors in DG75$^{doxHA\text{-}E2}$ could limit EBNA2 occupancy in DG75 at these LCL unique sites.

Figure 2. Cell line specific chromatin signatures predispose specific sites for EBNA2 binding.
(A) Intersection of EBNA2 binding sites identified in LCLs and DG75$^{doxHA\text{-}E2}$ CBF1 wt cells. (B) Anchor plots showing EBNA2 signal intensities for LCLs and for DG75$^{doxHA\text{-}E2}$ CBF1 wt at sites identified in both cell lines (LCL/DG75$^{doxHA\text{-}E2}$ shared) or unique to either cell line (LCL unique or DG75$^{doxHA\text{-}E2}$ unique). (C) Signals associated with active chromatin and enhancer state (H3K4me1, H3K4me3, and H3K27ac) at EBNA2 binding sites in CD19+ B cells and DG75 cell line. Using data provided by public resources (28, 29) the mean normalized signal for each histone modification and peak subset was calculated for the region flanking all EBNA2 peak centers for 20 kb in each direction, applying the same workflow for CD19+ B cells and DG75 data sets. Please note that signal intensities for the same histone modification should not be compared between the two cell lines since the experiments were conducted by different laboratories using different antibodies.

The comparison of ChIP-seq data between DG75$^{doxHA\text{-}E2}$ CBF1 wt and ko cells identified 1,789 EBNA2 binding sites in CBF1 proficient and 271 in CBF1 deficient DG75 with an almost complete overlap of 243 CBF1 independent EBNA2 peaks (Fig. 3A). A small group of 28 binding sites were only identified in CBF1 deficient cells and were not analyzed further. 1,546 EBNA2 sites were not detected in CBF1 deficient cells and thus defined as "CBF1 dependent". The mean EBNA2 signal intensity at EBNA2 binding sites was elevated 1.4-fold in wt compared to ko cells (Fig. 3B and C). Remarkably, CBF1 independent peaks, compared to dependent peaks, showed significantly enriched EBNA2 binding in both, CBF1 wt and ko, cell lines (2.5 and 3.8-fold, respectively, Fig. 3D and E). The quantitative re-analysis of the subclasses of EBNA2 peaks in LCLs confirmed that CBF1 independent peaks are characterized by stronger EBNA2 enrichment (Fig. 3D, E and F, right panel). Since CBF1 independent EBNA2 binding obviously contributes to EBNA2 occupancy in LCLs, we conclude that our CBF1

deficient B cell line is a valid model system to study mechanisms which drive EBNA2 chromatin interactions.

Figure 3. EBNA2 can access more than 15% of its chromatin binding sites in CBF1 deficient DG75 B cells.
(A) Intersection of EBNA2 binding sites identified in CBF1 proficient or deficient cells 24 h post doxycycline induction. 1,546 peaks that were identified in CBF1 proficient but not in CBF1 deficient cells were defined as "CBF1 dependent" EBNA2 peaks. 243 EBNA2 peaks identified in CBF1 deficient and proficient DG75 cells were defined as "CBF1 independent". (B-E) Comparison of EBNA2 signal distributions at CBF1 independent or dependent peaks. (B) Anchor and (C) scatter plots (mean + 95% CI) depicting signal distributions at EBNA2 peak subsets. Regions flanking the peak center for 2 kb in each direction were analyzed (Data underlying panel B). Absolute means and SEMs are indicated below. (D) Anchor and (E) scatter plots (mean + 95% CI) as shown in B and C but depicting EBNA2 signal intensities for the two different subsets of EBNA2 peaks as defined in A. Statistical significance for differences of all means were assessed applying unpaired two-tailed t-test for log values with Welch's correction (**** $p < 0.0001$); absolute means and SEMs are indicated below. (F) List of EBNA2 mean signal intensities at CBF1 independent and dependent peaks.

To better characterize CBF1 dependent and independent EBNA2 binding sites prior to EBNA2 binding we could use H3K4me1, H3K4me3, and H3K27ac ChIP-Seq data published for DG75 (29). Signal intensities of H3K4me1, H3K4me3, and H3K27ac were separately analyzed for the CBF1 dependent and independent peak subpopulations and compared to the mean signal peak intensities of the respective chromatin modification in DG75 (Fig. 4). All three activation marks showed almost the same high enrichment profiles for both subpopulations, indicating that chromatin signatures are most probably not the trigger either for CBF1 dependent or independent binding.

Figure 4. CBF1 independent and dependent EBNA2 binding sites are significantly enriched for activated chromatin marks in DG75 cells prior to EBNA2 binding.
Based on published data sets on histone modification in DG75, the two EBNA2 peak subsets (CBF1 independent dark blue; CBF1 dependent light blue) were separately analyzed for histone activation marks, typically found in enhancer regions. These data were compared to signal intensities of all peaks for the respective chromatin

modification (red). (A) Anchor plots depict H3K4me1, H3K4me3, and H3K27ac at the respective peak centers and 20 kb flanking regions. (B) Data underlying panel (A) were used to generate boxplots showing the signal distributions encompassing the entire 40 kb genomic region. The significance of differences of means was assessed by unpaired two-tailed t-tests with Welch's correction (**** $p < 0.0001$, *** $p < 0.001$). The differences of means for CBF1 independent compared to CBF1 dependent E2 peaks for H3K4me1 (-0.3004 ± 0.7957; p=0.706), H3K4me3 (0.4323 ± 1.411; p= 0.7595), and H3K27ac (-0.5184 ± 0.3501: p= 0.1396) were not statistically significant. Box plot whiskers extend to 1.5x interquartile range. (C) Table summarizing means and SEMs of histone modifications analyzed in (A) and (B).

CBF1 independent EBNA2 peaks are significantly enriched for EBF1 binding motifs and EBF1 signal intensities in LCLs.

To further investigate CBF1 independent EBNA2 binding to chromatin, *de novo* motif enrichment analyses of the two subclasses of EBNA2 binding sites were performed separately. Strikingly, the motif of EBF1, an important player in B cell development, was identified as the only and also highly enriched TF motif in the CBF1 independent EBNA2 peak subset, while CBF1 and EBF1 motifs, as well as a CBF1/EBF1 composite core motif, show up in the top five motifs of the CBF1 dependent EBNA2 peak set (Fig. 5A). In order to look at peak sets of similar size 243 out of 1546 CBF1 dependent peaks we randomly selected and re-analyzed. For this reduced set, only the CBF1 and EBF1 motifs were significantly enriched (data not shown). Since the vast majority of EBNA2 binding sites are also present in LCLs, we could use publicly available ChIP-seq data for EBF1 in LCLs to investigate EBF1 enrichment at CBF1 independent compared to dependent sites (Fig. 5B and C). Average CBF1 signal enrichment at EBNA2 binding sites did not significantly differ between CBF1 independent and dependent sites. However, EBF1 signal was highly and significantly enriched at CBF1 independent compared to CBF1 dependent sites, indicating a potential role for EBF1 in mediating CBF1 independent EBNA2 binding to chromatin. Further quantitative correlation analyses focusing on signal intensities of EBNA2, CBF1, EBF1, and PU.1/SPI1 (Fig. 5D and 5E) were performed to rank these co-occurring factors in a quantitative manner. PU.1/SPI1 was included since it had been suggested to serve as DNA anchor for EBNA2 in the past. As expected, CBF1 showed the highest correlation in signal

distribution with EBNA2 at EBNA2 peaks (r_s = 0.46) as well as genome wide (r_s = 0.5). Most strikingly, EBF1 highly correlated with EBNA2 signals at EBNA2 peaks (r_s = 0.4) as well as genome wide (r_s = 0.42). However, PU.1/SPI1 and EBNA2 signal intensities correlated weakly at EBNA2 peaks (r_s = 0.19) as well as genome wide (r_s = 0.17). A genome wide correlation, including all 84 TF ChIP-seq data sets provided by ENCODE for LCLs (28), revealed that CBF1 indeed represents the best EBNA2 correlating TF, immediately followed by EBF1. Other TFs, including PU.1/SPI1 show moderate or weak signal correlation (data not shown).

Figure 5. The EBF1 binding motif is highly enriched at CBF1 independent binding sites and the EBF1 signal correlates with EBNA2 binding signal distributions.

(A) *De novo* identified DNA sequence motifs and the respective E-values at CBF1 independent and dependent EBNA2 binding sites as discovered by MEME-ChIP (30). The analysis was performed for different sized data sets. TFs predicted to recognize the respective motifs, as assigned by TOMTOM (using the hocomoco v10 data base), are listed. If multiple TFs with comparable significances were assigned to one motif, the motif was designated as “core motif” for this subset. (B) CBF1 independent (dark blue) and dependent (light blue) EBNA2 binding sites were compared for CBF1 and EBF1 enrichment in LCLs. The average signal intensities for all EBF1 and all CBF1 peaks in LCLs are shown as reference for comparison (green), respectively. (C) The underlying data of panel B was used to generate box plots depicting signal distributions. An unpaired two-tailed t-test with Welch’s correction (**** $p < 0.0001$) was performed to determine significant differences between means. Box plot whiskers extend to 1.5x interquartile range. (D) Scatter plots of CBF1, PU.1, and EBF1 versus EBNA2 signal intensities for EBNA2 peaks in LCLs. For each transcription factor the maximal signal intensity was set to 1 to plot signal intensities as relative signal. Each dot represents one EBNA2 peak. Correlation analyses were performed and Spearman correlation coefficients (r_s) were calculated for each pair. A perfect correlation results in a line (upper left panel) and r_s = 1 for EBNA2. Spearman correlation coefficients (rs) were calculated for E2 (1.0), CBF1 (0.46), PU.1/SPI1 (0.19) and EBF1 (0.4). (E) Genome wide quantitative correlation study of EBNA2, CBF1, PU.1, and EBF1 binding intensities represented as matrix. The human genome was divided in 100 bp bins and mapped reads per bin were counted. A correlation coefficient using Spearman correlation was calculated for each TF pair and is displayed and color coded in the matrix.

EBF1 recruits EBNA2 to CBF1 independent binding sites.

To directly test if EBF1 can bind EBNA2 we performed co-immunoprecipitation (Co-IP) studies in DG75$^{doxHA\text{-}E2}$ CBF1 wt and ko cells. These Co-IP experiments revealed that EBF1 binds to EBNA2 in both CBF1 proficient as well as CBF1 deficient cells (Fig. 6).

Figure 6. EBNA2 and EBF1 form protein complexes in CBF1 proficient and deficient DG75 B cells.
DG75$^{doxHA\text{-}E2}$ CBF1 wt and CBF1 ko B cells were transfected with EBF1 expression plasmids or empty vector controls. EBNA2 expression was induced by Dox treatment directly after transfection or cells were left untreated. Total cellular extracts were harvested after 24 h and subjected to immunoprecipitation (IP) using EBF specific antibodies and then assayed by Western blot (WB) using EBF1 and EBNA2 specific antibodies. Total cell lysates (L) represent 1% of the cells used for IP (n=2, one representative experiment is shown).

Since, CBF1 was neither required nor inhibitory for EBF1/EBNA2 complex formation we asked if EBNA2 requires EBF1 to bind to either CBF1 independent or dependent chromatin sites. To this end, EBF1 protein levels were strongly reduced by siRNA mediated knock down (Fig. 7A and B). EBNA2 binding to chromatin was tested by ChIP followed by quantitative PCR (ChIP-qPCR) for six selected enhancer loci, three CBF1 independent and three CBF1 dependent sites, which also bind CBF1 and EBF1 in LCLs (Fig. 7C and D). While EBNA2 binding to CBF1 independent peaks was significantly reduced after EBF1 knock-down, CBF1 dependent EBNA2 binding was not significantly changed at reduced EBF1 levels. Thus, although EBF1 can bind to CBF1 dependent peaks it does not contribute to EBNA2 recruitment in this context.

Figure 7. EBNA2 requires EBF1 to bind to its CBF1 independent binding sites.
DG75$^{doxHA\text{-}E2}$ CBF1 wt or CBF1 ko B cells were transfected with a mixture of scrambled non-targeting siRNAs (siCNTRL) or EBF1 specific siRNAs (siEBF1). 8 h post transfection, EBNA2 transcription was induced. 24 h post transfection, cells were harvested and analyzed by immunoblots and ChIP-qPCR. (A) Representative immunoblots showing expression levels of EBNA2, EBF1, CBF1, and GAPDH before and after knockdown (n=3). EBF1 negative Jurkat cell lysate served as a negative control. (B) Protein band intensities were quantified by densitometry. The change of

EBF1 protein expression in siRNA (siEBF1) treated compared to non-treated cells (CNTRL) is significant according to paired t-test when indicated. (C and D) EBNA2 (E2) binding signals and peak tracks as obtained in DG75$^{doxHA\text{-}E2}$ (DG75) and EBNA2, CBF1 and EBF1 binding peaks tracks in LCLs are shown for three CBF1 independent (C) and three CBF1 dependent (D) EBNA2 binding sites. ChIP-qPCR results for EBNA2 binding to chromatin before and after EBF1 knock are shown below the chromatin profiles. Standard deviations and p-values, based on Student's paired t-test, are indicated.

Discussion

EBNA2 can regulate cellular gene expression in CBF1 deficient B cells

Despite the ubiquitous expression of its anchor protein CBF1, EBNA2 is preferentially recruited to B cell specific enhancers and super enhancers (10, 19, 20, 31, 32). The underlying mechanism that recruits EBNA2 specifically to these sites in B cells is still not understood and hard to study in the constitutive presence of CBF1. Since it was expected and also shown by other labs that CBF1 knock-down is not compatible with long term proliferation of LCLs (15, 16), we used a CBF1 deficient EBV negative B cell line to study whether EBNA2 can activate cellular genes and bind to chromatin in the absence of CBF1. This CBF1 deficient B cell line had been generated by targeted homologous recombination in DG75, a somatic cell line derived from an EBV negative Burkitt's lymphoma (33). The proliferation of DG75 cells is driven by the reciprocal t (8;14) translocation which hyper-activates c-MYC expression and which renders proliferation of this cell line CBF1 independent (21).

EBNA2 target gene expression has been intensively studied by many groups using different experimental systems, different B cell lines, methodologies, and different statistical evaluations (21, 22, 34-43). Our genome wide gene expression studies confirm previously described EBNA2 cellular target genes e.g. CD21, SLAMF1, RHOH, HEY1 or CCR7 (22) also identify novel cellular EBNA2 target genes including long non-coding RNAs and micro RNAs. Notably, EBNA2 also controls a smaller but well defined set of CBF1 independent target genes. A selection of targets was validated by qPCR and confirmed the robust regulation of targets in both cell lines proving a strong biological activity of EBNA2 in CBF1 deficient B cells.

A direct comparison of these target gene collections across all studies is thus difficult and would be misleading. For selected EBNA2 target genes the comparison can be made. While some target genes were identified in all studies, others appear specifically in distinct B cell lines as exemplified by the EBNA2 target gene CXCR7 which is induced in LCLs and BL41, a Burkitt's lymphoma cell line, but not in BJAB, a human lymphoblastoid B cell line (22, 43). These findings suggest, that activation of a subset of EBNA2 target genes requires specific cellular factors that, unlike EBF1, are not ubiquitously expressed. The DG75 cell lines used here express extremely low levels of the cellular transcription factors IRF4 and BATF, which are both well expressed in LCLs (data not shown). In addition, chromatin signatures at enhancer positions that can be bound by EBNA2 are distinct for DG75 and naïve B cells. Thus, EBNA2 target gene activation is fine-tuned by multiple factors in B cells. We thus do not want to exclude that additional rate limiting transcription factors apart from EBF1 control EBNA2 functions. A comparative analysis of CBF1 proficient and deficient B cells with distinct transcription factor signatures will be required to identify these additional factors.

CBF1 dependent induced targets were strongly enriched for biological processes involved in immunoglobulin receptor binding functions and a broad array of enzymatic activities. While CBF1 independent EBNA2 induced targets were not significantly enriched for any biological processes, repressed and CBF1 independent targets could be assigned to multiple biological processes involving immune responses. Some of these repressed B cell specific genes like CD79A/mb1, CD79B/B29, VpreB3 have been described previously (21, 22, 44). These targets are well characterized EBF1 induced target genes in mice (18, 45-48) and have been confirmed in human cells (49). Recently, it has been demonstrated that EBNA2 promotes the formation of new CBF1 and EBF1 chromatin binding sites (16). We speculate that EBNA2 might redirect EBF1 to novel chromatin sites and thereby deplete EBF1 activities required for target gene activation.

EBF1 is a chromatin anchor for EBNA2

Several lines of evidence support a dynamic model for CBF1/DNA complex formation. Rather than functioning as a pre-bound DNA anchor, this dynamic model suggests that

CBF1 is recruited to its DNA binding sites, when complexed to cellular or viral binding partners. Notch (50), EBNA2 (16, 41), the EBV viral protein EBNA3C (51) and also RTA (52), the KSHV derived CBF1 binding protein, all promote CBF1/chromatin complex formation and influence chromatin site recognition. We propose that additional tissue-specific cellular or viral factors guide CBF1 associated activator or repressor proteins to functional regulatory elements in the cell.

Our genome-wide EBNA2 ChIP-Seq studies revealed that EBNA2 can bind to chromatin in a CBF1 independent manner. We used publicly available information on transcription factor occupancy in LCLs or peripheral human B cells to characterize different subpopulations of EBNA2 binding sites: i) EBNA2 binding sites shared by or unique to either LCLs or DG75 and ii) CBF1 independent and dependent binding sites. CBF1 independent binding sites are found in CBF1 proficient and CBF1 deficient cells. The total number of EBNA2 binding sites found in DG75 cells was significantly smaller than the number of binding sites found in LCLs, although EBNA2 was expressed abundantly in DG75 transfectants. Most EBNA2 binding sites initially identified in DG75 cells were shared by LCLs. In LCLs, CBF1 independent binding sites score as strong EBNA2 binding sites.

In silico transcription factor binding analysis predicted CBF1 and EBF1 to be bound at CBF1 dependent binding sites while CBF1 independent EBNA2 binding sites where predicted to bind EBF1 only. Thus these latter binding sites might have low affinity for CBF1 suggesting that EBF1 might be a B cell specific chromatin co-factor for EBNA2, which enhances complex formation also in CBF1 proficient LCLs and DG75 at sites with low affinity for CBF1.

For our study, we re-analyzed publicly available primary data sets and correlated signal intensities of transcription factors either at a genome wide level or by focusing on EBNA2 binding sites. These quantitative correlation studies on CBF1, PU.1/SPI1, EBF1, and EBNA2 signal intensities revealed a strong positive correlation of CBF1 and EBF1 to EBNA2 and weak correlation of CBF1 and EBF1 to each other. Surprisingly, PU.1/SPI1 binding activity correlated neither with EBNA2 nor CBF1 nor EBF1 binding activity. A physical interaction of PU.1/SPI1 and EBNA2 has been described, but was never characterized in detail (53, 54). Transient promoter reporter studies had

previously suggested that both, PU.1/SPI1 and CBF1, are critical for transactivation of the viral LMP1 promoter by EBNA2 (12, 14, 55). However, inactivation of the LMP1 promoter PU.1/SPI1 binding site in the viral genome did not grossly change the transformation potential of the viral mutants. LMP1 expression and proliferation was diminished but not abolished while inactivation of the EBF1 binding site ablated LMP1 expression (15). Until today, there is no experimental proof indicating that EBNA2 is recruited to chromatin by PU.1/SPI1 (16). If the pioneer factor PU.1/SPI1 does not serve as chromatin anchor for EBNA2, it could facilitate the access of transcription factors to compacted chromatin or prevent chromatin silencing at the respective enhancer regions (56).

Here we show that EBNA2 and EBF1 can form complexes in cells and thus provide the first evidence that EBF1 interacts with a viral protein. Only a few cellular binding partners of EBF1 have been described so far. EBF1 can bind DNA as a homodimer (57), but can further interact and cooperate with other transcription factors like MEF2C (58), the deoxygenase TET2, an enzyme involved in the DNA demethylation process (59), or the histone acetyltransferase CBP (60). EBF1 also binds to CNOT3, a subunit of the CCR4-NOT complex (61) which regulates multiple steps in RNA metabolism including transcription, nuclear RNA export and RNA decay (62) and thereby also modulates target gene profiles of EBF. In addition, two multi-zinc finger proteins, ZNF423 and ZNF521, antagonize the biological activity of EBF1 and thereby might promote tumorigenesis (63). It should be mentioned, that in B cells with a single exception (CNOT3), these interactions have been described after expressing at least one binding partner ectopically or using cross-linking reagents before co-immunoprecipitations have been performed (58). Thus, it appears that EBF1 protein-protein interactions are particular difficult to detect at the endogenous expression levels in B cells. While EBNA2/CBF1 interactions can be readily detected in LCLs, we and others have tried and failed to detect EBNA2/EBF1 complexes from LCLs until today (16).

In order to define the contribution of EBF1 to EBNA2 chromatin binding, EBF1 protein expression was downregulated by siRNA. These knock down experiments proved that EBNA2 needs EBF1 to bind efficiently to CBF1 independent chromatin sites in both,

CBF1 proficient and deficient cells. In contrast, EBNA2 binding to CBF1 dependent sites was not impaired by EBF1 siRNA knock down and thus was defined to be EBF1 independent although EBF1 is present.

In summary, the genetic ablation of CBF1 expression in B cells provides novel valuable insights into the molecular mechanisms of EBNA2 action. We could differentiate two functionally distinct subclasses of EBNA2 binding sites and characterize them in detail. Since EBNA2/EBF1 complex formation could be demonstrated in CBF1 proficient and CBF1 deficient cells heterotrimeric complexes might be formed and EBF1 can serve as co-factor of EBNA2. Whether these complexes activate or repress transcription might depend on their composition and the chromatin context of enhancer and promoters they bind to. Any working hypothesis to be tested will have to take into account the dimeric nature of EBNA2 and EBF1 as well as the fact, that CBF1 and EBF1 are co-expressed and also their binding motifs might overlap (64). Our future studies will need to explore the architecture of these complexes in order to understand, how pre-formed EBNA2/CBF1 complexes can use EBF1 to guide EBNA2 to B cell specific enhancers and thereby provide B cell specificity to EBNA2 activities.

Acknowledgements:

We thank Sophie Beer for critical reading of the manuscript.

Material and Methods

Plasmids

pcDNA3 (pCDNA3) and EBF1-myc expression plasmid (pCDNA3.EBF1-5xmyc) were kindly provided by Mikael Sigvardsson (65). pCKR74.2 is a Dox (doxycycline) inducible HA- (haemagglutinin) tagged EBNA2 expression plasmid (pCKR74.2) based on pRTR (66, 67).

Cell lines and cell culture conditions

The cells were maintained as suspension cultures in RPMI 1640 medium (Gibco Life Technologies) supplemented with 10 % FCS (fetal calf serum, Bio&Sell), 4 mM L-Glutamine and 1 x penicillin/streptomycin (Gibco Life Technologies). The DG75$^{ER/EBNA2}$ CBF1 wt and ko cells (SM295 and SM296) have been described (21, 22). The ER/EBNA2 (estrogen receptor hormone binding domain EBNA2) fusion protein was activated by cultivating the cells in cell culture medium supplemented with 1 µM ß-estradiol. The DG75$^{doxHA\text{-}E2}$/CBF1 wt (CKR128-34) and the DG75$^{doxHA\text{-}E2}$/CBF1 ko (CKR178-10) cell lines carry the Dox inducible HA-EBNA2 expression plasmid (pCKR74.2). They were cultivated in 1 µg/ml puromycin containing media. EBNA2 expression was induced by doxycycline treatment (1µg/ml).

Genome wide expression analysis by application of the Human Gene 2.0 ST array (Affymetrix) and relative quantification of transcripts by real-time RT-PCR

Total RNA was extracted from 1x10^7 cells induced for 24 h with 1 µM ß-estradiol using the Qiagen RNeasy Mini Kit. Expression analysis starting from 100ng of total cellular RNA was performed using the Ambion® WT Expression Kit (Applied Biosystems) and subsequently the GeneChip® WT Terminal Labeling and Hybridization Kit (Affymetrix) followed by the the GeneChip Human Gene 2.0 ST array (Affymetrix) according to the manufacturer's protocol. Affymetrix CEL files have been processed in Bioconductor/R using robust multiarray average (RMA) for normalization and summarization and limma for differential expression and significance. Quality has been checked using the array QualityMetrics package. Additional filtering based on the fold change between the two conditions was applied with different stringency, individually described in the legend of the tables and figures. Analyzation and Visualization of the Microarray was performed using Genesis, available at http://genome.tugraz.at. Real-time RT-PCR analysis was performed as described previously (68). Primers used for RT-qPCR were designed applying Primer3 software (http://primer3.ut.ee/) and selection of mature transcripts was ensured by amplification across exon-exon junctions. Primers used for real-time RT-PCR are summarized in S1 Table. All data were normalized for the relative abundance of the GAPDH transcript.

Gene ontology analysis

GOrilla is a tool to identify and visualize enriched GO terms in ranked lists of genes (http://cbl-gorilla.cs.technion.ac.il/). Enrichment is defined as E = (b/n) / (B/N), with N = the total number of genes, B = the total number of genes associated with a specific GO term, n = the number of genes in the top of the user's input list and b = the number of genes in the intersection. The threshold for n is selected by GOrilla by maximizing E and statistical significance is computed taking into account the multiple hypothesis tests arising due to the maximization.

All GO terms for which B < 10 were ignored. GO terms with a q-value (FDR) ≤ 10^{-4} were selected and ranked for their enrichment score given by GOrilla.

As induction and repression was on average 8-fold stronger in DG75ER/EBNA2, CBF1 wt cells than in CBF1 ko cells, principal component analysis (PCA) was used to identify genes regulated on average or differentially between wt and ko (Fig S5). PCA was performed for all genes significantly regulated in CBF1 wt or ko cells (limma q < 0.01). The first principal component corresponded to average regulation, while the second principal component represented CBF1 dependence. Genes were first ranked according to the first principal component, i.e. top entries corresponded to genes that were induced on average in CBF1 wt and ko cells. This was repeated after reversing the list to analyze genes repressed on average. Furthermore, from each of these two lists, the top 2000 genes were selected and both were ranked according to the second principal component. Both lists were additionally reversed. Therefore, in these four additional lists, genes that are either induced or repressed on average were ranked according to their degree of CBF1 dependence.

Immunoprecipitation (IP)

$1x10^7$ $DG75^{doxHA\text{-}E2}$/CBF1 wt or $DG75^{doxHA\text{-}E2}$/CBF1 ko cells were lysed in 500 µl NP-40 lysis buffer (1% NP-40, 150 mM NaCl, 10 mM Tris-HCL pH 7.4, 1mM EDTA pH 8.0, 3% Glycerol) supplemented with complete protease inhibitor cocktail (Roche) for 1h (30 min rolling at 4°C, 30 min on ice). Precleared protein lysates were used for co-

immunoprecipitation by adding 100 µl of hybridoma supernatant (E2: α-HA R1 3F10; E.Kremmer) or 1 µg of purified antibody (α-EBF Santa Cruz Biotechnology, sc-137065) at 4°C under rotation overnight. Subsequently, 50 µl of 50% suspension of pre-blocked, equilibrated protein G-coupled Sepharose beads (GE Healthcare) were added to the lysates and incubated for 2h at 4°C under rotation. Immunoprecipitates were washed 5 times with NP-40 lysis buffer, Laemmli buffer was added to the beads, and the samples were boiled, submitted to electrophoresis by SDS-PAGE and analyzed by immunoblotting.

Immunoblotting (Western Blot)

5x 10^6 cells were lysed in 200 µl NP-40 lysis buffer (1% NP-40, 150 mM NaCl, 10 mM Tris-HCL pH 7.4, 1mM EDTA pH 8.0, 3% Glycerol) for 2 h on ice. 30 µg of total cell lysate were submitted to SDS-PAGE under reducing conditions. Immunoblotting was performed on polyvinylidene difluoride (PVDF) membranes. Western blots were probed with the following primary antibodies: α-E2 (R3 supernatant; IgG2A; E. Kremmer), α-CBF1 (RBP-J 7A11, E. Kremmer), α-EBF (Santa Cruz Biotechnology, sc-137065) and α-GAPDH (EMD Millipore MAB374). HRP-coupled secondary antibodies (Santa Cruz Biotechnology) and an ECL kit (GE Healthcare) were used for visualization. For subsequent quantification of protein levels, exposed films were scanned in transmission mode and protein band intensities were determined by densitometry using *ImageJ* software (http://rsbweb.nih.gov/ij/) (69).

Transfection

5x 10^6 DG75 cells were transfected by electroporation at 250 V and 950 µF in 250 µl reduced serum media (Opti-MEM, Gibco Life Technologies; without supplements) using 0.4 cm-electrode-gap cuvettes (Bio-Rad) and the Bio-Rad Gene Pulser.

siRNA knockdown in DG75 cells

5x 10^6 cells were transfected with 100 pmol control siRNA-A or EBF1 siRNA (both Santa Cruz Biotechnology, sc-37007 and sc-10695) by electroporation. 24 h after transfection, 1x 10^7 induced, siRNA treated cells were harvested for chromatin isolation and 5x10^6 cells for protein isolation.

Chromatin immunoprecipitation

This ChIP protocol is based on reference (59) with minor modifications as indicated below. In brief, 2x 10^7 DG75$^{doxHA\text{-}E2}$ cells were harvested and washed twice in ice cold PBS, resuspended in 20 ml RPMI 1640 (Gibco Life Technologies) and formaldehyde (1% final) was added for cross-linking. The reaction was stopped by addition of glycine (125 mM final) after 7 min and gentle shaking for 5 min at RT. Cells were pelleted and washed twice in ice cold PBS. Nuclei were isolated by washing the cells 3x with 10 ml of ice cold Lysis Buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM MgCl2, 0.5% NP-40, 1x proteinase inhibitor cocktail (PIC, Roche)) and subsequent centrifugation (300 g for 10 min at 4 °C). Nuclei were resuspended in 1 ml Sonication Buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, pH 8.0, 0.5% SDS, 1x PIC) and incubated on ice for 10 min. Chromatin was sheared to an average size of 200-300 bp by four rounds of sonication for 10 min (30 sec pulse, 30 sec pause) using a Bioruptor® device (Biogenode). Cell debris was separated by centrifugation at maximum speed for 10 min at 4 °C and chromatin containing supernatants were stored at -80 °C or directly used for IP. To prepare input DNA, 25 µl aliquots (1/10 of the amount used per IP) were saved at -80 °C. For IPs 250 µl chromatin (equals 5x 106 cells) were diluted 1:4 with IP Dilution Buffer (12.5 mM Tri-HCl, pH 8.0, 212.5 mM NaCl, 1.25 % Triton X-100, 1 x PIC) and incubated with 100 µl of hybridoma supernatant on a rotating platform at 4 °C overnight. A combination of EBNA2 and HA-tag specific antibodies (⅓ α-E2 R3 (rat IgG2a, , ⅓ α-E2 1E6 (rat IgG2a), and ⅓ α-HA R1-3F10 (rat IgG1)) was used to precipitate EBNA2 and an isotype-matched unspecific antibody mixture (⅔ α- GST 6G9 (rat IgG2a) and ⅓ α-CD23 Dog-CD3 (rat IgG1) both by E. Kremmer) was used as negative control. Protein G sepharose (GE Healthcare) was equilibrated with IP Dilution Buffer, added to the lysate and incubated at 4°C for 4 h with constant rotation. Beads were extensively

washed with: 2x Wash Buffer I (20 mM Tris-HCl, pH 8.0, 2 mM EDTA, pH 8.0, 1% Triton X-100, 150 mM NaCl, 0.1% SDS, 1x PIC), 1x Wash Buffer II (20 mM Tris-HCl, pH 8.0, 2 mM EDTA, pH 8.0, 1% Triton X-100, 500 mM NaCl, 0.1% SDS, 1x PIC), 1x Wash Buffer III (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0, 250 mM LiCl, 1% NP-40, 1% sodium deoxycholate, 1x PIC) for 5 min under rotation, and 2x with TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, pH 8.0) for 1 min. Protein-DNA complexes were eluted with 2x 150 µl Elution Buffer (25 mM Tris-HCl, pH 7.5, 10 mM EDTA, pH 8.0, 1% SDS) at 65 °C for 15 min. Input samples were adjusted to 300 µl with Elution Buffer. Eluates and input samples were incubated with Proteinase K (1.5 µg/µl final, Roche) for 1 h at 42 °C. Cross-linking was reversed by incubation at 65 °C overnight. DNA was recovered using QIAquick PCR purification kit (Qiagen).

The EBNA2 specific ChIP in LCL was performed as described above with the following modifications: Protein-protein interactions were fixated by adding disuccinimidyl glutarate (DSG, Pierce #20593, 2 mM final, using freshly prepared 0.5 M stock solution in DMSO) for 23 min at RT and prior to formaldehyde (1% final) cross-link for additional 7 min. Sonication Buffer was composed of 50 mM Tris-HCl, pH 8.0, 5 mM EDTA, pH 8.0, 0.5% SDS, 0.5% Triton X-100, 0.05% sodium deoxycholate, and 1x PIC. IP Dilution Buffer was composed of 12.5 mM Tri-HCl, pH 8.0, 187.5 mM NaCl, 1.25 mM EDTA, pH 8.0, 1.125 % Triton X-100, and 1 x PIC. For EBNA2 specific IP 50 µl of α-E2 R3 (rat IgG2a) and 50 µl α-E2 1E6 (rat IgG2a) hybridoma supernatant were applied and the same volume of isotype-matched nonspecific antibody (α- GST 6G9 (rat IgG2a) E. Kremmer) was used as negative control.

Whole-Genome Chromatin Immunoprecipitation DNA Sequencing (ChIP-Seq)

For sequencing purposes DNA concentration was measured using the Qubit® dsDNA HS Assay Kit (Thermo Fisher). A maximum of 100 ng ChIP or input derived DNA were used for library preparation (NEBNext® Ultra™ DNA Library Prep Kit for Illumina®) and subsequently subjected to deep sequencing using a HiSeq 1500 device (Illumina).

Chromatin Immunoprecipitation quantitative Polymerase Chain Reaction (ChIP-qPCR)

The amount of recovered DNA in input samples and after IP with specific antibody or an unspecific isotype-matched IgG control was quantified by qPCR using primers listed in S1 Table.

qPCR was performed using LightCycler 480 SYBR Green I Master (Roche) on a LightCycler 480 II instrument (Roche) as described previously (68). 2 technical replicates were analyzed for each biological replicate. Amplification was always conducted at 63°C. To account for differences in amplification efficiencies a standard curve was generated for each primer pair using serial dilutions of sheared DNA (input) as template. DNA quantities detected in input samples were adjusted to the amount of chromatin used per IP by multiplication with 20. Values obtained from IP samples with unspecific IgG control were subtracted from the DNA amounts recovered by IP with specific antibody. The percent of input was calculated as (DNA from specific IP corrected for IgG control background/ DNA input) x 100. To validate the ChIP, qPCR at a known (ChIP-Seq) positive locus was performed. To compromise divergent EBNA2 inducibility in wildtype and knockout cells, the percent input was calculated relative to a known negative locus (ChIP-Seq; percent input at tested locus/percent input of known negative locus). To display the change in binding, the mean relative input of the wildtype cells treated with control siRNA was set to one. A paired t-test was performed.

Bioinformatics

All bioinformatic analyses of ChIP-Seq data were conducted by using the galaxy bioinformatics platform (70) hosted and maintained by the Bioinformatics Department of the University of Freiburg. For all sequenced samples, at least 17 million reads were obtained and biological duplicates of E2 ChIP and input samples were sequenced. Reads were mapped to the human genome using Bowtie2 (71). For all samples at least 95% of reads were mappable to the human genome including at least 69% of uniquely mapping reads with one distinct location (S2 Table). Biological duplicates of mapped reads were merged and subsequently significant EBNA2 binding sites were identified using MACS2 (72) and normalizing ChIP to input samples (S2 Table). In a second step, the peaks were further filtered and “negative peaks” (negative amplitude, significantly

higher read count in the input sample), peaks located at black-listed regions (73), peaks with a very low enrichment score, and such located on chromosomes not included in the ENCODE data for GM12878 (e.g. chrY, chrUn) were excluded (S2 Table). Normalized EBNA2 signal tracks were generated by subjecting duplicate-merged ChIP and input read files to bamCompare of the deepTool package (74) and normalizing ChIP to input samples by subtraction as well as normalizing to fragments (reads) per kb per million (RPKM) to account for genome coverage. Mean signal intensities at specific peak sets were calculated using computeMatrix of the deepTools package. The details of all analyses steps are captured in a Galaxy workflow which can be downloaded at github (https://github.com/bgruening/galaxytools/tree/master/workflows/peak_calling) and re-run and analyzed in Galaxy. Data provided by public resources are listed in S3 Table.

Flow Cytometry

Inducibility of EBNA2 expression in DG75$^{doxHA\text{-}E2}$/CBF1 wt and ko cell lines was evaluated by monitoring the expression of the eGFP surrogate marker of pCKR74.2. Cells were induced for 16 h or 24 h with doxycycline, washed and fixed with 0.5% PFA in PBS. For quantification of induced cells, the FACSCalibur system (BD Biosciences) and CellQuest Pro software (BD Biosciences) were applied.

References

1. Aster JC, Pear WS, Blacklow SC. The Varied Roles of Notch in Cancer. Annual review of pathology. 2016 Dec 05. PubMed PMID: 27959635.
2. Wang H, Zang C, Liu XS, Aster JC. The role of Notch receptors in transcriptional regulation. Journal of cellular physiology. 2015 May;230(5):982-8. PubMed PMID: 25418913. Pubmed Central PMCID: 4442318.
3. Grossman SR, Johannsen E, Tong X, Yalamanchili R, Kieff E. The Epstein-Barr virus nuclear antigen 2 transactivator is directed to response elements by the J kappa recombination signal binding protein. Proceedings of the National Academy of Sciences of the United States of America. 1994;91(16):7568-72.
4. Yalamanchili R, Tong X, Grossman S, Johannsen E, Mosialos G, Kieff E. Genetic and biochemical evidence that EBNA 2 interaction with a 63-kDa cellular GTG-binding protein is essential for B lymphocyte growth transformation by EBV. Virology. 1994;204(2):634-41.

5. Zimber-Strobl U, Strobl LJ, Meitinger C, Hinrichs R, Sakai T, Furukawa T, et al. Epstein-Barr virus nuclear antigen 2 exerts its transactivating function through interaction with recombination signal binding protein RBP-J kappa, the homologue of Drosophila Suppressor of Hairless. The EMBO journal. 1994;13(20):4973-82.
6. Henkel T, Ling PD, Hayward SD, Peterson MG. Mediation of Epstein-Barr virus EBNA2 transactivation by recombination signal-binding protein J kappa. Science. 1994 Jul 1;265(5168):92-5. PubMed PMID: 8016657.
7. Hayward SD, Liu J, Fujimuro M. Notch and Wnt signaling: mimicry and manipulation by gamma herpesviruses. Sci STKE. 2006 May 16;2006(335):re4. PubMed PMID: 16705130.
8. Longnecker RM, Kieff E, Cohen JI. Epstein-Barr virus. In: Knipe DM, Howley PM, Cohen JI, Griffin DE, Lamb RA, Martin MA, et al., editors. Fields Virology. 2. 6 ed. Philadelphia: Lippincott Williams and Wilkins; 2013. p. 1898-959.
9. Kempkes B, Ling PD. EBNA2 and Its Coactivator EBNA-LP. Current topics in microbiology and immunology. 2015;391:35-59. PubMed PMID: 26428371.
10. Zhao B, Zou J, Wang H, Johannsen E, Peng CW, Quackenbush J, et al. Epstein-Barr virus exploits intrinsic B-lymphocyte transcription programs to achieve immortal cell growth. Proceedings of the National Academy of Sciences of the United States of America. 2011 Sep 6;108(36):14902-7. PubMed PMID: 21746931. Pubmed Central PMCID: 3169132.
11. Carotta S, Wu L, Nutt SL. Surprising new roles for PU.1 in the adaptive immune response. Immunological reviews. 2010 Nov;238(1):63-75. PubMed PMID: 20969585.
12. Sjoblom A, Jansson A, Yang W, Lain S, Nilsson T, Rymo L. PU box-binding transcription factors and a POU domain protein cooperate in the Epstein-Barr virus (EBV) nuclear antigen 2-induced transactivation of the EBV latent membrane protein 1 promoter. J Gen Virol. 1995;76(Pt 11):2679-92.
13. Sjoblom A, Nerstedt A, Jansson A, Rymo L. Domains of the Epstein-Barr virus nuclear antigen 2 (EBNA2) involved in the transactivation of the latent membrane protein 1 and the EBNA Cp promoters. J Gen Virol. 1995;76(Pt 11):2669-78.
14. Laux G, Adam B, Strobl LJ, Moreau-Gachelin F. The Spi-1/PU.1 and Spi-B ets family transcription factors and the recombination signal binding protein RBP-J kappa interact with an Epstein-Barr virus nuclear antigen 2 responsive cis-element. The EMBO journal. 1994;13(23):5624-32.
15. Murata T, Noda C, Narita Y, Watanabe T, Yoshida M, Ashio K, et al. Induction of Epstein-Barr Virus Oncoprotein LMP1 by Transcription Factors AP-2 and Early B Cell Factor. Journal of virology. 2016 Apr;90(8):3873-89. PubMed PMID: 26819314. Pubmed Central PMCID: 4810554.
16. Lu F, Chen HS, Kossenkov AV, DeWispeleare K, Won KJ, Lieberman PM. EBNA2 Drives Formation of New Chromosome Binding Sites and Target Genes for B-Cell Master Regulatory Transcription Factors RBP-jkappa and EBF1. PLoS pathogens. 2016 Jan;12(1):e1005339. PubMed PMID: 26752713. Pubmed Central PMCID: 4709166.
17. Boller S, Ramamoorthy S, Akbas D, Nechanitzky R, Burger L, Murr R, et al. Pioneering Activity of the C-Terminal Domain of EBF1 Shapes the Chromatin Landscape for B Cell Programming. Immunity. 2016 Mar 15;44(3):527-41. PubMed PMID: 26982363.

18. Boller S, Grosschedl R. The regulatory network of B-cell differentiation: a focused view of early B-cell factor 1 function. Immunological reviews. 2014 Sep;261(1):102-15. PubMed PMID: 25123279. Pubmed Central PMCID: 4312928.
19. Zhou H, Schmidt SC, Jiang S, Willox B, Bernhardt K, Liang J, et al. Epstein-Barr virus oncoprotein super-enhancers control B cell growth. Cell host & microbe. 2015 Feb 11;17(2):205-16. PubMed PMID: 25639793.
20. McClellan MJ, Wood CD, Ojeniyi O, Cooper TJ, Kanhere A, Arvey A, et al. Modulation of enhancer looping and differential gene targeting by epstein-barr virus transcription factors directs cellular reprogramming. PLoS pathogens. 2013 Sep;9(9):e1003636. PubMed PMID: 24068937.
21. Maier S, Santak M, Mantik A, Grabusic K, Kremmer E, Hammerschmidt W, et al. A somatic knockout of CBF1 in a human B-cell line reveals that induction of CD21 and CCR7 by EBNA-2 is strictly CBF1 dependent and that downregulation of immunoglobulin M is partially CBF1 independent. Journal of virology. 2005 Jul;79(14):8784-92. PubMed PMID: 15994772.
22. Maier S, Staffler G, Hartmann A, Hock J, Henning K, Grabusic K, et al. Cellular target genes of Epstein-Barr virus nuclear antigen 2. Journal of virology. 2006 Oct;80(19):9761-71. PubMed PMID: 16973580. Pubmed Central PMCID: 1617228.
23. Tveito S, Andersen K, Karesen R, Fodstad O. Analysis of EpCAM positive cells isolated from sentinel lymph nodes of breast cancer patients identifies subpopulations of cells with distinct transcription profiles. Breast cancer research : BCR. 2011;13(4):R75. PubMed PMID: 21816090. Pubmed Central PMCID: 3236339.
24. Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, et al. Genome-wide analysis of estrogen receptor binding sites. Nature genetics. 2006 Nov;38(11):1289-97. PubMed PMID: 17013392.
25. Rae JM, Johnson MD, Scheys JO, Cordero KE, Larios JM, Lippman ME. GREB 1 is a critical regulator of hormone dependent breast cancer growth. Breast cancer research and treatment. 2005 Jul;92(2):141-9. PubMed PMID: 15986123.
26. Nilsson S, Gustafsson JA. Estrogen receptor transcription and transactivation: Basic aspects of estrogen action. Breast cancer research : BCR. 2000;2(5):360-6. PubMed PMID: 11250729. Pubmed Central PMCID: 138658.
27. Gustafsson JA. Novel aspects of estrogen action. Journal of the Society for Gynecologic Investigation. 2000 Jan-Feb;7(1 Suppl):S8-9. PubMed PMID: 10732321.
28. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, et al. ENCODE data in the UCSC Genome Browser: year 5 update. Nucleic acids research. 2013 Jan;41(Database issue):D56-63. PubMed PMID: 23193274.
29. Kretzmer H, Bernhart SH, Wang W, Haake A, Weniger MA, Bergmann AK, et al. DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. Nature genetics. 2015 Nov;47(11):1316-25. PubMed PMID: 26437030.
30. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011 Jun 15;27(12):1696-7. PubMed PMID: 21486936. Pubmed Central PMCID: 3106185.
31. Wood CD, Veenstra H, Khasnis S, Gunnell A, Webb HM, Shannon-Lowe C, et al. MYC activation and BCL2L11 silencing by a tumour virus through the large-scale

reconfiguration of enhancer-promoter hubs. eLife. 2016;5. PubMed PMID: 27490482. Pubmed Central PMCID: 5005034.
32. Gunnell A, Webb HM, Wood CD, McClellan MJ, Wichaidit B, Kempkes B, et al. RUNX super-enhancer control through the Notch pathway by Epstein-Barr virus transcription factors regulates B cell growth. Nucleic acids research. 2016 Jun 2;44(10):4636-50. PubMed PMID: 26883634. Pubmed Central PMCID: 4889917.
33. Ben-Bassat H, Goldblum N, Mitrani S, Goldblum T, Yoffey JM, Cohen MM, et al. Establishment in continuous culture of a new type of lymphocyte from a "Burkitt like" malignant lymphoma (line D.G.-75). Int J Cancer. 1977 Jan;19(1):27-33. PubMed PMID: 188769.
34. Wang F, Gregory CD, Rowe M, Rickinson AB, Wang D, Birkenbach M, et al. Epstein-Barr virus nuclear antigen 2 specifically induces expression of the B-cell activation antigen CD23. Proceedings of the National Academy of Sciences of the United States of America. 1987 May;84(10):3452-6. PubMed PMID: 3033649.
35. Calender A, Cordier M, Billaud M, Lenoir GM. Modulation of cellular gene expression in B lymphoma cells following in vitro infection by Epstein-Barr virus (EBV). Int J Cancer. 1990 Oct 15;46(4):658-63. PubMed PMID: 1698730.
36. Knutson JC. The level of c-fgr RNA is increased by EBNA-2, an Epstein-Barr virus gene required for B-cell immortalization. Journal of virology. 1990 Jun;64(6):2530-6. PubMed PMID: 2159528.
37. Burgstahler R, Kempkes B, Steube K, Lipp M. Expression of the chemokine receptor BLR2/EBI1 is specifically transactivated by Epstein-Barr virus nuclear antigen 2. Biochem Biophys Res Commun. 1995;215(2):737-43.
38. Sakai T, Taniguchi Y, Tamura K, Minoguchi S, Fukuhara T, Strobl LJ, et al. Functional replacement of the intracellular region of the Notch1 receptor by Epstein-Barr virus nuclear antigen 2. Journal of virology. 1998;72(7):6034-9.
39. Johansen LM, Deppmann CD, Erickson KD, Coffin WF, 3rd, Thornton TM, Humphrey SE, et al. EBNA2 and activated Notch induce expression of BATF. Journal of virology. 2003 May;77(10):6029-40. PubMed PMID: 12719594.
40. Pegman PM, Smith SM, D'Souza BN, Loughran ST, Maier S, Kempkes B, et al. Epstein-Barr virus nuclear antigen 2 trans-activates the cellular antiapoptotic bfl-1 gene by a CBF1/RBPJ kappa-dependent pathway. Journal of virology. 2006 Aug;80(16):8133-44. PubMed PMID: 16873269. Pubmed Central PMCID: 1563820.
41. Mohan J, Dement-Brown J, Maier S, Ise T, Kempkes B, Tolnay M. Epstein-Barr virus nuclear antigen 2 induces FcRH5 expression through CBF1. Blood. 2006 Jun 1;107(11):4433-9. PubMed PMID: 16439682.
42. Zhao B, Maruo S, Cooper A, M RC, Johannsen E, Kieff E, et al. RNAs induced by Epstein-Barr virus nuclear antigen 2 in lymphoblastoid cell lines. Proceedings of the National Academy of Sciences of the United States of America. 2006 Feb 7;103(6):1900-5. PubMed PMID: 16446431.
43. Lucchesi W, Brady G, Dittrich-Breiholz O, Kracht M, Russ R, Farrell PJ. Differential gene regulation by Epstein-Barr virus type 1 and type 2 EBNA2. Journal of virology. 2008 Aug;82(15):7456-66. PubMed PMID: 18480445. Pubmed Central PMCID: 2493322.
44. Boccellato F, Anastasiadou E, Rosato P, Kempkes B, Frati L, Faggioni A, et al. EBNA2 Interferes with the Germinal Center Phenotype by Downregulating BCL6 and

TCL1 in Non-Hodgkin's Lymphoma Cells. Journal of virology. 2007 Mar;81(5):2274-82. PubMed PMID: 17151114.
45. Sigvardsson M, Clark DR, Fitzsimmons D, Doyle M, Akerblad P, Breslin T, et al. Early B-cell factor, E2A, and Pax-5 cooperate to activate the early B cell-specific mb-1 promoter. Molecular and cellular biology. 2002 Dec;22(24):8539-51. PubMed PMID: 12446773. Pubmed Central PMCID: 139876.
46. Hagman J, Travis A, Grosschedl R. A novel lineage-specific nuclear factor regulates mb-1 gene transcription at the early stages of B cell differentiation. The EMBO journal. 1991 Nov;10(11):3409-17. PubMed PMID: 1915300. Pubmed Central PMCID: 453069.
47. Hagman J, Belanger C, Travis A, Turck CW, Grosschedl R. Cloning and functional characterization of early B-cell factor, a regulator of lymphocyte-specific gene expression. Genes & development. 1993 May;7(5):760-73. PubMed PMID: 8491377.
48. Akerblad P, Rosberg M, Leanderson T, Sigvardsson M. The B29 (immunoglobulin beta-chain) gene is a genetic target for early B-cell factor. Molecular and cellular biology. 1999 Jan;19(1):392-401. PubMed PMID: 9858563. Pubmed Central PMCID: 83897.
49. Bohle V, Doring C, Hansmann ML, Kuppers R. Role of early B-cell factor 1 (EBF1) in Hodgkin lymphoma. Leukemia. 2013 Mar;27(3):671-9. PubMed PMID: 23174882.
50. Krejci A, Bray S. Notch activation stimulates transient and selective binding of Su(H)/CSL to target enhancers. Genes & development. 2007 Jun 01;21(11):1322-7. PubMed PMID: 17545467. Pubmed Central PMCID: 1877745.
51. Kalchschmidt JS, Gillman AC, Paschos K, Bazot Q, Kempkes B, Allday MJ. EBNA3C Directs Recruitment of RBPJ (CBF1) to Chromatin during the Process of Gene Repression in EBV Infected B Cells. PLoS pathogens. 2016 Jan;12(1):e1005383. PubMed PMID: 26751214. Pubmed Central PMCID: 4708995.
52. Carroll KD, Bu W, Palmeri D, Spadavecchia S, Lynch SJ, Marras SA, et al. Kaposi's Sarcoma-associated herpesvirus lytic switch protein stimulates DNA binding of RBP-Jk/CSL to activate the Notch pathway. Journal of virology. 2006 Oct;80(19):9697-709. PubMed PMID: 16973574. Pubmed Central PMCID: 1617261.
53. Bheda A, Yue W, Gullapalli A, Shackelford J, Pagano JS. PU.1-dependent regulation of UCH L1 expression in B-lymphoma cells. Leukemia & lymphoma. 2011 Jul;52(7):1336-47. PubMed PMID: 21504384. Pubmed Central PMCID: 4435811.
54. Yue W, Davenport MG, Shackelford J, Pagano JS. Mitosis-specific hyperphosphorylation of Epstein-Barr virus nuclear antigen 2 suppresses its function. Journal of virology. 2004 Apr;78(7):3542-52. PubMed PMID: 15016877.
55. Johannsen E, Koh E, Mosialos G, Tong X, Kieff E, Grossman SR. Epstein-Barr virus nuclear protein 2 transactivation of the latent membrane protein 1 promoter is mediated by J kappa and PU.1. Journal of virology. 1995;69(1):253-62.
56. Zaret KS, Carroll JS. Pioneer transcription factors: establishing competence for gene expression. Genes & development. 2011 Nov 01;25(21):2227-41. PubMed PMID: 22056668. Pubmed Central PMCID: 3219227.
57. Treiber N, Treiber T, Zocher G, Grosschedl R. Structure of an Ebf1:DNA complex reveals unusual DNA recognition and structural homology with Rel proteins. Genes &

development. 2010 Oct 15;24(20):2270-5. PubMed PMID: 20876732. Pubmed Central PMCID: 2956205.
58. Kong NR, Davis M, Chai L, Winoto A, Tjian R. MEF2C and EBF1 Co-regulate B Cell-Specific Transcription. PLoS genetics. 2016 Feb;12(2):e1005845. PubMed PMID: 26900922. Pubmed Central PMCID: 4762780.
59. Guilhamon P, Eskandarpour M, Halai D, Wilson GA, Feber A, Teschendorff AE, et al. Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2. Nature communications. 2013;4:2166. PubMed PMID: 23863747. Pubmed Central PMCID: 3759038.
60. Zhao F, McCarrick-Walmsley R, Akerblad P, Sigvardsson M, Kadesch T. Inhibition of p300/CBP by early B-cell factor. Molecular and cellular biology. 2003 Jun;23(11):3837-46. PubMed PMID: 12748286. Pubmed Central PMCID: 155219.
61. Yang CY, Ramamoorthy S, Boller S, Rosenbaum M, Rodriguez Gil A, Mittler G, et al. Interaction of CCR4-NOT with EBF1 regulates gene-specific transcription and mRNA stability in B lymphopoiesis. Genes & development. 2016 Oct 15;30(20):2310-24. PubMed PMID: 27807034. Pubmed Central PMCID: 5110997.
62. Miller JE, Reese JC. Ccr4-Not complex: the control freak of eukaryotic cells. Critical reviews in biochemistry and molecular biology. 2012 Jul-Aug;47(4):315-33. PubMed PMID: 22416820. Pubmed Central PMCID: 3376659.
63. Mesuraca M, Chiarella E, Scicchitano S, Codispoti B, Giordano M, Nappo G, et al. ZNF423 and ZNF521: EBF1 Antagonists of Potential Relevance in B-Lymphoid Malignancies. BioMed research international. 2015;2015:165238. PubMed PMID: 26788497. Pubmed Central PMCID: 4695665.
64. Miele L. Transcription factor RBPJ/CSL: a genome-wide look at transcriptional regulation. Proceedings of the National Academy of Sciences of the United States of America. 2011 Sep 06;108(36):14715-6. PubMed PMID: 21873209. Pubmed Central PMCID: 3169161.
65. Sigvardsson M, O'Riordan M, Grosschedl R. EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes. Immunity. 1997 Jul;7(1):25-36. PubMed PMID: 9252117.
66. Jackstadt R, Roh S, Neumann J, Jung P, Hoffmann R, Horst D, et al. AP4 is a mediator of epithelial-mesenchymal transition and metastasis in colorectal cancer. The Journal of experimental medicine. 2013 Jul 01;210(7):1331-50. PubMed PMID: 23752226. Pubmed Central PMCID: 3698521.
67. Bornkamm GW, Berens C, Kuklik-Roos C, Bechet JM, Laux G, Bachl J, et al. Stringent doxycycline-dependent control of gene activities using an episomal one-vector system. Nucleic acids research. 2005;33(16):e137. PubMed PMID: 16147984. Pubmed Central PMCID: 1201338. Epub 2005/09/09. eng.
68. Harth-Hertle ML, Scholz BA, Erhard F, Glaser LV, Dolken L, Zimmer R, et al. Inactivation of Intergenic Enhancers by EBNA3A Initiates and Maintains Polycomb Signatures across a Chromatin Domain Encoding CXCL10 and CXCL9. PLoS pathogens. 2013 Sep;9(9):e1003638. PubMed PMID: 24068939.
69. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nature methods. 2012 Jul;9(7):671-5. PubMed PMID: 22930834.

70. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. Galaxy: a platform for interactive large-scale genome analysis. Genome research. 2005 Oct;15(10):1451-5. PubMed PMID: 16169926. Pubmed Central PMCID: 1240089.
71. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nature methods. 2012 Apr;9(4):357-9. PubMed PMID: 22388286. Pubmed Central PMCID: 3322381.
72. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). Genome biology. 2008;9(9):R137. PubMed PMID: 18798982. Pubmed Central PMCID: 2592715.
73. Derrien T, Estelle J, Marco Sola S, Knowles DG, Raineri E, Guigo R, et al. Fast computation and applications of genome mappability. PloS one. 2012;7(1):e30377. PubMed PMID: 22276185. Pubmed Central PMCID: 3261895.
74. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research. 2014 Jul;42(Web Server issue):W187-91. PubMed PMID: 24799436. Pubmed Central PMCID: 4086134.

Supporting information:

S1 Figure

Control panels documenting estrogen responses in ER/EBNA2 expressing DG75 cells compared to estrogen treated untransfected parental cell lines.

(A) DG75 parental cells (DG75 CBF1 wt), CBF1 deficient (DG75 CBF1 ko), ER/EBNA2 expressing (DG75$^{ER/EBNA2}$ CBF1 wt), and CBF1 deficient ER/EBNA2 expressing DG75 cells (DG75$^{ER/EBNA2}$ CBF1 ko) were treated with estrogen for 24 h or were left untreated. Total cellular RNA was isolated and submitted to gene expression analysis using the Human Gene 2.0 ST array. All probe sets represent single transcripts. For each condition 3 biological replicates were examined. Each vertical column in the heatmap represents the results obtained from a single microarray. Horizontal rows represent data obtained for a particular probe set across all cell lines and conditions after normalization of expression values on a scale ranging from -2.0 to 2.0 for each probe set. Expression levels of 950 transcripts which change expression levels at least 2-fold ($p \leq 0.05$) in response to estrogen in DG75 ER/EBNA2 cells are displayed. The relative high, medium and low expression values are represented by red, white and blue, respectively. Vertical columns are ranked according to fold changes in ER/EBNA2 expressing DG75 from highest induction on top to highest repression levels at the bottom. (B) RNA expression levels of a panel of previously described estrogen responsive target genes in DG75 cells after estrogen treatment (RMA= robust multi array average). (C) RNA expression level of previously defined EBNA2 target genes in DG75 ER/EBNA2 cells after estrogen induction.

S2 Figure

Heatmap representing the 132 transcripts regulated at least 2-fold ($p \leq 0.001$) by EBNA2 in CBF1 deficient DG75$^{ER/EBNA2}$ cells.

Total cellular RNA was isolated and submitted to gene expression analysis using the Human Gene 2.0 ST array. All probe sets represent single transcripts. For each condition 3 biological replicates were examined. Each vertical column represents the results obtained by a single microarray. Horizontal rows represent data obtained for a particular probe set across all cell lines and conditions after normalization of expression values on a scale ranging from -2.0 to 2.0 for each probe set. The relative high, medium and low expression values are represented by red, white, and blue color, respectively. Vertical columns are ranked according to fold changes in ER/EBNA2 expressing DG75 CBF1 ko from highest induction on top to highest repression levels at the bottom. The transcript cluster ID and the assigned genes/transcripts are indicated. Note that not more than five assigned genes are listed (*). If no assignment was available the chromosomal position is indicated (**).

S3 Figure

Validation of gene array hybridization results by quantitative RT-PCR. (A) Relative transcript levels of EBNA2 target genes were quantified from total RNA samples of the indicated cell lines by RT-qPCR. All results were normalized to actin B transcript levels. (B) For comparison the expression levels measured by gene array hybridization are shown in parallel.

S4 Figure

Heatmap showing microRNAs regulated at least 1.5-fold ($p \leq 0.05$) by EBNA2 in DG75$^{ER/EBNA2}$ CBF1 wt cells (for all details see Figure S1).

S5 Figure

Identification of individual target gene subsets based on Principle Component Analysis

Since on average target gene expression changes in CBF1 positive cells were stronger than in CBF1 negative cells, principle component analysis on EBNA2 regulated genes was used to identify specific subpopulations: The first principle component (green arrow) describes the upregulation of genes in both cell lines, the second principle component (red arrow) describes the degree of CBF1 dependence. The scatter blots depict all genes (A) or the top 2000 (B) induced/repressed genes which are regulated in at least one cell line.

S6 Figure

Doxycycline inducible HA-EBNA2 expression in CBF1 proficient or deficient DG75 B cells. (A) pRTR$^{doxHA\text{-}E2}$ vector used to generate stable DG75 cell lines. The coding sequence for EBNA2 fused to a N-terminal HA-tag (HA-E2), plus a preceding intron of the beta-globin gene for enhanced expression, was cloned into the pRTR vector (Jackstadt et al., 2013, Bornkamm et al., 2005) using SfiI restriction sites. The

bidirectional promoter simultaneously drives the expression of HA-EBNA2 and the bicistronic reporter construct consisting of a truncated nerve growth factor receptor gene (tNGFR) and enhanced green fluorescent protein (eGFP) gene upon doxycycline induction. (B) Expression of HA-EBNA2 was induced with 1 µg/ml doxycycline (Dox) for 24 h and monitored by quantifying eGFP expression via flow cytometry and scored at least 89% with a maximum of 5% difference between DG75 CBF1 wt and ko cells. Data from one representative experiment (n=3) and percentages of induced cells are shown. (C) Western Blot analysis confirming the expression of HA-EBNA2 in DG75$^{doxHA\text{-}E2}$ cell lines 24 h post induction with 1 µg/ml Dox. The absence of CBF1 expression in the DG75$^{doxHA\text{-}E2}$ CBF1 ko cell line is confirmed. EBF1 and PU.1/SPI1 are shown for comparison. GAPDH serves as loading control.

S1 Table Primer qPCR

S2 Table Summary ChIP-Seq Results

S3 Table Public Resources Used for this Study

Figure 1

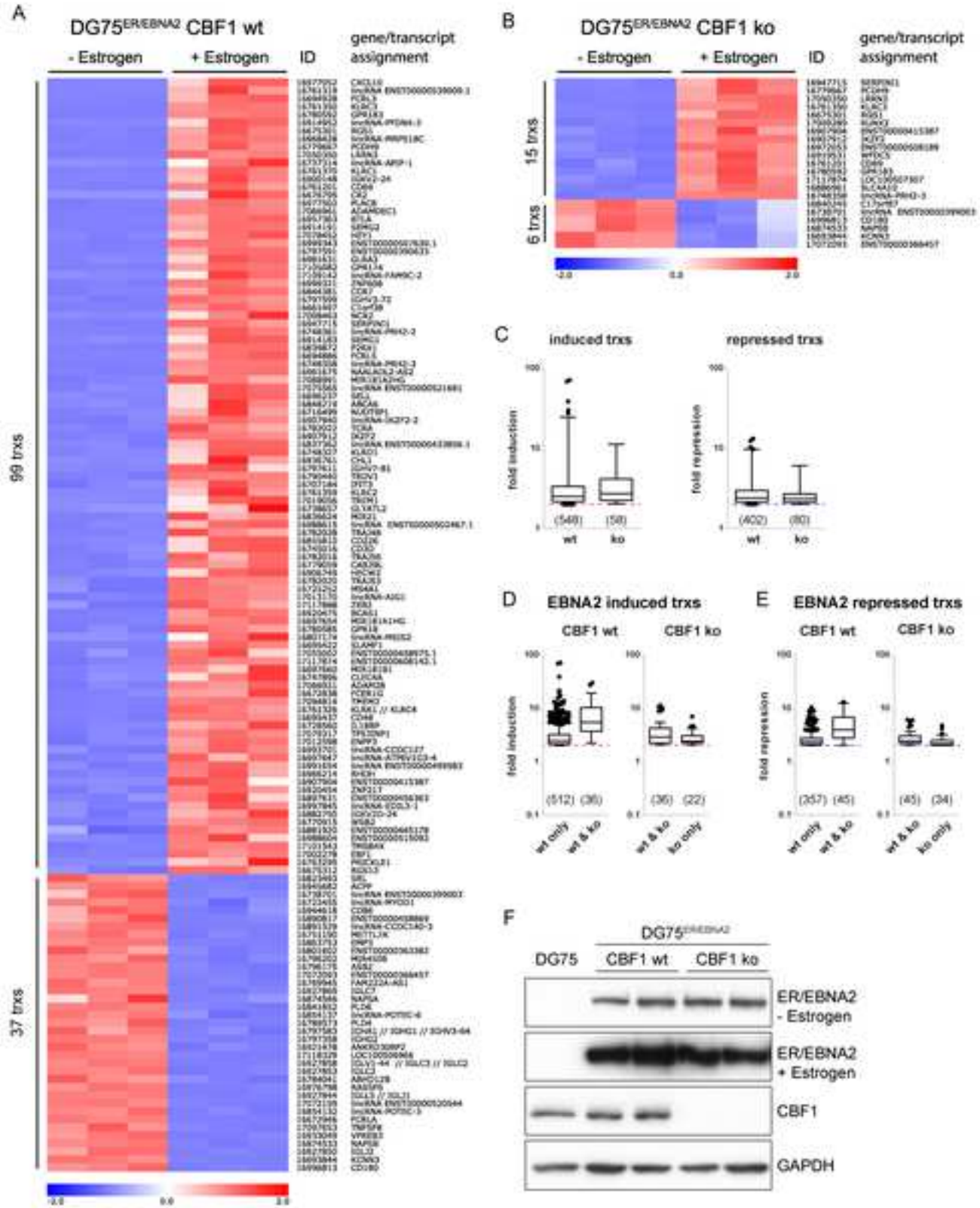


Figure 2

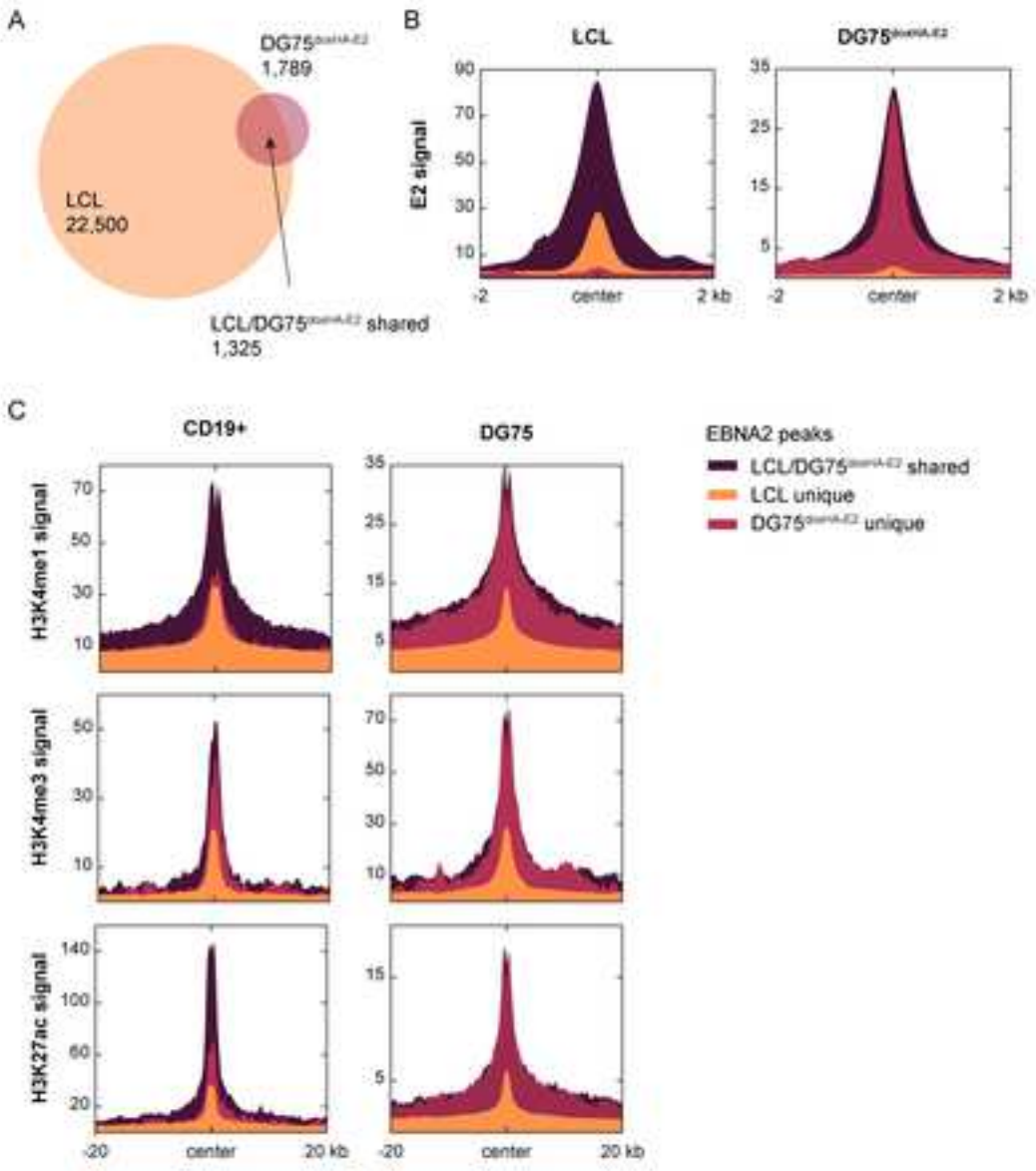
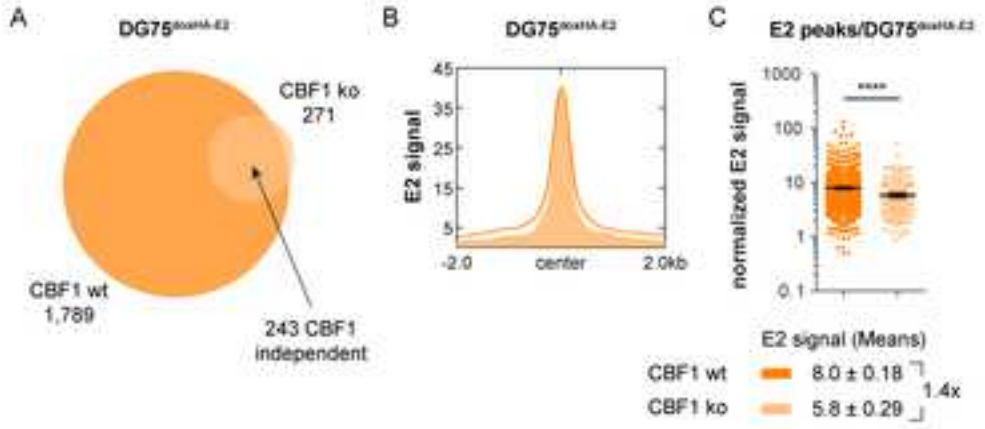
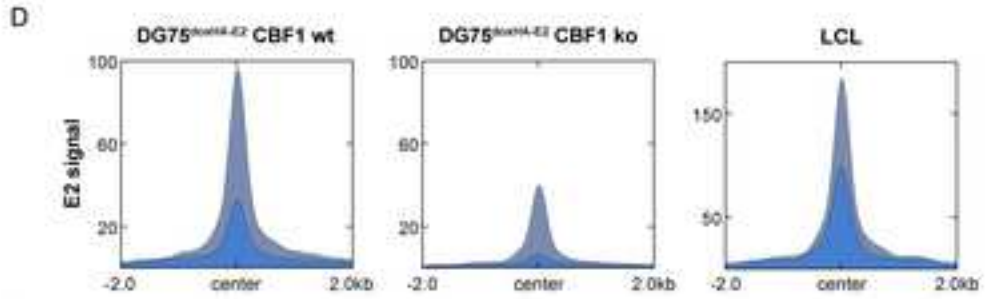
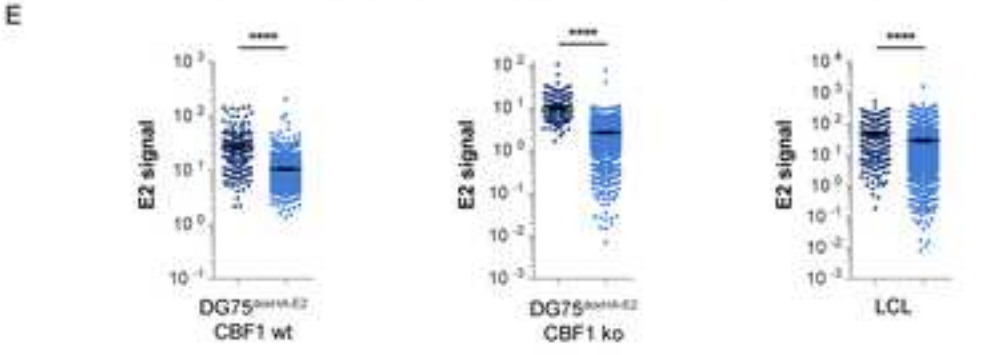


Figure 3





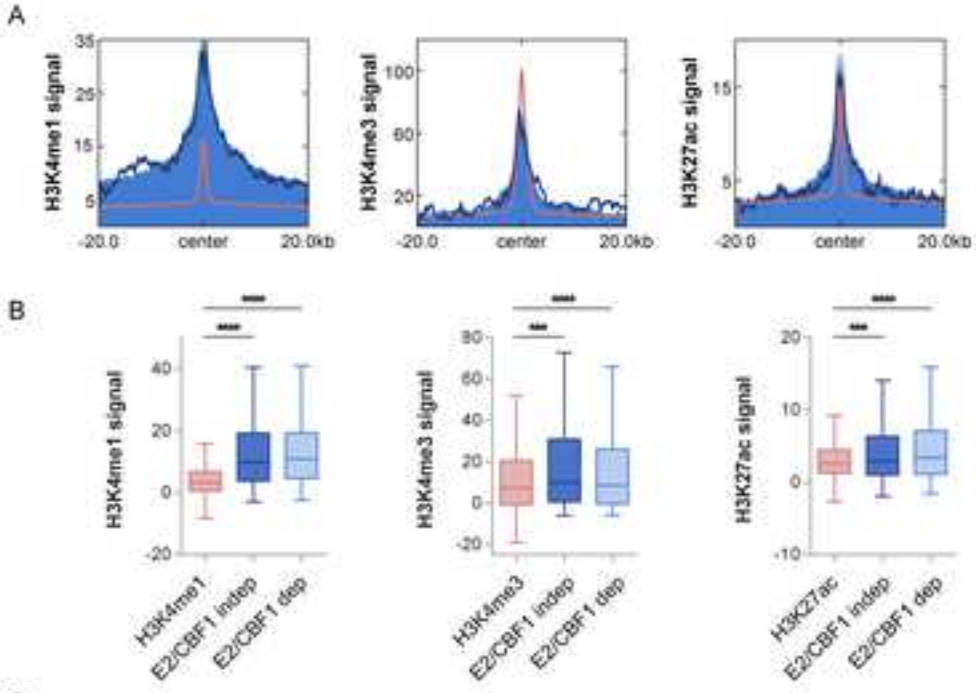


F

	DG75 CBF1 wt	DG75 CBF1 ko	LCL
E2 mean signal intensities:			
CBF1 independent	27.0 ± 1.6	10.2 ± 0.6	49.6 ± 4.6
CBF1 dependent	10.6 ± 0.2	2.7 ± 0.1	29.4 ± 1.7

Figure 4

Click here to download Figure Figure 4 Glaser et al.,-01.tif



C

Signal	H3K4me1	H3K4me3	H3K27ac
E2 Peaks			
Histone modification	4.5 ± 0.02	12.3 ± 0.09	3.3 ± 0.01
CBF1 independent	12.9 ± 0.74	17.0 ± 1.29	4.5 ± 0.32
CBF1 dependent	13.2 ± 0.28	16.6 ± 0.57	5.0 ± 0.15

Figure 5

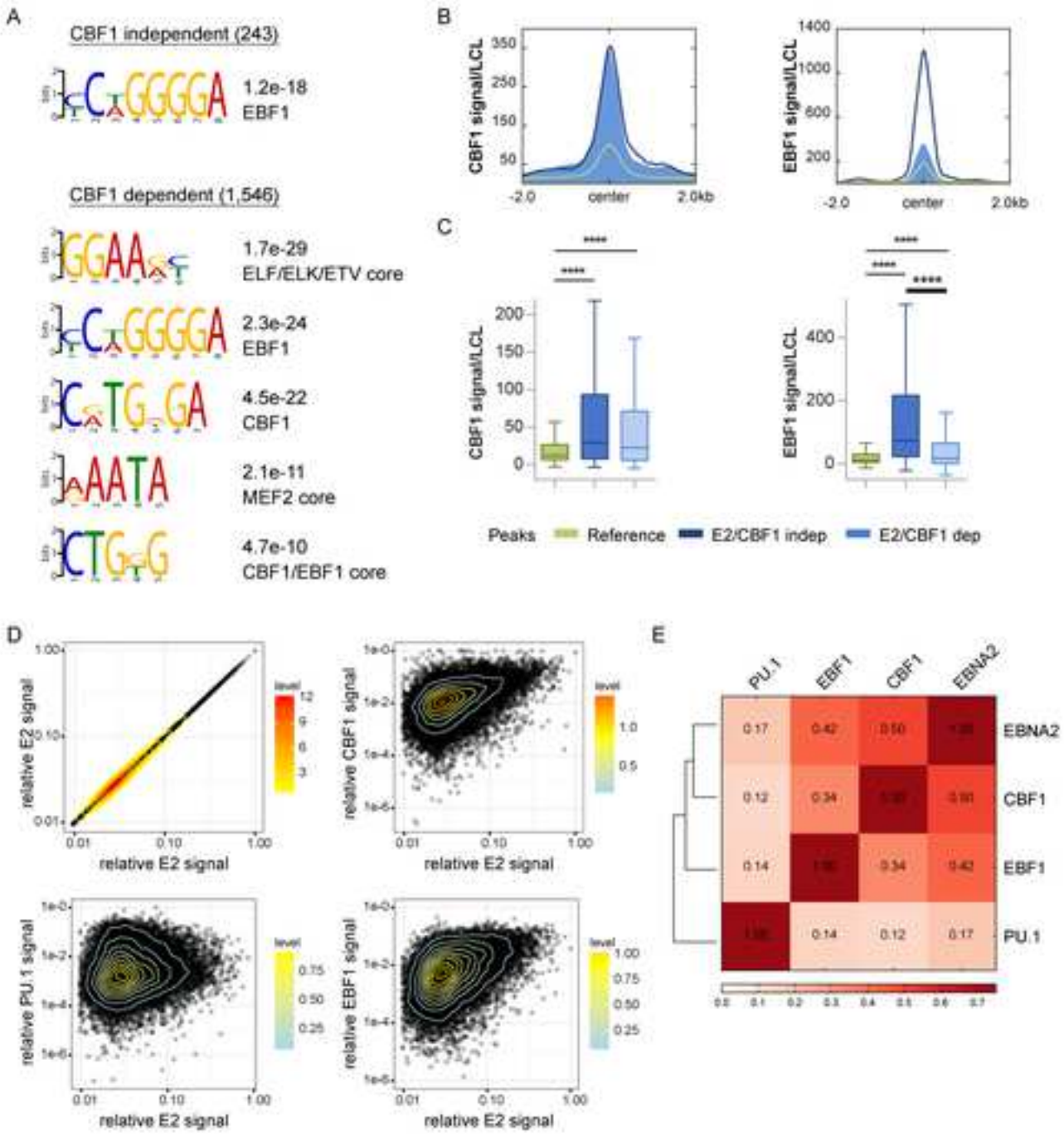


Figure 6

Click here to download Figure Figure 6 Glaser et al.,-01.tif

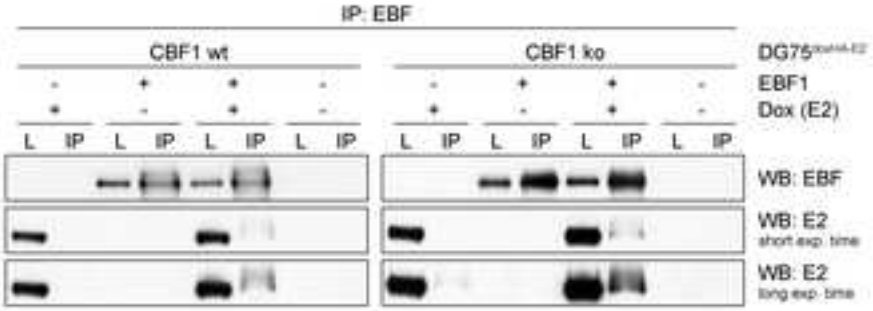
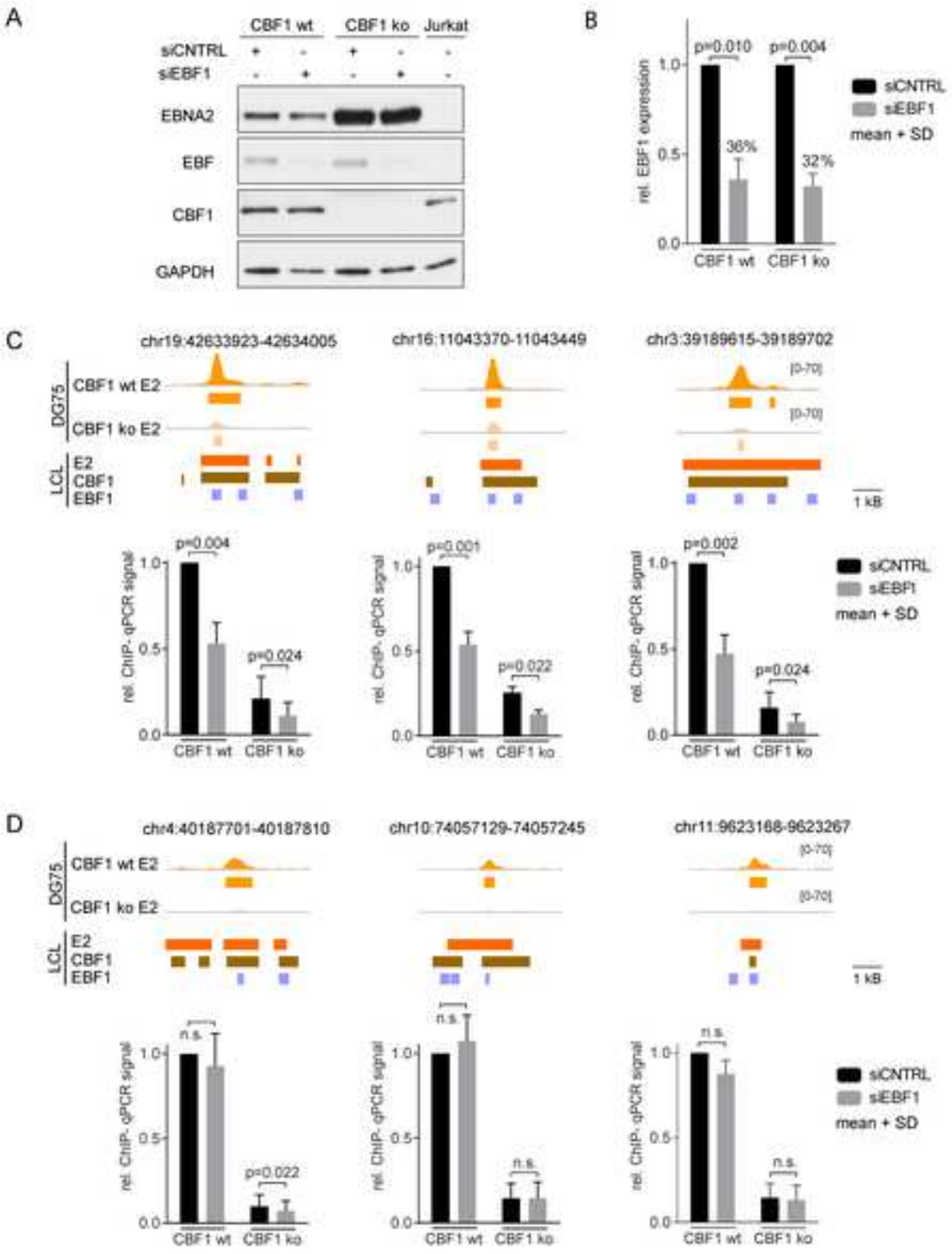


Figure 7



S1 Fig


Click here to access/download
Supporting Information
S1_Fig.tif

S2 Fig

Click here to access/download

Supporting Information

S2_Fig.tif

S4 Fig

Click here to access/download

Supporting Information

S4_Fig.tif

S5 Fig

Click here to access/download
Supporting Information
S5_Fig.tif

S1 Table