

## **Herpesviral induction of germline transcription factor DUX4 is critical for viral gene expression**

Stephanie Walter<sup>1</sup>, Vedran Franke<sup>2</sup>, Nir Drayman<sup>3</sup>, Emanuel Wyler<sup>2</sup>, Savaş Tay<sup>3</sup>, Markus Landthaler<sup>2</sup>, Altuna Akalin<sup>2</sup>, Armin Ensser<sup>1</sup> and Florian Full<sup>1,4</sup>

<sup>1</sup> Institute for Clinical and Molecular Virology, University Hospital Erlangen, 91054 Erlangen, Germany

<sup>2</sup>Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Helmholtz Society, Berlin, Germany

<sup>3</sup> The Pritzker School for Molecular Engineering, The University of Chicago, Chicago, IL, USA, 60637.

<sup>4</sup> Institute of Virology, Medical Center - University of Freiburg, Freiburg, Germany

\*Correspondence to:

Florian Full, PhD

Institute for Clinical and Molecular Virology

University Hospital Erlangen

91054 Erlangen

Germany

phone: +49 9131 8526783

florian.full@uk-erlangen.de

## **Abstract:**

DUX4 is a transcription factor and a master regulator of embryonic genome activation (EGA). During early embryogenesis, EGA is crucial for maternal to zygotic transition at the 8-cell stage in order to overcome silencing of genes and enable transcription from the zygotic genome. In adult somatic cells, DUX4 expression is largely silenced. Activation is likely pathogenic, and in adult muscle cells causes genetic disorder Facioscapulohumeral Muscular Dystrophy (FSHD).

We identified activation of DUX4 expression upon lytic replication of the herpesviruses HSV-1, HCMV, EBV and KSHV, but not of adenoviruses, negative strand RNA viruses or positive strand RNA viruses. We demonstrate by RNA-Seq analysis that DUX4 expression upon herpesviral replication leads to the induction of hundreds of DUX4 target genes including germline-specific retroelements as well as several members of the TRIM, PRAMEF and ZSCAN protein families. Moreover, we show that DUX4 expression is a direct consequence of herpesviral infection. DUX4 can be stimulated by overexpression of HSV-1 immediate early proteins, indicating active induction of EGA genes by herpesviral infection. We further show that DUX4 expression is critical for driving HSV-1 gene expression.

Our results show that viruses from alpha-, beta- and gamma-herpesvirus subfamilies induce DUX4 expression and downstream germline-specific genes and retroelements. We hypothesize that herpesviruses induce DUX4 expression in order to induce an early embryonic-like transcriptional program that prevents epigenetic silencing of the viral genome and facilitates herpesviral gene expression.

## Introduction:

Herpesviruses are a major health burden worldwide, with a prevalence of 60 – 100% dependent on the virus species and geography<sup>1–3</sup>. Herpesviruses cause a number of prevalent diseases, like oral and genital herpes, chickenpox, shingles and infectious mononucleosis. While they rarely induce life-threatening infections in healthy humans, infections in e.g. immunocompromised patients, transplant recipients or newborns can be severe<sup>4,5</sup>. Some herpesviruses such as Epstein-Barr virus (EBV) and Kaposi's sarcoma-associated herpesvirus (KSHV) are also known to be the causative agent of particular cancer types<sup>6</sup>. Due to their lifelong persistence in the host, the risk of reactivation and developing an lytic infection is constantly present. Treatment options are limited and therapies able to counteract the viral persistence or to clear the virus out of the body are unavailable. A better understanding of the mechanisms involved during herpesviral infection is essential.

In a recent study, we observed that the upregulation of the cellular protein Tripartite-motif 43 (TRIM43) upon HSV-1 infection is dependent on the embryonic transcription factor DUX4<sup>7</sup>, which was recently confirmed by Friedel et al.<sup>8</sup>. DUX4 is a germline transcription factor exclusively expressed from the 4-cell to the 8-cell state, a short phase during human embryonic development<sup>9–11</sup>. There, DUX4 is essential for embryonic genome activation (EGA, equivalent to zygotic genome activation in other mammals)<sup>10–12</sup>, a transcriptional activation event that leads to the induction of hundreds of target genes and allows the embryo to proceed further in development<sup>13,14</sup>. Concomitantly, it induces retroelements such as LTRs and ERVs<sup>9,10,15</sup> which function as alternative promoters, regulate EGA-related genes and mediate pluripotency<sup>16–18</sup>. After this limited period of activation, DUX4 is silenced throughout all adult tissues, except for spermatocytes<sup>19</sup>. An exception to this is the genetic disorder Facioscapulohumeral Muscular Dystrophy (FSHD). In FSHD, DUX4 silencing is lost and aberrant DUX4 expression in adult muscle cells leads to apoptosis and degeneration of the affected muscles<sup>20–23</sup>. The reasons for the onset of this disease are complex and not yet fully understood, though they are thought to lie within the unique genomic location of DUX4<sup>22,24</sup>. It resides within the macrosatellite repeat array D4Z4<sup>25,26</sup>, that harbors up to 100

copies of DUX4<sup>27</sup>, with only the most distal copy encoding for a protein<sup>21</sup>. During early embryonic development, this macrosatellite array is transformed into heterochromatin after the 8-cell stage<sup>9,10</sup> and its repression is strictly controlled by chromatin remodeling factors and repressors which have not been fully elucidated<sup>19,24,28</sup>.

Given the peculiar expression pattern of DUX4 and its upregulation in herpesviral infection, we aimed to determine whether DUX4 plays a role during herpesviral infection and to shed light on its potential functions. Our data reveal aspects of a mechanism in herpesviral infection that has hitherto not been characterized. We show that DUX4 expression during herpesviral replication is a general mechanism that is conserved among all herpesviral subfamilies. For HSV-1, we could demonstrate that the viral proteins ICP0 and ICP4 promote DUX4 expression and hundreds of DUX4 target genes, including several classes of endogenous retroelements, and that DUX4 is critical for HSV-1 gene expression. This points at so far unknown endogenous functions of DUX4 and extends our knowledge of this unique transcription factor.

## Results

We previously showed that DUX4 protein is expressed upon lytic infection with HSV-1<sup>7</sup>. To confirm the expression of DUX4 and DUX4 target genes TRIM49 and ZSCAN4 upon HSV-1 infection, we probed their expression by Western blot analysis in infected primary human foreskin fibroblast (HFF) cells and 293T cells (Fig. 1A-C). Next, we investigated whether this DUX4 expression is unique for HSV-1 or a general feature of herpesviral infection. We therefore assessed the expression of DUX4 upon infection with one member of the  $\alpha$ -,  $\beta$ -, and  $\gamma$ -herpesvirus subfamilies, namely HSV-1, HCMV and KSHV respectively. Upon HCMV infection of primary HFF cells, DUX4 as well as ZSCAN4 protein expression could be robustly detected by Western Blot (Figure 1D and E). In addition, DUX4 mRNA was strongly increased upon infection with HSV-1 and HCMV, as well as lytic reactivation of KSHV (Fig. 1 F-H). Moreover, we could also detect induction of known DUX4 target genes TRIM48, TRIM49, ZSCAN4, ZSCAN5a, ZSCAN5d and RFPL4A<sup>29,30</sup> for all three viruses (Figure 1I-K,

Fig. S1). This indicates that DUX4 is functional and acts as a transcriptional activator in herpesvirus infected cells. DUX4 expression was also induced upon infection of primary human melanocytes with HSV-1, emphasizing the physiological relevance of DUX4 induction (Fig. S2). In addition, reanalysis of several published next generation sequencing datasets from HSV-1 infected cells confirmed that observation<sup>7,31,32</sup>. Comparison of three independent RNA-seq datasets in 3 different cell types showed no expression of DUX4 and DUX4 target genes (e.g. FRG1) in uninfected cells (Fig 2A). Upon infection with HSV-1, however, transcription is induced for DUX4 as well as all known DUX4 target genes (Fig. 2A). Interestingly, we also confirmed the expression of HSAT-II satellite repeat transcripts. HSAT-II repeats are known DUX4 targets<sup>12,33</sup> and have recently been shown to play a role in HCMV replication<sup>34</sup>. (Fig. S3) Moreover, the reanalysis of a published Precision Run-On sequencing (PRO-seq.) dataset<sup>34</sup> showed no RNA-polymerase-II occupancy at the DUX4 locus in uninfected cells, indicating that the locus is completely silenced under normal conditions (Fig. 2A). In contrast, HSV-1 infection induced a robust PRO-seq signal at the DUX4 locus. Taken together, these data show that the induction of DUX4 expression after herpesviral infection is a common feature for human herpesviruses and is conserved among  $\alpha$ -, $\beta$ -, and  $\gamma$ -herpesviruses.

Next, we wanted to address the significance of herpesviral DUX4 induction for the infection itself, since DUX4 is a germline transcription factor that is not expressed in healthy adult tissue<sup>9,10,19</sup>. We first hypothesized that the induction of DUX4 might be part of an antiviral response of the cell to viral infection triggered by a herpesviral pathogen associated molecular pattern (PAMP), or by herpesviral induction of a cellular DNA-damage response (DDR). We reanalyzed several published RNA-sequencing datasets, investigating DUX4 mRNA expression in response to cellular changes, but found no evidence for DUX4 expression. In addition, triggering a DNA-damage response by treating cells with Bleocin or Etoposide did not induce DUX4 expression (Fig. S4A). The only evidence for DUX4 expression came from a Chromatin-IP-sequencing (ChIP-Seq.) dataset<sup>36</sup> investigating the binding of the HSV-1 infected cell protein 4 (ICP4) to the cellular genome. ICP4 is a potent

activator of viral transcription, part of the viral tegument, and essential for replication of HSV-1<sup>37,38</sup>. Our reanalysis showed a very strong binding of the ICP4 protein to the entire DUX4 locus (Figure 2A). Of note, the peak intensity of ICP4 binding at the DUX4 locus was among the highest in the entire host genome. This ICP4 binding was a strong indication for an active induction of DUX4 by HSV-1 infection. Moreover, experiments with phosphonoacetic acid (PAA), an inhibitor of the viral DNA-polymerase, showed that herpesviral DNA replication is dispensable for DUX4 expression (Fig. S4B), indicating that induction of the DUX4 gene takes place at the immediate-early stage of the viral gene expression cascade. Furthermore, after infection with UV-inactivated virus, DUX4 protein was still induced, indicating that incoming components of the virion contribute to DUX4 induction (Fig. S4B). Analysis of the kinetics of DUX4 expression upon HSV-1 infection further demonstrated that DUX4 protein could be detected as early as 4h post infection with protein levels constantly increasing over the course of infection (Fig. 2B). Reanalysis of RNA-sequencing data from Rutkowski et al. further demonstrated that known DUX4 target genes have variable dynamics during infection with most DUX4 target genes being upregulated at 3-4 hours post infection (Fig. S5). To further elucidate the involvement of HSV-1 tegument proteins in the induction of DUX4 expression we performed infection experiments with HSV-1 mutants depleted of the immediate early proteins ICP0, ICP4, ICP27 and gamma-34.5. We could show that wildtype (wt) HSV-1 infection as well as infection with HSV-1- $\Delta$   $\gamma$ 34.5 resulted in DUX4 expression, whereas DUX4 expression is abrogated in cells infected with HSV-1 lacking either ICP0 or ICP4 (Fig. 2C and D). In addition, the transient co-expression of ICP0 and ICP4 is sufficient to induce DUX4 expression in the absence of viral infection (Fig. 2E), confirming that ICP0 and ICP4 are necessary and sufficient for inducing DUX4 expression.

DUX4 is a transcription factor active in the nucleus where it executes its physiological function by activating its target genes. We used a recombinant HSV-1 expressing ICP4-YFP and VP26-RFP in order to visualize progression of infection and localization of DUX4. ICP4 is an immediate-early gene and part of the tegument, whereas VP26, as the small capsid protein of the virus, is expressed late in the infection cycle. By co-staining of cells with a

DUX4 antibody, we could demonstrate that only ICP4+/VP26- cells (YFP+/-/RFP- cells) show nuclear expression of DUX4, indicating activation of DUX4 within early stages of infection (Figure 3a). In contrast, ICP4+ and VP26+ cells (YFP+ and RFP+ cells) show either no DUX4 expression or an aberrant cytoplasmic localization of DUX4. This indicates that DUX4 is only briefly activated by ICP0 and ICP4 during the early phase of infection. To test this, we treated cells with PAA, which arrests viral infection at the stage of viral DNA replication, i.e. after immediate-early gene expression but before late gene expression (Figure 3B). PAA treated cells showed a strong increase in expression of DUX4 and its target genes by qRT-PCR analysis of HDF cells infected with HSV-1 (Figure 3B), as well as the DUX4 protein by WB analysis in 293T cells (Figure S4B), supporting the notion of a transient activation of DUX4 during the early stages of HSV-1 infection.

It is known that the physiological function of DUX4 during embryonic genome activation at the 2-8 cell stage is to directly bind to DNA and activate genes and retroelements that are necessary for developmental progression<sup>9,15</sup>. In particular endogenous retroelements act as promoters for downstream genes, and binding of DUX4 to retroelements activates transcription of respective genes<sup>18,39</sup>. In order to further analyze the role of DUX4 in herpesviral replication, we performed endogenous DUX4 ChIP-Seq experiments. HFF cells were infected with HSV-1 for 8h and 16h, and binding of DUX4 to both the host and viral genome analyzed. The ChIP-Seq experiments revealed about 11000 DUX4 binding sites within the host genome. Analysis of the binding sites from our endogenous DUX4 ChIP-Seq and comparison with the DUX4 ChIP-Seq performed by Young et al.<sup>18</sup> with overexpressed DUX4 protein in muscle cells showed almost identical consensus binding sites (Fig. 4A). We observed a strong correlation between DUX4 binding sites at 8h and 14h time points (Fig. 4B). Surprisingly, most DUX4 binding sites were not at transcriptional start sites or within exon regions of genes, but within intronic and intergenic regions (Fig. 4C). Comparison of DUX4-ChIP-Seq binding sites upon HSV-1 infection with RNA-Seq data published from Rutkowski et al. showed that a subset of genes with DUX4 binding sites get upregulated during the course of infection whereas other genes with DUX4 binding sites are either not

regulated or even get downregulated upon infection (Fig. 4D). In depth analysis of intronic / intergenic binding sites showed a strong preference of DUX4 binding to repetitive genetic elements (Fig 4E). We found predominant binding of DUX4 to long-terminal repeat (LTR-) elements and short interspersed nuclear elements (SINE) and only a small fraction of binding sites in actual genes (Fig. 4E). During early embryonic development and in particular EGA, it is known that activation of endogenous retroelements can generate alternative promoters for expression of genes, resulting in transcription of developmental genes that are essential for further embryonic development<sup>16,17</sup>. Comparison of DUX4 binding during EGA and herpesviral infection shows that a specific subset of retroelements is bound in both settings (Fig. 4F). Accordingly, reanalysis of RNA-seq data showed a DUX4-mediated activation of a specific subset of LTR-elements that are expressed during EGA. Both HSV-1 infection as well as DUX4 overexpression leads to an induction of the MLT- and THE-class of retroelements, indicating that the expression upon herpesviral infection is driven by DUX4. In addition, HSV-1 infection also induces the LTR12C class of retroelements. Expression of LTR12C retroelements is independent of DUX4 (Fig. 4F), and results in the expression of the previously described C10rf159 antisense transcript<sup>31</sup>, since this transcript starts at an LTR12C element within the first intron of its host gene. Moreover, comparison of genes harboring DUX4 binding sites showed a partial overlap between genes expressed during herpesviral infection and genes expressed during EGA (Fig. 4G)<sup>40</sup>. This indicates that herpesviral induction of the germline-specific transcription factor DUX4 activates transcription of a subset of EGA-specific genes and retroelements. In contrast to the cellular genome, we could not identify robust binding of DUX4 to the HSV-1 genome. The DUX4-ChIP conditions optimized for DUX4 binding to the cellular genome resulted in a high background for the viral genome, most likely due to different physical properties that affect sonication and differences in chromatin accessibility. However, we cannot fully exclude direct binding of DUX4 to the viral genome. We observed a weak potential DUX4 binding site in the promoter region of the viral ICP27 gene, which also contains a potential DUX4 consensus-binding site, and conventional ChIP-experiments followed by quantitative PCR showed an enrichment of



DUX4 compared to controls (Fig. S6), suggesting at direct binding of DUX4 to the ICP27 promoter region.

Our data revealed a robust expression of DUX4 during lytic replication of all human herpesvirus subfamilies and thus we wanted to address the physiological consequences in regard to herpesviral gene expression and replication. DUX4 is located within the 4q35 region of chromosome 4<sup>25,26</sup> and the gene is particularly complicated to target by CRISPR/Cas9. In healthy individuals the locus consists of 10-100 repeat units, and a copy of DUX4 resides in every unit<sup>27</sup>. However, only one of the D4Z4 repeat arrays, the most distal one adjacent to the telomeres is encoding for the functional protein<sup>21</sup>. In addition to multiple repeats on chromosomes 4 (4q allele) there is also another heterochromatic repeat array on chromosome 10 (10q allele)<sup>41,42</sup> which is likely to interfere with CRISPR/CAS9 based knockout strategies. We therefore used a transient approach in 293T cells to target DUX4 at the population level using 4 different guide RNAs targeting different sequences within the DUX4 locus. The guide RNAs had been successfully used to silence DUX4 in FSHD myocytes previously<sup>48</sup>. We observed a knockdown but not complete knockout of DUX4 upon HSV-1 infection by Western analysis one week after selection with puromycin (Fig. 4A). The knockdown of DUX4 resulted in reduced protein expression of HSV-1 immediate-early genes ICP0, ICP4, and ICP27 upon infection compared to wildtype-cells (Fig.4A). In order to assess the effect of DUX4 knockdown on the transcription of the entire viral genome, RNA-Seq from HSV-1 infected knockdown and wildtype cells was performed. Unsupervised clustering of wildtype and knockdown cell lines showed altered expression of most HSV-1 genes in knockdown cells (Fig. 4B). Whereas the expression of early genes like UL23 and UL54 is lower in wildtype cells, the expression of late genes like UL44/UL45 is higher, indicating that the viral expression cascade proceeds faster in wildtype cells in the presence of DUX4 (Fig. 4B). Analysis of the cellular transcriptome confirmed a reduction of DUX4 target genes in knockdown cells, both by RNA-Seq analysis and qRT-PCR for ZSCAN4 (Fig.4C, D). In addition, we also observed a decrease in expression of endogenous retroelement upon DUX4 knockdown compared to wildtype cells, demonstrating the dependence on DUX4 (Fig.

S7). In order to further assess the role of DUX4 in viral gene expression we optimized our CRISPR/CAS9 protocol and could achieve a complete knockout of DUX4 in 293T cells. Infection of DUX4 knockout cells resulted in an almost complete loss of most HSV-1 genes tested in Western Blot, like ICP0, ICP4, ICP27 and VP16. However, the expression levels of HSV-1 gD were not affected by knockout of DUX4, which is in accordance with published data showing that the expression of gD is independent of the viral expression cascade<sup>49</sup>.

## Discussion

In humans the female oocyte is produced during female gametogenesis in the embryo and is then stored in prophase I of the meiosis for up to 50 years. The oocyte transcription is halted by epigenetic mechanisms, and stored mRNAs mostly control development. After fertilization the zygote is formed, maternal to zygotic transition (MZT) takes place and with the onset of EGA the zygotic genome starts to control transcription<sup>13</sup>. In order to induce transcription, the embryo has to overcome silencing of the genome, which is regulated by repressive epigenetic modifications like DNA methylation, histone modifications and a shortage of the cellular transcription machinery<sup>14,43</sup>. DUX4 has been shown to be essential for EGA by activating hundreds of genes that are necessary for further development, and several endogenous retroelements<sup>9,10,15</sup>. Although the detailed function of most of DUX4 target genes remains elusive, it is thought that DUX4 induces several important factors that are involved in the creation of a permissive environment that allows transcription from the newly formed diploid genome.

Upon herpesviral infection the viral genome enters the nucleus and gets chromatinized, although the degree of chromatinization of the viral genome is discussed controversially, in particular for HSV-1. However, it is widely accepted in the field that herpesviruses are subjected to epigenetic silencing, and that they evolved strategies to prevent epigenetic silencing of their genome in order to allow transcription necessary for viral replication and virus transmission. We show that herpesviruses from all human subfamilies induce a robust expression of DUX4 upon lytic infection. Considering that alpha-, beta-, and gamma-herpesviruses were split into three separate lineages about 200 million years ago<sup>44</sup>, this hints at a highly conserved mechanism in the coevolution of herpesviruses with their respective hosts. For HSV-1, we demonstrate that DUX4 expression is induced by viral tegument/immediate-early proteins ICP0 and ICP4, indicating that this is an active induction by herpesviruses. Viral ICP0 and ICP4 induce expression of DUX4 at early stages of infection for a brief period, indicated by nuclear staining of DUX4 in ICP4+/VP26- cells. In ICP4+/VP26+ double positive cells we either observed no DUX4 staining or cytoplasmic

DUX4 staining. Blocking viral DNA replication with PAA results in higher levels of DUX4 and its target genes, suggesting DUX4 behaves similarly to the strong HSV-1 beta genes, which are upregulated early in infection and are shut down during the switch to viral DNA replication and late gene expression. Interestingly, this very short induction of DUX4 resembles the mechanism of DUX4 induction during EGA, where a short burst of DUX4 is sufficient for target gene activation and reprogramming of the embryo but prevents toxicity mediated by DUX4.

Moreover, we could demonstrate that herpesviral DUX4 induction is essential for efficient herpesviral gene expression. Depletion of DUX4 from cells results in a loss of most herpesviral protein expression, as shown for HSV-1. We hypothesize that herpesviruses evolved to partially mimic EGA by actively inducing DUX4 expression. EGA is conserved in all animals, and due to its significance in embryonic development there is very little room for the host to antagonize this viral mimicry of EGA<sup>13</sup>. Any interference with DUX4 function, for example by mutations in the coding sequence or by preventing DUX4 expression could lead to drastic changes in the embryonic development that are incompatible with life. Thus, from a viral perspective, it is beneficial to exploit a host gene which is essential for development for its own purpose in order to limit mutations that affect viral replication.

In addition to its role in EGA and in the development of FSHD, it was recently published that DUX4 also plays an important role in a variety of human cancers<sup>45</sup>. Reanalysis of almost 10000 cancer transcriptomes from The Cancer Genome Atlas (TCGA) showed DUX4 re-expression in many human cancers<sup>45</sup>. The authors speculate that expression of DUX4 and DUX4 target genes may contribute to tumorigenesis<sup>45</sup>. Interestingly, the human gamma-herpesviruses, Epstein-Barr Virus (EBV) and KSHV are classified as human carcinogens by the WHO and cause cancer in humans<sup>6</sup>. It would be very informative to see but beyond the scope of this manuscript whether DUX4 expression induced by EBV and KSHV also contributes to herpesviral oncogenesis. In addition the authors showed that DUX4 expression results in a downregulation of MHC-Class I expression, hinting at a possible immune evasion mechanism<sup>45</sup>. Although herpesviruses of all subfamilies already encode for

several immune evasion genes that interfere with presentation of viral peptides by MHC-I, the downregulation of MHC-I by DUX4 could also contribute to herpesviral immune-evasion that prevents T-cell killing of infected cells *in vivo*.

We could show that herpesviral DUX4 expression also leads to expression of several endogenous retroelements, including LTR-containing retrotransposons, LINE-1 elements and Alu-elements. Whereas cell type specific differences in the induction of retroelements exist, with more retroelements being transcribed in tumor cell lines than in primary cells, we identified an Alu-element 5' of the ZSCAN4 gene that serves as a binding site for DUX4 and drives ZSCAN4 expression in all cell lines tested. In addition it is known for years that herpesviral infection leads to the induction of endogenous HERV-W retroviruses and also to the reactivation of the HIV-LTR<sup>46,47</sup>, hinting at a role of herpesviral DUX4 in respective processes. It is hypothesized for years that some herpesviruses including HSV-1 have evolved a high GC-content in their genome in order to prevent insertion of endogenous retroelements that have a bias for AT-rich sequences. The active induction of DUX4 by HSV-1 proteins and the importance of DUX4 for herpesviral gene expression / replication supports this possible explanation for the high GC-content of HSV-1 genomes. It may help to preserve herpesviral genome integrity despite the DUX4-mediated induction of endogenous retroelements that could integrate into the viral genome.

Our data points at a very important if not essential role for DUX4 in herpesviral gene expression and replication. We could demonstrate that depletion of DUX4 in cells almost completely abrogates viral immediate early gene expression. As such it is tempting to speculate that DUX4 and DUX4 downstream genes could be targeted for therapy of herpesvirus-associated diseases. Since aberrant DUX4 expression is the cause of FSHD, several drugs targeting DUX4 are currently under development using a variety of different strategies. Blocking DUX4 means preventing herpesviral immediate early gene expression and subsequent viral replication from the beginning. Targeting a cellular protein has the major advantage that viral escape mutants and resistance-formation are far more unlikely. Therefore, DUX4 provides an attractive target for anti-herpesviral therapy.

## **Acknowledgments:**

We are grateful to David M. Knipe (Harvard Medical School, Boston, USA), Benedikt Kauffer (Freie Universität Berlin, Berlin, Germany), Alexander Steinkasserer (University Hospital Erlangen, Germany), Roger Everett (University of Glasgow, Scotland), Adam Whisnant and Lars Dölken (University of Würzburg, Germany) for providing plasmids and virus stocks. We also thank Dr. Benjamin Schmid and Dr. Philipp Tripal the Optical Imaging Center Erlangen (OICE) from the Friedrich-Alexander University Erlangen for support with microscopy. This study was supported by grants from the German Research Foundation (DFG En 423/5-1, to AE), the IZKF Erlangen (J57, to FF), the BMBF Junior Research Group “Duxdrugs” (01KI2017, to FF) and the NIH grant GM127527 (S.T.).

## Figure legends

**Figure 1:** Expression of DUX4 and DUX4 target genes upon infection with  $\alpha$ -,  $\beta$ -,  $\gamma$ -herpesviruses **A.** Western blot of primary HFF cells infected with HSV-1 for 24 h and 293T cells infected with HSV-1 for 18 h, both MOI 5. ICP0 was used as marker for infection. **B.** Western blot of HFF cells infected with HSV-1 for 24 h with MOI 5. **C.** Western blot of 293T cells infected with HSV-1 for 18 h with MOI 5. **D,E.** Primary HFF cells infected with HCMV for 6d with MOI 1. Western blot analysis of DUX4 (D) and ZSCAN4 (E). HCMV gB was used as marker for infection. **F-H.** qPCR analysis of DUX4 expression upon infection of 293T cells with HSV-1 MOI 5 for 18 h (**F**), upon expression of HFF cells with HCMV MOI 1 for 6d (**G**) and upon lytic reaction of KSHV in iSLK cells at 5d p.i. (**H**) **I.** mRNA expression of DUX4 target genes in 293T upon infection with HSV-1 for 18 h with MOI 10. **J.** mRNA expression of DUX4 target genes in HFF infected with HCMV for 6 d with MOI 1. **K.** mRNA expression of DUX4 target genes in iSLK cells with reactivated KSHV infection, harvested after 5 d. In all qPCR experiments mRNA of targets was normalized to HPRT1 mRNA expression, one representative out of at least n=3.

**Figure 2: A.** RNA-Seq. data of the DUX4 and neighbouring FRG1 and FRG2 loci in WI38 9 hpi, HFF 4 hpi and HEK 293T cells 18 hpi. The Precision Run-On Sequencing (Pro-Seq.) data show RNA POL II occupancy on the FRG1, FRG2 and DUX4 locus in infected and uninfected cells. ICP4 ChIP-Seq. data show ICP4 occupation at the DUX4 locus in cells infected with HSV-1. **B.** Western Blot of DUX4 and HSV-1 proteins in primary HFF infected with HSV-1 harvested at different hpi (MOI 10). **C.** Western Blot of DUX4 and HSV-1 proteins in primary HFF infected with HSV-1 and HSV-1 mutants harvested at 16 hpi (MOI 10). One representative experiment out of n=5. **D.** qRT-PCR of DUX4 expression relative to HPRT from primary HFF cells infected with HSV-1, HSV-1  $\Delta$ ICP0 (ICP4-YFP), HSV-1  $\Delta$ ICP34.5, HSV-1  $\Delta$ ICP4 for 24 h (MOI 10). **E.** Western blot analysis of DUX4 and HSV-1 protein in 293T cells transfected with HSV-1 IE proteins ICP0, ICP0 FXE, ICP4 (ICP4-YFP), ICP0 + ICP4 (ICP4-YFP) and ICP0 FXE + ICP4 (ICP4-YFP) for 48 h or infected with HSV-1 for 18 h (MOI 10). ICP0 FXE is a mutant with a deletion in the RING domain, which inhibits Ubiquitin E3 ligase activity. EV: empty vector control. One representative experiment out of n=3.

**Figure 3: A.** Immunofluorescence of HDF-TERT cells infected with a recombinant HSV-1 expressing ICP4-YFP and VP26-RFP. DUX4 protein was stained with a DUX4 specific antibody, nuclei were counterstained with DAPI. Nuclear DUX4 expression is only detectable in ICP4+/VP26- cells. **B.** qRT-PCR analysis of cellular genes DUX4, TRIM43, PRAMEF2 and HPRT as well as viral genes gC, ICP0, ICP8 and UL36 in HDF-TERT cells untreated or treated with PAA and infected with HSV-1 (MOI 0.1). Values are presented as fold induction (normalized to HPRT RNA) relative to uninfected control cells.

**Figure 4:** ChIP-Seq analysis of endogenous DUX4 induced by HSV-1 infection **A.-E.** ChIP-Seq of endogenous DUX4 in HFF cells infected with HSV-1 at 8 hpi and 14 hpi. **A.** Comparison of DUX4 consensus binding sequence in HSV-1 infected cells (upper panel) with the previously, published DUX4 consensus sequences (Geng et al.) in the lower panel. **B.** 8 hpi ChIP samples plotted against 14 hpi samples. Both are highly concordant. DUX4 binding increases at 14 hpi. **C.** Number of binding sites in different genomic regions. (TTS: transcription termination site, TSS: transcription start site) **D.** Different expression levels of DUX4 target genes during HSV-1 infection plotted over time, analysed from Rutkowski et al. 2015 **E.** Analysis of DUX4 binding sites that can be found within repetitive elements. (LC: low complexity repeats, SR: simple repeat) **F.** Repetitive elements regulated in DUX4 overexpressing cells compared to HSV-1 infected cells reanalysed from Full et al. 2019. On the right site repetitive elements are ascribed to their respective class and family. DUX4 binding is scaled to repeat size, number of peaks and activation. **G.** Expression pattern of embryonic genes during embryonic genome activation compared to expression pattern in DUX4 overexpressing and

HSV-1 infected cells. Embryonic data was published in (reference) and is normalized to average expression value of each gene. Data from DUX4 overexpression and HSV-1 infection is relative to mock. Shown are genes which have DUX4 binding sites in the proximity of the TSS. (oo: oocyte, pro: pronucleus, zy: zygote, 2c: 2-cell state, 4c: 4-cell state, 8c: 8-cell state, mo: morula)

**Figure 5: DUX4 expression is critical for herpesviral immediate early protein expression** **A.** DUX4 kd cells were generated with sgRNAs #1-4. HSV-1 protein expression in 293T WT cells and 293T cells with DUX4 KD. Cells were infected with HSV-1 (MOI 5) and analysed by western blot. Time points as indicated in graph. Representative experiment out of n=3. **B.** mRNA expression (RNA-Seq.) of DUX4 kd cells plotted against wildtype 293T cells. **C.** qPCR of ZSCAN4 mRNA expression in 293T WT and 293T DUX4 KD cells infected with HSV-1 (MOI 5) for 18h. **D.** RNA-Seq analysis of DUX4 CRISPR/Cas9 treated 293T cells and 293T WT cells infected with HSV-1 (MOI 5) for 18 h. The heat-map illustrates differences in the expression levels of HSV-1 RNAs in WT and KD cells. DUX4 KD cells generated with sgRNAs #1-4. mRNA expression is normalized to expression of vault RNA (VTRNA). One representative experiment out of n=3. **E.** Western blot of 293T WT cells and 293T DUX4 KO cells infected with HSV-1 for 18 h. Comparison of HSV-1 protein expression in WT cells and cells with complete DUX4 KO. Representative experiment out of n=5.



## Methods

### Cell culture and viruses

HEK 293T, primary HFF, Vero were maintained in Dulbecco's modified Eagle's medium (DMEM, Thermo Fisher Scientific) supplemented with 10% (vol/vol) heat-inactivated fetal bovine serum (FBS, Thermo Fisher Scientific), 2 mM GlutaMAX (Thermo Fisher Scientific), 1 mM HEPES (Thermo Fisher Scientific) and 1% gentamycin (vol/vol). iSLK harbouring recombinant (r) KSHV.219 were cultured in DMEM supplemented with 10% (vol/vol) FBS, 2 mM GlutaMAX (Thermo Fisher Scientific), 10 mM HEPES buffer (Thermo Fisher Scientific) and 1% penicillin-streptomycin (vol/vol), 1 µg/ml Puromycin and 250 µg/ml C418. All cells were kept under standard culture conditions and regularly checked for mycoplasma contamination (MycoAlert Kit, Lonza).

Lytic replication of KSHV in iSLK.219 was induced by adding 1 µg/ml doxycycline. Infection experiments were performed with HSV-1 strain KOS, HSV-1 GFP (F-strain, provided by B. Kauffer, Berlin) or HCMV (strain Ad-169). HSV-1 deltaICP0 (ICP4-YFP) and HSV-1 delta ICP27 were provided from R. Everett (University of Glasgow). HSV-1 T-VEC del ICP34.5/ICP47 IMLYGIC<sup>®</sup>, Talimogene laherparepvec was provided by L. Heinzerling (Friedrich-Alexander-Universität Erlangen-Nürnberg). HSV-1 expressing ICP4-YFP and VP26-RFP is a new recombinant virus generated in the background of strain 17 by recombining HSV-1 expressing ICP4-YFP (obtained from Matthew D. Weitzman) and HSV-1 expressing VP26-RFP (obtained from Oren Kobilier) and plaque purifying for 5 cycles.

### Reagents, plasmids and transfections

Transfections were performed with GenJet (SigmaGen Laboratories) or Lipofectamin 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol. Plasmids are listed in Table 3. KU60019 was purchased from Selleckchem.

### Western Blotting

Cells were lysed in RIPA HS buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM NaCl, 1% Triton X-100 (vol/vol), 0,1% SDS (vol/vol), 0,1% deoxycholic acid (DOC) with Aprotinin and Leupeptin, MG-132 and sodium metavanadate (Sigma-Aldrich). The cell pellet was centrifuged at 4°C, 14.000 rpm for 30 min. Samples were diluted with Laemmli-SDS sample buffer and heated for 5 min at 95°C. Antibodies used are listed in Table 4.

### qRT-PCR

RNA was extracted using the Direct-zol RNA Miniprep Plus kit from Zymo Research according to manufacturer's instructions. Reverse transcription was performed using the Super Script IV Kit (Thermo Fisher Scientific) according to manufacturer's instructions. qRT-PCR was carried out using TaqMan<sup>™</sup> Universal PCR Master Mix I (Applied Biosystems, Thermo Fisher Scientific) 0,1 µg as template on a 7500 Fast Real-Time PCR machine. Or reverse transcription and qRT-PCR were conducted in one step using 0.1 µg RNA as template with the Luna Universal Probe One-Step RT-qPCR kit (New England Biolabs) following manufacturer's protocols. Primers/probes are listed in Table 1. Expression levels for each gene were obtained by normalizing values to *HPRT1* or *VTRNA* and fold induction was calculated using the comparative CT method ( $\Delta\Delta CT$  method).

### CRISPR and sgRNAs

All sgRNAs used in this study were previously described<sup>48</sup> and are listed in Table 2. sgRNAs were cloned into LentiCRISPRv2 plasmid gifted from F. Zhang (Addgene plasmid #52961) and verified by sequencing. Lentiviruses were packaged with pMD2.G (Addgene plasmid #12259) and psPAX2 (Addgene plasmid #12260) (both gifted from D. Trono) into HEK 293T. HEK 293T were seeded into 12 well plates and lentiviral supernatants added at 70-80% confluency. Plates were centrifuged at 1200 rpm for 2 min after centrifugation culture medium was added and cells incubated over night at 37°C. The medium was changed to normal culture medium the next day and selection with 2 µg/ml puromycin in normal culture medium started on day 3.

### **Next generation Chip-Seq**

For CHIPmentation primary HFF cells were seeded in T175 flasks and infected with HSV-1 KOS (MOI 10) for different time points. CHIPmentation was conducted as previously described (Schmidl et al. 2015). Cells were sonicated using the Bioruptor (Diagenode) for 30 cycles. For the immunoprecipitation protein G Dynabeads (Thermo Fisher Scientific) were used. Samples were incubated with either 2.5 µg Anti-DUX4 (E5-5) (Abcam) or Normal Rabbit IgG (Cell Signaling Technology) as control. Samples were purified using AMPureXP beads (Beckman Coulter) according to manufacturer's description.

### **Next generation RNA-Seq**

HEK 293T wildtype cells and HEK 293T CRISPR/Cas knockout cells were seeded in T25 flasks. Cells infected for different time points with HSV-1 KOS (MOI 10). Cells were lysed in Trizol (Life Technologies by Thermo Fisher Scientific), and total RNA was isolated using the RNA clean and concentrator kit (Zymo Research), according to the manufacturer's instructions. Sequencing libraries were prepared using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB) with 9 cycles PCR amplification, and sequenced on a HiSeq 4000 device with 1x50 cycles. For quantification of viral gene expression alignments were done using hisat2<sup>49</sup> on the HSV-1 genome (strain 17, genbank accession no. NC\_001806) and readcounts per gene quantified using quasR<sup>51</sup>.

### **Bioinformatic analysis of ChIP-seq data**

ChIP-seq data processing was done using the PiGx-ChIP-seq pipeline (<https://doi.org/10.1093/gigascience/giy123>). In short, adapters and low quality bases were trimmed from reads using Trim-galore. The reads were mapped on the hg19 version of the human genome, combined with HSV-1 genome, using Bowtie2 (22388286) with k = 1 parameter. bigWig tracks were created by extending reads to 200, collapsing them into pileups, and normalizing to reads per million. Peak calling was done with MACS2 (<https://github.com/taoliu/MACS>) using the default parameters. Motif discovery was done using MEME (25953851) with the default parameters, on the top 100 peaks (sorted by q value), in a region of +/- 50bp around the peak center. Peak annotation was done using the hg 19 ENSEMBL GTF file, downloaded on 17.03.2017. from the ENSEMBL database (29155950). Peaks were annotated based on the following hierarchy of functional categories: tss -> exon -> intron -> intergenic (eg. if a peak overlapped multiple categories, it was annotated by the class that is highest in the hierarchy). Peaks were overlapped with the hg19 Repeatmasker repeat annotation, downloaded from the UCSC database (30407534) on 03.02.2015.

## Transposon Quantification for KD samples

Transposon expression was quantified using TeTranscripts (26206304) version 2.2.0, using the hg19 version of the human genome and a custom formatted transposon GTF file, made available by the Hammell laboratory from the following link: [http://labshare.cshl.edu/shares/mhammelllab/www-data/TEtranscripts/TE\\_GTF/](http://labshare.cshl.edu/shares/mhammelllab/www-data/TEtranscripts/TE_GTF/). Before visualization, the read counts per TF repName repeatmasker category were multiplied by the following size factors 10000/(number of repeats), and 1000/(average repeat length). The transposon expression was visualized relative to the uninfected wild type sample, using the ComplexHeatmap package (27207943).

## Comparison of HSV-1 gene expression with the embryonic expression profile

Expression profiles for HSV-1 and DUX4 ectopically expressed, Hela cells were taken from the following publication (30420784). The embryonic expression profiles were downloaded from the ARCHS4 database (29636450). The data originated from the following repository GSE44183. Data was visualized using the ComplexHeatmap function.

## Comparison of repeat expression during HSV-1 infection and DUX4 overexpression.

Expression profiles obtained from HSV-1 and DUX4 ectopically expressed, Hela cells were taken from the following publication (30420784). DUX4 binding data extracted from the supplementary data from the following publication 24278031. RNA - seq data was mapped using STAR, and the repeats were quantified by counting the number of, uniquely mapping, spliced reads overlapping with the each transposon category. Expression was visualized using ComplexHeatmap. Prior to the visualization the reads were normalized to uninfected samples.

## Statistical analysis

*P* values were calculated using an unpaired Student's test. *P* <0.05 was considered statistically significant.

Table 1 Oligonucleotides

Real time quantitative PCR probes		
DUX4L: 5'-TCAGCCAGAATTCACGGAAG-3'	FAM/ZEN/3IABkFQ	IDT
5'-CCATTCTTCTGGGCATCC-3'		
ZSCAN4: 5'-CTAGTCACACATCAGCTCAGT-3'	FAM/ZEN/3IABkFQ	IDT
5'-TTCAGTCTCTTGCCTTGTGTC-3'		
ZSCAN5D:	FAM/ZEN/3IABkFQ	IDT
TRIM49B: 5'-GCTGCTGAGGAACACCA-3'	FAM/ZEN/3IABkFQ	IDT
5'-ATTGCTTCTAGCCTTAAATTCACAT-3'		
TRIM48: 5'-ATCACTGGACTGAGGGACA-3'	FAM/ZEN/3IABkFQ	IDT
5'-GGGCGGATTTTGACGGT-3'		

<b>TRIM53</b>	FAM/ZEN/3IABkFQ	IDT
<b>RFPL4A</b>	FAM/ZEN/3IABkFQ	IDT
<b>ICP0: 5'-CGGACACGGAAGTTCGA-3'</b>  <b>5'-CGCCCCGCAACTGC-3'</b>	FAM/ZEN/3IABkFQ	IDT
<b>HPRT1: 5'-CGAGATGTCATGAAGGAGATGG-3'</b>  <b>5'-TCAGCAAAGAACTTATAGCCCC-3'</b>	HEX/ZEN/3IABkFQ	IDT
<b>VTRNA1-1: 5'-TTTAATTGAAACAAGCAACCTGTCT-3'</b>	FAM/NFQ	Thermo Scientific Fisher

Table 2 gRNAs

**Small guide RNAs for CRISPR/Cas**

---

**sgRNA DUX4 #1 (E1-3; v1)**

**5'-CACCGCACCCGGGCAAAGCCGGG-3'**

**5'-AAACCCGGCTTTTGCCCGGGTGC-3'**

**sgRNA DUX4 #2 (E1-4):**

**5'-CACCGCTGGAAGCACCCCTCAGCG-3'**

**5'-AAACCGCTGAGGGGTGCTTCCAGC-3'**

**sgRNA DUX4 #3 (p8)**

**5'-CACCGTCGGACAGCACCCCTCCCCG-3'**

**5'-AAACCGGGGAGGGTGCTGTCCGAC-3'**

---

Table 3 Plasmids

**Plasmids**

---

<b>pcDNA 3.1 FLAG</b>	Invitrogen/Thermo Fisher Scientific
<b>pcDNA 3.1 ICP27</b>	gift from A. Steinkasserer (Erlangen)
<b>pCS2-mkgDUX4</b>	gift from S. Tapscott (Addgene plasmid #21156)
<b>pcW57.1-DUX4-CA</b>	gift from S. Tapscott (Addgene plasmid #99283)

<b>pIE-EYFP-ICP4</b>	gift from R. Everett (Glasgow)
<b>pUC18-ICP0</b>	gift from D. Knipe (Boston)

Table 4 Antibodies

<b>Primary Antibodies</b>			
Anti-DUX4 (E5-5)	rabbit polyclonal	Abcam	Ab124699
Anti-DUX4 (clone 9A12)	mouse monoclonal	EMD Millipore	MABD116
VP16 (14-5)	mouse monoclonal	Santa Cruz Biotechnology Inc	Sc-7546
HSV-1 ICP0 antibody (11060)	mouse monoclonal	Santa Cruz Biotechnology Inc	Sc-53070
HSV-1/2 ICP27 antibody	mouse monoclonal	Santa Cruz Biotechnology Inc	Sc-69806
HSV-1 gD antibody	mouse monoclonal	Santa Cruz Biotechnology Inc	Sc-21719
HSV-1 ICP4 (H943) antibody	mouse monoclonal	Santa Cruz Biotechnology Inc	Sc-69809
KAP1 antibody	rabbit polyclonal	Bethyl Laboratories Inc	A300-274A
Purified Anti-Tif1 $\beta$ Phospho (S473) (Poly6446)	rabbit polyclonal	Biolegend	644602
TIF1 $\beta$ (H-300)	rabbit polyclonal	Santa Cruz Biotechnology Inc	Sc-33186
Purified Anti-ATM Phospho (Ser1981) antibody	rabbit polyclonal	Biolegend	10H11.E12
Phospho KAP1 (S824) antibody	rabbit polyclonal	Bethyl Laboratories Inc	A300-767A
Anti-HA.11	mouse monoclonal	Covance	MMS-101R
THE GFP antibody	rabbit polyclonal	Genescript	A01704-40
Anti-HCMV gB (27-287)	mouse monoclonal	Hybridoma	
Anti-ZSCAN4	rabbit polyclonal	Sigma-Aldrich/Merck Millipore	HPA006491

THE $\beta$ -Actin antibody, HRP	mouse monoclonal	Genescript	A00730-40
GAPDH antibody, Biotin	goat	Genescript	A00915
B-Tubulin antibody	mouse monoclonal	Genescript	A01717-40
THE $\alpha$ -Tubulin antibody	mouse monoclonal	Genescript	A01410-40
Normal Rabbit IgG	Isotype control	Cell Signaling Technology	#2729
HSP 90 (4F10)	mouse monoclonal	Santa Cruz Biotechnology Inc	Sc-69703
<b>Secondary Antibodies</b>			
Goat anti-mouse HRP	SouthernBiotech		1036-05
Swine anti-rabbit Immunoglobulins/HRP	Dako/Agilent		P039901-2
DyLight649 Streptavidin	Jackson		016-490-084
Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	Invitrogen/Thermo Fisher Scientific		A31571
Goat anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	Invitrogen/Thermo Fisher Scientific		A21245
Donkey anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 555	Invitrogen/Thermo Fisher Scientific		A31570

## Reference

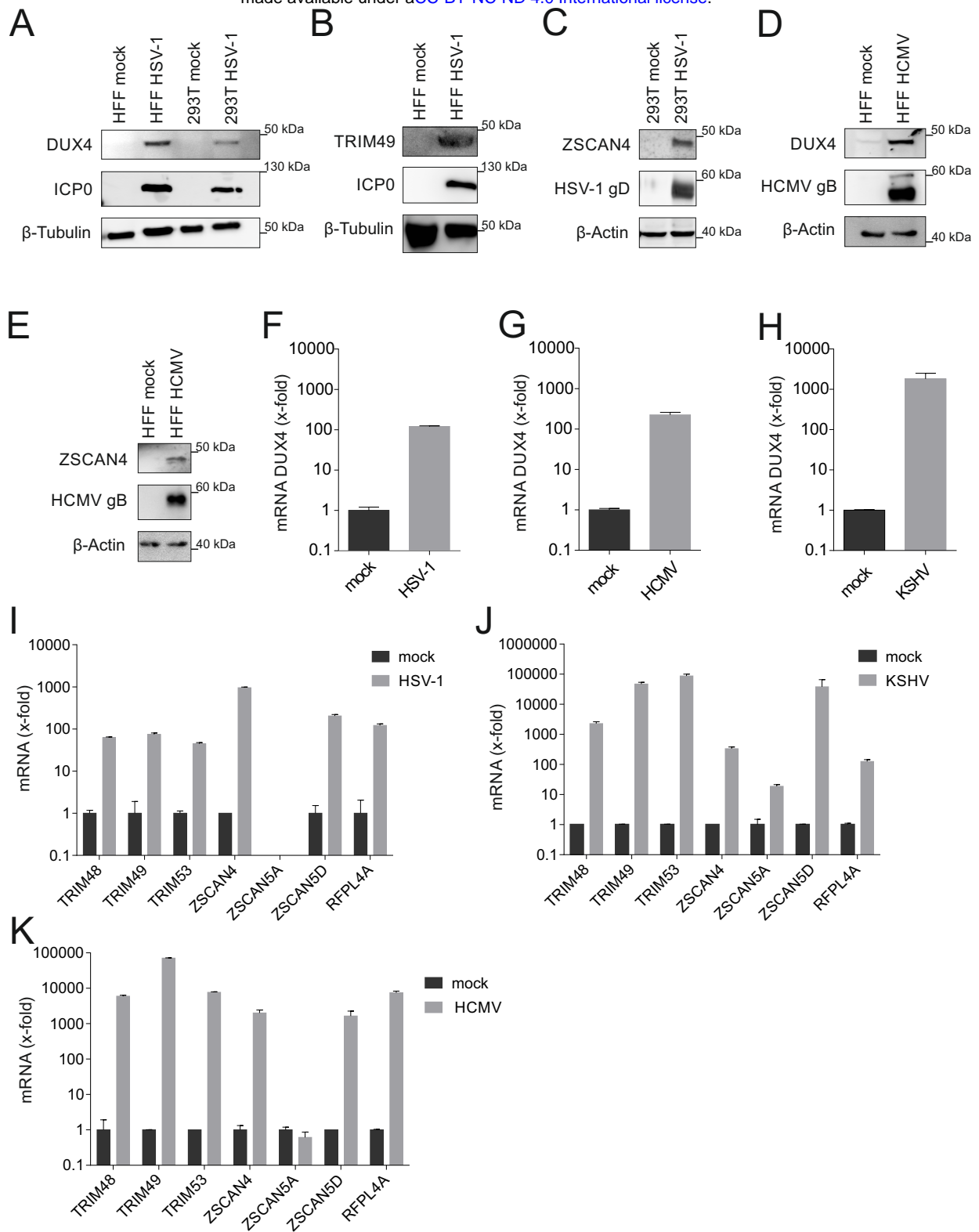
1. Pebody, R. G. *et al.* The seroepidemiology of herpes simplex virus type 1 and 2 in Europe. *Sex. Transm. Infect.* **80**, 185–191 (2004).
2. Zuhair, M. *et al.* Estimation of the worldwide seroprevalence of cytomegalovirus: A systematic review and meta-analysis. *Rev. Med. Virol.* **29**, e2034 (2019).
3. Bradley, H., Markowitz, L. E., Gibson, T. & McQuillan, G. M. Seroprevalence of herpes simplex virus types 1 and 2--United States, 1999-2010. *J. Infect. Dis.* **209**, 325–333 (2014).
4. Corey, L. & Wald, A. Maternal and neonatal herpes simplex virus infections. *N. Engl. J. Med.* **361**, 1376–1385 (2009).
5. Stern, L. *et al.* Human Cytomegalovirus Latency and Reactivation in Allogeneic Hematopoietic Stem Cell Transplant Recipients. *Front. Microbiol.* **10**, 1186 (2019).
6. Plummer, M. *et al.* Global burden of cancers attributable to infections in 2012: a synthetic analysis. *Lancet. Glob. Heal.* **4**, e609-16 (2016).
7. Full, F. *et al.* Centrosomal protein TRIM43 restricts herpesvirus infection by regulating nuclear lamina integrity. *Nat. Microbiol.* **4**, 164–176 (2019).
8. Friedel C.C. *et al.* Dissecting Herpes Simplex Virus 1-Induced Host Shutoff at the RNA Level. *J. Virol.* **95**, e01399-20 (2021).
9. Hendrickson, P. G. *et al.* Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).
10. De Iaco, A. *et al.* DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* **49**, 941–945 (2017).
11. Whiddon, J. L., Langford, A. T., Wong, C. J., Zhong, J. W. & Tapscott, S. J. Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* **49**, 935–940 (2017).
12. Hendrickson, P. G. *et al.* Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).
13. Lee, M. T., Bonneau, A. R. & Giraldez, A. J. Zygotic Genome Activation During the Maternal-to-Zygotic Transition. *Annu. Rev. Cell Dev. Biol.* **30**, 581–613 (2014).
14. Eckersley-Maslin, M. A., Alda-Catalinas, C. & Reik, W. Dynamics of the epigenetic landscape during the maternal-to-zygotic transition. *Nat. Rev. Mol. Cell Biol.* **19**, 436–450 (2018).
15. Whiddon, J. L., Langford, A. T., Wong, C. J., Zhong, J. W. & Tapscott, S. J. Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* **49**, 935–940 (2017).
16. Macfarlan, T. S. *et al.* ES cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
17. Wang, J. *et al.* Primate-specific endogenous retrovirus-driven transcription defines naive-like stem cells. *Nature* **516**, 405–409 (2014).
18. Young, J. M. *et al.* DUX4 Binding to Retroelements Creates Promoters That Are Active in FSHD Muscle and Testis. *PLoS Genet.* **9**, (2013).
19. Snider, L. *et al.* Facioscapulohumeral dystrophy: Incomplete suppression of a



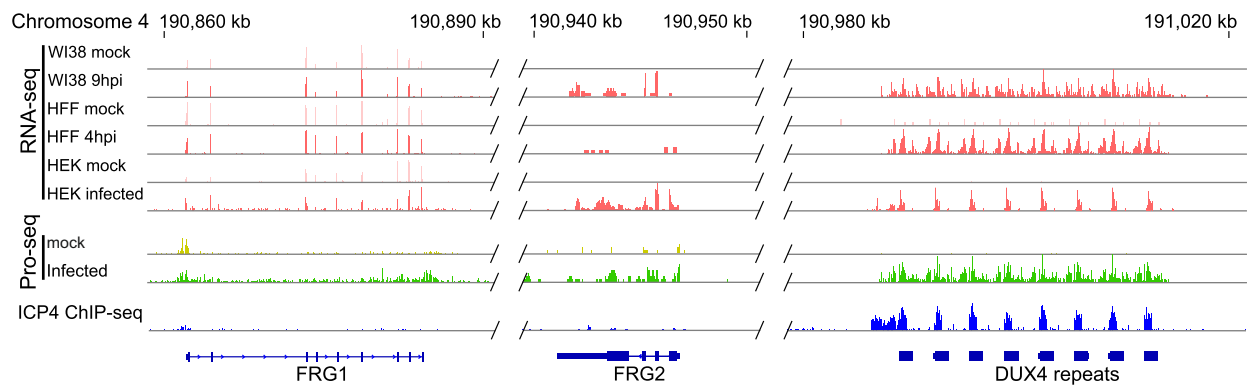
- retrotransposed gene. *PLoS Genet.* **6**, 1–14 (2010).
20. Dixit, M. *et al.* DUX4, a candidate gene of facioscapulohumeral muscular dystrophy, encodes a transcriptional activator of PITX1. *Proc. Natl. Acad. Sci.* **104**, 18157–18162 (2007).
  21. Lemmers, R. J. L. F. *et al.* A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science (80-. )*. **329**, 1650–1653 (2010).
  22. Himeda, C. L. & Jones, P. L. The Genetics and Epigenetics of Facioscapulohumeral Muscular Dystrophy. *Annu. Rev. Genomics Hum. Genet.* **20**, 265–291 (2019).
  23. Kowalijow, V. *et al.* The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul. Disord.* **17**, 611–623 (2007).
  24. Campbell, A. E., Belleville, A. E., Resnick, R., Shadle, S. C. & Tapscott, S. J. Facioscapulohumeral dystrophy: activating an early embryonic transcriptional program in human skeletal muscle. *Hum. Mol. Genet.* **27**, R153–R162 (2018).
  25. Hewitt, J. E. *et al.* Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophthly. *Hum. Mol. Genet.* **3**, 1287–1295 (1994).
  26. Gabriëls, J. *et al.* Nucleotide sequence of the partially deleted D4Z4 locus in a patient with FSHD identifies a putative gene within each 3.3 kb element. *Gene* **236**, 25–32 (1999).
  27. Deutekom, J. C. T. va. *et al.* FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.* **2**, 2037–2042 (1993).
  28. Zeng, W. *et al.* Specific loss of histone H3 lysine 9 trimethylation and HP1 $\gamma$ /cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet.* **5**, (2009).
  29. Yao, Z. *et al.* DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum. Mol. Genet.* **23**, 5342–5352 (2014).
  30. Van Den Heuvel, A. *et al.* Single-cell RNA sequencing in facioscapulohumeral muscular dystrophy disease etiology and development. *Hum. Mol. Genet.* **28**, 1064–1075 (2019).
  31. Wyler, E. *et al.* Widespread activation of antisense transcription of the host genome during herpes simplex virus 1 infection. *Genome Biol.* **18**, 1–19 (2017).
  32. Rutkowski, A. J. *et al.* Widespread disruption of host transcription termination in HSV-1 infection. *Nat. Commun.* **6**, (2015).
  33. Shadle, S. C. *et al.* DUX4-induced bidirectional HSATII satellite repeat transcripts form intranuclear double-stranded RNA foci in human cell models of FSHD. *Hum. Mol. Genet.* **00**, 1–15 (2019).
  34. Nogalski, M. T. *et al.* A tumor-specific endogenous repetitive element is induced by herpesviruses. *Nat. Commun.* **10**, (2019).
  35. Birkenheuer, C. H., Danko, C. G. & Baines, J. D. Herpes Simplex Virus 1 Dramatically Alters Loading and Positioning of RNA Polymerase II on Host Genes Early in Infection. *J. Virol.* **92**, e02184-17 (2018).
  36. Fox, H. L., Dembowski, J. A. & DeLuca, N. A. A herpesviral immediate early protein promotes transcription elongation of viral transcripts. *MBio* **8**, 1–16 (2017).
  37. Dixon, R. A. & Schaffer, P. A. Fine-structure mapping and functional analysis of temperature-sensitive mutants in the gene encoding the herpes simplex virus type 1



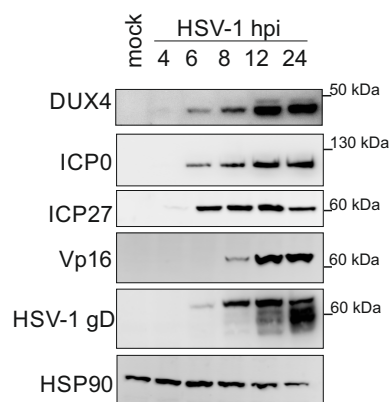
- immediate early protein VP175. *J. Virol.* **36**, 189–203 (1980).
38. Everett, R. D. Trans activation of transcription by herpes virus products: requirement for two HSV-1 immediate-early polypeptides for maximum activity. *EMBO J.* **3**, 3135–3141 (1984).
  39. Geng, L. N. *et al.* DUX4 Activates Germline Genes, Retroelements, and Immune Mediators: Implications for Facioscapulohumeral Dystrophy. *Dev. Cell* **22**, 38–51 (2012).
  40. Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500**, 593–597 (2013).
  41. Bakker, E. *et al.* The FSHD-linked locus D4F104S1 (p13E-11) ON 4q35 has a homologue on 10qter. *Muscle Nerve* **18**, S39–S44 (1995).
  42. Cacurri, S. *et al.* Sequence homology between 4qter and 10qter loci facilitates the instability of subtelomeric KpnI repeat units implicated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.* **63**, 181–190 (1998).
  43. Messerschmidt, D. M., Knowles, B. B. & Solter, D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes Dev.* **28**, 812–828 (2014).
  44. McGeoch, D. J., Cook, S., Dolan, A., Jamieson, F. E. & Telford, E. A. R. Molecular phylogeny and evolutionary timescale for the family of mammalian herpesviruses. *J. Mol. Biol.* **247**, 443–458 (1995).
  45. Chew, G. L. *et al.* DUX4 Suppresses MHC Class I to Promote Cancer Immune Evasion and Resistance to Checkpoint Blockade. *Dev. Cell* **50**, 658–671.e7 (2019).
  46. Perron, H. *et al.* Herpes simplex virus ICP0 and ICP4 immediate early proteins strongly enhance expression of a retrovirus harboured by a leptomeningeal cell line from a patient with multiple sclerosis. *J. Gen. Virol.* **74**, 65–72 (1993).
  47. Mosca, J. D. *et al.* Activation of human immunodeficiency virus by herpesvirus infection: Identification of a region within the long terminal repeat that responds to a trans-acting factor encoded by herpes simplex virus 1. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7408–7412 (1987).
  48. Himeda, C. L., Jones, T. I. & Jones, P. L. CRISPR/dCas9-mediated transcriptional inhibition ameliorates the epigenetic dysregulation at D4Z4 and represses DUX4-fl in FSH muscular dystrophy. *Mol. Ther.* **24**, 527–535 (2016).
  49. Smiley JR, Johnson DC, Pizer LI, Everett RD. The ICP4 binding sites in the herpes simplex virus type 1 glycoprotein D (gD) promoter are not essential for efficient gD transcription during virus infection. *J Virol.* **66**, 623–31 (1992).
  50. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
  51. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M. B. QuasR: quantification and annotation of short reads in R. *Bioinformatics* **31**, 1130–1132 (2015).



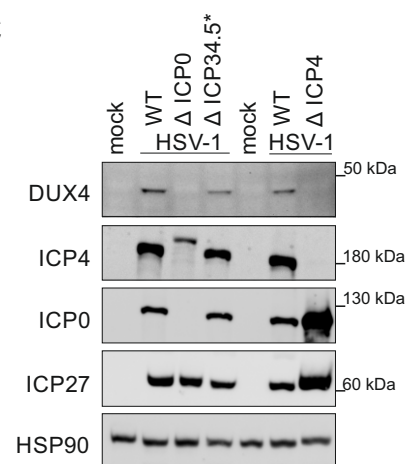
**A**



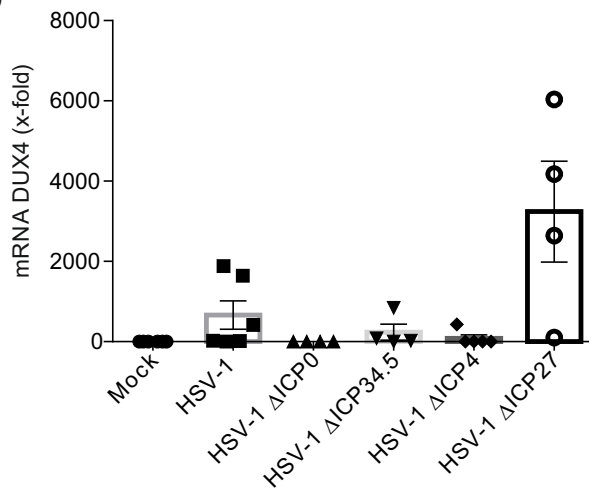
**B**



**C**



**D**



**E**

