OXFORD

Human Molecular Genetics, 2023, Vol. 32, 11, 1864–1874
https://doi.org/10.1093/hmg/ddad021
Advance access publication date 2 February 2023
Original Article

Human DUX4 and porcine DUXC activate similar early embryonic programs in pig muscle cells: implications for preclinical models of FSHD

Yee Nip[1,†], Sean R. Bennett[1,†], Andrew A. Smith[1], Takako I. Jones[2], Peter L. Jones[2,*] and Stephen J. Tapscott[1,3,*]

[1]Division of Human Biology, Fred Hutchinson Cancer Center, Seattle, WA 98109, USA
[2]Department of Pharmacology, University of Nevada, Reno School of Medicine, Reno, NV 89557, USA
[3]Department of Neurology, University of Washington School of Medicine, Seattle, WA 98105, USA
*To whom correspondence should be addressed at: Stephen J. Tapscott, Fred Hutchinson Cancer Center, 1100 Fairview Ave N, Seattle, WA 98109, USA. Tel: 1-206-667-4499; Email: stapscot@fredhutch.org; Peter L. Jones, 1664 N Virginia St. MS-0318, Reno, NV 89557, USA. Tel: 1-775-784-1566; Email: peterjones@med.unv.edu
†Authors contributed equally

Abstract

Human DUX4 and its mouse ortholog Dux are normally expressed in the early embryo—the 4-cell or 2-cell cleavage stage embryo, respectively—and activate a portion of the first wave of zygotic gene expression. DUX4 is epigenetically suppressed in nearly all somatic tissue, whereas facioscapulohumeral dystrophy (FSHD)-causing mutations result in its aberrant expression in skeletal muscle, transcriptional activation of the early embryonic program and subsequent muscle pathology. Although DUX4 and Dux both activate an early totipotent transcriptional program, divergence of their DNA binding domains limits the use of DUX4 expressed in mice as a preclinical model for FSHD. In this study, we identify the porcine DUXC messenger ribonucleic acid expressed in early development and show that both pig DUXC and human DUX4 robustly activate a highly similar early embryonic program in pig muscle cells. These results support further investigation of pig preclinical models for FSHD.

Introduction

Facioscapulohumeral dystrophy (FSHD) is caused by the misexpression of the DUX4 transcription factor in skeletal muscle (1–3). DUX4 is normally expressed in the early human embryo at the 4-cell stage and induces part of the first wave of zygotic gene expression, including genes associated with totipotency and pluripotency (4–6). Aberrant expression of DUX4 in FSHD skeletal muscle induces transcription of this early embryonic program and leads to muscle pathology (7–10). These and other advances in understanding the pathophysiology of FSHD have led to multiple therapeutic approaches and the need for appropriate preclinical models (11,12).

DUX4 is a primate-specific retrogene generated from the retrotransposition of the parental DUXC gene early in the primate lineage, whereas the mouse ortholog of human DUX4 (hDUX4), mouse Dux (mDux), is a retrogene thought to have arisen through an independent retrotransposition of the DUXC messenger ribonucleic acid (mRNA) in the rodent lineage (13,14). Several mouse models expressing hDUX4 have been generated and each has useful characteristics regarding preclinical models for therapies that target hDUX4 and/or its transcriptional activity (15–21). Although hDUX4 and mDux induce expression of a similar early embryonic totipotent program when expressed in human and mouse cells, respectively, expression of hDUX4 in mouse cells does not robustly induce this program, in part due to a divergence of the DNA binding regions and binding site sequences between mouse and human (5), thereby limiting the utility of expressing hDUX4 in mouse muscle as a model of disease progression in FSHD. While xenotransplants of human muscle into mice (22–25) partly circumvent this concern, these models have limitations for disease progression as well.

Recently, we showed that hDUX4 and canine DUXC (cDUXC) show higher conservation in their homeodomains compared to hDUX4 and mDux (26). Consistent with this, hDUX4 expressed in canine cells robustly induced its early embryonic transcriptional program, suggesting that dogs, and possibly other Laurasiatherians, might have some advantages over mice for preclinical models measuring the transcriptional program and consequences of hDUX4 in skeletal muscle. Among the Laurasiatherians, pigs have substantial advantages as biomedical models for human disease and preclinical therapeutic studies (27,28). Pig models of other neuromuscular diseases have been used for studying the pathophysiology and for evaluating therapies (29–32), including AAV-mediated CRISPR-Cas9 gene editing in a model for Duchenne muscular dystrophy (DMD) (33).

In the current study, we used a previously published RNA sequencing dataset (34) to identify a porcine DUXC (pDUXC) mRNA expressed in pre-implantation embryos at the time of zygotic gene activation. The pDUXC protein encoded by this mRNA maintains higher conservation of the homeodomains with hDUX4 than mDux and maintains conservation of the DUXC-family C-terminal domain. Expression of pDUXC in pig skeletal muscle cells robustly induces genes that are enriched for the early 2-cell/4-cell-like totipotent gene signature, indicating that pDUXC likely has a similar early embryonic role as hDUX4 and mDux. Importantly, expression of hDUX4 in pig muscle cells robustly activates a similar transcriptional program, and motif inference indicates that hDUX4 and pDUXC maintain similar DNA binding

Received: August 29, 2022. **Revised:** January 10, 2023. **Accepted:** January 29, 2023

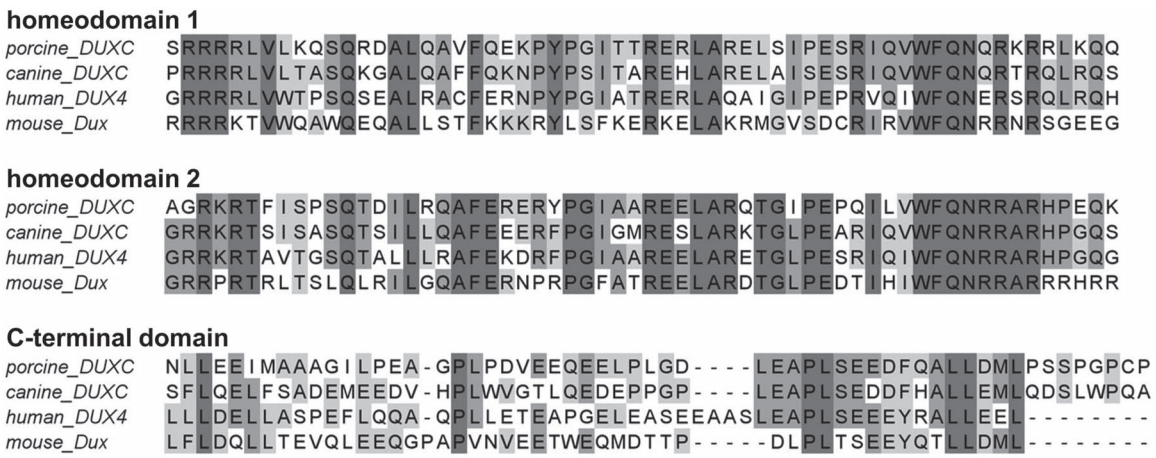

Figure 1. Conservation among DUX family members. The protein sequences of the two homeodomains and C-terminal domain of DUXC family members were aligned with ClustalWS Multiple Sequence Alignment.

motifs. Together, these findings support consideration of the pig as a candidate preclinical model for FSHD.

Results

Identification of pDUXC gene

The current *Sus scrofa* reference genome (Sscrofa11.1) lacks annotation of a DUXC gene (35). We identified an ~300 kb region in the subtelomeric region of chromosome 17 (chr17:63,201,754-63,485 ,686) with multiple areas of homology to a DUXC-like sequence indicating the presence of multiple repeats of a DUXC-like gene. Several predicted transcripts in this region encoded a DUXC-like protein that were identical with the exception of alternative first exons and an alternative splice acceptor site in the predicted exon 4 (Supplementary Material, Fig. S1A).

We mapped RNA sequencing reads from pre-implantation pig embryos (34) to this region and identified expression of a DUXC mRNA between the 2- and 8-cell stages that contained a novel first exon and the same additional four exons of the predicted DUXC-like genes (Supplementary Material, Fig. S1B), which we will call pDUXC. The pDUXC protein encoded by this transcript has two homeodomains and a C-terminal region with homology to other members of the DUXC family (Fig. 1; Supplementary Material, Fig. S1C). The first pDUXC homeodomain has 79% similarity (62% identity) to hDUX4, and the second homeodomain has 88% similarity (68% identity), whereas the comparison to mouse shows 65 and 75% similarity (39 and 63% identity), respectively.

pDUXC activates hDUX4-target genes in human myoblasts

To determine whether pDUXC will induce transcription of hDUX4 targets in human myoblasts, we transduced human myoblasts with a doxycycline inducible codon altered pDUXC transgene or an inducible codon altered hDUX4 (36) and assessed expression of a set of seven genes directly regulated by hDUX4 (7). pDUXC robustly activated this set of genes to a degree similar to hDUX4 (Supplementary Material, Fig. S2), indicating that pDUXC is a functional ortholog of hDUX4 and at least a portion of its DNA binding and transcriptional specificity is conserved.

pDUXC activates an early embryonic transcriptional program

To characterize the full transcriptional program of pDUXC and hDUX4 in pig myoblasts, we transduced primary cultures of pig myoblasts with doxycycline inducible pDUXC and hDUX4 transgenes or a GFP transgene control. Both pDUXC and hDUX4 showed substantial cell death at 24 and 72 hours after induction compared to GFP (Supplementary Material, Fig. S3). To determine the pDUXC transcriptome, we harvested RNA following 20 hours of induction and performed RNA sequencing.

Compared to cells with an induced GFP, pDUXC changed the expression of 2238 genes using a standard 2-fold change threshold (base mean ≥ 100, adjusted *P*-value < 0.05 corresponding to H_0: $|\log_2 FC| \leq 1$; Fig. 2A) with 1511 upregulated and 727 down (Supplementary Material, Table S1). With a more stringent 4-fold change threshold, pDUXC-induced 659 genes and decreased 51 (base mean ≥ 100, adjusted *P*-value < 0.05 corresponding to H_0: $|\log_2 FC| \leq 2$).

mDux and hDUX4 drive an early totipotent program that is part of the transcriptome of the 2-cell-like (2C-like) or 4C/8C-like state in mouse and human embryos, respectively (5,37). Of the 284 most highly induced genes of the mouse 2C-like gene signature ($\log_2 FC > 2$) (37), 96 have an annotated porcine ortholog (Supplementary Material, Table S2), and 48 (50%) of these were induced by pDUXC (adjusted *P*-value < 0.05 corresponding to H_0: $|\log_2 FC| \leq 1$; hypergeometric *P*-value = 3.00e-8; Fig. 2B). In addition, many members of the gene families induced by hDUX4 and mDux in the early embryo, such as the MBD3L, PRAMEF, TRIM and ZSCAN families, are not annotated in the pig genome and thus excluded from the comparison to the mouse 2C-like transcriptome analysis. However, sequence homology indicates that many members of these families are among the unannotated genes highly induced by pDUXC (see Fig. 2B), including MBD3L5, PRAMEF12, TRIM43 and ZSCAN5B.

Finally, gene ontology (GO) term analysis of pDUXC activated genes revealed an enrichment of genes related to regulation of developmental process, regulation of apoptosis and programmed cell death and embryo development (adjusted *P*-value < 0.05; Supplementary Material, Table S3). Together, these analyses indicate that pDUXC activates a conserved early totipotent program when expressed in myoblasts, consistent with its expression in pig embryos between the 2- and 8-cell stages.

hDUX4 and pDUXC activate similar transcriptional programs in porcine myoblasts

RNA sequencing indicated that the induced expression of hDUX4 (172.1 +/− 9.5 SD TPM) was approximately 16-fold lower than the

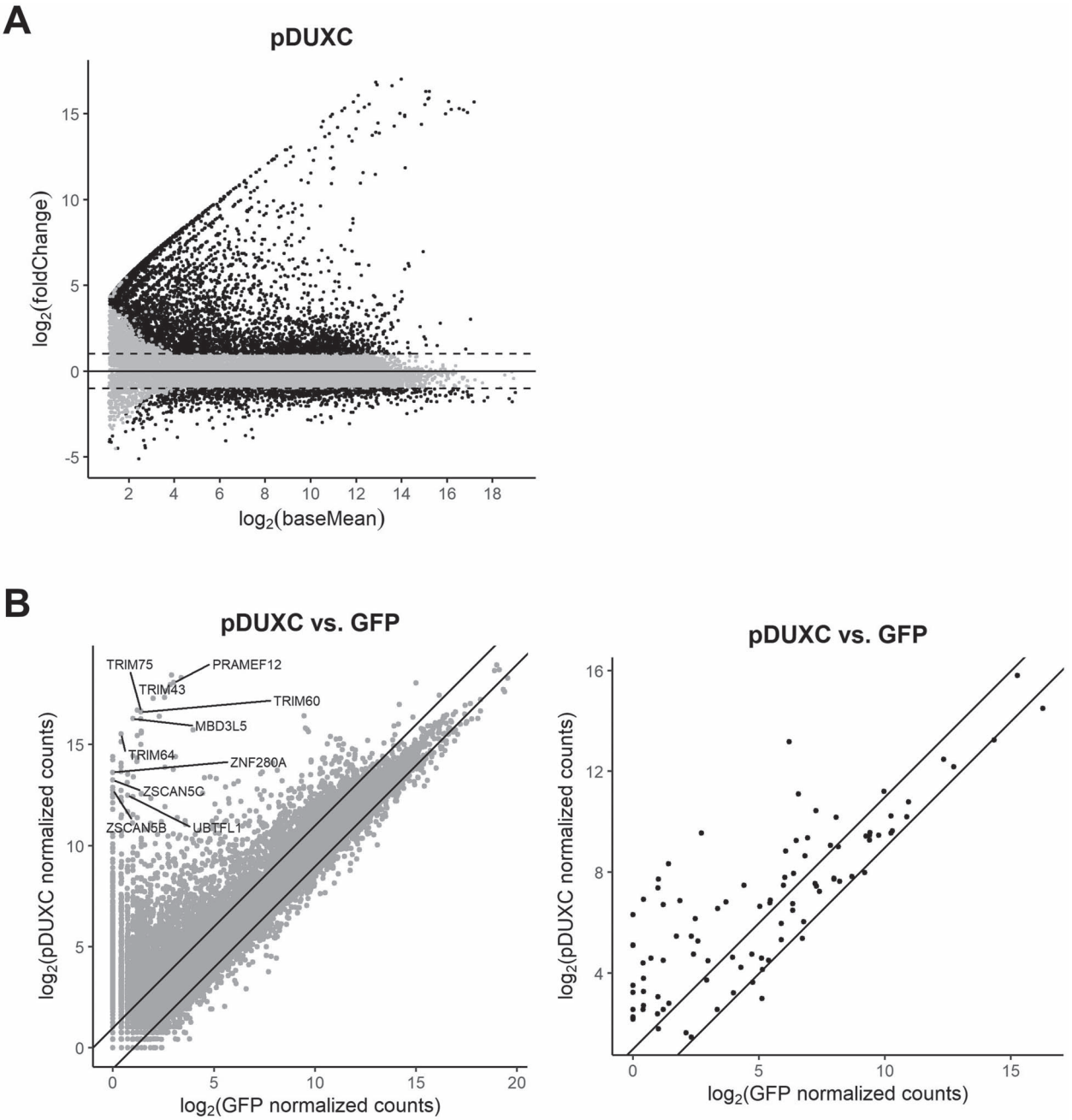

Figure 2. Transcriptome analysis of pDUXC in porcine myoblasts. (**A**) MA plot of gene expression in porcine myoblasts expressing pDUXC compared to GFP control. The *x*-axis is the $\log_2$ of the mean of normalized counts (baseMean) of each gene between two conditions (A value), and the *y*-axis is the $\log_2$ fold change (M value) where positive values indicate higher expression in the pDUXC expressing condition compared to the GFP expressing condition. Black dots represent the differentially expressed genes with adjusted *P*-value < 0.05 corresponding to H_0: $|\log_2 FC| \leq 1$. (**B**) Scatter plot of gene expression in porcine myoblasts expressing pDUXC compared to GFP, with normalized counts in the GFP condition on the x-axis and pDUXC condition on the *y*-axis. Left panel shows all genes, and right panel shows mouse 2C genes with a pig ortholog.

induced expression of pDUXC (2822.0 +/− 122.8 SD TPM) in the pig myoblasts. Despite its relatively lower level of expression at the mRNA level, hDUX4 activated a robust transcriptional program. Compared to an induced GFP, hDUX4 changed the expression of 1240 genes using a standard 2-fold threshold (base mean ≥ 100, adjusted *P*-value < 0.05 corresponding to H_0: $|\log_2 FC| \leq 1$; Fig. 3A) with 833 upregulated and 407 down (Supplementary Material, Table S1). With a more stringent 4-fold change threshold, hDUX4 increased 381 genes and decreased 99 (base mean ≥ 100, adjusted *P*-value < 0.05 corresponding to H_0: $|\log_2 FC| \leq 2$). Of the 96 mouse 2C-like genes (37) that have an annotated porcine homolog (Supplementary Material, Table S2), 29 (30%) were upregulated by hDUX4 (adjusted *P*-value < 0.05 corresponding to H_0: $|\log_2 FC| \leq 1$; hypergeometric *P*-value = 1.93e-8; Fig. 3B), including MBD3L5, PRAMEF12, TRIM43 and ZSCAN5B, which are orthologs of hDUX4 targets. GO term analysis of the hDUX4-induced genes showed an enrichment for genes related to developmental process and embryo development (Supplementary Material, Table S3).

Finally, there was a strong correlation between the genes induced by pDUXC and hDUX4 (Pearson $R = 0.70$ for genes with a base mean expression ≥ 100) (Fig. 3C), although hDUX4 had overall lower levels of induction possibly related to its lower level

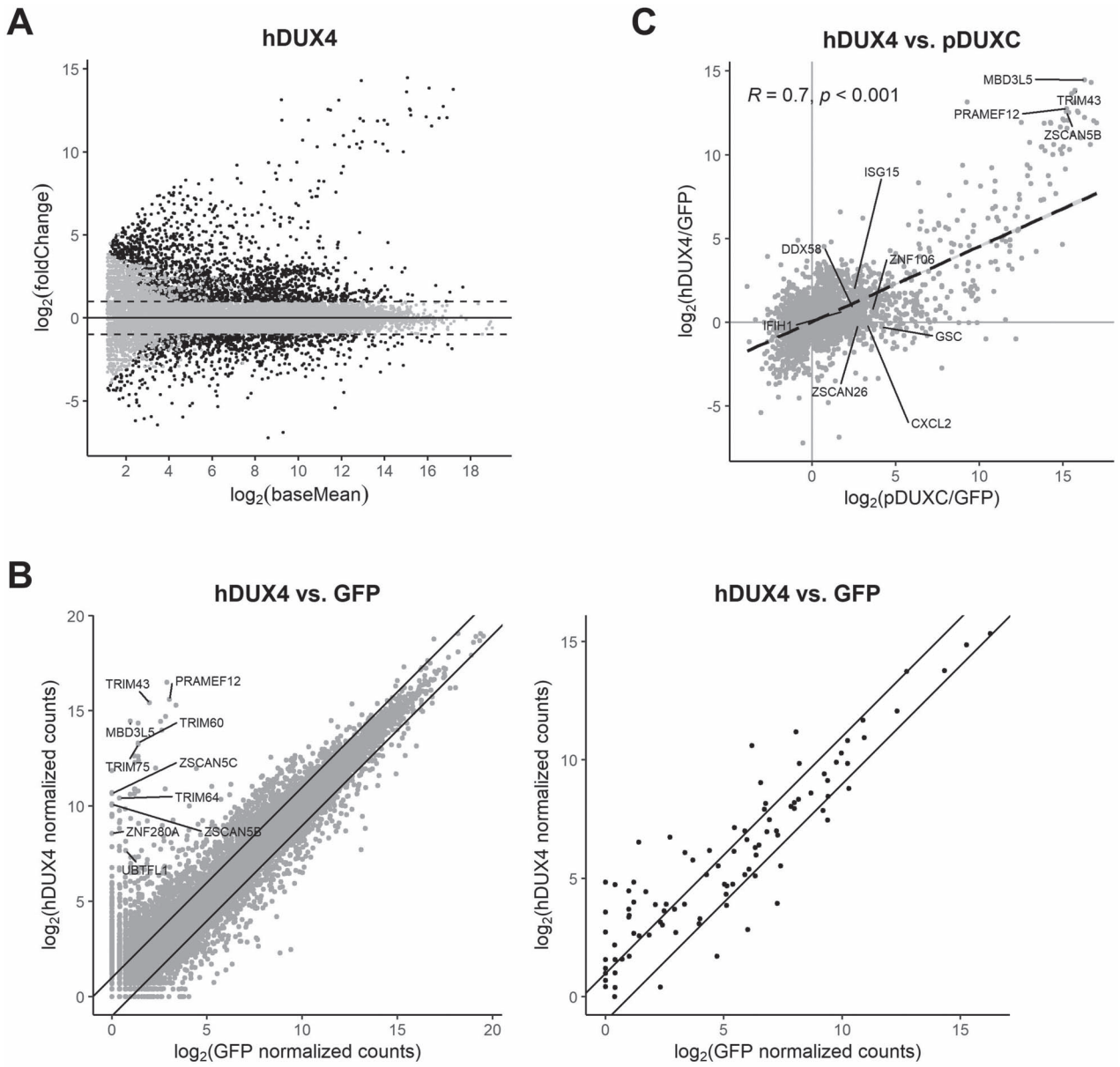

Figure 3. Transcriptome analysis of hDUX4 compared to pDUXC in porcine myoblasts. (**A**) MA plot of gene expression in porcine myoblasts expressing hDUX4 compared to GFP control. The x-axis is the $\log_2$ of the mean of normalized counts (baseMean) on each gene between two conditions (A value), and the y-axis is the $\log_2$ fold change (M value) where positive values indicate higher expression in the hDUX4 expressing condition compared to the GFP expressing condition. Black dots represent the differentially expressed genes with adjusted *P*-value < 0.05 corresponding to H_0: $|\log_2 FC| \leq 1$. (**B**) Scatter plot of gene expression in porcine myoblasts expressing hDUX4 compared to GFP, with normalized counts in the GFP condition on the x-axis and hDUX4 condition on the y-axis. Left panel shows all genes; right panel shows mouse 2C genes with a pig ortholog. (**C**) Scatter plot showing the linear relationship of the $\log_2$ fold change between the pDUXC (x-axis) and hDUX4 (y-axis) comparison models ($R = 0.7$, $P < 0.001$). The dashed line is the linear regression $y = 0.026 + 0.45x$.

of expression compared to pDUXC. A lack of RNA sequencing reads mapping back to endogenous pDUXC (<0.50 TPM) confirms that our data were not confounded by hDUX4 activation of endogenous pDUXC.

hDUX4 induces the expression of multiple classes of endogenous retrotransposons in human cells. Similarly, mapping the RNA-seq reads to the annotations from the UCSC Genome Browser RepeatMasker track showed that both pDUXC- and hDUX4-induced repetitive elements in pig myoblasts (Fig. 4A and B; Supplementary Material, Table S4), including LTR subfamilies similar to those activated by hDUX4 in human cells, LINE1 elements and simple repeats similar to the pericentromeric satellite repeats activated in human cells. Although there is some variation in the degree of activation for each repeat element, overall there is a moderate correlation ($R = 0.6$) between the repeat elements induced by pDUXC and hDUX4 in pig myoblasts (Fig. 4C). Together, these data demonstrate that hDUX4 and pDUXC activate similar transcriptional programs in pig myoblasts.

pDUXC target genes show enrichment for a motif that is similar to the cDUXC and hDUX4 binding motifs

To identify candidate binding motifs for pDUXC, we took a computational approach, due to a lack of a pDUXC antibody for ChIP-seq. We identified a 300 nucleotide segment of DNA near the transcriptional start site (TSS) (−250 to +50 relative to the TSS

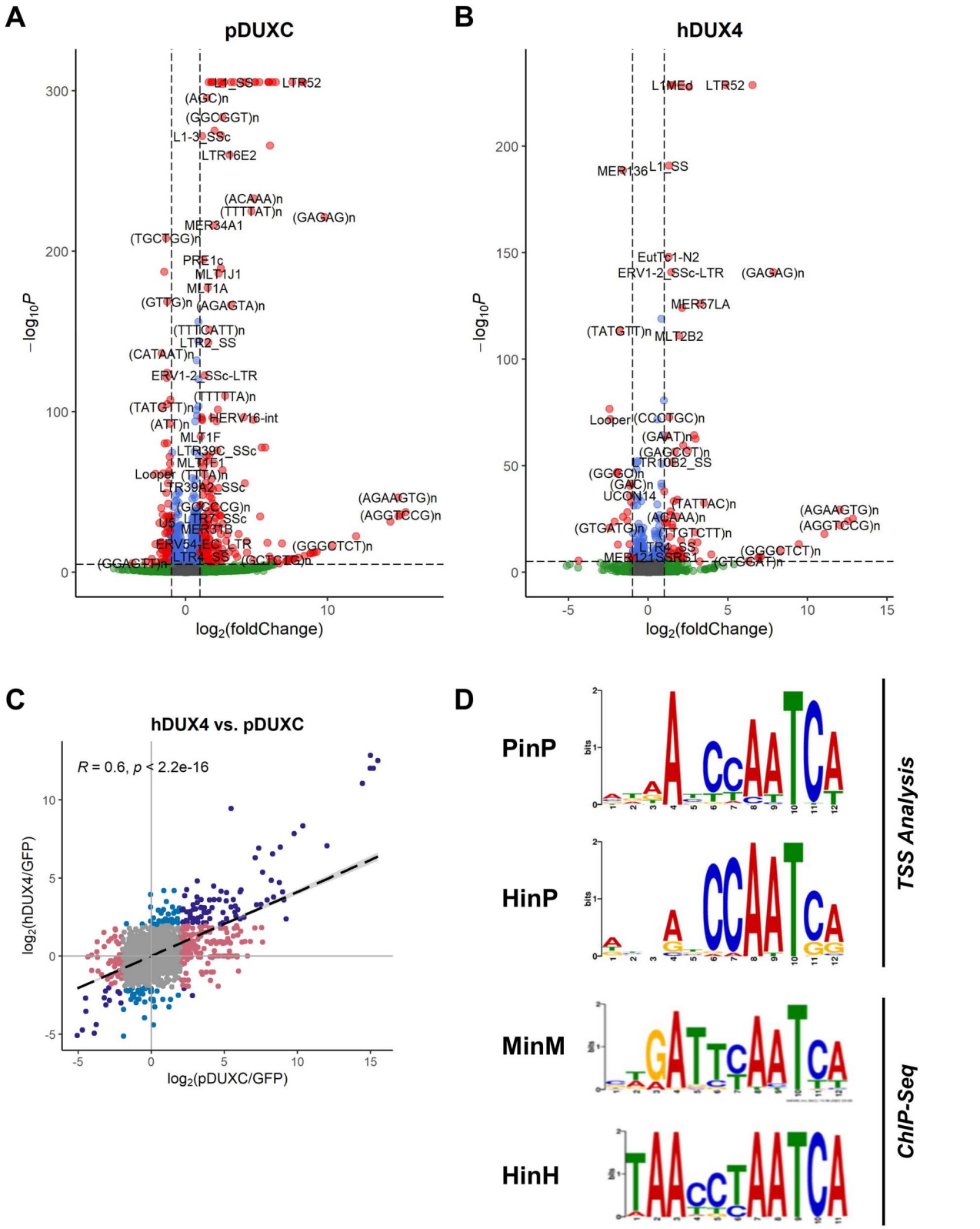

Figure 4. Repetitive element activation and candidate DNA binding motifs for pDUXC and hDUX4 in porcine myoblasts. (**A**) Volcano plot of repetitive element activation by pDUXC relative to GFP control. Green, $|\log_2 FC| > 2$; blue, $-\log_{10}P < 10e^{-6}$; red, both. (**B**) Same as A, for hDUX4. (**C**) Scatter plot showing correlation of repetitive element activation by pDUXC and hDUX4 ($R = 0.6$, $P < 2.2e\text{-}16$). Red, significantly changed by pDUXC; blue, by hDUX4, purple, by both. The dashed line is the linear regression $y = 0.013 + 0.41x$. (**D**) Candidate DNA binding motifs for pDUXC and hDUX4 expressed in porcine myoblasts was identified through MEME computational analysis of the genomic region −250 nt to +50 nt relative to the TSS as determined by RNA sequencing reads of the top 146 genes induced by pDUXC in pig cells or the highly overlapping top 106 genes induced by hDUX4 in pig cells. Binding site sequences determined by ChIP-seq for hDUX4 in human cells and mDux in mouse cells and presented in a prior study (26) are shown for comparison. PinP, pDUXC expressed in pig cells; HinP, hDUX4 expressed in pig cells; MinM, mDux expressed in mouse cells; HinH, hDUX4 expressed in human cells.

as determined by RNA-seq reads) of the top 146 genes induced by pDUXC (110 of which were also induced > 2-fold by hDUX4) and used a *de novo* motif discovery algorithm (MEME) to identify a candidate pDUXC binding site (Fig. 4D), which showed a strong similarity to the binding site previously determined for hDUX4 in human cells based on ChIP-seq (5,26). Using a similar pipeline, we analyzed the top 106 genes induced by hDUX4 (89 of which were also induced > 2-fold by pDUXC and 66 were in the pDUXC analysis) and identified a candidate binding motif for hDUX4 in pig cells (Fig. 4D).

Divergence of pDUXC and hDUX4 in innate immune response

GO term analysis of the pDUXC activated genes also showed an enrichment for genes related to regulation of the immune response (Supplementary Material, Table S3). Interestingly, this is in contrast to prior reports that hDUX4 expressed in human cells suppresses the immune response (8,38). Notable genes of the innate immune response like DDX58, IFIH1 and ISG15 were induced by pDUXC in porcine cells (Supplementary Material, Table S1), whereas these genes were not induced by hDUX4 in porcine cells, nor when expressed in human cells (8). These findings suggest an evolutionary divergence of function between the two DUXC proteins, with the caveat that hDUX4 was expressed at much lower levels than pDUXC. Future studies will be necessary to evaluate this further.

Homeodomain region confers pDUXC specificity on a subset of genes

To determine whether the differential activity of pDUXC and hDUX4 on a subset of genes could be attributed to divergence of their homeodomains or C-terminal regions, we generated a doxycycline inducible vector expressing a chimeric transcript encoding the pDUXC homeodomains (HD, aa1-171) fused to the C-terminal domain of hDUX4 (CTD, aa157-424) and transduced pig myoblasts with the inducible pDUXC, hDUX4 and pDUXChd:hDUX4ctd vectors. Reverse transcription–quantitative polymerase chain reaction (RT-qPCR) with primers in a region of the transcript common to the three transgenes showed an approximately 10-fold higher expression of the doxycycline-induced pDUXC compared to the hDUX4, whereas the pDUXChd:DUX4ctd chimera was induced nearly to the level of pDUXC (Fig. 5A, pCW57.1 primers). A set of genes that showed induction by both pDUXC and hDUX4 in the prior RNA-seq analysis showed robust induction by all three vectors (Fig. 5A), although the absolute level of induction by hDUX4 was lower than for pDUXC or pDUXChd:hDUX4ctd and consistent with the lower level of the hDUX4 transgene expression. In contrast, a set of genes that showed relatively preferential induction by pDUXC compared to hDUX4 in the prior RNA-seq analysis were also preferentially induced by pDUXC in this independent experiment, whereas the pDUXChd:hDUX4ctd chimera activated these genes to approximately the same degree as pDUXC (Fig. 5B). Therefore, the relatively inefficient induction of these genes by hDUX4 can be partly rescued by substituting the pig homeodomains and suggests that the homeodomain binding site divergence likely contributes to the specificity of pDUXC for a subset of genes.

Discussion

Currently, there are several preclinical models for FSHD, each with specific advantages and limitations. Several groups have generated mice that express hDUX4, either by inserting a shortened D4Z4 repeat and surrounding sequences (19) or inducible transgenes (15–18,21), or using AAV to deliver hDUX4 (20). Together these models have utility to test interventions that suppress hDUX4 mRNA expression, stability, translation and protein activity. However, because of the rapid divergence of mDux and hDUX4 first homeodomain and binding specificity (5), these preclinical models are limited regarding the consequences of hDUX4 expression, specifically how the hDUX4 transcriptional program results in disease pathophysiology and progression. To some extent, xenografts of human FSHD muscle into mice can partly address this limitation (22–24), but are also limited by the degree of engraftment, functional assessment, lack of an immune response and dynamic variables of regeneration. Therefore, although each existing preclinical model has strong utility for specific aspects of FSHD, none of the existing models accurately recapitulates disease initiation and progression.

Previously we showed that cDUXC showed less divergence from hDUX4 than mDux (26) and suggested that Laurasiatherians might have advantages compared to rodents as preclinical models for DUX4-induced pathology. In the current study, we identify the pDUXC mRNA expressed in the cleavage-stage pig embryo and show that its expression in pig muscle cells activates the same totipotent transcriptional program as hDUX4 and mDux in human and mouse cells, respectively; the same transcriptional program that hDUX4 induces in FSHD muscle cells (7,9,10). Most importantly, expression of hDUX4 in pig muscle cells robustly activates the same totipotent transcriptional program. The conservation of the transcriptional program when hDUX4 is expressed in pig muscle cells supports further evaluation of pigs as a preclinical model for FSHD.

Despite the lower expression of hDUX4 in the pig muscle cells, the correlation with genes induced by pDUXC was much better than prior studies comparing the hDUX4 and mDux in mouse muscle cells (5) and similar to the comparison of hDUX4 and cDUXC in dog muscle cells (26). A notable exception was a set of genes in the innate immune signaling pathways that were induced in pig muscle by pDUXC but not by hDUX4. Our prior studies in human muscle cells showed that although hDUX4 induces double-stranded RNA and retrotransposon expression sufficient to result in the phosphorylation of EIF2a/PKR (39,40), the presence of hDUX4 suppresses the induction of many genes in the innate immune response pathways (8), such as a broad class of interferon stimulated genes (ISGs). Our current findings suggest the cloned version of pDUXC might not suppress these pathways and warrants future study. Furthermore, the induction of a subset of genes preferentially induced by pDUXC, including some genes in the immune signaling pathways, by a chimeric protein consisting of the pig homeodomain region and the human C-terminal region, suggests that the preferential induction of subsets of genes by pDUXC might be due, in part, to divergence of the homeodomain sequence between pig and human.

Over the last decade, pigs have become an increasingly important model for a broad spectrum of human diseases and preclinical therapies (27,28,41). For example, in the neuromuscular diseases, genetically engineered pigs recapitulate the disease pathology and progression for amyotrophic lateral sclerosis caused by a mutation in SOD1 (29) and for DMD caused by dystrophin mutations (31,32,42). The DMD pigs have advanced studies of disease progression (43) and, most importantly, provide a valuable preclinical model for systemic therapies, such as the AAV delivery of Cas9-mediated exon excision vectors (33). Compared to rodents, the pig immune system has some features more similar to humans (44). These similarities of pig and human immune

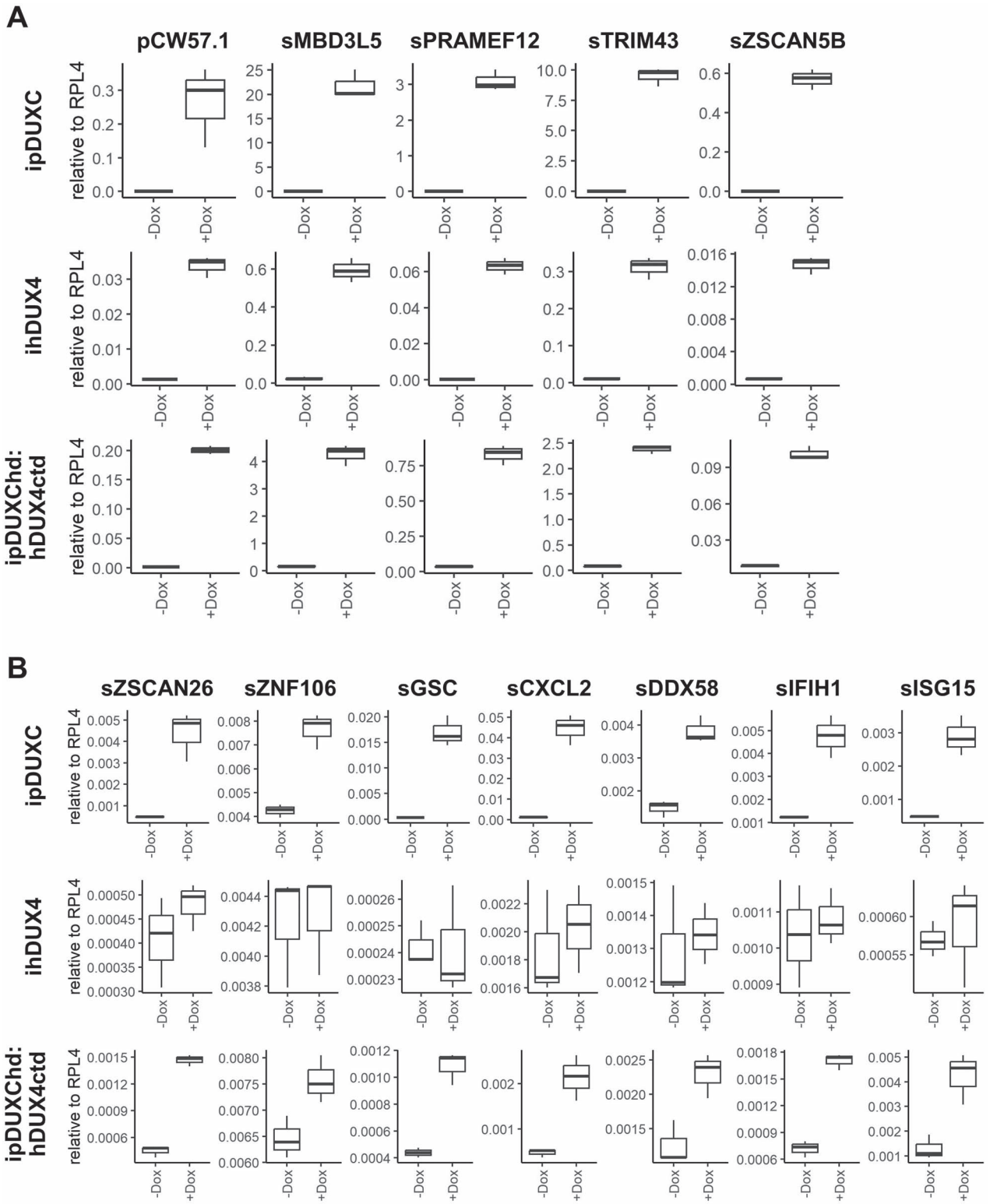

Figure 5. Homeodomain region confers pDUXC specificity on a subset of genes. Pig myoblasts were transduced with the doxycycline inducible pDUXC, hDUX4 or pDUXChd:hDUX4ctd constructs and the indicated mRNA levels measured by RT-qPCR before (–) and after (+) treatment with doxycycline. (**A**) pCW57.1 primers amplify a portion of the vector sequence contained in the mRNA transcript for each transgene and permit measurement of relative transgene expression levels; the genes sMBD3L5, sPRAMEF12, sTRIM43 and sZSCAN5B are all genes that were induced by pDUXC and hDUX4 in the RNA-seq. (**B**) Genes preferentially induced by pDUXC compared to hDUX4 in the RNA-seq data were also preferentially induced by pDUXC in this new dataset using RT-qPCR. In contrast, the chimeric protein pDUXChd:hDUX4ctd showed target induction similar to pDUXC, indicating that part of the target specificity of pDUXC relative to hDUX4 maps to the homeodomains.

responses might further make the pig a good model for FSHD, which is characterized by an immune infiltration in skeletal muscle (45,46). In addition, similar to the human population, pre-existing neutralizing antibodies to different AAV serotypes exist in sampled pig populations (47) and provide an opportunity for measuring their possible implications for future human studies (48); whereas studies determining pre-existing immunity to different Cas9 proteins remain to be investigated. Finally, the development of mini-pig strains with breeding, growth and behavioral characteristics conducive for medical research, such as the Göttingen minipig (49), has greatly facilitated the use of these pigs for disease models. Our findings provide the basis for evaluating the FSHD transcriptome in future pig models of FSHD and also provide hope that hDUX4 expression in pigs might faithfully recapitulate the transcriptional profile and pathophysiology of FSHD.

Materials and Methods

Primary myoblast isolation from porcine skeletal muscle and cell culture

The semimembranosus (SM) muscle was dissected from a 1-year-old wild-type male Göttingen pig at Marshall BioResources (North Rose, NY), rinsed with 1× PBS + 0.5× antibiotic-antimycotic (ABM) and shipped overnight to the Jones Lab (UNR, Reno, NV) in a moistened container on a cold pack for live cell isolation. The SM muscle (~5 g) was rinsed with cold ePBS (25 mM glucose, 14 mM sucrose, 0.5× ABM in 1× PBS), manually minced with sterilized scissors in ePBS and centrifuged at 800 x g for 1 minute, and the tissue pellet was resuspended with 15 mL digestion buffer (0.025% Trypsin–EDTA, 0.01% DNaseI, 0.2% Collagenase type IA in HBSS, Ca^{2+}, Mg^{2+}-free) and digested at room temperature for 30 minutes with gentle stirring at 100 RPM. The tissue lysate was neutralized with an equal volume of 10% FBS/DMEM, diluted to 50 mL with DMEM and then sequentially filtered through a 500-μm mesh and then a 100-μm mesh to remove muscle fiber debris. Cells were pelleted at 800× g for 2 minutes, resuspended in 1× PBS and processed for removal of red blood cells by the standard protocol. Remaining cells were plated on 10% Matrigel-coated cell culture plates and cultured with HMP (20% FBS, 1x chicken embryonic extract, 1x ABM in Ham's F-10) with 10 ng/mL bFGF for 2–3 days. Once cell number was $>1 \times 10^6$, myogenic cells were isolated using MACS sorting with human CD56-conjugated magnetic beads per manufacturer's instruction. The CD56-positive cells (>95%) were plated on 10% Matrigel-coated plate and expanded in HMP + 10 ng/mL bFGF media. Myogenic differentiation was confirmed by culturing in DMEM with 1% heat-inactivated horse serum, 10μg/mL insulin and 10μg/mL transferrin. These primary porcine myoblasts were cultured in F-10 media supplemented with 15% fetal bovine serum, 1× Gibco Antibiotic-Antimycotic and 1% chicken embryo extract for further experiments. Immortalized human myoblasts (MB135) were cultured in F-10 media supplemented with 10% fetal bovine serum, 1% penicillin/streptomycin, 10 ng/mL rhFGF and 1uM dexamethasone.

Expression of pDUXC, hDUX4, pDUXChd:hDUX4ctd and GFP

The pDUXC transgene was cloned into the pCW57.1 lentiviral vector, which has a doxycycline-inducible promoter. pDUXC was codon altered (ca) to enhance transgene expression. Primary porcine myoblasts were transduced with pCW57.1-pDUXCca, followed by puromycin selection. Stable cell lines expressing hDUX4ca (36) pDUXChd:hdDUX4ctd and 3xFLAG-NLS-GFP were generated in a similar manner.

Reverse transcription-quantitative PCR

Inducible pDUXCca, hDUX4ca, pDUXChd:hDUX4ctd and GFP cells were treated, in triplicate, with 1ug/mL doxycycline for 20 hours and harvested for RNA with the NucleoSpin RNA Kit (Takara) according to the manufacturer's protocol. RNA quality was verified by NanoDrop 2000 (Thermo Scientific). RNA was treated with DNaseI, Amplification Grade (Invitrogen). Reverse transcription was performed in a 20 μL reaction: 500 ng whole RNA, 1uL dNTP (10 mM), 1uL oligo dT primer (10 mM), 4 μL 5× SSIV Buffer, 1 μL DTT (100 mM), 1 μL RNaseOUT and 1 μL SSIV RT enzyme. Thermal cycling conditions for reverse transcription were as follows: 50°C for 40 minutes, 55°C for 30 minutes and 80°C for 10 minutes. Complementary DNA (cDNA) was treated with 1uL of RNaseH and incubated at 37°C for 20 minutes, then diluted 1:5 with 80 μL of RNase-free H_2O. Quantitative real-time PCR (qPCR) was performed on the QuantStudio™ 7 Flex Real-Time PCR System in a 10uL reaction: 2 μL cDNA (~10 ng), 5 μL 2x iTaq Universal SYBR Green Supermix, 0.3 μL forward and reverse primer (10 μM) and 2.4 μL H_2O. qPCR primers were synthesized by Integrated DNA Technologies (IDT) and are listed below. Thermal cycling conditions for qPCR were as follows: 50°C for 2 minutes and 95°C for 10 minutes; 40 cycles of 95°C for 15 seconds and 60°C for 60 seconds. qPCR reactions were run in technical triplicates, including RT controls. Median CT values of the technical triplicates were used for analysis. Gene expression was normalized to the housekeeping gene RPL4 (ribosomal protein L4).

qPCR primer sequences

Target	Forward primer	Reverse primer
sRPL4 (50)	AGGAGGCTGTTCTGCTTCTG	TCCAGGGATGTTTCTGAAGG
sMBD3L5	CATCATCGTCTGGAGAAGCTCA	CCAGGATGAGCAGTTGTCTTTG
sTRIM43	TCAGGAATTGGGAGGATTATGTG	TCTTTGCTCTCCTCGTCTAGT
sPRAMEF12	AGATGCTCAGGTGCCTGGAT	AGGAGCTCAGGGCTAAACTT
sZSCAN5B	AAGGGTCCTGAGACAGAAGA	CCTTCTTGGGAAGTGAGACTG
sZSCAN26	AGACAGCCCCTCTGAAAGCA	ATCTTGGTCTCCTTACCCTTCTCT
sZNF106	AGAGGGATCTGCTGTGTGAAG	CCGCATGTGTTCATCCATCTC
sGSC	GAGGAGAAAGTGGAGGTCTGGT	CGTCTTGTTCCATTTCTCGGC
sCXCL2	ACCAAACGGAAGTCATAGCCA	CATCAGTTGGCACTGCTCTTG
sDDX58	ATAAGAATGGTGGAGAAATGCCA	TCTTTGGTTTGGGGTGCAGT
sIFIH1	GCTGTGAAAGCAATGCAGAATC	AGACTTGGCTGATCTGTGGC
sISG15	ACGGCCATGGGTAGGGA	ATCTGCTGCTTCAGGTCCG
hRPL27	GCAAGAAGAAGATCGCCAAG	TCCAAGGGGATATCCACAGA
h3.Y	GGACCTGCGCTTCCAGAG	CATGTCTCGGGGCATAATTG
hKHDC1L	TGAATCAGGTGGGAGCACAG	CAATGCAGCGAAGGTACGTG
hLEUTX	AAGGAGGAGACTCCCTCAGC	AAAGAGAGTGGAGGCCCAAG
hMBD3L2	GACGCAGAAAAGGGGACGAG	CTGTCCAGAGATTCACCGGC
hPRAMEF12	TCACCTCTCAGTTCCTCAAGC	CAGGCATTCGGTCATTACG
hTRIM43	ACCCATCACTGGACTGGTGT	CACATCCTCAAAGAGCCTGA
hZSCAN4	TGGAAATCAAGTGGCAAAAA	CTGCATGTGGACGTGGAC
pCW57.1	TGACTGGATATGTTGTGTTTTAC	CAACCCCGGATCCTTAGTG

RNA sequencing

RNA samples were submitted in biological triplicates for bulk RNA sequencing (100-bp single ended sequencing on NextSeq 2000, about 50 million reads per sample). Reads were aligned to the reference genome RefSeq Sscrofa 11.1 (Accession # GCF_000003025.6) using Rsubread, and differential expression was called using DESeq2 using GFP samples as reference levels. TSS of pDUXC and hDUX4-activated genes were identified through the RNA sequencing tracks, and *de novo* motif analysis was performed with MEME software with a 300-bp sequence (−250 bp to +50 bp of TSS). Repetitive element annotation from the UCSC Genome Browser RepeatMasker track was used to identify RNA-seq reads mapping to repetitive elements (51–55).

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We thank Marshall Bioresources for providing the skeletal muscle from the Göttingen pig for derivation of primary muscle cell cultures for these experiments.

Conflict of Interest statement. P.L.J. and T.I.J. are founders and CSO (P.L.J.) and Head of Translational Medicine (T.I.J.) of Renogenyx and P.L.J., T.I.J. and S.J.T. serve on the Renogenyx Board of Directors, a company developing therapeutics for FSHD. S.J.T. consults for several pharmaceutical companies developing therapies for FSHD. The authors declare no other competing interests.

Funding

The National Institutes of Health (AR045203 to S.J.T.; P50AR065139 to Wellstone Muscular Dystrophy Specialized Research Center, S.J.T. and J. Chamberlain; AR062587 to P.L.J.), the Friends of FSH Research and the Chris Carrino Foundation for FSHD.

Data availability

RNA sequencing data have been deposited to the GEO database, GSE210772.

References

1. Tawil, R., Van Der Maarel, S.M. and Tapscott, S.J. (2014) Facioscapulohumeral dystrophy: the path to consensus on pathophysiology. *Skelet. Muscle*, **4**, 1–15.
2. Lemmers, R.J., van der Vliet, P.J., Klooster, R., Sacconi, S., Camano, P., Dauwerse, J.G., Snider, L., Straasheijm, K.R., van Ommen, G.J., Padberg, G.W. *et al.* (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*, **329**, 1650–1653.
3. Snider, L., Geng, L.N., Lemmers, R.J., Kyba, M., Ware, C.B., Nelson, A.M., Tawil, R., Filippova, G.N., van der Maarel, S.M., Tapscott, S.J. *et al.* (2010) Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet.*, **6**, e1001181.
4. Hendrickson, P.G., Doráis, J.A., Grow, E.J., Whiddon, J.L., Lim, J.-W., Wike, C.L., Weaver, B.D., Pflueger, C., Emery, B.R. and Wilcox, A.L. (2017) Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.*, **49**, 925–934.
5. Whiddon, J.L., Langford, A.T., Wong, C.J., Zhong, J.W. and Tapscott, S.J. (2017) Conservation and innovation in the DUX4-family gene network. *Nat. Genet.*, **49**, 935–940.
6. De Iaco, A., Planet, E., Coluccio, A., Verp, S., Duc, J. and Trono, D. (2017) DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.*, **49**, 941–945.
7. Yao, Z., Snider, L., Balog, J., Lemmers, R.J., Van Der Maarel, S.M., Tawil, R. and Tapscott, S.J. (2014) DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum. Mol. Genet.*, **23**, 5342–5352.
8. Geng, L.N., Yao, Z., Snider, L., Fong, A.P., Cech, J.N., Young, J.M., van der Maarel, S.M., Ruzzo, W.L., Gentleman, R.C., Tawil, R. *et al.* (2012) DUX4 activates germline genes, retroelements, and immune mediators: implications for facioscapulohumeral dystrophy. *Dev. Cell*, **22**, 38–51.
9. Wang, L.H., Friedman, S.D., Shaw, D., Snider, L., Wong, C.J., Budech, C.B., Poliachik, S.L., Gove, N.E., Lewis, L.M., Campbell, A.E. *et al.* (2019) MRI-informed muscle biopsies correlate MRI with pathology and DUX4 target gene expression in FSHD. *Hum. Mol. Genet.*, **28**, 476–486.
10. Wong, C.J., Wang, L.H., Friedman, S.D., Shaw, D., Campbell, A.E., Budech, C.B., Lewis, L.M., Lemmers, R., Statland, J.M., Maarel, S.M. *et al.* (2020) Longitudinal measures of RNA expression and disease activity in FSHD muscle biopsies. *Hum. Mol. Genet.*, **29**, 1030–1043.
11. Schatzl, T., Kaiser, L. and Deigner, H.P. (2021) Facioscapulohumeral muscular dystrophy: genetics, gene activation and downstream signalling with regard to recent therapeutic approaches: an update. *Orphanet J. Rare Dis.*, **16**, 129.
12. Wang, L.H. and Tawil, R. (2021) Current therapeutic approaches in FSHD. *J. Neuromuscul. Dis.*, **8**, 441–451.
13. Clapp, J., Mitchell, L.M., Bolland, D.J., Fantes, J., Corcoran, A.E., Scotting, P.J., Armour, J.A. and Hewitt, J.E. (2007) Evolutionary conservation of a coding function for D4Z4, the tandem DNA repeat mutated in facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **81**, 264–279.
14. Leidenroth, A., Clapp, J., Mitchell, L.M., Coneyworth, D., Dearden, F.L., Iannuzzi, L. and Hewitt, J.E. (2012) Evolution of DUX gene macrosatellites in placental mammals. *Chromosoma*, **121**, 489–497.
15. Bosnakovski, D., Chan, S.S.K., Recht, O.O., Hartweck, L.M., Gustafson, C.J., Athman, L.L., Lowe, D.A. and Kyba, M. (2017) Muscle pathology from stochastic low level DUX4 expression in an FSHD mouse model. *Nat. Commun.*, **8**, 550.
16. Bosnakovski, D., Shams, A.S., Yuan, C., da Silva, M.T., Ener, E.T., Baumann, C.W., Lindsay, A.J., Verma, M., Asakura, A., Lowe, D.A. *et al.* (2020) Transcriptional and cytopathological hallmarks of FSHD in chronic DUX4-expressing mice. *J. Clin. Invest.*, **130**, 2465–2477.
17. Giesige, C.R., Wallace, L.M., Heller, K.N., Eidahl, J.O., Saad, N.Y., Fowler, A.M., Pyne, N.K., Al-Kharsan, M., Rashnonejad, A., Chermahini, G.A. *et al.* (2018) AAV-mediated follistatin gene therapy improves functional outcomes in the TIC-DUX4 mouse model of FSHD. *JCI Insight*, **3**, e123538.
18. Jones, T.I., Chew, G.L., Barraza-Flores, P., Schreier, S., Ramirez, M., Wuebbles, R.D., Burkin, D.J., Bradley, R.K. and Jones, P.L. (2020) Transgenic mice expressing tunable levels of DUX4 develop characteristic facioscapulohumeral muscular dystrophy-like pathophysiology ranging in severity. *Skelet. Muscle*, **10**, 8.
19. Krom, Y.D., Thijssen, P.E., Young, J.M., den Hamer, B., Balog, J., Yao, Z., Maves, L., Snider, L., Knopp, P., Zammit, P.S. *et al.* (2013) Intrinsic epigenetic regulation of the D4Z4 macrosatellite repeat in a transgenic mouse model for FSHD. *PLoS Genet.*, **9**, e1003415.
20. Wallace, L.M., Garwick, S.E., Mei, W., Belayew, A., Coppee, F., Ladner, K.J., Guttridge, D., Yang, J. and Harper, S.Q. (2011) DUX4, a candidate gene for facioscapulohumeral muscular dystrophy, causes p53-dependent myopathy in vivo. *Ann. Neurol.*, **69**, 540–552.
21. Jones, T. and Jones, P.L. (2018) A cre-inducible DUX4 transgenic mouse model for investigating facioscapulohumeral muscular dystrophy. *PLoS One*, **13**, e0192657.
22. Sakellariou, P., O'Neill, A., Mueller, A.L., Stadler, G., Wright, W.E., Roche, J.A. and Bloch, R.J. (2016) Neuromuscular electrica

l stimulation promotes development in mice of mature human muscle from immortalized human myoblasts. *Skelet. Muscle*, **6**, 4.
23. Mueller, A.L., O'Neill, A., Jones, T.I., Llach, A., Rojas, L.A., Sakellariou, P., Stadler, G., Wright, W.E., Eyerman, D., Jones, P.L. *et al.* (2019) Muscle xenografts reproduce key molecular features of facioscapulohumeral muscular dystrophy. *Exp. Neurol.*, **320**, 113011.
24. Chen, J.C., King, O.D., Zhang, Y., Clayton, N.P., Spencer, C., Wentworth, B.M., Emerson, C.P., Jr. and Wagner, K.R. (2016) Morpholino-mediated knockdown of DUX4 toward facioscapulohumeral muscular dystrophy therapeutics. *Mol. Ther.*, **24**, 1405–1411.
25. Oliva, J., Galasinski, S., Richey, A., Campbell, A.E., Meyers, M.J., Modi, N., Zhong, J.W., Tawil, R., Tapscott, S.J. and Sverdrup, F.M. (2019) Clinically advanced p38 inhibitors suppress DUX4 expression in cellular and animal models of facioscapulohumeral muscular dystrophy. *J. Pharmacol. Exp. Ther.*, **370**, 219–230.
26. Wong, C.J., Whiddon, J.L., Langford, A.T., Belleville, A.E. and Tapscott, S.J. (2022) Canine DUXC: implications for DUX4 retrotransposition and preclinical models of FSHD. *Hum. Mol. Genet.*, **31**, 1694–1704.
27. Bendixen, E., Danielsen, M., Larsen, K. and Bendixen, C. (2010) Advances in porcine genomics and proteomics—a toolbox for developing the pig as a model organism for molecular biomedical research. *Brief. Funct. Genomics*, **9**, 208–219.
28. Lunney, J.K., Van Goor, A., Walker, K.E., Hailstock, T., Franklin, J. and Dai, C. (2021) Importance of the pig as a human biomedical model. *Sci. Transl. Med.*, **13**, eabd5758.
29. Crociara, P., Chieppa, M.N., Vallino Costassa, E., Berrone, E., Gallo, M., Lo Faro, M., Pintore, M.D., Iulini, B., D'Angelo, A., Perona, G. *et al.* (2019) Motor neuron degeneration, severe myopathy and TDP-43 increase in a transgenic pig model of SOD1-linked familiar ALS. *Neurobiol. Dis.*, **124**, 263–275.
30. Duque, S.I., Arnold, W.D., Odermatt, P., Li, X., Porensky, P.N., Schmelzer, L., Meyer, K., Kolb, S.J., Schümperli, D. and Kaspar, B.K. (2015) A large animal model of spinal muscular atrophy and correction of phenotype. *Ann. Neurol.*, **77**, 399–414.
31. Stirm, M., Fonteyne, L.M., Shashikadze, B., Lindner, M., Chirivi, M., Lange, A., Kaufhold, C., Mayer, C., Medugorac, I., Kessler, B. *et al.* (2021) A scalable, clinically severe pig model for Duchenne muscular dystrophy. *Dis. Model. Mech.*, **14**, dmm049285.
32. Klymiuk, N., Blutke, A., Graf, A., Krause, S., Burkhardt, K., Wuensch, A., Krebs, S., Kessler, B., Zakhartchenko, V., Kurome, M. *et al.* (2013) Dystrophin-deficient pigs provide new insights into the hierarchy of physiological derangements of dystrophic muscle. *Hum. Mol. Genet.*, **22**, 4368–4382.
33. Moretti, A., Fonteyne, L., Giesert, F., Hoppmann, P., Meier, A., Bozoglu, T., Baehr, A., Schneider, C., Sinnecker, D. and Klett, K. (2020) Somatic gene editing ameliorates skeletal and cardiac muscle failure in pig and human models of Duchenne muscular dystrophy. *Nat. Med.*, **26**, 207–214.
34. Cao, S., Han, J., Wu, J., Li, Q., Liu, S., Zhang, W., Pei, Y., Ruan, X., Liu, Z. and Wang, X. (2014) Specific gene-regulation networks during the pre-implantation development of the pig embryo as revealed by deep sequencing. *BMC Genomics*, **15**, 1–13.
35. Warr, A., Affara, N., Aken, B., Beiki, H., Bickhart, D.M., Billis, K., Chow, W., Eory, L., Finlayson, H.A. and Flicek, P. (2020) An improved pig reference genome sequence to enable pig genetics and genomics research. *Gigascience*, **9**, giaa051.
36. Jagannathan, S., Shadle, S.C., Resnick, R., Snider, L., Tawil, R.N., van der Maarel, S.M., Bradley, R.K. and Tapscott, S.J. (2016) Model systems of DUX4 expression recapitulate the transcriptional profile of FSHD cells. *Hum. Mol. Genet.*, **25**, 4419–4431.
37. Akiyama, T., Xin, L., Oda, M., Sharov, A.A., Amano, M., Piao, Y., Cadet, J.S., Dudekula, D.B., Qian, Y. and Wang, W. (2015) Transient bursts of Zscan4 expression are accompanied by the rapid derepression of heterochromatin in mouse embryonic stem cells. *DNA Res.*, **22**, 307–318.
38. Chew, G.L., Campbell, A.E., De Neef, E., Sutliff, N.A., Shadle, S.C., Tapscott, S.J. and Bradley, R.K. (2019) DUX4 suppresses MHC class I to promote cancer immune evasion and resistance to checkpoint blockade. *Dev. Cell*, **50**, 658, e657–671.
39. Shadle, S.C., Bennett, S.R., Wong, C.J., Karreman, N.A., Campbell, A.E., van der Maarel, S.M., Bass, B.L. and Tapscott, S.J. (2019) DUX4-induced bidirectional HSATII satellite repeat transcripts form intranuclear double stranded RNA foci in human cell models of FSHD. *Hum. Mol. Genet.*, **28**, 3997–4011.
40. Shadle, S.C., Zhong, J.W., Campbell, A.E., Conerly, M.L., Jagannathan, S., Wong, C.J., Morello, T.D., van der Maarel, S.M. and Tapscott, S.J. (2017) DUX4-induced dsRNA and MYC mRNA stabilization activate apoptotic pathways in human cell models of facioscapulohumeral dystrophy. *PLoS Genet.*, **13**, e1006658.
41. Aigner, B., Renner, S., Kessler, B., Klymiuk, N., Kurome, M., Wunsch, A. and Wolf, E. (2010) Transgenic pigs as models for translational biomedical research. *J. Mol. Med. (Berl)*, **88**, 653–664.
42. Echigoya, Y., Trieu, N., Duddy, W., Moulton, H.M., Yin, H., Partridge, T.A., Hoffman, E.P., Kornegay, J.N., Rohret, F.A., Rogers, C.S. *et al.* (2021) A dystrophin Exon-52 deleted miniature pig model of Duchenne muscular dystrophy and evaluation of exon skipping. *Int. J. Mol. Sci.*, **22**, 13065.
43. Tamiyakul, H., Kemter, E., Kosters, M., Ebner, S., Blutke, A., Klymiuk, N., Flenkenthaler, F., Wolf, E., Arnold, G.J. and Frohlich, T. (2020) Progressive proteome changes in the myocardium of a pig model for Duchenne muscular dystrophy. *iScience*, **23**, 101516.
44. Pabst, R. (2020) The pig as a model for immunology research. *Cell Tissue Res.*, **380**, 287–304.
45. Frisullo, G., Frusciante, R., Nociti, V., Tasca, G., Renna, R., Iorio, R., Patanella, A.K., Iannaccone, E., Marti, A., Rossi, M. *et al.* (2011) CD8(+) T cells in facioscapulohumeral muscular dystrophy patients with inflammatory features at muscle MRI. *J. Clin. Immunol.*, **31**, 155–166.
46. Statland, J.M., Shah, B., Henderson, D., Van Der Maarel, S., Tapscott, S.J. and Tawil, R. (2015) Muscle pathology grade for facioscapulohumeral muscular dystrophy biopsies. *Muscle Nerve*, **52**, 521–526.
47. Calcedo, R., Franco, J., Qin, Q., Richardson, D.W., Mason, J.B., Boyd, S. and Wilson, J.M. (2015) Preexisting neutralizing antibodies to adeno-associated virus capsids in large animals other than monkeys may confound in vivo gene therapy studies. *Hum. Gene Ther. Methods*, **26**, 103–105.
48. Watano, R., Ohmori, T., Hishikawa, S., Sakata, A. and Mizukami, H. (2020) Utility of microminipigs for evaluating liver-mediated gene expression in the presence of neutralizing antibody against vector capsid. *Gene Ther.*, **27**, 427–434.
49. Simianer, H. and Kohn, F. (2010) Genetic management of the Gottingen Minipig population. *J. Pharmacol. Toxicol. Methods*, **62**, 221–226.
50. Cinar, M.U., Islam, M.A., Uddin, M.J., Tholen, E., Tesfaye, D., Looft, C. and Schellander, K. (2012) Evaluation of suitable reference genes for gene expression studies in porcine alveolar

macrophages in response to LPS and LTA. *BMC. Res. Notes*, **5**, 1–14.
51. Blighe, K., Rana, S. and Lewis, M. (2022) EnhancedVolcano: publication-ready volcano plots with enhanced colouring and labeling. *R package version 1.16.0*. *https://github.com/kevinblighe/EnhancedVolcano* in press.
52. Smit, A.F., Hubley, R. and Green, P. (1996-2010) Repeat-masker Open-3.0. *http://www.repeatmasker.org*., in press.
53. Smit, A.F. (1999) Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.*, **9**, 657–663.
54. Smit, A.F. (1996) The origin of interspersed repeats in the human genome. *Curr. Opin. Genet. Dev.*, **6**, 743–748.
55. Jurka, J. (2000) Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet.*, **16**, 418–420.