

GENERAL ARTICLE

DUX4-induced bidirectional HSATII satellite repeat transcripts form intranuclear double-stranded RNA foci in human cell models of FSHD

Sean C. Shadle^{1,2}, Sean R. Bennett¹, Chao-Jen Wong¹, Nancy A. Karreman¹, Amy E. Campbell¹, Silvère M. van der Maarel³, Brenda L. Bass⁴ and Stephen J. Tapscott^{1,*}

¹Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA, ²Molecular and Cellular Biology Program, University of Washington, Seattle, WA 98105, USA, ³Department of Human Genetics, Leiden University Medical Center, 2333 ZA Leiden, the Netherlands, and ⁴Department of Biochemistry, University of Utah, Salt Lake City, UT 84112, USA

*To whom correspondence should be addressed at: Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, Seattle, WA 98109, USA. Tel: +1 2066674499;
Email: stapscot@fredhutch.org

Abstract

The DUX4 transcription factor is normally expressed in the cleavage-stage embryo and regulates genes involved in embryonic genome activation. Misexpression of DUX4 in skeletal muscle, however, is toxic and causes facioscapulohumeral muscular dystrophy (FSHD). We recently showed DUX4-induced toxicity is due, in part, to the activation of the double-stranded RNA (dsRNA) response pathway and the accumulation of intranuclear dsRNA foci. Here, we determined the composition of DUX4-induced dsRNAs. We found that a subset of DUX4-induced dsRNAs originate from inverted *Alu* repeats embedded within the introns of DUX4-induced transcripts and from DUX4-induced dsRNA-forming intergenic transcripts enriched for endogenous retroviruses, *Alu* and LINE-1 elements. However, these repeat classes were also represented in dsRNAs from cells not expressing DUX4. In contrast, pericentric human satellite II (HSATII) repeats formed a class of dsRNA specific to the DUX4 expressing cells. Further investigation revealed that DUX4 can initiate the bidirectional transcription of normally heterochromatin-silenced HSATII repeats. DUX4-induced HSATII RNAs co-localized with DUX4-induced nuclear dsRNA foci and with intranuclear aggregation of EIF4A3 and ADAR1. Finally, gapmer-mediated knockdown of HSATII transcripts depleted DUX4-induced intranuclear ribonucleoprotein aggregates and decreased DUX4-induced cell death, suggesting that HSATII-formed dsRNAs contribute to DUX4 toxicity.

Introduction

The Double Homeobox 4 (DUX4) transcription factor is normally expressed in the testis, likely the germline cells (1), and in the cleavage-stage embryo coincident with embryonic genome

activation (EGA) (2–5). In contrast, the misexpression of DUX4 in skeletal muscle causes facioscapulohumeral muscular dystrophy (FSHD) (1,6,7). In both embryonic stem cells and in skeletal muscle cells, expression of DUX4 activates the transcription of hundreds of genes that are characteristically expressed in the

Received: August 12, 2019. Revised: September 19, 2019. Accepted: October 3, 2019

© The Author(s) 2019. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

cleavage-stage embryo. DUX4 also activates the expression of normally silenced repetitive elements that are transcribed during EGA such as endogenous retroviruses (ERVs) and the pericentric human satellite II (HSATII) repeats (3,4,8,9). The expression of DUX4 in skeletal muscle cells is toxic and causes apoptosis (10–13); however, the specific molecular pathways leading to cell toxicity are not fully understood. Previous studies have suggested multiple, non-mutually exclusive mechanisms for DUX4 toxicity including an increased sensitivity to oxidative stress (12,14), interference with PAX3/PAX7 (12,15) and the formation of insoluble TDP-43 nuclear aggregates due to impaired protein turnover (16).

More recently, we reported that the expression of DUX4 in skeletal muscle cells resulted in the accumulation of intranuclear foci of double-stranded RNAs (dsRNAs) (13). The accumulation of DUX4-induced dsRNAs correlated with PKR and eIF2 α phosphorylation, both proapoptotic characteristics of the cellular innate immune response typically triggered by viral dsRNAs (17). Indeed, knockdown of EIF2AK2 (PKR) or RNASEL, another mediator of the dsRNA response pathway, decreased the amount of cell death in cells overexpressing DUX4 (13). However, the composition of the DUX4-induced intranuclear dsRNA foci was unknown.

In this study, we identified the composition of the dsRNAs induced following DUX4 expression in human skeletal muscle cells. In contrast to the mainly intronic origin of endogenous human dsRNAs (18–20), we found that a large portion of DUX4-induced dsRNAs originate from intergenic, non-protein coding regions of the genome, and were highly enriched for HSATII satellite sequences. We found that DUX4 initially induces the transcription of one strand of HSATII, which forms single-stranded intranuclear foci that co-localize with EIF4A3. Subsequently, the induction of the complementary HSATII strand results in dsRNA foci that co-localize with the dsRNA-binding ADAR1 enzyme. Finally, gapmer-mediated depletion of HSATII transcripts results in the loss of dsRNA foci and corresponding ribonucleoprotein aggregates, diminishes PKR phosphorylation and decreases DUX4-induced cell death, suggesting that HSATII transcription contributes to DUX4 toxicity. Moreover, our results provide the first comprehensive study of double-stranded RNAs in non-affected and DUX4-expressing human muscle cells.

Results

Double-stranded RNA immunoprecipitation and sequencing identifies regions of dsRNA induced by DUX4

Previous studies have identified double-stranded RNAs by high-throughput sequencing of immunoprecipitated RNAs (dsRIP-seq) that were pulled down with the widely used J2 dsRNA antibody (19,21,22). To identify dsRNAs induced by DUX4, we performed dsRIP-seq on a doxycycline (DOX)-inducible DUX4 human myoblast cell line (MB135-iDUX4) with both the J2 and the independent monoclonal dsRNA-recognizing antibody, K1. After read mapping, we used the MACS2 peak-calling algorithm to compare the enrichment of the J2 or K1 dsRNA IPs to the mock IgG immunoprecipitation in order to identify dsRNA-enriched regions within each $-/+$ DOX condition (Fig. 1A). A correlation heatmap of read count data showed high similarity between the K1 and J2 antibodies within each condition (Supplementary Material, Fig. S1), indicating that both antibodies identified a largely overlapping set of dsRNAs. Therefore, we combined the J2

and K1 data to identify candidate dsRNA regions for subsequent analyses.

To describe the genomic regions producing dsRNAs induced by DUX4, we compared transcripts differentially enriched between $-$ DOX and $+$ DOX conditions. Using an absolute log₂ fold-change threshold of 4.0 and an FDR-adjusted P -value of 10^{-10} , 870 regions showed increased dsRNA enrichment in DUX4 expressing cells (see Supplementary Material, Fig. S2, for chromosomal locations) compared to 193 regions that showed decreased enrichment (Fig. 1B and Supplementary Material, Table S1), confirming a general induction of dsRNA enriched regions following DUX4 expression.

To verify that we successfully enriched for dsRNAs, we next examined dsRNA-specific RNA editing. The ADAR enzymes convert adenosine to inosine specifically in dsRNAs, which manifests as A to G mismatching in RNA-seq reads (23). We therefore examined mismatches in a strand-specific RNA-seq dataset of MB135-iDUX4 myoblasts expressing DUX4. Within DUX4-induced dsRNA regions, the A–G mismatch type accounted for $>60\%$ of all 12 possible called mismatches compared to the $\sim 22\%$ frequency of A–G mismatches called in the entire genome-mapped reads (Fig. 1C), confirming that the K1 and J2 antibodies successfully enriched for dsRNAs. We further confirmed that the A–G mismatches represented editing events rather than rare polymorphisms by topo-cloning and comparing the DNA sequence of MB135-iDUX4 cells with the RNA sequence in selected dsRNA regions (Fig. 1D and E and Supplementary Material, Fig. S3). Combined, the above results validate our dsRNA IP strategy and indicate, as previously observed (13), that DUX4 expression leads to the induction of dsRNA transcripts that are subject to dsRNA-specific ADAR editing.

DUX4-induced double-stranded RNAs are enriched for non-coding intergenic RNAs

We next determined the annotated genomic features associated with regions of DUX4-induced dsRNAs. Constitutive dsRNAs, which were expressed in the presence or absence of DUX4 expression (Fig. 2A), recapitulated the known intronic enrichment as well as slight 3-prime UTR enrichment for A-I edited dsRNA (Fig. 2B) (18–20). Conversely, DUX4-induced dsRNAs had a profile shifted largely toward intergenic regions (Fig. 2B), either upstream or downstream of the annotated gene, such as in the example shown upstream of TP53BP2 (Fig. 2C) or previous examples downstream of DUXA (Fig. 1D) and ZNF679 (Fig. 1E), or embedded within larger DUX4-induced long non-coding intergenic transcripts that extended upwards of 500 kb in length (Fig. 2D and Supplementary Material, Fig. S4).

Based on previously published ChIP-seq (8), many of these transcripts showed evidence for binding by DUX4 at endogenous retroviral LTR elements near the beginning of the transcript. Indeed, DUX4 bound significantly closer ($P = 6.83 \times 10^{-10}$) to DUX4-induced dsRNA regions than constitutive dsRNA regions (Supplementary Material, Fig. S5), suggesting a direct activation for many of the identified DUX4-induced dsRNAs.

Collectively, these data indicate that, in contrast to the largely intronic origin of dsRNAs in mammalian cells, a substantial portion of DUX4-induced dsRNAs are comprised of sequences embedded within non-coding, intergenic DUX4-induced transcripts. At least some of these transcripts appear to be directly activated by DUX4 as they originate in close proximity to DUX4 binding sites.

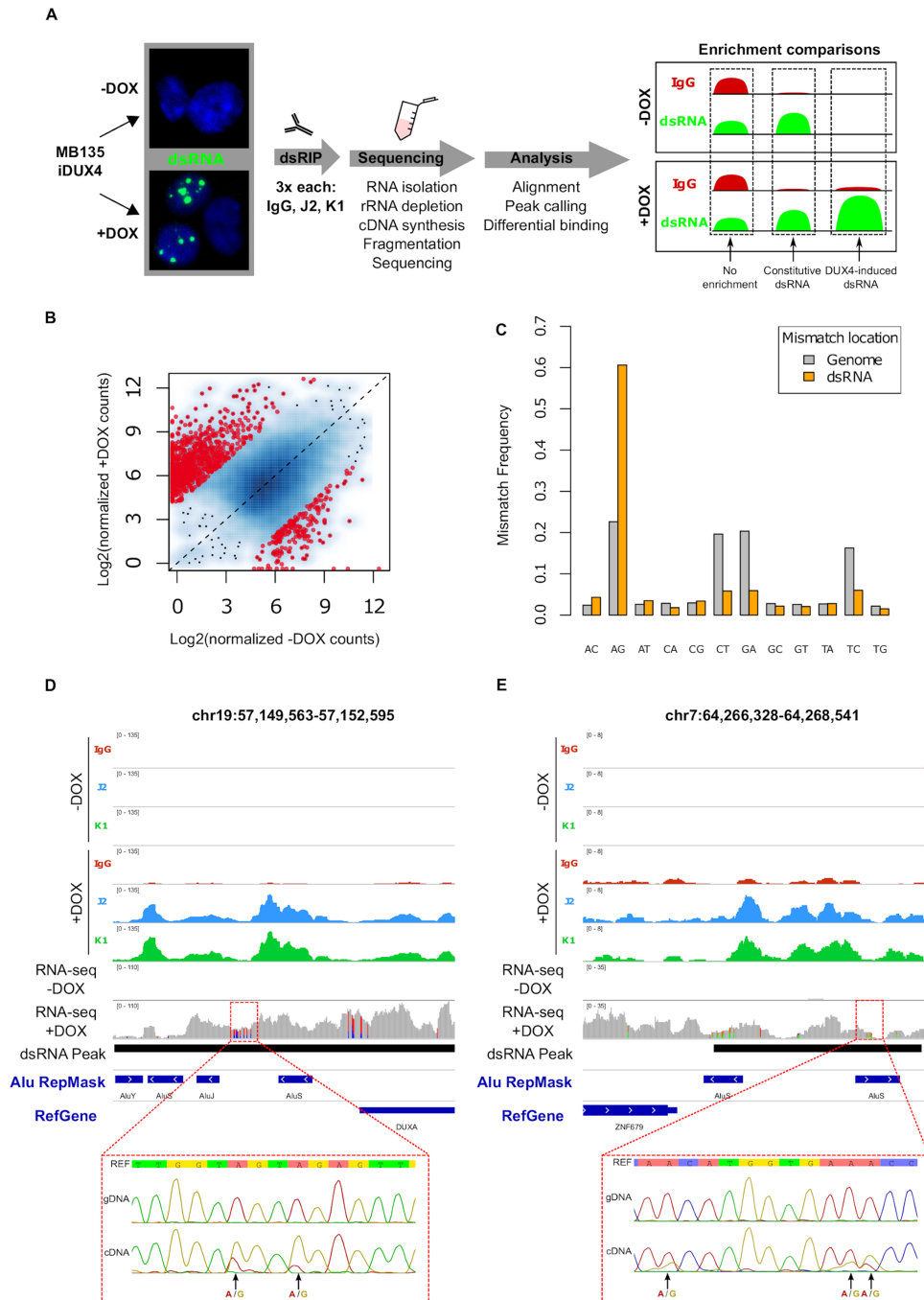


Figure 1. Double-stranded RNA immunoprecipitation and sequencing of DUX4-induced dsRNAs enriches for A-I edited transcripts. **(A)** Schematic of experimental outline for dsRNA immunoprecipitation and sequencing. Briefly, $-/+$ DOX-induced MB135-iDUX4 cells were lysed and subjected to immunoprecipitation of either matched antibody isotype control (IgG) or J2 or K1 antibodies, performed in triplicate. Libraries were constructed from isolated RNA and informatics analysis consisted of broad peak calling followed by differential enrichment testing. Within each ($-/+$ DOX) condition, dsRNA 'enrichment' was determined against the IgG control, which essentially represents total RNA abundance. Comparison of enrichment between conditions provides a way to measure level of induction following DUX4 expression. Three separate cultures of MB135-iDUX4 cells were used per condition. **(B)** Scatterplot comparing normalized levels of dsRNAs in $-/+$ DOX conditions, depicted as the log2 counts plus a pseudocount value of 1.0. Red points highlight dsRNAs that show evidence for statistically significant differential enrichment and are limited to values with a threshold absolute log2-fold change of 4.0 and FDR-adjusted P-value of 1×10^{-10} . **(C)** Mismatch frequencies of each of the 12 possible mismatch types within DUX4-induced dsRNAs in the $+DOX$ -stranded RNA-seq dataset within the indicated locations. **(D and E)** Browser screenshots in the indicated hg38 regions of normalized read counts for $-/+$ DOX immunoprecipitations. These tracks represent merged read counts for the indicated triplicate datasets. 'RNA-seq $-/+$ DOX' refers to the stranded RNA-seq dataset and is shown as raw read counts such that mismatches (highlighted in non-gray colors) can be visualized. DUX4-induced dsRNA regions are depicted by black bars. Sanger sequencing results of amplicons designed across the annotated regions (red dashed lines) in either genomic DNA (gDNA) or complementary DNA (cDNA) of $+DOX$ MB135-iDUX4 cells, with AG mismatches, highlighted. Here, the reference strand is shown as the strand of origin for the transcripts, based on the stranded RNA-seq dataset.

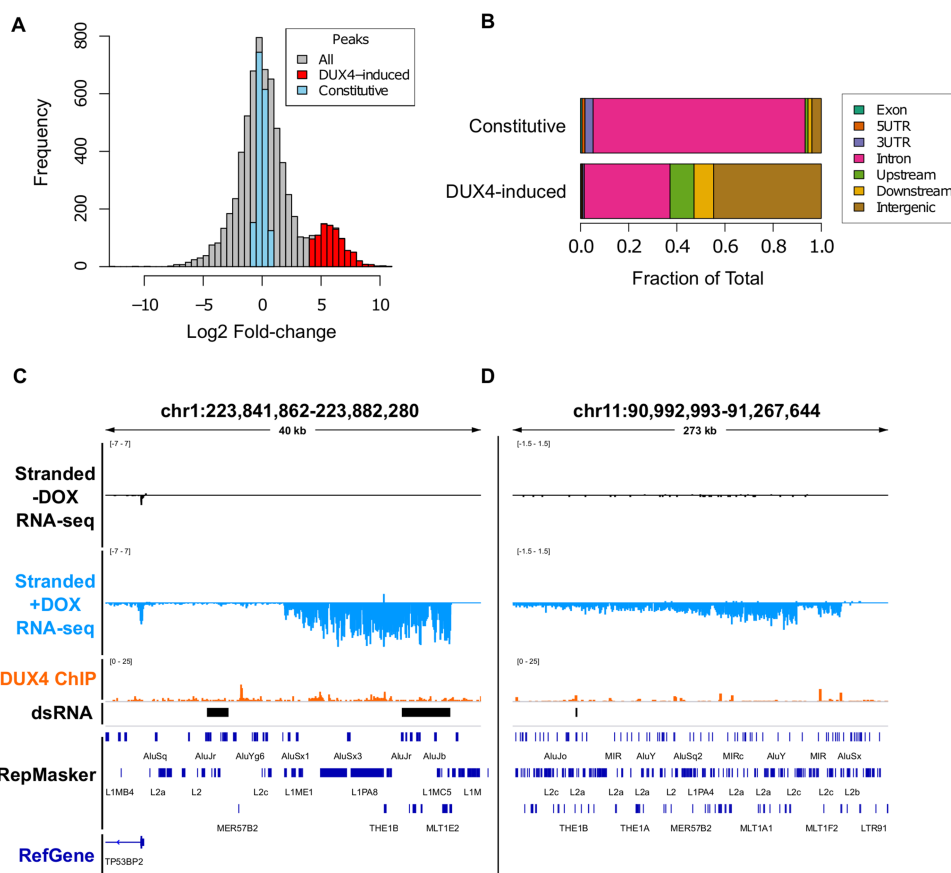


Figure 2. DUX4-induced double-stranded RNAs are enriched for non-coding intergenic RNAs. (A) Histogram of consensus 'peaks' from DiffBind analysis of dsRIP-seq experiment with constitutive peaks (blue, $|\log_2 \text{fold change}| < 1.0$ and FDR-adjusted P -value > 0.05), and DUX4-increased (red, $\log_2 \text{fold change} > 4.0$ and FDR-adjusted P -value $\leq 1 \times 10^{-10}$). (B) Genomic feature distribution overlap of each dsRNA, defined as in Figure 2A. See Materials and Methods for feature definitions and prioritizations. (C and D) Browser screenshots in the indicated hg38 regions of normalized stranded ribonucleic acid (RNA) MB135-iDUX4 RNA-seq data in two intergenic locations. These tracks represent merged read counts for the indicated duplicate datasets.

DUX4-induced dsRNAs are enriched for repeat sequences including *Alu*, LINE-1, HERVL/MaLR and the pericentric HSATII repeat

Transcripts that contain repetitive sequences are known to form dsRNAs. For example, *Alu* SINE repeats are vastly overrepresented in human dsRNAs (18) where they often form inverted, hybridized pairs within a single RNA transcript (20,24). LTR-containing ERVs have also been shown to form dsRNAs in tumor cells treated with 5-azacytidine (25,26). To determine the set of repeats activated by DUX4, we used the Dfam database of repetitive elements (27), which identified the classes of repeats induced in our stranded total RNA sequencing dataset (Fig. 3A and Supplementary Material Table S2). This analysis confirmed the previously published observation that DUX4 activates expression of HSATII pericentric satellite repeats, MaLR sequences such as THE1D, HERVL and a subset of LINE-1 repeats (3,4,8,9,28).

We next determined repeat subfamilies that were enriched by the K1 or J2 antibodies compared to the IgG control in DOX treated MB135-iDUX4i cells (Supplementary Material, Table S3). Intersecting all dsRNA enriched repeats with DUX4-induced repeats identified the set of repeats that were both strongly induced by DUX4 and were also enriched for dsRNAs (red dots, Fig. 3B and C and Supplementary Material, Fig. S6A and B). Validating this analysis, *Alu* repeat subfamilies were enriched in the dsRIP-seq datasets compared to the IgG control, though, as

expected, they were not specifically upregulated following DUX4 expression (Supplementary Material, Fig. S7A). LINE-1, ERVL-MaLR and HERVL repeats were also modestly enriched in the K1 and J2 RIPs (Supplementary Material, Fig. S7B and C and Supplementary Material, Table S3) and showed increased expression by DUX4 (Supplementary Material Table S2). However, the most highly induced repeat by DUX4 that was also enriched for dsRNA was the HSATII class of pericentric satellite repeats (Fig. 3A and C and Supplementary Material, Fig. S6A and B).

Compared to non-DUX4-expressing cells where 0.00–0.01% of dsRIP-seq reads mapped to HSATII, $\sim 4.1\%$ of K1 dsRIP-seq reads from DUX4-expressing myoblasts aligned to HSATII repeats. For unknown reasons, the J2 antibody showed more modest affinity than K1 for HSATII (compare Fig. 3B and Supplementary Material, Fig. S6A), though these two antibodies do have differences in dsRNA epitope preferences (29). Intriguingly, HSATII repeat transcripts are known to form intranuclear foci after aberrant de-repression in certain cancer cells, appearing as distinct spots (30), reminiscent of our DUX4-induced dsRNA foci. For these reasons, we next focused on determining whether DUX4-induced intranuclear dsRNA foci were comprised of HSATII transcripts.

DUX4 binds to and activates HSATII transcription

Satellite repeats can be transcribed during certain cellular stresses such as heat shock (31). However, our previously

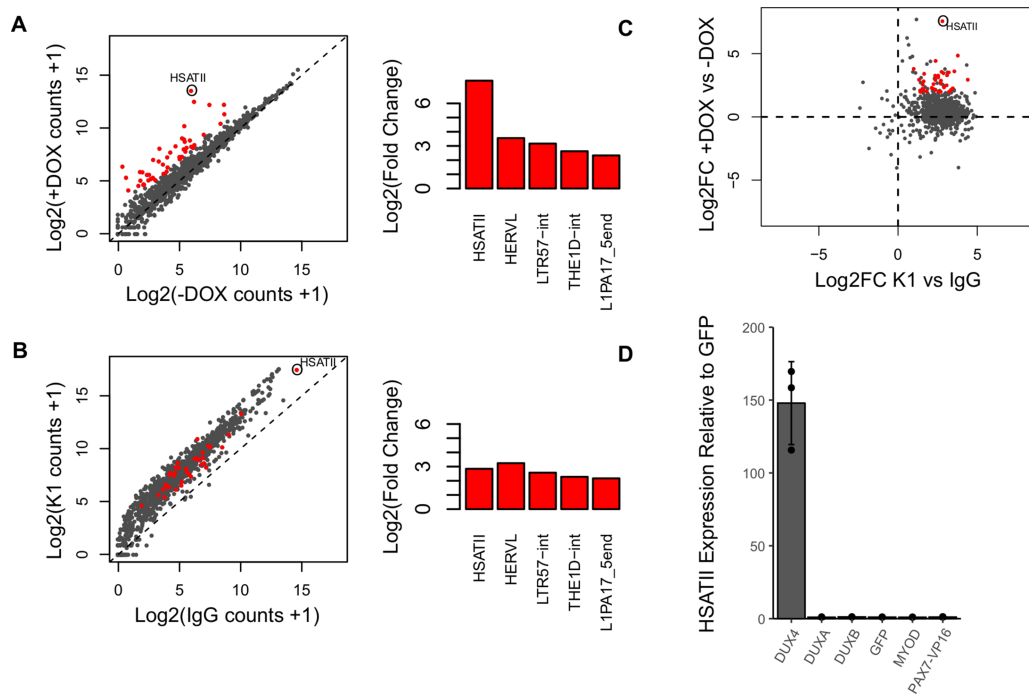


Figure 3. DUX4-induced dsRNAs are enriched for repeat sequences including HSATII. (A) Scatterplot depicting log₂ normalized read counts of Dfam predicted repeat class subfamilies within stranded RNA-seq dataset of cells $-/+$ DOX is shown on the left. Repeat subfamilies are highlighted as red when $|\log_2$ -fold change > 2.0 and FDR-adjusted P -value < 0.01 . DESeq2 moderated log₂-fold change values of select DUX4-induced repeat subfamilies are shown on the right. (B) Scatterplot depicting log₂ normalized read counts of Dfam predicted repeat class subfamilies within K1 versus IgG dsRIP-seq datasets in the +DOX condition. Significantly differentially expressed repeats (red dots) from (A) are considered enriched and highlighted as red in this plot if they meet the criteria of $|\log_2$ -fold change > 1.0 and FDR-adjusted P -value < 0.01 . DESeq2 moderated log₂-fold change values of select repeat subfamilies are shown on the right. (C) Scatterplot of Dfam predicted repeat class subfamilies showing DESeq2 moderated log₂-fold change values (Log₂FC) in +DOX versus $-$ DOX stranded RNA-seq (y axis) compared to moderated log₂-fold change values in K1 versus IgG RNA immunoprecipitations in the +DOX condition (x axis). Red points are highlighted as in panel B. (D) RT-qPCR showing levels of endogenous HSATII RNA expression relative to GFP following transfection of HEK293T cells with the indicated expression vectors (x axis). All indicated expression vectors were cloned into the pCS2 backbone. Error bars represent standard deviation of the mean of three separate cultures, depicted as individual points.

published ChIP-seq indicated that DUX4 directly binds HSATII repeats (3,9). The consensus HSATII sequence contains close matches to the known DUX4 binding motif (Supplementary Material, Fig. S8A) and variation within unmapped HSATII repeat sequences (9,32) creates the potential for large regions of arrayed near-perfect matches for DUX4 binding sites within pericentric regions. Indeed, Sanger sequencing of cloned HSATII Polymerase Chain Reaction (PCR) amplicons verified that these repeats often contain multiple copies of the consensus DUX4 binding motif (Supplementary Material, Fig. S8A), suggesting that DUX4 binding to the repeats might directly induce HSATII transcription, leading to dsRNA formation.

To more directly test whether DUX4 can activate HSATII transcription, we cloned a multicopy HSATII repeat that contained DUX4-binding motifs into the promoterless pGL3 luciferase reporter vector in both the forward and reverse orientation, relative to the consensus HSATII sequence (Supplementary Material, Fig. S8B). Co-transfection of HEK293T cells with this HSATII-luciferase reporter and a DUX4 expression vector or control GFP expression vector showed robust activation of luciferase RNA expression by DUX4, which was dependent on the presence of the HSATII sequence but largely independent of the orientation (Supplementary Material, Fig. S8C). Further demonstrating the specificity of HSATII activation by DUX4, transcriptional upregulation of endogenous HSATII was observed following overexpression of DUX4 but was not observed after overexpression of the related double homeobox proteins, DUXA or DUXB or other transcription factors MYOD or PAX7 fused to

the VP16 (PAX7-VP16) transactivation domain (Fig. 3D). These results provide additional evidence for the specificity of the binding and transcriptional activation of HSATII repeats by DUX4.

Temporally controlled bidirectional HSATII transcription forms DUX4-induced nuclear dsRNA foci

Analysis of the stranded DUX4-induced RNA-seq data demonstrated that, although reads predominantly mapped to the reverse complement of the consensus HSATII sequence (hereafter referred to as 'reverse' HSATII transcripts), a minor proportion of 'forward' transcripts were also evident (Fig. 4A). As a direct tandem repeat, unidirectional transcription of HSATII would not be predicted to lead to the formation of intramolecular dsRNA. Moreover, HSATII repeats are known to be bidirectionally transcribed in the context of some tumor cells (33). For these reasons, we postulated that bidirectional HSATII transcription might lead to the accumulation of intermolecularly formed HSATII dsRNA foci in DUX4 expressing cells and that transcription of the forward strand might be the rate-limiting step for dsRNA formation.

To determine the relationship of forward and reverse HSATII transcripts to the intranuclear dsRNA foci, we performed RNA fluorescence *in situ* hybridization (RNA-FISH) with probes detecting the forward or reverse HSATII transcripts (30) in a time course following DOX induction of MB135-iDUX4 cells. The probes to either the forward or the reverse HSATII transcripts revealed

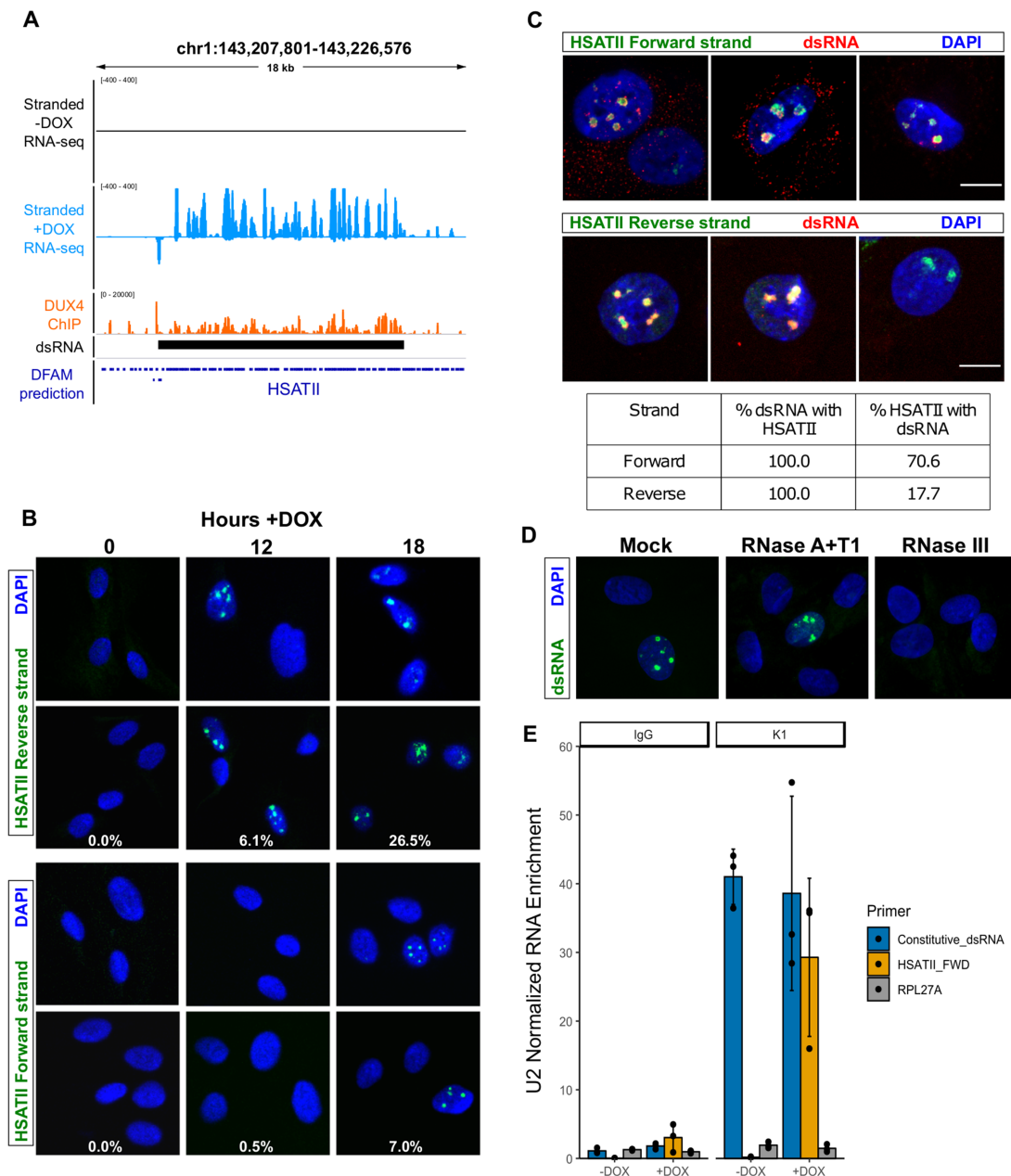


Figure 4. DUX4-induced HSATII transcripts are bidirectionally transcribed and form dsRNA. (A) Browser screenshot in the indicated hg38 region of normalized stranded ribominus MB135-iDUX4 RNA-seq data. Forward and reverse read counts are indicated above and below the horizontal line, respectively. Shown at the bottom are Dfam predicted HSATII repeat locations. (B) Confocal images of RNA fluorescence in situ hybridization of MB135-iDUX4 cells treated with doxycycline for the indicated times and hybridized with probes that detect the indicated HSATII strands, where forward and reverse is relative to the consensus sequence. Images are representative from two independent time course experiments conducted on separate days. Estimates for percent of foci positive nuclei are indicated for each time point and are from more than 150 random nuclei counted per time point. (C) Combined immunofluorescence in MB135-iDUX4 cells at 18 h +DOX, using the K1 dsRNA antibody and RNA-FISH with probe detecting forward (upper) or reverse (lower) HSATII transcripts. Images are representative from two independent, combined IF RNA-FISH experiments conducted on separate days. Note that with the reverse probe a larger subset of DUX4 expressing cells contained HSATII foci, but no dsRNA foci (top), whereas another class contained HSATII foci that co-localized with dsRNA foci (bottom). Scale bars indicate 10 μ m. Estimates for percentages of nuclear foci co-occurrence of the indicated HSATII strand and K1 antibody are shown in table below and were made from counts of more than 250 randomly selected nuclei per time point. (D) K1 immunofluorescence in MB135-iDUX4 cells at 18 h +DOX treated with the indicated RNase enzyme or mock treated prior to the immunofluorescence. No dsRNA foci were detected after treatment with RNase III, a double-stranded RNase. Experiment was performed twice, and a representative image is shown. (E) RT-qPCR of RNAs immunoprecipitations using K1 or IgG as control in MB135-iDUX4 cells \pm DOX for \sim 18 h, as indicated. Primers were designed to detect a constitutive dsRNA from our dsRIP-seq dataset or forward HSATII after strand-specific RT-qPCR. To account for differences in total RNA levels and reverse transcription efficiency, IPs were normalized to the amount of U2 RNA, which serves as an abundant background RNA species that should not be enriched by K1 levels. Error bars represent the standard deviation of the mean for three independent immunoprecipitations, which are shown as individual data points.

intracellular foci at the 18 h time point following DOX treatment (Fig. 4B), indicating that DUX4 induces bidirectional expression of HSATII. Interestingly, the probe targeting the reverse HSATII

transcript identified foci in a substantial percentage of nuclei at both 12 and 18 h following DUX4-induction, whereas the probe to the forward transcript identified obvious foci only

at the 18 h time point and in a smaller proportion of nuclei (Fig. 4B).

Next, we performed immunofluorescence using the K1 antibody combined with HSATII RNA-FISH to determine whether these HSATII transcripts might overlap dsRNA foci. This experiment revealed that the DUX4-induced nuclear dsRNA foci nearly perfectly coincided with HSATII RNA foci (Fig. 4C). These K1-stained foci were confirmed as double stranded because they were not evident following treatment with RNase III, a double-stranded ribonuclease, but were still evident following treatment with the single stranded RNases A and T1 (Fig. 4D).

Because reverse strand HSATII foci formation was relatively more common and preceded forward strand HSATII foci formation, we predicted that forward transcripts would be more highly associated with dsRNA foci. Indeed, in a majority (~70.6%) of nuclei, the forward HSATII foci coincided with dsRNA, whereas the reverse HSATII foci were more frequent and coincided with K1-positive dsRNA aggregates in a minority of nuclei (~17.7%) (Fig. 4C). Because all K1 dsRNA foci showed positive FISH signal for both forward and reverse HSATII RNA strands, our data indicate that DUX4-induced nuclear dsRNA aggregates were formed via bidirectional transcription of HSATII repeats, with the less abundant forward HSATII strand acting as a limiting factor for dsRNA formation.

To independently confirm that transcription of the forward HSATII strand is the rate-limiting step in the formation of DUX4-induced dsRNA foci, we performed strand-specific RT-qPCR following K1 dsRIP. This experiment revealed that forward HSATII transcripts were highly enriched by K1 compared to IgG in DUX4 expressing cells (Fig. 4E), whereas the enrichment of HSATII in a non-strand-specific RT-qPCR was more modest (Supplementary Material, Fig. S9). This result is expected if forward HSATII transcripts are limiting for the formation of dsRNA. Combined, our results strongly suggest that DUX4-induced intranuclear dsRNA foci are comprised of temporally controlled bidirectional HSATII repeat transcripts that form upon expression of the forward HSATII strand.

HSATII transcription is associated with EIF4A3 and ADAR1 aggregation and correlates with the formation of intranuclear BMI1 foci and phosphorylation of H2AX

We next sought to determine whether nuclear proteins may co-aggregate as a consequence of HSATII-containing dsRNA foci formation in DUX4 expressing cells. We previously showed that the exon junction complex factor EIF4A3 accumulates in intranuclear foci following DUX4 induction (13). In this prior study, EIF4A3 aggregates appeared earlier than dsRNA foci, but nearly all dsRNA foci co-localized with EIF4A3 foci upon their appearance. We therefore performed combined immunofluorescence and RNA-FISH of DUX4-induced cells with an antibody to EIF4A3 and the FISH probe to the HSATII reverse strand. This showed strong co-localization of EIF4A3 and reverse HSATII foci (Supplementary Material, Fig. S10A), indicating that EIF4A3 accumulates with reverse HSATII transcripts prior to the appearance of forward HSATII transcripts and the formation of dsRNA.

Although we cannot directly assess ADAR editing of HSATII repeats due to polymorphic variations from the consensus sequence (9), FISH combined with an antibody to ADAR1 indicated a strong redistribution of ADAR1 (which specifically binds dsRNA) to HSATII foci in DUX4 expressing cells (Supplementary Material, Fig. S10B). Nearly 100% of these ADAR1 foci also co-localized with K1 dsRNA nuclear-staining (Supplementary Material, Fig. S10C), suggesting that DUX4-induced HSATII

dsRNA is bound by ADAR1 and leads to ADAR1 accumulation as intranuclear foci. Cells with either EIF4A3 or ADAR1 foci had lost their normal pan-nuclear distribution, indicating a possible sequestration of these proteins in the foci.

We also wondered about other consequences of HSATII dsRNA formation in DUX4 expressing cells. Derepression of HSATII repeats in cancer cells has been associated with the formation of polycomb bodies (30,34) and with double-strand DNA breaks (35). In our model system, DUX4 induction of HSATII repeats was similarly associated with the formation of polycomb bodies and foci of DNA damage, as determined by immunodetection of the PRC1 complex protein BMI1 and foci of phospho-H2AX (Supplementary Material, Fig. S11A). The BMI1 foci did not co-localize with induced HSATII RNA foci (Supplementary Material, Fig. S11B) but did show strong co-localization with the phospho-H2AX foci suggesting that the PRC1 complex was accumulating at sites of DNA damage, although it is possible that they might also be associated with demethylated HSATII DNA regions in the genome as has been suggested (30).

HSATII transcripts are necessary for DUX4-induced dsRNA foci and contribute to toxicity

Locked nucleic acid gapmer oligonucleotides can degrade cellular transcripts via RNase H-dependent cleavage of complementary transcripts following hybridization, allowing for effective knockdown of nuclear transcripts, including satellite repeat RNAs (36). To determine whether DUX4-induced HSATII transcripts are necessary for dsRNA formation, we transfected MB135-iDUX4 cells with two separate pooled HSATII-targeted gapmers designed against both the forward and reverse HSATII RNA strands prior to DOX induction of DUX4. Our results demonstrated that transfection of HSATII-gapmers can deplete DUX4-induced HSATII transcripts as compared to a non-targeting control gapmer (Fig. 5A), with no discernible effect on DUX4 expression (Supplementary Material, Fig. S12) or DUX4 target gene expression (Fig. 5B). The F1R2 gapmer pool was more effective at depleting HSATII transcripts than the F2R1 pool (Fig. 5A). Importantly, gapmer-mediated HSATII knockdown diminished the formation of dsRNA and EIF4A3 foci following DUX4 expression (Fig. 5C), particularly in the case of the more effective F1R2 gapmer pool. These results indicate that HSATII expression is necessary for the formation of DUX4-induced intranuclear dsRNA and EIF4A3 foci.

Our prior observation that activation of the dsRNA response contributed to DUX4-induced toxicity led us to hypothesize that HSATII transcription might be toxic to DUX4 expressing cells. Therefore, we next tested whether gapmer-mediated HSATII knockdown could decrease DUX4 toxicity. Indeed, as compared to a control gapmer, transfection of gapmers targeting HSATII forward and reverse transcripts enhanced survival following DUX4 expression when assessing the percentage of live cells compared to no DUX4 expression (Fig. 5D), particularly with the F1R2 combination. HSATII gapmers also appeared to moderately retard myoblast proliferation leading to decreased numbers of total cells in the non-DUX4 expressing conditions (Fig. 5E). This rescue of cell toxicity was similar to our previous report that knockdown of PKR or eIF2 α partially rescued DUX4 toxicity (13). Importantly, gapmer transfection also diminished DUX4-induced, pro-apoptotic PKR phosphorylation (Fig. 5F), most notably with the F1R2 gapmer combination, which occurs upon detection of dsRNAs by PKR. In addition, F1R2 and F2R1 gapmer combinations also diminished the increase of eIF2 α

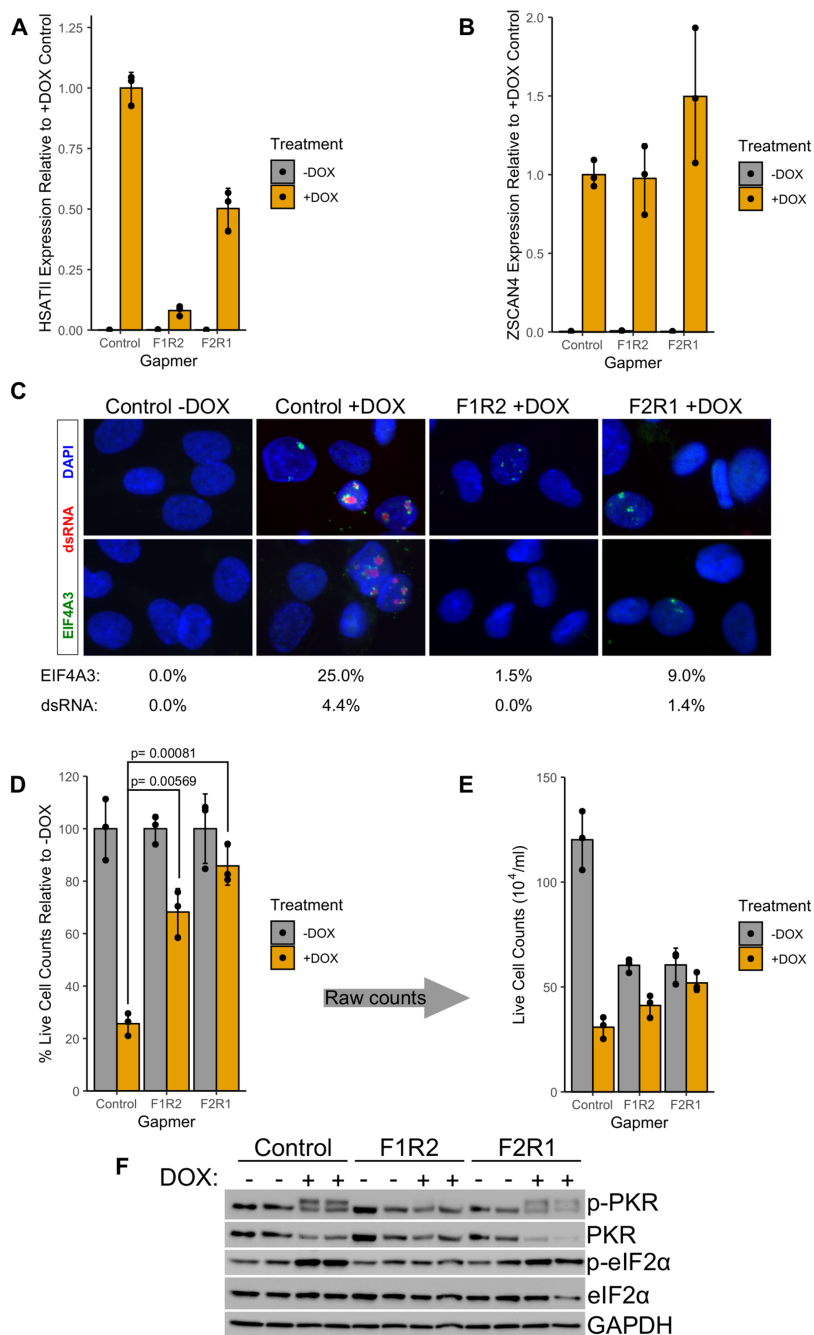


Figure 5. Gapmer-mediated knockdown of HSATII transcripts confirms their contribution to dsRNA foci and increases cell survival following DUX4 expression. (A) RT-qPCR showing levels of total HSATII RNA at 24 h relative to the +DOX condition, normalized to RPL27A. Error bars represent the standard deviation of the mean for three independently cultured samples, shown as individual data points. (B) RT-qPCR showing levels of total ZSCAN4 (direct DUX4 transcriptional target) RNA at 24 h relative to the +DOX condition, normalized to RPL27A. Error bars represent the standard deviation of the mean for three independently cultured samples, shown as individual data points. (C) Representative examples of K1 (dsRNA) and EIF4A3 immunofluorescence images in MB135-iDUX4 cells at 18 h $-/+$ DOX. Percentages of K1 or EIF4A3 foci-positive nuclei in the indicated conditions are based on counts from more than 500 randomly selected nuclei performed blinded to the experimental condition. Experiment was performed three times. (D) Trypan blue exclusion-based live cell counts in gapmer-transfected samples at 24 h following DUX4 induction. The data are presented as the percentage of each gapmer's paired $-$ DOX sample for each time point. P values were calculated using two-sided Welch's T test. No multiple-testing adjustment was used. Each dot represents the average relative live cell count from an independently cultured sample and error bars depict standard deviation from the mean. (E) Raw live cell counts used to make Figure 5D that shows the reduction in live cell numbers due to HSATII knockdown (compare gray bars). Each dot represents the average live cell count from an independently cultured sample and error bars depict standard deviation from the mean. (F) Western blot showing levels of phosphorylated and unphosphorylated PKR and eIF2 α in the indicated conditions as in Figure 5D. Each lane is an independently transfected sample. GAPDH serves as a sample loading control.

phosphorylation that occurs in DUX4 expressing cells. As eIF2 α is a phosphorylation target of activated PKR, this result confirms the reduction of PKR activity in DUX4 expressing cells following

HSATII depletion. In sum, these data indicate that DUX4-induced HSATII transcripts are necessary for the observed intranuclear dsRNA foci formation and that depletion of HSATII transcripts

decreases dsRNA-induced PKR phosphorylation and cell death following overexpression of DUX4.

Endogenous DUX4 in FSHD muscle cells induces dsRNA-forming transcripts including HSATII

In cultures of FSHD skeletal muscle cells, DUX4 is expressed in only a minor fraction of the muscle nuclei (1). To determine whether the expression of endogenous DUX4 in FSHD muscle resulted in transcriptional upregulation of the identified dsRNA regions, we used RT-qPCR and found expression of genic as well as intergenic RNAs in the DUX4-induced dsRNA regions in both FSHD1 (MB073) and FSHD2 (MB200) cultured muscle cells, but not in control (MB135) muscle cells (Supplementary Material, Fig. S13). Sequencing cDNA amplicons from differentiated FSHD2 cells identified A-G mismatches compared to the gDNA, indicating ADAR editing and dsRNA formation of these transcripts in FSHD myotubes (Supplementary Material, Fig. S14).

We also found elevated HSATII expression in differentiated FSHD myotubes compared to a control non-FSHD cell line (Fig. 6A). Consistent with our previous demonstration that a small proportion of FSHD muscle cells formed intranuclear dsRNA foci (13), RNA-FISH revealed rare intranuclear foci of HSATII using the probe to the reverse strand in FSHD myotubes, but not in control cells (Fig. 6B). Therefore, as with ectopic expression of DUX4 in MB135-iDUX4 cells, endogenous DUX4 expression in FSHD myotubes also led to dsRNA expression and intranuclear foci of HSATII RNA.

DUX4-induced dsRNA-forming RNAs including HSATII are expressed in pre-implantation embryos

Recent studies indicated that DUX4 and its mouse ortholog Dux activate a subset of the early cleavage-stage transcription program in human and mouse embryos (2–4). In humans, EGA involves the expression of HSATII and HERVL retrotransposons and in mice involves the expression of MERVL and GSAT major pericentric satellite repeats. To determine whether cleavage-stage expression of DUX4 correlated with dsRNA formation, we re-analyzed published early embryo RNA-seq data (3) across several different stages of development. As previously reported, we identified DUX4 expression at the cleavage stage of embryonic development and this correlated with increased expression of many of the same DUX4-induced dsRNA-forming regions in myoblasts, including HSATII (Fig. 6C and Supplementary Material, Table S4). Analysis of base mismatches between the RNA-seq and the reference genome sequence revealed a dramatic enrichment of A-G mismatches over other mismatch types across DUX4-induced dsRNA regions (Fig. 6D). Interestingly, both forward and reverse strand transcripts of HSATII were at near equal levels in the cleavage stage (Supplementary Material, Fig. S15), perhaps because the samples contained a heterogeneous pool of embryos at different cleavage cell numbers. Thus, our analysis verified that many of the DUX4-induced dsRNA-containing transcripts identified in DUX4 expressing myoblasts, including HSATII, are also present in the early embryo.

Discussion

When DUX4 is expressed in induced pluripotent stem cells or skeletal muscle cells, it activates genes characteristic of a totipotent developmental expression program (2–4). For

example, DUX4 activates the expression of ZSCAN4 that has roles in both telomere elongation (37) and DNA demethylation through the degradation of UHRF1 and DNMT1 (38), as well as KDM4E, which is a lysine demethylase that can relieve chromatin repression (39). DUX4 also activates the expression of repetitive elements, including LINE-1, HERVL and MaLR ERVs and HSATII satellite repeats (3,4,8,9,28). In this study, we show that DUX4 creates intergenically derived dsRNAs from a subset of these repeats. Most dramatically, bidirectional transcription of HSATII repeats results in the formation of intranuclear dsRNA foci and gapmer knockdown of the HSATII RNAs partially rescues DUX4 toxicity together with decreasing the DUX4-induced phosphorylation of PKR and eIF4 α . These results demonstrate that the toxic dsRNA response induced by DUX4 is mostly mediated by the induction of HSATII transcripts. Although we used forced expression of DUX4 in skeletal muscle cells to identify these dsRNA-forming transcripts, analysis of RNA and RNA-seq showed that these dsRNA-forming regions were also transcribed and ADAR-edited in FSHD muscle cells and cleavage-stage embryos expressing the endogenous DUX4. Along with our previous results (13), these findings solidify a mechanism of dsRNA formation in FSHD that can contribute to cell toxicity.

In mice, during the cleavage stage of early embryogenesis, pericentric 'major satellite' repeats are expressed first in the forward, then reverse direction (relative to the consensus sequence) where they have been hypothesized to form dsRNAs that direct repressive heterochromatin modifications to the pericentromeres (36,40). However, the mechanism of regulating these transcripts remained unknown. Gapmer-mediated depletion of these repeat transcripts blocked chromocenter formation and interrupted development at the four-cell stage, indicating a necessary role for satellite RNAs in establishing pericentric heterochromatin (36). While transcription of satellite repeats appears to be necessary for heterochromatin establishment, whether dsRNA is required is less clear (41). As with DUX4 in human embryogenesis, mouse Dux is expressed in the cleavage stage and, interestingly, is coincident with transcription of GSAT (major satellite) pericentric repeats (3,4). In the case of HSATII, our study shows that DUX4 appears to first initiate HSATII transcription in predominantly one direction, producing an RNA encoding the reverse complement of the consensus HSATII sequence, and then later induces HSATII transcription in the opposite direction coincident with the formation of multiple distinct intranuclear dsRNA foci. Given that this bidirectional expression pattern of HSATII by DUX4 parallels the temporal dynamics of mouse expression of GSAT repeats, we postulate that DUX4, and by analogy mouse Dux, has a primary role in regulating the bidirectional transcription of pericentric satellite repeats in early human and mouse embryogenesis.

In contrast to its potential function in the early embryo, the expression of HSATII RNA in FSHD muscle could have both direct and indirect biological consequences that could alter RNA processing and chromatin repression. Previously we reported that DUX4 expression induces intranuclear focal aggregates of EIF4A3 (13), and here we showed that EIF4A3 accumulates with nuclear HSATII RNA foci. Although it remains unclear whether the association of EIF4A3 with HSATII RNA is due to splicing of the HSATII transcripts or another recruitment mechanism, the formation of EIF4A3 foci correlates with the inhibition of nonsense-mediated RNA decay (NMD). This suggests that the sequestration of EIF4A3 might contribute to DUX4-mediated inhibition of NMD and could act in concert with the DUX4-mediated depletion of UPF1 (13,42). Similar to the association of EIF4A3 with the predominantly expressed reverse HSATII strand, the association of ADAR1 foci

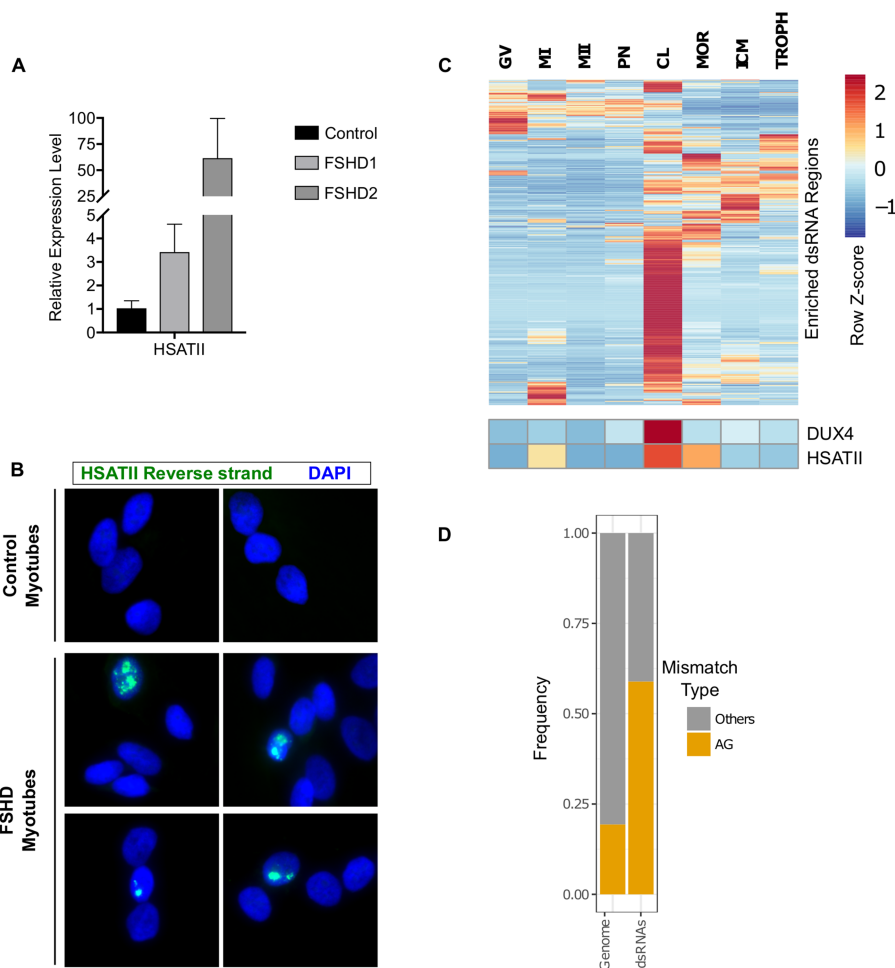


Figure 6. Endogenously expressed DUX4 induces dsRNA forming transcripts, including HSATII. (A) RT-qPCR data showing transcript expression levels of HSATII in control (MB135), FSHD1 (MB073) or FSHD2 (MB200) cells grown in differentiation medium. Data are normalized to RPL27 levels and shown relative to the control cell line. Data are depicted as the mean values of three experiments performed on independent days. Error bars represent the standard deviation of the mean. (B) RNA-FISH using probes targeting the reverse HSATII transcript in control (MB135) or FSHD (MB073) differentiated myotubes. Consistent with rare DUX4 expression in FSHD muscle cells, we observed a small subset of HSATII-positive nuclei in the FSHD cells, but not control cells. Images are representative from two independent experiments conducted on separate days. (C) Heatmap of DUX4-induced dsRNA, DUX4 and HSATII expression across various stages of human embryo development. Read counts were normalized and depicted as the within row z score. Data are from Hendrickson et al. (3). (D) Frequency plot of AG compared to all other possible mismatches in cleavage stage samples from Hendrickson et al. (3). Whole genome refers to all mismatches called, regardless of location and dsRNAs are mismatches called within DUX4-induced dsRNA regions. GV, germinal vesicle; MI, metaphase I; MII, metaphase II; PN, pronuclear stage; CL, cleavage stage; MOR, morula; ICM, inner cell mass; TROPH, trophectoderm.

with HSATII dsRNA could also contribute to toxicity in DUX4-expressing cells. For example, ADAR1 aggregation might limit the enzyme's activity on non-HSATII containing endogenous dsRNAs, which would be expected to lead to activation of cellular antiviral pathways. In accordance with this, ADAR1 knockout leads to dsRNA-associated apoptosis in a human cell line that can be rescued by simultaneous knockout of RNASEL (43), paralleling our previous finding that DUX4 toxicity can be alleviated by knockdown of RNASEL (13). Future work is required to determine whether the HSATII ribonucleoprotein foci sufficiently sequester EIF4A3 and ADAR1 to prevent their normal cellular function.

DUX4-mediated induction of HSATII RNA might have other consequences as well. Because of prior reports that HSATII transcription has been associated with DNA instability and the formation of intranuclear foci of components of the PRC1 complex (polycomb bodies), which can potentially lead to de-repression of LINE-1 elements or other repeats (30), we investigated whether the induction of HSATII transcripts correlated with the forma-

tion of polycomb bodies or evidence of DNA breaks. Consistent with these prior reports, DUX4-induced HSATII expression correlated with the formation of polycomb bodies, as indicated by the evidence of BMI1 foci. The co-localization of the BMI1 foci with phospho-H2AX foci suggests that the DNA breaks might be the driver of the location of the BMI1 foci, although further work will be necessary to determine whether the DNA damage is a consequence of HSATII repeat expression. In this regard it is interesting to note that, in mice, expression of pericentric major satellite repeats can lead to genome instability, likely by destabilizing DNA replication forks (35). This causes activation of the DNA damage response and cell death, which can be alleviated by knockout of p53. Whether HSATII expression causes these effects in human cells is currently unknown, but its demonstrated activation in tumor cell lines (33,44) suggests that its expression may similarly trigger genome instability.

A possible function for non-HSATII intergenic dsRNAs induced by DUX4 remains more speculative. It is interesting to note the relative enrichment of LINE-1 and ERV repeats within

many of the intergenic DUX4-induced dsRNAs, classes of repeats that are expressed in the cleavage-stage embryo, but generally repressed in most tissues (45). Similar to the proposed role of HSATII repeats, the double-stranded secondary structure of these transcripts might also facilitate heterochromatin formation. This has been postulated as a function of dsRNA-forming lncRNAs, which can recruit methyltransferase activity to DNA (46). Interestingly as well, the DUX4-induced long intergenic transcripts often contain a region of dsRNA enrichment and are reminiscent of the previously characterized 'very long intergenic non-coding' (vlinc) RNAs observed in a multitude of pluripotent cells and also cancer lines (47,48), though the nature of the transcriptional regulation of these vlincRNAs is presently not well understood.

The enrichment of retroelement-derived dsRNAs in a subset of DUX4 target genes might also serve a function. Although retroelements are typically subject to DNA methylation and other silencing mechanisms (49), their presence can have a positive effect on gene expression at early stages of development (50–52). Silencing of these retroelements at later stages might then provide precise temporal control of DUX4 targets whose continued expression would counteract later differentiation processes. It is possible, therefore, that the dsRNAs in DUX4-induced genes might contribute to the subsequent silencing of these genes through regional dsRNA-mediated heterochromatin formation.

In summary, our work confirms the induction of dsRNA containing transcripts by DUX4. Similar to constitutively expressed dsRNAs, many of the DUX4-induced dsRNAs are enriched for inverted Alu repeats in genic transcripts. In contrast to constitutive dsRNAs, a large fraction of DUX4-induced dsRNAs are intergenic and are modestly enriched for LTR-containing ERVs and LINE-1 elements, but highly enriched for HSATII satellite repeats. DUX4 induction of HSATII repeats in a bidirectional, temporally controlled pattern mirrors the dynamics of mouse major satellite repeat expression in the early embryo and suggests that human DUX4 activation of HSATII repeat transcripts in the cleavage-stage embryo might play a similar role in establishing pericentric heterochromatin via dsRNA formation. Our results from depleting DUX4-induced HSATII transcripts indicate that HSATII dsRNAs contribute to DUX4 toxicity in somatic cells and might also contribute to DUX4 toxicity in FSHD and is consistent with our prior observation that knockdown of RNASEL or PKR also partially rescues DUX4 toxicity (13). The expression of DUX4 and bidirectional HSATII transcripts in the early embryo suggests a role for these dsRNAs in early development that merits further study.

Materials and Methods

Cell culture

De-identified human primary myoblast cell lines from the Fields Center for FSHD Neuromuscular Research at the University of Rochester Medical Center were immortalized by retroviral transduction of CDK4 and hTERT (53). All immortalized human myoblasts were cultured in F10 medium (Gibco/ThermoFisher Scientific) supplemented with 20% Fetal Bovine Serum (GE Healthcare Life Sciences) and 1% penicillin/streptomycin (Gibco/ThermoFisher Scientific), 10.0 ng/ml recombinant human FGF (Promega) and 1.0 μ M dexamethasone (Sigma-Aldrich). Myoblasts were differentiated by culturing in Dulbecco's Modified Eagle Media (Gibco/ThermoFisher Scientific) containing 1% horse serum (Gibco/ThermoFisher Scientific), 1% peni-

collin/streptomycin (Gibco/ThermoFisher Scientific), 10 μ g/ml insulin and 10 μ g/ml transferrin for 48–72 h. All '+DOX' labels mean that cells were grown for ~18 h in the presence of 1.0 μ g/ml of DOX hyclate (Sigma-Aldrich), unless otherwise noted.

Gapmer LNA transfections. Gapmers were transfected into MB135-iDUX4 cells on the day prior to DOX induction using Lipofectamine RNAiMAX reagent (ThermoFisher Scientific) following the manufacturer's instructions. Gapmers were ordered from QIAGEN. Sequences of gapmers are shown below (a '+' indicates a locked nucleic acid modification in the following base):

```
Control_gfp +g+a+g+aAAGTGTGACA+a+g+t+g
HSATII_R1 +t+g+a+tTCCATTCGATT+c+c+a+t
HSATII_R2 +c+a+t+tCGATGATTCC+a+t+t+c
HSATII_F1 +a+t+g+gAATCGAATGGA+a+t+c+a
HSATII_F2 +g+a+a+tGGAATCATCG+a+a+t+g
```

dsRNA Immunoprecipitation. Human myoblast MB135-iDUX4 cells were treated $-/+$ DOX and washed in Phosphate-Buffered Saline (PBS), trypsinized and counted prior to lysis. Approximately 1.2×10^6 cells were used for each IP. Lysis was performed in 1.0 ml total volume by sonication using a Diagenode Bioruptor on light setting (5 min total, 30 s on/off at 4°C) in a buffer composed of 15 mM Tris pH 7.5, 0.1 M NaCl, 5 mM MgCl₂, 0.5% Triton X-100, 1 mM dithiothreitol and 40 U/ml RNase inhibitor (ThermoFisher Scientific). Lysates were precleared using 40.0 μ l of protein G Dynabeads (ThermoFisher Scientific) for 1 h prior to an overnight incubation at 4°C with either J2, K1 or an isotype-matched anti-GFP (IgG) control antibody. A total of 40.0 μ l of protein G Dynabeads were added the following morning for 1 h to bind the antibody, and beads were subsequently washed four times with 1.0 ml of cold lysis buffer. After the final wash, 1.0 ml of TRIzol (ThermoFisher Scientific) was added directly to the beads for RNA extraction.

RNA extraction and library preparation. TRIzol RNA extractions were performed as per manufacturer's recommendations. Ribosomal RNA depletion was performed using the NEBNext rRNA depletion kit (New England Biolabs). First strand synthesis was achieved using SuperScript III reverse transcriptase (ThermoFisher Scientific) with random hexamers and using the following thermocycler conditions: 25°C, 10 min; 50°C, 30 min; 55°C, 30 min; and 85°C, 5 min. Second strand synthesis was achieved using the NEB second strand cDNA synthesis kit, following manufacturer's recommendations. The double-stranded cDNA was fragmented to ~250 bp average size using a Diagenode Bioruptor sonicator in 120 μ l total volume for three 10 min intervals on medium intensity at 4°C, 20 s on/off, replenishing with pre-cooled water between each interval. Libraries were made from the AmpureXP (Beckman Coulter) purified cDNA using the Ovation Ultralow system V2 kit and following the manufacturer's instructions. Stranded total RNA-seq libraries were made from 1.0 μ g of MB135-iDUX4 RNA using the KAPA-stranded RNA-seq kit with RiboErase (HMR) and following the manufacturer's instructions. Libraries were sequenced using 100 bp single-end sequencing on the Illumina HiSeq 2500 platform by the FHCRC Genomics facility.

dsRIP sequence alignment and differential peak calling. We essentially treated our dsRIP-seq data as diffuse (e.g. histone mark) ChIP-seq to find locations of enrichment. First, dsRNA-IP reads were aligned to hg38 using BWA version 0.7.12 with the option -n

6. Next, peaks were called by comparison to the IgG background control condition using MACS2 version 2.1.0 using the options: `-broad, -broad-cutoff 0.01, -q 0.01`. We then used Diffbind version 2.3.8 to differentially call dsRNA enrichment between +/- DOX conditions with 'minOverlap' set at 0.5, which treated J2 and K1 antibodies as replicates for comparison purposes. The significance threshold was set at $1e-10$. We extracted the consensus peakset from the Diffbind dba report, setting the significance threshold as 1 to capture all dsRNA-enriched loci. We filtered our list of enriched dsRNAs to eliminate dsRNAs that overlapped known rRNAs, tRNAs and snRNAs (obtained from UCSC table browser repeatMasker track filter reclass: 'snRNA OR rRNA OR tRNA OR snoRNA') using bedtools 'intersect'. We used the Integrative Genomics Viewer version 2.3.98 for data visualization. We used the 'kpPlotRegions' function from the karyoploteR package to plot location of DUX4-induced dsRNA-enriched regions.

Feature annotation. We used bedtools 'annotate' and GenCode v24 feature annotations downloaded from the UCSC genome browser to classify per-base feature overlaps of dsRNAs with 'intergenic' being defined as the absence of a feature overlap and 'downstream' and 'upstream' being defined as 5 kb 3-prime or 5-prime of an annotated gene feature, respectively. Overlapping features were prioritized as: coding exons > 5' UTR > 3' UTR > intron > upstream 5 kb > downstream 5 kb > intergenic.

Stranded RNA-seq. Stranded RNA-seq libraries were aligned using TopHat version 2.1.1 with Bowtie version 2.2.9 and the iGenomes UCSC gene transfer format file with the following options: `-read-mismatches 8 -read-edit-dist 8`.

RNA editing detection. We used the python script, REDIttools *de novo* version 1.0.4 (54) to call editing sites on our aligned samples. Replicate BAM files were merged using samtools before analysis. The following REDIttools parameters were used to limit false positives: `-E` (exclude positions with multiple changes), `-a t` (two-tailed Fisher's exact test), `-c 10` (minimum read coverage of 10), `-T 6-6` (six bases were trimmed from each end of the reads), `-q 25` (minimum quality score of 25), `-n 0.1` (minimum editing frequency of 0.1) and `-v 3` (minimum number of reads supporting the variation set at 3). The SNP147 GTF file was downloaded from the UCSC genome table browser to exclude known polymorphic sites from consideration. We used a P-value threshold value of <0.1 .

Repeat analysis. For repeat counting in our stranded, ribosome-depleted RNA-seq dataset we used the Tetranscripts script from TEToolkit package (55) with the multi-mode option and a custom GTF file that was made from the current (November 7, 2016) release of Dfam non-redundant repeat matches. This approach enabled counting within predicted HSATII repeats (which are not available in the hg38 repeatMasker annotation used). We performed differential expression analysis on the resulting count table with DESeq2 using default options. For differential expression of repeats in our dsRIP-Seq data, we performed the above analysis in K1 or J2 versus IgG in DOX-treated samples.

Confirmation of A-I editing. We amplified dsRNA regions containing putative A-I RNA editing sites using Phusion High Fidelity DNA polymerase (ThermoFisher Scientific) from genomic DNA or cDNA of DOX-induced MB135-iDUX4 cells. Purified amplicons were Sanger sequenced by the FHCRC genomics facility on a 3730xl DNA Analyzer and visualized using Geneious Pro software, version 5.0.4.

RNA isolation and real time qPCR. For immunoprecipitation samples, RNA was isolated using TRIzol reagent (ThermoFisher Scientific) according to the manufacturer's protocol and using GlycoBlue (ThermoFisher Scientific) as a co-precipitant. For all other samples, RNA was isolated with the RNeasy kit (QIAGEN), according to the manufacturer's protocol. Purified RNA was treated with DNaseI (ThermoFisher Scientific), heat inactivated and reverse transcribed into cDNA using Superscript III (ThermoFisher Scientific) following the manufacturer's instructions. Quantitative PCR was performed with SYBR green reagent (ThermoFisher Scientific). For RT-qPCR of RNA immunoprecipitation, random hexamers (ThermoFisher Scientific) were used for non-strand-specific detection. Data were calculated relative to +DOX input and normalized to U2 background RNA values for each independent immunoprecipitation. For HSATII forward strand specific RT-qPCR, we used the 'SATII_pericentric_1L' oligo to prime cDNA synthesis with a 5-prime linker ATGGATCAGGAGAAGACTGA sequence. For qPCR we used the 'SATII_pericentric_1R' and linker sequence as the primers. Samples were subsequently normalized to U2 signal in the respective random hexamer primed reactions, as above.

HSATII RNA-FISH and combined immunofluorescence RNA-FISH. Locked nucleic acid, FITC-conjugated HSATII probes were purchased from QIAGEN and are based on the sequence used in previous publications (30,56,57). Probe 1: 5' FAM-ATTCCATTCAGATTCATTCGATC detects the reverse HSATII transcript. Probe 2, which detects the forward HSATII transcript, is the reverse complement of probe 1 with 5' FAM modification. LNA positions were not disclosed. We performed HSATII RNA-FISH essentially as described previously (30) with modifications. Briefly, cells grown on coverslips were rinsed in cold Cytoskeletal (CSK) buffer (100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 10 mM Pipes pH 6.8) and then permeabilized in ice cold extraction buffer [CSK, 0.1% TX-100 and 10 mM ribonucleoside vanadyl complex (NEB)] for 5 min. After a second CSK rinse, cells were fixed in 4% paraformaldehyde (Electron Microscopy Sciences) in PBS for 10 min at room temperature. Cells were dehydrated in series of 70% ethanol, 90% ethanol and 100% ethanol and then air dried. Probes were diluted to 5.0 pmol/ml and denatured at 75°C for 5 min in whole chromosome painting (WCP) buffer (50% formamide, 2xSSC, 10% dextran sulfate) and hybridized overnight at 37°C. Coverslips were washed with 50% formamide/2xSSC at 37°C and then twice with 2xSSC/0.1% TX100 for 5 min before mounting using prolong gold antifade with DAPI (ThermoFisher Scientific).

For combined immunofluorescence and RNA-FISH we first fixed cells on coverslips using 4% paraformaldehyde for 7 min at room temperature. Cells were permeabilized with 0.5% TX100 in PBS for 5 min prior to overnight incubation with primary antibodies at 4°C. Samples were incubated with appropriate secondary antibody for 1 h at room temperature prior to fixation of the antibody interaction via a second crosslinking step using 4% paraformaldehyde for 10 min at room temperature. Cells were incubated with probes, which were denatured as above for 2 h at 37°C prior to sequential washes with 15% formamide 2xSSC for 20 min at 37°C, 2xSSC for 20 min at 37°C and 2xSSC for 5 min at room temperature. Cells were imaged using a Leica TCS SP5 II confocal microscope, where indicated, or a Zeiss AxioPhot. Image channel merging and processing was performed using ImageJ software.

Analysis of human embryo RNA-seq data. We aligned raw reads to genome build hg38 using TopHat2, allowing two mismatches per read and maximum of 20 multiple alignments. The counts

for enriched dsRNA regions were computed by summarizeOverlaps() from the GenomicAlignments Bioconductor package with 'IntersectStrict' mode. Next, we applied the countRepeats R package (<https://github.com/TapscottLab/countRepeats>), which is designed specifically for counting reads of repetitive elements originating from forward/reverse transcripts. To obtain DUX4 counts, we aligned the raw reads to the customized D4Z4 'genome' comprised of DUX4 exon sequences.

Cloning of HSATII and testing DUX4 activation. An ~1.5 kb PCR product was cloned from MB135 cell genomic DNA using the 'SATII pericentric' primers that were synthesized with flanking NheI and HindIII restriction sites such that the amplicon could be readily ligated into the pGL3 basic vector. The opposite orientation clone was made using the original vector as template by PCR with the flanking restriction sites swapped to reverse the orientation of the HSATII insert in the same position. We tested for luciferase expression by RT-qPCR using SYBR green reagent and the comparative C_T method.

Antibodies (product ID, epitope, company). J2 anti-dsRNA, SCICONS English & Scientific Consulting; K1 anti-dsRNA, SCICONS English & Scientific Consulting; ab180573 anti-EIF4A3, Abcam; anti-hADAR, polyclonal rabbit antibody obtained from B. Bass; ab126783 anti-BMI1, Abcam; JBW301 anti-phosphohistone H2A.X (S139), Millipore; D7F7 anti-PKR, Cell Signaling Technologies; ab32036 anti-phospho-PKR (T446), Abcam; sc-133132 anti-eIF2 α , Santa Cruz Biotechnology; ab32157 anti-phospho-eIF2 α (S51), Abcam; GTX28245 anti-GAPDH, GeneTex.

Primers

A-G_conf1_chr19_1_L AACCAACCTTCATCCAGTC
 A-G_conf1_chr19_1_R CACAAAACTGCAGCCACAT
 A-G_conf1_chr7_1_L TCACACCTTATGGCACAAGAA
 A-G_conf1_chr7_1_R GCATCTGCTTCAGGGTTTTC
 A-G_conf1_chr1_2_L TCTGGCAGCTGCTTCTAGTTC
 A-G_conf1_chr1_2_R ATGGATGGTCCACATCACCT
 chr18_1_dsRNA_L GTGATGGAATGTGGGAAAAG
 chr18_1_dsRNA_R ATCACTCGGAGTCTCTTCG
 ZSCAN4_1_dsRNA_L CAGGCTCAAGTGATCCTTCC
 ZSCAN4_1_dsRNA_R GCACCTTTGGAGGCTGAAAG
 Zscan4_RT_L TGGAAATCAAGTGCCAAAAA
 Zscan4_RT_R CTGCATGTGGACGTGGAC
 RPL27-1 L GCAAGAAGAAGATCGCCAAG
 RPL27-1R TCCAAGGGGATATCCACAGA
 DUX4_ex2-3 F2 CGGAGAACTGCCATTCTTTC
 DUX4_ex2-3 R2 CAGCCAGAATTCACGGAAAG
 DUX4CA_F1 TAGGGGAAGAGGTAGACGGC
 DUX4CA_R1 CGGTTCCGGGATTCGGATAG
 Luciferase_F CCAGGTATCAGGCAAGGATATG
 Luciferase_R GTTCGTCTTCGTCCAGTAAG
 SATII_pericentric_1L TGAATGGAATCGTCATCGAA
 SATII_pericentric_1R CCATTCGATAATTCGGCTTG
 Constitutive_dsRNA1L AAAATTGCTCAGCCTTGTC
 Constitutive_dsRNA1R TGTCTGCTAAATCCCCTTG
 U2_1L ATCGCTTCTCGGCCTTTT
 U2_1R CGTTCCTGGAGGTAAGTCTGCAA

Data Availability. MB135-iDUX4 dsRIP-seq and total stranded RNA-seq data were deposited into the Gene Expression Omnibus (GEO) under the SuperSeries accession number GSE114940. DUX4 ChIP-seq data used in this study are available from GEO under accession number GSE33838. Early embryo RNA-seq data were obtained from GEO accession number GSE72379.

Supplementary Material

Supplementary Material is available at HMG online.

Funding

The National Institute of Arthritis and Musculoskeletal and Skin Diseases (R01AR045203 to S.J.T.); The National Institute of Neurological Disorders and Stroke (P01NS069539 to S.v.d.M.); the National Institute of General Medical Sciences (R01GM044073 to B.L.B. and T32GM007270 to S.C.S.); Friends of FSH Research (to S.J.T.). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The authors thank Mathew Thayer for helpful comments and assistance with the FISH protocol and Lisa Kursel for assistance with confocal imaging. They thank Rebecca Resnick, James Thomas and Christine Beck for critical reading of the manuscript and Bob Eisenman for helpful comments.

Conflict of Interest statement. None declared.

References

- Snider, L., Geng, L.N., Lemmers, R.J.L.F., 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.
- 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.
- 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., Wilcox, A.L. et al. (2017) Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.*, **49**, 925–934.
- 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.
- Campbell, A.E., Belleville, A.E., Resnick, R., Shadle, S.C. and Tapscott, S.J. (2018) Facioscapulohumeral dystrophy: activating an early embryonic transcriptional program in human skeletal muscle. *Hum. Mol. Genet.*, **27**, R153–R162.
- Lemmers, R.J.L.F., van der Vliet, P.J., Klooster, R., Sacconi, S., Camaño, 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.
- Snider, L., Asawachaicharn, A., Tyler, A.E., Geng, L.N., Petek, L.M., Maves, L., Miller, D.G., Lemmers, R.J.L.F., Winokur, S.T., Tawil, R. et al. (2009) RNA transcripts, mi RNA-sized fragments and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum. Mol. Genet.*, **18**, 2414–2430.
- 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 facioscapulo-humeral dystrophy. *Dev. Cell*, **22**, 38–51.
9. Young, J.M., Whiddon, J.L., Yao, Z., Kasinathan, B., Snider, L., Geng, L.N., Balog, J., Tawil, R., van der Maarel, S.M. and Tapscott, S.J. (2013) DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLoS Genet.*, **9**, e1003947.
 10. Kowaljow, V., Marcowycz, A., Anseu, E., Conde, C.B., Sauvage, S., Mattéotti, C., Arias, C., Corona, E.D., Nuñez, N.G., Leo, O. et al. (2007) The DUX4 gene at the FSHD1A locus encodes a pro-apoptotic protein. *Neuromuscul. Disord.*, **17**, 611–623.
 11. 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 p 53-dependent myopathy in vivo. *Ann. Neurol.*, **69**, 540–552.
 12. Bosnakovski, D., Xu, Z., Gang, E.J., Galindo, C.L., Liu, M., Simsek, T., Garner, H.R., Agha-Mohammadi, S., Tassin, A., Coppée, F. et al. (2008) An isogenetic myoblast expression screen identifies DUX4-mediated FSHD-associated molecular pathologies. *EMBO J.*, **27**, 2766–2779.
 13. 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 ds RNA and MYC mRNA stabilization activate apoptotic pathways in human cell models of facioscapulohumeral dystrophy. *PLoS Genet.*, **13**, e1006658.
 14. Bosnakovski, D., Choi, S.H., Strasser, J.M., Toso, E.A., Walters, M.A. and Kyba, M. (2014) High-throughput screening identifies inhibitors of DUX4-induced myoblast toxicity. *Skelet. Muscle*, **4**, 4.
 15. Banerji, C.R.S., Panamarova, M., Hebaishi, H., White, R.B., Relaix, F., Severini, S. and Zammit, P.S. (2017) PAX7 target genes are globally repressed in facioscapulohumeral muscular dystrophy skeletal muscle. *Nat. Commun.*, **8**, 2152.
 16. Homma, S., Beermann, M.L., Boyce, F.M. and Miller, J.B. (2015) Expression of FSHD-related DUX4-FL alters proteostasis and induces TDP-43 aggregation. *Ann. Clin. Transl. Neurol.*, **2**, 151–166.
 17. Schlee, M. and Hartmann, G. (2016) Discriminating self from non-self in nucleic acid sensing. *Nat. Rev. Immunol.*, **16**, 566–580.
 18. Athanasiadis, A., Rich, A. and Maas, S. (2004) Widespread A-to-I RNA editing of Alu-containing mRNAs in the human transcriptome. *PLoS Biol.*, **2**, e391.
 19. Blango, M.G. and Bass, B.L. (2016) Identification of the long, edited ds RNAome of LPS-stimulated immune cells. *Genome Res.*, **26**, 852–862.
 20. Blow, M., Futreal, P.A., Wooster, R. and Stratton, M.R. (2004) A survey of RNA editing in human brain. *Genome Res.*, **14**, 2379–2387.
 21. Lybecker, M., Zimmermann, B., Bilusic, I., Tukhtubaeva, N. and Schroeder, R. (2014) The double-stranded transcriptome of *Escherichia coli*. *Proc. Natl. Acad. Sci.*, **111**, 3134–3139.
 22. Whipple, J.M., Youssef, O.A., Aruscavage, P.J., Nix, D.A., Hong, C., Johnson, W.E. and Bass, B.L. (2015) Genome-wide profiling of the *C. elegans* ds RNAome. *RNA*, **21**, 786–800.
 23. Bahn, J.H., Lee, J.H., Li, G., Greer, C., Peng, G. and Xiao, X. (2011) Accurate identification of A-to-I RNA editing in human by transcriptome sequencing. *Genome Res.*, **22**, 142–150.
 24. Kim, D.D.Y., Kim, T.T.Y., Walsh, T., Kobayashi, Y., Matise, T.C., Buyske, S. and Gabriel, A. (2004) Widespread RNA editing of embedded Alu elements in the human transcriptome. *Genome Res.*, **14**, 1719–1725.
 25. Chiappinelli, K.B., Strissel, P.L., Desrichard, A., Li, H., Henke, C., Akman, B., Hein, A., Rote, N.S., Cope, L.M., Snyder, A. et al. (2015) Inhibiting DNA methylation causes an interferon response in cancer via ds RNA including endogenous retroviruses. *Cell*, **162**, 974–986.
 26. Roulois, D., Loo Yau, H., Singhanian, R., Wang, Y., Danesh, A., Shen, S.Y., Han, H., Liang, G., Jones, P.A., Pugh, T.J. et al. (2015) DNA-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. *Cell*, **162**, 961–973.
 27. Hubley, R., Finn, R.D., Clements, J., Eddy, S.R., Jones, T.A., Bao, W., Smit, A.F.A. and Wheeler, T.J. (2016) The Dfam database of repetitive DNA families. *Nucleic Acids Res.*, **44**, D81–D89.
 28. Jagannathan, S., Shadle, S., Resnick, R., Snider, L., Tawil, R.N., van der, 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.
 29. Schonborn, J., Oberstraß, J., Breyel, E., Tittgen, J., Schumacher, J. and Lukacs, N. (1991) Monoclonal antibodies to double-stranded RNA as probes of RNA structure in crude nucleic acid extracts. *Nucleic Acids Res.*, **19**, 2993–3000.
 30. Hall, L.L., Byron, M., Carone, D.M., Whitfield, T.W., Pouliot, G.P., Fischer, A., Jones, P. and Lawrence, J.B. (2017) Demethylated HSATII DNA and HSATII RNA foci sequester PRC1 and MeCP2 into cancer-specific nuclear bodies. *Cell Rep.*, **18**, 2943–2956.
 31. Jolly, C., Metz, A., Govin, J., Vigneron, M., Turner, B.M., Khochbin, S. and Vourc'h, C. (2004) Stress-induced transcription of satellite III repeats. *J. Cell Biol.*, **164**, 25–33.
 32. Altemose, N., Miga, K.H., Maggioni, M. and Willard, H.F. (2014) Genomic characterization of large heterochromatic gaps in the human genome assembly. *PLoS Comput. Biol.*, **10**, e1003628.
 33. Bersani, F., Lee, E., Kharchenko, P.V., Xu, A.W., Liu, M., Xega, K., Mac Kenzie, O.C., Brannigan, B.W., Wittner, B.S., Jung, H. et al. (2015) Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proc. Natl. Acad. Sci.*, **112**, 15148–15153.
 34. Brueckmann, N.H., Pedersen, C.B., Ditzel, H.J. and Gjerstorff, M.F. (2018) Epigenetic reprogramming of pericentromeric satellite DNA in premalignant and malignant lesions. *Mol. Cancer Res.*, **16**, 417–427.
 35. Zhu, Q., Hoong, N., Aslanian, A., Hara, T., Benner, C., Heinz, S., Miga, K.H., Ke, E., Verma, S. and Soroczynski, J. (2018) Heterochromatin-encoded satellite RNAs induce breast cancer. *Mol. Cell*, **70**, 842–853.e7.
 36. Probst, A.V., Okamoto, I., Casanova, M., El Marjou, F., Le Baccon, P. and Almouzni, G. (2010) A Strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev. Cell*, **19**, 625–638.
 37. Zalzman, M., Falco, G., Sharova, L.V., Nishiyama, A., Thomas, M., Lee, S.-L., Stagg, C.A., Hoang, H.G., Yang, H.-T., Indig, F.E. et al. (2010) Zscan 4 regulates telomere elongation and genomic stability in ES cells. *Nature*, **464**, 858.
 38. Dan, J., Rousseau, P., Hardikar, S., Veland, N., Wong, J., Autexier, C. and Chen, T. (2017) Zscan 4 inhibits maintenance DNA methylation to facilitate telomere elongation in mouse embryonic stem cells. *Cell Rep.*, **20**, 1936–1949.
 39. Jack, A.P.M., Bussemer, S., Hahn, M., Pünzler, S., Snyder, M., Wells, M., Csankovszki, G., Solovei, I., Schotta, G. and Hake, S.B. (2013) H3K56me3 is a novel, conserved heterochro-

- matic mark that largely but not completely overlaps with H3K9me3 in both regulation and localization. *PLoS One*, **8**, e51765.
40. Santenard, A., Ziegler-Birling, C., Koch, M., Tora, L., Bannister, A.J. and Torres-Padilla, M.E. (2010) Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. *Nat. Cell Biol.*, **12**, 853–862.
 41. Casanova, M., Pasternak, M., El Marjou, F., Le Baccon, P., Probst, A.V. and Almouzni, G. (2013) Heterochromatin reorganization during early mouse development requires a single-stranded noncoding transcript. *Cell Rep.*, **4**, 1156–1167.
 42. Feng, Q., Snider, L., Jagannathan, S., Tawil, R., van der Maarel, S., Tapscott, S.J. and Bradley, R.K. (2015) A feedback loop between nonsense-mediated decay and the retrogene DUX4 in facioscapulohumeral muscular dystrophy. *Elife*, **4**, 1–13.
 43. Li, Y., Banerjee, S., Goldstein, S.A., Dong, B., Gaughan, C., Rath, S., Donovan, J., Korennykh, A., Silverman, R.H. and Weiss, S.R. (2017) Ribonuclease I mediates the cell-lethal phenotype of double-stranded RNA editing enzyme ADAR1 deficiency in a human cell line. *Elife*, **6**.
 44. Ting, D.T., Lipson, D., Paul, S., Brannigan, B.W., Akhavanfard, S., Coffman, E.J., Contino, G., Deshpande, V., Iafrate, A.J., Letovsky, S. et al. (2011) Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* (80-), **331**, 593–596.
 45. Gerdes, P., Richardson, S.R., Mager, D.L. and Faulkner, G.J. (2016) Transposable elements in the mammalian embryo: pioneers surviving through stealth and service. *Genome Biol.*, **17**, 1–17.
 46. Weinberg, M.S. and Morris, K.V. (2016) Transcriptional gene silencing in humans. *Nucleic Acids Res.*, **44**, 6505–6517.
 47. Kapranov, P., St Laurent, G., Raz, T., Ozsolak, F., Reynolds, C.P., Sorensen, P.H.B., Reaman, G., Milos, P., Arceci, R.J., Thompson, J.F. et al. (2010) The majority of total nuclear-encoded non-ribosomal RNA in a human cell is 'dark matter' unannotated RNA. *BMC Biol.*, **8**, 149.
 48. St Laurent, G., Shtokalo, D., Dong, B., Tackett, M.R., Fan, X., Lazorthes, S., Nicolas, E., Sang, N., Triche, T.J., McCaffrey, T.A. et al. (2013) VlnC RNAs controlled by retroviral elements are a hallmark of pluripotency and cancer. *Genome Biol.*, **14**, R73.
 49. Rowe, H.M. and Trono, D. (2011) Dynamic control of endogenous retroviruses during development. *Virology*, **411**, 273–287.
 50. Grow, E.J., Flynn, R.A., Chavez, S.L., Bayless, N.L., Wossidlo, M., Wesche, D.J., Martin, L., Ware, C.B., Blish, C.A., Chang, H.Y. et al. (2015) Intrinsic retroviral reactivation in human preimplantation embryos and pluripotent cells. *Nature*, **522**, 221–246.
 51. Peaston, A.E., Evsikov, A.V., Graber, J.H., De Vries, W.N., Holbrook, A.E., Solter, D., Knowles, B.B. and Harbor, B. (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell*, **7**, 597–606.
 52. Jachowicz, J.W., Bing, X., Pontabry, J., Bošković, A., Rando, O.J. and Torres-Padilla, M.E. (2017) LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.*, **49**, 1502–1510.
 53. Stadler, G., Chen, J.C.J., Wagner, K., Robin, J.D., Shay, J.W., Emerson, C.P., Jr. and Wright, W.E. (2011) Establishment of clonal myogenic cell lines from severely affected dystrophic muscles-CDK4 maintains the myogenic population. *Skelet. Muscle*, **1**, 12.
 54. Picardi, E. and Pesole, G. (2013) REDIttools: high-throughput RNA editing detection made easy. *Bioinformatics*, **29**, 1813–1814.
 55. Jin, Y., Tam, O.H., Paniagua, E. and Hammell, M. (2015) TETranscripts: a package for including transposable elements in differential expression analysis of RNA-seq datasets. *Bioinformatics*, **31**, 3593–3599.
 56. Silahatoglu, A.N., Tommerup, N. and Vissing, H. (2003) FISHing with locked nucleic acids (LNA): evaluation of different LNA/DNA mixmers. *Mol. Cell. Probes*, **17**, 165–169.
 57. Swanson, E.C., Manning, B., Zhang, H. and Lawrence, J.B. (2013) Higher-order unfolding of satellite heterochromatin is a consistent and early event in cell senescence. *J. Cell Biol.*