Nucleic Acids Research, 2025, **53**, gkaf643
https://doi.org/10.1093/nar/gkaf643
Gene regulation, Chromatin and Epigenetics

Nucleolar FRG2 lncRNAs inhibit rRNA transcription and cytoplasmic translation, linking FSHD to dysregulation of muscle-specific protein synthesis

Valentina Salsi[1,†], Francesca Losi[1,†], Bruno Fosso[2], Marco Ferrarini[1], Sara Pini[1], Marcello Manfredi[3,4], Gaetano Vattemi[5], Tiziana Mongini[6], Lorenzo Maggi[7], Graziano Pesole[2,8,9], Anthony K. Henras[10], Paul D. Kaufman[11], Brian McStay[12], Rossella Tupler[1,11,*]

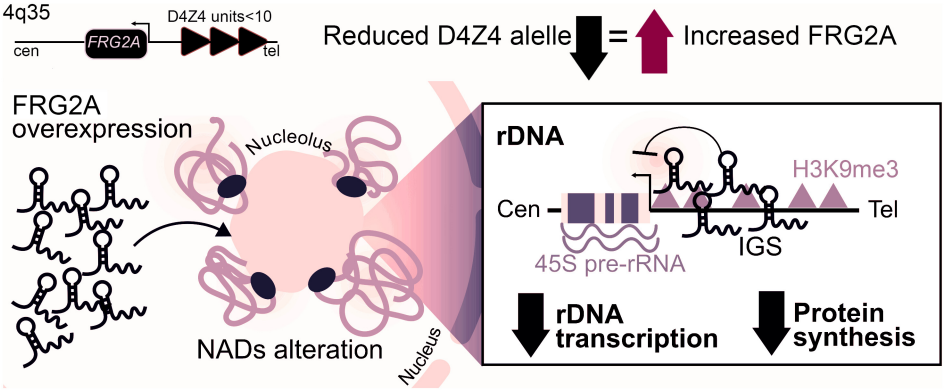
[1]Department of Biomedical, Metabolic and Neural Sciences University of Modena and Reggio Emilia, Modena 41125, Italy
[2]Department of Biosciences, Biotechnology and Environment, University of Bari A. Moro, Bari 70121, Italy
[3]Department of Translational Medicine, University of Piemonte Orientale, Novara 28100, Italy
[4]Institute for Molecular and Translational Cardiology (IMTC), IRCCS Policlinico San Donato, Milan 20097, Italy
[5]Department of Neurosciences, Biomedicine and Movement Sciences, Section of Clinical Neurology, University of Verona, Verona 37134, Italy
[6]Department of Neurosciences "Rita Levi Montalcini", Center for Neuromuscular Diseases, University of Turin, Turin 10124, Italy
[7]Neuroimmunology and Neuromuscular Diseases Unit, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan 20133, Italy
[8]Institute of Biomembranes, Bioenergetics and Molecular Biotechnologies, Consiglio Nazionale delle Ricerche, Bari 70126, Italy
[9]Consorzio Interuniversitario Biotecnologie, Trieste 34148, Italy
[10]Molecular, Cellular and Developmental Biology Unit (MCD), Centre for Integrative Biology (CBI), University of Toulouse, CNRS, UPS, Toulouse 31400, France
[11]Department of Molecular, Cell and Cancer Biology, University of Massachusetts Chan Medical School, Worcester, MA 01605, United States
[12]Centre for Chromosome Biology, College of Science and Engineering, University of Galway, Galway H91 W2TY, Ireland

[*]To whom correspondence should be addressed. Email: rossella.tupler@unimore.it
[†]These authors contributed equally to this work.

Abstract

Facioscapulohumeral muscular dystrophy (FSHD) is a hereditary myopathy linked to deletions of the tandemly arrayed D4Z4 macrosatellite at human chromosome 4q35. These deletions cause local chromatin changes and anomalous expression of nearby transcripts FRG2A, DBET, and D4Z4. We discovered that FRG2A is part of a family of long noncoding RNAs (lncRNAs) expressed in skeletal muscle cells, with levels varying among patients. FRG2A localizes in the nucleolus and associates with repetitive DNA at ribosomal DNA (rDNA) loci and centromeres. Elevated FRG2A expression in FSHD cells alters the three-dimensional architecture of heterochromatin at the nucleolar periphery and reduces rDNA transcription and translation rates, resulting in decreased synthesis of skeletal muscle proteins. We also show that myoblasts from FSHD patients display reduced synthesis of skeletal muscle proteins during differentiation. Our results support a disease model in which nucleolar accumulation of D4Z4-driven lncRNA impairs protein synthesis and contributes to muscle wasting.

Graphical abstract



Received: April 14, 2025. **Revised:** June 15, 2025. **Editorial Decision:** June 16, 2025. **Accepted:** June 19, 2025

Introduction

Two percent of the genome is composed of Tandemly Arrayed Repeats of large size (TARls) commonly known as macrosatellites. The organization, number, and arrangement of these repeats was impossible to interrogate until the advent of long-read sequencing and the assembly of a complete human genome from telomere to telomere (T2T) [1]. Recent research suggests that TARls play essential roles in the maintenance of genome homeostasis by altering genome architecture, epigenome markings, and the regulation of gene expression [2].

Although TARls are not sequence-related, they share several common characteristics: they extend in tandem over hundreds of kilobases, encompassing significant portions of the genome; they are rich in CpG sites, making them often subject to regulation by DNA methylation; they frequently express both noncoding and coding RNAs. To date, the relationships between TARls copy number variations and human physiology and disease are still unclear [2].

The pathogenesis of facioscapulohumeral muscular dystrophy (FSHD) (MIM 158900), a common genetic myopathy, is currently linked to deletions of complete copies of the tandemly arrayed D4Z4 macrosatellite at the 4q35 telomere [3, 4]. A pathogenic threshold number of D4Z4 repeats (≤10) has been established and used for FSHD diagnostics [5]. Importantly, large clinical datasets based on a standardized clinical classification [6] have been generated and these studies concluded that FSHD clinical presentation varies widely [7, 8]. A rough genotype–phenotype correlation between the number of repeats and the severity of the disease holds well among those who carry deleted alleles with 1–3 repeat units. However, this correlation is often broken by high levels of intrafamilial and interindividual clinical heterogeneity, especially in patients with intermediate-sized alleles (4–10 repeats) [8]. Furthermore, D4Z4 alleles with 10 or fewer repeats have been detected in healthy carriers (4.6%) [9] and in individuals with other myopathies. Therefore, D4Z4 array deletions contribute to a complex mechanism that only in certain, still unknown, conditions, lead to muscle wasting.

The prevailing model for FSHD attributes the disease to D4Z4 contraction leading to loss of repressive histone marks and ectopic expression of the *DUX4* retrogene from the last repeat in individuals with a permissive haplotype [10, 11]. However, accumulating clinical, epidemiological, and technological advances suggest that this hypothesis has significant limitations in explaining the incomplete penetrance observed in individuals carrying the FSHD molecular signature [12, 13]. Gene regulation at the 4q35 telomere is complex. D4Z4 deletions cause epigenetic changes *in cis* at the 4q subtelomere, resulting in transcriptional derepression of several proximal genes beyond *DUX4* [14, 15].

Consistent with this view, we recently reported that the 4q subtelomere is subdivided into discrete domains, each with characteristic chromatin features associated with distinct gene expression profiles [16]. These discrete domains undergo diverse region-specific chromatin changes upon treatment with chromatin enzyme inhibitors or genotoxic drugs [16]. Upon DNA damage, the 4q35 telomere-proximal *FRG2* (hereafter named FRG2A), DBE-T, and D4Z4-derived transcripts are induced to levels inversely correlated with the D4Z4 repeat number. All these transcripts are normally barely detectable but stabilized through post-transcriptional mechanisms and bound to chromatin [12]. Among these, *FRG2A* exhibits the strongest response to external stimuli, showing the higher variations in its levels in respect to the other 4q35 genes. Additionally, *FRG2A* transcript levels are abnormally elevated in FSHD muscle samples [14] and increase during myoblast differentiation to myotubes [17]. Several studies also reported that *FRG2A* expression during myogenic differentiation is regulated by long-range chromosomal interactions involving the 4q35 subtelomere [18, 19]. However, although *FRG2A* expression is regulated by environmental, developmental, and genetic stimuli, the biological features of *FRG2A* gene products are poorly understood, and no evidence has been provided regarding the function or localization of *FRG2A*-encoded proteins.

Here, we demonstrate that *FRG2A* functions as a long noncoding RNA (lncRNA). To explore its role, we mapped its chromatin interaction sites and defined its RNA interactome. We showed that FRG2A is functionally important for nuclear architecture, serving as a scaffold for heterochromatin interactions at the nucleolar periphery (NP). We also report that these functions are altered in FSHD cells, observing an unexpected impairment of nucleolar activity and muscle-specific protein synthesis, a feature that could contribute to muscle wasting and may be shared by other myopathies.

Materials and methods

Cell culture

HeLa and HEK293 cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10% fetal bovine serum (FBS), 1% glutamine, and 1% penicillin–streptomycin. Trypsin-EDTA 1× in Phosphate Buffer Saline (PBS) was used to collect cells. Human chromosome hybrids (CHO/Hyb) were obtained from the Coriell Institute for Medical Research and maintained following the supplier's instructions: chromosome 3/CHO hybrid (GM10253), chromosome 4/CHO hybrid (GM10115), chromosome 10/CHO hybrid (GM10926) chromosome 20/CHO hybrid (GM13140), and Chromosome 22/CHO hybrid (GM10888). Control and FSHD-derived primary myoblasts were selected from The Italian National Registry for FSHD [20]. Molecular and clinical features are reported in Supplementary Data S1. cultured in DMEM, supplemented with 20% FBS, 1% glutamine, 1% penicillin–streptomycin, 2 ng/ml epidermal growth factor (EGF), and 25 ng/ml of fibroblast growth factor (FGF). Trypsin 2.5% in HBSS, diluted 1:10 in PBS, was used to collect primary myoblasts. Myoblasts were purified from muscular biopsies (Supplementary Table S1) as previously described [21]. For each FSHD individual, a healthy relative was selected as a control, based on both sample availability and degree of relatedness. This family-matched design was used to minimize inter-individual variability unrelated to disease status and to strengthen the statistical power of within-family comparisons. The research was performed following ethical approval from the Comitato Etico Area Vasta Emilia Nord (approval number: 264/2024/TESS/UNIMO).

FRG2 expression profiling with NGS

Copy DNA (cDNA) from human primary fibroblasts (HPFs) and human biopsies was used to amplify FRG2-exon4 with adapters-containing-pan-FRG2s primer pairs (Supplementary Table S4). Polymerase Chain Reaction (PCR)

products were purified and subjected to NGSelect Amplicons service based on Illumina Platforms and provided by Eurofins Genomics. Read qualities were assessed by Fastqc (Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. [Online] Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)[22]. Individual reports were merged by MultiQC [23] and quality metrics were visually inspected.

Reads were aligned to the T2T human genome assembly using Bowtie2 [24]. Sample-level alignment statistics were computed by samtools idxstats [25]. To evaluate the number of reads aligning to each paralog, we generated (i) a BED file containing the genomic coordinates of the paralogs; (ii) a BED file for each sample containing the genomic coordinates of read alignments, defined considering a window of 600 bp downstream the leftmost mapping position of the SAM files. For each sample we then compared the two sets of coordinates using bedtools intersect [26] with -c option and reported the number of reads aligning to each paralog.

CRISPR–Cas9 editing

Editing of HeLa and HEK293T cell lines was carried out using pSpCas9(BB)-2A-GFP (PX458) plasmid, which encodes Cas9 linked to 2A-EGFP and a cloning backbone for sgRNA. gRNA spacer was designed using the CHOPCHOP [27] web tool (Supplementary Table S4). Sense and antisense oligonucleotide carrying gRNA spacer targeting 5′ and 3′ end of *FRG2* were selected without paralog specificity. Single strand oligonucleotides with 40 bp homology arms were used as repair template to provide knock-in at *FRG2* region (Supplementary Table S4). Oligonucleotides were designed (IDT) with phosphodiester modification at 5′ and 3′ ends of the molecule to impair degradation. We transfected HeLa and HEK293T at 80% of confluency with Lipofectamine™ 2000 (Invitrogen) following the supplier's protocol. A total amount of 2 μg of plasmid DNA was transfected in each well of a six-well tissue culture dish. Repair template was added to the transfection solution at the final concentration of 300 nM. Cells were incubated overnight with transfection medium, then medium was changed. Thirty-six hours after transfection, cells were trypsinized and resuspended at 1×10^6 cells/ml. Cells were sorted with a FACS ARIA III cell sorter based on expression of the EGFP tag on Cas9. Editing characterization of the four bulk populations was carried out with NGSelect Amplicons (Eurofins Genomics) on PCR products obtained by using adapters-containing primers set amplifying all *FRG2*'s paralogs (pan-edited *FRG2*s in Supplementary Table S4). Bioinformatic analysis was performed as previously described for *FRG2* expression profiling with NGS. Design of primers and repair template used in each knock-in setting are displayed in Supplementary Table S4.

RNA extraction and real-time quantitative PCR (RT-qPCR)

Total cellular RNA was obtained from cell lines and HPMs by using PureLink RNA Mini Kit (Thermo Fisher Scientific cat #12183018A), according to the manufacturer's instructions. DNAse digestion and cDNA synthesis were performed by using Maxima H-cDNA Synthesis Master Mix, with dsDNase (Thermo Fisher M1482). Specific mRNA expression was assessed by qRTPCR (iTaq Universal SYBR® Green Supermix, BIORAD #1725120 in a CFX connect Real Time Machine BIORAD) using primers listed in Supplementary Table S4, normalized over RPLP0 and GAPDH housekeeping mRNAs.

Protein extraction and immunoblotting

Protein extraction was carried out starting from 5×10^5 cells per sample. Cells were incubated in 50 μl of extraction buffer (50 mM Tris–HCl, 400 mM NaCl, 1% NP-40, and 1× Protease Inhibitor Cocktail) on an inverting wheel for 30 min in a cold chamber. After centrifugation 12 470 × *g* for 30 min at 4°C, supernatant fractions containing extracted protein were collected. Then, immunoblotting was performed to assess the presence of FRG2 proteins. Protein extracts were separated on 12% SDS–PAGE gel and then electrotransferred onto a nitrocellulose membrane, and subjected to overnight incubation with primary antibody, listed in Supplementary Table S4. Secondary antibodies conjugated with horseradish peroxidase were used and visualized by chemiluminescence using a ChemiDoc system (Bio-Rad).

RNA fractionation

RNA fractionation was performed starting from 1×10^6 cells. After collection, cells were lysed with 175 μl of cold cytoRNA solution [50 mM Tris–HCl pH 8.0, 140 mM NaCl,1.5 mM $MgCl_2$, 0.5% NP-40, 2 mM vanadyl ribonucleoside complex (VRC); Sigma] and incubated 5 min on ice. Cell suspensions were centrifuged at 4°C and 300 × *g* for 2 min and the supernatant, corresponding to the cytoplasmic fraction, was transferred into a new tube and stored on ice. Pelleted nuclei were extracted with 175 μl of cold nucRNA solution (50 mM Tris–HCl pH 8.0, 500 mM NaCl, 1.5 mM $MgCl_2$, 0.5% NP-40, and 2 mM VRC; Sigma) and incubated 5 min on ice. The lysed nuclei were centrifuged at 4°C and 16 360 × *g* for 2 min and the supernatant, corresponding to the nuclear-soluble fraction, was transferred into a new tube and stored in ice. The remaining pellet was collected corresponding to the chromatin-associated fraction. Total RNA from the cytoplasmic and nuclear fractions was extracted by using PureLink RNA MiniKit (Invitrogen) following the manufacturer's instructions for the RNA extraction from aqueous solutions. The pellet containing the chromatin-associated fraction was extracted with the standard procedure. cDNAs were obtained as described above.

Chromatin RNA immunoprecipitation

Chromatin RNA immunoprecipitation (ChRIP) was performed as described [28] using antibodies anti-H3K4me3, H3K9me3, H3K27me3 or negative control IgG antibodies, as reported in Supplementary Table S4. A total of 3×10^6 human primary myoblasts (HPMs) were used for each IP. RNA was extracted and real-time quantitative PCR (RT-qPCR) was performed as described above. Ten percent input was used to calculate the percentage of transcript bound to chromatin compared to the negative control Immunoglobulin G (IgG) .

Chromatin isolation by RNA purification

Chromatin isolation by RNA purification (ChIRP) was performed in HeLa cells by using the Magna ChIRP™ RNA Interactome Kit (Sigma, cat #17-10494) strictly following manufacturer'instructions. We designed targeted biotinylated antisense oligonucleotide probes as shown in Supplementary Fig. S19, complementary to the predicted FRG2A

mature transcript (NMNM_001286820.2) by using Stellaris (ChIRP) Probe Designer tool, generating a custom ChIRP probe set consisting of 10 specific antisense oligonucleotides (Supplementary Fig. S19). ChIRP probes were divided in two sets: (i) *FRG2* EVEN, containing 5 out of 10 antisense oligonucleotides (sets 2, 4, 6, 8, and 10), and (ii) *FRG2* ODD, containing the remaining five antisense oligonucleotides (sets 1, 3, 5, 7, and 9). To rule out artifacts due to direct probes hybridization to genomic DNA, we also used two controls: (iii) LacZ probe set, which would pull down potential contaminating DNA and (iv) TERC probe set, which was also used as a positive control for the effective success of the immunoprecipitation protocol. The isolated RNA and DNA samples were then subjected to next-generation sequencing, while the protein fraction was subjected to mass-spectrometry analysis.

ChIRP bioinformatic analysis

ODD and EVEN sequencing data were initially analyzed in parallel by applying the same workflow. Raw data were trimmed by using fastp [29, 30] (-q 20 -u 20 -l 25 -g 10 -w 15) to remove low-quality sequences. Retained high-quality reads were mapped against the T2T human genome [1] by using bowtie2 [24] (–no-unal –no-discordant). By using samtools, the obtained SAM files were binary compressed in BAM (view -bS) and sorted. Optical and PCR duplicates were marked by using PICARD (Broad Institute. Picard Toolkit. 2019. GitHub Repository: https://broadinstitute.github.io/picard/), MarkDuplicates (–REMOVE_DUPLICATES false –ASSUME_SORT_ORDER coordinate). Following ODD and EVEN alignment data per each experiment, were intersected by using the bedtools [31] function intersect (-wa) in both direction (i.e. ODD to EVEN and EVEN to ODD). The obtained bam files were sorted and merged by samtools. Finally, merged bam files were sorted and indexed by samtools. The average insert size is 190 bp and 205 bp for EVEN and ODD probe sets respectively. Peaks calling was performed by using the macs3 [32] callpeak function (-f BAMPE -g 3.05e + 9 -B -q 0.01). Regardless of the analyzed experiment, peak calling was performed by using input genomic DNA as control. Finally, the annotatePeaks.pl Perl script, belonging to the Homer suite (HOMER Software and Data Download. http://homer.ucsd.edu/homer) [33], was used to deduce annotations associated with the genomic regions in which peaks fell. Raw data and processed files for the ChIRP sequencing have been deposited in the NCBI Gene Expression Omnibus under accession number: PRJNA1122505.

NADs annotation and NAD finder overlap analyses

Sequencing data belonging to the Bioproject PRJNA354332 [34] and regarding WGS and Nucleolus Sequencing (NS) in untreated HeLa cells and relative replicates were retrieved from the SRA repository by using the sratoolkit prefetch and fastq-dump tools (–split-files –gzip). Raw data were trimmed by using fastp [29] (-q 20 -u 20 -l 25 -g 10 -w 15) to remove low-quality sequences. Retained high-quality reads were mapped against the T2T human genome [1] by using bowtie2 (–no-unal). By using samtools [25], the obtained sam files were binary compressed (view -bS), sorted (sort), and indexed (index). NS replicates were used by the R library HMMtBroadPeak (binsize = 50 000, background = 25, pseudocount = 1) to identify NADs using WGS data as control. HMMtBroadPeak relies on the same strategy implemented in NADfinder [35], but it is optimized to overcome memory leaps. In particular, per each bin, the number of reads falling within is counted, and bins supported by a numbered reads at least equal to the background parameter are retained and normalized to CPM (count per million). Then $\log_2$ transform is applied to the ratio among NS and WGS normalized counts.

Transformed values are finally used to define NAD peaks by running the Hidden Markov Model implemented in the R HMMt library (Devailly, GitHub repository: https://github.com/gui11aume/HMMt).

Overlaps between FRG2 peaks and inferred NADs were identified by using the bedtools intersect function (-wa -f). In particular, an increasing value of the required overlap (1 bp, 50% and 90% peak region) among peaks and NAD for reporting was applied to reduce the effect of casual co-localization. Moreover, in order to further evaluate the non-causality of peaks-NAD co-localization, 10 000 simulations of randomly extracted regions were produced by using bedtools random and the overlaps were measured as previously described. Finally, the R function ppois was applied to infer the probability to observe a number of overlaps greater than the one observed between FRG2 peaks and NADs, using the Poisson cumulative density function.

HeLa and primary cells proteomics

After cell lysis with RIPA buffer, proteins were digested with trypsin and peptides were analyzed on an Ultimate 3000 RSLC nano coupled directly to an Orbitrap Exploris 480 with a High-Field Asymmetric Waveform Ion Mobility Spectrometry System (FAIMS) (all Thermo Fisher Scientific). Samples were injected onto a reversed-phase C18 column (15 cm × 75 μm i.d., Thermo Fisher Scientific) and eluted with a gradient of 6%–95% mobile phase B over 41 min by applying a flow rate of 500 nl/min, followed by an equilibration with 6% mobile phase B for 1 min. The acquisition time of one sample was 41 min and the total recording of the MS spectra was carried out in positive resolution with a high voltage of 2500 V and the FAIMS interface in standard resolution, with a CV of −45 V. The acquisition was performed in data-independent mode (DIA): precursor mass range was set between 400 and 900, isolation window of 8 m/z, window overlap of 1 m/z, HCD collision energy of 27%, orbitrap resolution of 30 000 and RF Lens at 50%. The normalized AGC target was set to 1000, the maximum injection time was 25 ms, and microscan was 1. For DIA data processing, DIA-NN (version 1.8.1) was used: the identification was performed with "library-free search" and "deep learning-based spectra, RTs and IMs prediction" enabled. Enzyme was set to Trypsin/P, precursors of charge state 1–4, peptide lengths 7–30, and precursor m/z 400–900 were considered with maximum two missed cleavages. Carbamidomethylation on C was set as fixed modification and oxidation on M was set as variable modification, using a maximum of two variable modifications per peptide. False Discovery Rate (FDR) was set to 1%.

Myoblasts differentiation

HPMs from CTRL and FSHD patients were induced to differentiate between passages P3 and P5. To induce differentiation, myoblasts were plated at a confluence of 15 000 cells/cm^2. Twenty-four hours after seeding, growth medium

was replaced by differentiation media composed of DMEM supplemented with 10% FBS, 1% glutamine, and 1% penicillin–streptomycin, without the addition of EGF and FGF. Medium was changed every 2–3 days. The resulting myotube cultures were used for analysis at day 8 (D8) of myogenic differentiation, after assessing their complete differentiation as previously described [16].

Ribo-seq data analysis

First, we accessed (April 2024) the RiboSeq data portal (https://rdp.ucc.ie/) to select RiboSeq experiments exploiting HeLa cells (Toggles: Available in Trips-Viz,Organism: Homo sapiens, Library-Type: Ribo-Seq, Inhibitor: None). Following considering the selected samples, RiboSeq profiles for ENST00000226798 (*FRG1* gene), ENST00000504750 (*FRG2* gene), ENST00000443774 (*FRG2B* gene), and ENST00000308062 (*FRG2C* gene) transcripts were extracted as CSV files from Trips-Viz [67]. Moreover, the RiboSeq profiles of ENST00000395240 (*ALDOA* gene) and ENST00000501122 (*NEAT1* gene) transcripts were also accessed as examples of protein-coding and non-coding genes, respectively.

RNA-FISH and immunofluorescent-RNA-FISH (immuno-RNA-FISH)

RNA-FISH experiments were performed as previously described [36] in HeLa cells and HPMs using locked nucleic acid (LNA) fluorescent probes (FRG2A/B-LNATY563 and FRG2A-LNA665), designed by Genglobe LNA designer (Qiagen). Probes sequences and localization are shown in Supplementary Fig. S19. Where indicated, cells were treated with 1 μg/ml Actinomycin D (ActD) for 1–4 h. Briefly, cells were fixed with 3% paraformaldehyde (PFA) in PBS for 12 min at room temperature and permeabilized with 0.5% Triton X-100 in PBS. As controls, RNase A or H treatments were also performed by incubating the slides with 1 mg/ml RNase A/H in PBS at room temperature for 30 min. Slides were then stored in 70% ethanol (EtOH), at −20°C for at least one night. The next day, slides were dehydrated in 80%, 95%, and 100% EtOH, 3 min each, and left to dry. The hybridization solution was prepared as follows: 0.2 μM LNA probe, 50% formamide, 2× saline sodium citrate (SSC), 10% dextran sulphate, 10 mM VRC, 2 mg/ml bovine serum albumin (BSA). The mix was denatured at 75°C for 7 min. Hybridization was performed in a humid chamber for 35 min at 37°C, protected from light. Coverslips were washed three times in 0.1× SSC heated at 60°C for 5 min, rinsed in 2× SSC, and mounted by using the ProLong™ Diamond Antifade Mountant with DAPI (Invitrogen™, cat #P36962).

For immunofluorescence-RNA-FISH (immuno-RNA-FISH), slides were first subjected to IF protocol as follows. After fixation in PFA 3% and permeabilization with 0.5% Triton X-100 in PBS, coverslips were incubated in blocking solution (1% BSA in PBS) for 15 min, and then with primary antibodies (diluted 1:100) in blocking solution for 2 h at room temperature. After three washes in PBS, the cells were incubated with the secondary antibodies, diluted 1:250 in blocking solution, for 45 min. Slides were post-fixed in 3% PFA/PBS for 10 min at room temperature, washed in 2× SSC, and then RNA-FISH protocol was followed as described above. A list of the antibodies used in this paper is provided in Supplementary Table S4.

Images of RNA-FISH and immuno-RNA-FISH experiments were acquired by using the Leica TCS SP8 AOBS confocal microscope system equipped with a ×63 oil immersion lens. Emission wavelengths were: 405 nm for DAPI signal, 488 nm for green staining, 550 nm for red fluorescence, and 647 nm for magenta staining.

Nucleoli isolation

Nucleoli were isolated from HeLa cells as previously described [35] with some variations. Briefly, 1×10^8 HeLa cells were washed three times with cold PBS. Cells were scraped and collected by centrifugation at 218 × *g* at 4°C for 5 min. Cells were subsequently resuspended in 5 ml of cold Buffer A (10 mM Hepes pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 0.5 mM DTT, EDTA-free protease inhibitors, Roche) and incubated 15 min on ice. Cells were then dounce homogenized 20 times using a tight pestle and centrifuged at 218 × *g* at 4°C for 5 min. The supernatant was collected as cytoplasmic fraction, while the pellet, which contained nuclei, was resuspended with 3 ml of S1 solution (0.25 M sucrose, 10 mM $MgCl_2$, EDTA-free protease inhibitors). Resuspended nuclei were layered over 3 ml of S2 solution (0.35 M sucrose, 0.5 mM $MgCl_2$, EDTA-free protease inhibitors) and centrifuged at 1430 × *g* at 4°C for 5 min. The resulting nuclear pellet was resuspended with 3 ml of S2 solution and sonicated with 10-s bursts (with 10-s intervals between each burst) using Soniprep 150 (MSE) fine probe sonicator set at power setting 5. The sonicated sample was layered over 3 ml of S3 solution (0.88 M sucrose, 0.5 mM $MgCl_2$, EDTA-free protease inhibitors) and centrifuged at 3000 × *g* at 4°C for 10 min. The supernatant was collected as the nucleoplasmic fraction. The pellet, which contained nucleoli, was resuspended in 0.5 ml of S2 solution, centrifuged at 1430 × *g* at 4°C for 5 min, resuspended in 0.1 ml of S2 solution, and used to assess nucleoli purity. DNA and RNA were extracted from each fraction by using DNeasy® Blood and Tissue Kit (QIAGEN, cat#69504) and Pure Link RNA Mini Kit (Invitrogen™ cat #12183018A), respectively. Enrichment of target DNA sequences and RNAs were assessed by RT-qPCR, as previously described, using primers in Supplementary Table S4.

Sequence of primers and probes

The sequences of the primers and LNA fluorescent probes used in this study are provided in Supplementary Table S4.

Short-interfering RNA gene silencing

FRG2A/B transcripts were silenced using short interfering RNAs (siRNAs). We designed two FRG2A/B-specific siRNAs (Supplementary Fig. S19 for sequences and localization) and silencer select siRNA (Invitrogen Silencer™ Select Pre-Designed siRNA Catalog number: 4392420) and a scramble siRNA was used as a control. Lipofectamine™ RNAiMAX Transfection Reagent (Invitrogen, cat #13778 075) was diluted in Opti-MEM™ I Reduced Serum Medium (Gibco™, cat #31985062). CTRL and FRG2A/B-specific siRNAs were diluted in Opti-MEM™ I Reduced Serum Medium, at 20 pM final concentration. Diluted siRNAs and Lipofectamine™ RNAiMAX Transfection Reagent were mixed together and incubated at room temperature for 15 min, to allow complexes to form. StealthTM FRG2A/B/control siRNA - Lipofectamine™ RNAiMAX Transfection Reagent complexes were added to each well containing cells and medium and incubated at 37°C in a humidified CO_2 incubator for 48 h.

Northern blot

Ribosomal RNA maturation pathways were evaluated by northern blotting, as described [37]. Briefly, 4 μg of total RNA extracted from HeLa cells and HPMs was diluted in 5 volumes of Glyoxal mix (6 ml of Dimethyl sulfoxide (DMSO), 2 ml of deionized glyoxal (Fluka, cat. #50 649), 1.2 ml of 10× Boric Acid/Phosphate/Tris/EDTA (BPTE), 600 μl of 80% glycerol, 40 μl of 10 mg/ml EtBr), incubated at 55°C for 1 h, and chilled on ice for 10 min. Then, an electrophoresis run was performed in a 1.2% agarose gel prepared in BPTE buffer [for 10× concentration: 30 g of 100 mM PIPES (SIGMA, cat. #P6757), 60 g of 300 mM BIS-TRIS (SIGMA, cat. #B9754), and 20 ml of 0.5 M EDTA pH 8.0], for 5 h at 100 V. When migration was sufficient, the gel was rinsed twice with H_2O and soaked for 20 min in 75 mM NaOH and twice in 0.5 M Tris–HCl pH 7.4, 1.5 M NaCl for 10 min. RNA was transferred to a Hybond N+ (GE) membrane, O/N at room temperature. After a pre-hybridization step with transfer-RNA (tRNA) from *Escherichiacoli*, hybridization was performed by using ^{32}P-labeled Internal Transcribed Spacer (ITS)1- and ITS2-specific probes, incubated O/N at 42°C. Membranes were washed three times in 2× SCC + 0.1% SDS an twice in 0.5× SCC + 0.1% SDS. Radioactive signals were detected by using Typhoon Trio PhosphorImager and quantified using the MultiGauge software.

Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) assays were performed in HPMs as described earlier in using specific antibodies, as listed in Supplementary Table S4. Enrichments of DNA sequences immunoprecipitated from at least three independent ChIP experiments were analyzed by quantitative real-time PCR (qPCR), as described above. Enrichment of amplified DNA sequences (primers listed in Supplementary Table S4) in immunoprecipitates was calculated as the ratio between the DNA amount in immunoprecipitation samples and that in the total input chromatin.

Nascent RNA detection

Global RNA synthesis in CTRL and FSHD HPMs was visualized and quantified by using Click-iT™ RNA Alexa Fluor™ 488 Imaging Kit (Invitrogen™, cat #C10329). This technology exploits the incorporation of the uridine analogue 5-ethynyl uridine (EU) within nascent RNA transcripts. Then, through a click-it chemistry, Alexa Fluor™ 488 fluorophore was coupled to the incorporated EU, to visualize the newly synthesized transcripts.

Briefly, the day before labeling, HPMs were plated into multi-24 well plates (2×10^5 cells per well), with 12 mm øcoverslips placed at the bottom. Cells were incubated 1 h with EU at 1 mM final concentration, under normal cell culture conditions. After, cells were fixed in 3.7% formaldehyde in PBS and permeabilized with 0.5% Triton X-100 in PBS. The Click-iT® reaction was carried out following the manufacturer's instructions. Then, since the ScanR software required labeled nucleoli to quantify EU nucleolar signals, immunofluorescence with anti-fibrillarin antibody was performed, as previously described.

Nascent RNA capture assay

Newly synthesized RNA in HPMs was also measured by using the Click-iT™ Nascent RNA Capture Kit (Invitrogen™, cat #C10365), another technology based on EU incorporation. Total RNA or mRNA labeled with EU is isolated and used in a click-it reaction, which binds a biotin azide to the incorporated EU, creating a biotin-based handle for capturing nascent RNA transcripts on streptavidin magnetic beads.

The day before labeling, HPMs were plated into a T75 flask. The next day, cells were incubated with 0.2 mM EU for 1 h under normal cell culture conditions. Then, total RNA was extracted as described above. EU-labeled RNA was biotinylated through the click-it reaction with 0.5 mM biotin azide, following the manufacturer's instructions. EU-biotin-labeled RNA was isolated by using Dynabeads® MyOne™ Streptavidin T1 magnetic beads and retro-transcribed as previously described. The cDNAs obtained were quantified by RT-qPCR to assess the amount of rRNAs with primers listed in Supplementary Table S4.

This experiment was performed both in CTRL and FSHD HPMs, as well as in FRG2A/B-silenced FSHD HPMs, treated with siRNAs 48 h before EU incorporation, as previously reported.

Nascent protein synthesis detection

Nascent protein synthesis was analyzed in HPMs by using the Click-iT™ HPG Alexa Fluor™ 488 Protein Synthesis Assay Kit (Invitrogen™, cat #C10428). Click-iT® HPG (L-homopropargylglycine) is an amino acid analog of methionine, which is incorporated into proteins during active protein synthesis. Detection of the incorporated amino acid is allowed by a click-it reaction with Alexa Fluor® 488.

The day before labeling, HPMs were plated into a multi-24 well plates (2×10^5 cells per well), with 12 mm øcoverslips placed at the bottom. Cells were incubated with HPG at 50 μM final concentration for 30 min, under normal cell culture conditions. After, cells were fixed in 3.7% formaldehyde in PBS and permeabilized with 0.5% Triton X-100 in PBS. The Click-iT® reaction was carried on following the manufacturer's instructions.

High content imaging-based assay

Stimulated Emission Depletion (STED) images were acquired using the Leica SP8 White Light Laser (WLL) confocal microscope, equipped with a 775 nm depletion laser. The depletion power was set to 30% for the ATTO647-N fluorophore and 100% for FRG2A/B-LNATY563.

For centromere localization analysis and quantification of EU and HPG incorporation, images were obtained using a Nikon Confocal Microscope A1 equipped with a ×60 oil immersion lens. and analyzed using the ScanR Analysis software (Olympus).

To count the nucleolus-associated centromeres in HPMs, each centromere was segmented based on Centromere Protein B (CENP-B) signal, using an edge detection algorithm. Following segmentation, all CENP-B foci were counted and those colocalizing with FBL or in which FBL was within a range of 140 nm were considered as nucleolus-associated centromeres. The analysis was performed in CTRL and FSHD HPMs, as well as in FSHD HPMs after FRG2A/B siRNA treatment. In the latter, we also measure FRG2A/B mean intensity within the nucleoli as the ratio between the

mean fluorescence intensity inside and around the nucleolus.

Images from the Click-iT™ RNA Alexa Fluor™ 488 Imaging Kit assay and Click-iT™ HPG Alexa Fluor™ 488 Protein Synthesis Assay Kit followed a similar analysis. To quantify EU incorporation within the nucleoli as a measure of ribosomal RNA (rRNA) transcription, nucleoli were segmented based on FBL signal. EU incorporation was quantified as the ratio between the EU signal inside the nucleolus and from the surrounding nuclear region. To quantify HPG incorporation, nuclei were segmented based on DAPI fluorescent signal. Then, HPG mean intensity was quantified within a selected area surrounding the nuclei.

Statistical significance was tested by using the Mann–Whitney test, when comparing CTRL versus FSHD HPMs and two-way ANOVA for siRNAs treatments.

Results

The T2T-CHM13 human genome assembly contains 14 *FRG2*-like sequences

As the first step to investigate the biological role of *FRG2A*, we interrogated the T2T-CHM13 human genome assembly and found 14 *FRG2* paralogs (*FRG2s*; Fig. 1A), while only six were annotated in the previous GRCh38 assembly. In the T2T genome, we detected six complete copies (2918 bp from Transcription Start Site to polyA signal), located on chromosomes 4 (*FRG2A*), 10 *(FRG2B*), 20 (*FRG2EP*), and 22 (*FRG2C-1*) and in two different positions on chromosome 3 (*FRG2C* centromeric and *FRG2FP* telomeric). Out of the 14 *FRG2* paralogs, four have predicted protein coding potential (*FRG2A*, *FRG2B*, *FRG2C*, and *FRG2C-1*), while the others are annotated as "unprocessed pseudogenes" based on CAT-Liftoff gene annotation [38]. Predicted proteins encoded by paralogs with coding potential display distinct amino acid sequences and lengths (284aa FRG2C; 281aa FRG2A and B; 164 aa FRG2C-1) (Supplementary Fig. S1). Notably, the various loci encoding FRG2 transcripts are predominantly positioned within heterochromatic regions, including telomeres and pericentromeric domains.

To determine which paralog encode stable transcripts, we performed NGS sequencing of PCR-enriched FRG2 cDNAs, comparing muscular biopsies as well as fibroblasts from healthy control individuals and FSHD patients from four families (Supplementary Table S1). Taking advantage of chromosome-specific single nucleotide polymorphisms (SNPs), we identified the genomic loci encoding each FRG2 transcript (Fig. 1B and Supplementary Supplementary Fig. S2). In the fibroblast samples FRG2C transcripts were detected almost exclusively. In contrast, transcripts from many *FRG2* paralogs were detected in the muscle samples, and their relative expression levels differed among individuals. These Supplementary Data suggest regulation mechanisms that are both tissue-specific and subject-specific. Consistent with previous reports [14, 16, 17] this analysis confirmed that *FRG2A* is expressed only in muscle cells and significantly overexpressed in FSHD-derived samples in comparison to healthy control family members ($P = 0.042$, Fig. 1B). The other paralogs do not exhibit significant or consistent expression changes across FSHD and control samples in our RNA-seq dataset.

We observed that five out of the nine *FRG2* paralogs that produce transcripts—namely *FRG2FP, FRG2GP*, *FRG2HP*, *FRG2IP*, and *FRG2JP*—lack predicted coding potential. Most of these genes are in the pericentromeric region of chromosome 16 and comprise only incomplete sequences without an open reading frame (ORF).

Additionally, *FRG2EP* and *FRG2FP*, despite being transcribed, would encode very short proteins of just 10 and 79 amino acids, respectively, if translated. We thus concluded that *FRG2* paralogs can produce transcripts in primary cells and in muscle biopsies, some of them lacking coding potential.

FRG2 paralogs encode a family of heterochromatin-associated long noncoding RNAs

Having detected both coding and noncoding FRG2 transcripts, we next tested whether we could detect *FRG2*-encoded proteins. To overcome the lack of available antibodies that recognize FRG2 proteins, we created four different edited cell lines with FLAG epitope tag sequences in frame with *FRG2* ORFs. Via CRISPR–Cas9 technology (Supplementary Fig. S3A), we placed a FLAG epitope sequence at either the N-terminus or the C-terminus of *FRG2* genes, in both HeLa and HEK293 cells. By PCR from genomic DNA in bulk populations, we observed the presence of Flag-tagged-knocked-in alleles at *FRG2A* (4q35), *FRG2B* (10q26), *FRG2C* (3p12), *FRG2C-1* (22p11) coding-predicted genes, and at *FRG2EP* (20q11) and *FRG2FP* (3q29) (Supplementary Fig. S3B–E). Specific primers for each coding-predicted paralog were designed and tested using CHO-monohybrid cell lines that contain a single human chromosome (Supplementary Fig. S3F). In all the edited cell lines, both the endogenous (Supplementary Fig. S3G) and the Flag-tagged FRG2 transcripts (Supplementary Fig. S3H) were easily detectable and, as measured by RT-PCR, their expression increased significantly after doxorubicin treatment [16] (Supplementary Fig. S3I). Because HeLa and HEK293 cells are aneuploid and were edited as bulk populations, the allelic configuration of the FLAG insertion (monoallelic versus biallelic) cannot be determined.

However, immunoblotting using the antibody against the Flag epitope failed to detect any FRG2 proteins (Supplementary Fig. S4). The inability to detect FRG2 proteins by immunoblotting raised the question whether FRG2 transcripts are translated into proteins. We thus interrogated multiple databases for evidence of FRG2 protein expression. First, we examined the Genome Wide Information on Protein Synthesis (GWIPS) database [39] that catalogs translational potential via ribosome profiling and searched within available RiboSeq data. This database did not detect FRG2 RNA association with ribosomes (Supplementary Fig. S5) or the presence of *FRG2*-derived peptides. We also failed to detect FRG2 peptides within available proteomic databases from human myoblasts [40, 41]. Likewise, the Protein Atlas [42] and Uniprot [43] databases lack *FRG2*-derived proteins identified in proteome samples. We also performed an untargeted proteomic analysis in HeLa cells with cells treated with doxorubicin that induce stabilization of FRG2 RNAs [16]. More than 6000 human proteins were identified, but FRG2s were not detected (Supplementary Data S1). Collectively, these Supplementary Data suggested that FRG2 transcripts might be rarely if ever translated into proteins.

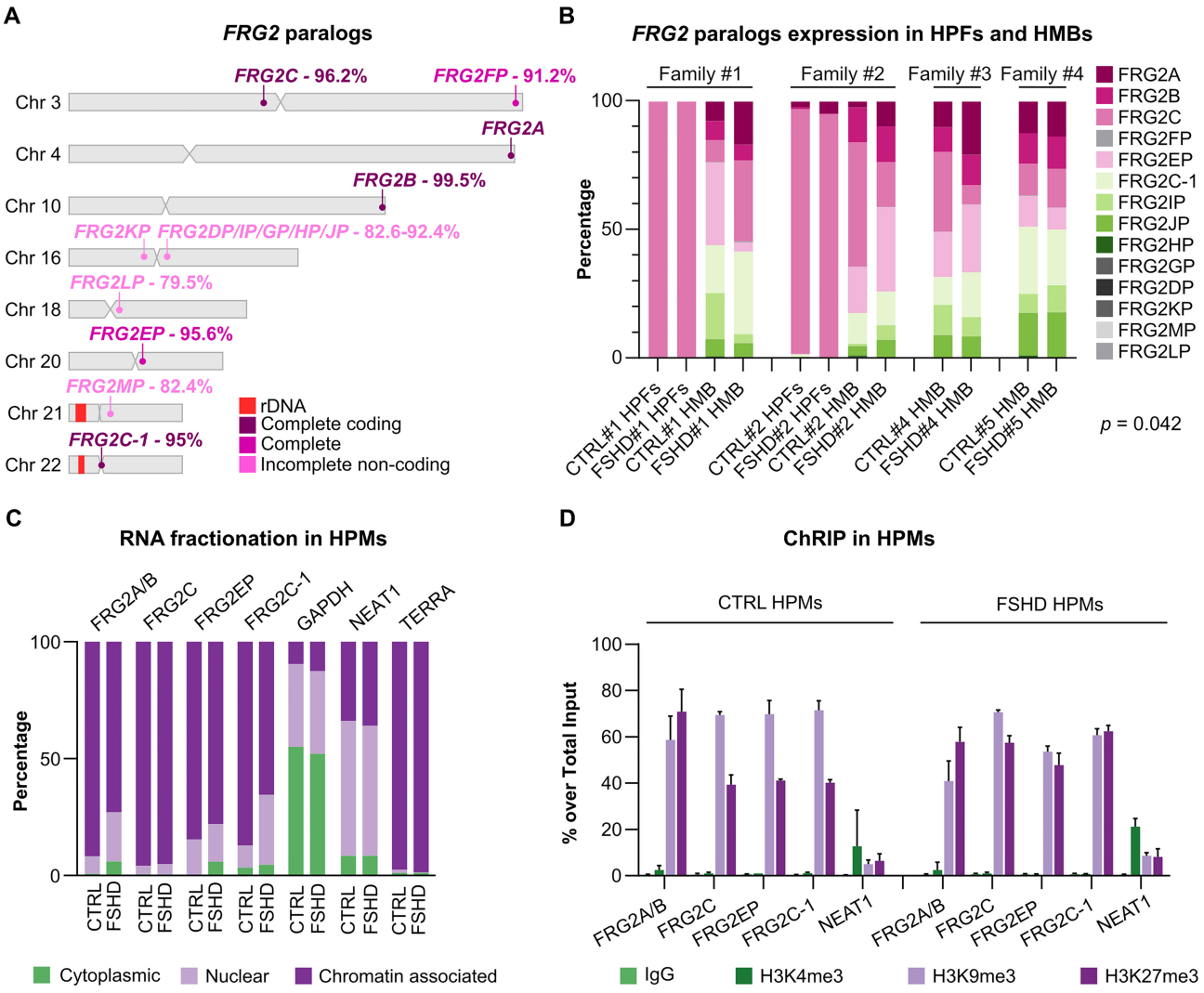


Figure 1. The human genome encodes 14 FRG2 paralogs. (**A**) Chromosome ideograms showing the location of *FRG2* paralogs and percentage identity to *FRG2A* on chromosome 4. The different colors indicate *FRG2* paralogs that are complete (2918 bp) with coding potential, (*FRG2A, FRG2B, FRG2C,* and *FRG2C-1*), complete without coding potential, (*FRG2FP* and *FRG2EP*) or partial and noncoding, (*FRG2KP, FRG2DP, FRG2IP, FRG2GP, FRG2HP, FRG2JP, FRG2LP,* and *FRG2MP*) according to T2T-CHM13 annotation. (**B**) Genomic origin of FRG2s transcripts: NGS of PCR-enriched FRG2 cDNAs obtained from HPFs and human muscular biopsies (HMBs) from FSHD subjects and healthy relatives (four families). The analysis demonstrated muscle-specific expression of *FRG2A* and its overexpression in FHSD HMBs compared to controls. The statistical significance was assessed using a paired *t*-test, comparing the mean number of reads in controls ($n = 4$) and FHSD ($n = 4$) HMBs ($P = 0.042$). (**C**) RNA fractionation in primary myoblasts (HPMs) from CTRL and FSHD subjects showed protein-coding FRG2 transcripts enriched in the chromatin-associated fraction of RNA. GAPDH, NEAT1 [77], and TERRA [45] were used as positive controls for cytoplasmic, nuclear, and chromatin-associated RNA enrichment, respectively. (**D**) ChRIP in CTRL and FSHD HPMs showed FRG2s transcripts enrichment with heterochromatin marks H3K9me3 and H3K27me3. NEAT1, associated with H3K4me3-marked euchromatin, was used as a negative control [77].

We therefore focused on FRG2-derived RNAs to learn about their potential as regulatory molecules. First, we performed cell fractionation experiments using HeLa cells and primary biopsy-derived myoblasts from control or FSHD-affected individuals (Fig. 1C and Supplementary Fig. S6A), and isolated RNAs from the cytoplasmic, nuclear, and chromatin fractions. In both primary myoblasts and HeLa cells, transcripts from *FRG2* paralogs were highly enriched in the chromatin-associated fraction. To profile the types of chromatin interacting with FRG2s transcripts, we performed ChRIP [28], a chromatin-based RNA enrichment method (conceptually a "pull-down" on chromatin), using antibodies raised against histone marks typical of euchromatin or heterochromatin (Fig. 1D and Supplementary Fig. S6B). We found that FRG2s transcripts are enriched in H3K9me3- and H3K27me3-marked but not in H3K4me3-marked fractions, indicating that they are preferentially associated with heterochromatin [44]. Therefore, we concluded that FRG2A and its paralogs constitute a family of lncRNAs associated with heterochromatin.

Trans-interactions of FRG2A/B RNAs are enriched at highly repetitive loci: rDNA and centromeric satellites

Among *FRG2* paralogs, *FRG2A* is the only one that has been linked to a human disease, FSHD. Notably, the amount of FRG2A transcript (hereafter referred as FRG2A-t) is elevated in muscle biopsies from FSHD affected subjects compared to healthy family members (Fig. 1B) and inversely correlates with

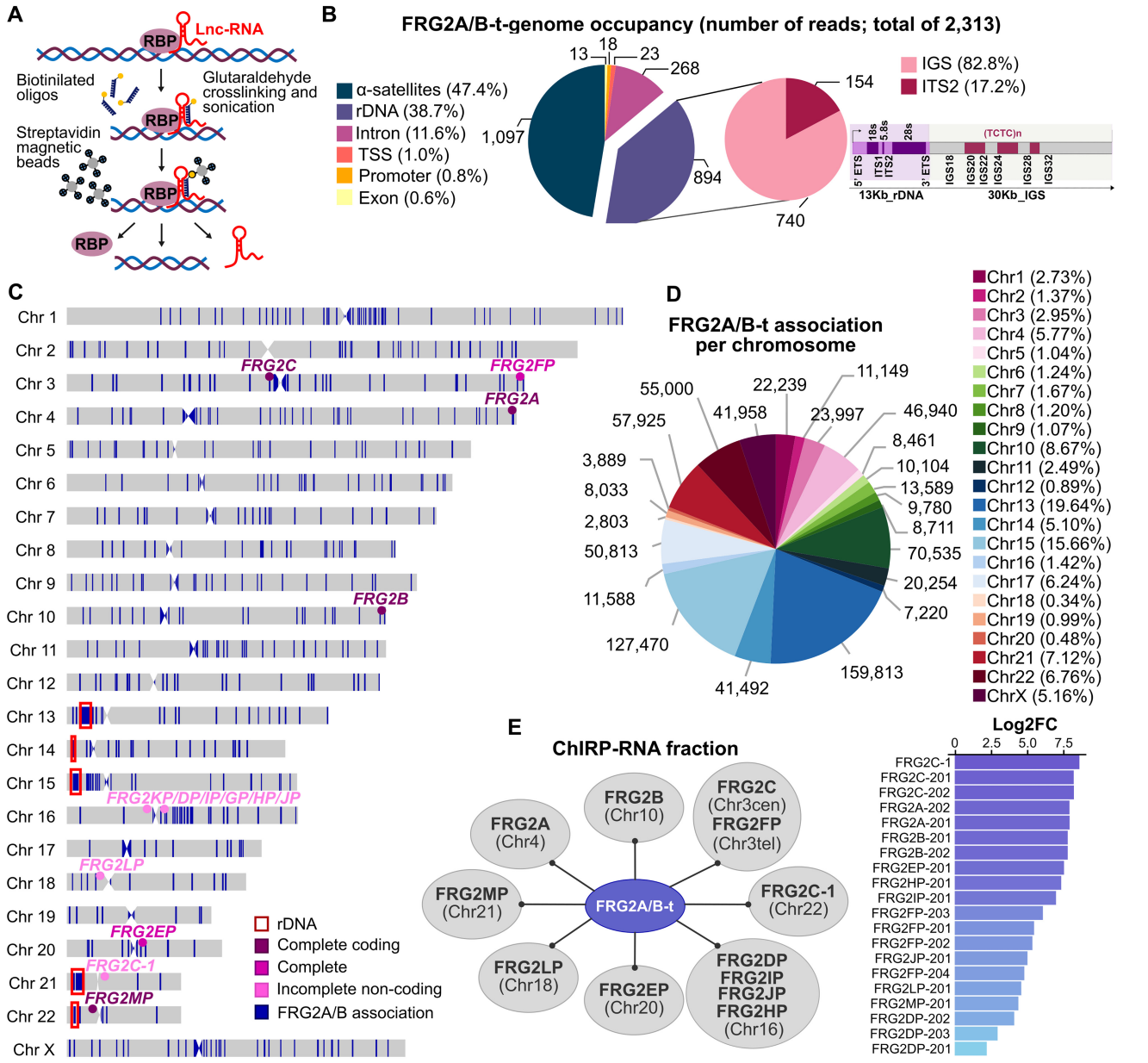


Figure 2. FRG2A/B are chromatin-associated RNAs preferentially enriched at rDNA and centromeric loci. (**A**) Schematic diagram of the ChIRP workflow. (**B**) Pie chart showing relative representation of various genomic regions obtained by the FRG2A/B-t ChIRP-DNA sequencing. A total of 2313 genomic target sites were found, with the enrichment of highly repetitive elements clustered into two main regions: the centromeric α-satellites (47.4%) and the ribosomal DNA arrays (38.7%). Within rDNA arrays, FRG2A/B-t was found to be predominantly associated with the IGS region (82.8%) and the ITS2 (17.2%), as illustrated in the schematic representation of the rDNA unit shown to the right. (**C**) Chromosome ideograms overlapped with FRG2A/B-t-association track displaying the enrichment of FRG2A/B-t at the centromeres and rDNA arrays. *FRG2* genes and rDNA arrays localization are indicated. (**D**) Pie-chart showing FRG2A/B-t association sites per chromosome, detected via ChIRP-DNA sequencing. Values are in bp and percentages are indicated in the legend. (**E**) FRG2 paralog transcripts interacting with FRG2A/B-t identified by ChIRP-RNA sequencing. Bar plot (right) shows $\log_2$ fold-change (ChIRP versus input) for each enriched transcript. Transcript IDs follow the CAT LiftOff annotation based on the T2T-CHM13 reference genome, and the numeric suffix (e.g. -201) indicates specific transcript isoforms derived from alternative splicing or promoter usage.

the number of D4Z4 repeats at 4q35 [14, 16]. To investigate *FRG2A* function, we obtained a comprehensive FRG2A-t interactome by using ChIRP (chromatin isolation by RNA purification, i.e. a "pull-down" on RNA) analysis [45, 46], which can simultaneously map RNA–RNA (ChIRP RNA-seq), RNA–protein (ChIRP-MS), and RNA–chromatin interactions (Fig. 2A) by oligonucleotides-mediated purification of a target lncRNA. The experiment targeted both FRG2A (Chr4q35) and FRG2B (Chr10q26) transcripts (FRG2A/B-t) due to the 99.5% identity between the two genes (Fig. 1A). This analysis uncovered a total of 2 313 genomic target sites for FRG2A/B-t, spanning 813 763 bp in HeLa cells (Fig. 2B and Supplementary Data S2).

Interestingly, most of FRG2A-t-associated regions occurred with two types of repetitive elements: 47.4% of read coverage was from centromeric satellites and 38.7% from ribosomal DNA (rDNA) arrays (Fig. 2B). Other types of repetitive regions such as telomeres or macrosatellites were not detected, nor were *in cis* interactions with the D4Z4 macrosatellite at the 4q35 and 10q26 subtelomeres. Figure 2C illustrates

that FRG2A/B-t preferentially associating *in trans* with the centromeres of all the human chromosomes, and within the rDNA arrays on all the acrocentric chromosomes. We also observed a preferential (>5% of the total reads) association *in trans* with chromosome 17 and X centromeres and *incis* with chromosome 4 and 10 centromeres (Fig. 2D).

Within the centromeres, the ChIRP data indicated that the most frequent FRG2A/B-t associations occurred at tandem arrays such as satellite DNA, including higher order repeats (HORs). HORs are multimeric structures formed from tandemly arranged repeat units that are characteristic of alpha satellite DNA within centromeric regions, and represent the sites of kinetochore assembly and CENP-A/B enrichment at the active centromeric repeat arrays [47] (Fig. 2B and Supplementary Fig. S7).

At rDNA loci, 82.8% of rDNA ChIRP reads mapped to the IGS region, and 17.2% to the 5.8S 5′ region overlapping the ITS2 region (Fig. 2B and Supplementary Fig. S8). The FRG2A/B-t associated regions within the IGS span from IGS20/22 to IGS30/32, corresponding to CT-microsatellites (Supplementary Fig. S8) [48].

To determine whether specific DNA motifs are enriched among FRG2A/B-t-associated sites, we searched the ChIRP peaks for consensus sequences, using the Geneious tool [49]. We analyzed all FRG2A/B-t-enriched loci collectively, and specific subsets of sites: all acrocentric chromosomes, individual acrocentric chromosomes, the most enriched centromeres, and individual centromeres from chromosomes 4, 10, 17, and X, as shown in Supplementary Fig. S9.

The most prominent motif identified was the (TCTC)*n* repeat. This sequence is prevalent in the genome-wide alignment consensus of FRG2A/B-t-associated genomic regions and was also found in the consensus binding site of each individual acrocentric chromosome. In contrast, the FRG2A/B-t associated centromeric regions diverged more substantially and did not share a large common consensus sequence. Instead, these centromeric regions are enriched for binding sequences for CENP-B, known as the CENP-B box [47] consistent with our finding that FRG2A/B-t binds to HOR sequences. These observations, together with the lack of direct sequence complementarity between FRG2A/B-t and its genomic targets (Supplementary Table S2) support the idea that structural features rather than primary sequence specificity drive FRG2A/B-t association. This raises the possibility that FRG2A/B-t interacts with higher-order chromatin conformations or secondary DNA structures, such as noncanonical DNA motifs or specific topological states of chromatin, which may be enriched in these repetitive regions. Alternatively, it may serve as a scaffold for proteins that recognize specific sequences and/or chromatin states. Notably, the specificity of FRG2A-t and FRG2B-t purification by ChIRP is supported by the fact that, among all FRG2 paralogs present in the genome, only the chromatin loci corresponding to FRG2A and FRG2B were enriched in the ChIRP-bound fraction (Supplementary Data S2). This indicates that the probes selectively recovered the target RNA–chromatin complexes at the expected loci, without detectable cross-binding to other FRG2 paralogs. In contrast, the ChIRP-RNA fraction obtained using the same oligonucleotide probes contained all FRG2s transcripts (FRG2s-t). In particular, ChIRP-RNA analysis revealed that multiple transcripts derived from distinct FRG2 paralogs are significantly and uniformly enriched, suggesting that chromatin association is a shared feature among these RNAs, rather than being driven by a single gene transcript. These data suggest that this newly discovered lncRNA family may participate in a large number of multivalent chromatin interactions. (Fig. 2E and Supplementary Data S2).

FRG2s RNAs localize in the nucleolus

The ChIRP data raised the possibility that FRG2 RNAs could be linked to nucleolar function. We therefore assessed their localization via RNA-FISH experiments in primary myoblasts from control and FSHD samples. We used two different fluorescent LNA probes: one specifically recognizing FRG2A-t (LNA647-FRG2A) and one recognizing both FRG2A-t and FRG2B-t (LNATY563-FRG2A/B) (Fig. 3C–H and Supplementary Fig. S10A–J). Confocal imaging of LNATY563-FRG2A/B revealed that FRG2A/B-t signal was superimposable with the nucleolus both in control and FSHD cells (Fig. 3C–H). To confirm the nucleolar localization of FRG2A/B-t, we performed RNA-FISH in Hela-KYOTO cells [50], which express a nucleolin-GFP fusion protein to facilitate visualization of the nucleolus. Supplementary Figure S10K–N shows the co-staining between LNATY563-FRG2A/B and nucleolin-GFP. RNase A (Supplementary Fig. S10, panels O–R) or RNase H treatment (Supplementary Fig. S10S–V) abolished the signal and confirmed that we were detecting FRG2A/B-t as DNA–RNA hybrids [51]. We concluded that FRG2A/B-t are enriched in the nucleolus.

To corroborate these findings, we isolated nucleoli from HeLa cells [35] (Fig. 3A and B). DNA and RNA were extracted from cytoplasmic, nuclear, nucleoplasmic and nucleolar fractions and the qPCR analysis of the DNA and RNA content in the different fractions was used to validate the purification of nucleoli (Fig. 3A and B). Analysis of the purified RNAs revealed that FRG2s transcripts were enriched in the nucleolar-RNA fraction (Fig. 3A).

To evaluate the precise localization of FRG2A/B-t within the nucleolus we performed immuno-RNA FISH in primary myoblasts using LNATY563-FRG2A/B and specific antibodies against upstream-binding factor (UBF), fibrillarin (FBL), and nucleolar protein Nop52 (NOP52) to visualize the fibrillar center (FC), dense fibrillar component (DFC), and granular component (GC) regions, respectively (Supplementary Fig. S10W–V'). We also performed these experiments after actinomycin-D (ActD)-induced nucleolar segregation [52], which helps to discriminate nucleolar compartments. In ActD-treated cells, FRG2A/B-t staining was visible as discrete dots symmetrically organized around the nucleolus (Fig. 3I–N and Supplementary Fig. S11A–L) that were superimposable with nucleolar caps, usually appearing as one or more dots surrounding the central GC body.

Co-staining experiments revealed that FRG2A/B-t was not primarily associated with the UBF-positive FC compartment (Supplementary Fig. S11M–T) and only partially colocalized with the NOP52-positive GC (Supplementary Fig. S11U–B'). Instead, the highest degree of overlap was observed between FRG2A/B-t and FBL (Supplementary Fig. S11C'–J'). Notably, FSHD cells did not differ from control cells, displaying a comparable localization of nucleolar components and FRG2A/B-t.

To better characterize the spatial relationships among FBL, UBF, and FRG2A/B-t, we performed simultaneous immunofluorescence labeling of these three nucleolar species in

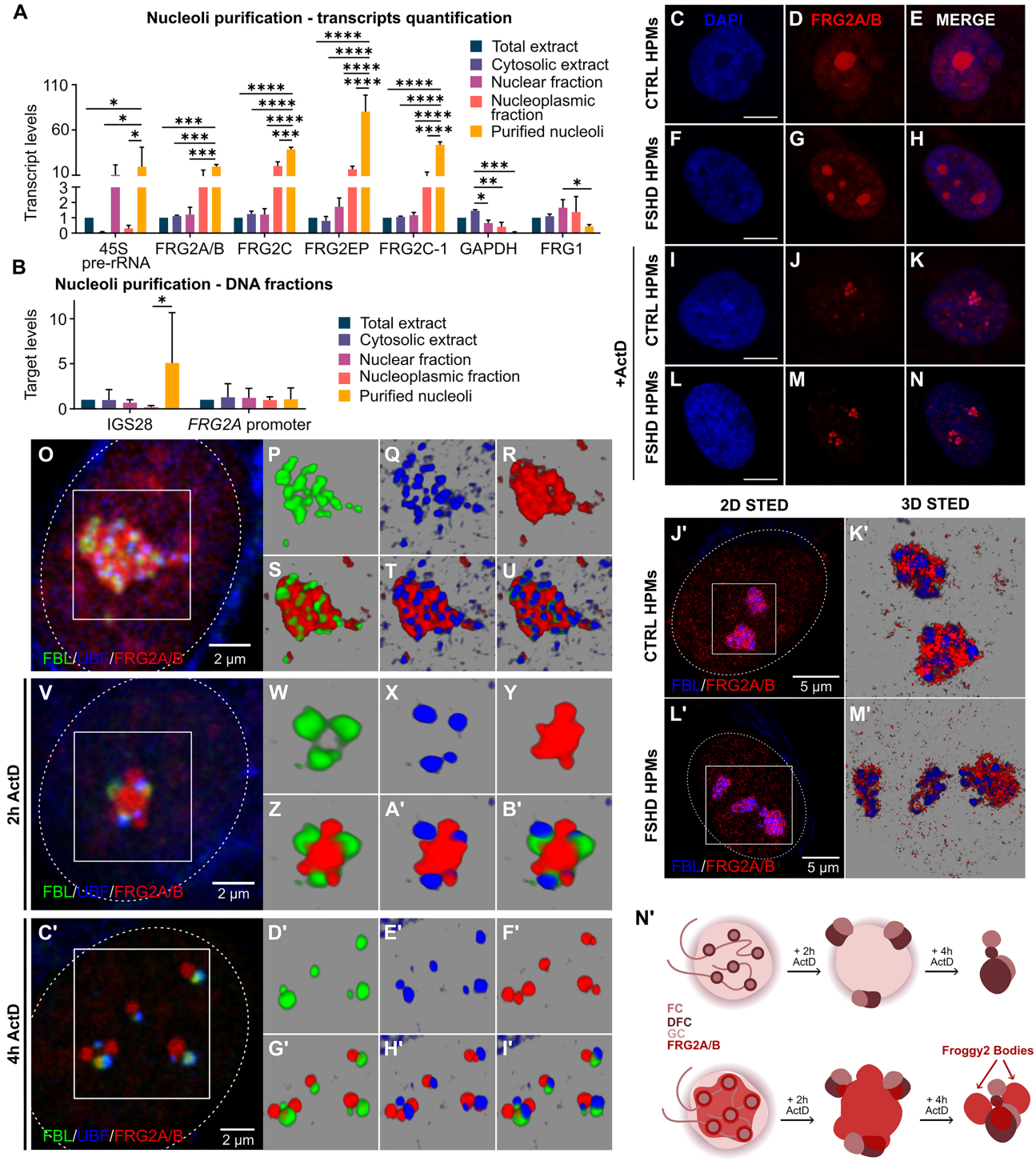


Figure 3. FRG2A/B RNA defines a previously undescribed nucleolar domain. (**A**) RT-qPCR of RNA from HeLa cell fractionation confirmed FRG2 paralogs enrichment in the nucleolar fraction, while protein-coding transcripts (GAPDH and FRG1) showed no enrichment (two-way ANOVA, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$). (**B**) RT-qPCR on DNA fractions from HeLa cell fractionation validated assay success. The nucleolar target (rDNA region IGS28) was enriched in the nucleolar fraction, whereas the non-nucleolar region (*FRG2A* promoter) showed no enrichment (two-way ANOVA, $*P < 0.05$). (**C–N**) RNA-FISH with FRG2A/B-specific fluorescent LNA probe (red) in HPMs from CTRL and FSHD subjects. Nucleolar segregation was induced in HPMs (I–N) by actinomycin D (ActD, 1 µg/ml, 4 h); scale bar: 5 µm. (**O–I'**) Immuno-RNA-FISH with anti-UBF, anti-FBL, and FRG2A/B-specific LNA probe in FSHD-HPMs, untreated (O–U) or ActD-treated (1 µg/ml) for the indicated times (**V–**I'). Panels P–U, W–B', and D'–I' show 3D renderings of respective confocal images (O, V, and C'). Dotted circles indicate nuclei, white squares highlight zoomed regions. (**J'–M'**) STED confocal images of immuno-RNA-FISH with anti-FBL (blue) and FRG2A/B LNA probe (red) in untreated CTRL and FSHD-HPMs. Panels K' and M' display 3D renderings of images (J' and L'). Circles indicate nuclei, white squares mark zoomed regions. (**N'**) Schematic representation of nucleolar structure in untreated and ActD-treated cells, showing FRG2A/B-t localization and Froggy2 Bodies formation.

untreated cells as well as after 2 and 4 h of ActD treatment. The 3D rendering of confocal images shown in Fig. 3O–U demonstrates that, in untreated cells, FRG2A/B-t overlaps and surrounds FBL and UBF foci, suggesting its localization in both the DFC and the GC. Following nucleolar segregation (Fig. 3V–I'), FRG2A/B-t becomes clearly visible in distinct bodies at 2 h post-treatment, which we termed "Froggy2 Bodies" (Fig. 3V–B'). By 4 h, these structures are spatially separated from both UBF and FBL, yet they remain in close proximity to FBL, maintaining a tight association with partial overlap (Fig. 3C'–I').

This observation was validated using stimulated emission depletion (STED) microscopy with FRG2A/B-t and FBL (Fig. 3J'–N'). Super-resolution images confirmed that FRG2A/B-t does not perfectly colocalize with FBL even in untreated cells; instead, it surrounds FBL-positive structures, partially overlapping them while extending into the GC.

Additionally, we tested the colocalization of FRG2A/B-t with other nuclear bodies [53], such as PML bodies (PML), speckles (SC35), paraspeckles (SFPQ), and Cajal bodies (SMN), as shown in Supplementary Fig. S12. To perform this analysis, we treated cells with ActD, to evaluate the possibility that FRG2A/B-t interacts with these nuclear bodies after nucleolar disruption. Our results show that FRG2A/B-t is not present in any of these well-characterized nuclear bodies suggesting that it is localized in different, previously unidentified structures.

To further explore the subnuclear localization of FRG2A/B-t, we assessed its colocalization with various nuclear bodies, including PML bodies (PML), speckles (SC35), paraspeckles (SFPQ), and Cajal bodies (SMN), as illustrated in Supplementary Fig. S12. To investigate whether FRG2A/B-t relocalizes upon nucleolar disruption, we treated cells with actinomycin D (ActD). This analysis revealed that FRG2A/B-t does not accumulate at these structures following nucleolar stress, suggesting that it predominantly resides in distinct nuclear compartments.

Based on these observations, we concluded that FRG2A/B-t is associated with nucleoli, with a unique localization at the interface between the FBL-positive dense FCs and the GC. This hypothesis is further supported by the FRG2A/B-RNAs association profile obtained by ChIRP (Supplementary Text, Supplementary Data S2, Supplementary Table S3, Fig. S14, and Supplementary Data S5), showing a prevalence of noncoding-Alu containing RNAs [54, 55] and a partial overlap with the 47S rRNA interactome [56]. We therefore set out to test how FRG2A/B-t might contribute to the structure and function of nucleoli.

FRG2A/B-t promote centromere–nucleolar associations in FSHD cells

The genomic association profile of FRG2A/B-t revealed frequent associations with rDNA arrays and with the centromeric regions of all chromosomes (Fig. 2). In mammalian cells, these repetitive, heterochromatic regions are enriched around the periphery of nucleoli, as part of a subset of the genome known as nucleolus-associated domains (NADs) [57–59]. This prompted us to investigate the interaction between FRG2A/B-t and NADs. We first analyzed existing NADs mapping data [34] from HeLa cells by using the NAD finder tool [35] and the T2T-CHM13 genome assembly. Graphing the NADs (Supplementary Data S3, Fig. 4, and Supplementary Fig. S13) together with the genomic association profile of FRG2A/B-t obtained by ChIRP (Fig. 2C) revealed a significant (Supplementary Fig. S13) extensive overlap between these two features (Fig. 4A and B). Quantifying these data, we found that 86.3% of FRG2A/B-t- associated regions overlap with NADs in HeLa cells (Fig. 4A). The overlap included NADs on the rDNA-containing acrocentric chromosome arms (50.4%) and centromeric regions (49.6%) (Fig. 4A and B).

To address these issues, we investigated the localization of centromeres and nucleoli by immunofluorescence, using antibodies raised against kinetochore protein CENP-B and the nucleolar protein FBL, in HPMs (Fig. 4C–J). We measured the distance between CENP-B foci and nucleoli (FBL staining for DFC), and we detected a significant increase in the number of centromeres present at the NP in FSHD samples (Fig. 4K). To confirm that this difference results from *FRG2A* overexpression rather than any other difference in genotype, we performed siRNA-mediated depletion of FRG2A/B-t in FSHD cells (Supplementary Fig. S15), which reduced the frequency of centromere-nucleolus associations (Fig. 4L–G'). Together, these experiments indicate that FRG2A/B-t contributes to the clustering of centromeres at nucleoli and indicates that this is a dynamic process.

To better visualize these interactions, we performed immuno-RNA FISH experiments on isolated nucleoli from HeLa cells (Fig. 4H'–M'). The purified nucleoli clearly displayed FRG2A/B-t enrichment that was highly overlapping with FBL, and surrounded by foci of centromeric CENP-B.

Together, our ChIRP and microscopy data indicated that FRG2A/B-t associates with rDNA repeats and centromeres, two main components of NADs. Furthermore, FRG2A/B-t promotes association of centromeres with the NP demonstrating an unforeseen contribution to nuclear compartmentalization. In this scenario, the increase in nucleoli-centromere contacts in FSHD cells could alter the structure of the perinucleolar chromatin, with FRG2A/B-t potentially acting as a molecular linker that drives changes in the epigenetic or transcriptional state of the associated loci.

The increased levels of FRG2A/B-t in FSHD cells impair nucleolar function

The localization of FRG2A/B-t within the DFC region of the nucleolus suggested functional role(s) in rRNA transcription or processing [60]. To test this possibility, we first examined processing of rRNA transcripts in FSHD cells, but no defects were observed (Supplementary Fig. S16), suggesting that post-transcriptional processing of rRNAs is not a regulatory target of FRG2 RNAs [60]. This is consistent with the observation that FRG2A/B is associated with the rDNA arrays (Fig. 2C), but not with any rRNA species (Supplementary Data S2).

We therefore turned our attention to the potential roles of FRG2A/B-t in the regulation of rDNA transcription. Each repeat unit of rDNA contains a 13 kb region that is transcribed to produce the 45S rRNA precursor RNA (45S pre-rRNA), which is processed into mature 18S, 5.8S, and 28S rRNAs. Between each *45S* gene, rDNA repeats also contain a 30kb-long intergenic spacers (IGS), which include short repetitive sequences [48] (Supplementary Fig. S8).

Several lncRNAs regulate rDNA transcription and help to maintain nucleolar structures and functions [61–63]. We

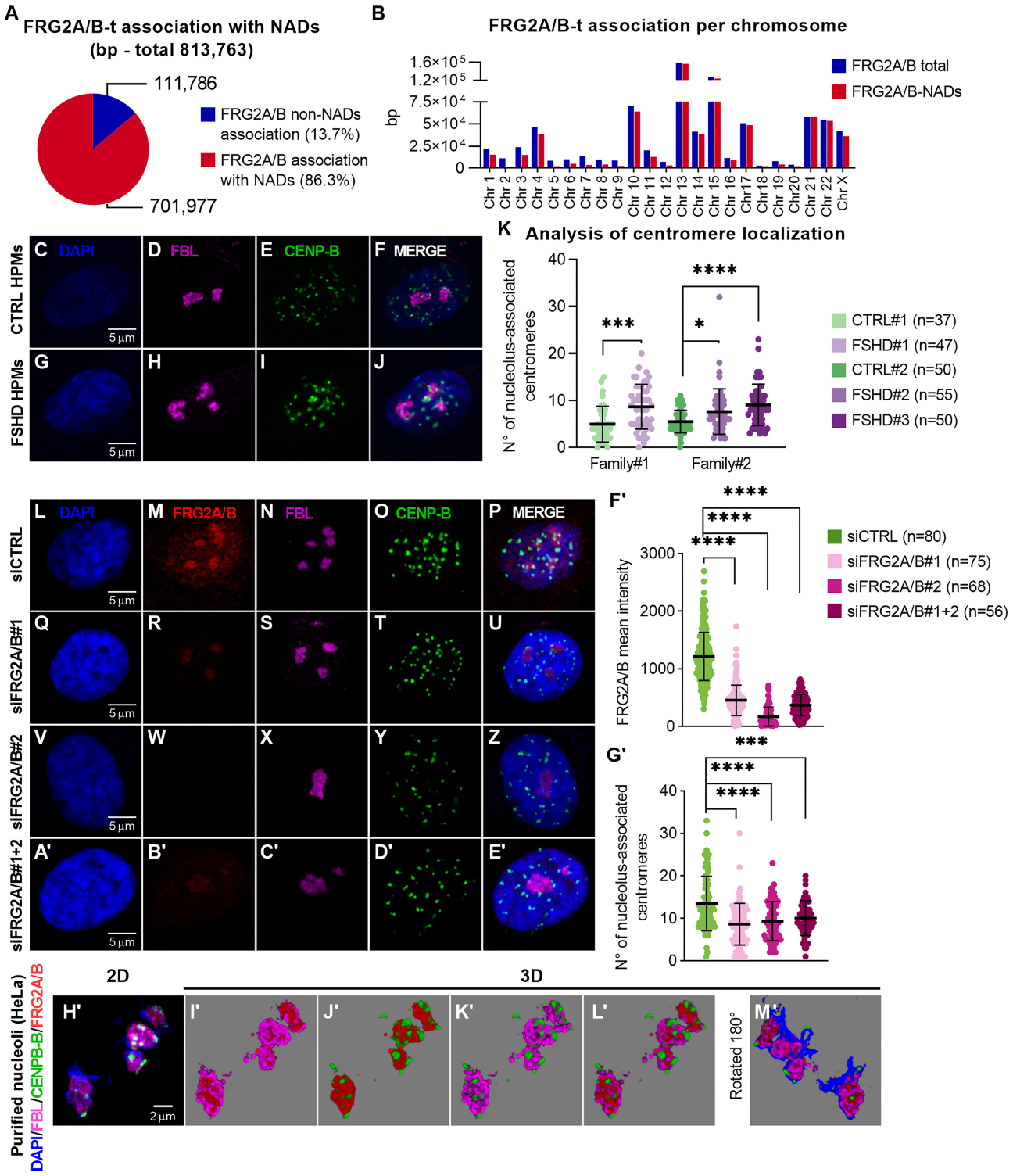


Figure 4. FRG2A/B-t mediates centromere association with the NP in FSHD cells. (**A**) Pie chart quantifying FRG2A/B-t-associated regions overlapping with NADs in HeLa cells (from existing data). Statistical analysis in Supplementary Fig. S13. (**B**) The plot shows the total number of base pairs (bp) associated with FRG2A/B-t across each chromosome, along with the subset overlapping with NADs (FRG2A/B-NADs). (**C–J**) Immunofluorescence with anti-FBL (magenta) and anti-CENP-B (green) in CTRL and FSHD HPMs showed increased nucleolar-centromere association in FSHD. (**K**) Quantification of nucleolus-associated centromeres per cell (from C to J), defined as CENP-B foci colocalizing with FBL or within 140 nm from FBL. Two families (CTRL and FSHD) were analyzed using ScanR software. Mann–Whitney test ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$). (**L–E'**) Immuno-RNA-FISH with anti-FBL (magenta), anti-CENP-B (green), and FRG2A/B probe (red) in FSHD HPMs after FRG2A/B siRNA silencing (siFRG2A/B#1, #2, or both). Scrambled siRNA (siCTRL) was used as control. FRG2A/B signal was reduced (R, W, B'), and centromere localization was restored (T, Y, D'). (**F'**) FRG2A/B mean intensity quantification (from L to E') confirmed signal reduction (two-way ANOVA, $*P < 0.05$, $***P < 0.001$, $****P < 0.0001$). (**G'**) Quantification of nucleolus-associated centromeres in FRG2A/B-silenced cells showed a significant reduction (two-way ANOVA, $*P < 0.05$, $***P < 0.001$, $****P < 0.0001$). (**H'–M'**) Immuno-RNA-FISH with anti-FBL (magenta) and anti-CENP-B (green), and with FRG2A/B-specific probe (red), in nucleoli purified from HeLa cells. Panels I'–L' show 3D rendering of (H' and J'), and (M') shows a rotated 3D view.

thus reasoned that the association of FRG2A/B-t with IGS regions might have a similar role. To explore this possibility, we measured the levels of 45S pre-rRNA transcript in muscle samples and fibroblasts from FSHD affected patients and unaffected siblings. As shown in Fig. 5A and Supplementary Fig. S17A, we observed significantly lower levels of 45S pre-rRNA transcripts in FSHD myoblasts but not in fibroblasts.

This reduction in pre-rRNA levels in FSHD patients' cells was accompanied by a significant increase of heterochromatin-associated histone mark H3-K9me3 in FSHD cells in comparison with control cells (Fig. 5B). These elevations were observed within the transcribed region and the IGS, including at the rDNA promoter (IGS42.9, Supplementary Fig. S8). These findings suggest a double-negative regulatory pathway, whereby the reduction in D4Z4 chromatin leads to increased expression of a regulatory lncRNA, which in turn suppresses nascent rRNA transcription. To clarify for the reader, this double-negative logic means that a decrease in D4Z4 repeats paradoxically results in reduced rRNA synthesis, due to the intermediate upregulation of a repressive lncRNA. This counterintuitive mechanism is schematized as ↓D4Z4 → ↑lncRNA → ↓nascent rRNA.

To test the functional contribution of FRG2A/B-t to the regulation of rDNA transcription, we examined the effects of their siRNA-mediated depletion. We observed increased levels of 45S pre-rRNA transcripts after silencing both in HeLa cells and in HPMs derived from FSHD patient biopsies (Fig. 5C and Supplementary Fig. S17B).

To bolster the previous steady-state measurements, we pulse-labeled living cells with EU followed by "click-it chemistry" [64] to label and capture newly synthesized RNAs. Nucleolar incorporation of the EU was quantified as the ratio of nucleolar EU/nuclear EU (Fig. 5D and Supplementary Fig. S18A–H). Via biotinylation of EU-labeled-newly synthesized RNAs, we performed a nascent RNA capture assay and quantified specific transcripts by RT-qPCR (Fig. 5E). These experiments detected significantly lower levels of nascent 45S pre-rRNA in FSHD samples compared to control samples.

Paralleling the steady state readout, the nascent 45S pre-rRNA levels were increased in FSHD cells upon silencing of FRG2A/B (Fig. 5F). Furthermore, these assays confirmed that FRG2A/B are actively transcribed in myoblasts derived from FSHD patients but not in myoblasts from healthy family members (Fig. 5E). Together, these results establish a previously unanticipated regulatory role in rRNA synthesis for these lncRNAs. Interestingly, in addition to the transcriptional changes, we observed that FSHD cells display a significant increase in the average number of nucleoli per cell, and a reduction of the total area occupied by nucleoli per cell (Fig. 5G and H, Supplementary Data S6, and Supplementary Fig. S10 U'–X'). This is consistent with previous data indicating that the size of the nucleolus positively correlates with rRNA synthesis and that the downregulation of rRNA gene transcription leads to a reduction in nucleolar size [65].

We hypothesized that these altered nucleolar characteristics would correlate with changes in protein synthesis capacity. To test this, we measured HPG incorporation as a label for newly synthesized proteins. Indeed, we found that the global rate of protein synthesis in FSHD cells was reduced compared with that of cells from their healthy relatives (Fig. 6A). The involvement of FRG2A/B-t in this process was confirmed because siRNA-mediated depletion of the transcript reversed the protein synthesis defects (Fig. 6B) Based on these findings, we conclude that the negative regulation of rRNA transcription imposed by excessive levels of FRG2A/B-t results in the impairment of protein synthesis in FSHD cells.

A compensatory ribosomal protein response in FSHD cells disrupting muscle protein synthesis

Our Supplementary Data show a significant decline in pre-rRNA levels and total protein synthesis in FSHD cells, a phenotype mediated by FRG2A/B-t. These findings prompted us to employ mass spectrometry to comprehensively profile the proteomic alterations underlying these defects.

We compared FSHD and control samples from family #2 (CTRL#2 and #3 and FSHD #2 and #3), first using undifferentiated myoblasts. Gene Ontology (GO) [66] analysis of differentially expressed proteins (DEPs) in FSHD myoblasts revealed a marked decrease in protein levels from pathways linked to ribosome biogenesis, translation, protein transport, and chromatin remodeling, coupled with a significant increase in levels of splicing-related factors (Fig. 6C and Supplementary Data S4).

To investigate how these alterations are affected during differentiation, we next compared protein levels in cells derived from healthy and FSHD patients that were differentiated from myoblasts towards a myotube state. Myoblasts were transferred to differentiation medium and analyzed 8 days after induction. We next identified the proteins that were expressed differently during the transition from myoblast to myotube in FSHD versus controls. We observed that differentiation triggered upregulation of ribosomal proteins and rRNA processing factors in FSHD samples. This suggests that differentiating FSHD myoblast cells attempt to counteract their translation deficits by enhancing ribosome assembly.

However, despite this compensatory response, many muscle-specific proteins, particularly those involved in sarcomere organization, contraction, and other high-demand translational processes, remained underrepresented in FSHD myotubes (Fig. 6D). This selective deficit suggests that the translational machinery in FSHD cells is unable to fully support the synthesis of functionally important muscle proteins.

Our findings indicate that the initial ribosomal impairment in FSHD myoblasts leads to an attempted rescue during differentiation, but this compensation is incomplete and may result in the production of non-standard ribosomal complexes. We hypothesize that in FSHD cells, the reduced availability of rRNA could result in excess, unassembled ribosomal proteins, or could promote formation of noncanonical, "specialized" ribosomes with altered composition and function [67]. In this manner, FSHD might share some characteristics with ribosomopathies, in which defects in ribosome function disproportionately affect specific protein classes [68]. In FSHD, this appears to selectively impact muscle proteins, providing a mechanistic link between the ribosomal dysfunction and the muscle pathology observed in the disease (Fig. 6E).

Discussion

Here, we discover a previously unrecognized family of regulatory lncRNAs. These FRG2 family members are localized to nucleoli. We show that the FRG2A/B-t RNAs are upregulated

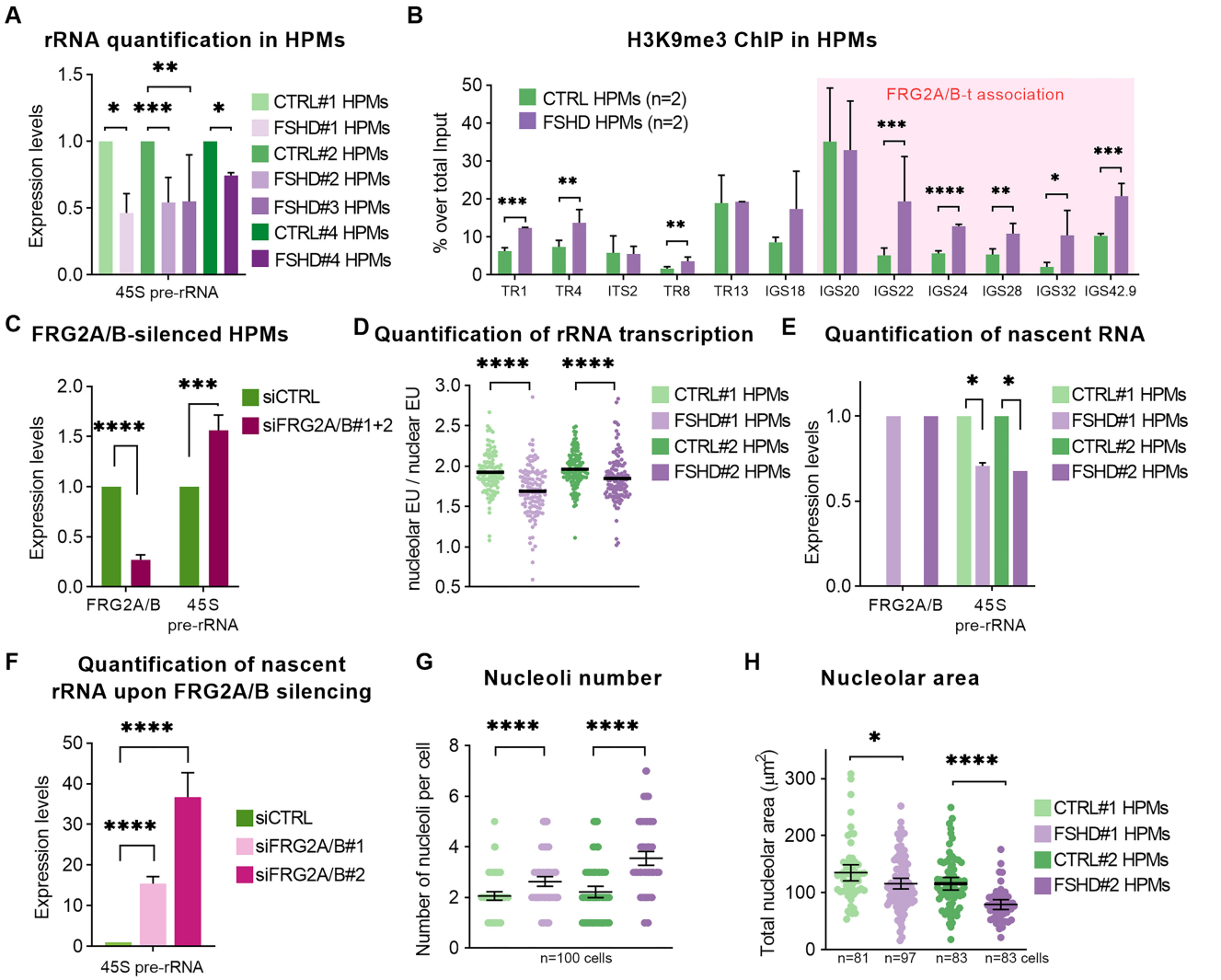


Figure 5. FRG2A/B-t inhibits rDNA transcription in FSHD myoblasts. (**A**) RT-qPCR of HPMs cDNA of CTRL and FSHD subjects. Statistical significance was tested by two-way ANOVA analysis. *P*-value (**P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001). (**B**) Chromatin-immunoprecipitation (ChIP) performed in CTRL and FSHD-derived HPMs showing H3K9me3 enrichment at IGS and rDNA promoter. Data from two CTRLs and two FSHD patients were averaged ($n = 2$). FRG2A/B-associated loci are indicated by the pink box (see Supplementary Fig. S8). Statistical significance was tested via two-way ANOVA analysis. (**C**) RT-qPCR of FSHD HPMs cDNA upon FRG2A/B siRNA-mediated silencing. Statistical significance was tested by using the two-way ANOVA analysis. *P*-value (**P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001). (**D**) Via EU incorporation, rRNA transcription was quantified as the ratio between nucleolar/nuclear EU fluorescent signals. Statistical significance was tested by using the Mann–Whitney test. (**E** and **F**) RT-qPCR of nascent RNAs captured by EU incorporation in CTRL and FSHD HPMs (E), and upon siFRG2A/B treatments of FSHD-HPMs (F). Statistical analysis was performed by using the two-way ANOVA test. (**G** and **H**) The number of nucleoli per cell and the total nucleolar area per cell were measured in FSHD and CTRL HPMs. Statistical significance was tested by using the Mann–Whitney test. Each FSHD sample was compared to a healthy control selected from a related family member (matched control), chosen based on availability and degree of relatedness. This family-based design ensures consistency across comparisons.

in FSHD patients and repress rDNA transcription, thereby diminishing translational capacity. This regulatory cascade follows a double-negative logic: loss of D4Z4 repression leads to increased expression of FRG2 lncRNAs, which associate with chromatin at rDNA and centromeric regions, and are linked to reduced nascent rRNA synthesis.

Protein synthesis is a key molecular process for skeletal muscle maintenance and growth and requires ribosome biogenesis. However, ribosome biogenesis is reduced under catabolic states, in particular during decreased skeletal muscle contractile activity [69]. This suggests that maintenance above a certain threshold has the potential to reduce or prevent muscle atrophy during disuse, aging and disease [69, 70].

In the case of FSHD, the association of FRG2A/B-t with ribosomal DNA and its impact on protein synthesis suggests a link between nucleolar dysfunction and the development of FSHD pathology. We have demonstrated that FRG2A-t is expressed in muscle cells and overexpressed in cells derived from FSHD patients. Therefore, conditions that increase levels of FRG2A/B-t are expected to impair ribosome biogenesis and protein synthesis, with negative effects on muscle functions eventually leading to disease. Supporting this prediction, genotoxic agents and DNA damage are known to contribute to muscle wasting and aging when they persist over time [71, 72].

While the primary function of the nucleolus is ribosome biogenesis [73], additional biological processes occur in

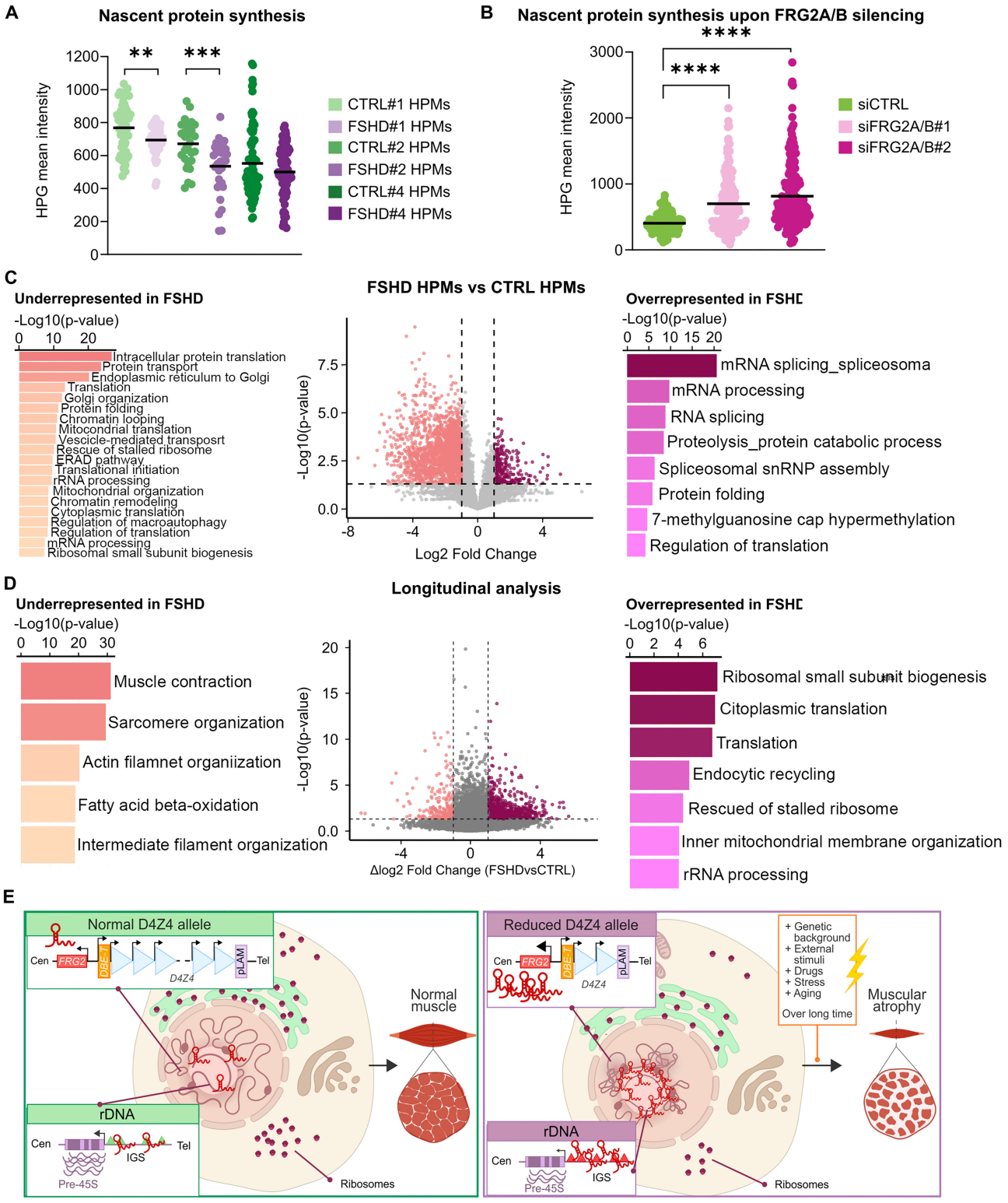


Figure 6. Mass spectrometry revealed impaired synthesis of ribosomal and muscle-specific proteins in FSHD. (**A** and **B**) Quantification of nascent proteins in CTRL and FSHD HPMs at basal conditions (A) and after FRG2A/B silencing in FSHD-HPMs (B). HPG incorportion in newly translated proteins was quantified as cytoplasmic HPG mean intensity. Statistical significance was tested with the Mann–Whitney test $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$). (**C** and **D**) Mass spectrometry analysis in CTRL and FSHD HPMs and myotubes. Histograms show the most significant misregulated GO molecular function terms ($-\log_{10}$ *P*-value > 5, by using DAVID annotation [49]). Upregulated and downregulated proteins in FSHD are in purple and pink, respectively. (**C**) Volcano plot comparing FSHD versus CTRL HPMs. (**D**) Volcano plot of the longitudinal analysis: DEPs identified in myoblast-to-myotube transition in FSHD and controls were used to compare conditions by calculating the delta $\log_2$FC. (**E**) A hypothesis for the contribution of FRG2 lncRNAs to FSHD pathogenesis. In individuals carrying a normal-sized D4Z4 allele at 4q35, FRG2A is expressed at low levels, and has no effects on genome architecture or rDNA arrays, allowing normal muscle growth and function. In subjects carrying a reduced D4Z4 allele (≤10 repeats), FRG2A is overexpressed, altering nucleolar chromatin interactions and impairing rDNA transcription, reducing protein synthesis and potentially contributing to muscle atrophy over time.

nucleoli, for example the processing of the signal recognition particle (SRP) [74]. It is also clear that the NP is an important hub for organizing the large heterochromatin subset of the genome known as NADs [34, 59, 57]. Our studies show that FRG2 lncRNAs regulate interactions at the NP promoting centromere clustering and rDNA transcription. Similarly, another lncRNA that contributes to chromosome anchoring to the NP is Firre, which is required for the inactivation of chromosome X (Xi) [75]. How lncRNAs regulate both nuclear architecture and gene expression are major questions for nuclear cell biology.

The discovery of the FRG2 lncRNA family is a foundational step towards a new paradigm in which repetitive elements exert *trans*-functions via lncRNAs.

This raises the intriguing possibility that distinct FRG2 paralog RNAs could co-associate with the same *trans*-acting regulatory complexes, potentially acting in different combinations to influence chromatin states or transcriptional activity at other genomic loci. While this remains speculative, it aligns with emerging evidence that lncRNAs can function as scaffolds or guides in multi-component regulatory assemblies.

Given that FRG2A lncRNA predominantly localizes to the interior of nucleoli, while the chromatin changes are observed at the NP, we cannot exclude the possibility that these effects are indirect. Such changes may be mediated by intermediate RNA-binding proteins or alterations in nucleolar organization that impact perinucleolar heterochromatin structure.

An open question raised by our findings concerns the molecular composition and function of the Froggy2 bodies. While our data clearly show the accumulation of FRG2A/B t RNAs in these structures, it remains to be determined which molecular partners they associate with. Future work will be needed to identify their constituents and to define whether these bodies represent functional hubs for chromatin regulation, RNA processing, or nuclear organization

Noting that one contributing factor to FSHD disease is the deletion of D4Z4 macrosatellites near the *FRG2A* locus, we speculate that FRG2A represents just one example of repeat-regulated lncRNAs interacting *in trans* with many genomic loci and representing architectural and regulatory hubs in the mammalian cell nucleus.

Other lncRNAs play key roles in the formation of membrane-less nuclear bodies through liquid–liquid phase separation, serving as a scaffold to anchor proteins and mRNAs [76, 77], but in the previously studied cases there is a single locus encoding the lncRNAs. When encoded or regulated by repetitive loci, the potential for complexity escalates.

In this scenario, we speculate that the effects of *FRG2A* expression on translation might be affected not only by copy number variations of D4Z4 but also by the variable number of rDNA repeats. That is, the effects of FRG2 lncRNAs on rDNA or other TARls targets could depend on the ratio of the numbers of each element. Indeed, we propose that the FRG2 lncRNAs not only contribute to FSHD disease phenotypes but also provide a measure of the intra- and inter-familiar phenotypic variability associated with this disease. These ideas may propel the discovery of new roles for repetitive elements in cell biology and provide robust systems to test the effects of TARls Copy number variations (CNVs) in shaping cell functions and their response to external stimuli.

Acknowledgements

We are indebted to all patients and their families for their participation in this study. We thank Jhonathan Vinet for technical assistance and imaging; Serena Carra and Floriana Della Ragione for providing protocols and reagents; Donata Orioli for her comments and assistance in editing the manuscript, Matteo Chiara for his bioinformatic training to SP and Caterina Manzari for her technical help in NGS-libraries preparation. We extend our gratitude to Professor Michael R. Green for his contributions to our research, and we honor his memory.

Author contributions: Valentina Salsi (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Visualization [equal], Writing—original draft [equal]), Francesca Losi (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Investigation [equal], Visualization [equal], Writing—original draft [equal]), Bruno Fosso (Data curation [equal], Formal analysis [equal], Resources [equal]), Marco Ferrarini (Formal analysis [equal], Investigation [supporting]), Sara Pini (Data curation [equal], Formal analysis [equal]), Marcello Manfredi (Funding acquisition [supporting], Investigation [supporting], Resources [supporting]), Gaetano Vattemi (Resources [supporting]), Tiziana Mongini (Resources [supporting]), Lorenzo Maggi (Resources [supporting]), Graziano Pesole (Funding acquisition [supporting], Writing—review & editing [supporting]), Anthony K. Henras (Funding acquisition [supporting], Investigation [supporting], Resources [equal], Supervision [supporting], Writing—review & editing [supporting]), Paul D. Kaufman (Funding acquisition [supporting], Supervision [supporting], Writing—review & editing [equal]), Brian McStay (Funding acquisition [supporting], Resources [equal], Supervision [supporting], Writing—review & editing [supporting]), and Rossella Tupler (Conceptualization [equal], Funding acquisition [lead], Resources [lead], Supervision [equal], Writing—review & editing [lead]).

Supplementary data

Supplementary data is available at NAR online.

Conflicts of interest

The authors declare no competing interests.

Funding

The research leading to these results has received funding from the European Union -NextGenerationEU through the Italian Ministry of University and Research under PNRR - M4C2-I1.3 Project PE_00000019 “HEAL ITALIA” to Rossella Tupler CUP E93C22001860006 of University of Modena and Reggio Emilia. The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them; National Institutes of Health [RO1NS0475840 to R.T., U01 CA260669 to P.D.K., R35 GM15220 1 to P.D.K.]; FSHD Global Research Foundation to R.T.; Cariplo-Telethon Alliance [GJC23060 to R.T.]; FAR-FOMO 2021 to R.T.; Wellcome Trust to B.MS.; Investigator Award [223049/Z/21/Z to B.MS.]; LSH-Puglia

[T4-AN-01 H93C22000560003 to G.P.]; INNOVA [PNC-EJ-2022-23683266 PNC-HLS-DA to G.P.]; ELIXIRNextGenIT [IR0000010 to G.P.]; REA1COM to A.K.H.; DUKKED to A.K.H.; and MUR "Departments of Excellence 2023-2027", AGING Project to M.M. Funding to pay the open access publication charges for this article was provided by Cariplo-Telethon Alliance GJC23060.

Data availability

Correspondence and requests for materials should be addressed to Prof. Rossella Tupler (rossella.tupler@unimore.it). The datasets generated during and/or analyzed during the current study are available in the NIH repository, BioProject: PRJNA1122505 and ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD053349.

References

1. Nurk S. The complete sequence of a human genome. *Science* 2022;**376**:44–53. https://doi.org/10.1126/science.abj6987
2. Dumbovic G, Forcales SV, Perucho M. Emerging roles of macrosatellite repeats in genome organization and disease development. *Epigenetics* 2017;**12**:515–26. https://doi.org/10.1080/15592294.2017.1318235
3. Winokur ST, Bengtsson U, Feddersen J *et al.* The DNA rearrangement associated with facioscapulohumeral muscular dystrophy involves a heterochromatin-associated repetitive element: implications for a role of chromatin structure in the pathogenesis of the disease. *Chromosome Res* 1994;**2**:225–34. https://doi.org/10.1007/BF01553323
4. Hewitt JE, Lyle R, Clark LN *et al.* Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystropothhy. *Hum Mol Genet* 1994;**3**:1287–95. https://doi.org/10.1093/hmg/3.8.1287
5. Wijmenga C, Hewitt JE, Sandkuijl LA *et al.* Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet* 1992;**2**:26–30. https://doi.org/10.1038/ng0992-26
6. Ricci G, Ruggiero L, Vercelli L *et al.* A novel clinical tool to classify facioscapulohumeral muscular dystrophy phenotypes. *J Neurol* 2016;**263**:1204–14. https://doi.org/10.1007/s00415-016-8123-2
7. Ricci G, Mele F, Govi M *et al.* Large genotype–phenotype study in carriers of D4Z4 borderline alleles provides guidance for facioscapulohumeral muscular dystrophy diagnosis. *Sci Rep* 2020;**10**:1–12. https://doi.org/10.1038/s41598-020-78578-7
8. Ruggiero L, Mele F, Manganelli F *et al.* Phenotypic variability among patients with D4Z4 reduced allele facioscapulohumeral muscular dystrophy. *JAMA Netw Open* 2020;**3**:e204040. https://doi.org/10.1001/jamanetworkopen.2020.4040
9. Salsi V, Chiara M, Pini S *et al.* A human pan-genomic analysis reconfigures the genetic and epigenetic make up of facioscapulohumeral muscular dystrophy. medRxiv, https://doi.org/10.1101/2023.06.13.23291337. 02 August 2023, preprint: not peer reviewed
10. Salsi V, Vattemi GNA, Tupler RG. The FSHD jigsaw: are we placing the tiles in the right position? *Curr Opin Neurol* 2023;**36**:455–63.
11. Snider L, Geng LN, Lemmers RJLF *et al.* Facioscapulohumeral dystrophy: incomplete suppression of a retrotransposed gene. *PLoS Genet* 2010;**6**:1–14. https://doi.org/10.1371/journal.pgen.1001181
12. Ricci G, Zatz M, Tupler R. Facioscapulohumeral muscular dystrophy: more complex than it appears. *CMM* 2014;**14**:1052–68. https://doi.org/10.2174/1566524014666141010155054
13. Ricci G, Scionti I, Sera F *et al.* Large scale genotype-phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain* 2013;**136**:3408–17. https://doi.org/10.1093/brain/awt226
14. Gabellini D, Green MR, Tupler R. Inappropriate gene activation in FSHD: a repressor complex binds a chromosomal repeat deleted in dystrophic muscle. *Cell* 2002;**110**:339–48. https://doi.org/10.1016/S0092-8674(02)00826-7
15. Zeng W, Chen YY, Newkirk DA *et al.* Genetic and epigenetic characteristics of FSHD-associated 4q and 10q D4Z4 that are distinct from non-4q/10q D4Z4 homologs. *Hum Mutat* 2014;**35**:998–1010. https://doi.org/10.1002/humu.22593
16. Salsi V, Losi F, Salani M *et al.* Posttranscriptional RNA stabilization of telomeric RNAs FRG2, DBE-T, D4Z4 at human 4q35 in response to genotoxic stress and D4Z4 macrosatellite repeat length. *Clin Epigenet* 2025;**17**:73. https://doi.org/10.1186/s13148-025-01881-5
17. Rijkers T, Deidda G, Van Koningsbruggen S *et al.* FRG2, an FSHD candidate gene, is transcriptionally upregulated in differentiating primary myoblast cultures of FSHD patients. *J Med Genet* 2004;**41**:826–36. https://doi.org/10.1136/jmg.2004.019364
18. Bodega B, Ramirez GDC, Grasser F *et al.* Remodeling of the chromatin structure of the facioscapulohumeral muscular dystrophy (FSHD) locus and upregulation of FSHD-related gene 1 (FRG1) expression during human myogenic differentiation. *BMC Biol* 2009;**7**:41. https://doi.org/10.1186/1741-7007-7-41
19. Gaillard MC, Broucqsault N, Morere J *et al.* Analysis of the 4q35 chromatin organization reveals distinct long-range interactions in patients affected with Facio-Scapulo-Humeral Dystrophy. *Sci Rep* 2019;**9**:12552. https://doi.org/10.1038/s41598-019-46861-x
20. Bettio C, Salsi V, Orsini M *et al.* The Italian National Registry for FSHD: an enhanced data integration and an analytics framework towards Smart Health Care and Precision Medicine for a rare disease. *Orphanet J Rare Dis* 2021;**16**:470. https://doi.org/10.1186/s13023-021-02100-z
21. Soriano-Arroquia A, Clegg PD, Molloy AP *et al.* Preparation and culture of myogenic precursor cells/primary myoblasts from skeletal muscle of adult and aged humans. *J Vis Exp* 2017;**86**:e55047. https://doi.org/10.3791/55047
22. Van der Auwera GA, Carneiro MO, Hartl C *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma* 2013;**43**:11.10.1–11.10.33.
23. Ewels P, Magnusson M, Lundin S *et al.* MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinforma Oxf Engl* 2016;**32**:3047–8. https://doi.org/10.1093/bioinformatics/btw354
24. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;**9**:357–9. https://doi.org/10.1038/nmeth.1923
25. Danecek P, Bonfield JK, Liddle J *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* 2021;**10**:giab008. https://doi.org/10.1093/gigascience/giab008
26. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;**26**:841–2. https://doi.org/10.1093/bioinformatics/btq033
27. Labun K, Montague TG, Krause M *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res* 2019;**47**:W171–4. https://doi.org/10.1093/nar/gkz365
28. Mondal T, Subhash S, Kanduri C. Chromatin RNA immunoprecipitation (ChRIP). *Methods Mol Biol Clifton NJ* 2018;**1689**:65–76. https://doi.org/10.1007/978-1-4939-7380-4_6
29. Santini F, Deligianni X, Paoletti M *et al.* Fast open-source toolkit for water T2 mapping in the presence of fat from multi-echo spin-echo acquisitions for muscle MRI. *Front Neurol* 2021;**12**:630387. https://doi.org/10.3389/fneur.2021.630387

30. Wingett SW, Andrews S. FastQ screen: a tool for multi-genome mapping and quality control. *F1000Res* 2018;**7**:1338. https://doi.org/10.12688/f1000research.15931.1
31. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;**26**:841–42. https://doi.org/10.1093/bioinformatics/btq033
32. Zhang Y, Liu T, Meyer CA *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;**9**:841–2. https://doi.org/10.1186/gb-2008-9-9-r137
33. Heinz S, Benner C, Spann N *et al.* Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* 2010;**38**:576–89. https://doi.org/10.1016/j.molcel.2010.05.004
34. Peng T, Hou Y, Meng H *et al.* Mapping nucleolus-associated chromatin interactions using nucleolus hi-C reveals pattern of heterochromatin interactions. Nat Commun 2023;**14**:313 https://doi.org/10.1038/s41467-023-36021-1
35. Vertii A, Ou J, Yu J *et al.* Two contrasting classes of nucleolus-associated domains in mouse fibroblast heterochromatin. *Genome Res* 2019;**29**:1235–49. https://doi.org/10.1101/gr.247072.118
36. Fioriniello S, Csukonyi E, Marano D *et al.* MeCP2 and major satellite forward RNA cooperate for pericentric heterochromatin organization. *Stem Cell Rep* 2020;**15**:1317–32. https://doi.org/10.1016/j.stemcr.2020.11.006
37. Jaafar M, Contreras J, Dominique C *et al.* Association of snR190 snoRNA chaperone with early pre-60S particles is regulated by the RNA helicase Dbp7 in yeast. *Nat Commun* 2021;**12**:6153. https://doi.org/10.1038/s41467-021-26207-w
38. Shumate A, Salzberg SL. Liftoff: accurate mapping of gene annotations. *Bioinformatics* 2021;**37**:1639–43. https://doi.org/10.1093/bioinformatics/btaa1016
39. Kiniry SJ, Michel AM, Baranov PV. The GWIPS-viz browser. *Curr Protoc Bioinformatics* 2018;**62**:e50. https://doi.org/10.1002/cpbi.50
40. Kollipara L, Buchkremer S, Weis J *et al.* Proteome profiling and ultrastructural characterization of the Human RCMH cell line: myoblastic properties and suitability for myopathological studies. *J Proteome Res* 2016;**15**:945–55. https://doi.org/10.1021/acs.jproteome.5b00972
41. Nishimura Y, Bittel AJ, Stead CA *et al.* Facioscapulohumeral muscular dystrophy is associated with altered myoblast proteome dynamics. *Mol Cell Proteomics* 2023;**22**:100605. https://doi.org/10.1016/j.mcpro.2023.100605
42. Uhlén M, Fagerberg L, Hallström BM *et al.* Proteomics. Tissue-based map of the human proteome. *Science* 2015;**347**:1260419. https://doi.org/10.1126/science.1260419
43. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res* 2021;**49**:D480–9. https://doi.org/10.1093/nar/gkaa1100
44. Millán-Zambrano G, Burton A, Bannister AJ *et al.* Histone post-translational modifications—cause and consequence of genome function. *Nat Rev Genet* 2022;**23**:563–80. https://doi.org/10.1038/s41576-022-00468-7
45. Chu HP, Cifuentes-Rojas C, Kesner B *et al.* TERRA RNA antagonizes ATRX and protects telomeres. *Cell* 2017;**170**:86–101.e16. https://doi.org/10.1016/j.cell.2017.06.017
46. Quinn JJ, Ilik IA, Qu K *et al.* Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. *Nat Biotechnol* 2014;**32**:933–40. https://doi.org/10.1038/nbt.2943
47. Altemose N, Logsdon GA, Bzikadze AV *et al.* Complete genomic and epigenetic maps of human centromeres. *Science* 2022;**376**:eabl4178. https://doi.org/10.1126/science.abl4178
48. Hori Y, Shimamoto A, Kobayashi T. The human ribosomal DNA array is composed of highly homogenized tandem clusters. *Genome Res* 2021;**31**:1971–82. https://doi.org/10.1101/gr.275838.121
49. Kearse M, Moir R, Wilson A *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012;**28**:1647–9. https://doi.org/10.1093/bioinformatics/bts199
50. Mediani L, Antoniani F, Galli V *et al.* Hsp90-mediated regulation of DYRK3 couples stress granule disassembly and growth via mTORC1 signaling. *EMBO Rep* 2021;**22**:e51740. https://doi.org/10.15252/embr.202051740
51. Pallan PS, Egli M. Insights into RNA/DNA hybrid recognition and processing by RNase H from the crystal structure of a non-specific enzyme–dsDNA complex. *Cell Cycle* 2008;**7**:2562–9. https://doi.org/10.4161/cc.7.16.6461
52. Boulon S, Westman BJ, Hutten S *et al.* The nucleolus under stress. *Mol Cell* 2010;**40**:216–27. https://doi.org/10.1016/j.molcel.2010.09.024
53. Sleeman JE, Trinkle-Mulcahy L. Nuclear bodies: new insights into assembly/dynamics and disease relevance. *Curr Opin Cell Biol* 2014;**28**:76–83. https://doi.org/10.1016/j.ceb.2014.03.004
54. Caudron-Herger M, Pankert T, Seiler J *et al.* Alu element-containing RNAs maintain nucleolar structure and function. *EMBO J* 2015;**34**:2758–74. https://doi.org/10.15252/embj.201591458
55. Lubelsky Y, Ulitsky I. Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* 2018;**555**:107–11. https://doi.org/10.1038/nature25757
56. Tsue AF, Kania EE, Lei DQ *et al.* Multiomic characterization of RNA microenvironments by oligonucleotide-mediated proximity-interactome mapping. *Nat Methods* 2024;**21**:2058–71. https://doi.org/10.1038/s41592-024-02457-6
57. Bizhanova A, Kaufman PD. Close to the edge: heterochromatin at the nucleolar and nuclear peripheries. *Biochim Biophys Acta* 2021;**1864**:194666. https://doi.org/10.1016/j.bbagrm.2020.194666
58. Bersaglieri C, Kresoja-Rakic J, Gupta S *et al.* Genome-wide maps of nucleolus interactions reveal distinct layers of repressive chromatin domains. *Nat Commun* 2022;**13**:1483. https://doi.org/10.1038/s41467-022-29146-2
59. Németh A, Längst G. Genome organization in and around the nucleolus. *Trends Genet* 2011;**27**:149–56. https://doi.org/10.1016/j.tig.2011.01.002
60. Smirnov E, Hornáček M, Kováčik L *et al.* Reproduction of the FC/DFC units in nucleoli. *Nucleus* 2016;**7**:203–15. https://doi.org/10.1080/19491034.2016.1157674
61. Santoro R, Schmitz KM, Sandoval J *et al.* Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. *EMBO Rep* 2010;**11**:52–8. https://doi.org/10.1038/embor.2009.254
62. Yan Q, Zhu C, Guang S *et al.* The functions of non-coding RNAs in rRNA regulation. *Front Genet* 2019;**10**:290. https://doi.org/10.3389/fgene.2019.00290
63. Pirogov SA, Gvozdev VA, Klenov MS. Long noncoding RNAs and stress response in the nucleolus. *Cells* 2019;**8**:668. https://doi.org/10.3390/cells8070668
64. Jao CY, Salic A. Exploring RNA transcription and turnover *in vivo* by using click chemistry. *Proc Natl Acad Sci USA* 2008;**105**:15779–84. https://doi.org/10.1073/pnas.0808480105
65. Ogawa LM, Buhagiar AF, Abriola L *et al.* Increased numbers of nucleoli in a genome-wide RNAi screen reveal proteins that link the cell cycle to RNA polymerase I transcription. *MBoC* 2021;**32**:956–73. https://doi.org/10.1091/mbc.E20-10-0670
66. Sherman BT, Hao M, Qiu J *et al.* DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res* 2022;**50**:W216–21. https://doi.org/10.1093/nar/gkac194
67. Xue S, Barna M. Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat Rev Mol Cell Biol* 2012;**13**:355–69. https://doi.org/10.1038/nrm3359

68. Farley-Barnes KI, Ogawa LM, Baserga SJ. Ribosomopathies: old concepts, new controversies. *Trends Genet* 2019;**35**:754–67. https://doi.org/10.1016/j.tig.2019.07.004
69. Figueiredo VC, McCarthy JJ. Regulation of ribosome biogenesis in skeletal muscle hypertrophy. *Physiology* 2019;**34**:30–42. https://doi.org/10.1152/physiol.00034.2018
70. Millward DJ, Garlick PJ, James WP *et al.* Relationship between protein synthesis and RNA content in skeletal muscle. *Nature* 1973;**241**:204–5. https://doi.org/10.1038/241204a0
71. Bou Saada Y, Zakharova V, Chernyak B *et al.* Control of DNA integrity in skeletal muscle under physiological and pathological conditions. *Cell Mol Life Sci* 2017;**74**:3439–49. https://doi.org/10.1007/s00018-017-2530-0
72. Bou Saada Y, Dib C, Dmitriev P *et al.* Facioscapulohumeral dystrophy myoblasts efficiently repair moderate levels of oxidative DNA damage. *Histochem Cell Biol* 2016;**145**:475–83. https://doi.org/10.1007/s00418-016-1410-2
73. Correll CC, Bartek J, Dundr M. The nucleolus: a multiphase condensate balancing ribosome synthesis and translational capacity in health, aging and ribosomopathies. *Cells* 2019;**8**:869. https://doi.org/10.3390/cells8080869
74. Politz JC, Yarovoi S, Kilroy SM *et al.* Signal recognition particle components in the nucleolus. *Proc Natl Acad Sci USA* 2000;**97**:55–60. https://doi.org/10.1073/pnas.97.1.55
75. Hacisuleyman E, Goff LA, Trapnell C *et al.* Topological organization of multi-chromosomal regions by Firre. *Nat Struct Mol Biol* 2014;**21**:198–206. https://doi.org/10.1038/nsmb.2764
76. Mattick JS, Amaral PP, Carninci P *et al.* Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat Rev Mol Cell Biol* 2023;**24**:430–47. https://doi.org/10.1038/s41580-022-00566-8
77. West JA, Davis CP, Sunwoo H *et al.* The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol Cell* 2014;**55**:791–802. https://doi.org/10.1016/j.molcel.2014.07.012

Received: April 14, 2025. **Revised:** June 15, 2025. **Editorial Decision:** June 16, 2025. **Accepted:** June 19, 2025