BRAIN
ORIGINAL ARTICLE

Pleiotropic effects of *MORC2* derive from its epigenetic signature

Fatemeh Peymani,[1,2,†] Tomohiro Ebihara,[2,†] Dmitrii Smirnov,[1,2,†] Robert Kopajtich,[1,2] Masahiro Ando,[3] Enrico Bertini,[4] Rosalba Carrozzo,[5] Daria Diodato,[4] Felix Distelmaier,[6] Fang Fang,[7] Daniele Ghezzi,[8] Maja Hempel,[9,10] Katarzyna Iwanicka-Pronicka,[11] Thomas Klopstock,[12,13,14] Sarah L. Stenton,[1,2] Costanza Lamperti,[8] Zhimei Liu,[7] Aysylu Murtazina,[15] Yuji Okamoto,[3,16] Yasushi Okazaki,[17] Dorota Piekutowska-Abramczuk,[11] Agnés Rötig,[18] Oxana Ryzhkova,[15] Christian Schlein,[19] Olga Shagina,[15] Hiroshi Takashima,[3] Polina Tsygankova,[15] Michael Zech,[1,2] Thomas Meitinger,[1] Masaru Shimura,[2,20] Kei Murayama[17,20,‡] and Holger Prokisch[1,2,21,‡]

[†,‡]These authors contributed equally to this work.

Heterozygous missense mutations in *MORC2* have been implicated in various clinical entities, ranging from early-onset neurodevelopmental disorders to late-onset neuropathies. The mechanism underlying the phenotypic heterogeneity and pleiotropic effects of *MORC2* has remained elusive.

Here, we analysed blood and fibroblast DNA methylation, transcriptomes, proteomes and phenotypes of 53 *MORC2* patients.

We identified a *MORC2*-specific DNA methylation episignature that is universal across all *MORC2*-associated phenotypes and conserved across different tissues. The *MORC2* episignature consists mainly of DNA hypermethylation in promoter regions, leading to transcriptional repression of target genes resulting in a *MORC2*-specific RNA signature. Concomitant downregulation of three disease-associated genes—*ERCC8*, *NDUFAF2* and *FKTN*—at different levels mirrors the variable biochemical defects and clinical manifestations observed in *MORC2* patients. Silencing of *NDUFAF2* accounts for the Leigh syndrome manifestation, whereas dysmorphic features are due to the repression of *ERCC8*.

Overall, we showed that pathogenic *MORC2* variants cause specific episignature, whereby methylation level variability and its repression impact on target genes explains the pleiotropy and predicts phenotypic heterogeneity in *MORC2*-related disorders. We predict that epigenetic variation may underlie pleiotropy in other Mendelian disorders.

1 Institute of Human Genetics, School of Medicine and Health, Technical University of Munich, Munich 81675, Germany
2 Institute of Neurogenomics, Computational Health Center, Helmholtz Munich, Neuherberg 85764, Germany
3 Department of Neurology and Geriatrics, Kagoshima University Graduate School of Medical and Dental Sciences, Kagoshima 890-8544, Japan
4 Neuromuscular Disorders Research Unit, Bambino Gesù Children's Hospital IRCCS, Rome 00165, Italy
5 Laboratory of Medical Genetics, Translational Cytogenomics Research Unit, Bambino Gesù Children's Hospital, IRCCS, Rome 00165, Italy
6 Department of General Pediatrics, Neonatology and Pediatric Cardiology, University Children's Hospital, Heinrich-Heine-University Düsseldorf, Düsseldorf 40225, Germany

Received January 11, 2025. Revised March 21, 2025. Accepted April 08, 2025. Advance access publication April 30, 2025
© The Author(s) 2025. Published by Oxford University Press on behalf of the Guarantors of Brain.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

7 Department of Neurology, Beijing Children's Hospital, Capital Medical University, Beijing 100005, China
8 Unit of Medical Genetics and Neurogenetics, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milano 20133, Italy
9 Institute of Human Genetics, Heidelberg University, Heidelberg 69120, Germany
10 Institute of Human Genetics, University Hospital Heidelberg, Heidelberg 69120, Germany
11 Department of Medical Genetics, The Children's Memorial Health Institute, Warsaw 04-730, Poland
12 Department of Neurology, Friedrich-Baur-Institute, LMU University Hospital, Ludwig-Maximilians-Universität (LMU) München, Munich 81377, Germany
13 German Center for Neurodegenerative Diseases (DZNE), Munich 81377, Germany
14 Munich Cluster for Systems Neurology (SyNergy), Munich 81377, Germany
15 Research Centre for Medical Genetics, Moscow 115522, Russia
16 Department of Physical Therapy, Kagoshima University of School of Health Sciences, Kagoshima 890-8544, Japan
17 Diagnostics and Therapeutics of Intractable Diseases, Intractable Disease Research Center, Graduate School of Medicine, Juntendo University, Tokyo 113-8421, Japan
18 UMR1163, Université Paris Descartes, Sorbonne Paris Cité, Institut Imagine, Paris 75015, France
19 Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg 20246, Germany
20 Department of Metabolism, Chiba Children's Hospital, Chiba City 266-0007, Japan
21 German Center for Child and Adolescent Health (DZKJ), partner site Munich, Munich 81377, Germany

Correspondence to: Holger Prokisch
Institute of Neurogenomics, Helmholtz Munich
Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)
Ingolstädter Landstraße 1, Neuherberg 85764, Germany
E-mail: holger.prokisch@helmholtz-munich.de

Keywords: pleiotropy; MORC2; episignature; multi-omics; Leigh syndrome; CMT

Introduction

Pleiotropy, the simultaneous presentation of different phenotypes due to variants in a single gene, and phenotypic heterogeneity are common, yet often unexplained, features of monogenic disorders.[1] *MORC2* (Microrchidia CW-type zinc finger protein 2)-related disorders are a prime example of this phenomenon. Heterozygous missense mutations in *MORC2* have been identified in association with a number of disorders, including Charcot–Marie–Tooth disease type 2Z (CMT2Z),[2-4] spinal muscular atrophy (SMA),[5] Leigh syndrome, mitochondrial disease[6,7] and Cockayne syndrome.[8] In another study, patients with *MORC2* pathogenic variants have been grouped under the umbrella term DIGFAN, which stands for developmental delay, impaired growth, dysmorphic facies, axonal neuropathy[9] (Fig. 1 and Supplementary Tables 1 and 2).

At the molecular level, *MORC2* encodes a DNA-dependent ATPase chromatin remodeller that functions both dependently and independently of the human silencing hub (HUSH) complex.[10,11] Within the HUSH complex, MORC2 contributes to silencing the expression of the retroelements, thereby maintaining genome integrity. The HUSH complex regulates the deposition of the repressive epigenetic mark H3K9me3 (trimethylation of histone H3 at lysine 9) and recruits MORC2 to mediate chromatin compaction.[10,12,13] Independent of the HUSH complex, MORC2 is involved in gene silencing and facilitating DNA repair by modulating chromatin structure- chromatin compaction and decompaction.[13-15] Disease-causing variants in the ATPase module of *MORC2* result in its increased activity and hyper-repression of target genes.[9,10,12] However, the molecular mechanism linking increased MORC2 activity to the wide range of associated phenotypes, including mitochondrial dysfunction, remains poorly understood. Furthermore, the gain-of-function pathomechanism, the highly variable phenotypes, and the lack of functional assays pose challenges for interpreting the pathogenicity of novel variants of uncertain significance (VUS) in the *MORC2* gene.

To investigate the pathomechanism underlying *MORC2* gain-of-function pathogenic variants, we performed a multi-omics analysis of blood and fibroblast samples from *MORC2* patients. We discovered a *MORC2*-specific episignature that is conserved across tissues and consistent across all related phenotypes. This episignature not only serves as a functional assay for *MORC2* VUS but also identifies hitherto unknown MORC2 targets. Hypermethylation of promoter regions in target genes resulted in gene repression at levels similar to those seen in Mendelian disorders. The concurrent repression of disease-associated genes—*ERCC8*, *NDUFAF2*, *FKTN*—underpins the observed pleiotropy in patients. Among the targets, the methylation level of the bidirectional promoter of *ERCC8-NDUFAF2* exhibits the strongest correlation with disease severity, being higher in Leigh syndrome than in CMT. Overall, our study illustrates the link between epigenetic modifications and pleiotropy in Mendelian disorders.

Materials and methods

Study cohort

A cohort of 53 *MORC2* genetically diagnosed patients has been collected through both national sources and international collaborations. The cohort includes 39 newly identified cases and 14 previously reported ones.[3,6] Patient phenotypes were extracted and coded as Human Phenotype Ontology (HPO) terms (Supplementary Table 2). Written informed consent was obtained from all individuals or their guardians prior to evaluation and

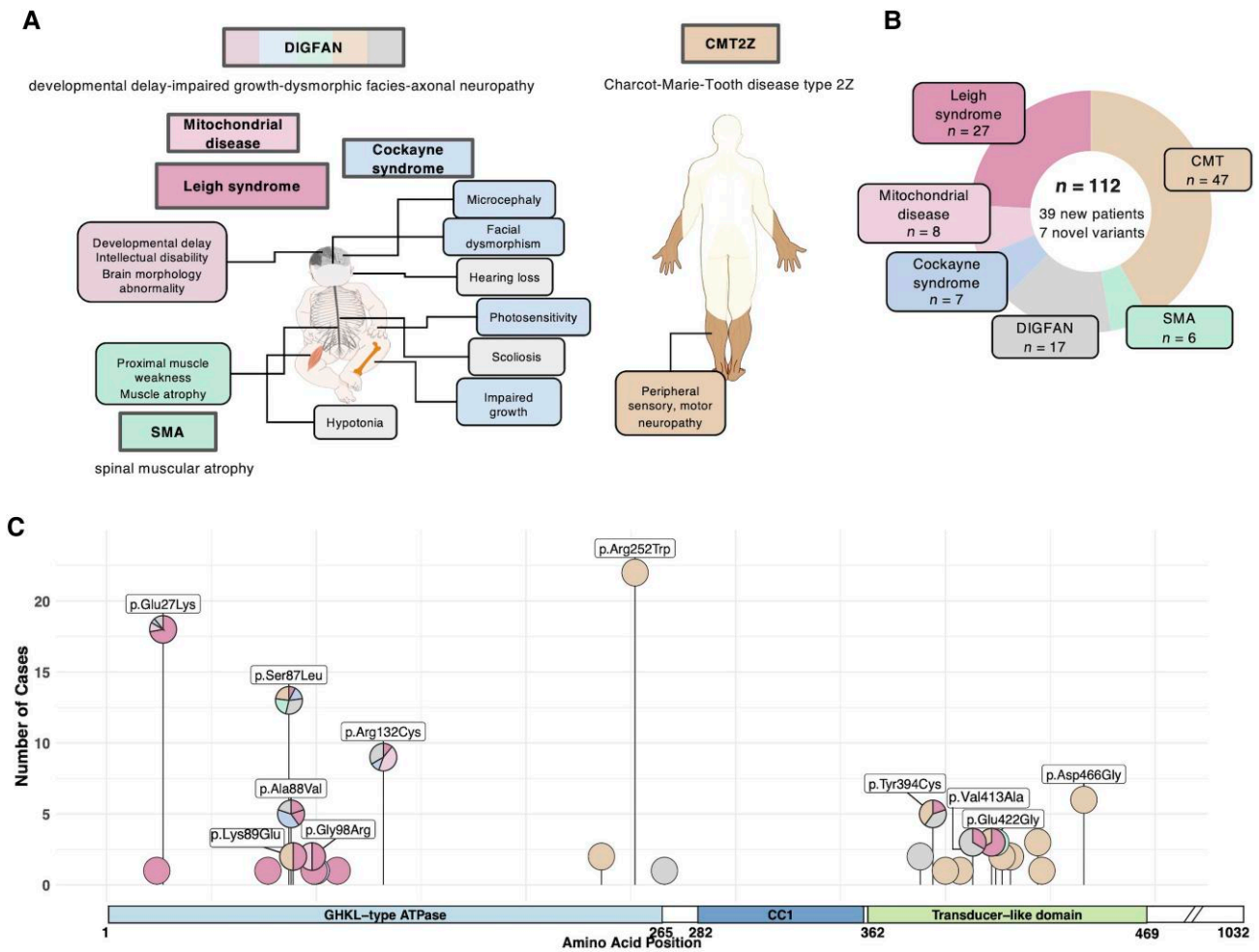

Figure 1 Phenotypic heterogeneity in *MORC2*-realted disorders. (**A**) Main phenotypic abnormalities observed in *MORC2*-related disorders are shown. (**B**) Donut chart showing the distribution of different disorders associated with *MORC2* pathogenic variants in a total of 112 patients, including both previously reported patients ($n = 73$) and new patients reported in this paper ($n = 39$). (**C**) Lollipop plot showing the distribution of pathogenic variants in *MORC2* and their associated disorders. Variants reported more than five times or associated with more than one disorder are labelled in the plot. The data include 112 patients described in **B**.

testing, in accordance with the Declaration of Helsinki, and approved by the ethical committees of the participating centres.

The classification of disease types for unreported cases was performed as follows. Initially, cases with developmental delay and bilateral basal ganglia or brainstem lesions on head MRI were classified as Leigh syndrome. Subsequently, cases with multisystem symptoms involving more than three organs and with defects in mitochondrial oxidative phosphorylation enzymes or elevated lactate were classified as mitochondrial disease.[16] Among the remaining cases, those with growth impairment (short stature, microcephaly) and characteristic facial features were classified as Cockayne syndrome.[17] Cases not classified into the above three categories were characterized by predominant peripheral neuropathy and were classified as CMT (Supplementary Tables 1 and 2).

A clinical-radiological score[18] was calculated to evaluate the clinical similarity of the cases to Cockayne syndrome. Clinical symptoms were aggregated for all patients; however, head imaging was omitted in cases where CNS symptoms were not the primary manifestation (Supplementary Table 2).

DNA methylation profiling was performed on both blood and fibroblasts in 45 patients, depending on material availability. One patient had both blood and fibroblasts, and one patient had two samples from different time points, resulting in 47 DNA methylation data points from *MORC2* patients ($n = 35$ blood and $n = 12$ fibroblast DNA methylation).

RNA sequencing (RNA-seq) and proteomics were performed on fibroblasts from 12 and 13 *MORC2* patients, respectively.

DNA methylation analysis pipeline

Processing of DNA methylation

Genomic DNA was extracted from blood and fibroblast samples. DNA methylation analysis was performed on bisulfite-converted DNA from both blood and fibroblast samples using the Infinium Methylation EPIC v1.0 and v2.0 BeadChip arrays, in accordance with the manufacturer's protocols at Helmholtz Munich Core Facility. DNA methylation data preprocessing was performed using the minfi package[19] in R version 4.3.1. Initial quality control involved filtering out samples based on several criteria: average detection *P*-value less than 0.05, discordance in recorded gender and predicted gender based on the methylation profile, methylation intensity deviating more than three standard deviations from the mean, and non-bimodal distribution in density plots. Quantile

normalization was applied on the CpG sites common between EPIC v1.0 and EPIC v2.0 arrays. In addition, the probes with a detection *P*-value greater than 0.01, probes overlapping with single nucleotide polymorphisms (SNPs), cross-reactive probes and probes located on the X and Y chromosomes were excluded from the analysis. Following these filtering criteria, 631 142 CpG sites in 512 samples ($n = 481$ blood and $n = 31$ fibroblast) remained for downstream analysis. Methylation beta values and M-values were calculated for all the CpG sites. The proportions of white blood cell types were estimated using the Houseman method.[20] Age was also inferred from methylation data using the Horvath method implemented in the methylclock package.[21]

Discovery of *MORC2*-specific episignature

An epigenome-wide association study (EWAS) was performed on blood DNA methylation. A cohort of 13 *MORC2* patients and 58 healthy controls were selected, and 80:20 per cent train-to-test splitting was performed. The EWAS was conducted on the training cohort consisting of 10 *MORC2* patients and 46 healthy controls. Linear regression on the methylation values (M-values), adjusting for disease status (*MORC2* versus control), age, gender, white blood cell type composition and the version of the EPIC array (EPIC v1.0 and v2.0) was performed using the limma package. CpG sites showing significant differences at Bonferroni-adjusted *P*-values < 0.05 were considered differentially methylated positions (DMPs).

Given the extensive number of CpG sites showing significant differences at Bonferroni-adjusted *P*-values < 0.05 ($n = 760$ CpG), we employed several criteria to refine the selection of CpG sites for inclusion in our diagnostic classifier. Following the approach by Levy *et al.*,[22] we computed a methylation score for each site by multiplying the negative log-transformed false discovery rate (FDR) by the absolute value of the log-fold change (logFC). The top 1000 probes, ranked by this score, were initially selected. The feature importance for the selected CpGs was calculated using the filterVarImp function from the Caret package. From this analysis, 220 CpG sites exhibiting an area under the curve (AUC) greater than 0.99 were chosen for training the classifier.

Unsupervised visualization was performed by calculating the Euclidean distances between samples, followed by multidimensional scaling on the M-values of 220 CpG sites to examine the data structure and clustering of *MORC2* patients and controls. Subsequently, a support vector machine (SVM) classifier was trained using the e1071 package based on the M-values of the same 220 CpG sites. The training was performed on the samples used for the EWAS analysis. A 10-fold cross-validation with a linear kernel was applied. The probability scores generated by the classifier were used to classify *MORC2* patients, with a threshold set at a score greater than 0.5. The classifier was tested on a blood sample cohort comprising three *MORC2* patients, 12 healthy controls and 368 samples with other genetic disorders, and on a fibroblast sample cohort consisting of six *MORC2* patients and 18 healthy or diseased controls. In addition, the phenotype specificity of the classifier was evaluated in a validation cohort of 49 DNA methylation samples collected after the discovery of episignature, including 28 samples from *MORC2* patients ($n = 22$ blood; $n = 6$ fibroblast DNA methylation), 22 of whom were clinically diagnosed with CMT.

Discovery of phenotype-specific episignature

A total of 43 samples from *MORC2* patients exhibiting the *MORC2*-specific episignature were clinically diagnosed with Leigh syndrome ($n = 13$), mitochondrial disease ($n = 6$) or CMT ($n = 24$). For the EWAS, data from the two most distinct phenotypes, typical CMT and Leigh syndrome, were selected. DNA methylation data for the majority of Leigh syndrome patients ($n = 10$) were derived from fibroblasts, whereas for CMT patients, the data were obtained mainly from blood samples ($n = 23$). To mitigate the confounding effect of tissue in EWAS, we combined the DNA methylation data of CMT patients with fibroblast DNA methylation of control cell lines. EWAS was performed comparing 10 Leigh syndrome patients against 32 samples comprising 16 CMT patients and 16 controls. Gender, age, tissue (blood or fibroblast) and the version of the EPIC array were included as covariates. CpG sites at Bonferroni-adjusted *P*-values < 0.05 ($n = 80$ CpG sites) were considered significant.

Similar to the discovery of the episignature, the multidimensional scaling plot was utilized to visualize the data structure using the 80 CpG sites identified in the previous step. Subsequently, the SVM classifier was trained on the set of samples used for EWAS, employing 5-fold cross-validation to classify *MORC2* patients with Leigh syndrome/mitochondrial disease versus those with CMT. The classifier was tested on additional samples, including nine patients with Leigh syndrome/mitochondrial disease and 11 patients with CMT. The classifier was further tested on DNA methylation of a cohort of 449 samples ($n = 3$ fibroblast and $n = 446$ blood).

Annotation of DNA methylation CpG sites

CpG site annotation was performed using the getAnnotation function from the IlluminaHumanMethylationEPICanno.ilm10b4.hg19 package in R. The spatial localization of CpG sites relative to genes was determined using the 'UCSC_RefGene_Group' column, which includes TSS1500 (200–1500 bases upstream of the TSS), TSS200 (0–200 bases upstream of the TSS), 5′UTR, 1st exon, ExonBnd (exon boundaries), Body (gene body) and 3′UTR. We defined the gene promoters as CpGs located in TSS1500 or TSS200. Since each CpG site can have multiple annotations based on its location relative to different transcripts of the gene, we considered the most frequent annotation for downstream analysis to ensure one annotation per CpG. We also repeated the analysis using all possible annotations for each CpG, but this did not show significant differences compared to the first analysis.

Gene annotations were obtained from the InfiniumAnnotation database (https://zwdzwd.github.io/InfiniumAnnotation/EPIC_hm450_hg19.html), which maps the genes to the CpG sites spanning from 1.5 kb upstream of the transcription start site to the transcription termination site, using GENECODE release 26. For CpG sites with multiple corresponding genes, all genes were considered for the analysis.

HUSH target sites and MORC2 binding sites were obtained from Seczynska *et al.*[23] and Tchasovnikarova *et al.*,[10] respectively. The CpGs located within the identified loci were marked as CpGs in the HUSH target site or MORC2 binding site.

Sample-based aggregated methylation value

The calculation of the methylation value for each sample based on a set of CpG sites was performed as follows. First, for each CpG, we calculated the methylation Z-score by subtracting the mean M-value across all DNA methylation samples from the M-value of that CpG in each sample and dividing by the standard deviation. Then, for each sample, we calculated the mean of the methylation Z-scores for the specific set of CpG sites (such as the 220 CpGs in the episignature or the 3019 CpGs mapped to HUSH targets).

Transcriptome analysis pipeline

RNA sequencing and data processing

RNA-seq was performed according to Yépez *et al.*[24] and Kremer *et al.*[25] Briefly, RNA was isolated from patient-derived skin fibroblasts, and library preparation was conducted using either the TruSeq Stranded mRNA or Illumina Stranded mRNA Ligation protocol. The RNA libraries were sequenced as 100 bp paired-end runs on Illumina HiSeq2500 or HiSeq4000 platforms. RNA-seq reads were demultiplexed and aligned to the hg19 genome assembly (UCSC Genome Browser build) using STAR version 2.7.0a. The DROP pipeline (version 1.3.4) was employed for read counting and quality control, using the GRCh37 primary assembly, release 29, of the GENCODE project as the reference genome. RNA-SEQC was utilized for quality control, and the summarizeOverlaps function from the GenomicAlignments R package was used to count reads. Genes with a 95th percentile FPKM below one were filtered out as not sufficiently expressed.

Discovery of *MORC2*-specific RNA signature

Sixty-four samples, comprising 12 *MORC2* samples, 18 healthy controls and 34 diseased controls, were selected for differential expression analysis. An 80:20 train-test split was performed on this cohort, resulting in a training set of 10 *MORC2* patients and 42 controls (13 healthy and 29 diseased), which was used for the identification of differentially expressed genes (DEGs). Principal component analysis (PCA) was conducted and revealed the impact of library preparation protocols.

DESeq2 version 1.34.0 was employed for the normalization and detection of DEGs. Briefly, DESeq2 estimates size factors to normalize for differences in sequencing depth across samples. A design formula incorporating disease status (*MORC2* versus others), sex and library protocol was used to model the count data using negative binomial generalized linear models (GLMs). The Wald test was applied to assess significant differences in gene expression between *MORC2* patients and controls while accounting for the effects of covariates. A threshold of Benjamini–Hochberg-adjusted *P*-value < 0.05 was used to determine statistically significant DEGs ($n = 106$ DEGs). Unsupervised clustering was performed on 106 DEGs, in the same way as performed for episignature, to examine the clusterization of *MORC2* patients and control samples.

A SVM classifier was trained on 106 DEGs using a methodology analogous to that employed for DNA methylation. The classifier was trained on normalized counts, where normalization was performed using DESeq2 by estimating size factors with the estimateSizeFactors function and applying variance stabilizing transformation (VST) with the vst function to stabilize the variance across the data, followed by scaling. The classifier was trained on the sample set used for DEGs detection with 10-fold cross-validation. The SVM probability scores were utilized similarly to the DNA methylation analysis, with a threshold of 0.5 set to classify *MORC2* patients. The classifier was tested on the test cohort of two *MORC2* patients and 10 healthy or diseased controls, as well as an additional set of 168 genetically undiagnosed samples.

Detection of expression outliers

Expression outliers were identified using the OUTRIDER package on the transcriptome of fibroblast cell lines of a cohort of 877 individuals consisting of 12 *MORC2* patients. OUTRIDER employs a denoising autoencoder to control for latent effects and provides an RNA z-score for each gene and sample, allowing for the detection of expression outliers. Unsupervised clustering has also been performed on the denoising autoencoder-derived normalized counts of 106 DEGs.

Proteomics analysis pipeline

Proteomics has been performed on fibroblast cell lines of 660 samples consisting of 13 *MORC2* patients, as described by Kopajtich *et al.*[26] In summary, the protocol included reduction, alkylation, tryptic digestion using Trypsin Gold (Promega), and TMT labelling and liquid chromatography-mass spectrometry measurements. In total, 9040 proteins have been detected. The proteins with missing values in more than 20% of the samples were filtered out, resulting in 6749 proteins. PROTRIDER[26] has been utilized for normalization and protein outlier detection. Similar to OUTRIDER, PROTRIDER employed a denoising autoencoder to mitigate the influence of latent confounders. Proteins with a Z-score < -2 were classified as protein outliers. Recurrent protein outliers were defined as those identified as outliers in at least two *MORC2* patients. The significance of recurrent protein outliers in *MORC2* patients was assessed using a one-sided Fisher exact test.

The cryo-electron microscopy structure in the active state (PDB ID: 6G2J) was used in PyMol V.2.5.4 (Schrödinger) to visualize the protein abundances in respiratory chain complex I.

Oxygen consumption rate measurement

The oxygen consumption rate was examined using skin fibroblasts according to the method previously described.[27] For each disease type and control (NHDF), 22 to 24 wells were used, and the measurement was conducted twice.

Results

Discovery of tissue-independent *MORC2* episignature

Pathogenic variants in genes involved in the epigenetic machinery have been shown to result in gene-specific methylation changes called episignatures. These episignatures serve as highly sensitive and specific biomarkers that can be used as functional assays to assess the functional impact of VUS in a growing number of Mendelian disorders.[22] Despite MORC2's recognized role in gene regulation, its genome-wide impact on DNA methylation remains unexplored. To assess the impact of pathogenic variants in *MORC2* on DNA methylation, we conducted an EWAS using 10 *MORC2* patients and 46 healthy controls (Supplementary Table 3). The analysis identified 760 DMPs when corrected for multiple testing (Bonferroni *P*-adjusted value < 0.05); 93% ($n = 707$) were hypermethylated in *MORC2* patients (Fig. 2A and Supplementary Table 4). The 760 DMPs were mapped to 480 genes (differentially methylated genes). From these 760 DMPs, we selected the 220 most differentiating CpG sites (see the 'Materials and methods' section for details) for the unsupervised analysis. This resulted in a clear separation between *MORC2* patients and controls (Supplementary Fig. 1A). To validate the specificity of the identified episignature, we trained a SVM classifier on blood DNA methylation data using 220 CpGs on the sample set used in the EWAS. This classifier was then tested on an independent cohort of three *MORC2* patients, 12 healthy controls, 281 patients with mitochondrial disease and 86 patients with other neurodevelopmental disorders. The

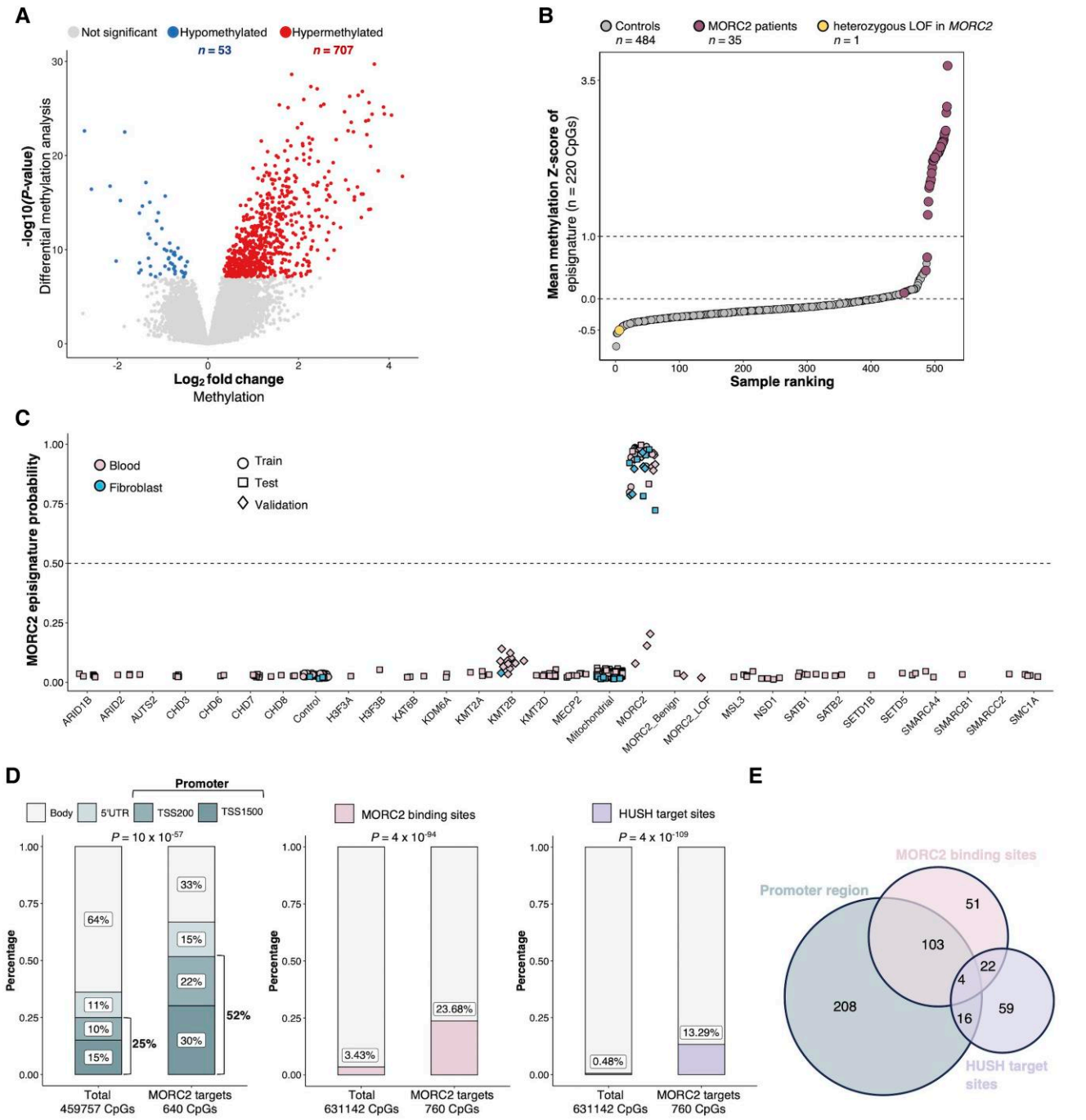

Figure 2 Tissue-independent *MORC2*-specific episignatures are mainly localized in gene promoters. (**A**) Volcano plot of differential methylation analysis comparing *MORC2* patients ($n = 10$) to healthy controls ($n = 46$). Shown in blue and red are the 760 significantly differentially methylated CpGs with a Bonferroni-adjusted *P*-value threshold of < 0.05. (**B**) The mean methylation *Z*-score of 220 CpGs (*MORC2* episignature) (*y*-axis) was calculated for each individual with blood DNA methylation. Each filled circle represents one individual. The controls (grey-filled circles) consist of healthy individuals ($n = 58$) and patients with mutations in other genes than *MORC2* ($n = 424$). The yellow-filled circle shows the individual with a benign stopgain variant in *MORC2* (p.Arg90Ter), and the pink-filled circles represent the *MORC2* patients. (**C**) Results of the support vector machine classifier trained on blood DNA methylation of 220 CpGs to predict *MORC2* patients. The *y*-axis shows the probability that the DNA methylation profile matches the *MORC2*-specific episignature. Each dot represents one individual, and the *x*-axis indicates the causal genes or groups. 'Mitochondrial': Patients with pathogenic variants in mitochondrial disease genes ($n = 301$). 'MORC2_Benign': Two patients with benign missense variants in *MORC2*. 'MORC2_LOF': One patient with a benign loss-of-function variant in *MORC2*. Training set: Blood DNA methylation of 10 *MORC2* patients and 46 controls. Test set: Blood DNA methylation of three *MORC2* patients, 12 controls, one MORC2_Benign, 367 patients with other genetic disorders, and fibroblast DNA methylation of six *MORC2* patients, four controls and 14 patients with mitochondrial disease. Validation set: Blood DNA methylation of 22 *MORC2* patients, 20 patients with other genetic disorders, and fibroblast DNA methylation of six *MORC2* patients and one patient with another genetic disorder. (**D**) Localization of the 760 CpGs identified in **A** relative to gene structure and known MORC2 target sites. 'Body' refers to CpG sites located within gene bodies and 3′-untranslated regions (3′UTRs). 'TSS' denotes the transcription start site. 'TSS200' represents CpG sites located 0–200 bases upstream of the TSS, and 'TSS1500' represents CpG sites located 200–1500 bases upstream of the TSS. Fisher's exact test showed significant enrichment of differentially methylated CpG sites. (**E**) Venn diagram illustrating the overlap among CpG sites located in promoter regions, MORC2 binding sites and HUSH target sites. Of the 760 differentially methylated CpG sites, 463 that belong to at least one of these categories are included in the diagram. LOF = loss-of-function.

classifier successfully identified all *MORC2* patients in the test set (Fig. 2C).

DNA methylation patterns vary significantly across tissues. To date, DNA methylation episignatures have been identified and applied in diagnostics on blood samples. However, blood is often not the most suitable tissue for functional studies, whereas fibroblasts are well-established in diagnostic routines for pathomechanistic investigations. We sought to evaluate if the *MORC2*-specific episignature found in blood could be detected across different tissues. Therefore, we tested the classifier on DNA methylation data from 24 fibroblast cell lines, including six from *MORC2* patients. All the *MORC2* patients were successfully classified, as with the blood samples (Fig. 2C). This finding demonstrates that the disease-specific methylation changes that build the episignature are conserved across different tissues and stronger than tissue-dependent differences.

Since the classifier was solely tested on *MORC2* patients with MD, we sought to evaluate its phenotype specificity. We further collected 28 samples from *MORC2* patients, including 22 with a CMT phenotype and 21 additional disease controls, including one patient with a benign missense variant in *MORC2*. All groups were correctly classified except for three *MORC2* CMT patients that were falsely classified as controls (AUC = 0.988) (Fig. 2C). Even though the probability from the classifier was a little bit elevated, it was still in the benign range. All the variants reported in these three patients are located in the transducer S5-like domain, representing three of seven variants in this domain that did not show the episignature. Further investigation is needed to clarify the pathogenicity of these variants. Overall, we demonstrated that the episignature is universal for all *MORC2*-related disorders investigated.

Leveraging on the high specificity of the *MORC2* episignature, we tested the classifier on the DNA methylation profile of a mitochondrial disease patient carrying a VUS in *MORC2* (NM_001303256.3: c.535C>T, p.Arg179Cys), classified according to the American College of Medical Genetics/Association for Molecular Pathology guidelines. The variant is located in the hotspot region of the ATPase module. The classifier showed a clear benign result, which was later supported by segregation analysis showing the presence of the variant in the healthy, unaffected mother. We concluded that this variant is not a significant contributor to the observed disease phenotype and reclassified it as benign (Fig. 2C).

To date, all pathogenic variants reported in *MORC2* are exclusively missense variants, with a suspected gain-of-function mechanism.[10,12] To verify this hypothesis, we assessed the impact of a heterozygous stop-gain variant (NM_001303256.3: c.268C>T, p.Arg90Ter) in *MORC2* in an individual with an unrelated phenotype of hypertension. The classifier correctly predicted this variant to be benign, supporting the hypothesis that gain-of-function, not loss-of-function (LOF), is the pathomechanism of pathogenic *MORC2* variants (Fig. 2B and C).

Predominant localization of *MORC2* episignature in gene promoters

Currently, the primary motivation to search for episignatures and the value of an identified episignature lies in its potential for functional validation of VUS and its impact on subsequent variant classification. However, the disease-specific methylation changes have broadly failed to provide further insights into the Mendelian disease pathomechanisms. To explore MORC2 pathomechanism, we analysed the spatial distribution of DMPs in the genome. In a previous study by Tchasovnikarova *et al.*,[10] MORC2 binding sites were identified and found to be enriched at transcription start sites (TSSs) and HUSH targets. Indeed, we observed a significant enrichment of DMPs in MORC2 binding sites[10] and gene promoter regions, as well as HUSH target sites[23] (Fig. 2D). Of the DMPs located at MORC2 binding sites (23%), 60% are located in gene promoter regions and 14% overlapped with HUSH target sites (Fig. 2E). This observation aligns with previous studies indicating that MORC2 binding sites are mainly located at TSSs and, to a lesser extent, at HUSH target loci within gene bodies.[10] Moreover, among the DMPs that do not overlap with known MORC2 target sites, 30% are also located in gene promoter regions, while 10% overlap with HUSH target sites. (Figure 2D and E). Overall, this suggests that although some of the effects of mutant MORC2 on DNA methylation are mediated through interactions with the HUSH complex, the majority are driven by the HUSH-independent function of MORC2.

Around 60% of hypermethylation sites have not been previously identified either as MORC2 targets or HUSH targets. This could represent unrecognized physiological targets of MORC2 or novel targets arising as a result of *MORC2* gain-of-function variants. To differentiate between the two alternatives, we took advantage of the *MORC2* LOF. We hypothesized that physiological targets of MORC2 should show a reduced level of methylation in the carrier of the LOF variant. Indeed, across all known HUSH target sites, the LOF variant in *MORC2* resulted in reduced methylation. This pattern is even more pronounced in the newly identified sites. While DMPs are hypermethylated in *MORC2* patients, they are hypomethylated in the heterozygous LOF variant carrier (Supplementary Fig. 2). Taken together, these data support the idea that increased methylation occurs mostly at MORC2 physiological targets rather than at novel binding sites arising from gain-of-function mutations.

Overall, our analysis demonstrates that the CpGs comprising the *MORC2* episignature are predominantly located in gene promoters, and the comparison of *MORC2* gain-of-function and LOF variants further extends the list of previously identified MORC2 targets.

MORC2 mutants drive specific gene repression, resulting in an RNA signature

Building on the observation that *MORC2* mutations induce DNA hypermethylation in gene promoters, we next explored the functional impact of this promoter hypermethylation on gene expression, leveraging the conserved episignature in fibroblast cell lines. Differential expression analysis of 10 *MORC2* patients and 42 control cell lines revealed 106 DEGs at FDR-adjusted *P*-value < 0.05 (Fig. 3A and Supplementary Tables 5 and 6). Notably, 97% of these genes (103 of 106) were downregulated, and nearly half exhibited a significant increase in promoter methylation, consistent with the concept that hypermethylation in promoter regions results in the silencing of gene expression. Furthermore, these promoter regions were also hypermethylated in blood samples, suggesting an early establishment of MORC2 gain-of-function specific methylation changes in patients.

We then asked whether the DEGs were specific and sensitive enough to classify *MORC2* patients. Using the same approach as for episignature discovery, we applied an unsupervised analysis on both DESeq2-normalized (size factor normalization followed by variance stabilizing transformation) and OUTRIDER-normalized (employing denoising autoencoder)[28] counts of 106 DEGs. A separation between *MORC2* patients and controls was observed, with the separation being more distinct when using OUTRIDER-normalized counts (Supplementary Fig. 1B and C). Subsequently, we trained a SVM

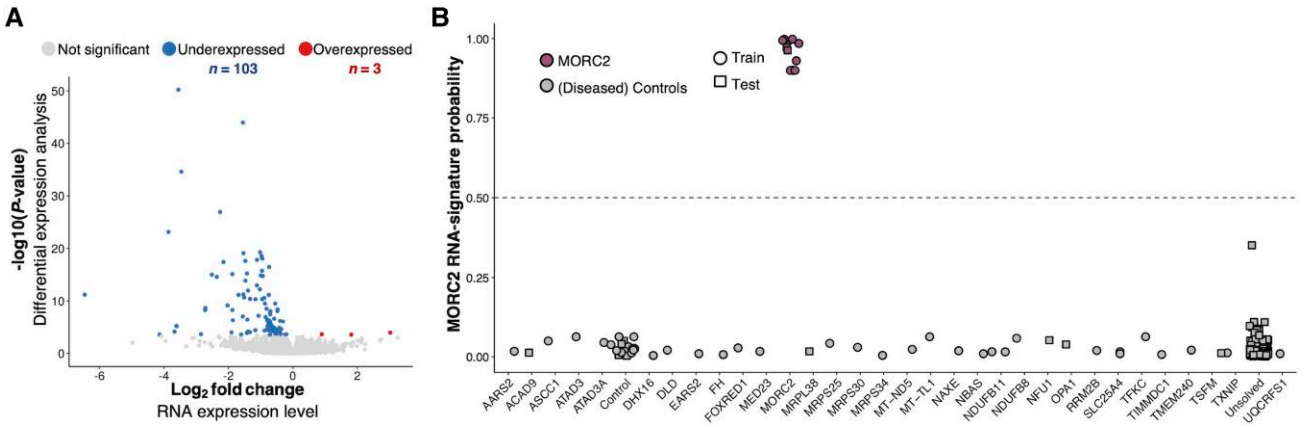

Figure 3 MORC2-specific RNA-signature in fibroblast cell lines. (**A**) Volcano plot of differential expression analysis in fibroblast cell lines comparing *MORC2* patients ($n = 10$) to healthy or diseased controls ($n = 42$). The 106 significant differentially expressed genes at a Benjamin–Hochberg-adjusted *P*-value threshold of <0.05 are shown in blue and red. (**B**) Results of the support vector machine classifier trained on the expression profile of the 106 genes identified in **A**. The *y*-axis shows the probability that the transcriptome profile matches the *MORC2*-specific RNA signature. Training set: 10 *MORC2* patients, 13 healthy controls, 29 diseased controls. Test set: Two *MORC2* patients, five healthy controls, five diseased controls and 168 genetically undiagnosed patients.

classifier on the same set of samples and tested it on an independent cohort of 180 samples (two *MORC2* patients, 10 healthy or disease controls and 168 genetically undiagnosed patients). The classifier successfully distinguished the *MORC2* patients from all other samples (Fig. 3B). In conclusion, heterozygous pathogenic variants in *MORC2* result in a hyperactive enzyme that causes promoter hypermethylation of MORC2 targets, leading to the silencing of a defined set of genes. This effect is strong and specific enough to establish an RNA-based *MORC2* classifier.

MORC2 downstream targets provide mechanistic insight into the pleiotropy of *MORC2*

Having identified *MORC2*-specific dysregulated genes, we questioned whether these targets share common features. Gene set enrichment analysis revealed a significant overrepresentation of C2H2-type zinc finger (ZNF) genes among both DMGs and DEGs, with C2H2-ZNF genes comprising 43% of the DEGs ($n = 47$) (Fig. 4). As a proof of concept, ZNF genes have previously been shown to be silenced in *MORC2* mutant cell lines. However, the repression has been attributed to histone methylation (H3K9me3) via the HUSH-MORC2 complex.[10,13,29] Our data showed that promoter hypermethylation of ZNF genes also contributes to their transcriptional repression. Despite this enrichment, none of the dysregulated C2H2-ZNF genes identified in our study have been associated with a genetic disorder.

We next asked whether the dysregulated genes at both methylation and expression levels have been associated with Mendelian disorders. Four disease-associated genes—*ERCC8, NDUFAF2, FKTN* and *NHLRC1*—were identified (Supplementary Fig. 3A). Notably, two of these genes, *ERCC8* and *NDUFAF2*, share the same bidirectional promoter. Three of the four genes—*ERCC8, NDUFAF2* and *FKTN*—were also identified as recurrent expression outliers in *MORC2* patients, with reduced expression levels typical of monogenic disorders caused by primary mutations in these genes (Fig. 5A–D).

To further validate the low expression level of the dysregulated disease genes, we performed proteomics on 13 fibroblast cell lines of *MORC2* patients. While ERCC8 and FKTN proteins were not detected in our proteomics assay, NDUFAF2 was detected and found as a recurrent protein outlier (Fig. 5C). The level of NDUFAF2 protein in *MORC2* patients' fibroblasts was as low as in patients with pathogenic variants in *NDUFAF2* itself (protein *Z*-score < –2).

NDUFAF2 encodes a mitochondrial respiratory chain complex I (CI) assembly factor that facilitates the last step of CI assembly, integrating the N module into CI. Pathogenic variants in *NDUFAF2* have previously been reported in patients with mitochondrial disease and Leigh syndrome and decreased CI levels and activity.[30-32] In addition to the reduced level of NDUFAF2, our proteomics data also show a significant reduction of CI N- and Q-module proteins, although their RNA expression is within the normal physiological range (Fig. 5C and Supplementary Table 7). This finding underscores the functional downstream effect of the NDUFAF2 dysfunction. Overall, the reduced level of NDUFAF2 and the consequent impairment of respiratory chain CI establish a functional link between *MORC2* mutations and mitochondrial dysfunction and explain the clinical presentation of Leigh syndrome.

Biallelic pathogenic variants in *ERCC8*, mostly LOF variants, have been reported in patients with Cockayne syndrome,[33,34] another phenotype observed in *MORC2* patients.[8] *ERCC8* encodes a WD repeat protein known as Cockayne syndrome A protein (CSA), which plays a role in DNA damage repair.[35] Accumulation of DNA damage has been observed in *Morc2a* mutant mice and has remained so far unexplained.[36] We conclude that the reduced expression of *ERCC8* contributes to the Cockayne syndrome features observed in *MORC2* patients, including signs of accelerated ageing.

FKTN, another dysregulated gene, encodes fukutin, an enzyme involved in the glycosylation of α-dystroglycan. Biallelic pathogenic variants in *FKTN* are associated with dystroglycanopathies, a spectrum of muscular dystrophies that can also be accompanied by brain and eye malformations as well as cardiomyopathy.[37,38] Muscular dystrophy has been reported in *MORC2* patients, and some have shown elevated serum creatine kinase levels, a biomarker of muscle damage.[9,39] Nevertheless, cobblestone lissencephaly, a hallmark feature of *FKTN*-related disorders on brain MRI, has not been observed so far in *MORC2* patients. Our findings suggest that *FKTN* downregulation may contribute to the muscular dystrophy observed in *MORC2* patients. Additionally, we suggest that *MORC2* variants be considered in patients presenting with *FKTN*-specific brain abnormalities.

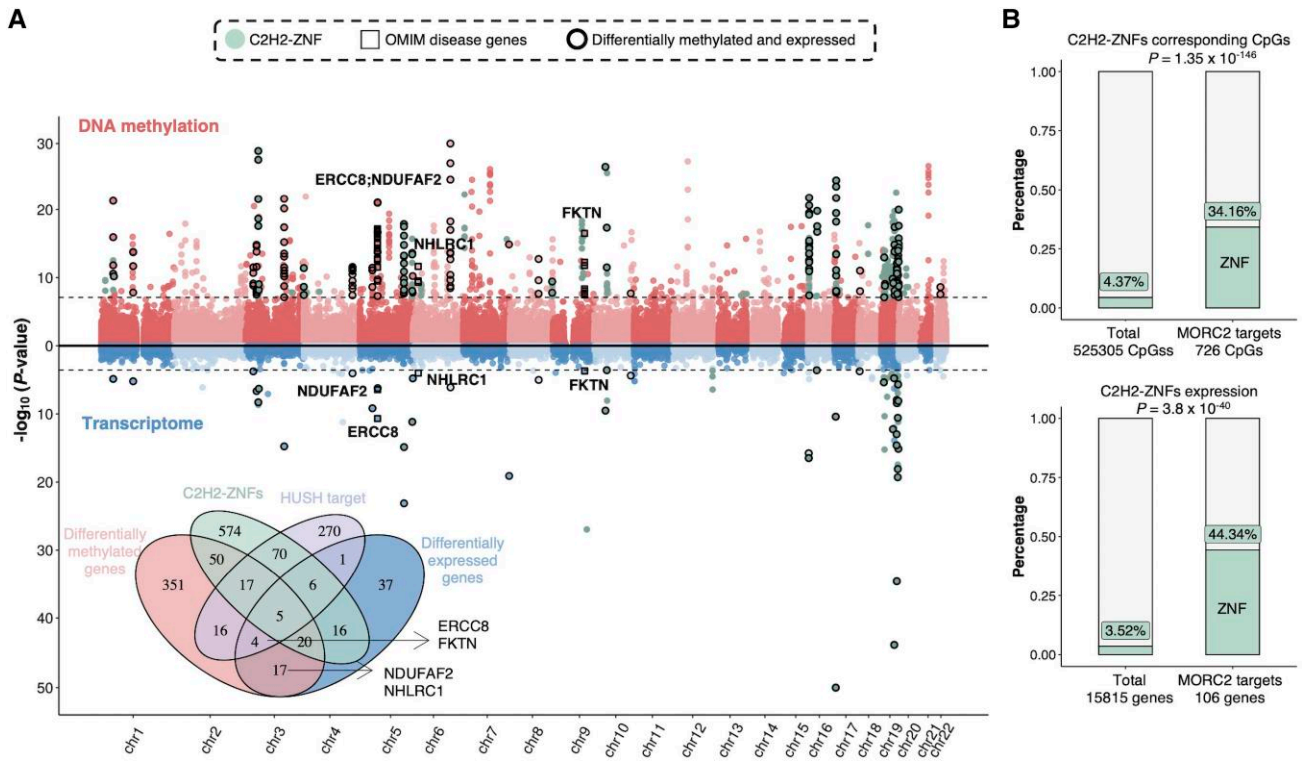

Figure 4 The downstream effect of aberrant DNA methylation on gene expression. (**A**) Mirror Manhattan plot displaying the results of differential methylation analysis (*top*, red-filled circles) and differential expression analysis (*bottom*, blue-filled circles). The *y*-axis represents $-\log_{10}$(*P*-value) for both analyses, and the *x*-axis denotes the genomic positions of CpG sites or genes. Threshold *P*-values were set at 7.93×10^{-8} (corresponding to Bonferroni-adjusted *P*-value < 0.05) for differential methylation and 2.7×10^{-4} (corresponding to Benjamin–Hochberg-adjusted *P*-value < 0.05) for differential expression analyses (indicated by dashed lines). Differentially methylated CpG sites mapped to C2H2-ZNF genes and differentially expressed C2H2-ZNF genes are highlighted in green. Genes that are both differentially methylated and differentially expressed are shown in bold. OMIM disease genes that are both differentially methylated and expressed are enclosed in rectangles and labelled. The Venn diagram at the *bottom* illustrates the overlap between C2H2-ZNFs, HUSH target genes,[23] differentially expressed genes and differentially methylated genes. (**B**) Enrichment of C2H2-ZNFs among both differentially methylated CpGs (*top*) and differentially expressed genes (*bottom*) with Fisher's exact test *P*-values of 1.35×10^{-146} and 3.18×10^{-40}, respectively. OMIM = Online Mendelian Inheritance in Man; ZNFs = zinc finger genes.

While mitochondrial CI deficiency was the predominant biochemical defect observed in *MORC2* patients, complex II (CII) deficiency was occasionally noted. In agreement with this, we observed promoter hypermethylation and reduced expression of *SDHAF4*, a CII assembly factor, in two patients with CII deficiency (Fig. 5E). This observation highlights the variability of MORC2 targets among patients, as evidenced in this case, leading to CII deficiency in a small subset of individuals.

Additionally, we observed a significant downregulation of *MORC2* itself, independent of any significant changes in DNA methylation (Supplementary Fig. 3B). Previous studies have identified MORC2 as a target of the HUSH complex, with its downregulation observed in *MORC2* mutant cell line.[10] Our findings align with previous findings, supporting the hypothesis that the reduced *MORC2* expression is likely mediated by histone methylation (H3K9me3) via the HUSH complex rather than through DNA methylation.

Finally, we observed only marginal promoter hypermethylation and reduced expression of the *NHLRC1* gene. The gene was not identified as an expression outlier (Supplementary Fig. 3C). The pathogenic variants in *NHLRC1* have been associated with myoclonus epilepsy, a phenotype not observed in our cohort but previously reported in some *MORC2* patients.[9] Further investigation of *NHLRC1* promoter methylation in those patients is warranted.

The level of methylation predicts phenotypic differences in *MORC2*-related disorders

MORC2 patients present with overlapping symptoms of different disorders. Besides MRI findings which are specific to Leigh syndrome, all other symptoms are observed in subsets of both CMT and mitochondrial disease *MORC2* patients (Fig. 6A and Supplementary Table 1). Our integrative multi-omics analysis identified the key genes driving the *MORC2*-related phenotypes, highlighting their composite nature. The *MORC2*-related disorder is the consequence of variable quantitative gene silencing of at least three genes—*NDUFAF2*, *ERCC8*, *FKTN*—alongside additional genes that are affected in some but not all patients. How much does the genotype affect the variability? Generally, there is some correlation between genotype and phenotype. Variants in the N-terminal GHKL domain are more frequently associated with severe neurodevelopmental disorders and MD, while those in the C-terminal GHKL domain and transducer-like domain are more commonly linked to CMT. However, several variants have been reported in association with variable clinical presentation, even among members of the same family (Fig. 1C and Supplementary Table 1).[40]

We therefore asked whether the methylation pattern may help in predicting patient phenotypes. We selected patients from the two most distinct *MORC2*-related disorders, patients with Leigh syndrome ($n = 10$) from the mitochondrial disease cohort and patients with the

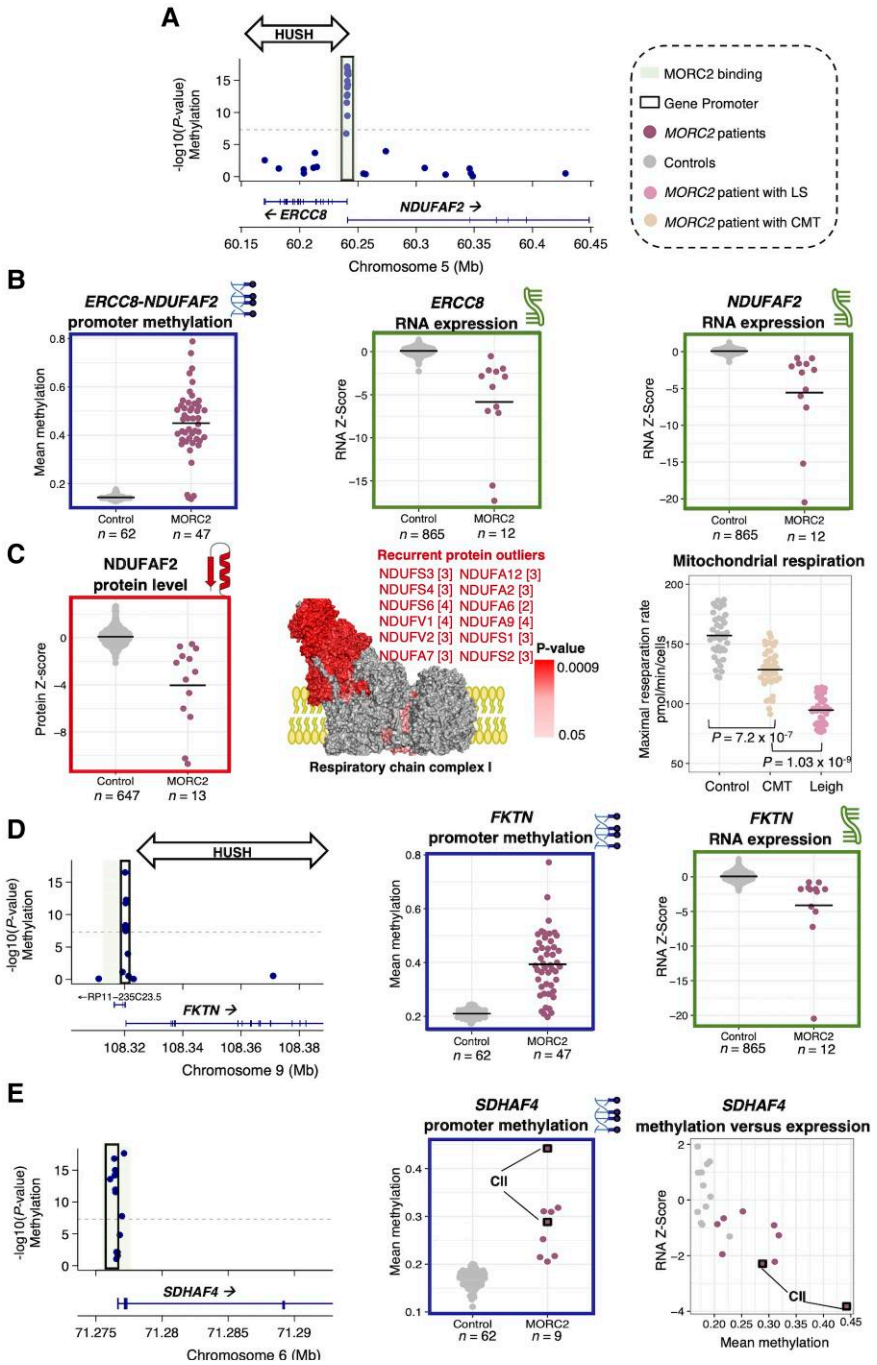

Figure 5 Hypermethylation and downregulation of disease-associated genes by *MORC2* gain-of-function variants. (**A**) LocusZoom plot for *ERCC8* and *NDUFAF2* genes based on the differential methylation analysis (from Fig. 2A). Each dot represents a CpG site, and the *y*-axis indicates the $-\log_{10}$(*P*-value) of the differential methylation analysis. (**B**) Promoter methylation and expression levels of *ERCC8* and *NDUFAF2* in *MORC2* patients and controls. *Left*: Mean methylation (*y*-axis) represents the average of the methylation level (beta value) across 14 CpGs located in the *ERCC8-NDUFAF2* promoter region. *Right*: RNA *Z*-scores derived from OUTRIDER results for *ERCC8* and *NDUFAF2* expression levels, respectively. (**C**) NDUFAF2 protein levels in *MORC2* patients and controls. *Left*: Protein *Z*-scores of NDUFAF2 derived from PROTRIDER results. *Middle*: Mitochondrial respiratory chain complex I with significant recurrent protein outliers (protein *Z*-score < −2) in *MORC2* patients highlighted and named in red (one-sided Fisher's exact test *P*-value < 0.05). The number of *MORC2* patients in whom the protein was identified as an outlier is written in brackets. Proteins shown in grey were either not detected or not significant. *Right*: Mitochondrial respiration rate in control and *MORC2* patients with Charcot–Marie–Tooth disease (CMT) or Leigh syndrome (LS). *P*-values were computed using an independent two-sample *t*-test. (**D**) LocusZoom plot for the *FKTN* gene based on the differential methylation analysis (from Fig. 2A). Promoter methylation and expression levels of *FKTN* in *MORC2* patients and controls are shown, respectively. The mean methylation represents the average of the methylation level (beta value) across 7 CpG sites located in the *FKTN* promoter region. (**E**) LocusZoom plot for *SDHAF4* gene-based based on the differential methylation analysis (from Fig. 2A). The promoter methylation level of *SDHAF4* in *MORC2* patients with measured mitochondrial respiratory chain complex II activity and controls are shown (*middle*). *MORC2* patients with complex II deficiency are indicated by rectangles. The mean methylation represents the average of the methylation level (beta value) across 8 CpGs located in the *SDHAF4* promoter. The correlation between the level of promoter methylation and the expression (RNA *Z*-score) is shown (*right*). The dashed line in the locusZoom plots represents the *P*-value threshold of 7.93×10^{-8}, corresponding to a Bonferroni-adjusted *P*-value < 0.05. MORC2 binding sites[10] and HUSH target sites[23] are shown as green rectangles and double-sided arrows in the locusZoom plots, respectively.

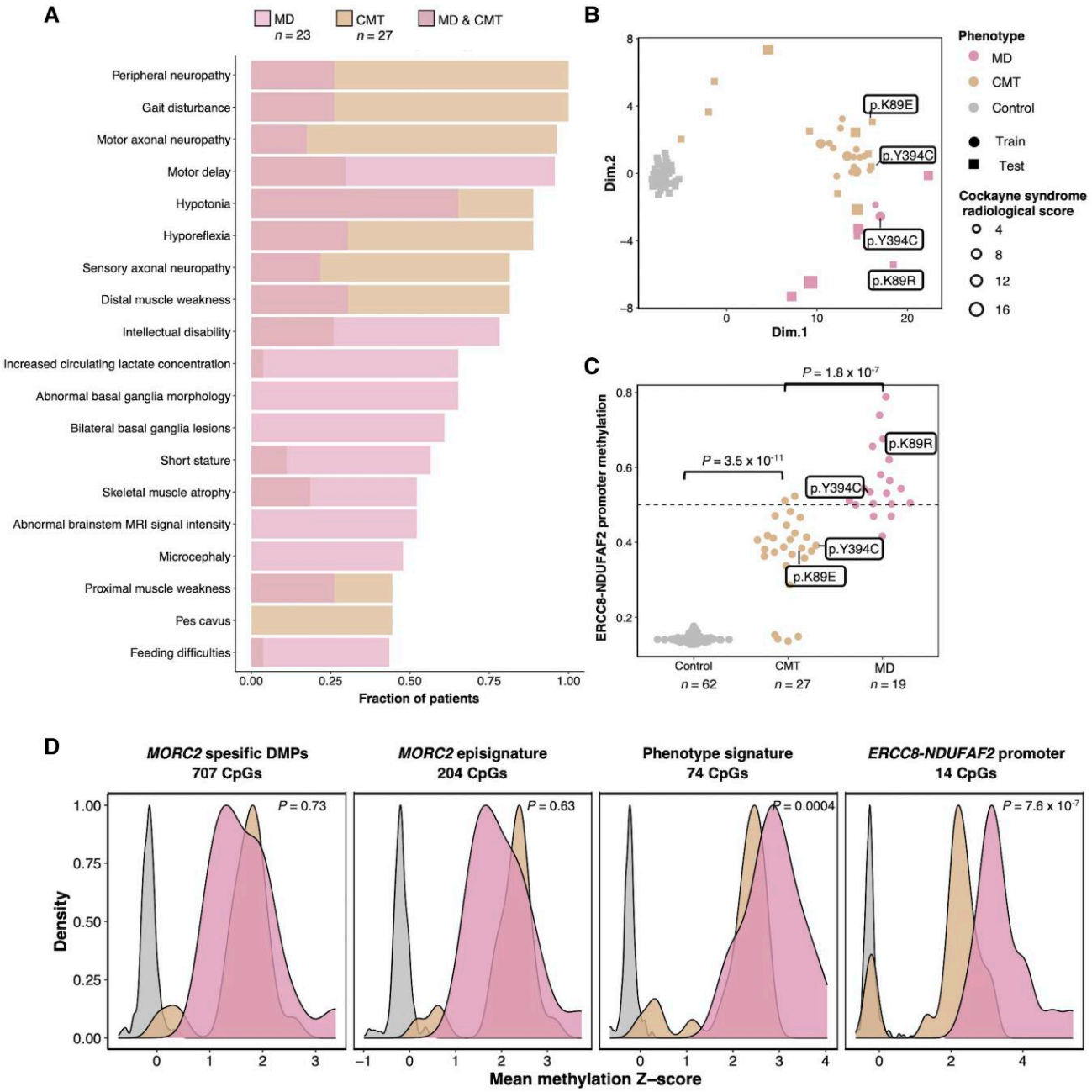

Figure 6 The phenotypic spectrum of *MORC2*-related disorder can be explained by DNA methylation. (**A**) Stacked bar plot showing the distribution of HPO terms in our in-house *MORC2* patients with clinical diagnoses of Charcot–Marie–Tooth disease (CMT) or mitochondrial disease (MD). The *x*-axis represents the fraction of patients with a clinical diagnosis of CMT or mitochondrial disease exhibiting specific phenotypic abnormalities. Only HPO terms reported in at least 40% of patients within each category are displayed. (**B**) Multidimensional scaling plot based on blood DNA methylation of 80 phenotype-specific differentially methylated positions (DMPs). The size of each point reflects the Cockayne syndrome radiological score, with larger points indicating a greater resemblance of patients' phenotype to Cockayne syndrome. The dataset includes *MORC2* patients with CMT phenotype ($n = 26$), *MORC2* patients with mitochondrial disease ($n = 8$) and controls ($n = 58$). (**C**) The methylation level of *ERCC8* and *NDUFAF2* promoter in controls and MORC2 patients with a clinical diagnosis of CMT or mitochondrial disease in blood and fibroblast. The *y*-axis represents the average methylation level (beta value) of 14 CpGs mapped to the *ERCC8-NDUFAF2* promoter. *P*-values were computed using an independent two-sample *t*-test. Patients with Leigh syndrome are included in the category of mitochondrial disease. The variants associated with different phenotypes are labelled in **B** and **C**. (**D**) Scaled density plots of the mean DNA methylation *Z*-scores of hypermethylated CpGs identified in the EWAS, selected as *MORC2* episignature, identified as phenotype specific and located in the *ERCC8-NDUFAF2* promoter. The dataset includes blood or fibroblast DNA methylation of 466 healthy or diseased controls, 27 *MORC2* patients with CMT and 19 *MORC2* patients with mitochondrial disease. *P*-values were computed to compare the level of DNA methylation between the *MORC2* patients with CMT and mitochondrial disease using an independent two-sample *t*-test. EWAS = epigenome-wide association study; HPO = Human Phenotype Ontology.

typical presentation of CMT ($n = 16$), as well as 16 controls for differential methylation analysis (see the 'Materials and methods' section for details) and identified 80 DMPs at a Bonferroni-adjusted *P*-value < 0.05 (Supplementary Tables 3 and 4). Notably, 88% of the 80 DMPs had already been identified as *MORC2*-specific DMPs. A multidimensional scaling plot using these 80 DMPs showed a clear separation between

MORC2 patients and controls in the first dimension. The second dimension separated *MORC2* patients with CMT from those with mitochondrial disease (Fig. 6B and Supplementary Fig. 4C). A classifier was trained to predict the phenotypes and tested on an independent *MORC2* cohort, including 11 patients with CMT and 9 with MD, achieving high accuracy with an AUC of 0.99 (Supplementary Fig. 4A and B). The classifier correctly predicted the phenotypes of two pairs of patients with the same affected amino acid but different Leigh syndrome and CMT phenotypes (Supplementary Fig. 4A). This finding suggests that DNA methylation adds information on top of the genetic variant in predicting the patient phenotype. Given that *MORC2*-related disorders exist as a continuum (Fig. 6A), we recommend interpreting the phenotype-classifier results as probabilities indicating whether the clinical presentation is more characteristic of mitochondrial disease rather than CMT.

Among the identified DMPs, a significant proportion (16%) were located in the *ERCC8-NDUFAF2* bidirectional promoter. Mitochondrial disease patients exhibited significantly higher methylation levels than those with CMT, while CMT patients still demonstrated significantly elevated methylation levels relative to controls (Fig. 6C). These findings suggest that mitochondrial dysfunction is also present in CMT patients. When comparing mitochondrial respiration, the fibroblast cell line from a CMT patient exhibited an intermediate phenotype, falling between that of a Leigh syndrome patient with CI deficiency and a healthy control (Fig. 5C, right).

We further investigated whether methylation levels across all DMPs were elevated in *MORC2* patients with mitochondrial disease compared to CMT. No significant differences in methylation levels were observed between phenotypes when considering all *MORC2*-specific DMPs. While a significant increase in DNA methylation was observed in mitochondrial disease patients across the 80 phenotype-specific DMPs, the elevation was more pronounced at CpGs within the *ERCC8-NDUFAF2* promoter (Fig. 6D). This suggests that methylation at this promoter plays a pivotal role in driving the phenotypic heterogeneity in *MORC2*-related disorders.

Discussion

DNA methylation episignatures are increasingly recognized as specific biomarkers for a subset of rare disorders.[41] However, their application has so far been predominantly limited to diagnostic biomarkers, with minimal use in understanding disease pathomechanisms. In this study, we demonstrate that patients with pathogenic gain-of-function variants in *MORC2* share a common DNA methylation pattern despite displaying varied biochemical defects and broad clinical presentations. Strikingly, promoter hypermethylation of the target genes, which constitutes the core of the episignature, is conserved across tissues and results in reduced RNA expression and protein levels with an effect size comparable to the Mendelian disorders due to pathogenic variants impacting the target genes themselves. The repression of known disease-associated genes and the degree of their downregulation explain the pleiotropy seen in *MORC2* patients. These *MORC2*-specific methylation changes further extend our knowledge of the molecular function of MORC2 and its downstream targets.

MORC2 patients have so far been classified as having distinct disorders, such as CMT, SMA, Cockayne syndrome, Leigh syndrome and mitochondrial disease, based on their predominant symptoms. However, the phenotypic presentation is broader than these classifications, with overlapping features of different disorders that blur the boundaries between categories. Sacoto *et al.*[9] attempted to address the phenotypic variability by proposing a broad category named DIGFAN, encompassing all *MORC2*-related disorders except CMT. However, patients rarely present with the full spectrum of DIGFAN features, and diagnosis remains reliant on identifying the underlying genetic cause in *MORC2*. In this study, we demonstrated that the composite repression of multiple disease-associated genes driven by MORC2 hyperactivity underlies its pleiotropic effects. The variability in target genes and the level of repression amongst affected individuals account for the observed phenotypic heterogeneity, even in patients with the same mutation. A consistent observation across patients is impaired energy metabolism driven by mitochondrial CI deficiency. These results suggest that *MORC2*-related disorders exist along a spectrum, ranging from mild CMT to severe Leigh syndrome. We propose moving away from splitting *MORC2*-related disorders into separate categories and instead lumping them into a single disease spectrum with a shared pathomechanism.

All *MORC2* patients diagnosed with mitochondrial disease or Leigh syndrome had an age of onset within the first 2 years of life, which is also reported for all *NDUFAF2* patients. Neurological abnormalities in *NDUFAF2* patients primarily involve the brainstem, cerebellum, upper cervical spinal cord and basal ganglia,[30] similar to MRI findings in *MORC2* patients. Infantile-onset brainstem neurodegenerative disease with early lethality is also the reported phenotype associated with a bi-allelic *ERCC8-NDUFAF2* deletion,[42] further supporting that *NDUFAF2* repression is the major cause of Leigh syndrome in *MORC2* patients.

The *MORC2* episignature was found to be conserved across tissues despite the large differences in DNA methylation between blood and fibroblasts. To date, DNA methylation episignatures have mainly been identified and validated in blood samples, with only the Sotos episignature replicated in fibroblasts on a small sample size.[43] Here, we provide an additional example of episignature stability across tissues. This cross-tissue stability has the advantage of allowing the performance of functional studies on different cell lines, including fibroblast cell lines, to assess the effect of potential treatment options. The conservation of the episignature across tissues raises the question of when the episignature is established. Correlation between the blood and fibroblast methylation pattern with the biochemical defects and clinical manifestation affecting different organs argues that the *MORC2* methylation signature is likely determined early in development.

MORC2 binds to DNA in a non-sequence-specific manner,[44] modulating chromatin structure to facilitate gene silencing. While a single-gene study has shown that MORC2 forms a complex with DNA methyltransferases (DNMTs) at gene promoters to mediate transcription repression,[15] a systematic analysis of its interactions with DNMTs is still lacking. The promoter hypermethylation due to *MORC2* mutations observed in this study supports the interaction between MORC2 and DNMTs. Further functional studies are needed to clarify the mechanisms by which MORC2 binds to gene promoters and interacts with chromatin modifiers such as DNMTs. Of note, not all changes in gene expression can be directly attributed to promoter hypermethylation. Some of these changes may be secondary effects mediated by the repression of ZNFs, some of which function as transcription factors and could indirectly influence RNA expression. However, we found no evidence that any repressed ZNFs regulate any known Mendelian disease genes.

How the methylation pattern gets established is beyond the focus of the current study. The observation of patients with the same mutation but different methylation patterns may reflect a complex mechanism with a number of components involved in the establishment of the patient-specific signature. Such variability may

stem from random influences, differences in the proteins involved, or polymorphisms in the target regions. We could not identify any polymorphisms in the promoter regions of the target genes that were associated with methylation levels, probably due to the small cohort of patients with the same genotype. A larger cohort would increase the power to investigate this question.

RNA-seq has recently been used as a complementary diagnostic tool to DNA sequencing by quantifying the direct effects of genetic variants on candidate gene expression and splicing.[24,25] To our knowledge, the *MORC2* RNA-signature discovered in this study is the first gene-specific RNA-signature for a monogenic disorder. *MORC2* might not be the only gene where downstream target genes can be linked to pleiotropic effects. Transcriptome studies from patients with disorders with known episignatures can indicate whether the observed RNA signature is unique to *MORC2* or can be extended to other genes. Screening for RNA-signature could broaden the diagnostic application of RNA-seq.

Episignatures hold significant potential as both therapeutic targets and biomarkers for assessing treatment efficacy. DNMT inhibitors, such as azacitidine and decitabine, have demonstrated clinical benefit in certain haematological malignancies.[45] However, their applicability in the treatment of *MORC2*-related disorders may be limited by their toxicity and off-target effects. Given the gain-of-function pathomechanism, direct targeting of the *MORC2* disease allele, either at the RNA level using antisense oligonucleotides or at the protein level with inhibitors like geldanamycin or radicicol,[46,47] presents another promising therapeutic strategy. Further research is warranted to evaluate the efficacy of these approaches, both individually and in combination, in the treatment of *MORC2*-related disorders.

The finding of MORC2 gain-of-function driven DNA methylation accounting for the phenotypic pleiotropy has implications beyond *MORC2*-related disorders. First, we hypothesize the existence of disease drivers in other disorders with episignatures, specifically when they are conserved across tissues. Outlier detection may help to identify these drivers. Second, we hypothesize an epigenetic contribution to other Mendelian disorders with known pleiotropy.

Data availability

The dataset generated and analysed in the current study is available on request.

Acknowledgements

We are grateful to the families and patients without whom this study would not have been possible.

Funding

This study was supported by the BMBF (German Federal Ministry of Education and Research) through the mitoNET German Network for Mitochondrial Diseases (grant number 01GM1906B), PerMiM Personalized Mitochondrial Medicine (grant number 01KU2016A), the European Joint Programme Rare Diseases (EJP RD) project GENOMIT (01GM1920A, genomit.eu), the German Center for Child and Adolescent Health (DZKJ) under the funding code 01GL2406B, the Bavarian Ministry of Health, Care and Prevention within its framework of DigiMed Bayern (grant number DMB-1805-0002), and the European Union's Horizon 2020 research and innovation program project Recon4IMD (grant number 101080997). K.M. acknowledges support from Grant-in-Aid for Research on Measures for Intractable Diseases (H27-Nanji-Ippan-028) from the Ministry of Health, Labor and Welfare, Japan. K.M., Ya.O. and Yu.O. acknowledge the support from a grant for the Practical Research Project for Rare/Intractable Diseases from the Japan Agency for Medical Research and Development (AMED, Fund ID: 23ek0109625, 23ek0109636, 24lk0221189s0301). M.A. and Yu.O. were supported in part by Grant-in-Aid from the Research Committee for Ataxic Disease, the Ministry of Health, Labor and Welfare, Japan (2016100002B). M.A. and Yu.O. acknowledge support from the Japan Agency for Medical Research and Development (grant numbers 201442014A and 201442071A). M.A. and Yu.O. have also received support from the Japan Society for the Promotion of Science (JSPS KAKENHI) (grant numbers JP18H02742, JP20K16604, JP21K15702, JP21H02842, JP22K15713, JP22K07495, JP22K07519, JP23K06931, and JP23K06966, JP24K18708). M.Z. receives research support from the German Research Foundation (DFG 458949627; ZE 1213/2-1). M.Z. acknowledges grant support from the European Joint Programme Rare Diseases (EJP RD Joint Transnational Call 2022) and the German Federal Ministry of Education and Research (BMBF, Bonn, Germany), awarded to the project PreDYT (PREdictive biomarkers in DYsTonia, 01GM2302), and from the German Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder, as well as by the Institute for Advanced Study, Technical University of Munich. M.Z.'s research is supported by a 'Schlüsselprojekt' grant from the Else Kröner-Fresenius-Stiftung (2022_EKSE.185). F.D. was supported by a grant of the German Research Foundation/Deutsche Forschungsgemeinschaft (DI 1731/2-3) and by a grant from the 'Elterninitiative Kinderkrebsklinik e.V.' (Düsseldorf; #701900167). C.L. is a member of the ERN-NMD and the interERN for mitochondrial disorders. C.S. was supported by the DFG (SCHL2276/2-1, 450149205-TRR333/1).

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Thaxton C, Goldstein J, DiStefano M, et al. Lumping versus splitting: How to approach defining a disease to enable accurate genomic curation. *Cell Genom*. 2022;2:100131.
2. Albulym OM, Kennerson ML, Harms MB, et al. MORC2 mutations cause axonal Charcot-Marie-Tooth disease with pyramidal signs. *Ann Neurol*. 2016;79:419-427.
3. Ando M, Okamoto Y, Yoshimura A, et al. Clinical and mutational spectrum of Charcot–Marie–Tooth disease type 2Z caused by MORC2 variants in Japan. *Eur J Neurol*. 2017;24:1274-1282.
4. Sevilla T, Lupo V, Martínez-Rubio D, et al. Mutations in the MORC2 gene cause axonal Charcot-Marie-Tooth disease. *Brain J Neurol*. 2016;139:62-72.
5. Schottmann G, Wagner C, Seifert F, Stenzel W, Schuelke M. MORC2 mutation causes severe spinal muscular atrophy-phenotype, cerebellar atrophy, and diaphragmatic paralysis. *Brain J Neurol*. 2016;139:e70.

6. Kistol D, Tsygankova P, Krylova T, et al. Leigh syndrome: Spectrum of molecular defects and clinical features in Russia. *Int J Mol Sci*. 2023;24:1597.
7. McCormick EM, Keller K, Taylor JP, et al. Expert panel curation of 113 primary mitochondrial disease genes for the Leigh syndrome Spectrum. *Ann Neurol*. 2023;94:696-712.
8. Stafki SA, Turner J, Littel HR, et al. The Spectrum of MORC2-related disorders: A potential link to Cockayne syndrome. *Pediatr Neurol*. 2023;141:79-86.
9. Guillen Sacoto MJ, Tchasovnikarova IA, Torti E, et al. De Novo variants in the ATPase module of MORC2 cause a neurodevelopmental disorder with growth retardation and Variable craniofacial dysmorphism. *Am J Hum Genet*. 2020;107:352-363.
10. Tchasovnikarova IA, Timms RT, Douse CH, et al. Hyperactivation of HUSH complex function by Charcot–Marie–Tooth disease mutation in MORC2. *Nat Genet*. 2017;49:1035-1044.
11. Xie HY, Zhang TM, Hu SY, Shao ZM, Li DQ. Dimerization of MORC2 through its C-terminal coiled-coil domain enhances chromatin dynamics and promotes DNA repair. *Cell Commun Signal*. 2019;17:160.
12. Douse CH, Bloor S, Liu Y, et al. Neuropathic MORC2 mutations perturb GHKL ATPase dimerization dynamics and epigenetic silencing by multiple structural mechanisms. *Nat Commun*. 2018;9:651.
13. Liu N, Lee CH, Swigut T, et al. Selective silencing of euchromatic L1s revealed by genome-wide screens for L1 regulators. *Nature*. 2018;553:228-232.
14. Shao Y, Li Y, Zhang J, et al. Involvement of histone deacetylation in MORC2-mediated down-regulation of carbonic anhydrase IX. *Nucleic Acids Res*. 2010;38:2813-2824.
15. Wang T, Qin ZY, Wen LZ, et al. Epigenetic restriction of Hippo signaling by MORC2 underlies stemness of hepatocellular carcinoma cells. *Cell Death Differ*. 2018;25:2086-2100.
16. Emmanuele V, Ganesh J, Vladutiu G, et al. Time to harmonize mitochondrial syndrome nomenclature and classification: A consensus from the North American mitochondrial disease consortium (NAMDC). *Mol Genet Metab*. 2022;136:125-131.
17. Wilson BT, Stark Z, Sutton RE, et al. The Cockayne Syndrome Natural History (CoSyNH) study: Clinical findings in 102 individuals and recommendations for care. *Genet Med*. 2016;18:483-493.
18. Spitz MA, Severac F, Obringer C, et al. Diagnostic and severity scores for Cockayne syndrome. *Orphanet J Rare Dis*. 2021;16:63.
19. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: A flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. *Bioinformatics*. 2014;30:1363-1369.
20. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13:86.
21. Pelegí-Sisó D, de Prado P, Ronkainen J, Bustamante M, González JR. Methylclock: A bioconductor package to estimate DNA methylation age. *Bioinformatics*. 2021;37:1759-1760.
22. Levy MA, Relator R, McConkey H, et al. Functional correlation of genome-wide DNA methylation profiles in genetic neurodevelopmental disorders. *Hum Mutat*. 2022;43:1609-1628.
23. Seczynska M, Bloor S, Cuesta SM, Lehner PJ. Genome surveillance by HUSH-mediated silencing of intronless mobile elements. *Nature*. 2022;601:440-445.
24. Yépez VA, Gusic M, Kopajtich R, et al. Clinical implementation of RNA sequencing for Mendelian disease diagnostics. *Genome Med*. 2022;14:38.
25. Kremer LS, Bader DM, Mertes C, et al. Genetic diagnosis of Mendelian disorders via RNA sequencing. *Nat Commun*. 2017;8:15824.
26. Kopajtich R, Smirnov D, Stenton SL, et al. Integration of proteomics with genomics and transcriptomics increases the diagnostic rate of Mendelian disorders. *bioRxiv*. [Preprint]. https://doi.org/10.1101/2021.03.09.21253187
27. Invernizzi F, D'Amato I, Jensen PB, Ravaglia S, Zeviani M, Tiranti V. Microscale oxygraphy reveals OXPHOS impairment in MRC mutant cells. *Mitochondrion*. 2012;12:328-335.
28. Brechtmann F, Mertes C, Matusevičiūtė A, et al. OUTRIDER: A statistical method for detecting aberrantly expressed genes in RNA sequencing data. *Am J Hum Genet*. 2018;103:907-917.
29. Douse CH, Tchasovnikarova IA, Timms RT, et al. TASOR is a pseudo-PARP that directs HUSH complex assembly and epigenetic transposon control. *Nat Commun*. 2020;11:4940.
30. Abu Hanna F, Zehavi Y, Cohen-Barak E, et al. Lack of mitochondrial complex I assembly factor NDUFAF2 results in a distinctive infantile-onset brainstem neurodegenerative disease with early lethality. *Orphanet J Rare Dis*. 2024;19:92.
31. Formosa LE, Dibley MG, Stroud DA, Ryan MT. Building a complex complex: Assembly of mitochondrial respiratory chain complex I. *Semin Cell Dev Biol*. 2018;76:154-162.
32. Hoefs SJG, Dieteren CEJ, Distelmaier F, et al. NDUFA2 Complex I mutation leads to leigh disease. *Am J Hum Genet*. 2008;82:1306-1315.
33. Bertola DR, Cao H, Albano LMJ, et al. Cockayne syndrome type A: Novel mutations in eight typical patients. *J Hum Genet*. 2006;51:701-705.
34. Calmels N, Botta E, Jia N, et al. Functional and clinical relevance of novel mutations in a large cohort of patients with Cockayne syndrome. *J Med Genet*. 2018;55:329-343.
35. Henning KA, Li L, Iyer N, et al. The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH. *Cell*. 1995;82:555-564.
36. Lee GS, Kwak G, Bae JH, et al. Morc2a p.S87L mutant mice develop peripheral and central neuropathies associated with neuronal DNA damage and apoptosis. *Dis Model Mech*. 2021;14:dmm049123.
37. Gaertner A, Burr L, Klauke B, et al. Compound heterozygous FKTN variants in a patient with dilated cardiomyopathy led to an aberrant α-dystroglycan pattern. *Int J Mol Sci*. 2022;23:6685.
38. Godfrey C, Escolar D, Brockington M, et al. Fukutin gene mutations in steroid-responsive limb girdle muscular dystrophy. *Ann Neurol*. 2006;60:603-610.
39. Sivera R, Lupo V, Frasquet M, et al. Charcot–Marie–Tooth disease due to MORC2 mutations in Spain. *Eur J Neurol*. 2021;28:3001-3011.
40. Duan X, Liu X, Wang G, et al. Characterization of genotype–phenotype correlation with MORC2 mutated axonal Charcot–Marie–Tooth disease in a cohort of Chinese patients. *Orphanet J Rare Dis*. 2021;16:244.
41. Kerkhof J, Rastin C, Levy MA, et al. Diagnostic utility and reporting recommendations for clinical DNA methylation episignature testing in genetically undiagnosed rare diseases. *Genet Med*. 2024;26:101075.
42. Sabharwal A, Gupta V, Kv S, et al. Whole genome sequencing followed by functional analysis of genomic deletion encompassing ERCC8 and NDUFAF2 genes in a non-consanguineous Indian family reveals dysfunctional mitochondrial bioenergetics leading to infant mortality. *Mitochondrion*. 2024;75:101844.
43. Choufani S, Cytrynbaum C, Chung BHY, et al. NSD1 mutations generate a genome-wide DNA methylation signature. *Nat Commun*. 2015;6:10207.

44. Fendler NL, Ly J, Welp L, et al. Identification and characterization of a human MORC2 DNA binding region that is required for gene silencing. *Nucleic Acids Res*. 2024;53: gkae1273.
45. Derissen EJB, Beijnen JH, Schellens JHM. Concise drug review: Azacitidine and decitabine. *Oncologist*. 2013;18: 619-624.
46. Sharma SV, Agatsuma T, Nakano H. Targeting of the protein chaperone, HSP90, by the transformation suppressing agent, radicicol. *Oncogene*. 1998;16:2639-2645.
47. Stebbins CE, Russo AA, Schneider C, Rosen N, Hartl FU, Pavletich NP. Crystal structure of an Hsp90–geldanamycin complex: Targeting of a protein chaperone by an antitumor agent. *Cell*. 1997;89:239-250.