Novel gene-intergenic fusion involving ubiquitin E3 ligase UBE3C causes distal hereditary motor neuropathy

Anthony N. Cutrupi,[1,2] Ramesh K. Narayanan,[1,2] Gonzalo Perez-Siles,[1,2] Bianca R. Grosz,[1,2] Kaitao Lai,[1,3] Alexandra Boyling,[1,2] Melina Ellis,[1,2] Ruby C. Y. Lin,[2,4] Brent Neumann,[5] Di Mao,[6] Motonari Uesugi,[6] Garth A. Nicholson,[1,7] Steve Vucic,[2,8] Mario A. Saporta[9,†] and Marina L. Kennerson[1,2,7,†]

[†]**These authors contributed equally to this work.**

Abstract

Distal hereditary motor neuropathies (dHMNs) are a group of inherited diseases involving the progressive, length-dependent axonal degeneration of the lower motor neurons. There are currently 29 reported causative genes and 4 disease loci implicated in dHMN. Despite the high genetic heterogeneity, mutations in the known genes account for less than 20% of dHMN cases with the mutations identified predominantly being point mutations or indels. We have expanded the spectrum of dHMN mutations with the identification of a 1.35 Mb complex structural variation (SV) causing a form of autosomal dominant dHMN (DHMN1 OMIM %182906). Given the complex nature of SV mutations and the importance of studying pathogenic mechanisms in a neuronal setting, we generated a patient-derived DHMN1 motor neuron model harbouring the 1.35 Mb complex insertion. The DHMN1 complex insertion creates a duplicated copy of the first 10 exons of the ubiquitin-protein E3 ligase gene (*UBE3C*) and forms a novel gene-intergenic fusion sense transcript by incorporating a terminal pseudo-exon from intergenic sequence within the DHMN1 locus. The *UBE3C* intergenic fusion (*UBE3C-IF*) transcript does not undergo nonsense-mediated decay and results in a significant reduction of wild type full length UBE3C (UBE3C-WT) protein levels in DHMN1 iPSC-derived motor neurons. An engineered transgenic *C. elegans* model expressing the UBE3C-IF transcript in GABA-ergic motor neurons shows neuronal synaptic transmission deficits. Furthermore, the transgenic animals are susceptible to heat stress which may implicate defective protein homeostasis underlying DHMN1 pathogenesis. Identification of the novel *UBE3C-IF* gene-intergenic fusion transcript in motor neurons

highlights a potential new disease mechanism underlying axonal and motor neuron degeneration. These complementary models serve as a powerful paradigm for studying the DHMN1 complex SV and an invaluable tool for defining therapeutic targets for DHMN1.

Author affiliations:

1 Northcott Neuroscience Laboratory, ANZAC Research Institute, Sydney, Australia

2 Faculty of Medicine and Health, University of Sydney, Sydney, NSW, Australia

3 Ancestry and Health Genomics Laboratory, Charles Perkins Centre, Faculty of Medicine and Health, University of Sydney, Australia

4 Centre for Infectious Diseases and Microbiology, Westmead Institute for Medical Research, Sydney, NSW, Australia

5 Monash Biomedicine Discovery Institute and Department of Anatomy and Developmental Biology, Monash University, Melbourne, VIC 3800, Australia

6 Institute for Integrated Cell-Material Sciences and Institute for Chemical Research, Kyoto University, Japan

7 Molecular Medicine Laboratory, Concord Repatriation General Hospital, Sydney, Australia

8 Brain and Nerve Research Centre, Concord Repatriation General Hospital, Sydney, Australia

9 Department of Neurology, University of Miami Miller School of Medicine, Miami, FL, 33136, USA

Correspondence to: Dr Anthony N. Cutrupi

Northcott Neuroscience Laboratory, ANZAC Research Institute

Gate 3 Hospital Rd, Concord Repatriation General Hospital, NSW, 2139, Australia

E-mail: anthony.cutrupi@sydney.edu.au

Correspondence may also be addressed to: Professor Marina L. Kennerson

E-mail: marina.kennerson@sydney.edu.au

Running title: Novel gene-intergenic fusion causes dHMN

Keywords: iPSC-derived motor neurons; distal hereditary motor neuropathy; gene-intergenic fusion; ubiquitin E3 ligase; ubiquitin proteasome system

Abbreviations: CMT = Charcot-Marie-Tooth; DEG = differentially expressed genes; dHMN = distal hereditary motor neuropathy; DHMN1 = distal hereditary neuropathy Type 1; Hi-C = chromatin conformation capture and paired-end sequencing; IPN = inherited peripheral neuropathy; iPSC = induced pluripotent stem cells; MNP = motor neuron progenitors; NbM = neural base medium; NF68 = Neurofilament 68 kDa, aka Neurofilament Light Chain; sMN = spinal motor neurons; SV = structural variation; TAD = topologically associated domains; UBE3C-IF = UBE3C intergenic fusion, UBE3C-WT = wild type UBE3C

Introduction

Distal hereditary motor neuropathies (dHMNs) are a group of clinically and genetically heterogeneous, inherited neurogenerative diseases predominantly involving the lower motor neurons of the peripheral nervous system.[1-4] Patients present with a slowly progressive, length-dependent degeneration or 'dying back' of lower motor neuron axons, causing denervation of the distal limb muscles resulting in muscle atrophy, paresis and chronic disability.[1, 2, 5] The dHMNs are comparatively rare to other forms of inherited peripheral neuropathy (IPN), with a prevalence of approximately 2 in every 100,000.[6] Currently, 29 reported causative genes and 4 disease loci have been implicated in dHMN with diverse roles in motor neuron biology and function[7] including: axonal transport[7-13], neuronal DNA/RNA processing and transcription[14-16], protein biosynthesis and post-translational processing[17-19], mitochondrial function and energy production[20, 21], cell survival, apoptosis and signalling[22-24], axonal growth and guidance[25, 26], and neuronal/axonal structural integrity and repair.[27, 28]

Despite the genetic heterogeneity, mutations in known genes account for less than 20% of known dHMN cases.[23] Furthermore, the types of mutations identified (point mutations or indels) have only been reported to affect the protein coding sequences of these genes. A large proportion of genetically undiagnosed dHMN families are likely to be caused by alternative mutational mechanisms that may occur in the non-coding region of DNA or involve structural variation (SV) rearrangements.

Structural variation is a broad term encompassing genomic rearrangements that disrupt chromosomal organization and genome architecture.[29] There are many types of SV[30-32] and range in size from 50 to millions of base pairs as defined by the Structural Variation Analysis Group.[33] SV is reported to contribute to a wide array of diseases with a genetic aetiology, including sporadic developmental syndromes to Mendelian diseases.[34] SV causing IPN is not unprecedented with over 20 cases reported to date across the spectrum of IPN, ranging from simple copy number variation to more complex chromosomal events (see Cutrupi *et al.,*[29] for a review). The most common subtype of IPN is Charcot-Marie-Tooth 1A (CMT1A) (OMIM #118220) which is caused by a 1.5 Mb tandem duplication of chromosome 17p11.2.[35] The duplication results in trisomy of the peripheral myelin protein gene *PMP22*.[36-39] Similarly, atypical genomic rearrangements occurring at the CMT1A locus including two duplications upstream of *PMP22*[40, 41] also produce CMT1A. More recently, duplication of a long-range enhancer for the *PMP22* gene[42] has been reported. Collectively, this suggests that gene dosage changes of genetic elements that control gene expression can also produce disease.

We previously reported a novel SV mutation segregating in a large Australian family (F-54) with autosomal dominant DHMN Type 1 (DHMN1, OMIM:#182960) that maps to chromosome 7q34-q36.2.[2] Using whole genome sequencing (WGS), we identified a 1.35 Mb duplication of chromosome 7q36.3 inserted in the reverse orientation into the DHMN1 locus. The inserted sequence fragment contains four protein-coding genes (*MNX1* (also known a*s HB9*)*, NOM1, RNF32, LMBR1*) and their regulatory elements, as well as the upstream regulatory elements and first 10 exons of the ubiquitin-protein E3 ligase gene (*UBE3C*). We hypothesised the DHMN1 complex insertion may produce neuropathy through aberrant expression of gene(s) by three

possible mechanisms: 1) gene dosage due to trisomy of the 1.35 Mb complex insertion or 2) position variegation of genes flanking the DHMN1 insertion breakpoints or 3) the genomic rearrangement introducing regulatory elements causing ectopic expression of flanking genes.

Investigating how the complex insertion causes the DHMN1 phenotype poses several challenges. Firstly, due to its size, current cloning techniques cannot reproduce the complex rearrangement. Secondly, the complex nature of the SV could mean that several genes show expression changes making identifying the causative gene difficult[29] and obscuring the precise mechanism by which the SV might produce the DHMN1 phenotype. Finally, studying SV in the context of peripheral nerve degeneration has been hampered by the invasive procedures needed to examine the appropriate neuronal tissues.[29, 43, 44]

To address these issues, we generated an *in vitro* human spinal motor neuron (sMN) model using induced pluripotent stem cells (iPSC) reprogrammed from DHMN1 patient fibroblasts harbouring the 1.35 Mb complex insertion. We show that the genomic rearrangement results in the production of a novel gene-intergenic fusion transcript in which the *UBE3C* partial copy is transcribed from the reverse strand and incorporates a terminal pseudo-exon from sequence within the DHMN1 locus. This *UBE3C* intergenic fusion (*UBE3C-IF*) transcript is not degraded by nonsense-mediated decay, and DHMN1 spinal motor neurons (sMN) harbouring the *UBE3C-IF* transcript show significant reduction of wild-type full length UBE3C (UBE3C-WT) protein levels. Transgenic *C. elegans* expressing the *UBE3C-IF* transcript in GABA-ergic motor neurons show neuronal synaptic transmission deficits and susceptibility to heat stress which may implicate defective protein homeostasis in DHMN1 pathogenesis.

Materials and methods

Tissue culture and cell line maintenance

HeLa cell lines were maintained in H-DMEM consisting of Dulbecco's Modified Eagle Medium (DMEM; Gibco, Life Technologies) supplemented with 10% (v/v) fetal bovine serum (FBS;

SAFC Biosciences) and 1% (v/v) Penicillin/Streptomycin (P/S; Gibco, Life Technologies). Primary fibroblasts were maintained in F-DMEM culture medium comprising DMEM, 10% (v/v) FBS, 1% (v/v) P/S and 1% (v/v) L-glutamine (Gibco, Life technologies). All cells were maintained in 5% CO_2, humidified air at 37°C. All research and cell culture procedures from patient skin biopsies was performed after informed consent and in accordance with relevant guidelines and regulations approved by the Sydney Local Health District Human Ethics Committee (HREC/17/CRGH/8).

Generation, culturing, and maintenance of iPSC lines

Reprogramming of patient and unrelated, neurologically normal control fibroblasts was performed by FUJIFILM Cellular Dynamics (Wisconsin, USA) and has been previously described.[45] In brief, DHMN1 patient fibroblasts harbouring the 1.35 Mb insertion were transfected with oriP/EBNA1(Epstein-Barr nuclear antigen-1)-based episomal vectors carrying reprogramming transgenes (*OCT4, OX2, NANOG, LIN28, c-MYC, KLF4* and *SV40LT*)[46, 47] and then seeded onto Matrigel (Corning)-coated plates in reprogramming medium. After 7 days, the medium was replaced with E8 medium (Gibco, Life Technologies) and cells were cultured for an additional 14 days. Single iPSC colonies were picked and propagated in E8 culture medium on Matrigel-coated plates. G-banded karyotyping confirmed iPSC lines were karyotypically normal (Wi-Cell; Wisconsin, USA). Pluripotency of iPSC lines was confirmed by in-house analysis of stem cell specific pluripotency genes for endogenous expression (FUJIFILM Cellular Dynamics).

iPSC lines were seeded on pre-treated (0.167 mg/mL Matrigel) 6-well plates and cultured in TeSR-E8 iPSC culture medium (StemCell Technologies) according to the manufacture's guidelines. Briefly, media was replaced daily until cells reached optimal morphology and approximately 75-80% confluence (4-5 days of culture). The cells were then passaged as aggregates (50-200 µm in diameter) using 0.5 mM EDTA and either re-seeded with split ratios (1:3 to 1:8) onto 6-well plates or cryopreserved using CryoStor CS10 Cryopreservation Reagent (Sigma). iPSC lines were maintained in 5% CO_2, humidified air at 37°C.

Motor neuron progenitor induction and sMN differentiation

We differentiated 3 control (C1, C2, and C3) and 3 DHMN1 patient (P1, P2, and P3) iPSC lines into sMN using previously described methods[48] that have been adapted and modified by our laboratory to increase culturing efficacy. The protocol is based on a combination of dual SMAD inhibition, WNT activation and NOTCH inhibition to generate highly pure and homogenous populations of expandable motor neuron progenitors (MNP) that can be differentiated into mature sMN. Detailed methods are included in Supplementary Methods. Media and supplements used are summarised in Supplementary Table S3. All adherent culturing was performed in 6-well plates (Corning-Costar) pre-treated with 0.167 mg/mL Matrigel. In brief, neural induction was initiated using a chemically defined neural base medium (NbM) supplemented with 2 μM Dorsomorphin (StemCell Technologies), 3 μM CHIR990021 (Sigma) and 2 μM SB431542 (StemCell Technologies). Cells were maintained in NbM for 6 days after which they were passaged with 1 U/mL Dispase (StemCell Technologies) and cultured in NbM supplemented with 0.1 μM Retinoic Acid (RA, Sigma), 0.5 μM Smoothened Agonist (SAG, StemCell Technologies), 2 μM Dorsomorphin, 1 μM CHIR990021 and 2 μM SB431542 for a further 6 days to generate MNP. Terminal differentiation of MNP to sMN was carried out in suspension using ultra-low attachment 6-well plates (Corning-Costar) to induce spheroid formation. Spheroids were maintained for 6 days in NbM supplemented with 0.5 μM RA and 0.1 μM SAG. Spheroids were dissociated into single cells using dissociation solution (1:1 0.25% Trypsin-EDTA & Accumax) (ThermoFisher) and maturation of sMN was carried out under adherent conditions (as described above) in NbM supplemented with 0.5 μM RA, 0.1 μM SAG, 0.1 μM Compound-E (StemCell Technologies), 2ng/mL BDNF, 2ng/mL GDNF, and 2ng/mL CNTF (Life Technologies). After 72 h, NbM was replaced and additionally supplemented with 0.02 μM SN38-P to purify cultures of proliferative progenitors and undifferentiated stem cells (Mao et al., 2018). sMN were cultured under these conditions for a further 4-9 days.

DHMN1 SV genotyping PCR assay

DNA was isolated from patient and control iPSC lines using the QuickExtract DNA Extraction Solution 1.0 (Epicenter Bio) according to the manufacturer's instructions. Multiplex PCR

amplification was performed in a 10 µL reaction containing 1X MyTaq HS Red Mix (Bioline), 10 ng DNA template, 4 pmol each of the forward and reverse primers and 8 pmol of the dual forward/reverse (multiplex) primer. Water was used as a blank control. Thermal cycling conditions and primers have been described previously [2]. Agarose gels (1.5%) were prepared in 1 X TAE buffer (Astral Scientific) with 0.01% SYBR Safe DNA gel stain (Bioline). PCR products (5 µL) were size fractionated for 30-35 min at 90 V (40 V cm^{-1}). Gels were visualised using a Safe Imager Transilluminator 2.0 (Invitrogen) and images captured using a Canon PhotoShot S5-IS digital camera with Hoya O(G) filter.

UBE3C-IF PCR and sanger sequencing

Total RNA was isolated from patient and control sMN using the RNeasy mini kit (Qiagen) according to the manufacturer's instructions. Yields were assessed using Nanodrop (Thermofisher). RNA (0.5 µg) was reverse transcribed using the iScript cDNA Synthesis kit (BioRad) in accordance with the manufacturer's instructions. PCR amplification was performed in a 10 µL reaction containing 1X MyTaq HS Red Mix, 25 ng cDNA template and 4 pmol each of the forward and reverse primers. Water was used as a blank control. Thermal cycling was performed as previously described.[2] PCR products were size fractionated by agarose gel electrophoresis and gels imaged as described above. For sequence validation, PCR products (5 µL) from samples P1 and C1, as well as 10 µM (10 µL) forward and reverse primers were provided to the Garvin Molecular Genetics Facility, Garvan Institute of Medical Research (Sydney, Australia) and sequencing was performed using BigDye Terminator Cycle Sequencing protocols. Electrophoresis primer pair: Forward – 5'-GGTACCCAAAGTCAGGAAGC-3', Reverse – 5'-CAAAGCAGCAGTTCGAGTCT-3'. Sequencing primer pair: Forward – 5' CTGCTCAACCTGGTGTGGA-3', Reverse – 5'-CAAAGCAGCAGTTCGAGTCT-3'.

UBE3C-IF qualitative RT-PCR (qRT-PCR)

Total RNA was isolated from DHMN1 and control tissues and reverse transcribed as described above. Quantitative RT-PCR was performed in 20 µL reaction volumes containing 1X TaqMan Gene Expression (Applied Biosystems) Assay, 1X TaqMan Gene Expression Mastermix

(Applied Biosystems) and 50ng cDNA template. Water was used as a negative control in all assays. Thermal cycling was performed with a StepOnePlus Real-Time PCR machine (Applied Biosystems) using the following cycling protocol: 95°C for 20 s; followed by 40 cycles of 95°C for 1 s and 60°C for 20 s. *GAPDH* was used as an internal housekeeping gene. The control sMN group was used as the reference group. TaqMan Gene Expression probes used in this study can be found in supplementary (Supplementary Table S1).

Immunohistochemistry

Cells were washed once in DPBS (Gibco), fixed in 4% (v/v) paraformaldehyde (PFA, Sigma) for 20 min at room temperature, washed once with DPBS, treated with permeablisation solution (DPBST) containing DPBS and 0.3% (v/v) Triton X-100 (Calbiochem) for 30 min at room temperature and blocked (DPBS- and 5% (w/v) bovine serum albumin (Sigma)) for 1 h at room temperature. Cells were incubated with primary antibodies overnight at 4°C. iPSC: anti-OCT4 (Cell Signalling, #2840, 1:400), anti-SOX2 (Cell Signalling, #3579, 1:400), and anti-NANOG (Cell Signalling, #4903, 1:400). MNP: anti-OLIG2 (Millipore, MABN50 1:100). sMN: anti-MNX1 (Sigma, HPA071717, 1:500), anti-68kDa NFL (Abcam, ab24520 1:1000) and anti-TUBB3 (Sigma, T2220, 1:1000). Details of all primary antibodies used in this study can be found in supplementary (Supplementary Table S2).

The cells were then incubated with Alexa Fluor (AF) secondary antibodies (Invitrogen) for 3 h at room temperature. The following secondary antibodies were used: AF goat anti-rabbit (488 or 555, 1:500), AF goat anti-mouse (488 or 555, 1:500) and AF goat anti-chicken (647, 1:500). Nuclei were counter-stained with 4,6-diamidino-2- phenylindole (DAPI, Molecular Probes, 1:5000) and mounted using Prolong Gold Antifade reagent (Invitrogen). Samples were imaged on a Leica SP8 confocal microscope and visualised using LAX software (Leica Microsystems). Images were processed using FIJI (version 2.0) for Mac OSX.

Quantification of MNX1 stained nuclei

sMN positive for MNX1 were quantified using CellProfiler (version 3.1.9). DAPI-stained nuclei were segmented via “Otsu”-based global thresholding and used to define regions of interest

(ROI) for the measurement of fluorescence of nuclei positive for MNX1. Global thresholding using the "Minimum cross entropy" method was applied to the MNX1 channel to determine a fluorescence value for positively stained nuclei. The "RelateObjects" function was used to overlay MNX1 images on their corresponding DAPI images. The percentage of positive nuclei was determined by dividing the number of $MNX1^+/DAPI^+$ co-positive nuclei by the number of $DAPI^+$ nuclei. Data were obtained for each line from three independent differentiations.

Quantification of neurite occupied area

Quantification was performed using CellProfiler (version 3.1.9). Images of TUBB3-stained axons were transformed to greyscale and segmented via "Otsu"-based global thresholding. Grey values (intensity) for each pixel were determined for each image using the "MeasureImageAreaOccupied" function. Area occupied by neurites was calculated by dividing grey pixels by the total number of pixels per image. Value is expressed as a percentage. Data were obtained for each line from three independent differentiations.

Western blot

Samples were lysed using 1 X radioimmunoprecipitation assay (RIPA) buffer (Merck) supplemented with 1X cOmplete EDTA-free protease inhibitor (Roche). Protein concentrations for each sample were analysed using Pearson's Bicinchoninic Acid (BCA) protein analysis kit (Thermofisher Scientific) in accordance with the manufacturer's protocols and measured on an EnSpire Multimode Plate Reader (Perkin Elmer). Western blot analysis was carried out using 20 µg protein lysate. Protein samples were size fractionated on 4-15% Mini PROTEAN TGX Precast Gels (BioRad) via SDS-polyacrylamide gel electrophoresis at 130 V for 1.5 h in running buffer (25 mM Tris, 192 mM glycine, SDS-PAGE 0.1 % SDS, pH 8.3) and then transferred to Immobilin-P polyvinylidene difluoride (PVDF) transfer membranes (Merck) at 70 V for 75 min in transfer buffer (BioRad). Membranes were blocked (5% w/v skim milk powder (Oxoid) in 1 X Tris-Buffer Saline (TBS)) for 1 h at room temperature followed by overnight incubation at 4°C with primary antibodies prepared in TBS-T (TBS, 0.1% v/v Tween20). Primary antibodies: MNX1 (Sigma Aldrich, HPA071717, 1:500), UBE3C (Invitrogen, #PA5-110540, 1:500).

Membranes stained with Ponceau S (Sigma Aldrich) were used to control for protein loading in these experiments. Details of all primary antibodies used in this study can be found in supplementary (Supplementary Table S2). The membranes were then incubated in anti-rabbit (Sigma Aldrich) and anti-mouse (Abcam) horseradish peroxidase (HRP) conjugated secondary antibodies (1: 5000) and signal detected using Clarity™ Western ECL Substrate solution (BioRad). Blots were visualised on a ChermiDOC XRS+ system (BioRad) and images processed using Image Lab software v5 (BioRad). Intensity of protein bands were quantified using FIJI (version 2.0).

NanoString nCounter gene expression assay

Gene expression quantification using the NanoString nCounter gene expression system (NanoString Technologies Inc) was outsourced to Westmead Medical Research Institute (Sydney, Australia). A 72-target custom panel (63 gene targets, 2 MNP markers, 2 sMN markers and 5 housekeeping genes) was designed to include all candidate genes 3 Mb on either side of the DHMN1 insertion breakpoints and genes within the DHMN1 complex insertion. The target-specific oligonucleotides were designed by NanoString Technologies (Washington, USA) and synthesized by Integrated DNA Technologies (Iowa, USA). Total RNA was isolated (sMN: D27, n = 5; MNP: D18, n = 4; iPSC: D0, n = 3) from patient and control tissues using the RNeasy mini kit (Qiagen) and diluted to a concentration of 10ng/µL. RNA yield was assessed using Nanodrop and integrity determined using TapeStation (Agilent). Gene count normalisation and expression analysis was performed using the nSolver Analysis Software package (version 4.0) (NanoString Technologies). The normalisation of the gene count data was performed using the recommended parameters described in the nSolver User Manual (Version 4.0). Background correction and low count gene filtering was performed via the in-built background thresholding function using the default threshold value of 20 counts. Details regarding assay preparation and execution can be found in Supplementary Methods.

RNA sequencing data processing and analysis

sMN total RNA from DHMN1 patient (n = 3) and controls (n = 3) was isolated using the RNeasy mini kit (Qiagen) according to the manufacturer's instructions. Yields were assessed using Nanodrop and RNA integrity determined using 2100 Bioanalyser (Agilent). Library preparation and mRNA sequencing was outsourced to Macrogen (Seoul, South Korea). In brief, polyA species selection and library preparation was carried out using the TruSeq Stranded mRNA LT Sample Prep Kit (Illumina). Paired-end (151 bp) sequencing was performed on an Illumina NovaSeq 6000 sequencer. Quality control (QC) was performed by Macrogen using FastQC[49] with adapter sequences and low-quality bases trimmed using Trimmomatic.[50] Trimmed reads were mapped to GRCh38/hg38 using the splice-aware aligner HISAT2.[51] Quantification of gene and transcript level abundances and the identification of novel and alternative splicing transcripts was carried out using StringTie.[52] Identification of fusion gene products was performed using Defuse[53], FusionCatcher[54] and Arriba[55] programs. Gene-level differential gene expression analysis was performed using DESeq2.[56] Differentially expressed genes (DEGs) were determined by $\log_2$FC ≥ abs 1.0 and a false discovery rate (FDR) adjusted p-value threshold of 0.05. Gene Ontology and enrichment analysis to determine biologically relevant pathways was performed using the Database for Annotation, Visualisation and Integrated Discovery (DAVID)[57, 58] and StringDB.[59-62]

Chromatin Conformation Capture and paired-end sequencing (Hi-C)

Hi-C library preparation and sequencing was outsourced to Dovetail Genomics (Santa Cruz, California). A library was prepared from motor neurons from a DHMN1 patient (n = 1) and sequenced (150 bp paired-end) using three lanes of an Illumina HiSeq X to generate ~300 Gb per lane. This would generate 900 million reads for a 10 kb binsize resolution (non-overlapping sequences) in which more than 80% of all possible bins would have 1000 or more reads (contacts). Mapping of Hi-C libraries and normalisation was performed using HiC-Pro pipeline.[63] Sequencing reads were mapped to the hg38 reference genome and a hg38 custom-built DHMN1 chromosome 7 with the 1.35 Mb insertion, using the Bowtie aligner and assessed for artefact

levels using the human genome and restriction enzyme cut sites. BAM files mapping the paired-end tags (PETs) were filtered to eliminate invalid PETs. Visualisation and analysis of genome-wide contact matrices from mapped and normalised Hi-C PET data was performed using HiGlass.[64] Visualisation, identification, and prediction of TADs was performed using TADtool.[65]

Construct generation and cloning

Q5 Site Directed Mutagenesis (SDM; New England Biolabs) was used to generate the pCMV6-Entry-UBE3C-IF plasmid. This was conducted in two stages. (1) the commercial template plasmid pCMV6-Entry-UBE3C (Origene #:RC215110) was amplified using primers which partially inserted the intergenic fusion sequence and simultaneously deleted the UBE3C coding sequence downstream of p.Val443 (F: 5'-aagaggatcattACGCGTACGCGGCCGCTC-3'; 5'-gaattgcttcctGACTTTGGGTACCATCATGCGGTGCT

G-3'). (2) Q5 site-directed mutagenesis was again conducted on the plasmid generated in step 1 to introduce the remaining intergenic fusion sequence alone (5'-caccagaataaaACGCGTACGCGGCCGCTC-3'; 5'-acatcttgttaaAATGATCCTCTTGAATTGCTTCCTGACTTTGGGTACC-3'). The UBE3C-IF coding sequence was then amplified from pCMV6-Entry-UBE3C-IF using primers with flanking *Xba* I and *Not* I restriction sites (F: 5'-aaaaaatctagaATGTTCAGCTTCGAAGGC-3'; R: 5'-aaaaaagcggccgcCTATTTATTCTGGTGACATCTTGTTAAAATG-3'). This amplicon was then inserted into the pPD157.60 vector containing the unc-25 promoter sequences using the *Xba* I and *Not* I restriction sites to form the expression plasmid pPD157.60-UBE3C-IF. The ISOLATE II Plasmid mini (Bioline) was used for the purification of expression plasmids according to manufacturer's instructions.

Overexpression of *UBE3C-IF* in HeLa cells.

HeLa cells (1 x 10^6) were seeded onto 6-well plates and maintained in H-DMEM at 37°C and 5 % CO_2 for 24 h. Cells were transfected with either empty vector or the expression plasmid pCMV6-Entry-UBE3C-IF (0.5 µg, 1.0 µg, 2.0 µg) via lipofection using Lipofectamine 3000

(Invitrogen) as per the manufactures protocol. Cultures were harvested 48 h post-transfection for western blot analysis as described above.

C. elegans methods

Transgenic *C. elegans* strains were generated by microinjecting a cocktail of the expression plasmid containing UBE3C-IF or the empty vector at a concentration of 50 ng/µL and plasmid PCFJ90, a co-injection marker, at a final concentration of 5 ng/µL. Three independent transgenic lines were generated, and the experiments were carried out using the line that stably inherited the human transgene.

Strain information and additional *C. elegans* methods can be found in Supplementary Methods.

Statistical analysis

For the statistical analysis of NanoString gene expression data, differential expression and significance was determined using the nSolver software (v4.0). The Differential Expression Call (DE Call) test function was used to predict differential expression for MNP and iPSC data where biological replication was limited. Statistical analysis of qRT-PCR data was carried out using the StepOne Plus Software (v2.1) (Applied Biosystems). For the immunofluorescence quantification data, significance was determined using a two-tailed Student's t-test. Statistical analysis of RNA-seq was performed in RStudio utilising a custom DESeq2 pipeline. The significance of differential gene expression was determined using the Wald Test assuming a $\log_2$FC threshold of abs 1.0 and a false discovery rate (FDR) adjusted *p*-value threshold of 0.05. The statistical analysis and significance of western blot experiments was determined using either a two-tailed Student's t-test (sMN experiments) or a one-way ANOVA followed by Dunnet's multiple comparison test (HeLa experiments). The following statistical thresholds were applied throughout the study: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$.

Data availability

All relevant data are included within this manuscript and supplementary material files. Raw RNA-seq data will be available upon reasonable request.

Results

Characterisation of DHMN1 patient-derived iPSC lines

Skin fibroblasts from three DHMN1 patients (2nd generation siblings, 2 male and 1 female; Supplementary Fig. S12) were reprogrammed by FUJIFILM Cellular Dynamics International (CDI) using non-integrative episomal plasmids and company in-house protocols. To fully validate the iPSC lines, in-house pluripotency analysis was performed, and molecular analysis confirmed the DHMN1 complex insertion was retained after the reprogramming process. These results were confirmed in the iPSC lines from three different patients and representative data is presented (Fig. 1). Normal karyotyping was observed (Fig. 1A) and genotyping DNA extracted from the patient-derived iPSC lines confirmed the presence of the SV mutation post-reprogramming (Fig. 1B). Pluripotency of the DHMN1 iPSC lines was initially confirmed by CDI as part of the service. Subsequent immunofluorescence experiments in our laboratory also confirmed the DHMN1 iPSC lines were positive for the pluripotency markers OCT4A, SOX2 and NANOG (Fig. 1C).

Generation of MNP and sMN by differentiation of DHMN1 patient-derived iPSC lines

To develop a motor neuron model for DHMN1, 3 patient and 3 control iPSC lines were differentiated into sMN from highly expandable populations of *OLIG2*$^{+}$ positive motor neuron progenitors (MNP) as initially reported by Du *et al.*[48] Generating sMN using MNP provided added practicality as they can be cryopreserved, and undergo thawing for differentiation, thereby reducing the overall time to produce sMN. In addition, culturing MNP produces highly pure populations of sMN, and avoids many of the pitfalls associated with other differentiation methods

including poor yields and low efficiency[48]. We modified the protocol to include a purification step whereby adherent single cell sMN were cultured in the presence of 0.02 μM SN38-P for 96 h (Fig. 2A). The SN38-P acts to purify the culture of any remaining proliferative, undifferentiated iPSC and MNP[66]. We have previously shown SN38-P treatment enriches cultures for mature sMN without the need for additional sorting protocols[67]. Immunohistochemistry confirmed that iPSC differentiated into ventralised neural stem cells expressing OLIG2 (MNP, Supplementary Fig. S1A). Additionally, the NanoString data demonstrated that *Nestin* (*NES),* also a marker of neural stem cells[68] was expressed at higher levels in differentiated cells compared to iPSC (Supplementary Fig. S1B). Together these data indicate successful differentiation of iPSC into MNP.

MNP that were cultured in suspension for 6 days formed 3D aggregates (neurospheres) of MNP actively differentiating into sMN. Following the disaggregation of these neurospheres, the sMN were matured under adherent 2D culturing conditions for an additional 7-12 days. The mature sMN demonstrated robust expression of the sMN marker MNX1, the neuronal cytoskeletal marker βIII-Tubulin (TUBB3) and the axonal marker NF68 (Fig. 2B). Similarly, the expression of the sMN markers *MNX1, ISL1* and *ChAT*, determined by the NanoString gene read counts, were higher in sMN when compared to the iPSC (Fig. 2C). Quantification of the $MNX1^{+}$ nuclei revealed comparable numbers of sMN generated for the different patients and controls [Fig. 2D(i and ii)]. Similarly, quantification of TUBB3 did not reveal any differences in the 2D neuronal networks between the patients and controls [Fig. 2E(i and ii)].

Genes within and flanking DHMN1 insertion breakpoints are dysregulated

Initial linkage studies mapped the DHMN1 locus to a 12.3 Mb region at chromosome 7q34-q36.3[69] containing 112 candidate genes. Identification of the SV mutation enabled prioritisation of possible candidate genes by 1) copy number, 2) disruption of the normal regulatory environment or, 3) introduction of new regulatory elements driving ectopic expression of nearby genes (see Kleinjan[70-72] for review). To analyse the local gene expression signatures associated

with the DHMN1 complex insertion, 63 candidate transcripts were selected for NanoString analysis in iPSC, MNP and sMN tissues (Supplementary Table S3). We hypothesised the genomic rearrangement causing the DHMN1 phenotype could affect genes lying within the DHMN1 complex insertion as well as genes flanking the insertion breakpoints – given that cis-regulatory control of gene expression has been reported to occur over distances of up to 1.45 Mb. [73] The custom assay was therefore designed to select genes localising within the DHMN1 complex insertion or in the genomic regions (3 Mb) flanking the DHMN1 insertion breakpoints (Fig. 3A). Genes not contained within this conservative 6 Mb genomic interval around the insertion breakpoints were excluded from analysis.

Analysis of the expression signatures of the 63 genes showed that 10 targets and 14 targets were differentially expressed in MNP and iPSC respectively following background correction and filtering of low-count genes. (DE Call test; Supplementary Table S4). For the sMN, 4 of the 63 gene targets tested showed differential expression (DE) ($p < 0.05$) following background correction and filtering of low-count genes (Fig. 3D, Supplementary Table S5). For the four targets showing differential expression, two were within the DHMN1 insertion sequence (*MNX1, UBE3C*), one was within the DHMN1 linkage locus (*TMEM176B*) and one target was outside the DHMN1 locus and proximal to the insertion sequence originating from chromosome 7q36.3 (*SHH*) (Fig. 3B, Supplementary Table S3). The assessment of target genes in the MNP and iPSC displayed minimal overlap with the expression changes in sMN with *UBE3C* being the only differentially expressed gene across the three tissue types (Fig. 3C).

The DHMN1 complex insertion alters 3D genome organisation

The structure and architecture of the 3D genome plays a crucial role in the spatiotemporal regulation of gene expression in tissues and cells. To assess how the 1.35 Mb complex insertion could change the genomic architecture and thereby impact downstream gene regulation at the DHMN1 locus, Hi-C analysis on patient ($n = 1$) sMN tissue was performed. Data was analysed using a custom, in-house, targeted Hi-C analysis pipeline. Topological associated domains (TAD) calling algorithms identified an altered TAD profile in patient sMN at the DHMN1 locus when

compared to the control sMN profile (Fig. 4A). Visualisation and assessment of genome contacts across the DHMN1 locus also predicted aberrant genomic interactions occurring within patient sMN as well as the formation of a neo-TAD (Fig. 4B). Overlaying the contact matrices with the genomic map of the DHMN1 locus showed predicted aberrant interactions with the neo-TAD overlapping with the significantly dysregulated expression of the candidate genes *MNX1* and *UBE3C* (Fig. 4C). The observed congruence between the gene dysregulation and Hi-C data further prioritised these genes as potential causative DHMN1 candidates.

RNA-seq excludes global gene dysregulation as a potential pathogenic mechanism and does not reveal biologically relevant pathways related to prioritised candidate genes

RNA-seq is an invaluable research tool enabling comprehensive quantitative analysis of the global transcriptome.[74, 75] One of the aims of this study was to compare global transcriptome profiles of DHMN1 and control sMN to elucidate genes, pathways and mechanisms that may be impacted by dysregulation of *UBE3C* or *MNX1*. To address this aim, RNA-seq was performed on polyA-selected mRNA transcripts extracted from patient (n = 3) and healthy control (n= 3) sMN cultures (day 12). Quality control, trimming and mapping statistics for each sample are summarised in supplementary data (Supplementary Table S6). An in-depth summary of data diagnostics is available in supplementary data (Supplementary Fig. S2 and S3). Following low count filtering, a total of 19,929 features were identified as expressed in all samples. Differential gene expression analysis of these expressed features revealed only 22 differentially expressed genes (DEGs) in DHMN1 sMN compared to healthy controls (Supplementary Table S7). The 22 genes did not localise to the DHMN1 locus and only one gene (*UNCX*) mapped to chromosome 7. The expression signatures of *UBE3C* and *MNX1* also showed the upregulation (Fig. S4 A and B) however, the changes did not reach statistical significance as observed in NanoString (Supplementary Fig. S4 A' and B'). Gene Ontology (GO) analysis and functional annotation clustering was performed using StringDB and DAVID algorithms to determine the biological relevance of the 22 DEGs (Supplementary Fig. S11A-D). After combining the output from both pipelines, a total of 30 significantly enriched terms were identified. The highest proportion of

terms (13 out of 30) corresponded to tissue/organ development and morphogenesis (Supplementary Fig. S11E). Of these, 50% (6 out of 12) involved the nervous system (Supplementary Fig. S11F). No GO terms relating to neuronal compartments (axon, soma, dendrites), motor neurons, known pathways involved in CMT/IPN, motor neuron degeneration or other potentially biologically relevant pathways were identified across either analysis.

RNA-seq reveals a novel gene-intergenic fusion involving the *UBE3C* partial gene duplication

The DHMN1 SV involves the duplication and insertion of a 1.35 Mb region of chromosome 7q36.3 into the disease locus (chromosome 7q36.2) in an inverted orientation resulting in partial duplication (first 10 exons) of the *UBE3C* gene. Given that we have observed upregulation of *UBE3C* in DHMN1 sMN, we hypothesised that this was due to the production of a novel fusion transcript involving the partial *UBE3C*. The RNA-seq dataset was interrogated for segregating, predicted, novel gene fusions involving the partial *UBE3C* with a nearby gene or non-coding genomic sequence on the negative strand. A summary of chromosome 7 gene fusions can be found in supplementary material. The Arriba and Defuse algorithms detected a novel fusion involving the partial *UBE3C* sequence that was present in the three DHMN1 patient sMN samples [Fig. 5A(i) and Supplementary Table S8]. *In silico* splice site prediction analysis (NNSPLICE 0.99[76]) of the intergenic sequence downstream of the *UBE3C* partial gene strongly (score of 1.0 out of 1.0) indicated a canonical acceptor site ~ 7 kb downstream of the final splice donor site of the partially duplicated *UBE3C* gene. The genomic sequence downstream of this splice acceptor site corresponded with the sequence of the novel terminal pseudo-exon (tPE) detected by Arriba and localised to intergenic DNA from the DHMN1 locus (Fig. 5B). RT-PCR and Sanger sequencing subsequently confirmed the presence of the *novel gene-intergenic fusion* sense transcript (*UBE3C-IF*) [Fig. 5A(ii) and B]. *UBE3C-IF* lacks 13 of 23 exons (aa 445-1083) of canonical *UBE3C-WT.* These exons comprises the complete HECT domain – the main catalytic domain of the protein which is responsible for receiving ubiquitin from cognate E2 enzymes UBE2D1 and ligating it to target substrates bound to the proteasome[77-79] (Supplementary Fig. S4). To determine if a truncated *UBE3C-IF* was transcribed in addition to the wild-type full-length *UBE3C* (*UBE3C-WT*), quantitative RT-PCR using TaqMan probes was

performed on sMN tissue from patients and controls. Significant differences in *UBE3C* expression were observed between patients and controls for a probe localising within the duplicated exons ($p < 0.01$). In contrast, no changes in expression between patients and controls were detected with a probe localising to the non-duplicated exons (Fig. 5D).

DHMN1 sMN protein lysates show a reduction in full-length UBE3C.

BLAST Protein[80-83] confirmed the peptide sequence spanning the exon 10/pseudo-exon junction predicted by Arriba aligned to wild-type, canonical UBE3C-WT (NP_055486.2) with 100% sequence identity up to the pseudo-exon junction in all patients (Supplementary Fig. S5A-C). The full peptide sequence of UBE3C-IF is predicted to be 460 amino acids (aa) corresponding to a molecular weight of approximately 50.6 kDa (based on an average molecular weight of 110 Da/aa; Supplementary Fig. S5D). To investigate if *UBE3C-IF* is translated into protein, western blot analysis was performed on sMN protein lysates blotted with a UBE3C-WT antibody raised against the N-terminal of the protein (aa 1-270). Chemiluminescent visualisation revealed 2 additional bands below UBE3C-WT (123 kDa) at ~ 55 kDa and ~ 50 kDa (Fig. 6A). These bands may correspond to shorter protein isoforms of UBE3C that are annotated in Ensembl (Supplementary Fig. S6) and overlap with the predicted molecular weight of UBE3C-IF. Densitometric quantification of the two bands was performed to determine if the UBE3C-IF co-migrated with the shorter UBE3C isoform however, no difference was observed between the DHMN1 and control sMN tissues (Supplementary Fig. S7B-B'). Although it was not possible to resolve a specific band corresponding to UBE3C-IF, densitometric quantification of the full length UBE3C-WT protein showed levels that were significantly reduced in DHMN1 sMN tissue compared to controls ($p < 0.0001$; Fig. 6B). Western blot analysis of MNX1 (which was also prioritised as a candidate with differential gene expression) showed no difference between DHMN1 and control sMN tissue (Supplementary Fig. S7A-A') and was therefore excluded as a pathogenic candidate gene.

Overexpression of *UBE3C-IF* in HeLa cells recapitulates reduction of full-length UBE3C-WT observed in patient sMN

To determine if the presence of UBE3C-IF results in the reduction of UBE3C-WT protein levels, HeLa cells were transfected with 0.5-2 µg of construct containing *UBE3C-IF* (or empty vector; EV) and endogenous UBE3C was analysed from protein lysates harvested for western blot analysis using the UBE3C-WT antibody. Chemiluminescent visualisation of the blot showed cells transfected with the UBE3C-IF construct displayed an additional band corresponding to ~ 50 kDa. This demonstrated that UBE3C-IF can be detected with an antibody raised against the N-terminal of the UBE3C-WT. Furthermore, this indicates that UBE3C-IF could co-migrate with a shorter, tissue-specific 50 kDa isoform of UBE3C observed in the sMN lysates (Fig. 6C). Whilst there was not a trend of decrease in endogenous UBE3C-WT protein levels with increasing amounts of UBE3C-IF construct, densitometric quantification showed that the cells transfected with 2.0 µg of UBE3C-IF had significantly reduced levels of UBE3C-WT (Fig. 6D) when compared to cells transfected with empty vector suggesting perturbed autoregulation at higher concentrations of UBE3C-IF.

Overexpression of *UBE3C-IF* in *C. elegans* affects synaptic transmission, causes susceptibility to heat stress but does not affect neuronal morphology or locomotion behaviour

Aldicarb (an acetylcholine esterase inhibitor) and levamisole (an acetylcholine receptor antagonist) are drugs commonly used for screening *C. elegans* mutants defective in synaptic transmission.[84] Using this method, we have previously successfully identified synaptic transmission deficits in a CMT model of *C. elegans.*[85] *UBE3C-IF* animals displayed resistance to aldicarb induced paralysis when compared to control animals (Fig. 7A). In the presence of 1mM aldicarb, the time taken for 50% of animals to paralyse was approximately 70 min for *oxIs12* and *oxIs12;EmptyVector* when compared to 90 min for *oxIs12;UBE3C-IF* animals. Transgenic animals were hypersensitive to levamisole when compared to *oxIs12* animals (Fig. 7B). However, there was no significant difference between animals carrying *UBE3C-IF* plasmid and

animals carrying the empty vector backbone. This suggested the hypersensitivity to levamisole may be due to the plasmid copy number and not the *UBE3C-IF* transcript. Our results show that the post-synaptic component (muscle) is not involved in the aldicarb resistant phenotype observed in *UBE3C-IF* animals thus implicating pre-synaptic compartment (axonal) deficits. This reflects the axonal presentation observed in DHMN1 patients. *oxIs12* animals carrying the transgene *unc-47::GFP,* allow visualisation of motor neurons in live *C. elegans*. Live imaging of age-synchronised day1 old *UBE3C-IF* and control animals showed intact axon and cell bodies with no signs of neurodegeneration (Fig. 7C and D). Correspondingly, the thrashing assay used to identify locomotion deficits associated with *UBEC-IF* overexpression showed no significant changes in locomotion between transgenic and control animals (Fig. 7E). Loss of *Hul5*, the yeast ortholog of *UBE3C*, results in reduced recovery rate following heat shock.[86] To determine if the truncated *UBE3C-IF* construct could negatively impact a heat-induced protein unfolding response (the heat shock assay), *UBE3C-IF* and control animals were exposed to heat stress (8 h at 35ºC). *UBE3C-IF* animals showed reduced survival rate (~50%) when compared to controls animals [*oxIs12* (~82%) and *oxIs12;EmptyVector* (~77%)] (Fig. 7F) suggesting reduced fitness in response to protein unfolding.

Discussion

Distal hereditary motor neuropathies have benefited from next generation sequencing gene discovery and genetic diagnosis; however, a large proportion of cases (up to 50%) remain genetically unsolved. In such instances, there is a strong precedent for looking beyond the exome and querying the remainder of the genome for SV mutations. Several cases of SV – both typical and atypical (see Cutrupi *et al.*[29] and Boyling *et al.*[87] for a review) – have been reported to cause inherited peripheral neuropathies (IPN). Therefore, in the cases where protein-coding mutations cannot be identified, SV represents a possible mutational mechanism that may account for some of the genetically undiagnosed dHMN and other IPN cases[29].

This study highlights the utility of this approach and represents the first available human neuronal model to examine the impact of the DHMN1 complex SV mutation. We have shown the 1.35 Mb

complex insertion contains a partial duplication of the *UBE3C* gene which forms a novel gene-intergenic fusion (*UBE3C-IF*) by incorporating a pseudo-exon from intergenic sequence from within the DHMN1 locus. The *UBE3C-IF* forms a stable transcript that is not degraded by nonsense mediated decay and can readily be detected in mRNA from patient derived sMN. Chromosomal rearrangements such as somatic SVs are a common driver of fusion gene formation. Several publications have described this mechanism in cancer where gene fusions have been extensively studied.[88-95] Typically, fusion genes involve two or more coding sequences (gene-gene fusions; e.g. *BCR-ABL*[88, 89] and *TMPRSS-ERG*[90]) resulting in chimeric mRNA transcripts that can form oncogenic "neo-antigens".[96] Fusions involving intergenic/non-coding sequences that result in cryptic (pseudo) exon formation are by comparison rare – only being reported by a few studies.[96-99] Interestingly, the majority of genomic breakpoints in fusion genes are intergenic or intronic and are therefore not typically present in mRNA or protein-coding sequences[93, 94, 100] making detection of gene-intergenic fusions difficult. This may account for the predominance of classical gene-gene fusions over gene-intergenic fusions in the literature to date. Thorough investigation of genetically unsolved and/or suspected SV cases using a combination of WGS and RNA-seq may therefore help to improve the diagnostic rate of gene-intergenic fusions in genetic disease and expand the spectrum of mutations causing inherited motor neuron disorders.

Studying the effects of these SV mutations is challenging as the size and complexity of the DNA rearrangement can increase the number of potential candidate causative genes. This makes interpretation of data from gene expression studies difficult particularly when alternative tissues are used in place of disease relevant tissue. In this study, we addressed this issue by generating a disease-relevant, tissue-specific model and applying rigorous, experimental-driven filtering to refine the list of probable disease-causing candidate genes. Our original hypothesis was that DHMN1 neuropathy is caused by transcriptional dysregulation. Using a positional cloning approach, we showed that using a targeted gene expression panel of 63 genes combined with Hi-C analysis was able to eliminate all but two high-priority candidate genes, *MNX1* and *UBE3C*. Western blot analysis showed no change in MNX1 and was therefore excluded as a causative candidate gene. The observed transcriptional upregulation of *MNX1* is likely a bystander effect of the SV mutation. MNX1 has been reported to autoregulate[101] and this may explain why a

corresponding difference was not observed at the protein level. In contrast, although *UBE3C* appeared to be upregulated in patient tissues, we showed that this was an artefact based on the primer/probe design detecting both *UBE3C-WT* and *UBE3C-IF* transcripts (Fig. 5D). The lack of meaningful gene dysregulation observed at the local transcriptional level was also observed globally with only 22 of 19,929 genes detected as differentially expressed. The absence of any discernible and meaningful aberrant transcriptional regulation provides further evidence against our original hypothesis and lends strong support for *UBE3C-IF* as the pathogenic candidate. This is further strengthened by the *UBE3C-IF* pseudo-exon using DNA sequences from within the previously mapped DHMN1 linkage region.[69]

The ubiquitin protein E3 ligases are a superfamily of over 600 genes responsible for the transfer of ubiquitin (Ub) to substrate proteins marked for degradation by the ubiquitin proteasome system (UPS[102, 103]). UBE3C is a member of the HECT (homologous to E6-AP carboxyl terminus) class of E3 ligases which comprises 28 members.[104] The activity of HECT E3 ligases is tightly regulated with respect to Ub linkage specificity and chain type, interaction with cognate E2 conjugating enzymes and substrate recognition.[105] E3 ligases (in particular HECT E3 ligases) are reported to have important functions in neuronal development and migration, synaptic transmission[102] and have been implicated in a variety of neurological and neurodevelopmental diseases including inherited peripheral neuropathy (see Ambrozkiewicz *et al.*,[102, 106] George *et al.,*[103] and Lescouzeres *et al.*[107] for review). We show that *UBE3C* is expressed in iPSC, MNP and sMN. This is not surprising given that the vast majority of E3 ligases (including *UBE3C*) are ubiquitously expressed[108]. Importantly however, these three tissues correspond to three distinct developmental stages suggesting that the *UBE3C* gene may be important in neuronal development. GO analysis of the 22 DEGs identified several enriched pathways related to nervous system development. However, given the lack of meaningful differential expression detected at both local and global transcriptional levels, it cannot be determined if this is related to *UBE3C* or an artefact of inducing neuronal differentiation. Furthermore, given that we show no observable change in *UBE3C-WT* mRNA expression, it may be unlikely that the term enrichment in DEGs is related to *UBE3C*.

E3 ligases are primary determinants of the substrate specificity of the UPS, however many of them are poorly characterised in this regard.[109] The specificity of E3 ligase-substrate interactions contributes to the large diversity in the spatiotemporal control of ubiquitination.[109, 110] In order to preserve this, E3 ligases are tightly regulated by post-transcriptional modifications such as ubiquitination[111] or by various homotypic and/or heterotypic interactions (reviewed in Balaji & Hoppe[112]). In this study, we observed that levels of full-length wild-type UBE3C are reduced in patient sMN harbouring the UBE3C-IF. Furthermore, we show overexpression of UBE3C-IF recapitulated the reduced endogenous UBE3C-WT levels in HeLa cells which was observed in patient sMN. Taken together, these data suggest that the presence of UBE3C-IF protein at higher concentrations affects UBE3C autoregulation resulting in reduced levels of endogenous UBE3C-WT protein. This is not unprecedented given that autoregulation has been observed among E3 ligase family members. For example, upon degradation of their respective substrates, the E3 ligases MDM2 and SIAH1 have been shown to self-regulate.[113, 114] This is triggered by increased cellular levels of these proteins following target substrate degradation. MDM2 and SIAH1 autoregulation begins with homodimerization proceeded by autoubiquitination resulting in proteasomal degradation.[113, 114] Recent evidence suggests that UBE3C can self-regulate through autoubiquitination[79] and it is therefore possible that the observed reduction in UBE3C-WT in patient sMN is due to autoregulation triggered by the presence of UBE3C-IF. This might also explain the absence of the UBE3C-IF protein signature in sMN western blot. Aside from the aforementioned co-migration of UBE3C-IF with a shorter tissue-specific isoform obscuring visualisation, it is possible that due to the rapid nature of UPS-mediated protein degradation, UBE3C-IF and UBE3C-WT are degraded simultaneously and too rapidly to observe UBE3C-IF without the use of proteasome inhibitors. This will be useful to investigate further in follow-up studies.

The maintenance of protein homeostasis is fundamental to the proper functioning of cells. To achieve this, cells must balance protein synthesis with protein degradation.[115] The transfer of Ub to targeted substrates is fundamental to this process. Mutation or dysfunction of genes within the system results in the accumulation of ubiquitinated inclusions which have been associated with several neurodegenerative diseases[116, 117] including amyotrophic lateral sclerosis (ALS)[118-121]. Furthermore, perturbed protein homeostasis is implicated in ALS. UBE3C associates with the

proteasome and assembles Lys-29 and Lys-48 linked polyUb chains.[77, 78, 115, 122] Knockdown experiments in yeast have shown that the absence of *Hul5p* (orthologous to human *UBE3C*) results in a reduction in the processivity of the proteasome making it less able to degrade stable proteins leaving behind partially degraded remnants.[123, 124] Recent studies have expanded on these findings and have shown the UBE3C protects against accumulation of partially degraded substrates arising due to partial proteolysis.[115]

To investigate the role of the *UBE3C* gene-intergenic fusion on neuron morphology and nervous system function, we generated transgenic *C. elegans* that either overexpressed *UBE3C-IF* transcript or the empty vector backbone in animals that carry the transgene *oxIs12*[*unc-47::GFP*]. Defects associated with the neuromuscular junction (NMJ) are reported for peripheral neuropathies.[125] UBE3C-IF animals displayed a resistance phenotype in the presence of 1 mM aldicarb (Fig. 7A). Aldicarb resistance is observed in mutants with less acetylcholine in the synaptic cleft, suggesting potential deficits in neurotransmitter release in UBE3C-IF animals. In *C. elegans*, overexpression of UBE3C-IF did not affect neuron morphology and animal locomotion (Fig. 7C-E). Upon exposure to heat stress, UBE3C-IF animals showed reduced survival suggesting a defective response to temperature-induced protein unfolding stress (Fig 8F). Interestingly, knockdown of Hulp5 (the yeast ortholog of *UBE3C*) was found to compromise the recovery of yeast after heat shock.[86] Similarly, knockout of UBE3C in human cells renders them more susceptible to heat shock following treatment with a heat shock protein 90 (hsp90) inhibitor.[115] Taken together, our DHMN1 in vivo model shows that the UBE3C-IF causes synaptic transmission deficits and defective stress response resulting in reduced animal survival.

Conclusions

This study has presented the first patient derived sMN model for DHMN1. Moreover, this study is the first report of a gene-intergenic fusion (*UBE3C-IF*) causing a motor neuron disorder and has therefore expanded the spectrum of mutations known to cause motor neuron diseases. The presence of UBE3C-IF in patient tissues results in a dominant-negative reduction of WT-UBE3C. This study has shed light on a new disease mechanism in motor neuron diseases and underscores

the importance of looking beyond the exome when studying inherited diseases with unknown genetic aetiology. We therefore propose that dominant loss-of-function mutations resulting in either 1) knockout of one allele of UBE3C or 2) mutations in the HECT domain that ablate E3 ligase function should be investigated in unsolved motor neuropathy families. Furthermore, we show the potential of *UBE3C-IF* using an *in vivo* model to elucidate pathogenic processes through highlighting synaptic dysfunction and compromised proteasome processivity as plausible disease processes underlying DHMN1 pathogenesis. Further follow-up is needed to define the precise mechanism of action by which UBE3C-IF and UBE3C-WT interact and to define the targets of UBE3C-WT. This work demonstrates the novelty of gene-intergenic fusions as an important and understudied mechanism for motor neuron disease that will provide essential new knowledge and inform avenues for treatment and therapies.

Acknowledgements

The authors thank the CMT families who have participated in the research and volunteered skin biopsies for the generation of iPSC under informed consent. We also acknowledge and thank Dr Renata Maciel for her invaluable technical assistance with iPSC and sMN work presented in this study. Some *C. elegans* strains were obtained from the Caenorhabditis Genetics Center (CGC) funded by NIH Office of Research Infrastructure programs (P40 OD010440). We would like to thank Paul Davis and Prof. Tim Schedl for assistance with obtaining WormBase strain accession numbers. Strains generated in this study can be obtained by contacting the corresponding author.

Funding

This work was supported by the National Health and Medical Research Council Project Grant (APP1046680) awarded to M.L.K and G.A.N, NHMRC Ideas Grant (APP1186867) awarded to M.L.K, G.P-S and R.K.N and a CMT Australia Grant awarded to M.L.K., G.P-S, M.A.S, R.CY.L and G.A.N. A postgraduate scholarship from Sydney Medical School (Pamela Jeanne Elizabeth Churm Postgraduate Research Scholarship) supported A.N.C. The Charcot-Marie-Tooth Association and NIH/NCATS KL2 Career Development Award (CTSI-KL2-FY19-0X) supported M.A.S.

Competing interests

The authors report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

1. Drew, A.P., Blair, I.P. and Nicholson, G.A. Molecular genetics and mechanisms of disease in distal hereditary motor neuropathies: insights directing future genetic studies. *Curr Mol Med*, 2011. **11**(8): p. 650-65.
2. Drew, A.P., Cutrupi, A.N., Brewer, M.H., Nicholson, G.A. and Kennerson, M.L. A 1.35 Mb DNA fragment is inserted into the DHMN1 locus on chromosome 7q34-q36.2. *Hum Genet*, 2016. **135**(11): p. 1269-1278.
3. Landrieu, P. and Baets, J. Chapter 49 - Early onset (childhood) monogenic neuropathies, in, S. Gérard & K. Christian, Editors*Handbook of Clinical Neurology*. 2013, Elsevier. p. 863-891.
4. Dierick, I., Baets, J., Irobi, J., Jacobs, A., De Vriendt, E., Deconinck, T., et al. Relative contribution of mutations in genes for autosomal dominant distal hereditary motor neuropathies: a genotype-phenotype correlation study. *Brain*, 2008. **131**(Pt 5): p. 1217-27.
5. Rossor, A.M., Kalmar, B., Greensmith, L. and Reilly, M.M. The distal hereditary motor neuropathies. *J Neurol Neurosurg Psychiatry*, 2012. **83**(1): p. 6-14.
6. Bansagi, B., Griffin, H., Whittaker, R.G., Antoniadi, T., Evangelista, T., Miller, J., et al. Genetic heterogeneity of motor neuropathies. *Neurology*, 2017. **88**(13): p. 1226-1234.

7. Kandel, E.R. Nerve cells and behaviour, in, E.R. Kandel, J.H. Schwartz, & T.M. Jessel, Editors*Principles of Neural Science*. 2000, McGraw Hill. p. 19-35.

8. Evgrafov, O.V., Mersiyanova, I., Irobi, J., Van Den Bosch, L., Dierick, I., Leung, C.L., et al. Mutant small heat-shock protein 27 causes axonal Charcot-Marie-Tooth disease and distal hereditary motor neuropathy. *Nat Genet*, 2004. **36**(6): p. 602-6.

9. Harms, M.B., Ori-McKenney, K.M., Scoto, M., Tuck, E.P., Bell, S., Ma, D., et al. Mutations in the tail domain of DYNC1H1 cause dominant spinal muscular atrophy. *Neurology*, 2012. **78**(22): p. 1714-20.

10. Puls, I., Jonnakuty, C., LaMonte, B.H., Holzbaur, E.L., Tokito, M., Mann, E., et al. Mutant dynactin in motor neuron disease. *Nat Genet*, 2003. **33**(4): p. 455-6.

11. Kahle, K.T., Flores, B., Bharucha-Goebel, D., Zhang, J.W., Donkervoort, S., Hegde, M., et al. Peripheral motor neuropathy is associated with defective kinase regulation of the KCC3 cotransporter. *Science Signaling*, 2016. **9**(439): p. 1-14.

12. Barwick, K.E., Wright, J., Al-Turki, S., McEntagart, M.M., Nair, A., Chioza, B., et al. Defective presynaptic choline transport underlies hereditary motor neuropathy. *Am J Hum Genet*, 2012. **91**(6): p. 1103-7.

13. Kennerson, M.L., Nicholson, G.A., Kaler, S.G., Kowalski, B., Mercer, J.F., Tang, J., et al. Missense mutations in the copper transporter gene ATP7A cause X-linked distal hereditary motor neuropathy. *Am J Hum Genet*, 2010. **86**(3): p. 343-52.

14. Grohmann, K., Schuelke, M., Diers, A., Hoffmann, K., Lucke, B., Adams, C., et al. Mutations in gene encoding imunoglobulin u-binding protein 2 cause spinal musclar atrophy with respiratory distress type 1. *Nature Genetics*, 2001. **29**: p. 75-77.

15. Chen, Y.Z., Bennett, C.L., Huynh, H.M., Blair, I.P., Puls, I., Irobi, J., et al. DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am J Hum Genet*, 2004. **74**(6): p. 1128-35.

16. Butterfield, R.J., Stevenson, T.J., Xing, L., Newcomb, T.M., Nelson, B., Zeng, W., et al. Congenital lethal motor neuron disease with a novel defect in ribosome biogenesis. *Neurology*, 2014. **82**(15): p. 1322-30.

17. Safka Brozkova, D., Deconinck, T., Griffin, L.B., Ferbert, A., Haberlova, J., Mazanec, R., et al. Loss of function mutations in HARS cause a spectrum of inherited peripheral neuropathies. *Brain*, 2015. **138**(Pt 8): p. 2161-72.

18. Antonellis, A., Ellsworth, R.E., Sambuughin, N., Puls, I., Abel, A., Lee-Lin, S.Q., et al. Glycyl tRNA synthetase muations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy. *American Journal of Human Genetics*, 2003. **72**: p. 1293-1299.

19. Zhao, Z., Hashiguchi, A., Hu, J., Sakiyama, Y., Okamoto, Y., Tokunaga, S., et al. Alanyl-tRNA synthetase mutation in a family with dominant distal hereditary motor neuropathy. *Neurology*, 2012. **78**(21): p. 1644-9.

20. Pitceathly, R.D., Murphy, S.M., Cottenie, E., Chalasani, A., Sweeney, M.G., Woodward, C., et al. Genetic dysfunction of MT-ATP6 causes axonal Charcot-Marie-Tooth disease. *Neurology*, 2012. **79**(11): p. 1145-54.

21. Aure, K., Dubourg, O., Jardel, C., Clarysse, L., Sternberg, D., Fournier, E., et al. Episodic weakness due to mitochondrial DNA MT-ATP6/8 mutations. *Neurology*, 2013. **81**(21): p. 1810-1818.

22. Stoll, M., Teoh, H., Lee, J., Reddel, S., Zhu, Y., Buckley, M., et al. Novel motor phenotypes in patients with VRK1 mutations without pontocerebellar hypoplasia. *Neurology*, 2016. **87**(1): p. 65-70.

23. Zhao, H., Race, V., Matthijs, G., De Jonghe, P., Robberecht, W., Lambrechts, D., et al. Exome sequencing reveals HINT1 mutations as a cause of distal hereditary motor neuropathy. *Eur J Hum Genet*, 2014. **22**(6): p. 847-50.

24. Maystadt, I., Rezsohazy, R., Barkats, M., Duque, S., Vannuffel, P., Remacle, S., et al. The nuclear factor kappaB-activator gene PLEKHG5 is mutated in a form of autosomal recessive lower motor neuron disease with childhood onset. *Am J Hum Genet*, 2007. **81**(1): p. 67-76.

25. Sumner, C.J., d'Ydewalle, C., Wooley, J., Fawcett, K.A., Hernandez, D., Gardiner, A.R., et al. A dominant mutation in FBXO38 causes distal spinal muscular atrophy with calf predominance. *Am J Hum Genet*, 2013. **93**(5): p. 976-83.

26. Choi, B.O., Kang, S.H., Hyun, Y.S., Kanwal, S., Park, S.W., Koo, H., et al. A complex phenotype of peripheral neuropathy, myopathy, hoarseness, and hearing loss is linked to an autosomal dominant mutation in MYH14. *Hum Mutat*, 2011. **32**(6): p. 669-77.

27. Auer-Grumbach, M., Weger, M., Fink-Puches, R., Papic, L., Frohlich, E., Auer-Grumbach, P., et al. Fibulin-5 mutations link inherited neuropathies, age-related macular degeneration and hyperelastic skin. *Brain*, 2011. **134**(Pt 6): p. 1839-52.

28. Sferra, A., Baillat, G., Rizza, T., Barresi, S., Flex, E., Tasca, G., et al. TBCE Mutations Cause Early-Onset Progressive Encephalopathy with Distal Spinal Muscular Atrophy. *Am J Hum Genet*, 2016. **99**(4): p. 974-983.

29. Cutrupi, A.N., Brewer, M.H., Nicholson, G.A. and Kennerson, M.L. Structural variations causing inherited peripheral neuropathies: A paradigm for understanding genomic organization, chromatin interactions, and gene dysregulation. *Mol Genet Genomic Med*, 2018. **6**(3): p. 422-433.

30. Alkan, C., Coe, B.P. and Eichler, E.E. Genome structural variation discovery and genotyping. *Nat Rev Genet*, 2011. **12**(5): p. 363-76.

31. Guan, P. and Sung, W.K. Structural variation detection using next-generation sequencing data: A comparative technical review. *Methods*, 2016. **102**: p. 36-49.

32. Stankiewicz, P. and Lupski, J.R. Structural variation in the human genome and its role in disease. *Annu Rev Med*, 2010. **61**: p. 437-55.

33. Sudmant, P.H., Rausch, T., Gardner, E.J., Handsaker, R.E., Abyzov, A., Huddleston, J., et al. An integrated map of structural variation in 2,504 human genomes. *Nature*, 2015. **526**(7571): p. 75-81.

34. Hurles, M.E., Dermitzakis, E.T. and Tyler-Smith, C. The functional impact of structural variation in humans. *Trends Genet*, 2008. **24**(5): p. 238-45.

35. Lupski, J.R., de Oca-Luna, R.M., Slaugenhaupt, S., Pentao, L., Guzzetta, V., Trask, B.J., et al. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell*, 1991. **66**(2): p. 219-32.

36. Lupski, J.R., Wise, C.A., Kuwano, A., Pentao, L., Parke, J.T., Glaze, D.G., et al. Gene Dosage Is a Mechanism for Charcot-Marie-Tooth Disease Type-1a. *Nature Genetics*, 1992. **1**(1): p. 29-33.

37. Patel, P.I., Roa, B.B., Welcher, A.A., Schoenerscott, R., Trask, B.J., Pentao, L., et al. The Gene for the Peripheral Myelin Protein-Pmp-22 Is a Candidate for Charcot-Marie-Tooth Disease Type-1a. *Nature Genetics*, 1992. **1**(3): p. 159-165.

38. Timmerman, V., Nelis, E., Vanhul, W., Nieuwenhuijsen, B.W., Chen, K.L., Wang, S., et al. The Peripheral Myelin Protein Gene Pmp-22 Is Contained within the Charcot-Marie-Tooth Disease Type-1a Duplication. *Nature Genetics*, 1992. **1**(3): p. 171-175.

39. Valentijn, L.J., Baas, F., Wolterman, R.A., Hoogendijk, J.E., van den Bosch, N.H., Zorn, I., et al. Identical point mutations of PMP-22 in Trembler-J mouse and Charcot-Marie-Tooth disease type 1A. *Nat Genet*, 1992. **2**(4): p. 288-91.

40. Weterman, M.A., van Ruissen, F., de Wissel, M., Bordewijk, L., Samijn, J.P., van der Pol, W.L., et al. Copy number variation upstream of PMP22 in Charcot-Marie-Tooth disease. *Eur J Hum Genet*, 2010. **18**(4): p. 421-8.

41. Zhang, F., Seeman, P., Liu, P., Weterman, M.A., Gonzaga-Jauregui, C., Towne, C.F., et al. Mechanisms for nonrecurrent genomic rearrangements associated with CMT1A or HNPP: rare CNVs as a cause for missing heritability. *Am J Hum Genet*, 2010. **86**(6): p. 892-903.

42. Pantera, H., Moran, J.J., Hung, H.A., Pak, E., Dutra, A. and Svaren, J. Regulation of the neuropathy-associated Pmp22 gene by a distal super-enhancer. *Hum Mol Genet*, 2018. **27**(16): p. 2830-2839.

43. Feuk, L., Marshall, C.R., Wintle, R.F. and Scherer, S.W. Structural variants: changing the landscape of chromosomes and design of disease studies. *Hum Mol Genet*, 2006. **15 Spec No 1**: p. R57-66.

44. Kleinjan, D.J. and van Heyningen, V. Position effect in human genetic disease. *Hum Mol Genet*, 1998. **7**(10): p. 1611-8.

45. Yu, J., Chau, K.F., Vodyanik, M.A., Jiang, J. and Jiang, Y. Efficient feeder-free episomal reprogramming with small molecules. *PLoS One*, 2011. **6**(3): p. e17557.

46. Black, J. and Vos, J.M. Establishment of an oriP/EBNA1-based episomal vector transcribing human genomic beta-globin in cultured murine fibroblasts. *Gene Ther*, 2002. **9**(21): p. 1447-54.

47. Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, II, et al. Human induced pluripotent stem cells free of vector and transgene sequences. *Science*, 2009. **324**(5928): p. 797-801.

48. Du, Z.W., Chen, H., Liu, H., Lu, J., Qian, K., Huang, C.L., et al. Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells. *Nat Commun*, 2015. **6**: p. 6626.

49. Anders, S. and Huber, W. Differential expression analysis for sequence count data. *Genome Biol*, 2010. **11**(10): p. R106.

50. Bolger, A.M., Lohse, M. and Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 2014. **30**(15): p. 2114-20.

51. Kim, D., Paggi, J.M., Park, C., Bennett, C. and Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*, 2019. **37**(8): p. 907-915.

52. Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T. and Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*, 2015. **33**(3): p. 290-5.

53. McPherson, A., Hormozdiari, F., Zayed, A., Giuliany, R., Ha, G., Sun, M.G., et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLoS Comput Biol*, 2011. **7**(5): p. e1001138.

54. Nicorici, D., Satalan, M., Edgren, H., Kangaspeska, S., Murumagi, A., Kallioniemi, O., et al. FusionCatcher – a tool for finding somatic fusion genes in paired-end RNA-sequencing data. 2014.

55. Uhrig, S., Ellermann, J., WWalther, T., Burkhardt, P., Fröhlich, M., Hutter, B., et al. Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Research*, 2021. **31**: p. 448-460.

56. Love, M.I., Huber, W. and Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 2014. **15**(12): p. 550.

57. Huang, D.W., Sherman, B.T., Tan, Q., Collins, J.R., Alvord, W.G., Roayaei, J., et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol*, 2007. **8**(9): p. R183.

58. Huang, D.W., Sherman, B.T. and Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 2009. **4**(1): p. 44-57.

59. von Mering, C., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P. and Snel, B. STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res*, 2003. **31**(1): p. 258-61.

60. von Mering, C., Jensen, L.J., Snel, B., Hooper, S.D., Krupp, M., Foglierini, M., et al. STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res*, 2005. **33**(Database issue): p. D433-7.

61. Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*, 2019. **47**(D1): p. D607-D613.

62. Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*, 2021. **49**(D1): p. D605-D612.

63. Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*, 2015. **16**: p. 259.

64. Kerpedjiev, P., Abdennur, N., Lekschas, F., McCallum, C., Dinkla, K., Strobelt, H., et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol*, 2018. **19**(1): p. 125.

65. Kruse, K., Hug, C.B., Hernandez-Rodriguez, B. and Vaquerizas, J.M. TADtool: visual parameter identification for TAD-calling algorithms. *Bioinformatics*, 2016. **32**(20): p. 3190-3192.

66. Mao, D., Chung, X.K.W., Andoh-Noda, T., Qin, Y., Sato, S.I., Takemoto, Y., et al. Chemical decontamination of iPS cell-derived neural cell mixtures. *Chem Commun (Camb)*, 2018. **54**(11): p. 1355-1358.

67. Perez-Siles, G., Cutrupi, A.N., Ellis, M., Kuriakose, J., La Fontaine, S., D, M., et al. Modelling the pathogenesis of X-lined distal hereditary motor neuropathy using patient-derived IPSC. *Disease Models and Mechanisms*, 2020.

68. Bernal, A. and Arranz, L. Nestin-expressing progenitor cells: function, identity and therapeutic implications. *Cell Mol Life Sci*, 2018. **75**(12): p. 2177-2195.

69. Gopinath, S., Blair, I.P., Kennerson, M.L., Durnall, J.C. and Nicholson, G.A. A novel locus for distal motor neuron degeneration maps to chromosome 7q34-q36. *Hum Genet*, 2007. **121**(5): p. 559-64.

70. Kleinjan, D.A. and van Heyningen, V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet*, 2005. **76**(1): p. 8-32.

71. Kleinjan, D.A. and Lettice, L.A. Long-range gene control and genetic disease. *Adv Genet*, 2008. **61**: p. 339-88.

72. Kleinjan, D.J. and Coutinho, P. Cis-ruption mechanisms: disruption of cis-regulatory control as a cause of human genetic disease. *Brief Funct Genomic Proteomic*, 2009. **8**(4): p. 317-32.

73. Long, H.K., Osterwalder, M., Welsh, I.C., Hansen, K., Davies, J.O.J., Liu, Y.E., et al. Loss of Extreme Long-Range Enhancers in Human Neural Crest Drives a Craniofacial Disorder. *Cell Stem Cell*, 2020. **27**(5): p. 765-783 e14.

74. Huang, W., Guo, Y., Du, W., Zhang, X., Li, A. and Miao, X. Global transcriptome analysis identifies differentially expressed genes related to lipid metabolism in Wagyu and Holstein cattle. *Sci Rep*, 2017. **7**(1): p. 5278.

75. Stark, R., Grzelak, M. and Hadfield, J. RNA sequencing: the teenage years. *Nat Rev Genet*, 2019. **20**(11): p. 631-656.

76. Reese, M.G., Eeckman, F.H., Kulp, D. and Haussler, D. Improved Splice Site Detection in Genie. *Journal of Computational Biology*, 1997. **4**(7): p. 311-223.

77. You, J. and Pickart, C.M. A HECT domain E3 enzyme assembles novel polyubiquitin chains. *J Biol Chem*, 2001. **276**(23): p. 19871-8.

78. You, J., Wang, M., Aoki, T., Tamura, T.A. and Pickart, C.M. Proteolytic targeting of transcriptional regulator TIP120B by a HECT domain E3 ligase. *J Biol Chem*, 2003. **278**(26): p. 23369-75.

79. Singh, S. and Sivaraman, J. Crystal structure of HECT domain of UBE3C E3 ligase and its ubiquitination activity. *Biochem J*, 2020. **477**(5): p. 905-923.

80. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. Basic local alignment search tool. *J Mol Biol*, 1990. **215**(3): p. 403-10.

81. Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Reserach*, 1997. **25**(17): p. 3389-3402.

82. Altschul, S.F., Wootton, J.C., Gertz, E.M., Agarwala, R., Morgulis, A., Schäffer, A.A., et al. Protein Database Searches Using Compositionally Adjusted Substitution Matrices. *FEBS J*, 2005. **272**(20): p. 5101-5109.

83. Papadopoulos, J.S. and Agarwala, R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics*, 2007. **23**(9): p. 1073-9.

84. Nonet, M.L., Saifee, O., Zhao, H., Rand, J.B. and Wei, L. Synaptic transmission deficits in Caenorhabditis elegans synaptobrevin mutants. *J Neurosci*, 1998. **18**(1): p. 70-80.

85. Narayanan, R.K., Brewer, M.H., Perez-Siles, G., Ellis, M., Ly, C., Burgess, A., et al. Charcot-Marie-tooth disease causing mutation (p.R158H) in pyruvate dehydrogenase kinase 3 (PDK3) affects synaptic transmission, ATP production and causes neurodegeneration in a CMTX6 C. elegans model. *Hum Mol Genet*, 2021. **31**(1): p. 133-145.

86. Fang, N.N., Ng, A.H., Measday, V. and Mayor, T. Hul5 HECT ubiquitin ligase plays a major role in the ubiquitylation and turnover of cytosolic misfolded proteins. *Nat Cell Biol*, 2011. **13**(11): p. 1344-52.

87. Boyling, A., Perez-Siles, G. and Kennerson, M.L. Structural Variation at a Disease Mutation Hotspot: Strategies to Investigate Gene Regulation and the 3D Genome. *Front Genet*, 2022. **13**: p. 842860.

88. Nowell, P.C. and Hungerford, D.A. A minute chromosome in human chronic granulocytic leukemia. *Science*, 1960. **132**(1497).

89. Rowley, J.D. A new consisten chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine flourescence giemsa staining. *Nature*, 1973. **243**(1): p. 290-293.

90. Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.-W., et al. Recurrent Fusion of TMPRSS2 and ETS Transcription Factor Genes in Prostate Cancer. *Science*, 2005. **310**(5784): p. 644-648.

91. Mitelman, F., Johansson, B. and Mertens, F. The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer*, 2007. **7**(4): p. 233-45.

92. Welch, J.S., Westervelt, P., Ding, L., Larson, D.E., Klco, J.M., Kulkarni, S., et al. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA*, 2011. **305**(15): p. 1577-84.

93. Annala, M.J., Parker, B.C., Zhang, W. and Nykter, M. Fusion genes and their discovery using high throughput sequencing. *Cancer Lett*, 2013. **340**(2): p. 192-200.

94. Yun, J.W., Yang, L., Park, H.Y., Lee, C.W., Cha, H., Shin, H.T., et al. Dysregulation of cancer genes by recurrent intergenic fusions. *Genome Biol*, 2020. **21**(1): p. 166.

95. Taniue, K. and Akimitsu, N. Fusion Genes and RNAs in Cancer Development. *Noncoding RNA*, 2021. **7**(1).

96. Hoogstrate, Y., Komor, M.A., Bottcher, R., van Riet, J., van de Werken, H.J.G., van Lieshout, S., et al. Fusion transcripts and their genomic breakpoints in polyadenylated and ribosomal RNA-minus RNA sequencing data. *Gigascience*, 2021. **10**(12).

97. Waddell, L.B., Bryen, S.J., Cummings, B.B., Bournazos, A., Evesson, F.J., Joshi, H., et al. WGS and RNA Studies Diagnose Noncoding DMD Variants in Males With High Creatine Kinase. *Neurol Genet*, 2021. **7**(1): p. e554.

98. Cmero, M., Schmidt, B., Majewski, I.J., Ekert, P.G., Oshlack, A. and Davidson, N.M. MINTIE: identifying novel structural and splice variants in transcriptomes using RNA-seq data. *Genome Biol*, 2021. **22**(1): p. 296.

99. Chen, X., Zhao, G., Zhong, P., Zhang, M., Chen, R. and Zhang, D. Chr2 30297612-ALK, A Novel Intergenic Fusion With Exon18 of ALK, Responds to Crizotinib. *Clin Lung Cancer*, 2020. **21**(6): p. e564-e566.

100. McPherson, A., Wu, C., Wyatt, A.W., Shah, S., Collins, C. and Sahinalp, S.C. nFuse: discovery of complex genomic rearrangements in cancer using high-throughput sequencing. *Genome Res*, 2012. **22**(11): p. 2250-61.

101. Arber, S., Han, B., Mendelsohn, M., Smith, M., Jessel, T.M. and Sockanathan, S. Requirement for the homeobox gene Hb9 in the consolidation of motor neuon identity. *Neuron*, 1999. **23**: p. 659-674.

102. Ambrozkiewicz, M.C. and Kawabe, H. HECT-type E3 ubiquitin ligases in nerve cell development and synapse physiology. *FEBS Lett*, 2015. **589**(14): p. 1635-43.

103. George, A.J., Hoffiz, Y.C., Charles, A.J., Zhu, Y. and Mabb, A.M. A Comprehensive Atlas of E3 Ubiquitin Ligase Mutations in Neurological Disorders. *Front Genet*, 2018. **9**: p. 29.

104. Rotin, D. and Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol*, 2009. **10**(6): p. 398-409.

105. Weber, J., Polo, S. and Maspero, E. HECT E3 Ligases: A Tale With Multiple Facets. *Front Physiol*, 2019. **10**: p. 370.

106. Ambrozkiewicz, M.C., Cuthill, K.J., Harnett, D., Kawabe, H. and Tarabykin, V. Molecular Evolution, Neurodevelopmental Roles and Clinical Significance of HECT-Type UBE3 E3 Ubiquitin Ligases. *Cells*, 2020. **9**(11).

107. Lescouzeres, L. and Bomont, P. E3 Ubiquitin Ligases in Neurological Diseases: Focus on Gigaxonin and Autophagy. *Front Physiol*, 2020. **11**: p. 1022.

108. Kaneko, M., Iwase, I., Yamasaki, Y., Takai, T., Wu, Y., Kanemoto, S., et al. Genome-wide identification and gene expression profiling of ubiquitin ligases for endoplasmic reticulum protein degradation. *Sci Rep*, 2016. **6**: p. 30955.

109. Iconomou, M. and Saunders, D.N. Systematic approaches to identify E3 ligase substrates. *Biochem J*, 2016. **473**(22): p. 4083-4101.

110. Jain, A.K. and Barton, M.C. Regulation of p53: TRIM24 enters the RING. *Cell Cycle*, 2009. **8**(22): p. 3668-74.

111. de Bie, P. and Ciechanover, A. Ubiquitination of E3 ligases: self-regulation of the ubiquitin system via proteolytic and non-proteolytic mechanisms. *Cell Death Differ*, 2011. **18**(9): p. 1393-402.

112. Balaji, V. and Hoppe, T. Regulation of E3 ubiquitin ligases by homotypic and heterotypic assembly. *F1000Res*, 2020. **9**.

113. Fang, S., Jensen, J.P., Ludwig, R.L., Vousden, K.H. and Weissman, A.M. Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself and p53. *J Biol Chem*, 2000. **275**(12): p. 8945-51.

114. Hu, G. and Fearon, E.R. Siah-1 N-Terminal RING domain is required for proteolysis function, and C-terminal sequences regulate oligomerization and binding to target proteins. *Molecular and Cellular Biology*, 1998. **19**(1): p. 724-732.

115. Chu, B.W., Kovary, K.M., Guillaume, J., Chen, L.C., Teruel, M.N. and Wandless, T.J. The E3 ubiquitin ligase UBE3C enhances proteasome processivity by ubiquitinating partially proteolyzed substrates. *J Biol Chem*, 2013. **288**(48): p. 34575-87.

116. Dantuma, N.P. and Bott, L.C. The ubiquitin-proteasome system in neurodegenerative diseases: precipitating factor, yet part of the solution. *Front Mol Neurosci*, 2014. **7**: p. 70.

117. Zheng, Q., Huang, T., Zhang, L., Zhou, Y., Luo, H., Xu, H., et al. Dysregulation of Ubiquitin-Proteasome System in Neurodegenerative Diseases. *Front Aging Neurosci*, 2016. **8**: p. 303.

118. Cheroni, C., Marino, M., Tortarolo, M., Veglianese, P., De Biasi, S., Fontana, E., et al. Functional alterations of the ubiquitin-proteasome system in motor neurons of a mouse model of familial amyotrophic lateral sclerosis. *Hum Mol Genet*, 2009. **18**(1): p. 82-96.

119. Chisholm, C.G., Lum, J.S., Farrawell, N.E. and Yerbury, J.J. Ubiquitin homeostasis disruption, a common cause of proteostasis collapse in amyotrophic lateral sclerosis? *Neural Regeneration Research*, 2022. **17**(10): p. 2218-2220.

120. Ruegsegger, C. and Saxena, S. Proteostasis impairment in ALS. *Brain Res*, 2016. **1648**(Pt B): p. 571-579.

121. Lambert-Smith, I.A., Saunders, D.N. and Yerbury, J.J. Proteostasis impairment and ALS. *Prog Biophys Mol Biol*, 2022.

122. Mastrandrea, L.D., You, J., Niles, E.G. and Pickart, C.M. E2/E3-mediated assembly of lysine 29-linked polyubiquitin chains. *J Biol Chem*, 1999. **274**(38): p. 27299-306.

123. Aviram, S. and Kornitzer, D. The ubiquitin ligase Hul5 promotes proteasomal processivity. *Mol Cell Biol*, 2010. **30**(4): p. 985-94.

124. Kohlmann, S., Schafer, A. and Wolf, D.H. Ubiquitin ligase Hul5 is required for fragment-specific substrate degradation in endoplasmic reticulum-associated degradation. *J Biol Chem*, 2008. **283**(24): p. 16374-83.

125. Bussmann, J. and Storkebaum, E. Molecular pathogenesis of peripheral neuropathies: insights from Drosophila models. *Curr Opin Genet Dev*, 2017. **44**: p. 61-73.

Figure legends

Figure 1 DHMN1 iPSCs harbouring the 1.35 Mb complex insertion display a normal karyotype and exhibit molecular features of pluripotency. (A) Karyotyping and G-band analysis confirm that iPSC lines have a normal 46, XY karyotype. **(B)** DHMN1 multiplex breakpoint genotyping assay for a representative patient (P1, closed diamond) and control (C1, open diamond) confirms the presence of the complex insertion in DHMN1 iPSC lines. The mutant junction fragments (501 bp, proximal; 412 bp, distal) and wild-type fragments (409 bp, proximal and 290 bp distal) were amplified from the respective DHMN1 and wild type

chromosomes. **(C)** Immunofluorescence staining confirms expression of pluripotency markers OCT4, SOX2 and NANOG for a representative patient iPSC sample (P2). Images are presented as pluripotency markers (green) overlayed on DAPI stained nuclei (blue). Scale bar 60 µm.

Figure 2 sMN derived from patient and control iPSC do not show changes in number of MNX1/HB9^{+} sMN or 2D neuronal networks. (A) Modified differentiation workflow used to generate MNP and sMN. Timeline of the differentiation process, media and factors used throughout the protocol. The yellow circle at day 21 indicates initiation of treatment with SN38-P **(B)** Immunofluorescence staining of iPSC-derived sMN for MN markers TUBB3, MNX1/HB9 and NF68. Scale bar 90 µm. **(C)** NanoString measurements comparing sMN and iPSC show increased expression of markers for motor neuron identity *ISL1, MNX1/HB9* and *ChAT*. **[D(i)]** Immunofluorescence staining of a representative patient and control sMN line positive for MNX1/HB9 and NF68. Scale bar 90 µm. **[D(ii)]** Comparison of the number of sMN generated from patient (n = 3) and control (n = 3) lines. Approximately 2500 cells were analysed per line. Data was pooled for each line from three independent rounds of differentiation. The number of sMN are represented by the percentage of nuclei stained for MNX1/HB9. **[E(i)]** Immunofluorescence staining of a representative patient and control sMN line positive for TUBB3. Scale bar 90 µm. **[E(ii)]** Quantification of TUBB3 fluorescence in patient (n = 3) and control (n = 3) sMN lines reveals no difference in 2D neuronal networks. Data was pooled for each line from three independent rounds of differentiation.

Figure 3 NanoString assay reveals nine differentially expressed genes (DEG) within a 6 Mb region flanking the DHMN1 insertion breakpoint sites. (**A**) Ideogram and localising the genomic region of chromosome 7q36.2-q36.3 involving the formation of the DHMN1 complex SV rearrangement. The green region represents the DHMN1 linkage region. The pink region represents the genomic interval at 7q36.3 defining the DHMN1 insertion sequence. Breakpoints are indicated by the vertical solid black lines. Letters a-d indicate genomic co-ordinates at the proximal (a) and distal (b) ends of the insertion site and the proximal (c) and distal (d) ends of the 7q36.3 duplicated sequence. Yellow bars indicate the genomic regions used to select genes for differential gene expression analysis with a custom designed NanoString panel. Horizontal black

lines show genomic distances. **(B)** Schematic representation of the DHMN1 locus showing the relative positions of differentially expressed genes within the queried genomic intervals. Arrows indicate orientation of genomic sequences. Letters a-d indicate genomic co-ordinates at the proximal (a) and distal (b) ends of the insertion site and the proximal (c) and distal (d) ends of the 7q36.3 duplicated sequence which is now in an inverted orientation within the DHMN1 locus. The insertion breakpoints are denoted by the vertical solid black lines. **(C)** Venn diagram depicting the overlap of statistically significant DEG from iPSC, MNP and differentiated sMN. **(D)** Mean normalized counts for the four genes showing statistically significant DEG between DHMN1 patient (n=3) and control (n=2) sMN. Data is represented as mean ± standard error (SE).

Figure 4 Hi-C analysis of DHMN1 sMN tissue identifies aberrant 3D genomic interactions and altered TAD profile at the DHMN1 locus in DHMN1 motor neurons. (A) Differential TAD profiles can be observed between patient (left panel, green box) and unaffected control (right panel, green box) sMN. **(B)** Contact matrices of DHMN1 (left panel, top and bottom) hg38 custom-built DHMN1 chromosome 7 and the hg38 reference genome (wild-type; right panel, top and bottom) sMN show 3D interaction profiles across the DHMN1 locus. Red boxes indicate differing regions of interaction. Red triangles depict the location of predicted TADs. A neo-TAD (yellow triangle) can be seen in patient sMN (left panel, bottom) suggesting the 1.35 Mb insertion changes the overall chromatin architecture at the DHMN1 locus creating new chromatin domains. **(C)** The Hi-C contact matrix when overlayed on a genomic map demonstrates the predicted neo-TAD (yellow triangle) overlaps with two significantly dysregulated candidate genes *MNX1* and *UBE3C* (red circles).

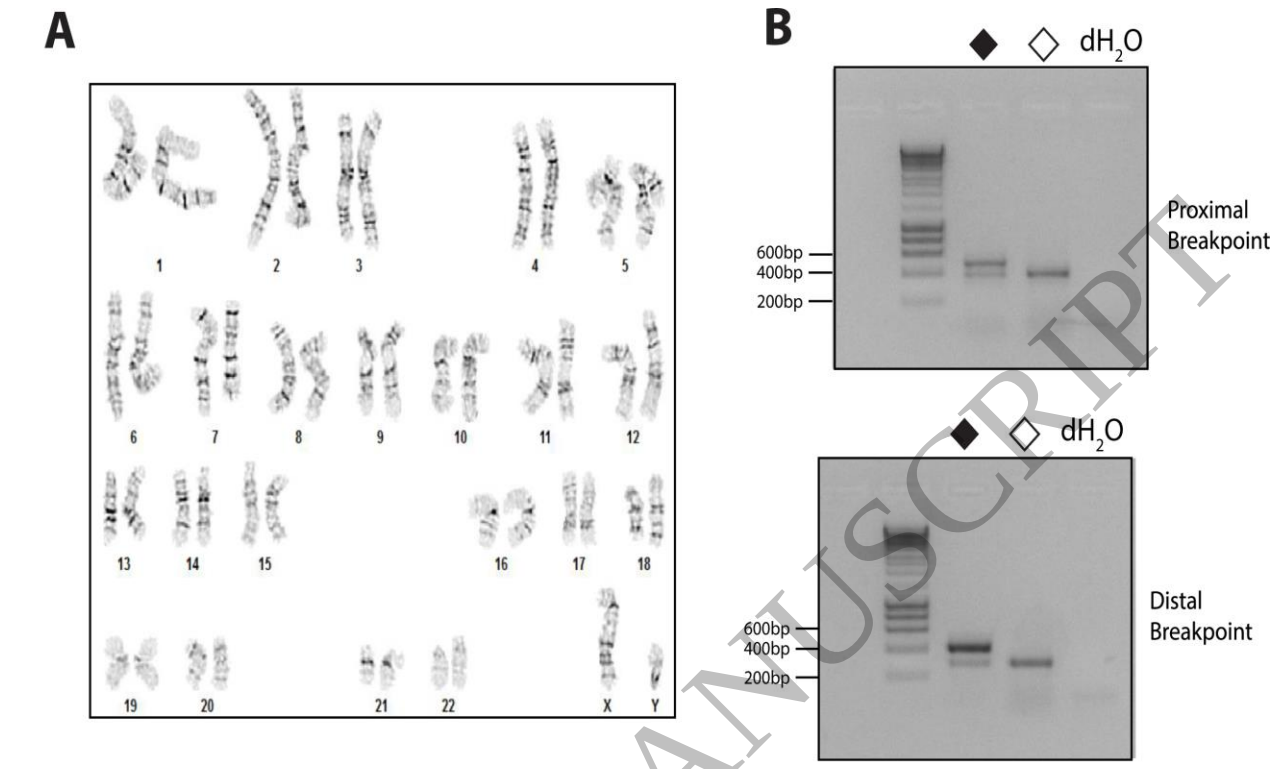
Figure 5 RNA-seq identifies a novel gene-intergenic fusion involving the partially duplicated *UBE3C*. [A(i)] DHMN1 genomic rearrangement showing the depth of coverage of exonic sequences for the novel fusion transcript. **[A(ii)]** Sanger sequencing of RT-PCR amplicons spanning the exon 10/pseudo-exon breakpoint confirms the presence of *UBE3C-IF* transcript in patient sMN mRNA. Coloured boxes indicate sequence alignments originating from the DHMN1 intergenic region (light green) and the partial *UBE3C* transcript exon 10 (light pink). The exact location of the end sequence from the partial *UBE3C* transcript exon 10 and start of the

DHMN1 intergenic sequence cannot be unambiguously defined due to a 5 bp (CCTGA) overlap (purple box). **(B)** RT-PCR validation of the *UBE3C* intergenic fusion transcript (*UBE3C-IF*) showing a 151 bp product amplified from the exon 10/pseudo-exon junction is present in patients (closed diamonds) and absent from controls (open diamonds). Expanded below to show a schematic of the fusion transcript. A terminal pseudo-exon (tPE) derived from intergenic DNA from within the DHMN1 locus is fused to exon 10 of *UBE3C*. **(C)** BLAST of tPE sequence from RNA-seq (green box) aligns with 100% sequence identity to the intergenic region at chr7:153,629,188-153,629,514 in the DHMN1 locus. Dashed black line indicates insertion site. **(D)** Schematic of the *UBE3C* transcript (refseq ID: NM_014671.1) showing duplicated and non-duplicated segments. Differential expression is observed between patient and control sMN when quantitative RT-PCR is performed using a TaqMan assay probe localising within the duplicated segment. Blue bars indicate probe location.

Figure 6 UBE3C-WT reduced in the presence of UBE3C-IF. (A) Representative western blot of UBE3C from sMN lysates. Ponceau stain is shown below (bottom panel). **(B)** Quantification of UBE3C-WT (123 kDa, top band) protein levels shows statistically significant reduction in DHMN1 sMN lysates compared to controls ($p < 0.0001$). UBE3C-WT was normalised to total protein (Ponceau). Quantification was performed using three independent experiments **(C)** A representative western blot of UBE3C-WT in HeLa cell lysates transfected with UBE3C-IF construct or empty vector (EV). UBE3C-IF can be detected using an antibody against UBE3C-WT (top band, middle panel). **(D)** Quantification of UBE3C-WT protein levels show a statistically significant reduction in cells transfected with 2.0 µg of UBE3C-IF construct when compared to EV ($p < 0.05$). Protein levels were normalised to Ponceau. Data is expressed as a ratio of UBE3C-WT expression in treatment conditions vs EV control. Quantification was performed using data from two independent experiments.

Figure 7 C. *elegans* transgenic animals overexpressing UBE3C-IF show synaptic transmission deficits and susceptibility to heat stress. (A and B) *C. elegans* strains were treated with 1 mM aldicarb and 0.2 mM Levamisole. In the presence of 1 mM aldicarb, *C. elegans* mutants carrying *UBE3C-IF* transgene showed significant resistance to aldicarb induced

paralysis when compared to control animals suggesting synaptic transmission deficits. Upon treatment with levamisole, transgenic *C. elegans* strains showed increased paralysis when compared to *oxIs12* animals. However, there was no significant difference between transgenic strains carrying empty vector and *UBE3C-IF* transgene suggesting that the aldicarb phenotype observed in the *oxIs12;UBE3C-IF* animals are due to pre-synaptic deficits. 20 to 25 animals were used per biological replicate with a minimum of three biological replicates per strain. The mean ± SEM is presented. * p-value < 0.05, two-tailed unpaired t-test. **(C)** Representative images of day-1 old control (*oxIs12 and oxIs12; EmptyVector*) and mutant (*oxIs12; UBE3C-IF*) *C. elegans* strains. The ventral nerve cord (VNC) of the transgenic strains were observed to be normal and similar to the *oxIs12* animals with all the cell bodies and axons intact. Neurodegeneration was not observed in *oxIs12* animals overexpressing the *UBE3C-IF* transgene. Scale bar – 0.3 mm. **(D)** Quantification of axonal neurodegeneration. No significant axonal degeneration was identified in transgenic animals when compared to controls. Animals used for live imaging are as follows: *oxIs12* (n=91), *oxIs12;EmptyVector* (n=79) and *oxIs12;UBE3C-IF* (n=78). ns - not significant. **(E)** Quantification of body thrash assay. Analysis of the swimming (thrashing) behaviour of day 4 old *C. elegans* strains showed no locomotion deficits associated with *UBE3C-IF* overexpression when compared to controls. ns - not significant. **(F)** Heat stress resistance assay. *C. elegans* strains were exposed to a temperature of 35ºC for 8 h and the survival percentage calculated following the heat exposure. Transgenic animals overexpressing the *UBE3C-IF* transgene showed a significant reduction in survival when compared to *oxIs12;EmptyVector* and *oxIs12* controls. Data was generated with 20 to 30 animals per genotype per replicate. A total of 4 experimental replicates were used for the heat stress resistance assay. Data is presented as mean ± SEM. One-way ANOVA and Tukey's multiple correction test. * adjusted p-value = 0.01 and ** adjusted p-value = 0.003.



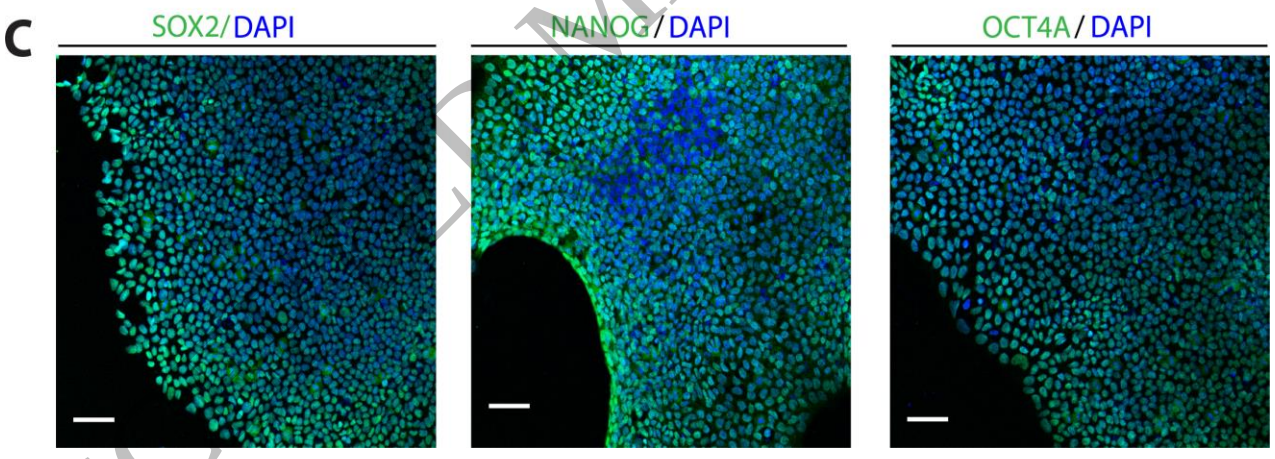


Figure 1
216x200 mm (.21 x DPI)

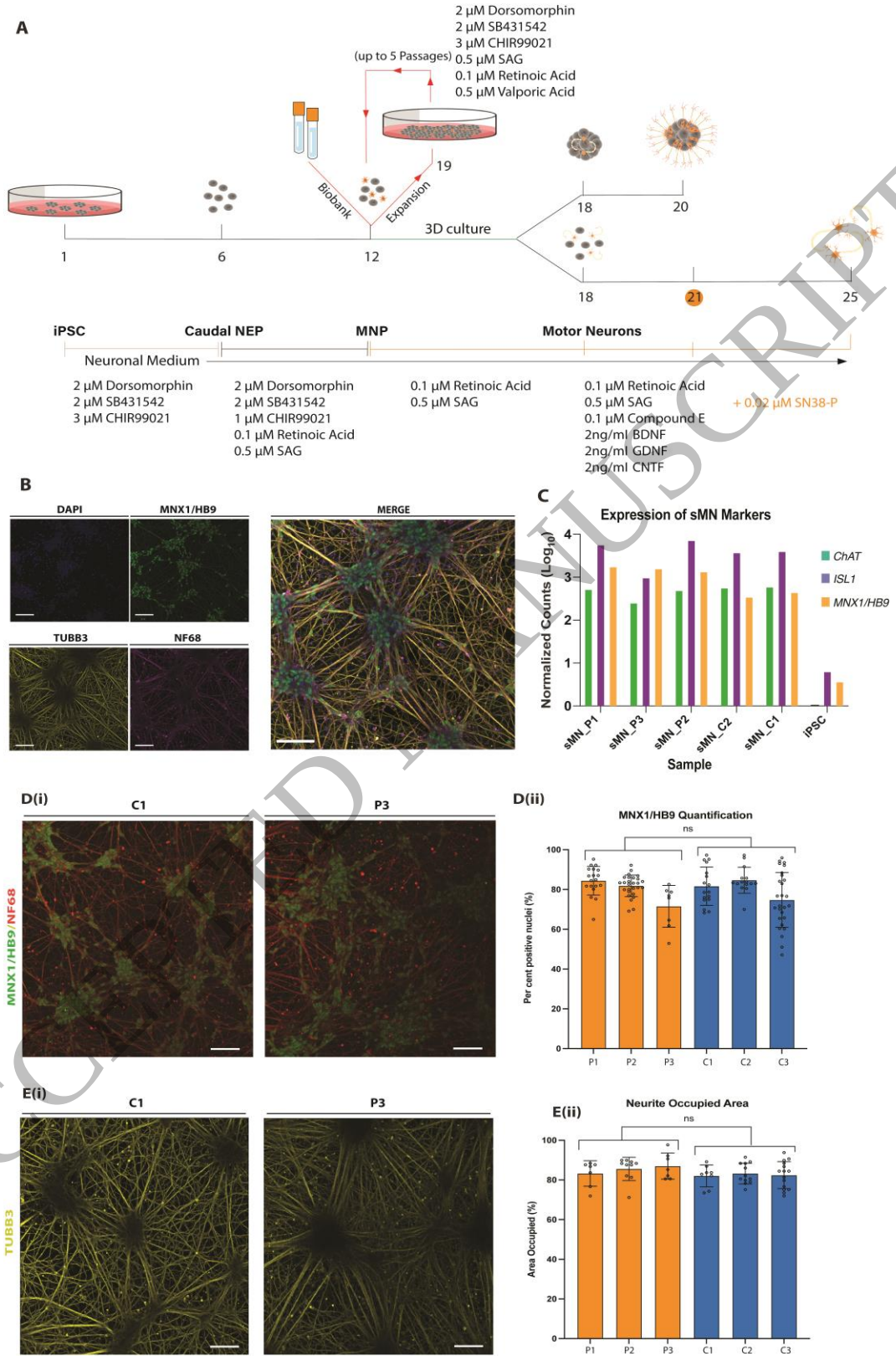


Figure 2
366x559 mm (.21 x DPI)

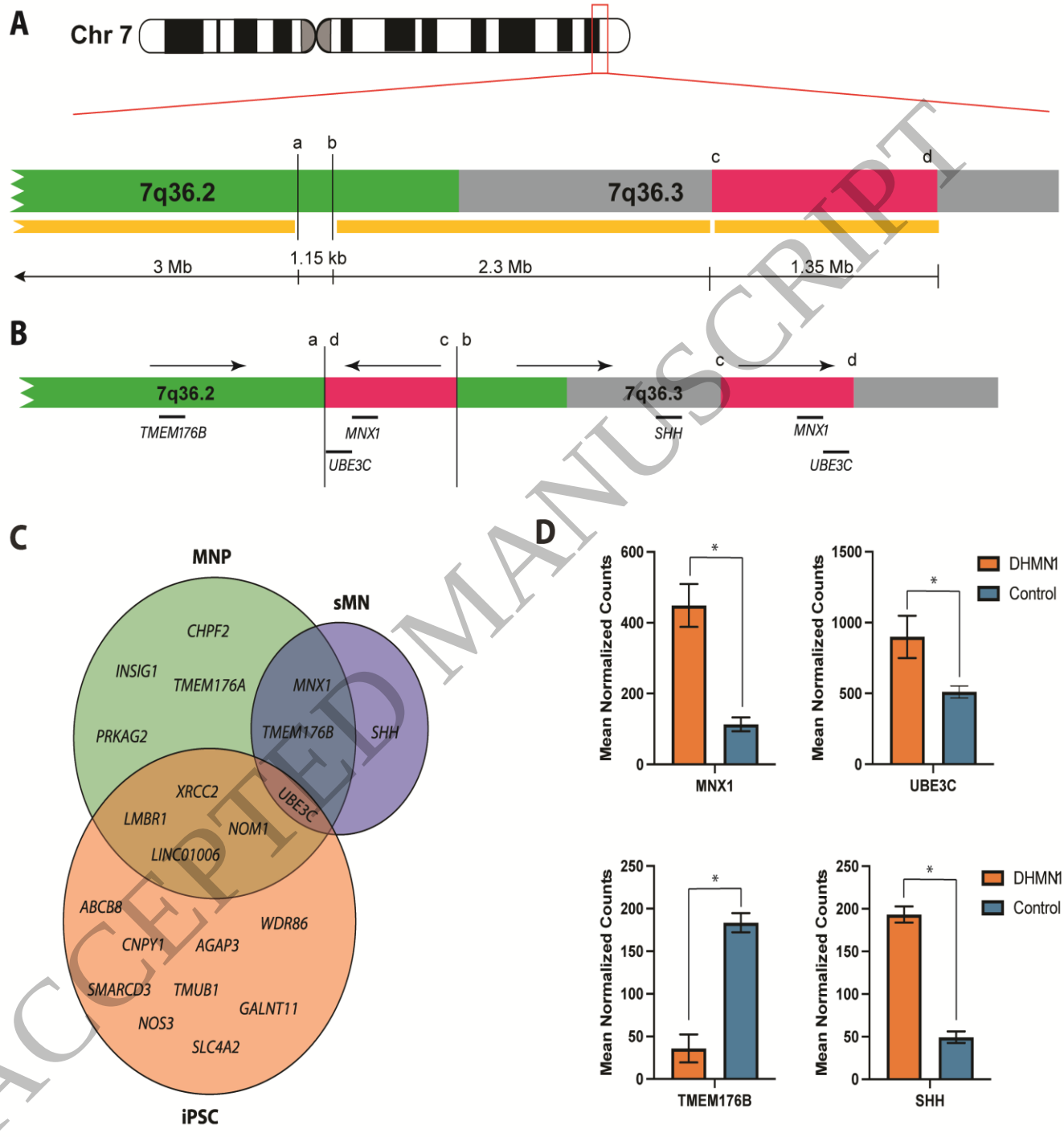


Figure 3
307x275 mm (.21 x DPI)

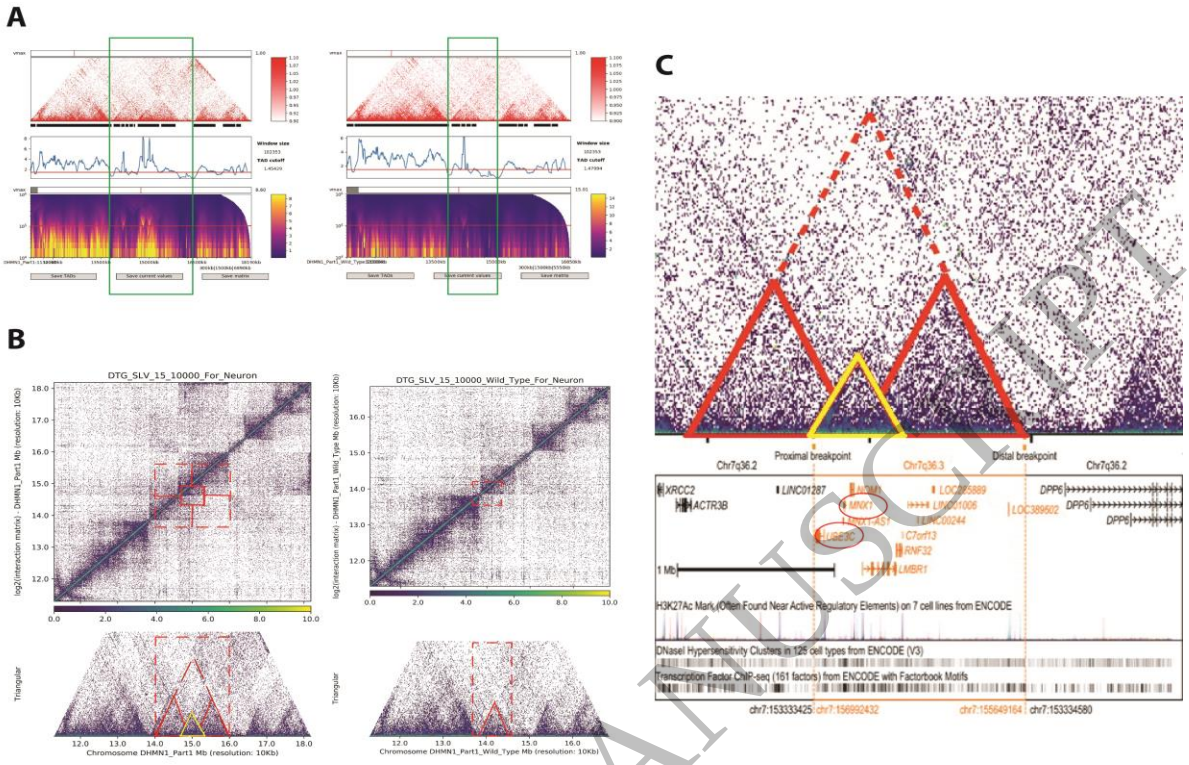


Figure 4
317x229 mm (.21 x DPI)

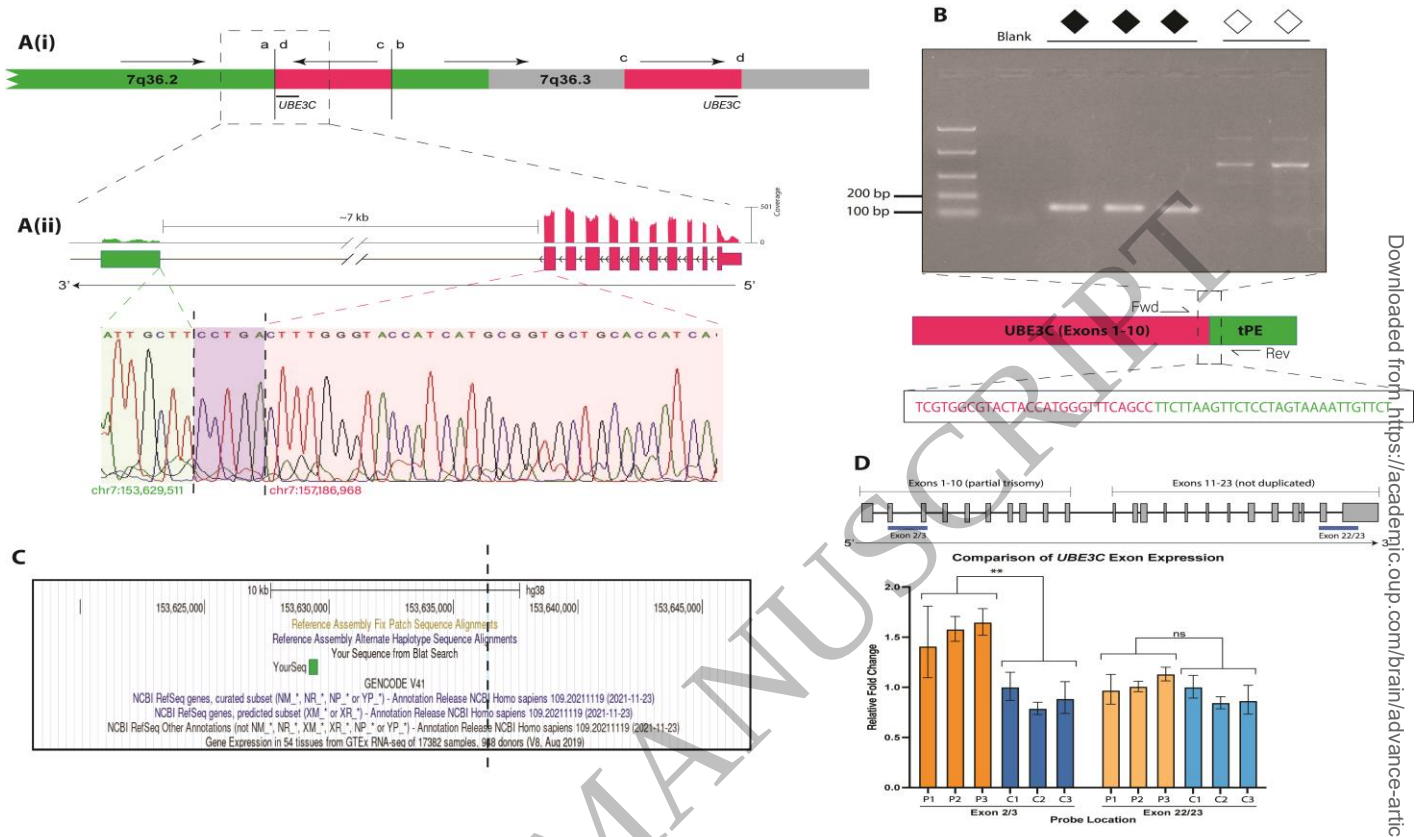


Figure 5

528x360 mm (.21 x DPI)

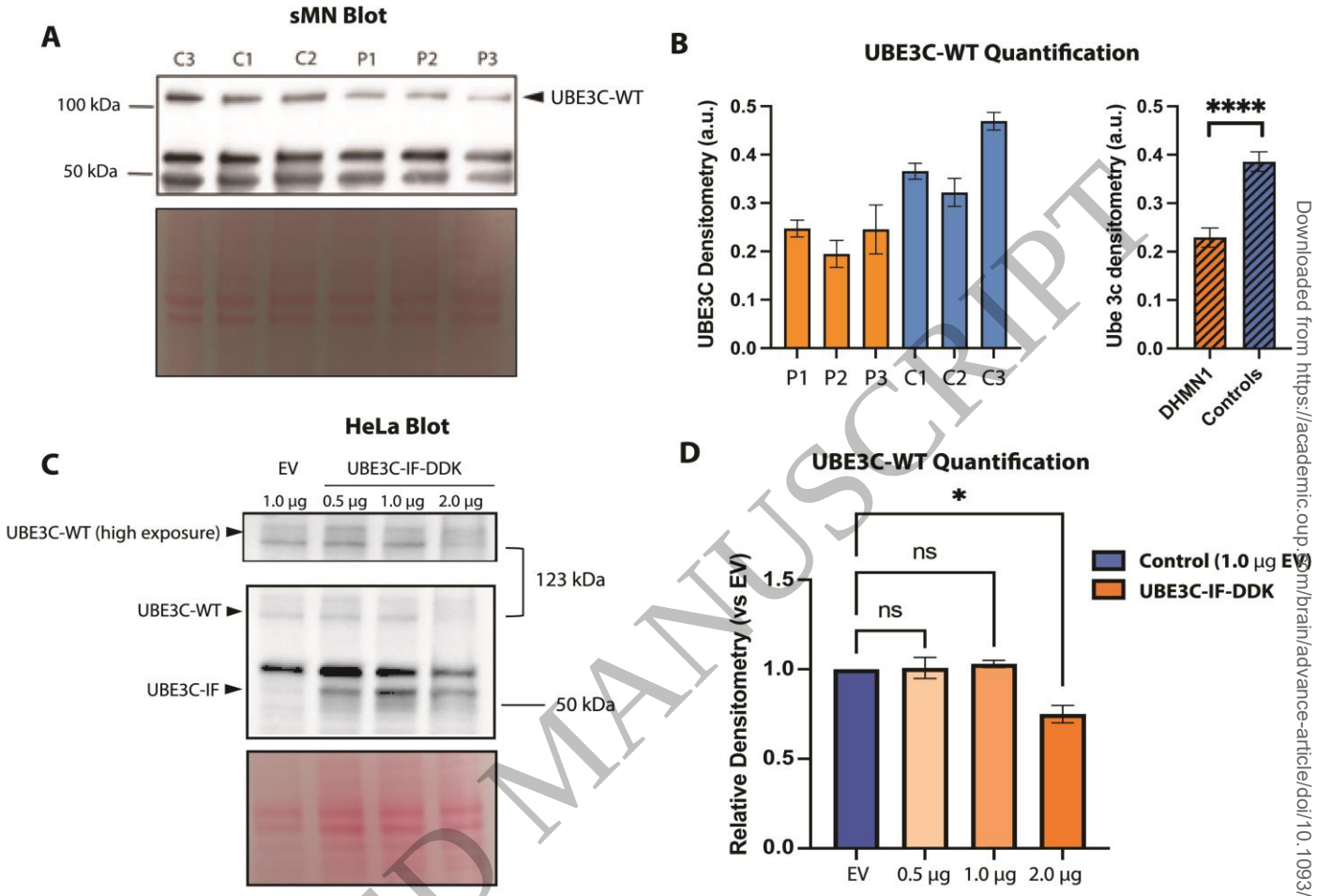


Figure 6
267x190 mm (.21 x DPI)

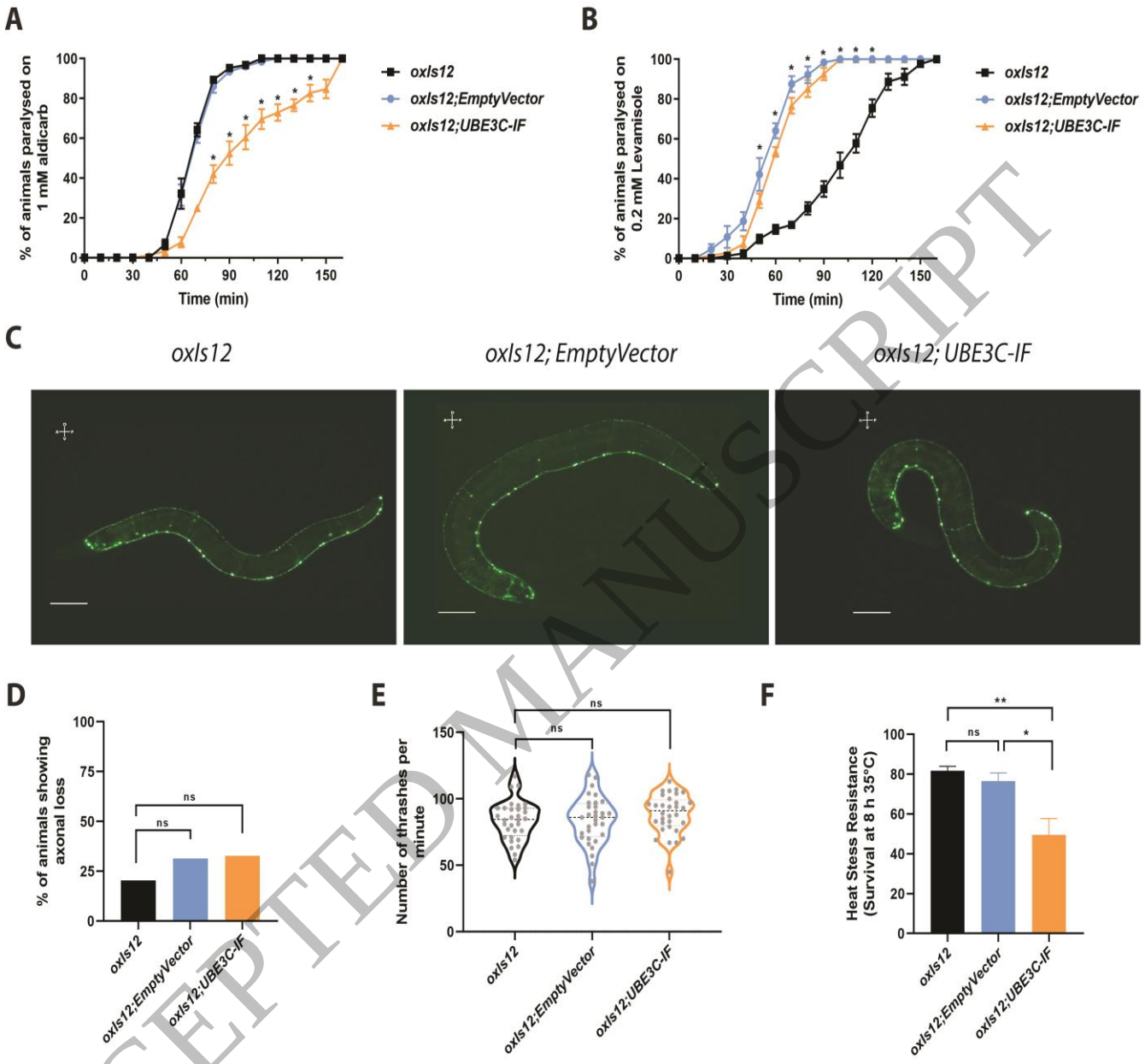


Figure 7
339x272 mm (.21 x DPI)