Received: 9 July 2021 | Revised: 16 September 2021 | Accepted: 22 September 2021

DOI: 10.1111/nan.12770

ORIGINAL ARTICLE

Neuropathology and Applied Neurobiology
JOURNAL OF THE BRITISH NEUROPATHOLOGICAL SOCIETY

WILEY

Image-based deep learning reveals the responses of human motor neurons to stress and *VCP*-related ALS

Colombine Verzat[1] | **Jasmine Harley**[2,3] | **Rickie Patani**[2,3] | **Raphaëlle Luisier**[1]

[1]Genomics and Health Informatics Group, Idiap Research Institute, Martigny, Switzerland

[2]Human Stem Cells and Neurodegeneration Laboratory, The Francis Crick Institute, London, UK

[3]Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, London, UK

Correspondence
Rickie Patani, Human Stem Cells and Neurodegeneration Laboratory, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK.
Email: rickie.patani@ucl.ac.uk

Raphaëlle Luisier, Genomics and Health Informatics Group, Idiap Research Institute, Martigny, Switzerland.
Email: raphaelle.luisier@idiap.ch

Funding information
MRC Senior Clinical Fellowship, Grant/Award Number: MR/S006591/1; Wellcome Trust; Medical Research Council; Cancer Research UK, Grant/Award Number: FC010110; Francis Crick Institute; Idiap Research Institute

Abstract

Aims: Although morphological attributes of cells and their substructures are recognised readouts of physiological or pathophysiological states, these have been relatively understudied in amyotrophic lateral sclerosis (ALS) research.

Methods: In this study, we integrate multichannel fluorescence high-content microscopy data with deep learning imaging methods to reveal—directly from unsegmented images—novel neurite-associated morphological perturbations associated with (ALS-causing) VCP-mutant human motor neurons (MNs).

Results: Surprisingly, we reveal that previously unrecognised disease-relevant information is withheld in broadly used and often considered 'generic' biological markers of nuclei (DAPI) and neurons (β III-tubulin). Additionally, we identify changes within the information content of ALS-related RNA binding protein (RBP) immunofluorescence imaging that is captured in VCP-mutant MN cultures. Furthermore, by analysing MN cultures exposed to different extrinsic stressors, we show that heat stress recapitulates key aspects of ALS.

Conclusions: Our study therefore reveals disease-relevant information contained in a range of both generic and more specific fluorescent markers and establishes the use of image-based deep learning methods for rapid, automated and unbiased identification of biological hypotheses.

KEYWORDS
amyotrophic lateral sclerosis, deep learning, human induced pluripotent stem cells, immunofluorescence, motor neurons

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is a relentlessly progressive and incurable neurodegenerative disease characterised by the loss of motor neurons (MNs). Key hallmarks of the disease include the mislocalisation and accumulation of ubiquitously expressed RNA binding proteins (RBPs) from the nucleus to the cytoplasm including TAR DNA-binding protein of 43 kDa (TDP-43), Fused in Sarcoma (FUS) and Splicing factor Proline and Glutamine rich (SFPQ) proteins.[1–4] What drives pathological mislocalisation and aggregation of RBPs in ALS remains unknown. However, alteration in liquid–liquid phase separation dynamics has been proposed to underlie this process.[5–9] RBPs are highly dynamic and have been shown to undergo changes in localisation in response to various stressors.[10–16] Notably, mitochondrial dysfunction and oxidative stress are recognised and robust phenotypes in ALS pathogenesis in vitro.[17] The role of RBPs in ALS and cellular stress highlights that a diverse and complex interplay exists.

All authors contributed equally to this work.

Cell shape and morphology are recognised readouts of a cell's physiological state or phenotype.[18] We previously reported common morphological descriptors that strongly discriminate sporadic ALS from control post-mortem tissue at single cell resolution,[19] further indicating that key information related to cellular state might be contained in cell shape in ALS. Dystrophic neurites are a common pathological feature in ALS, and disrupted synaptic integrity has been shown in valosin-containing protein (VCP) mutant human induced pluripotent stem cell (iPSC) cultures of MNs.[20] Taken together, these studies suggest that the neuronal processes (collectively termed neurites or the 'neuritome') may be a good cellular subcompartment to reveal ALS pathomechanisms. However, neurites are challenging to study both in tissue sections (as the arborisation of processes is not captured in tissue sections) and in vitro due to difficulty in accurate segmentation and association of neuronal processes with individual cells. Consequently, neuronal processes remain comparatively understudied in ALS, and it is still unknown how, and to what degree, the neuritome is affected, whether ALS-related stress insults modify this compartment or if cytoplasmic accumulation of RBPs in ALS MNs relates to other aberrant cellular phenotypes such as dystrophic neurites.

We previously generated a high-content imaging dataset of control and ALS-related VCP-mutant iPSC-derived MN cultures co-labelled with a combination of three fluorescent markers, specifically, (i) a nuclear-specific marker (DAPI), (ii) a neuron-specific marker of the neurites (β III-tubulin) and (iii) an antibody against one of five ALS-relevant RBPs: TDP-43, SPFQ, FUS, heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) or heterogeneous nuclear ribonucleoprotein K (hnRNPK).[16] In our previous study, we specifically analysed the spatio-temporal responses of the aforementioned ALS-related RBPs to different stressors (oxidative, heat and osmotic). Here, we applied deep learning methods to this rich imaging dataset to test in an automated fashion: (1) Whether aberrant cellular morphological phenotypes, including neuronal processes, associate with ALS; (2) whether these morphological phenotypes correlate to aberrant ALS-related RBP phenotypes; and (3) whether extrinsic stress insults in control MN cultures can recapitulate ALS phenotypic changes. Deep learning models such as convolutional neural networks (CNNs) are now widely used to efficiently perform image classification and image segmentation.[21–25] Such methods are able to analyse images without prior image segmentation, feature selection or human-directed training and automatically extract features from raw data, removing significant bias from this process. Importantly, CNN-based image classifier performance largely depends on whether sufficient information is contained in the provided set of images. DAPI and β III-tubulin capture complementary and non-overlapping information related to the nuclear shape and neuronal/neurite morphology, respectively. We hypothesised that comparing the performance of different classifiers trained with iterative combinations of fluorescent images can be used to identify which cellular compartment or specific RBP is most affected between any two given culture conditions. Additionally, we hypothesised that similar phenotypes between different MNs culture conditions can be quantified using the trained model predictions. We demonstrate the utility of this approach, which enables the discovery of novel phenotypes in ALS MN cultures and the identification of the relevant extrinsic stress condition that best approximates ALS pathogenesis. The advantage of our method is that it is highly versatile and can quickly guide the scientist towards the most promising hypothesis for further experimental validation. By providing our fluorescence microscopy raw images together with open-source implementations of the methods and trained models, we aim to allow other researchers to readily apply these methods and test additional hypotheses. In summary, we propose the use of deep learning methods to leverage the power of large image databases from ALS-related MN cultures to generate testable biological hypotheses automatically and rapidly, a method that could prove transformational in promoting innovative research directions, diagnostics and therapies.

Key points

- CNN-based image classifiers enable automatic detection of phenotypic changes in unsegmented images.
- Novel neurite-associated morphological perturbations in (ALS-causing) VCP-mutant human motor neurons.
- Heat stress recapitulates key aspects of ALS.

RESULTS

Repurposing image-based deep learning methods to test biological hypotheses

We previously studied the spatio-temporal responses of ALS-related RBPs to different stressors in control vs ALS-related VCP-mutant iPSC-derived electrically immature MN cultures using image-based analysis (Figure 1A and Table S1).[16] These MN cultures have been generated using our previously published protocol for the generation of highly enriched spinal cord MNs[20] and have been further characterised by the presence of MN-specific markers choline acetyltransferase (ChAT) and SMI-32 (Figure S1A). These cultures were immunolabelled after 1 h of exposure to oxidative stress, heat stress and osmotic stress, along with recovery timepoints from heat stress (2 h) and osmotic stress (1, 2 and 6 h). A combination of three specific markers was used: a nuclear marker (DAPI), a neuronal marker allowing precise identification of neurites (β III-tubulin) and an antibody against one of the following RBPs: TDP-43, SPFQ, FUS, hnRNPA1 or hnRNPK. Using this approach, we generated a large-scale imaging dataset of 156,577 images, which is publicly available in the Image Data Resource (IDR) (Figure 1B–D). In our previous study, we focused on nuclear-to-cytoplasmic ratio measurements of the aforementioned RBPs. Here, we aimed to capitalise on the richness of information contained within this high-dimensional image dataset to

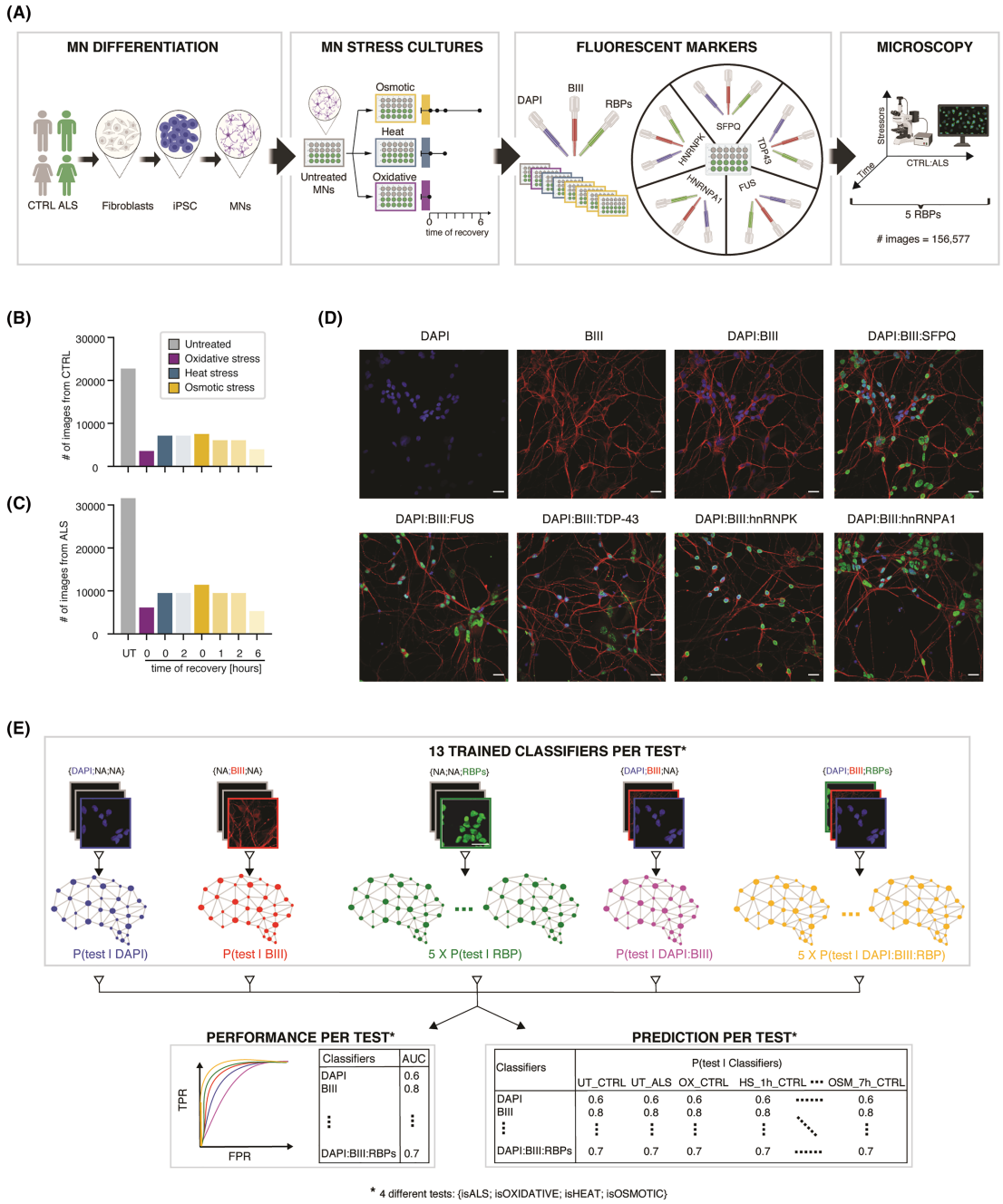

FIGURE 1 Overview of the high-dimensional immunofluorescent image dataset and paradigm to evaluate the relevance of markers in stress and ALS pathogenesis. (A) Experimental design for obtaining immunofluorescence microscopy images of motor neurons (MNs). Control ($n = 3$ cell lines) and VCP-mutant ($n = 4$ cell lines) induced pluripotent stem cells (iPSC)-derived MNs in different cellular stress (untreated, osmotic, heat and oxidative) and stress recovery (2 h after heat stress, 1, 2 and 6 h after osmotic stress) conditions were fluorescently labelled with DAPI, β III-tubulin (BIII) and key ALS-linked RNA-binding proteins (RBPs) and then imaged, resulting in 156,577 images. (B,C) Total number of images in CTRL and ALS cell lines grouped by stress conditions: untreated (UT) in grey, oxidative (OX) in purple, heat (HS, 2-h recovery) in blue and osmotic (OSM, 1-h recovery, 2-h recovery and 6-h recovery) in yellow. (D) Representative images of nuclear marker DAPI, neuronal marker β III-tubulin and ALS-linked RBPs in iPSC-derived motor neurons. Scale bars = 25 µm. (E) Fifty-two CNN-based classifiers have been trained in this study to discriminate (1) ALS from control MN cultures (isALS test), or untreated MN cultures from MN cultures exposed to (2) oxidative (isOXIDATIVE test), (3) heat (isHEAT test) and (4) osmotic stress (isOSMOTIC test) using 13 combinations of RGB images composed of different channels: Either a single channel was used (DAPI, BIII or RBP), either two channels (DAPI:BIII) or three channels (DAPI:BIII:RBP), and pitch-black images were assigned to the unused channels (Table S2). For each of the four tests (siALS, isOXIDATIVE, isHEAT and isOSMOTIC), the 13 classifiers' performance as obtained from the area under the receiver operating characteristic curve (AUC) were extracted and compared to uncover the importance of those markers in discriminating two conditions (Table S3). Additionally, 13 model predictions for each of the 4 tests have been extracted for each MN culture (Tables S4–S7)

test whether different MN stressors (including extrinsic stressors and endogenous ALS-causing mutations in the VCP gene) are characterised by detectable phenotypes in cellular compartments and/or RBP fluorescent images. Specifically, we hypothesised that ALS-related phenotypic changes will be recapitulated by one of our aforementioned stress conditions. CNN-based classifiers are powerful deep learning models that can be trained to discriminate images from different conditions by identifying complex relationships between pixels. Here, we trained the following 52 CNNs-based classifiers to recognise cellular phenotypes associated with (i) ALS, (ii) oxidative stress, (iii) heat stress or (iv) osmotic stress using 13 different combinations of immunolabelled images, ranging from DAPI fluorescent images only to the combination of three channels, that is, DAPI, β III and an RBP (Figure 1E and Table S2): 13 classifiers trained to discriminate untreated ALS from untreated control MN cultures (hereafter called *ALS classifiers*), 13 classifiers trained to discriminate untreated control MN cultures from control MN cultures exposed to oxidative stress (hereafter called *OX classifiers*), 13 classifiers trained to discriminate untreated control MN cultures from control MN cultures exposed to heat stress (hereafter called *HS classifiers*) and 13 classifiers trained to discriminate untreated control MN cultures from control MN cultures exposed to osmotic stress (hereafter called *OSM classifiers*). The CNN-based classifiers were obtained through transfer learning from MobileNetV2, which has been pre-trained using the ImageNet dataset.[21] The performance of each classifier was evaluated using the total area under the receiver operating characteristic (ROC) curve (AUC). AUC was calculated using 10-fold cross-validation, training on 90% of the dataset, testing on the remaining 10% of the dataset and repeating with 10 different train/test combinations (Table S3). The 52 trained classifiers assigned class (e.g., ALS and stress) probabilities for all the ~10 views from each cell culture (control vs ALS; untreated vs stressed; different timepoints) that were then averaged to obtain a final per culture classification probability (Tables S4–S7 and Figure S1B).

Noting that the information content of images determines the performance of a CNN-based classifier to discriminate between conditions, we harnessed distinct fluorescent markers (DAPI, β III-tubulin and RBPs) to capture different cryptic attributes that reveal cellular state. Against this background, we propose that the performance of the 13 different classifiers trained to identify a specific MN culture condition can reveal the relevant cellular compartment or RBP. During training, a classifier learns to identify a phenotype associated with a specific MN culture condition (ALS vs control; stressed vs untreated). Therefore, we propose that, once trained, this classifier can be used to predict whether similar phenotypes are shared among different conditions. Consequently, we use image-based deep learning methods in two novel ways that are expected to greatly facilitate and accelerate the process of hypothesis testing in biology. In the following sections, we first validate our approach by recapitulating previous findings. We next specifically demonstrate the utility of this approach by testing the following hypotheses: (1) ALS-causing VCP mutations result in previously unrecognised phenotypes contained within the information content of DAPI and/or β III-tubulin fluorescence images alone; (2) addition of ALS-related RBPs immunofluorescence images will improve phenotype detection in an RBP-specific manner; and (3) conventional extrinsic stressors can recapitulate phenotypic aspects of ALS.

Post-stress recovery of RBP- and neuritome-related phenotypes are closely correlated

Cell shape and morphology are recognised readouts of cell state or phenotype.[18] Here, we first sought to test whether oxidative, heat or osmotic stress are characterised by changes in cell shape. We analysed and compared the performance of CNN-based models trained to discriminate images of untreated control MN cultures from images of stress-treated control MN cultures either using the DAPI staining only (hereafter named *stress|DAPI* classifier), either the β III-tubulin immunolabelling only (hereafter named *stress|BIII* classifier) or the combination of two fluorescent markers (hereafter named *stress|DAPI:BIII* classifier). As shown in Figure 2A, *stress|DAPI* and *stress|BIII* classifiers outperform a random classifier ($AUC_{DAPI} > 0.5$ and $AUC_{\beta III} > 0.5$) across all stress conditions, indicating that DAPI and β III-tubulin fluorescent images capture relevant information related to stressed MNs. This supports the hypothesis that MNs stressed with oxidative, osmotic or heat exhibit previously unrecognised phenotypic changes in both nuclear and neuritome compartments. Furthermore, comparing the performance of these classifiers revealed that the β III-tubulin immunolabelling consistently leads to a significantly higher performance across all three stress conditions, suggesting that the compartment most affected by all three stressors is the neuritome. Although we cannot rule out the possibility that the increase in model performance between *stress|DAPI* and *stress|BIII* is due to the larger surface occupied by the neuronal processes compared with the nuclei, the minor increase in model performances across the three stress conditions between *stress|BIII* and *stress|DAPI:BIII* supports the hypothesis that DAPI-stained images are not major contributors in these classifiers. The greater performance of *osm|DAPI* compared with *ox|DAPI* and *heat|DAPI* finally suggests larger nuclear-related changes upon osmotic stress compared with the other stress insults.

We previously showed that different stressors affect the localisation of ALS-associated RBPs in control MNs.[16] Thus, we next aimed to test whether our image-based deep learning approach could shed light on the most relevant RBPs to each stress condition in order to replicate these previous findings. Several ALS-causing mutations occur in genes that encode RBPs, including TDP-43, FUS and hnRNPA1,[26–28] which typically exhibit subcellular mislocalisation of RBPs. Indeed, TDP-43 is mislocalised from the nucleus to the cytoplasm in the vast majority of cases.[29] More recently, we reported widespread SFPQ and FUS mislocalisation in various ALS models.[2,3] Furthermore, additional RBPs have been shown to be mislocalised in models of ALS, including hnRNPK, one of the most abundant hnRNPs.[30] Examining the performance of the *stress|RBP* models to discriminate untreated MN cultures from those exposed to osmotic, heat or oxidative stress and comparing these with the performance of

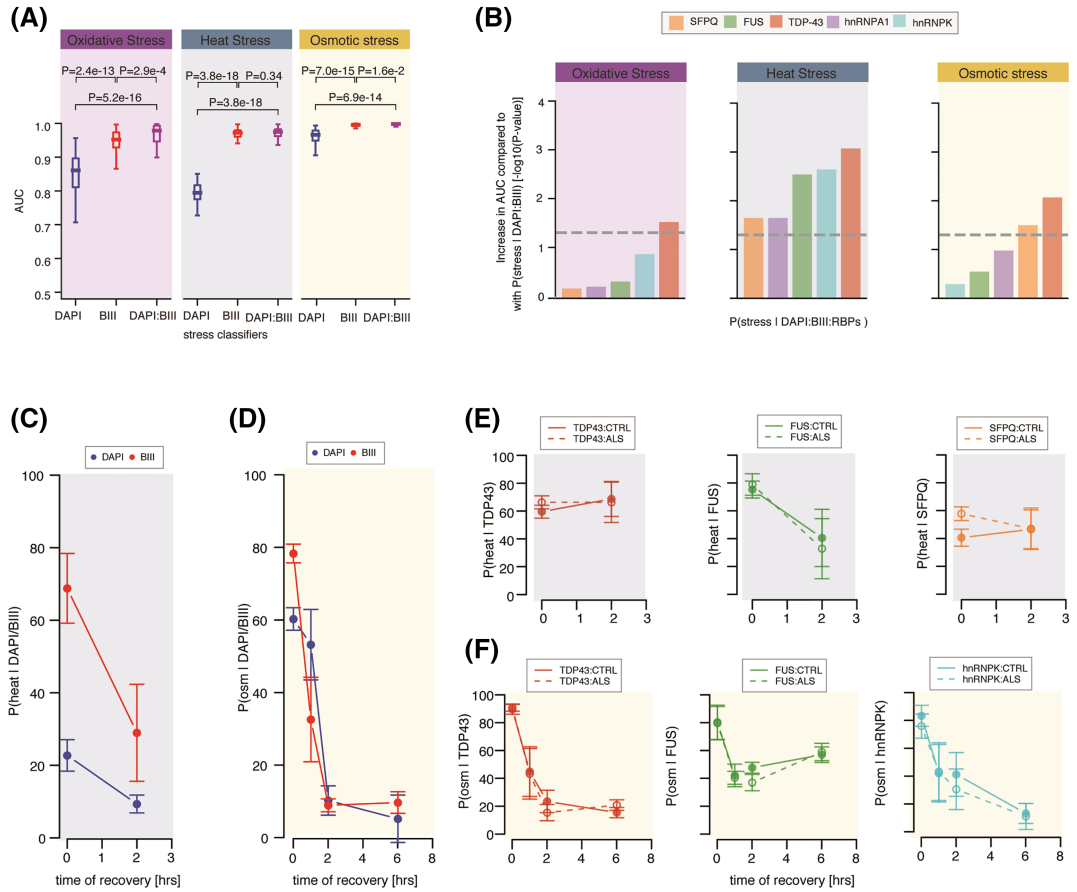

FIGURE 2 Different kinetics of stress recovery related to distinct cellular changes captured by CNN-based classifiers. (A) Boxplots showing the distributions of model performances as evaluated using the AUC for classifiers trained using DAPI, β III-tubulin (BIII) and the combination of DAPI and β III-tubulin markers, respectively, to discriminate untreated from stressed MN cultures. Each classifier was submitted to 10-fold cross-validation in 5 different subsets of the data, resulting in 50 points per classifier. Boxplots display the five-number summary of median, lower and upper quartiles, minimum and maximum values. *P*-values obtained from a one-sided Mann–Whitney test. (B) Bar graphs representing the increase in performance as obtained from −log10(*P*-values) of one-sided Mann–Whitney test comparing the AUCs from the *stress*|*DAPI:BIII* classifier and the AUCs from individual *stress*|*DAPI:BIII:RBP* classifiers for oxidative, heat and osmotic stresses. (C) Effect size (mean ± standard errors) of *heat*|*DAPI*-based (blue) and *heat*|*BIII*-based (red) classifier predictions of control MN cultures after 1 h of exposure to heat stress and 2 h of recovery from heat stress. Effect size of the treatment at each timepoint is obtained using linear mixed effects analysis accounting for idiosyncratic variations due to cell lines and experiment bias. (D) Effect size (mean ± standard errors) of *osm*|*DAPI* (blue) and *osm*|*BIII* (red) classifier predictions of control MN cultures 1 h after osmotic stress and 1, 2 and 6 h after recovery from osmotic stress. (E) Same as (C) for *heat*|*TDP43*, *heat*|*FUS* and *heat*|*SFPQ* classifier predictions. Solid lines = control MN cultures. Dashed lines = VCP-mutant MN cultures. (F) Same as (D) for *osm*|*TDP43*, *osm*|*FUS* and *osm*|*SFPQ* classifier predictions. Solid lines = control MN cultures. Dashed lines = VCP-mutant MN cultures

stress|*DAPI* models revealed significantly higher performances of all five *stress*|*RBPs* models compared with *stress*|*DAPI* irrespective of the stress (Figure S2A,B). This result indicates that, although these RBPs mostly localise to the nucleus,[16] their respective fluorescent images carry information beyond nuclear shape or texture as identified by DAPI. We also find that *stress*|*TDP43* exhibits the highest AUC across the three stressors. We next compared the performance of the *stress*|*DAPI:BIII:RBPs* models with the *stress*|*DAPI:BIII* models in each stress condition in order to test whether the integration of RBPs fluorescent images together with those of DAPI and β III-tubulin enables the identification of additional stress-related phenotypes. This analysis revealed that TDP-43 significantly increases the ability of the classifier to identify MN cultures under oxidative stress (Figure 2B). While CNN-based models are not suited to specifically address the subcellular localisation of RBPs, our finding that TDP-43 images, in conjunction with nuclear and neurite fluorescent markers, enable relevant oxidative-stress-related phenotypic information to be captured suggests that TDP-43 exhibits changes in localisation upon oxidative stress. This result is consistent with our prior finding that TDP-43—but not the other four RBPs analysed—exhibits a reduction in nuclear-to-cytoplasmic ratio upon oxidative stress.[16] We also find that all five *stress*|*DAPI:BIII:RBPs* perform significantly better than *stress*|*DAPI:BIII* to discriminate untreated from heated MN cultures. Furthermore, we find that the most informative RBPs to heat stress are TDP-43, FUS and hnnRNPK. While in our previous study we detected significant reduction in nuclear-to-cytoplasmic ratio for TDP-43 and FUS upon

heat stress, we can speculate that the present approach captures more subtle changes beyond the previously studied cellular relocalisation that could explain the detected relevance of hnRNPK to heat stress. Finally, while we previously found that all five RBPs exhibit nuclear-to-cytoplasmic relocalisation upon osmotic stress, here, we find that TDP-43 and SFPQ immunolabelling only contributes to significantly increase the *stress|DAPI:BIII* performance to identify MN cultures under osmotic stress. It is however important to note that *osm|DAPI:BIII* exhibits an AUC of ~1.0, implying that a significant improvement is difficult to achieve in this case and that, in the case of osmotic stress, this analysis may underestimate the contribution of the RBP immunolabelling. Altogether, these results indicate that the performance of a classifier is a reliable approach to prioritise which RBPs are most relevant to a specific cell culture condition.

Next, the extent of recovered cellular compartment- and RBP-related phenotypes after heat and osmotic stress were assessed using linear mixed effects analyses of the individual classifier predictions, accounting for idiosyncratic variations due to either individual cell lines or experiments. As shown in Figure 2C, 2 h after recovery from heat stress, the nuclear compartment has fully recovered, as predicted by *heat|DAPI*, while the neuritome compartment still exhibits some degree of aberrant phenotype, as predicted by *heat|BIII*. As opposed to heat stress, the nuclear compartment takes longer to recover after osmotic stress compared with the neuritome compartment; however, both compartments exhibit full recovery 6 h after treatment (Figure 2D). Next, looking at the RBP-related phenotypes, we find large heterogeneity in their predicted recovery pattern after both heat and osmotic stresses, with no complete recovery for any of the analysed RBPs 2 h after heat stress (Figure S2C,D) and long-term effects for several RBPs after osmotic stress (Figure S2E,F). In particular, we find that 2 h after heat stress, MN cultures still exhibit high *heat|TDP-43* and *heat|SFPQ* model predictions and lower (albeit still elevated) *heat|FUS* model prediction (Figure 2E). The results indicate that the TDP-43- and SFPQ-related phenotypes are still present at this stage and that the FUS-related phenotype is only partially resolved, partly reflecting on our previous study, where we did not detect reconstitution of nuclear TDP-43 and FUS to basal levels following 2 h of recovery from heat stress.[16] Our previous study also revealed slower nuclear relocalisation dynamics for TDP-43 and FUS after osmotic stress, with FUS exhibiting exceptionally aberrant nuclear-to-cytoplasmic distribution as long as 6 h post-stress.[16] Here, we find that TDP-43-related phenotype is fully resolved 2 h after treatment while FUS-related phenotype is not resolved 6 h after treatment (Figure 2F). We also find delayed hnRNPK-related phenotype recovery. Notably, we find that the recovery kinetics for most RBPs after both heat and osmotic stresses correlate over time with the neuritome-related phenotype, suggesting that changes in neuritome relate to change in RBP-related phenotype or vice versa. Finally, and in line with our previous study, we do not find any major difference between control and ALS-related VCP-mutant MNs cultures in their response to stress (Figure S2D,F). While these results at least in part recapitulate our previous findings, thereby confirming the validity of our approach, it is important to note that the trained classifiers do not necessarily capture a phenotype related to the previously studied nuclear-to-cytoplasmic relocalisation and that this may relate to more complex cellular response. Altogether, these results indicate that the performance of a classifier is a reliable approach to prioritise which RBPs are most relevant to a specific cell culture condition and that the CNN-based method can, at least in part, reproduce previous results showing slower TDP-43 and FUS relocalisation dynamics following heat and osmotic stress[16] that in some cases these might relate to changes in neuritome.

Heat stress-related changes in the MN neuritome resemble those occurring in ALS

We previously reported common morphological descriptors that strongly discriminate ALS from control tissue at the single cell level,[19] indicating that key information related to ALS cellular state might be contained in cellular shape. Having found that our approach is suitable to reproduce prior findings related to stress in MNs, we next sought to test whether ALS-related VCP-mutant MNs are characterised by changes in cellular shape in the nucleus, neuritome or the combination of both. This was achieved by comparing the performances of CNN-based classifiers trained to discriminate images of untreated control MN cultures from images of VCP-mutant MN cultures either using the DAPI staining only (hereafter named *ALS|DAPI* classifier), either the β III-tubulin immunolabelling only (hereafter named *ALS|BIII* classifier) or the combination of two markers (hereafter named *ALS|DAPI:BIII* classifier). As shown in Figure 3A, both *ALS|DAPI* and *ALS|BIII* classifiers outperform a random classifier ($AUC_{DAPI} > 0.5$ and $AUC_{\beta III} > 0.5$), indicating that, similarly to stress-related conditions, both compartments exhibit phenotypic changes associated with VCP mutation. Further comparing the performances of these classifiers revealed that the inclusion of the β III-tubulin immunolabelling leads to consistently significantly higher performance ($AUC_{DAPI+\beta III} = 0.85 > AUC_{\beta III} = 0.83 > AUC_{DAPI} = 0.65$), indicating that the compartment most affected by VCP mutation at this early disease stage is the neuritome. Notably, the consistently high model predictions among the four mutant cell lines according to both classifiers confirm the absence of experimental or cell line bias, as further confirmed by linear mixed model analysis (Figures 3B and S3A). Additionally, the significant correlation between the *ALS|DAPI* and *ALS|BIII* model predictions for each MN culture further indicates that the ALS-related changes identified by these classifiers in the nucleus or the neuritome respectively co-occur in the same MNs cultures (Pearson correlation coefficient = 0.67 and $P = 2.93e\text{-}33$; Figure 3C). Notably applying these classifiers to images from *TARDBP* G298S mutant MNs further validated that the neuritome compartment exhibits phenotypic changes when compared with control MNs (Figure S3B).

We next aimed to understand what information is used by these ALS classifiers to discriminate images from control and VCP-mutant MN cultures. Integrated gradient (IG) is one popular approach for CNN model interpretation enabling the visualisation of the relevant

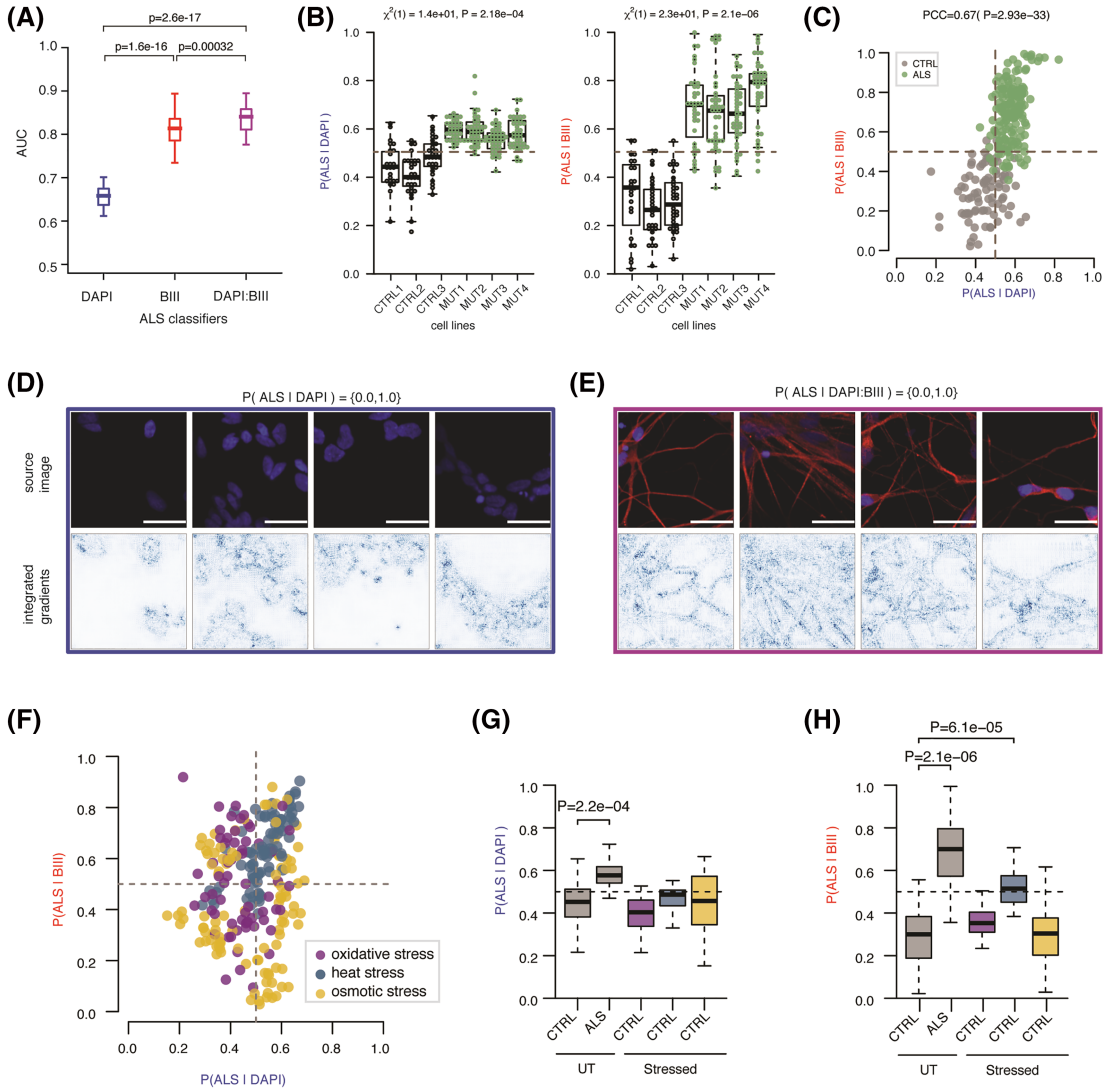

FIGURE 3 Heat stress-related changes in the MN neuritome resemble those occurring in ALS. (A) Boxplots showing the distributions of model performances as evaluated using the AUC for classifiers trained using DAPI, β III-tubulin (BIII) and the combination of DAPI and β III-tubulin markers to discriminate control from VCP-mutant MN cultures. Each classifier was submitted to 10-fold cross-validation in 5 different subsets of the data, resulting in 50 points per classifier. *P*-values are from a one-sided Mann–Whitney test. Data shown as in Figure 2A. (B) Distributions of the *ALS|DAPI* (*left*) and *ALS|BIII* (*right*) model predictions for the individual MN cultures originating from three control cell lines (*grey dots*) and four VCP-mutant cell lines (*green dots*). Linear mixed effects analysis of the relationship between each model prediction and VCP mutation to account for idiosyncratic variation due to cell line or experiment differences. VCP mutation significantly increases *ALS|DAPI* predictions, $\chi^2(1) = 14$ and $P = 2.28e^{-04}$, by about 0.13 ± 0.023 (standard errors), and *ALS|BIII* predictions, $\chi^2(1) = 23$ and $P = 2.1e^{-06}$, by about 0.4 ± 0.033 (standard errors). (C) Scatter plot of *ALS|DAPI* and *ALS|BIII* model predictions on individual control and VCP-mutant MN cultures. Grey = control MN cultures. Green = VCP-mutant MN cultures. PCC = Pearson correlation coefficient. (D,E) Randomly selected images with high model prediction according to the *ALS|DAPI* (D) and *ALS|DAPI:BIII* (E) classifiers. Images shown are the original image (*upper*) and the corresponding attribution magnitude image overlayed on the original image (*lower*). The magnitudes range from 0 (white), indicating no contribution of the pixel, to 1 (blue), indicating the strongest contribution of the pixel to the model prediction. Scale bars = 25 µm. (F) Scatter plot of the *ALS|DAPI* and *ALS|BIII* model predictions on individual control MN cultures 1 h after oxidative, heat and osmotic stress. Magenta = MN cultures 1 h after oxidative stress. Blue = MN cultures 1 h after heat stress. Yellow = MN cultures 1 h after osmotic stress. (G) Boxplots showing the distributions *ALS|*DAPI model predictions on untreated control and VCP-mutant MN cultures, and control MN cultures 1 h after oxidative, heat and osmotic stress. Magenta = MN cultures 1 h after oxidative stress. Blue = MN cultures 1 h after heat stress. Yellow = MN cultures 1 h after osmotic stress. Stress treatment effect analysis on model prediction obtained using linear mixed effects analysis. *P*-value is indicated when significant. (H) Same as (G) for *ALS|*BIII model

pixels for a specific image that contribute to its classification.[31] Looking at the IGs of randomly selected images with high *ALS|DAPI* model predictions showed relevant pixels mostly overlap with the outline of the nuclei, with some contribution from pixels located inside the nuclei (Figure 3D). This indicates that the ALS-related phenotype identified by the *ALS|DAPI* classifier primarily relates to nuclear

shape (including the size) rather than to other DAPI-related measurements such as texture or intensity. Next, looking at the IGs of randomly selected images with high *ALS|DAPI:BIII* model predictions showed relevant pixels primarily located at the edges of the neurites, indicating that relevant information mostly arises from the outline of the neurites rather than from the texture or the intensity of the β III-tubulin immunolabelling (Figure 3E). Altogether, these results indicate the network of neurites carries most ALS-related phenotype information.

Mitochondrial and oxidative stress are recognised and robust phenotypes in ALS development in vitro, and thus, in vitro models of cellular stress are important tools to investigate ALS.[17] However, it remains unknown which type of cellular stress is most physiologically relevant to study ALS pathogenesis. Thus, we next sought to test whether heat, osmotic or oxidative extrinsic stress insults induce similar nuclear and/or neuritome-related phenotypic changes in control MNs cultures as those captured by *ALS|DAPI* and *ALS|BIII* classifiers. First, looking at the scatter plot of these two model predictions for individual MN cultures colour-coded according to the stress conditions showed that heated MN cultures consistently score high according to both classifiers forming a coherent cluster, as opposed to oxidative and osmotic conditions (Figure 3F). We next quantified the extent of ALS-related nuclear vs neuritome phenotypes induced by each individual treatment (oxidative, heat and osmotic) by analysing the *ALS|DAPI* or *ALS|BIII* model predictions of individual stressed MN cultures using linear mixed modelling (see Section 4). This analysis revealed that heat stress, and to some extent osmotic stress, induces a minor however non-significant ALS-related nuclear phenotypic change in control MN cultures, as indicated by the increase in *ALS|DAPI* model predictions for these two MNs cultures (Figures 3G and S3C). This is in contrast with the neuritome compartment which *ALS|BIII* model predicts significant phenotypic changes in heated MNs cultures only (Figures 3H and S3D). Altogether, these results indicate that the neuritome is the compartment most affected by ALS-related VCP mutation and that heat stress induces similar neurite-associated changes in control MN cultures.

FUS immunolabelling best captures ALS-related phenotypes that are recapitulated by heat stress

Two thirds of RBPs are expressed in a cell type-specific and temporally regulated manner.[32] While previous studies showed RBP-specific mislocalisation in various ALS models,[2–4,33,34] it remains unknown which RBP is most predictive of ALS at a particular disease stage. Thus, we next sought to test whether different ALS-related RBPs capture distinct ALS-related information in VCP-mutant MN cultures by analysing the model performances of ALS classifiers trained with images immunolabelled with antibodies against TDP-43, FUS, SFPQ, hnRNPK or hnRNPA1 only (hereafter named *ALS|RBPs* classifiers). Because all five aforementioned RBPs exhibit predominant nuclear localisation,[16] we first tested whether *ALS|RBPs* classifiers exhibit significant improvement compared with the *ALS|DAPI* classifiers. This analysis showed that all five *ALS|RBP* classifiers outperform *ALS|DAPI* classifier ($AUC_{ALS|RBPs} > AUC_{ALS|DAPI}$), ruling out the possibility that the phenotypic changes captured by these classifiers simply overlap with those identified by the *ALS|DAPI* classifier and indicating that they identify ALS-related phenotypes beyond changes in the nuclear shape (Figure 4A). Comparing their individual performances further revealed large differences in the individual RBP-based classifiers' ability to discriminate ALS from control MN cultures, with *ALS|TDP-43* exhibiting the best performance and *ALS|SFPQ* the least ($AUC_{ALS|SFPQ} = 0.7 < AUC_{ALS|hnRNPK} = 0.73 < AUC_{ALS|FUS} = 0.79 < AUC_{ALS|hnRNPA1} = 0.85 < AUC_{ALS|TDP43} = 0.9$). Examining the IGs for randomly selected images with high *ALS|RBP* model predictions indicated that the relevant pixels in all five *ALS|RBPs* classifiers are excluded from the nuclear areas as opposed to the most relevant pixels of the *ALS|DAPI* classifier that are most commonly localised at the inner nuclear membrane or inside the nucleus (Figure 4B). This demonstrates that the better the performance of the classifier, the less relevant the intranuclear pixels. For example, relevant pixels in the *ALS|TDP-43* classifier are fully excluded from the nuclear area. Altogether, these results suggest that the different performance of *ALS|RBPs* classifiers in identifying ALS MNs cultures result from distinct RBPs localisation rather than nuclear shape. We previously showed that considering DAPI and β III-tubulin together significantly increases the performance of both *ALS|DAPI* and *ALS|BIII* classifiers. We next analysed how the performance of the ALS classifier would change by adding a third channel composed of RBPs immunolabelling and compared the performances of the *ALS|DAPI:BIII* classifier with the five *ALS|DAPI:BIII:RBP* classifiers. This analysis indicates that all five RBPs significantly increase the ability of the trained CNN to discriminate VCP mutant from control MN cultures; however, while hnRNPK leads to the most modest improvement, TDP-43 and FUS immunolabelling lead to the highest increases in classification performance (Figure S4A). These results support the hypothesis that TDP-43 and FUS exhibit changes in localisation that are detected by these classifiers. Indeed, our prior studies have shown that in the same experimental model used here and at the same development stage, TDP-43 and FUS mislocalisation already occur[16,20,35]; thus, we can expect similar RBP mislocalisation. These results further suggest that all five RBPs exhibit mislocalisation at different degrees, which previous studies could not demonstrate possibly due to lower sensitivity.

We next sought to test whether the ALS-related changes identified by the individual *ALS|RBPs* models are recapitulated in any of the extrinsic stressor cultures conditions. We first considered the effect size, as obtained by linear fixed effect analysis (Section 4), of each extrinsic stressor culture condition on the predictions of the seven ALS classifiers (ALS|DAPI, ALS|BIII, *ALS|hnRNPK*, *ALS|SFPQ*, *ALS|FUS*, *ALS|hnRNPA1* and *ALS|TDP-43*) trained using a single channel. This revealed that oxidative and osmotic stress recapitulate only one or two ALS-related phenotypes out of the seven captured by the ALS classifiers. However, heat stress induced more than a 10% increase in ALS prediction across five out of seven classifiers (Figure 4C). *ALS|TDP-43* classifier, which performs best in ALS MN classification, is

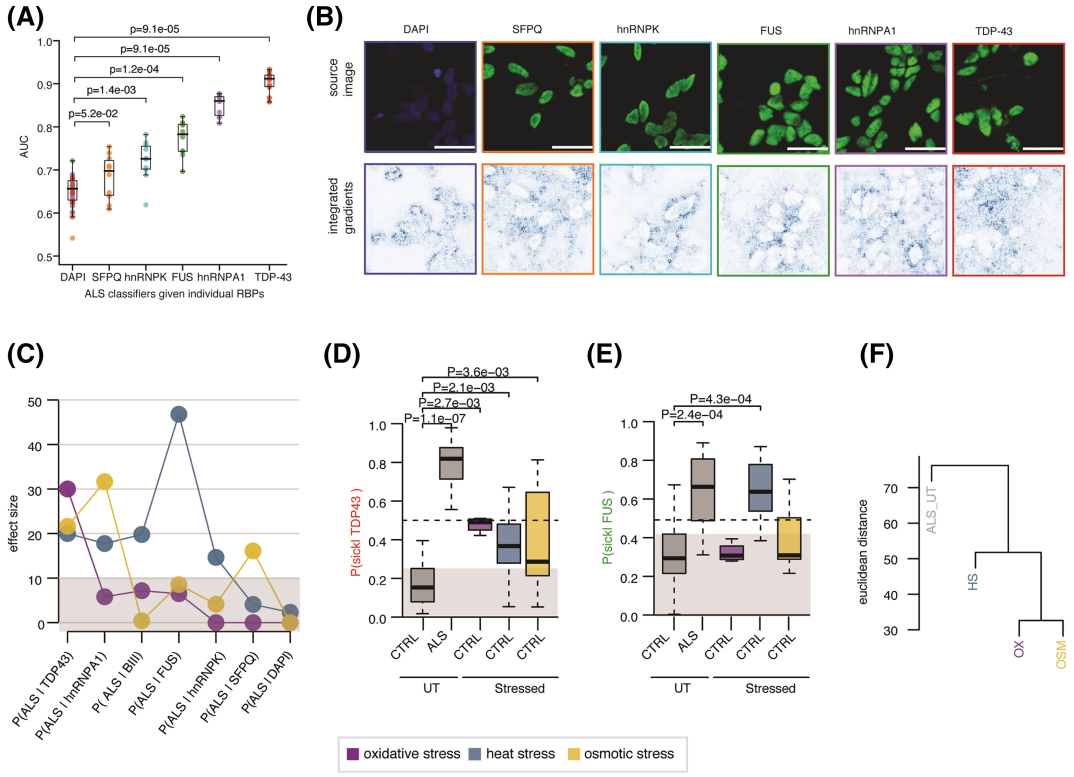

FIGURE 4 Heat stress as the most physiologically relevant stress condition to study ALS pathogenesis. (A) Boxplots showing the distributions of model performances as evaluated using the AUC for classifiers trained using the ALS-related RBPs markers SFPQ, hnRNPK, FUS, hnRNPA1 and TDP-43 to discriminate control from VCP-mutant MN cultures, and compared with *ALS|DAPI* classifier performance. Each classifier was submitted to 10-fold cross-validation in 5 different subsets of the data, resulting in 50 points per classifier. *P*-values are from a one-sided Mann–Whitney test. Data shown as in Figure 2A. (B) Randomly selected images with high model prediction according to the *ALS|DAPI* and *ALS|RBPs* classifiers. Images shown are the original image (*upper*) and the corresponding attribution magnitude image overlayed on the original image (*lower*). The magnitudes range from 0 (white), indicating no contribution of the pixel, to 1 (blue), indicating the strongest contribution of the pixel to the model prediction. Scale bars = 25 μm. (C) Comparisons of the effect sizes of each stress treatment on control MNs cultures *ALS|DAPI*, ALS|BIII and the five *ALS|RBPs* model predictions 1 h after treatment. Magenta = oxidative stress. Blue = heat stress. Yellow = osmotic stress. Stress treatment effect analysis on model prediction obtained using linear mixed effects analysis. (D) Boxplots showing the distributions of *ALS*|TDP-43 model predictions on untreated control and ALS MN cultures, and control MN cultures 1 h after oxidative, heat and osmotic stress. Magenta = oxidative stress. Blue = heat stress. Yellow = osmotic stress. Stress treatment effect analysis on model prediction obtained using linear mixed effects analysis. *P*-values obtained from linear mixed models are indicated when significant. (E) Same as (D) however for *ALS*|FUS. (F) Unsupervised hierarchical clustering of treatment effect sizes on *ALS|DAPI*, ALS|BIII and the five *ALS|RBPs* model prediction groups. Oxidative and osmotic stresses cluster together while heat stress is the closest to the ALS group. Euclidean distance and ward clustering

indeed the unique model that leads to significant, however modest, model prediction across the three stress conditions (Figures 4D and S4B,C). Additionally, we find that heat stress is the unique condition that leads to similarly high disease model prediction in control MN cultures and in untreated VCP-mutant MN cultures, given the *ALS|FUS* and *ALS|hnRNPK* classifiers (Figures 4E and S4B,C). Hierarchical clustering of the untreated ALS MN cultures together with the three stress conditions according to the effect size of each classifier (Euclidean distance and Ward clustering) eventually confirmed that heat stress induces overall the most similar cellular changes to ALS (Figure 4F). Altogether, these results confirm that MNs exposed to heat stress most closely resemble ALS cells with respect to phenotypes captured by the majority of ALS classifiers. Additionally, it shows that while TDP-43 is the RBP that carries the strongest information related to ALS, it is the FUS immunolabelling that captures most similar phenotypes between heat stress and ALS MNs cultures.

DISCUSSION

In this study, we combine multichannel fluorescence high-content microscopy data with deep learning imaging methods to unveil—directly from unsegmented images—novel neurite-associated morphological perturbations. This approach can be used to leverage existing high-content imaging datasets to gain new phenotypic insight into the original biological questions asked, as established by this study. We uncovered a surprising degree of previously unrecognised disease-

relevant information in broadly used and often considered 'generic' biological markers of nuclei (DAPI) and neurons (β III-tubulin) in the context of our human stem cell model of VCP-related ALS. Additionally, we reveal changes associated with ALS-related RBP immunofluorescence imaging that can be captured in VCP-mutant MN cultures. We were also able to systematically examine whether heat, oxidative or osmotic stress induce similar modifications that could therefore reinforce their utility in modelling aspects of MN dysfunction in ALS. Our study establishes the use of CNN-based methods for rapid, automated and unbiased testing of biological hypotheses.

CNN-based methods are now widely used for image classification and segmentation and have been successfully applied to medical imaging data for disease detection and prediction.[36–38] Here, CNN-based image classifiers have been trained to identify stress- and ALS-related phenotypic changes in unsegmented images of multiple MN cultures. We further showed that the performance of such classifiers is a reliable approach to prioritise which RBPs are most relevant to a specific cell culture condition, although refined analysis will be required to interpret their precise relevance to the underlying disease/stress process. CNN-based classifiers face challenges with interpretability and are not suited to specifically address the subcellular localisation of RBPs as previously described with conventional methods based on image segmentation.[16,35] Nevertheless, we showed that the phenotypes identified in the DAPI or β III-tubulin fluorescent images are indeed contained in the outlines of the nuclei and in the edges of the neurites, respectively. Furthermore, we could demonstrate that training a classifier with fluorescent images of a given RBP, in conjunction with nuclear and neurite fluorescent markers, enables the recapture of previously found phenotypes related to RBP cellular localisation. Specifically, we could reproduce previous results showing TDP-43 and FUS mislocalisation in ALS iPSC-derived MNs.[3,20,35] Additionally, our study suggests that SFPQ, hnRNPA1 and hnRNPK also exhibit mislocalisation however at different degrees. Notably, hnRNPA1 is a component of RNA transport granules in neurons,[39] and we can speculate that the extent of cytoplasmic relocalisation for this primarily nuclear RBP [16] may be too subtle to be captured by analysing its nuclear-to-cytoplasmic ratio.

DAPI and β III-tubulin are often considered 'generic' biological markers, and the usage of their fluorescent images are most often intended for nuclear or neuronal segmentation. Our study uncovers previously unrecognised disease-relevant information that is contained within DAPI and β III-tubulin fluorescent images. Given that DAPI and β III-tubulin are broadly used markers, and given that our method can be easily implemented in other biological conditions, future studies might use DAPI and β III-tubulin fluorescent images from different biological systems and experimental paradigms to reveal innovative research directions. For example, this approach could be useful in interrogating the presence and onset of aberrant cellular morphologies in time course experiments and across neurological disorders.

Morphological attributes of cells and their substructures are recognised readouts of physiological or pathophysiological states[18]; however, these have been relatively understudied in ALS research. Here, we demonstrate that the neuritome compartment exhibits aberrant phenotypes in ALS pathogenesis, as evidenced by the high efficiency of deep learning classifier to identify ALS MN cultures uniquely based on β III-tubulin fluorescent images. We also show modest (albeit significant) perturbations in the nuclear compartment given the predictive value of the DAPI fluorescent images in identifying ALS MN cultures. While it remains unclear whether these are strictly pathogenic events, the similar phenotypes detected in the neuritome of MN cultures exposed to heat stress suggest that these events relate to a form of MN stress. Through a thorough comparison of heat, oxidative and osmotic stress-induced changes in both cellular shape and ALS-related RBP immunolabelling, we further demonstrate that neuritome-associated perturbations were also detected in control MNs cultured in three different stress conditions. These findings support the notion that the neuronal processes exhibit large perturbations across various stress conditions and argue for increased focus on this cellular subcompartment in future research, such as testing these methods on ALS human pathological tissue sections. Another striking finding is the correlation between recovery kinetics of the neuritome compartment after osmotic and heat stresses, and those of several RBP-related phenotypes. Assuming that the RBP-related phenotypes captured by the CNN-based classifier relate to an RBP change in cellular localisation, this result suggests that previously observed stress-induced RBPs mislocalisations are coupled to global changes in the neuritome.[16] Indeed, neurite degeneration has been shown to occur upon oxidative stress through the cytoplasmic sequestration of two proteins (PRMT1 and Nd1-L) in in vitro models of FUS mutant-related ALS.[40] Furthermore, TDP-43 mislocalisation and aggregation has also been demonstrated in dystrophic neurites,[1,33] while we recently reported an increase in wild-type FUS within neuronal processes in VCP-mutant MNs.[35] Finally, a regulatory role for FUS has also been shown in synaptic formation and function,[40–43] and aberrant FUS activity in the axonal compartment has been evidenced in a FUS mutant ALS mouse model.[44] Altogether, these studies support the hypothesis of an association between RBP mislocalisation and aberrant neuronal processes in ALS. The finding that several RBP-related phenotypes present similar recovery patterns as the neuritome further suggests that additional ALS-related RBPs might exhibit similar aberrant neurite localisation in ALS. Future work will directly address the nature of these perturbations using classic approaches that necessitate nuclear and neurite segmentation and the acquisition of hundreds of measurements from each cellular compartment.

Several lines of evidence support the hypothesis that cellular stress is one central mechanism by which MN death occurs in ALS and in vitro models of cellular stress are therefore important tools to investigate ALS disease.[17] It remains however unknown which type of extrinsic cellular stress most closely approximates ALS pathogenesis, and relatively little is known about the effect of thermal stimulation, hyperosmolarity or arsenite-induced oxidative stress on the neuritome compartment. Here, we find that iPSC-derived MNs exposed to heat stress, as opposed to hyperosmolarity or arsenite-induced oxidative stress, closely recapitulate the phenotypes of ALS MN cultures

captured by several classifiers. This result suggests that heat stress more closely approximates ALS pathogenesis compared with osmotic and oxidative stressors. This is in line with previous non-mammalian studies that have defined heat stress as being relevant to the study of neurodegeneration. Heat stress-induced stress granules sequester more misfolded proteins, are less dynamic and have increased protein poly-ubiquitination compared with arsenite-related stress granules.[45–48] The lack of similarity between oxidative stress and ALS here may be attributed to the use of arsenite, which although widely used as an inducer of oxidative stress, may not truly approximate physiological oxidative stress that has been reported in ALS. As the pathogenic cascades underlying ALS are multifactorial and not fully determined, heat stress could indeed be a useful cellular model to study disease mechanisms. Our study demonstrates the ability of heat stress to induce subtle ALS-related cellular changes associated within the neuritome compartment and within the FUS immunofluorescent images. In particular, our study demonstrates the ability of heat stress to induce subtle ALS-related cellular changes associated within the neuritome compartment and within the FUS fluorescent images. Interestingly, we previously found that heat stress alone caused cell death in an iPSC-derived model of MNs.[16] Furthermore, heat stress,[49,50] ageing and neurodegeneration[51–53] all associate with intron retention.[2,4,54,55] Future research will investigate the molecular mechanisms by which elevated temperatures lead to similar responses as those observed in patient-specific ALS MN cultures.

MATERIALS AND METHODS

Compliance with ethical standards

Informed consent was obtained from all patients and control controls in this study. Experimental protocols were all carried out according to approved regulations and guidelines by UCLH's National Hospital for Neurology and Neurosurgery and UCL's Institute of Neurology joint research ethics committee (09/0272). Cell culture, stress treatments, immunohistochemistry and image acquisition were performed as in Harley and Patani.[16] Indeed, these data are utilised in the current manuscript and no additional experiment was required.

High-content imaging dataset

The imaging dataset used in this study consists of fluorescence microscopy images of iPSC-derived MNs as previously reported.[16] The neurons either came from control cell lines or cell lines with the ALS-related VCP mutation and underwent experimentation after 6 days of terminal differentiation. Details of iPSC lines are provided in Table S1. To induce stress, the cultures were subject to 1 h of oxidative stress, 1 h of osmotic stress and 1 h of heat stress. To examine recovery, the cultures were subject to 1 h of stress and then returned to untreated conditions for 2 h following heat stress and 1, 2 and 6 h following osmotic stress. Following stress treatments or recovery, cultures were fixed and then immunostained with a combination of three markers, specifically a nuclear-specific marker (DAPI), a neuron-specific marker allowing to outline the neurites (β III-tubulin) and an antibody against TDP-43, SPFQ, FUS, hnRNPA1 or hnRNPK. The dataset is divided in different experiments (repeats done different days), each with several 96-well plates. Each well corresponds to one cell line, one stress condition and one combination of fluorescent markers. Each well has several non-overlapping fields of view (ranging from 10 to 12), and each field of view has several planes or z-stacks (ranging from 3 to 5) with 1-µm steps, generating a large-scale imaging dataset of 156,577 images (Figure 1A,B). The dataset is publicly available in IDR and can be found under study idr0112 and using the direct link https://idr.openmicroscopy.org/webclient/?show=screen-3001.

Image pre-processing

All images went through pre-processing steps described in Figure S1. Raw images are 16-bit images. Sixteen-bit raw z-stack images (1080×1080 pixels) from the same field of view were first merged using Maximum Intensity Projection (MIP), where the pixel with maximum intensity across all z-stacks is selected at each location in the image. Following conversion of MIP images to 8-bit images, channels were merged together to form an RGB image. We created 13 types of RGB images, either composed of one, two or three channels, to train image classifiers with 13 different combinations of immunostained images (Figure 1E). For images with three channels, DAPI was assigned to blue channel, β III-tubulin to the red channel and the RBP to green channel. For images with one or two channels, pitch-black images were assigned to the remaining channels so that the image would still be considered RGB. Images were then enhanced using Python Image Library Pillow ImageOps[56] auto contrast function, to normalise image contrast. This function calculates a histogram of the input image, removes 0.1% of the lightest and darkest pixels from the histogram and remaps the image so that the darkest pixel becomes black (0) and the lightest becomes white (255). In the fourth step, the enhanced images were divided into 16 smaller images of size 270×270 pixels, which allowed better resolution and more images. Structures at this scale proved to be more distinguishable with IGs and yielded similar results than with whole images. This division also made sense for the fifth step, which consisted in resizing images to 224×224 pixels. Finally, images were normalised using mean intensity = [0.485, 0.456, 0.406] and std = [0.229, 0.224, 0.225] across the images from the ImageNet dataset. The last two steps were added in order to fulfil the requirements when using pre-trained models, which expect input images to be normalised in the same way as the dataset on which they were trained.

Data augmentation

In order to improve accuracy and reduce overfitting, we performed five augmentations on each image of the training set as follows and

as previously described[57]: (1) 90° rotation, (2) one horizontal mirror, (3) one vertical mirror, (4) 90° rotation of horizontal mirror and (5) 90° rotation of vertical mirror. This results in a sixfold increase of the number of available images for training (five rotations + original).

CNN-based image classifiers training

We trained 52 CNN-based classifiers to discriminate (1) ALS from control MNs cultures (isALS test), untreated MNs cultures from MNs cultures exposed to (2) oxidative (isOXIDATIVE test), (3) heat (isHEAT test) and (4) osmotic stress (isOSMOTIC test) using 13 combinations of RGB images composed of different channels: Either a single channel was used (DAPI, BIII or RBP), either two channels (DAPI:BIII) or three channels (DAPI:BIII:RBP), pitch-black images being assigned to the unused channels (Table S2). Instead of training a full new neural network, we performed transfer learning that can be used to address the relatively small number of available images by introducing information from another domain.[58] For training, images were fed into torchvision MobileNetV2 model, which has been pre-trained on ImageNet.[59,60] MobileNetV2 is a CNN based on a streamlined architecture that uses depth-wise separable convolutions to build lightweight deep neural networks and that is effective for fine-grained image classification. MobileNetV2 is a lightweight neural network of 3.5 million parameters, as opposed to the widely used ResNet that contained 11.7 million parameters, making it suitable for fine-tuning with limited number of images.[61] All layers of the pre-trained CNN classifier were fine-tuned on our dataset, allowing the training of a highly accurate model with a relatively small training dataset.[62] The last layer was modified so that it turned the features into predictions for two classes instead of the thousand classes from ImageNet. Training was performed by stochastic gradient descent with learning rate 0.001, batch size 32, using the cross entropy loss function. The training was stopped after 10 epochs. A 10-fold cross-validation scheme has been used to evaluate the accuracy of the classification predictions generated by the trained classifiers: The images were shuffled randomly and divided into 10 stratified folds, preserving the percentage of samples for each class; each fold was used once as a test dataset, while the remaining folds were used for the training dataset. ROC curves were generated to evaluate the model's ability to distinguish two cell culture conditions. ROC curves plot the true positive rate (sensitivity) vs the false positive rate (1 − specificity). The area under the ROC curve was used as the performance measure or classification accuracy. The classification accuracy over all folds is reported in Table S3. We evaluated the 52 trained models on all MN cultures when the right combination of markers were available (typically FUS, DAPI and BIII are available for some MNs cultures while SFPQ, DAPI and BIII are available for others). The probabilities to belong to one of the four tested conditions (iALS, isOXIDATIVE, isOSMOTIC and isHEAT) outputted by the classifiers are then aggregated by computing the average probability over all of the 16 cropped images originating from a single image, thereby obtaining a single probability per original-sized images. A single probability per MN culture is reported for each of the four tests by averaging the signal over all images (typically seven per MN culture) as reported in Tables S4–S7.

Model explainability and IG

The IG is a widely used interpretability algorithm that allows to identify what pixels of an image have the strongest effect on the model's predicted class probabilities and therefore allowing to visualise which parts or the image are important for classification,[31] by computing the gradient of the model's prediction output to its input features. We used the Captum Insights method[63] to obtain the IG for randomly selected images associated with high classifier prediction scores.

Model prediction data analysis

We used R and lme4[64] to perform linear mixed effects analysis of the relationship between individual model predictions and either stress MNs culture condition (at the corresponding time after treatment) or VCP-mutant MNs cultures, accounting for idiosyncratic variation due to either cell line or experiments. As fixed effects, we either entered the stress condition or the VCP-mutation variable into the model. As random effects, we had intercepts for either cell lines (CTRL1, CTRL2, CTRL3, MUT1, MUT2, MUT3 and MUT4) and experiments. *P*-values were obtained by likelihood ratio tests of the full model with the effect in question against the model without the effect in question, that is, comparing a full linear model fitting the classifier predictions using both the fixed effect (the stress or disease) and the random effects (cell line and experiments), with a reduced linear model, which only considers the random effects.

Data and software availability

We provide raw images, complete source code and trained models to readily reproduce figures, tables and other results that involve computation in order to facilitate the development and evaluation of additional profiling methods. The image data that support the findings of this study will be uploaded to the IDR.[65,66] As this requires some more time given the size of the data (1 TB), we have uploaded a subset of the data on Zenodo under the accession number 4664177. The 52 trained deep learning models have been uploaded on Zenodo under the accession number 4664252. The code for obtaining the models as well as the Jupyter notebooks are available at https://github.com/idiap/als-classification.

ACKNOWLEDGEMENTS

This work was supported by the Idiap Research Institute and by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC010110), the UK Medical Research Council (FC010110) and the Wellcome Trust (FC010110). R.P. holds an MRC

Senior Clinical Fellowship (MR/S006591/1). We thank Olivier Bornet, Bastien Crettol and Philip Abbet for their technical support in generating the repositories for the code, the models and the dataset.

CONFLICT OF INTEREST

The authors declare no competing interests.

ETHICS STATEMENT

Experimental protocols were all carried out according to approved regulations and guidelines by UCLH's National Hospital for Neurology and Neurosurgery and UCL's Institute of Neurology joint research ethics committee (09/0272).

AUTHOR CONTRIBUTIONS

Conceptualisation: R.L., R.P.; formal analysis: C.V., R.L.; investigation: C.V., J.H.; writing—original draft: R.L., R.P.; writing—review and editing: R.L., R.P., C.V., J.H.; resources: R.P.; visualisation: C.V., J.H.; funding acquisition: R.L., R.P.; supervision: R.L., R.P.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/nan.12770.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in IDR at https://idr.openmicroscopy.org/webclient/?show=screen-3001, reference number idr0112.

ORCID

Colombine Verzat https://orcid.org/0000-0001-5796-0924
Rickie Patani https://orcid.org/0000-0002-3825-7675
Raphaëlle Luisier https://orcid.org/0000-0002-5657-2943

REFERENCES

1. Neumann M, Sampathu DM. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science*. 2006;314(5796):130-133.
2. Luisier R, Tyzack GE. Intron retention and nuclear loss of SFPQ are molecular hallmarks of ALS. *Nat Commun*. 2018;9(1):2010.
3. Tyzack GE, Luisier R. Widespread FUS mislocalization is a molecular hallmark of amyotrophic lateral sclerosis. *Brain*. 2019;142(9):2572-2580.
4. Hogan AL, Grima N, Fifita JA, et al. SFPQ intron retention, reduced expression and aggregate formation in central nervous system tissue are pathological features of amyotrophic lateral sclerosis. *Cold Spring Harbor Laboratory*. 2020; 2020.09.22.309062 doi:https://doi.org/10.1101/2020.09.22.309062.
5. Li YR, King OD, Shorter J, Gitler AD. Stress granules as crucibles of ALS pathogenesis. *J Cell Biol*. 2013;201(3):361-372.
6. Ramaswami M, Taylor JP, Parker R. Altered ribostasis: RNA-protein granules in degenerative disorders. *Cell*. 2013;154(4):727-736.
7. Aulas A, Vande Velde C. Alterations in stress granule dynamics driven by TDP-43 and FUS: a link to pathological inclusions in ALS? *Front Cell Neurosci*. 2015;9:423.
8. Dewey CM, Cenik B. TDP-43 aggregation in neurodegeneration: are stress granules the key? *Brain Res*. 2012;1462:16-25.
9. Wolozin B. Regulated protein aggregation: stress granules and neurodegeneration. *Mol Neurodegener*. 2012;7(1):56.
10. Sama RRK, Ward CL. FUS/TLS assembles into stress granules and is a prosurvival factor during hyperosmolar stress. *J Cell Physiol*. 2013;228(11):2222-2231.
11. Dewey CM, Cenik B. TDP-43 is directed to stress granules by sorbitol, a novel physiological osmotic and oxidative stressor. *Mol Cell Biol*. 2011;31(5):1098-1108.
12. van der Houven van Oordt W, Diaz-Meco MT. The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. *J Cell Biol*. 2000;149(2):307-316.
13. Younas N, Zafar S. SFPQ and Tau: critical factors contributing to rapid progression of Alzheimer's disease. *Acta Neuropathol*. 2020;140(3):317-339.
14. Courteau L, Crasto J. Hexokinase 2 controls cellular stress response through localization of an RNA-binding protein. *Cell Death Dis*. 2015;6(8):e1837.
15. Fukuda T, Naiki T, Saito M, Irie K. hnRNP K interacts with RNA binding motif protein 42 and functions in the maintenance of cellular ATP level during stress conditions. *Genes Cells*. 2009;14(2):113-128.
16. Harley J, Patani R. Stress-specific spatiotemporal responses of RNA-binding proteins in human stem cell-derived motor neurons. *Int J Mol Sci*. 2020;21(21):8346.
17. Tyzack G, Hall C, Yao Z, Lakatos A, Gandhi S, Patani R. Cell autonomous and non-cell autonomous mechanisms of disease in VCP-related ALS. In: *13th European Meeting on Glial Cells in Health and Disease*; 2017.
18. Prasad A, Alizadeh E. Cell form and function: interpreting and controlling the shape of adherent cells. *Trends Biotechnol*. 2019;37(4):347-357.
19. Hagemann C, Tyzack GE. Automated and unbiased discrimination of ALS from control tissue at single cell resolution. *Brain Pathol*. 2021;31(4). https://doi.org/10.1111/bpa.12937
20. Hall CE, Yao Z. Progressive motor neuron pathology and the role of astrocytes in a human stem cell model of VCP-related ALS. *Cell Rep*. 2017;19(9):1739-1749.
21. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:1097-1105.
22. Hou L, Samaras D. Patch-based convolutional neural network for whole slide tissue image classification. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2016;2016:2424-2433.
23. Kraus OZ, Ba JL, Frey BJ. Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics*. 2016;32(12):i52-i59.
24. Ehteshami Bejnordi B, Veta M. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*. 2017;318(22):2199-2210.
25. Campanella G, Hanna MG. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med*. 2019;25(8):1301-1309.
26. Sreedharan J, Blair IP. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science*. 2008;319(5870):1668-1672.
27. Vance C, Rogelj B. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science*. 2009;323(5918):1208-1211.
28. Kim HJ, Kim NC. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*. 2013;495(7442):467-473.
29. Chen H-J, Mitchell JC. The heat shock response plays an important role in TDP-43 clearance: evidence for dysfunction in amyotrophic lateral sclerosis. *Brain*. 2016;139(5):1417-1432.

30. Moujalled D, Grubman A. TDP-43 mutations causing amyotrophic lateral sclerosis are associated with altered expression of RNA-binding protein hnRNP K and affect the Nrf2 antioxidant pathway. *Hum Mol Genet*. 2017;26(9):1732-1746.
31. Sundararajan M, Taly A, Yan Q. Axiomatic attribution for deep networks. arXiv [csLG] (2017).
32. McKee AE, Minet E. A genome-wide in situ hybridization map of RNA-binding proteins reveals anatomically restricted expression in the developing mouse brain. *BMC Dev Biol*. 2005;5(1):14.
33. Barmada SJ, Skibinski G. Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis. *J Neurosci*. 2010;30(2):639-649.
34. Deshaies J-E, Shkreta L. TDP-43 regulates the alternative splicing of hnRNP A1 to yield an aggregation-prone variant in amyotrophic lateral sclerosis. *Brain*. 2018;141(5):1320-1333.
35. Harley J, Hagemann C, Serio A, Patani R. FUS is lost from nuclei and gained in neurites of motor neurons in a human stem cell model of VCP-related ALS. *Brain*. 2020;143(12):e103.
36. Roth HR, Lu L. Improving computer-aided detection using convolutional neural networks and random view aggregation. *IEEE Trans Med Imaging*. 2016;35(5):1170-1181.
37. Dou Q, Chen H. Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. *IEEE Trans Med Imaging*. 2016;35(5):1182-1195.
38. Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2018;172:1122-1131.e9.
39. Elvira G, Wasiak S. Characterization of an RNA granule from developing brain. *Mol Cell Proteomics*. 2006;5(4):635-651.
40. Jun M-H, Ryu HH. Sequestration of PRMT1 and Nd1-L mRNA into ALS-linked FUS mutant R521C-positive aggregates contributes to neurite degeneration upon oxidative stress. *Sci Rep*. 2017;7(1):40474.
41. Sephton CF, Tang AA. Activity-dependent FUS dysregulation disrupts synaptic homeostasis. *Proc Natl Acad Sci U S A*. 2014;111(44):E4769-E4778.
42. Groen EJN, Fumoto K. ALS-associated mutations in FUS disrupt the axonal distribution and function of SMN. *Hum Mol Genet*. 2013;22(18):3690-3704.
43. Udagawa T, Fujioka Y. FUS regulates AMPA receptor function and FTLD/ALS-associated behaviour via GluA1 mRNA stabilization. *Nat Commun*. 2015;6(1):7098.
44. López-Erauskin J, Tadokoro T, Baughn MW, et al. ALS/FTD-linked mutation in FUS suppresses intra-axonal protein synthesis and drives disease without nuclear loss-of-function of FUS. *Neuron*. 2018;100(4):816-830.e7.
45. Cherkasov V, Hofmann S. Coordination of translational control and protein homeostasis during severe heat stress. *Curr Biol*. 2013;23(24):2452-2462.
46. Kroschwald S, Maharana S. Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules. *Elife*. 2015;4:e06807.
47. Markmiller S, Fulzele A, Higgins R, Leonard M, Yeo GW, Bennett EJ. Active protein neddylation or ubiquitylation is dispensable for stress granule dynamics. *Cell Rep*. 2019;27:1356-1363.e3.
48. Mateju D, Franzmann TM. An aberrant phase transition of stress granules triggered by misfolded protein and prevented by chaperone function. *EMBO J*. 2017;36(12):1669-1687.
49. Ninomiya K, Adachi S. LncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J*. 2020;39(3):e102729.
50. Erhardt S, Stoecklin G. The heat's on: nuclear stress bodies signal intron retention. *EMBO J*. 2020;39:e104154.
51. Adusumalli S, Ngian Z-K, Lin W-Q, Benoukraf T, Ong C-T. Increased intron retention is a post-transcriptional signature associated with progressive aging and Alzheimer's disease. *Aging Cell*. 2019;18(3):e12928.
52. Ong C-T, Adusumalli S. Increased intron retention is linked to Alzheimer's disease. *Neural Regen Res*. 2020;15(2):259-260.
53. Sznajder ŁJ, Thomas JD. Intron retention induced by microsatellite expansions as a disease biomarker. *Proc Natl Acad Sci U S A*. 2018;115(16):4234-4239.
54. Tyzack GE, Neeves J, Klein P, et al. An aberrant cytoplasmic intron retention programme is a blueprint for ALS-related RBP mislocalization. bioRxiv 2020 2020.07.20.211557 doi:https://doi.org/10.1101/2020.07.20.211557.
55. Wang Q, Conlon EG, Manley JL, Rio DC. Widespread intron retention impairs protein homeostasis in C9orf72 ALS brains. *Genome Res*. 2020;30(12):1705-1715.
56. Clark A. Pillow (PIL fork) documentation. 2015.
57. Datta K, Hossain I. *Training multiscale-CNN for large microscopy image classification in one hour*. Lecture Notes in Computer Science; 2019:463-477.
58. Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. A survey on deep transfer learning. In: *Artificial Neural Networks and Machine Learning—ICANN 2018*. Springer International Publishing; 2018:270-279.
59. Wang W, Li Y. A novel image classification approach via dense-MobileNet models. *Mob Inf Syst*. 2020;2020:1-8.
60. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: a large-scale hierarchical image database. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*; 2009:248-255.
61. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016:770-778.
62. Mormont R, Geurts P, Maree R. Comparison of deep transfer learning strategies for digital pathology. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2018 doi:https://doi.org/10.1109/cvprw.2018.00303
63. Kokhlikyan N, Miglani V, Martin M, *et al*. Captum: a unified and generic model interpretability library for PyTorch. *arXiv [csLG]* (2020).
64. Bates D, Maechler M, Bolker B, Walker S. lme4: linear mixed-effects models using Eigen and S4. In: *R package version 1.1–7*. Vol.2014;2015.
65. Williams E, Moore J, Li SW, et al. Image Data Resource: a bioimage data integration and publication platform. *Nat Methods*. 2017;14(8):775-781.
66. Williams E, Moore J, Li SW, et al. The Image Data Resource. 2018.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Verzat C, Harley J, Patani R, Luisier R. Image-based deep learning reveals the responses of human motor neurons to stress and *VCP*-related ALS. *Neuropathol Appl Neurobiol*. 2022;48(2):e12770. doi:10.1111/nan.12770