

## ORIGINAL ARTICLE

# Transcriptional signatures of synaptic vesicle genes define myotonic dystrophy type I neurodegeneration

Antonio Jimenez-Marin<sup>1,2</sup>  | Ibai Diez<sup>3,4,5</sup>  | Garazi Labayru<sup>6,7,8</sup>  |  
Andone Sistiaga<sup>6,7,8</sup>  | Maria C. Caballero<sup>9</sup> | Pol Andres-Benito<sup>7,10,11</sup>  |  
Jorge Sepulcre<sup>3,4,5</sup>  | Isidro Ferrer<sup>7,10,11,12</sup>  | Adolfo Lopez de Munain<sup>6,7,13,14</sup> |  
Jesus M. Cortes<sup>1,15,16</sup> 

<sup>1</sup>Computational Neuroimaging Group, Biocruces-Bizkaia Health Research Institute, Barakaldo, Spain

<sup>2</sup>Biomedical Research Doctorate Program, University of the Basque Country (UPV/EHU), Leioa, Spain

<sup>3</sup>Department of Radiology, Division of Nuclear Medicine and Molecular Imaging, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

<sup>4</sup>Gordon Center for Medical Imaging, Department of Radiology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

<sup>5</sup>Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

<sup>6</sup>Neuroscience Area, Biodonostia Research Institute, San Sebastián, Spain

<sup>7</sup>Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Institute Carlos III, Madrid, Spain

<sup>8</sup>Personality, Assessment and Psychological Treatment Department; Psychology Faculty, University of the Basque Country (UPV/EHU), San Sebastian, Spain

## Abstract

**Aim:** To delineate the neurogenetic profiles of brain degeneration patterns in myotonic dystrophy type I (DM1).

**Methods:** In two cohorts of DM1 patients, brain maps of volume loss (VL) and neuropsychological deficits (NDs) were intersected to large-scale transcriptome maps provided by the Allen Human Brain Atlas (AHBA). For validation, neuropathological and RNA analyses were performed in a small series of DM1 brain samples.

**Results:** Twofold: (1) From a list of preselected hypothesis-driven genes, confirmatory analyses found that three genes play a major role in brain degeneration: dystrophin (*DMD*), alpha-synuclein (*SNCA*) and the microtubule-associated protein tau (*MAPT*). Neuropathological analyses confirmed a highly heterogeneous Tau-pathology in DM1, different to the one in Alzheimer's disease. (2) Exploratory analyses revealed gene clusters enriched for key biological processes in the central nervous system, such as synaptic vesicle recycling, localization, endocytosis and exocytosis, and the serotonin and dopamine neurotransmitter pathways. RNA analyses confirmed synaptic vesicle dysfunction.

**Conclusions:** The combination of large-scale transcriptome interactions with brain imaging and cognitive function sheds light on the neurobiological mechanisms of brain degeneration in DM1 that might help define future therapeutic strategies and research into this condition.

**Abbreviations:** AAL, automated anatomical labelling; AD, Alzheimer's disease; AHBA, Allen Human Brain Atlas; ALE, activation likelihood estimation; C, connector hubs; CALCAP, California Computerized Assessment Package; CNS, central nervous system; CTG, cytosine-thymine-guanine; D, Dice index; DDS, data-driven strategy; DM1, myotonic dystrophy type I; DS, differential stability; DSM-V, Diagnostic and Statistical Manual of Mental Disorders, fifth edition; EPI, echo planar imaging; FA, fractional anisotropy; FEW, family-wise error; FWHM, full width at half maximum; GM, grey matter; GO, gene ontology; GWAS, Genome-Wide Association Studies; HCs, Healthy Controls; HDS, hypothesis-driven strategy; IQ, intelligence quotient; kDa, kilo Dalton; MA, middle age; MIRS, Muscular Impairment Rating Scale; MNI, Montreal National Institute; MPRAGE, magnetization-prepared rapid acquisition with gradient echo; MRI, magnetic resonance imaging; mRNA, messenger-RNA; N, Neg-corr genes; ND, neuropsychological deficits; NFT, neurofibrillary tangles; NSA, number of signals averaged; OMMYD-4, Outcome Measure for Myotonic Dystrophy Type 1, fourth international workshop; P, pos-corr genes; RAVLT, Rey Auditory Verbal Learning Test; RIN, RNA integrity number; RNA, ribonucleic acid; ROCF, Rey-Osterrieth Complex Figure; RT, reaction time; SEM, standard error of the mean; SP, sampling proportion; SPIR, spectral pre-saturation with inversion recovery; TBSS, tract-based spatial statistics; TE, echo time; TFCE, threshold-free cluster enhancement; TFE, turbo field echo; TR, repetition time; VBM, voxel-based morphometry; VL, volume loss; WAIS III, Wechsler Adult Intelligence Scale III; WM-TA, white matter tract alterations.

Adolfo Lopez de Munain and Jesus M. Cortes had equal last-author contribution.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Neuropathology and Applied Neurobiology* published by John Wiley & Sons Ltd on behalf of British Neuropathological Society

<sup>9</sup>Donostia University Hospital, Donostia-San Sebastian, Spain

<sup>10</sup>Department of Pathology and Experimental Therapeutics, University of Barcelona, Barcelona, Spain

<sup>11</sup>Institute of Biomedical Research of Bellvitge (IBIDELL), Hospitalet de Llobregat, Spain

<sup>12</sup>Institute of Neurosciences, University of Barcelona, Hospitalet de Llobregat, Spain

<sup>13</sup>Neurology Department, Donostia University Hospital, Donostia-San Sebastian, Spain

<sup>14</sup>Neurosciences Department, University of the Basque Country (UPV/EHU) Donostia-San Sebastian, Spain

<sup>15</sup>Cell Biology and Histology Department, University of the Basque Country (UPV/EHU), Leioa, Spain

<sup>16</sup>IKERBASQUE, The Basque Foundation for Science, Bilbao, Spain

#### Correspondence

Jesus M. Cortes, Computational Neuroimaging Group, Biocruces-Bizkaia Health Research Institute, Barakaldo, Spain.  
Email: [jesus.m.cortes@gmail.com](mailto:jesus.m.cortes@gmail.com)

#### Funding information

Department of Economic and Infrastructure Development of the Basque Country, Elkartek, Grant/Award Number: KK-2018/00032 and KK-2018/00090; Hezkuntza, Hizkuntza Politika Eta Kultura Saila, Eusko Jaurlaritza, Grant/Award Number: PRE\_2016\_1\_0187 and PRE\_2019\_1\_0070; Eusko Jaurlaritza, Grant/Award Number: SAIO08-PE08BF01; Secretaría de Estado de Investigación, Desarrollo e Innovación, Grant/Award Number: DPI2016-79874-R; Instituto de Salud Carlos III, Grant/Award Number: FIS PI17/000809, PI17/01231 and PI17/01841; "la Caixa" Foundation, Grant/Award Number: HR17-00268 and LCF/PR/HR19/52160007; Centro de Investigación Biomédica en Red sobre Enfermedades Neurodegenerativas, Grant/Award Number: grant 609

#### KEYWORDS

Allen Human Brain Atlas, DM1, neuropsychological deficits, structural neuroimaging, synaptic vesicles, volume loss

## INTRODUCTION

Myotonic dystrophy type 1 (DM1) is a complex multisystem disease that affects skeletal muscles,<sup>1</sup> heart,<sup>2</sup> lungs,<sup>3</sup> endocrine system,<sup>4</sup> regulation of sleep cycles<sup>5</sup> and other aspects of brain activity.<sup>6</sup> Epidemiologically, DM1 is the most common adult-onset muscular dystrophy in humans, with a reported prevalence of 1/7,400 people worldwide<sup>7</sup> and is about three times higher in Gipuzkoa,<sup>8</sup> Northern Spain (where this study was performed). Neuroimaging studies have shown the brain damage in DM1 patients, including grey matter (GM) atrophy mainly affecting the frontal and parietal lobes<sup>9</sup> but also, in the hippocampus<sup>10</sup> and other

subcortical structures.<sup>11</sup> White matter tract alterations (WM-TAs) have been widely reported in DM1, in both cross-sectional<sup>12-20</sup> and longitudinal studies.<sup>21</sup> These indicate widespread WM-TA throughout the whole brain, but that is more severe in the frontal, temporal and subcortical tracts. In relation to neuropsychological performance, executive function, attention and visuoconstruction have been shown significant association with WM-TA.<sup>16,22</sup> To the date, the neurogenetic profiles of such patterns of WM-TA or the ones corresponding to GM damage have not been yet determined.

Using an approach that combines magnetic resonance imaging (MRI) and large-scale brain transcriptomics, we aimed here to assess

to what extent *the structural damage in the DM1 brain represents a neurogenetic signature*. In contrast to other neurodegenerative diseases in which a large number of candidate genes are implicated, for example, about 700 genes in Alzheimer's disease (AD),<sup>23</sup> DM1 is a monogenic disorder caused by a mutation in the gene encoding the myotonic dystrophy protein kinase (*DMPK*).<sup>24</sup> However, although the disease is monogenic, its phenotype is mainly due to an abnormal activity of the ribonucleic acid (RNA)-binding protein muscleblind-like 1 and 2 genes (*MBNL1*, *MBNL2*)<sup>25,26</sup> and *CUGBP* which regulates the expression of many other genes, such as the chloride channel 1 gene (*CLCN1*) that regulates chloride conductance during muscle development,<sup>27</sup> the insulin receptor gene (*INSR*),<sup>28</sup> the bridging integrator 1 gene (*BIN1*)<sup>29</sup> or other genes directly related to the main symptoms of the condition.<sup>30</sup> The neuropsychological profile of DM1 is mainly associated with cognitive difficulties including visuospatial processing and executive functioning.<sup>31,32</sup> Strikingly, the *DMPK* pathogenic genotype has also been associated with other genes that encode proteins implicated in brain neurodegeneration, majorly to tau deposits,<sup>33</sup> but also amyloid beta ( $A\beta$ )<sup>34</sup> or alpha synuclein.<sup>35,36</sup> Thus, it is suspected that the neuropsychological deficits (NDs) and brain damage found in DM1 patients might present some similarities to those in other neurodegenerative diseases, although this issue has yet to be fully addressed. Therefore, despite the monogenic origin of DM1, the gene-to-gene interactome scales up to implicate multiple systems in the brain and body. To date, a precise association between the entire transcriptome and the brain neurodegeneration and cognitive deterioration in DM1 patients remains unexplored.

Some studies have assessed the relationship between genetics and structural brain alterations in DM1, confirming that the number of pathogenic repeats of the cytosine-thymine-guanine (CTG) triplet in the *DMPK* gene (a parameter used to quantify the molecular severity of the disease) was associated with both more GM and white matter atrophy, and to the NDs of these patients.<sup>20,37</sup> Here, we performed an intersection analysis of neuroimaging phenotypes and the Allen Human Brain Atlas (AHBA) of large-scale transcriptional human data,<sup>38</sup> following a similar methodology to that used previously.<sup>39–41</sup> Our hypothesis was that identifying the genes whose expression coincided more closely with the brain damage found in DM1 patients, we might better understand the gene relationships associated with the brain damage that arises in this pathology. Similarly, we also assessed the relationship between gene transcription in specific anatomical regions and the brain maps of NDs in these patients.

## METHODS

### Participants , two cohorts

A total of  $N = 95$  subjects were recruited to a cross-sectional study, 35 of whom were DM1 patients who were treated at the Neurology Department of the Donostia University Hospital (Gipuzkoa, Spain), while 60 subjects participated as healthy controls (HCs). All patients and HCs were recruited from the vicinity of the Donostia University

Hospital, and the two groups were matched for age, gender and education. The imaging data from the DM1 patients and HCs were acquired at two different Institutions. At one, 19 DM1 patients (mean age 53.3 years [SD  $\pm$  8.1 years]; 9 males, 10 females) and 29 HC (52.2 [ $\pm$ 8.1] years; 12 males, 17 females) were examined and at the second, 16 DM1 patients (48.8 [ $\pm$ 7.7] years; 7 males, 9 females) and 31 HC (47.6 [ $\pm$ 7.6] years; 14 males, 17 females). For the mean values, and the comparisons between groups and cohorts, see Table 1 and Table S1.

The DM1 patients were only included if they had molecular confirmation of their DM1 diagnosis, indicating the expansion from 50 to thousands of CTG repeats in the *DMPK* gene.<sup>24</sup> The diagnosis was obtained when patients were between 18 and 40 years old, and therefore categorized as adult-onset DM1, as proposed by the fourth Outcome Measure for Myotonic Dystrophy Type 1 (OMMYD-4). Patients were excluded if at least one of the following criteria was met: history of a major psychiatric or somatic disorder in accordance with Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-V) criteria; acquired brain damage; alcohol or drug abuse; the presence of corporal paramagnetic devices like pacemakers or metal prosthesis that might compromise the MRI studies; and the presence of brain abnormalities that could affect the volumetric analysis. HCs satisfied the same exclusion criteria but the number of the CTG repeats in their *DMPK* gene ranged from 5 to 34.<sup>24</sup> DM1 patients were recruited from the Neuromuscular Unit in the Neurology Department of the Hospital Universitario Donostia, while HCs were recruited from their healthy relatives and general population.

### Demographic , clinical and neuropsychological variables

The demographic variables of the subjects recorded were their age, gender and years of education. The clinical variables were the CTG expansion size and Muscular Impairment Rating Scale (MIRS) score.<sup>42</sup> Neuropsychological variables corresponded to composite values from different cognitive domains obtained through a comprehensive neuropsychological evaluation performed by an experienced neuropsychologist who was blind to the patient's clinical condition (CTG expansion size and MIRS results). The neuropsychological assessment included several subtests from the Wechsler Adult Intelligence Scale III (WAIS III),<sup>43</sup> including: digit span, vocabulary, block design, object assembly, arithmetic and similarities. Other cognitive tests used were as follows: Stroop test, California Computerized Assessment Package (CALCAP), Raven's progressive matrices, Rey Auditory Verbal Learning Test (RAVLT),<sup>44</sup> Word Fluency,<sup>45,46</sup> Rey-Osterrieth Complex Figure test (ROCF)<sup>47</sup> and Benton's Judgement of Line Orientation.<sup>48</sup> The patients' raw scores were converted into standardized  $t$  values based on the normative scores for the Spanish population in each test. Finally, the different neuropsychological scores were reduced into six different domains: visuospatial (block design and ROCF copy), verbal memory (RAVLT immediate recall, RAVLT delayed recall, Total RAVLT), attention (digit span, STROOP word, STROOP

**TABLE 1** Demographic, clinical and neuropsychological variables

	First cohort	Second cohort	t	p	Effect size (Cohen)
<b>Demographic</b>					
N	19	16			
Age, years	53.30 (8.09)	48.75 (7.72)	1.69	0.10	0.57
Males, n (%)	9 (47.37)	7 (43.75)	0.05 <sup>a</sup>	0.83	-0.07
Education, years	16.37 (4.87)	13.75 (4.85)	1.59	0.12	0.54
<b>Clinical</b>					
N	19	16			
CTG expansion size	522.79 (448.44)	827.13 (433.08)	-2.03	0.05	-0.69
MIRS	2.47 (0.96)	3.53 (0.83)	-3.37	<b>0.002</b>	-1.17
Time between MRI and CTG measures, years	0.14 (0.14)	9.40 (7.23)	-5.6	<b>0.000</b>	-1.93
<b>Neuropsychological</b>					
N	18	16			
Visuospatial	-0.35 (1.21)	-1.17 (1.34)	1.86	0.07	0.64
Verbal memory	-0.04 (2.33)	-0.74 (1.99)	0.94	0.36	0.32
Attention	-2.50 (1.92)	-2.62 (2.14)	0.19	0.85	0.06
Executive functioning	-0.91 (2.08)	-1.84 (2.40)	1.21	0.23	0.42
Visual memory	-0.34 (0.97)	-1.01 (1.14)	1.83	0.08	0.64
Intelligence (IQ)	101.11 (11.51)	91.19 (13.61)	2.3	<b>0.03</b>	0.79

Note: Mean values (standard deviations in brackets) of the different variables separated by cohort. The neuropsychological variables coincide with the different composite-scores from the different cognitive domains. All the neuropsychological variables were calculated using normative data from a healthy Spanish population. All domains are represented in Z scores (Mean =0; SD =1) except IQ (Mean =100; SD =15). Smaller values indicate worse performance. Bold values indicate  $p < 0.05$ .

<sup>a</sup>For gender differences, the  $\chi^2$  test was used.

colour, simple reaction time (RT), election RT, Sequential 1 RT, Sequential 2 RT), executive functioning (total RAVEN, semantic fluency, phonemic fluency, STROOP colour-word, STROOP interference), visual memory (ROCF delayed recall), and intelligence (estimation of IQ was based on vocabulary, block design, object assembly, arithmetic and similarities subtests).

### **MRI acquisition and pre-processing**

For the first cohort, MRI was conducted on a 3 Tesla scanner (TrioTim, Siemens) using a high-resolution 3D sequence of magnetization-prepared rapid acquisition with gradient echo (MPRAGE) and applying the following parameters: Sagittal 3D T1 weighted acquisition, repetition time (TR) = 2,300 ms, echo time (TE) = 2.86 ms, inversion time = 900 ms, flip angle = 9°, matrix = 192 × 192 mm<sup>2</sup>, slice thickness = 1.25 mm, voxel dimensions = 1.25 × 1.25 × 1.25 mm<sup>3</sup>, number of signals averaged (NSA) = 1, slices = 144, no gap, total scan duration = 7 min and 22 s. Diffusion-weighted images were acquired using an echo planar imaging (EPI) sequence: 1.75 × 1.75 × 2 mm voxels; 77

axial slices; b value of 1,000 s/mm<sup>2</sup>; 64 direction diffusion-weighted and 1 baseline image; TR = 10,000 ms, TE = 92 ms; angle 90, acquisition matrix size = 122 × 122.

For the second cohort, MRI was conducted on a 1.5 Tesla scanner (Achieva Nova, Philips), using a high-resolution volumetric turbo field echo (TFE) sequence with the following parameters: Sagittal 3D T1 weighted acquisition, TR = 7.2 ms, TE = 3.3 ms, inversion time = 0 ms, flip angle = 8°, matrix = 256 × 232 mm<sup>2</sup>, slice thickness = 1 mm, voxel dimensions = 1 × 1 × 1 mm<sup>3</sup>, NSA = 1, slices = 160, no gap, total scan duration = 5 min and 34 s. Diffusion-weighted images were acquired using a single shot spectral presaturation with inversion recovery (SPIR): 1.75 × 1.75 × 2 mm voxels; 60 axial slices; b value of 800 s/mm<sup>2</sup>; 32 direction diffusion-weighted and 1 baseline image; TR = 9,967 ms, TE = 66 ms; angle 90, acquisition matrix size = 128 × 128.

To perform voxel-based GM morphometric comparisons between the subjects, DM1 and HCs, we performed voxel-based morphometry (VBM) following a procedure similar to that used previously,<sup>49,50</sup> an optimized VBM protocol<sup>51</sup> carried out with the FSL v6.01 software. First, skull removal was performed, followed by

GM segmentation and registration to the MNI 152 standard space using non-linear registration.<sup>52</sup> The resulting images were averaged and flipped along the *x*-axis to create a left-right symmetric, study-specific GM template. Second, all native GM images were non-linearly registered to this study-specific template and 'modulated' to correct for local expansion (or contraction) due to the non-linear component of the spatial transformation. The modulated GM images were then smoothed with an isotropic Gaussian kernel at  $\sigma = 3$ , and finally, the partial GM volume estimates normalized to the subject's head size were compared.

WM-TA were assessed with tract-based spatial statistics (TBSS) from FSL<sup>53</sup> using fractional anisotropy (FA) images. First, all individual FA images were normalized to a common template using a non-linear transformation. Next, the mean image across all subjects was computed and skeletonized to get the mean FA skeletons, which represent regions with high confidence bundles common to all subjects, thus removing some of the individual subject-specific tract-based heterogeneities. FA images for each subject were then projected onto the mean skeleton.

### **Imaging statistical analysis**

For the VBM analyses, a generalized linear model was fitted for each voxel and image using the FSL software, controlling for age and head size with two different contrasts: DM1 < HC and DM1 > HC. All the results were obtained with two-tailed tests, correcting for multiple comparisons using the Monte Carlo simulation cluster-wise correction implemented in the AFNI v19.3.00 software, and using 10,000 iterations to estimate the probability of false positive clusters with a *p* value < 0.05. Each cohort was analysed separately and in combination, as explained below, although statistical comparisons between the two cohorts were not performed.

For the TBSS analyses, group comparison was performed using the *randomise* tool in FSL, a nonparametric permutation test for finding significant statistical differences between groups at the voxel level. For multiple comparison correction, we used threshold-free cluster enhancement (TFCE)<sup>54</sup> with a number of iterations of 5,000 and family-wise error (FWE) corrected *p* = 0.05, thus ensuring that the chance of false positives is no more than 5%, or equivalently, ensuring 95% confidence of no false positives. Group comparison was performed using two different contrasts, DM1 > HC and HC > DM1.

### **Transcriptomics brain maps**

To build brain maps of transcription, we took advantage of the publicly available data from the AHBA (<http://human.brain-map.org/>).<sup>38</sup> The dataset consisted of MRI images and a total of 58,692 microarray-based transcription profiles of about 20,945 genes sampled over 3,702 different regions across the brains of six humans. Of the total 3,702 sampling sites, 2,728 were located in cortical and subcortical GM, 368 in the cerebellum, 586 in the brain stem, and

15 in white matter. Our analyses here are restricted to the 2,728 sites covering the cortical and subcortical GM. To pool all the transcription data into a single brain template, we followed a similar procedure to that employed elsewhere.<sup>55</sup> First, to re-annotate the probes to genes we made use of the re-annotator toolkit.<sup>56</sup> Second, we removed those probes with insufficient signal by looking at the sampling proportion (SP), which was calculated for each brain as the ratio between the samples with a signal greater than the background noise divided by the total number of samples. Probes with a SP lower than 70% in any of the six brains were removed from the analysis, thereby ensuring sufficient sampling power in all the brains. After that, we chose the value of the probe for each gene with the maximum differential stability (DS), accounting for the reproducibility of gene expression across brain regions and individuals, and calculated using spatial correlations similar to those employed previously.<sup>57</sup> For this the automated anatomical labelling (AAL) atlas was used<sup>58</sup> from which the cerebellum was excluded, resulting in 90 different anatomical regions.<sup>1</sup> Finally, to remove the intersubject differences, the transcription values for each gene and brain were transformed into Z scores and pooled together from the six different brains, obtaining a single map using the MNI coordinates provided in the dataset. Finally, to eliminate the spatial dependencies of the transcription values at the sampling sites (that is, to correct for the fact that nearest sites have more correlated transcription), we finally obtained a single transcription value for each region in the AAL atlas by calculating the median of all the values belonging to the given region.

### **Association between VL and transcriptomics**

Volume loss (VL) was defined as the *t* statistic resulting from the group comparison DM1 < HC. To associate VL with transcriptomics, we transformed the *t* statistics map to the same AAL atlas as that used for the transcription values, calculating the median of the *t* statistics between all values in each region. For each gene, we calculated a similarity index using the Pearson correlation coefficient between VL and transcriptomics, the two variables represented in vectors with a dimension equal to the number of regions in the atlas. Therefore, a high contribution to this association is to have a high value in the product of the two vector components for a given region. In addition, the similarity index (i.e., spatial correlation) is proportional to the sum of all product terms across all brain regions, thus capturing systematically similar patterns throughout the brain of the gene expression vector and the one for VL. This procedure was done separately for the two cohorts.

### **Association between NDs and transcriptomics**

For the association between NDs and transcriptomics, we first built brain maps of ND using the BrainMap meta-analysis platform (<http://www.brainmap.org/>). In particular, the Sleuth tool v3.0.3<sup>59</sup> was used to search all the papers in the database using the name of

each neuropsychological domain as a keyword. In this way, all the co-activation coordinates that resulted from studies based on functional imaging when the participant in the scanner was performing a task related to each neuropsychological domain were obtained. Next, the GingerAle tool v3.0.2<sup>60</sup> was used to pool all the coordinates onto a single co-activation brain map for each neuropsychological domain, representing these as a Z score after applying the activation likelihood estimation (ALE) method. The brain maps were then transformed to the same AAL atlas by calculating the median of all the Z scores belonging to the same region in the atlas. Finally, the association between ND and transcriptomics was assessed for each gene through the similarity index, equivalent to the Pearson correlation coefficient between ND and transcriptomics; the two variables were represented in the vectors with a dimension equal to the number of regions in the atlas. Transcriptomics correlates were only obtained for those neuropsychological domains that were most affected, identified by choosing a mean Z score < -2, previously standardized to the normative scores based on a Spanish population. This procedure was done separately for the two cohorts.

### DM1 relevant genes and GO

Relevant genes were identified by combining the results from two different strategies. The first one was confirmatory, a hypothesis-driven strategy (HDS) that involved reviewing previous studies to identify genes that play a relevant role in some neurobiological aspects of DM1 (Table S2). The second was exploratory, a data-driven strategy (DDS) consisting of identifying the genes with the strongest transcription correlation with the parameter of interest, either VL or ND. In particular, we chose those genes with similarity index values where  $z < -2$  or  $z > 2$ , i.e., outliers of the correlation distribution in both the negative and positive tails. Genes in the positive tail ( $z > 2$ ) were designated as *pos-corr* genes (P), whereas those in the negative tail ( $z < -2$ ) were considered *neg-corr* genes (N). For example, when assessing the VL, the P genes were systematically expressed more strongly than other genes in the brain regions with more pronounced atrophy, whereas the N genes were expressed much more weakly than the rest. Although this procedure was performed separately for the two cohorts, the P and N genes finally adopted were those present in the two cohorts. We also considered those genes that were expressed similarly to the P and N genes as relevant genes, which were dubbed *connector hubs* (C). To identify these, we made use of the 'gene-expression connectivity matrix' that represents the similarity in expression between gene pairs obtained from the Pearson pairwise correlations for each entry.<sup>2</sup> For those genes absent in the groups of P and N genes, the strength towards the two P and N tails was calculated separately. PC genes were identified as those genes with a Z score in connectivity strength towards the P tail >2, and similarly, NC genes were identified by having a Z score in the strength towards the N tail >2. For the hypothesis-driven genes given in Table S2 that were defined as P or N genes, their statistical significance was assessed by surrogate data testing. The BrainSMASH tool<sup>61</sup> was used

to build null-distributions by generating 10,000 random maps with the same spatial autocorrelation as that for VL or ND.

Finally, we pooled together all the relevant genes (P genes, N genes, PC genes, NC genes and the genes that were common to PC and NC) to perform unsupervised K-means clustering with the *Silhouette* strategy to identify the optimal number of clusters. For each cluster, we performed a gene ontology (GO) *biological process*<sup>62</sup> and *Reactome pathways*<sup>63</sup> overrepresentation test using PANTHER v15.0 (<http://pantherdb.org/>), with the entire Homo Sapiens genome as the reference list and applying a Fisher's exact test with Bonferroni correction ( $p < 0.05$ ). To make the data more readily interpretable, we only reported  $\geq 2$ -fold enrichment.

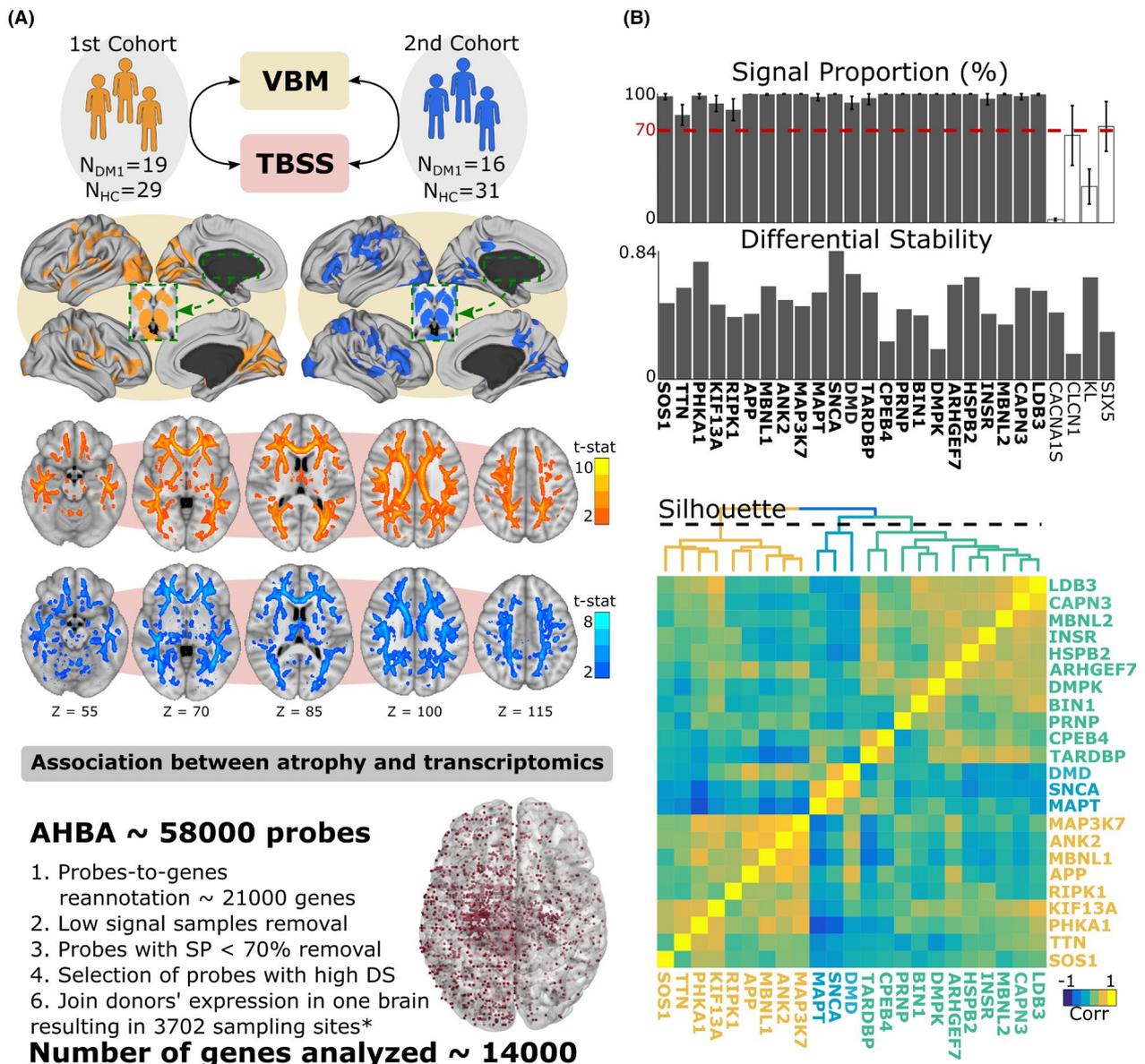
### Neuropathological analyses

Human brain samples were collected from DM1 patients after approval of the local ethical committee, treated and stored following the guidelines of the Spanish legislation (Real Decreto de Biobancos RD1716/2011; see Supplementary Material for a detailed protocol).

Brain samples were analysed from seven DM1 patients with ages 60, 58, 74, 67, 58, 68 and 59 at death. All of them were diagnosed by genetically-confirmed DM1 and had clinical manifestations. Due to the external origin of the cases, not all brain regions were available for study in every case; in some cases, only 20 different areas were suitable for the neuropathological study. Although frozen samples were obtained from different regions, only the hippocampus was available in every case for biochemical studies.

Gene expression was analysed in the hippocampus of four cases with neurofibrillary tangles (NFT) pathology stages III–VI of Braak (NFT control, aged 76, 85, 70 and 85), four middle-aged (MA) individuals (aged 53, 48, 55 and 39, and three DM1 cases [I13–131, I13–133, I13–134], aged 67, 68, and 59, stages III–VI). The selection of the first group was chosen to discriminate possible alterations linked to NFT in DM1 which might not be present in current cases with NFT pathology not linked to DM1 (AD-related pathology).

RNA from hippocampus was extracted following the instructions of the supplier (RNeasy Mini Kit, Qiagen® GmbH, Hilden, Germany). RNA integrity (RIN) and 28S/18S ratios were determined with the Agilent Bioanalyzer (Agilent Technologies Inc, Santa Clara, CA, USA). RIN values were between 5.5 and 7.5. Samples were treated with DNase digestion, and RNA concentration was evaluated using a NanoDrop™ Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). The normality of distribution of fold-change values was analysed with the Kolmogorov–Smirnov test. Pearson's correlation coefficient was used to assess a possible linear association between two continuous quantitative variables. Statistical analysis of the expression data between groups was made using one-way analysis of variance (ANOVA) followed by Tukey post-test or Kruskal–Wallis test followed by Dunns post hoc test when required using the SPSS software (IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 21.0; Armonk, NY: IBM Corp). Outliers were detected using



**FIGURE 1** Methodological scheme for the association between transcriptomics and atrophy in DM1, measured as brain volume loss (VL). (A) Two cohorts of DM1 patients were recruited (orange and blue) and we obtained the brain maps of the VL for each, comparing the images with a group of HCs using voxel-based morphometry (VBM) and tract-based spatial statistics (TBSS), correcting for multiple comparisons. We aimed to characterize the association between VL and transcriptomics, assessing the similarity in the spatial patterns of VL across brain regions and the spatial patterns of gene transcription from the AHBA dataset, pre-processed following a pipeline that is summarized in six main steps (for further details, see Methods). After running the AHBA pipeline, about 14 K genes finally had transcription values used in the analysis from the 58-K probes originally available. The red regions in the brain correspond to the sites at which transcription was sampled. \*Of the total 3,702 sampling sites, 2,728 were located in cortical and subcortical grey matter, 368 in the cerebellum, 586 in the brain stem, and 15 in white matter. Our analyses here are restricted to the 2,728 sites covering the cortical and subcortical grey matter. (B) The sampling proportion (SP) and differential stability (DS) for the 27 preselected hypothesis-driven genes included in the list of candidates relevant to DM1 (obtained by reviewing the literature). The *CACNA1S*, *CLCN1*, *KL* and *SIX5* genes did not have a mean SP value above 70% and thus, they were excluded from further analyses. Of the remaining 23 genes, the maximum DS corresponded to *SNCA* and the minimum DS to *DMPK* (see arrows). By examining the spatial similarity in the transcription values, the remaining 23 candidate genes were clustered into three groups. The blue one formed by *DMD*, *SNCA* and *MAPT* played a major role in the characterization of VL

the GraphPad software QuickCalcs ( $p < 0.05$ ). All data were expressed as mean values  $\pm$  standard error of the mean (SEM). Differences between cases were considered statistically significant at  $*p < 0.05$ ; trends were indicated with \* and the exact  $p$  value.

## RESULTS

A total of 35 DM1 patients and 60 HCs participated in this study, examined in two different cohorts at two distinct centres. In the

first of these, 19 DM1 patients were recruited (9 males) with an average age of 53.3 years (range 42.1–69.5 years) and with 16.4 years of education (range 8–25 years). With respect to the clinical variables, these patients had an average MIRS score of 2.5 (range 1–4) and a mean CTG expansion of 522.8 (range 65–1,733). These patients were compared with 29 well-matched HCs. In the second cohort, 16 DM1 patients were recruited (7 males), with an average age of 48.8 years (range 36.0–61.0 years) and 13.2 years of education (range 6–22 years). The mean MIRS score in this cohort was 3.5 (range 2–5) and with a CTG expansion size of 827.1 (range 267–1,833). These patients were compared to 31 well-matched HCs, and all the demographic, clinical and neuropsychological variables collected in the study are detailed in Table 1 and S1. As explained in the Methods, the two cohorts were considered independently for the imaging analyses, obtaining the corresponding VL brain maps and identifying the list of relevant genes for each cohort separately. Importantly, the two brain profiles of VL (represented by the vector of statistics  $t$  per brain region) had a high spatial correlation across brain regions ( $r = 0.82$ ), thus indicating that the statistics after group comparison remain quite stable in the two cohorts despite MRI scanner differences, which ultimately guaranteed a common list of relevant genes present in the two cohorts. It was this list of common genes shared by the two cohorts the one analysed further.

The VL brain maps from each cohort consisted of several widespread cortical and subcortical regions (Figure 1A). There was bilateral subcortical VL in the hippocampus, thalamus, basal ganglia and cerebellum, while cortical VL was evident in multiple regions that included part of the temporal, occipital, parietal cortices, precuneus, pars orbitalis and medial prefrontal cortex (for details, see Tables S3 and S4). Importantly, the VL maps for the two cohorts differed more in the cortical regions (Dice index,  $D = 0.3$ ), and they were more similar in the subcortical ones ( $D = 0.8$ ), indicating that subcortical atrophy was more prevalent in DM1 and less variable than cortical atrophy.

The WM-TA brain maps from each cohort showed alterations for most of the white matter tracts (Figure 1A; for significance values and statistics for each tract, see Tables S5 and S6). When looking for differences across cohorts, we found high similarity between the two maps ( $D = 0.7$ ), indicating again that despite MRI scanner differences, the results of WM-TA were quite robust from the two cohorts.

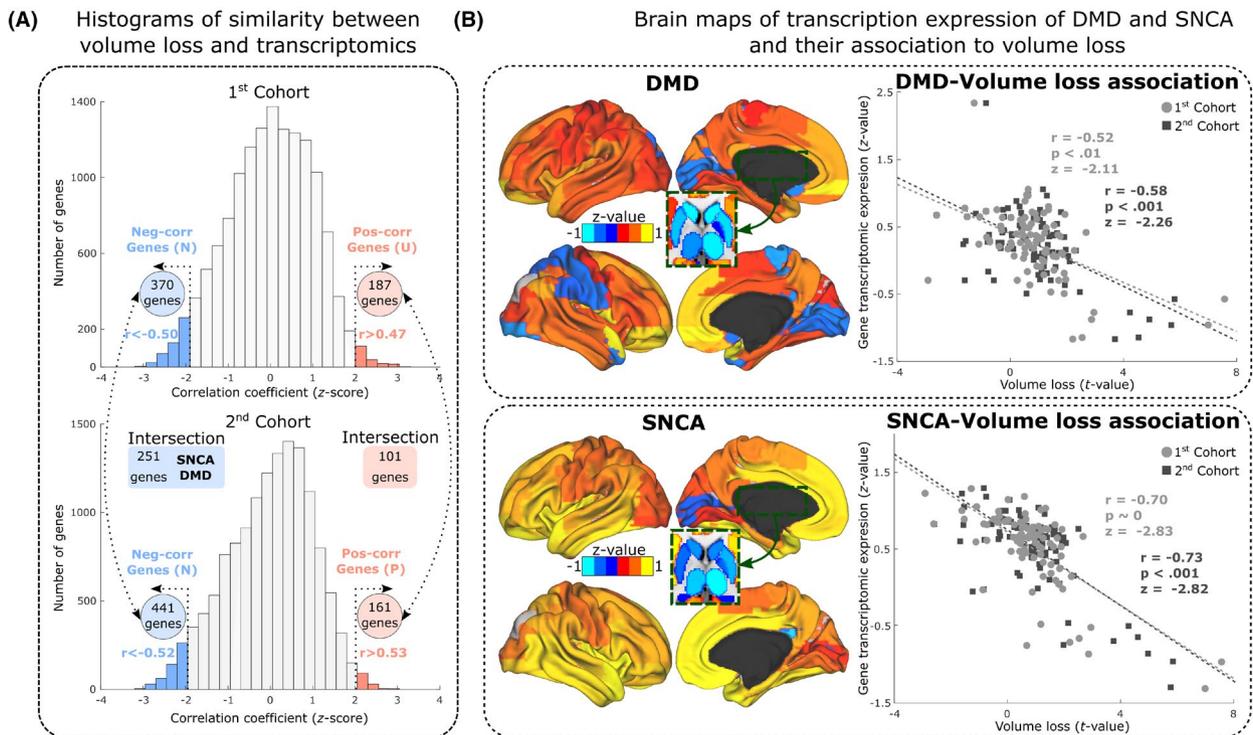
The main steps in the pipeline were followed to analyse the transcriptomics through AHBA dataset (Figure 1A), providing brain maps of transcriptional activity for each gene that were spatially correlated with the brain maps of VL and subsequently, with those of ND. A list of the 27 most relevant hypothesis-driven genes in DM1 was drawn up (Figure 1B; for details on references supporting the selection of each gene, see Table S2). Of these 27 preselected genes, four genes were discarded based on their SP and DS values (*KL*, *SIX5*, *CLCN1*, *CACNA1S*). These 4 genes had less than 70% SP (Figure 1B), which implied insufficient transcription signal across all the sampling sites. Therefore, the final list of the most relevant hypothesis-driven

genes contained: *DMPK*, *HSPB2*, *INSR*, *CPEB4*, *ANK2*, *ARHGEF7*, *SOS1*, *PHKA1*, *MBNL1*, *KIF13A*, *APP*, *MAPT*, *SNCA*, *MBNL2*, *RIPK1*, *PRNP*, *TARDBP*, *MAP3 K7*, *BIN1*, *DMD*, *LDB3*, *TTN* and *CAPN3*. When the pairwise gene-to-gene similarity in transcription was evaluated across the brain for these 23 genes, Silhouette maximization identified three clusters (in yellow, blue and green in Figure 1B), indicative of functional similarities in the transcription signals among the 23 most relevant genes.

We next applied a DDS that involved identifying the genes from the entire transcriptome with maximal association in the VL maps, resulting in a total of 370 N genes and 187 P genes from the first cohort and 441 N-genes and 161 NP-genes from the second one. The genes common to the two cohorts were those finally used in the analysis, a total of 251 N genes and 101 P genes (Figure 2A). Interestingly, two of the genes in the list of the 23 most relevant hypothesis-driven genes also appeared in the list of N genes, *SNCA* and *DMD*, displaying a similar transcription pattern as the *MAPT* gene (blue cluster in Figure 1B). The spatial correlation of *SNCA* and *DMD* transcription with the amount of VL in the different brain regions proved to be negative for the two genes and smaller in the two cohorts:  $< -0.52$  for *DMD* and  $< -0.70$  for *SNCA* (Figure 2B and Table S7). In addition to identifying the N and P genes, we also searched for the NC and PC genes (Figure 3A) that represent *hubs* towards the N and P tails of the expression similarity matrix. We found 452 NC genes, 396 PC genes and 238 genes connecting both N and P tails. Remarkably, in the group of PC genes, we found *LDB3*, *CAPN3* and *HSPB2* that were in the list of hypothesis-driven genes, belonging to the same cluster of expression similarity (green cluster, Figure 1B). In addition, the preselected gene *PHKA1* was also common to the groups of NC and PC genes.

Pooling the N, P, NC and PC genes together, along with those common to both the NC and PC categories, we adopted a DDS to achieve unsupervised clustering of the expression similarity matrix, identifying two major clusters after Silhouette maximization (in blue and red in Figure 3B). Importantly, the N and P genes fully segregated into the two differentiated clusters, with all the N genes belonging to the blue cluster and all the P genes to the red one, thereby confirming the different functional roles of the groups of genes in the N and P tails. These two clusters were used separately for gene enrichment analysis (the list of genes included in each cluster are given in Tables S8 and S9). The search for the *GO biological processes* and *Reactome pathways* confirmed the differentiated roles of these two clusters, with the *neg-corr* genes more related to neuronal and synaptic function, involving key synaptic vesicle events such as recycling, localization, endocytosis and exocytosis but also, the dynamics of serotonin and dopamine neurotransmitter release (Figure 3C). By contrast, the cluster of *pos-corr* genes was more related to non-neuronal activities, such as interferon signalling, endothelial cell differentiation, angiogenesis, blood transport and cell development.

To assess how the transcriptomics correlated with ND, we focused on the neuropsychological domains in which the composite score reflected strong impairment, satisfying  $z < -2$ , which was only the case for the attention category (first cohort  $z = -2.5$ ,

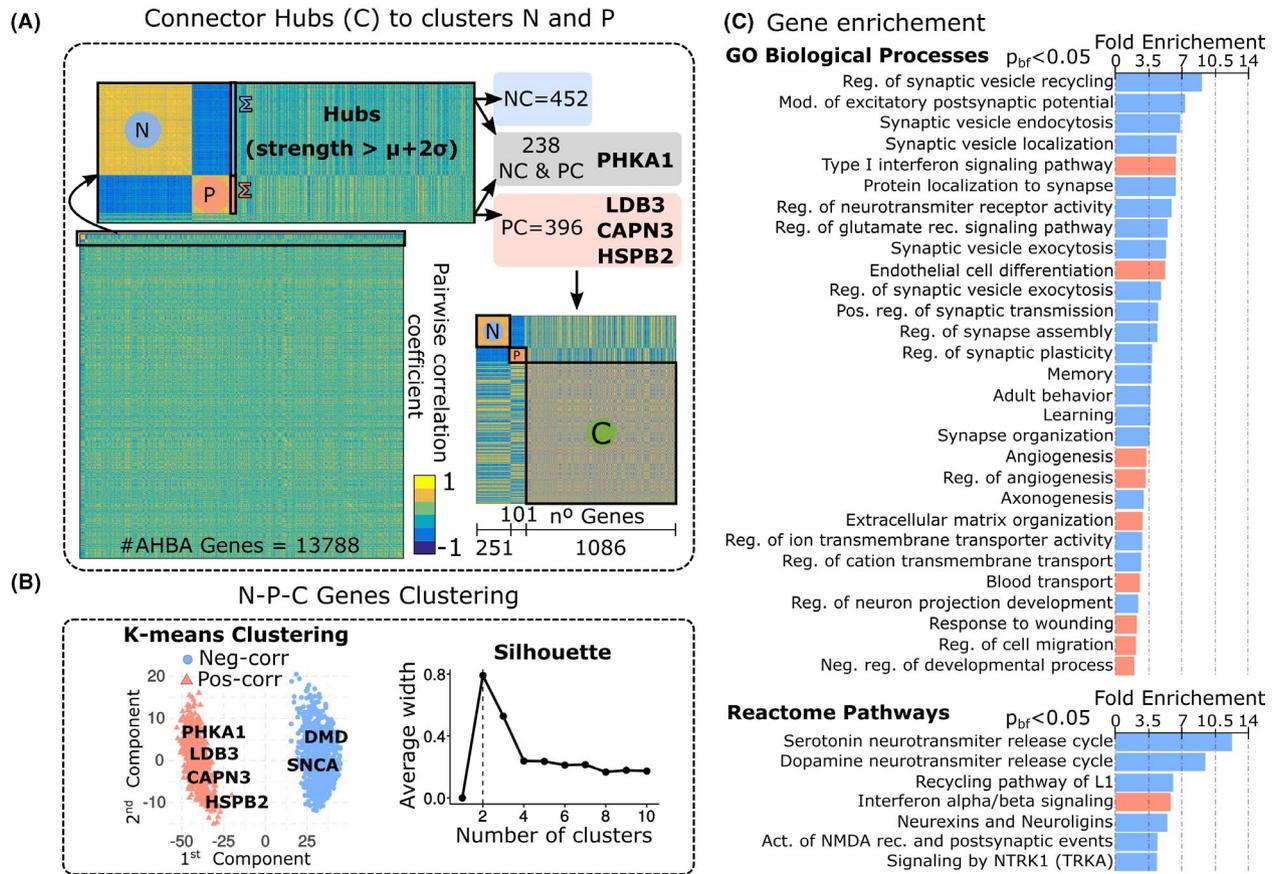


**FIGURE 2** Data-driven strategy to determine the association between the transcriptome and VL in DM1. (A) Histogram of the spatial-correlation values (measured as the Z-score) between volume loss (VL) and transcriptional activity for all the genes in both cohorts. For both cohorts the N genes ( $z < -2$ ) and P genes ( $z > 2$ ) are coloured in blue and red, respectively. The final list of genes used for further analyses are those that are common to the two cohorts, consisting of 251 N genes and 101 P genes. From all the genes that provide a maximum association between VL and transcriptomics (Tables S6 and S7), only two genes were in the panel of preselected genes: *SNCA* and *DMD*. (B) Brain maps of transcription in the brain regions for the two genes *DMD* and *SNCA*, which provided a high spatial correlation ( $r$ ) with the VL brain maps for both cohorts

second cohort  $z = -2.62$ : see Table 1 and Figure 4A for the distribution of all the Z score values from the two cohorts). As indicated in the Methods, the attention composite was obtained by averaging the Z scores for the following domains: digit span, STROOP word, STROOP colour, simple RT, election RT, sequential 1 RT and sequential 2 RT.

We then obtained brain co-activation maps using 'Attention' as a keyword for the search in GingerAle (Figure 4B). As for VL, we calculated the association between the attention co-activation maps and transcriptomics by calculating the spatial Pearson correlation between the two Z score vectors (one value per brain region in the atlas, see the histogram of all the correlations in Figure 4C). There was a total of 217 N genes (including *SNCA*), 54 P genes, 336 NC genes (including *MAPT* from the list of preselected hypothesis-driven genes), 265 PC genes, and 156 genes common to both the NC and PC gene sets (including *PHKA1*). Pooling together all classes of genes, we followed a DDS of unsupervised clustering and identified two clusters after Silhouette maximization (Figure 4C, Tables S10 and S11). Like VL, the two clusters were highly segregated and incorporated all *neg-corr* genes in one cluster, which was enriched in genes related to neuronal and synaptic function (Figure 4D), with all the *pos-corr* genes in the other cluster enriched in non-brain-related activities.

To validate some of the predictions done by our computational strategy linking neurodegeneration and neuropsychology with transcriptomics, we also performed neuropathological analyses of brain samples from DM1 patients. Neuropathological details of the cases are shown in Table 2. Overall, moderate cerebral atrophy with enlarged ventricles was found in every DM1 case. The main common microscopical alterations were pre-tangles, neurofibrillary tangles, and neuropil threads in selected regions of the brain stem and cerebrum. NFTs were stained with antibodies AT8 (P-tau Ser202/Thr205), P-tau Thr181, P-tau Ser422, 3Rtau, and 4Rtau. The distribution was variable depending on the case. Cases I09-254 and I13-132 were categorized as stages I/II of Braak; cases I09-255, I13-131, and I13-134 as stages III/IV; and cases I09-256 and I13-133 as stages V/VI.  $\beta$ -amyloid deposition was absent in five cases, and present in cases I13-134 and I13-133 corresponding to Thal phases 1 and 3, respectively. Mild astrocytic gliosis and microgliosis were observed in the same regions with NFTs. Dystrophic neurites, reactive astrocytes and active microglia were found around  $\beta$ -amyloid deposits in senile plaques. Lewy bodies were found in one case (I13-133) corresponding to stage 3 of Braak of Parkinson's pathology. TDP-43 proteinopathy was absent in every case. Mild to moderate small blood vessel disease, characterized by arteriolosclerosis and arteriolar hyalinosis, was



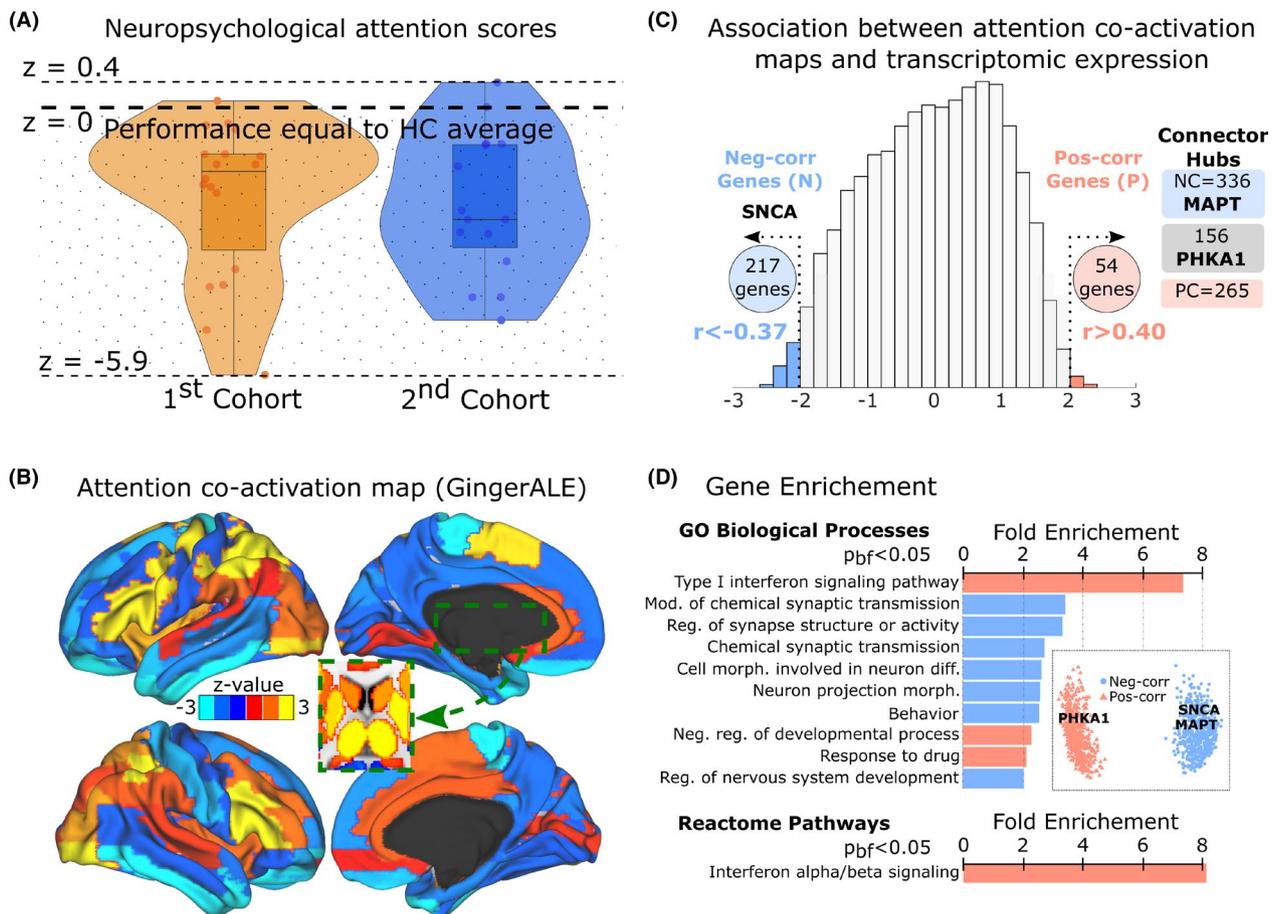
**FIGURE 3** Functional description of the genes with the highest association with volume loss (VL). (A) An all-to-all gene-expression similarity matrix identified the connector hub genes. A total of 1,086 genes were found, equal to the sum of the NC =452 (blue), PC =396 (red) and 238 common NC and PC genes (grey). (B) Two clusters were finally found that pooled all gene classes, the blue one contains the original N = 251 genes and the red one containing the original p = 101 genes. (C) Gene enrichment for *GO biological process* and *Reactome pathways*: in blue are the neg-corr genes and in red, the pos-corr genes. Abbreviations: Act., activation; CNS, central nervous system; Mod., modulation; Neg., negative; Pos., positive; Rec., receptor/s; Reg., regulation

common. This was accompanied by mild status cribosus in the striatum, thalamus and white matter in four cases. The white matter of the cerebrum was reduced in size and showed mild to moderate myelin pallor in five cases (I09-255, I13-131, I13-132, I13-133 and I13-134); the corpus callosum was thinner. Myelin changes, as seen with Klüver-Barrera were accompanied by loss of nerve fibres as visualized with anti-neurofilament antibodies. Axonal ballooning was absent. No inflammatory changes were observed in the white matter, with the exception of a few pigment-laden perivascular macrophages. Corpora amylacea were seldom observed in the subventricular region, perivascular spaces and subpial parenchyma. One case had suffered from a recent infarction in the left occipital cortex. Three cases had mild Purkinje cell loss (I13-131, I13-132 and I13-133), and two had mild atrophy of the inferior olives accompanied by astrocytic gliosis; representative tau pathology in Figure 5 lower panel.

Western blots of sarkosyl-insoluble fractions stained with anti-Tau-P-Ser214 antibodies revealed three bands of 68 kDa, 64 kDa, 60 kDa and an upper band of 73 kDa in one case with AD stage V. This pattern was in striking contrast with DM1 cases I09-254, I09-255 and I09-256 processed in the same membrane that apparently showed

two main bands of 58 and 55 kDa united by a smear, and no bands of upper molecular weights. The incubation of other membranes with anti-4Rtau and anti-3Rtau showed similar alterations. Unique bands of low molecular between 50 and 60 kDa were in striking contrast with the bands between 60 and 70 kDa observed in AD. Next, the seven cases with DM1 were run in the same gel and blotted with anti-P-tau-Ser214 antibodies. Two bands of 58 and 55 kDa characterized cases I09-255 and I09-256; one band of about 59/60 kDa characterized case I09-254. However, four bands of about 64, 60, 58 and 55 kDa were obtained in cases I13-131, I13-132 and I13-134. Finally, an additional band of about 68 kDa and bands of molecular weight below 50 kDa were identified in case I13-133. In agreement with the different phospho-tau bands depicted with the antibody P-tau-Ser214, the blots using anti-4Rtau and anti-3Rtau equally disclosed individual patterns of bands particularly linked to 3Rtau. Biochemical details of tau bands in every case are shown in Table 2; representative Western blots in Figure 5 upper panel.

We also studied the expression of selected genes from hippocampal tissue from DM1 patients. The study of messenger-RNA (mRNA) levels of genes related to synaptic vesicles revealed down-regulated expression of *SNAP25* ( $p = 0.035$ ) and a trend to reduced



**FIGURE 4** Data-driven strategy to define the association between the transcriptome and attention co-activation maps in DM1. (A) Attention scores measured as Z-scores for the two cohorts. Because the Z-scores were normalized to the values in the HCs, negative values of  $z$  indicate worse performance than the HCs. (B) Attention co-activation maps built with the GingerALE tool and projected onto the atlas. (C) Histograms of the spatial correlations between the Z scores of the attention maps and the transcriptional activity for each gene. The tail of N genes ( $z < -2$ , coloured in blue) includes the SNCA gene from the list of preselected genes, whereas the tail of the P genes ( $z > 2$ , red) does not include any of these. Following a procedure similar to that described in Figure 3A, we identified the PC genes (red), NC genes (blue, including MAPT), and those common to the NC and PC (grey, including PHKA1). (D) After pooling all classes of genes together and clustering, two groups were defined: one including all the neg-corr genes (blue, with SNCA and MAPT) and one with the pos-corr genes (red, with PHKA1). Gene enrichment for the tags GO biological process and Reactome pathways. As in Figure 3C, the two clusters also represented two separated functions: the neg-corr one correlated with neuronal functions, while the pos-corr correlated to non-brain functions. Abbreviations: Mod., Modulation; Neg., Negative; Pos., Positive; Reg., Regulation

expression of *SYP* and *SYN1* ( $p = 0.06$  and  $p = 0.08$ , respectively) between DM1 and NFT cases. Regarding the expression of genes related to structural components of synapse, *HOMER1* mRNA was significantly reduced in NFT cases compared with MA ( $p = 0.037$ ). *PCLO* and *ABLIM2* showed a tendency to decrease ( $p = 0.07$  and  $p = 0.07$ , respectively) in NFT cases compared with MA cases. None of this group of genes was altered in DM1. The study of the expression of GABAergic- and glutamatergic-related genes showed significantly reduced levels of *GABRD* transcript in NFT cases when compared with MA ( $p = 0.05$ ) and a tendency to reduced expression of *GABRG2* in NFT cases compared with MA ( $p = 0.10$ ) and DM1 cases ( $p = 0.09$ ) (Figure 6). Very striking, from the genes analysed in Figure 6, genes SNCA, SNAP25, SYP, SYN1, SLC17A7 (from synaptic vesicles; panel A), FRMPD4, PCLO, BSN, ARPC5L (from synapse structural components; panel B), and GABRA3, GRIA1 (from GABAergic- and glutamatergic-related genes; panel C), all of them appears in the neg-corr gene list

used in the enrichment obtained after linking VL or ND with transcriptomics (Table S8). The tendency showed in Figure 6 of this set of genes is to be downregulated with respect to MA and NFL, in agreement with the negative-association of these genes with VL.

## DISCUSSION

The AHBA provides information on the transcriptome across the brain in unprecedented detail, covering about 3,702 sampling sites and allowing activity patterns to be built for about 20,500 genes as a specific signature for each anatomic region. To date, its use has shed light on several fundamental aspects of the brain and their association with the transcriptome, such as myelination,<sup>64</sup> hierarchical cortical organization,<sup>65</sup> visuomotor integration<sup>41</sup> or large-scale connectivity.<sup>39,66,67</sup> In addition, it has provided data regarding pathologies,

TABLE 2 Neuropathological details of cases studied

case	age	NFT Braak	SP Thal	LB Braak	TDP-43	others	Western blotting P-tau-Ser214, tau bands
I09-254	60	Stage II	No	No	No	—	58/59 kDa,
I09-255	58	Stage IV	No	No	No	sbvd, sc, wmd, inf oliv	58 kDa, 55 kDa
I09-256	74	Stage V	No	No	No	inf oliv	58 kDa, 55 kDa
I13-131	67	Stage III/IV	No	No	No	sbvd, sc, wmd, Pc↓	weak 68 kDa, 64 kDa, 60 kDa, 58 kDa, 55 kDa
I13-132	58	Stage I-II	No	No	No	sbvd, wmd, Pc↓	64 kDa, 60 kDa, 58 kDa, 55 kDa
I13-133	68	Stage V/VI	Thal 3	LB 3	No	sbvd, sc, wmd, Pc↓	68 kDa, 64 kDa, 60 kDa, 58 kDa, 55 kDa, lower bands
I13-134	59	Stage III/IV	Thal 1	No	No	sbvd, sc, wmd, infarct	66 kDa, 64 kDa, 58 kDa, 55 kDa

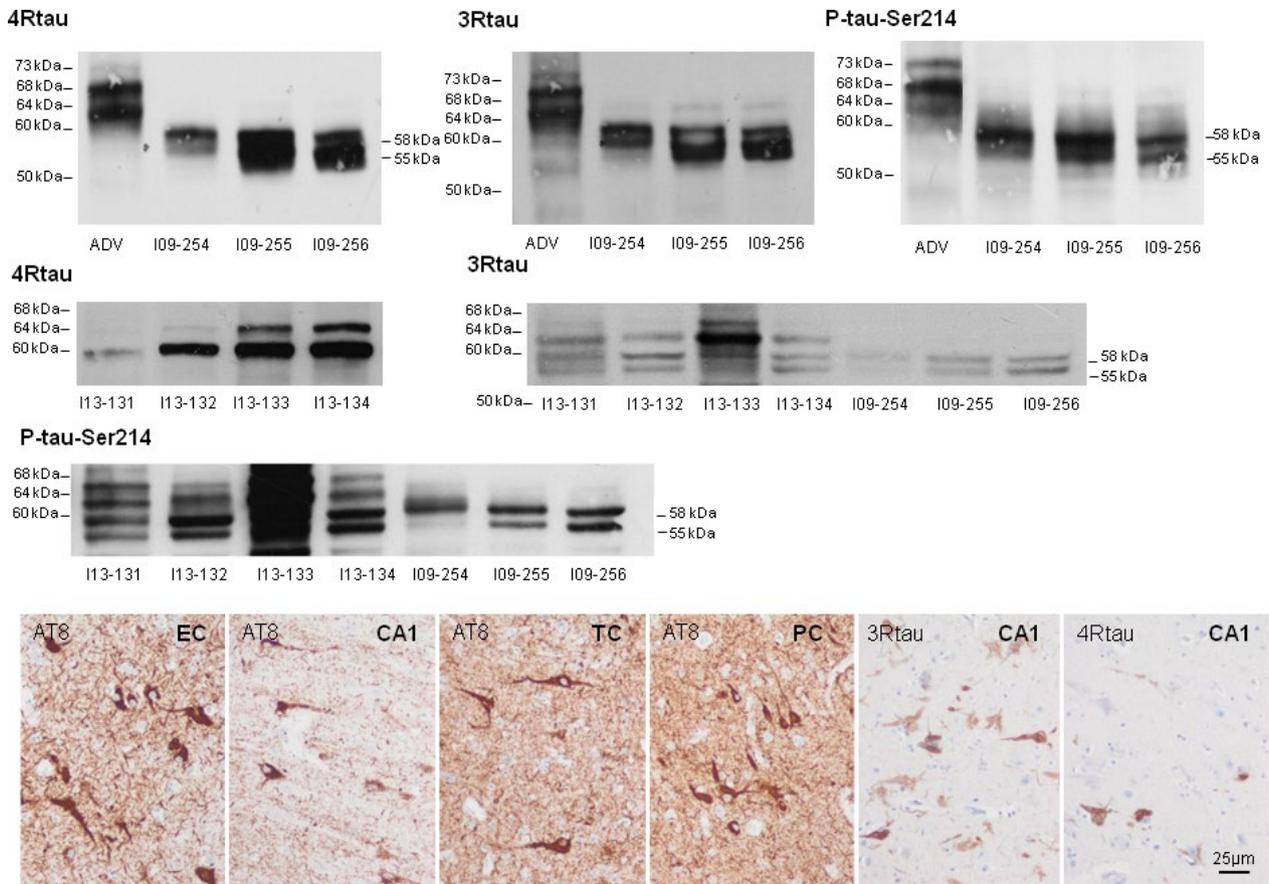
Abbreviations: *inf oliv*: mild to moderate nerve cell loss and mild astrocytosis in the inferior olives; *infarct*: acute occipital infarction; *LB Braak*: Lewy body pathology according to Braak stages; *NFT Braak*: Braak stages of neurofibrillary tangle pathology; *Pc↓*: mild to moderate loss of Purkinje cells in the cerebellum; *sbvd*: small blood vessel disease (arteriolosclerosis, arteriolar hyalinosis); *sc*: status cribosus and dilatation of perivascular spaces; *SP Thal*: phases of senile plaques according to Thal scale; *TDP-43*: TDP-43 proteinopathy; *wmd*: mild to moderate demyelination of the cerebral white matter.

suggesting novel molecular mechanisms underlying some disorders, such as autism spectrum disorder<sup>68</sup> or functional neurological disorder.<sup>69</sup> It is important to emphasize that our methodology, assessing the relationships between brain images associated with neurodegeneration and the entire transcriptome is complementary to other techniques, such as genome-wide association studies (GWAS),<sup>70</sup> simultaneously addressing genotype–phenotype associations from hundreds of thousands to millions of genetic variants in a data-driven manner. However, it is also important to note that such a vast number of multiple comparisons require very large samples to achieve statistical power, which is an important limitation for monogenic disorders like DM1.

Monogenic disorders are paradigmatic disease models of neurogenetic origin, whereby a mutation in a single gene can cause the disease. However, the alterations causing DM1 appear to propagate systemically to a larger proportion of the gene-to-gene interactome, affecting the function of several other genes. Here, we adopted a novel approach using the AHBA to study the transcriptomic correlates of brain damage and ND in DM1 patients, which to our knowledge has yet to be addressed. Our first finding is related precisely to the mutation located in the un-translated region of the DMPK gene, which triggers the disease and that we found to have the lowest DS (0.19) among the panel of the 23 hypothesis-driven preselected genes. Thus, *DMPK* expression was less reproducible in different brain regions and individuals, and consequently, it was more poorly enriched in terms of brain-related biological processes.<sup>57</sup> By contrast, the *SNCA* gene has the highest DS value (0.84). Our second finding is related to the clusters of spatial similarity among the hypothesis-driven genes in DM1. In particular, we found three clusters, one including the *DMPK* gene together with *LDB3*, *CAPN3* and *HSPB2*, the three latter genes associated with VL in DM1 patients. The second cluster contains several genes with a similar expression to *PHKA1*, which was associated with VL and ND in these patients. Finally, the third cluster contains the genes *SNCA*, *DMD* and *MAPT*, and as we show, it plays the most relevant role in VL and ND in DM1.

In relation to DM1 and *MAPT*, it has been shown that the alteration of Tau exon 2 splicing may produce predominant expression of low molecular weight Tau isoforms in DM1.<sup>71,72</sup> This is accompanied by abnormal splicing of Tau exon 6 in brain but not in muscle.<sup>73</sup> Mis-splicing of exon 10 has also been reported in some cases with DM1.<sup>74</sup> Splicing alterations of Tau in DM1 differ from those occurring during fetal development; mis-splicing of Tau in DM1 is resulting from a disease-associated mechanism.<sup>75</sup> The most common band pattern of phospho-Tau in Western blots of sarkosyl-insoluble fractions is the presence of a reduced band of 68/69 kDa and overexpression of Tau 55 kDa.<sup>72</sup> An abnormal pattern of phospho-Tau is also observed in the present series with marked individual variability. Low molecular phospho-Tau isoforms of about 58 and 55 kDa, as revealed with the P-Tau-Ser214 antibody, are present in almost every case, and they are the unique bands in two cases. Another case shows a predominant band of about 60 kDa. Three cases show four bands of about 64, 60, 58 and 55 kDa. One case is characterized by five bands of 68/69, 64, 60, 58 and 55 kDa. This is associated with variable individual patterns of 3RTau and 4RTau isoforms in sarkosyl-insoluble fractions.

The exploratory strategy DDS revealed the existence of two tails in the distribution of spatial-similarity between transcription activity and VL or ND. The cluster containing the pos-corr genes provided an enrichment towards non-brain-related functions, mainly interferon signalling, endothelial cell differentiation, angiogenesis or blood transport, processes known to be affected in DM1.<sup>2,76–78</sup> Pos-corr genes indicated that in those brain regions where more VL or ND existed, the genes are systematically transcribed more strongly, and vice versa, in the regions where the genes are more weakly expressed, VL or ND was less pronounced. It is important to emphasize that the association between gene expression and imaging or neuropsychology has not been found within a specific brain region, but it takes into consideration global multivariate VL and ND brain maps. Therefore, our relations between VL, ND and transcriptomics are in this sense fully multivariate, making it difficult to sort out towards

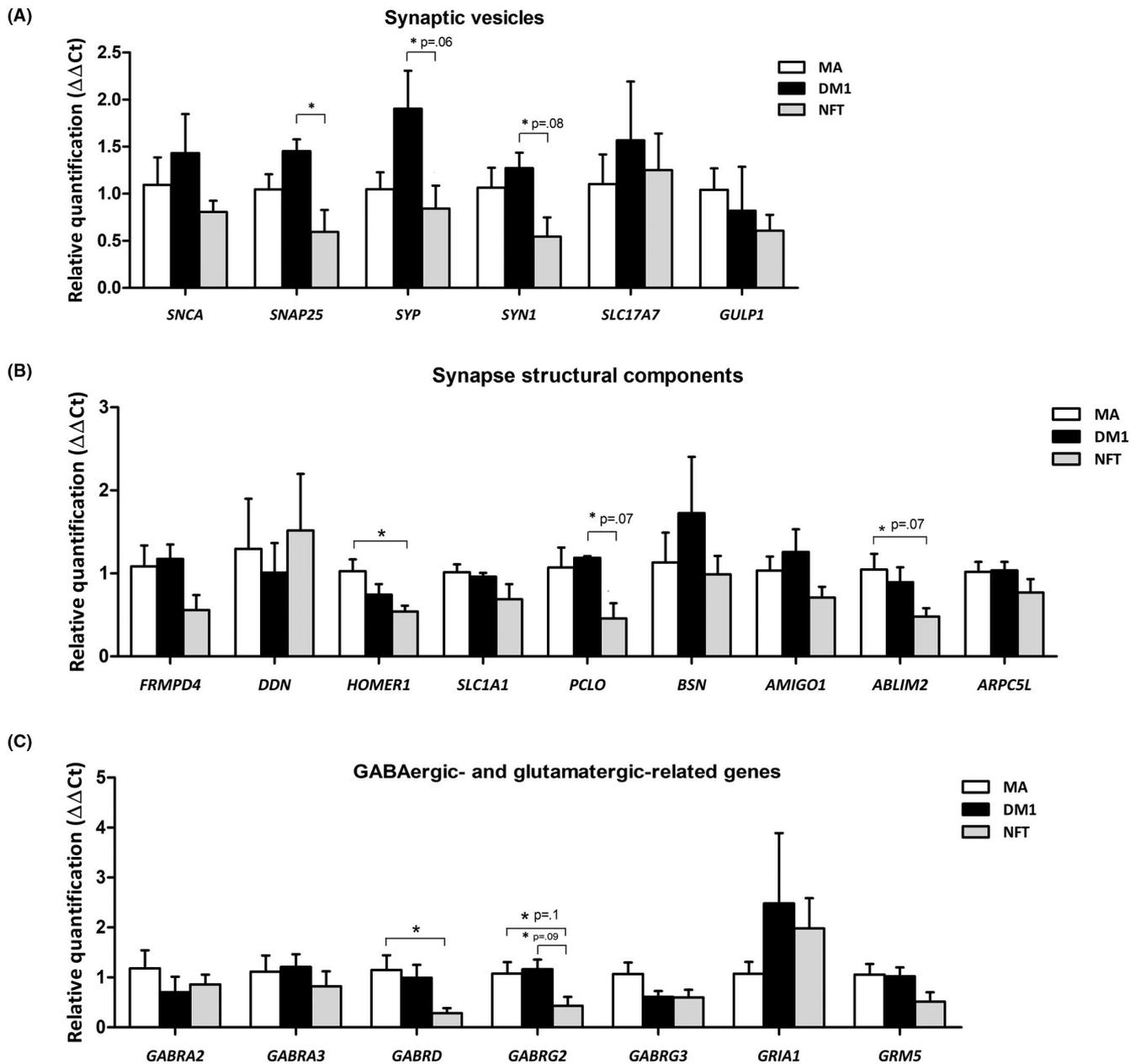


**FIGURE 5** Neuropathological analyses of brain samples revealed a high heterogeneous Tau-pathology in DM1, different to the one in AD. Upper panel: Gel electrophoresis and Western blotting of sarkosyl-insoluble fractions of the hippocampus incubated with antibodies against 3Rtau, 4Rtau, and phosphorylated tau at serine 214 (P-tau-Ser214) in DM1 cases, and one AD for control. AD is characterized by three bands of 68, 64 and 60 kDa and an upper weak band of 73 kDa. DM1 cases are characterized by the presence of lower bands of 58 and 55 kDa, and occasional upper bands of 60 and 64 kDa; the band of 68/69 kDa is barely present. This particular pattern, including differences from one case to another, is associated with low molecular weight 3Rtau and 4Rtau isoforms suggesting complex altered Tau splicing. Lower panel: Representative images of NFTs and neuropil threads in the entorhinal cortex (EC), CA1 area of the hippocampus (CA1), temporal cortex (TC), and parietal cortex (PC), visualized with the antibody AT8, and anti-3Rtau and 4Rtau antibodies. Paraffin sections slightly counterstained with haematoxylin; bar =25 µm

which specific region the triple correspondence (VL, ND and genetics) might converge.

The cluster containing the neg-corr genes incorporated *DMD*, *MAPT* and *SNCA*, providing instructions to produce respectively the proteins dystrophin, Tau and alpha-synuclein. These results are consistent with previous studies on DM1 patients, for instance, the existence of abnormalities in dystrophin,<sup>79,80</sup> Tau-positive degenerative neurites,<sup>81</sup> in agreement with our series where all showed phospho-Tau tangles and recently supported by common molecular pathways between the *DMPK* and *MAPT* genes,<sup>33</sup> and the presence of alpha-synuclein Lewy bodies,<sup>36</sup> also found in one of the seven patients assessed in this study by the neuropathological analyses. Moreover, the enrichment of this cluster was towards brain-related biological activities, such as synaptic vesicle processes (recycling, localization, endocytosis and exocytosis), and serotonin and dopamine neurotransmitter release (Table S12). These findings are in agreement with previous work in transgenic DM1 mice (DMSXL) showing synaptic protein

dysregulation,<sup>82,83</sup> events mediated by *RAB3A* upregulation and *SYN1* hyperphosphorylation. Importantly, the same authors in<sup>82</sup> validated these results from human tissues from DM1 patients. Here, we extend these results in a small series of three cases of hippocampal tissues from DM1 patients and assess the expression by RT-PCR of 22 genes related to synaptic vesicle function, synapse structure components and GABAergic and glutamatergic functions. We found down-regulated genes in DM1 more than in HCs (middle-aged individuals with no NFT pathology), although in a non-significant way, and more than in a second control group of AD-related pathology with similar NFT stage in a significant way for some of the genes such as *SYP* and *SYN1*. In relation to synapse structure components, our analyses have shown significant alterations in *PCLO* between DM1 and AD related with similar NFT control group, and the same occurred for *GABRG2* gene when looking into the function of GABAergic- and glutamatergic-related vesicles. Therefore, the comparison between DM1 cases and NFT cases linked to AD-related pathology with similar NFT Braak



**FIGURE 6** RNA expression in hippocampal tissues from DM1 patients revealed protein dysfunction of synaptic vesicle processes in DM1. mRNA expression of synapse-related genes was analysed in the hippocampus in middle-aged (MA) control cases, myotonic dystrophy 1 (DM1) cases, and cases with neurofibrillary tangles (NFT control) linked to AD-related pathology at middle and advanced stages (III-VI) of Braak and Braak. (A) Synaptic vesicles coding genes. (B) Synapse structural components coding genes. (C) GABAergic- and glutamatergic-related coding genes. All data were expressed as mean values  $\pm$  SEM. Differences between groups are statistically significant at  $*p < 0.05$ ; trends are indicated with \* and the exact  $p$  value

staging shows substantial differences. Whether these changes are due to different Tau species in DM1 and AD-related pathology or to other disease-specific factors deserves further study. It is important to emphasize that our analyses are complementary to those performed in the transgenic DM1 mice.<sup>82</sup> Although the animal model showed that the CTG expansion affected synaptic vesicle proteins, neurotransmission and behaviour, these mice did not have apparent Tau pathology, and therefore, synaptic alterations were probably linked to CTG expansion rather than Tau.

Our human data showed the presence of NFT in all tissues together with an aberrant synaptic vesicle dysfunction.

The major limitation of using AHBA transcriptomics data is that this atlas was obtained from brain samples from healthy young adults, and not from DM1 patients' brain tissues, so the relations studied here between *pathological* VL or ND and *healthy* transcriptomics still might shed some light on systemic neurological aspects. Indeed, our results showed the participation of some of the genes previously reported to play a role in the pathophysiology of DM1.

Despite this limitation, we have presented here neuropathological analyses of brain tissues and samples from DM1 patients to confirm a high heterogeneity of phospho-Tau fingerprints in DM1 patients, together with synaptic vesicles downregulated genes mediating DM1 significantly more than other AD-related with similar NFT diseases.

In conclusion, we have studied two different cohorts of DM1 patients, each one well-matched to a group HCs, and by employing a DDS that addresses all the hypothesis-driven preselected genes for DM1, *DMD*, *SNCA* and *MAPT* were seen to have a major influence on brain damage and ND. Moreover, we found an enrichment of key biological processes in the CNS, confirmed by a RNA analysis of brain samples from DM1 patients, such as synaptic vesicle cycling, recycling and dynamics, and also serotonin and dopamine neurotransmitter signalling. Further studies should clarify whether the interactions between *DMD*, *SNCA* and *MAPT* can be generalized to other degenerative or developmental conditions, or if these are specific to DM1.

### ACKNOWLEDGEMENTS

We wish to thank Prof. Virginia Arechavala for providing us with an updated list of relevant genes in DM1, some of which were considered in our study. J.M.C. is funded by Ikerbasque: The Basque Foundation for Science and from the Ministerio de Economía, Industria y Competitividad (Spain) and FEDER (grant DPI2016-79874-R), and from the Department of Economic and Infrastructure Development of the Basque Country (Elkartek Program, KK-2018/00032 and KK-2018/00090). A.L.d.M. was founded by the Institute of Health Carlos III co-founded by Fondo Europeo de Desarrollo Regional-FEDER (grant PI17/01841), CIBERNED (grant 609), and La Caixa Foundation (grant HR17-00268). A.S. was founded by the Institute of Health Carlos III co-founded by Fondo Europeo de Desarrollo Regional-FEDER (grant PI17/01231), and the Basque Government (grant SAIO08-PE08BF01). A.J.M. was partially funded by Euskampus Fundazioa and a predoctoral grant from the Basque Government (PRE\_2019\_1\_0070). G.L. was founded by a predoctoral grant from the Basque Government (PRE\_2016\_1\_0187). I.F. was founded from 'la Caixa' Foundation under the agreement LCF/PR/HR19/52160007 and was also supported by the Ministry of Economy and Competiveness, Institute of Health Carlos III (co-funded by European Regional Development Fund, ERDF, a way to build Europe): FIS PI17/000809.

### CONFLICT OF INTERESTS

The authors declare no competing interests.

### AUTHOR CONTRIBUTIONS

G.L., A.S. and A.L.d.M. recruited the participants; G.L. and A.S. performed the neuropsychological evaluations; M.C.C. carried out the autopsies of cases with myotonic dystrophy; P.A.B. carried out the PCR validation of selected genes; I.F. carried out the neuropathological study of the cases, and the biochemical study of tau; A.J.M. processed the MRI data; A.J.M., I.D., J.S. and J.M.C. designed and supervised the

imaging and statistical analyses; A.J.M. performed all the analyses; A.J.M., I.F. and J.M.C. made the figures; A.L.d.M. and J.M.C. supervised the research; all the authors wrote the manuscript and agreed with its submission; A.L.d.M. and J.M.C. had equal last-author contribution.

### ETHICAL APPROVAL

All participants were informed about the study and offered their signed informed consent. The study was approved by the Ethical Committee of the Donostia University Hospital (code DMRM-2017-01), and it was carried out in accordance with the tenets for human research laid out in the Helsinki Declaration.

### PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/nan.12725>.

### DATA AVAILABILITY STATEMENT

All neuroimaging data supporting the findings of this study are available from the corresponding author upon reasonable request. The genetic data used are available from the Allen Human Brain Atlas (<https://human.brain-map.org>).

### ORCID

Antonio Jimenez-Marin  <https://orcid.org/0000-0002-6053-6101>

Ibai Diez  <https://orcid.org/0000-0001-5769-0178>

Garazi Labayru  <https://orcid.org/0000-0002-7304-320X>

Andone Sistiaga  <https://orcid.org/0000-0003-2511-0911>

Pol Andres-Benito  <https://orcid.org/0000-0003-3000-0338>

Jorge Sepulcre  <https://orcid.org/0000-0002-4199-2304>

Isidro Ferrer  <https://orcid.org/0000-0001-9888-8754>

Jesus M. Cortes  <https://orcid.org/0000-0002-9059-8194>

### ENDNOTES

<sup>1</sup> AAL regions were eroded with a Gaussian kernel with a full width at half maximum (FWHM) equal to 2 mm, thereby eliminating false-positive sampling sites, i.e., those that do not belong to the region of interest but to one in the neighborhood.

<sup>2</sup> To calculate the correlations, the values corresponding to the different regions in the atlas were considered as observations.

### REFERENCES

1. Avaria M, Patterson V. Myotonic dystrophy: relative sensitivity of symptoms signs and abnormal investigations. *Ulster Med J*. 1994;63(2):151-154.
2. Lund M, Diaz LJ, Ranthe MF, et al. Cardiac involvement in myotonic dystrophy: a nationwide cohort study. *Eur Heart J*. 2014;35(32):2158-2164.
3. Bogaard JM, van der Meché FG, Hendriks I, Ververs C. Pulmonary function and resting breathing pattern in myotonic dystrophy. *Lung*. 1992;170(3):143-153.
4. Ørngreen MC, Arlien-Søborg P, Duno M, Hertz JM, Vissing J. Endocrine function in 97 patients with myotonic dystrophy type 1. *J Neurol*. 2012;259(5):912-920.
5. Laberge L, Gagnon C, Dauvilliers Y. Daytime sleepiness and myotonic dystrophy. *Curr Neurol Neurosci Rep*. 2013;13(4):340.

6. Modoni A, Silvestri G, Pomponi MG, Mangiola F, Tonali PA, Marra C. Characterization of the pattern of cognitive impairment in myotonic dystrophy type 1. *Arch Neurol*. 2004;61(12):1943-1947.
7. Harper PS. Myotonic dystrophy—the facts: a book for patients and families. Oxford ; New York: Oxford University Press; 2002. 113 p. (Oxford medical publications).
8. de Munain AL, Blanco A, Emparanza JI, et al. Prevalence of myotonic dystrophy in Guipuzcoa (Basque Country, Spain). *Neurology*. 1993;43(8):1573.
9. Antonini G, Mainero C, Romano A, et al. Cerebral atrophy in myotonic dystrophy: a voxel based morphometric study. *J Neurol Neurosurg Psychiatry*. 2004;75(11):1611-1613.
10. Weber YG, Roeblyng R, Kassubek J, et al. Comparative analysis of brain structure, metabolism, and cognition in myotonic dystrophy 1 and 2. *Neurology*. 2010;74(14):1108-1117.
11. Schneider-Gold C, Bellenberg B, Prehn C, et al. Cortical and subcortical grey and white matter atrophy in myotonic dystrophies Type 1 and 2 is associated with cognitive impairment, depression and daytime sleepiness. *PLoS One*. 2015;10(6):e0130352.
12. Fukuda H, Horiguchi J, Ono C, Ohshita T, Takaba J, Ito K. Diffusion tensor imaging of cerebral white matter in patients with myotonic dystrophy. *Acta Radiol*. 2005;46(1):104-109.
13. Ota M, Sato N, Ohya Y, et al. Relationship between diffusion tensor imaging and brain morphology in patients with myotonic dystrophy. *Neurosci Lett*. 2006;407(3):234-239.
14. Minnerop M, Weber B, Schoene-Bake J-C, et al. The brain in myotonic dystrophy 1 and 2: evidence for a predominant white matter disease. *Brain*. 2011;134(12):3530-3546.
15. Wozniak JR, Mueller BA, Bell CJ, Muetzel RL, Lim KO, Day JW. Diffusion tensor imaging reveals widespread white matter abnormalities in children and adolescents with myotonic dystrophy type 1. *J Neurol*. 2013;260(4):1122-1131.
16. Caso F, Agosta F, Peric S, et al. Cognitive impairment in myotonic dystrophy type 1 is associated with white matter damage. *PLoS One*. 2014;9(8):e104697.
17. Baldanzi S, Cecchi P, Fabbri S, et al. Relationship between neuropsychological impairment and grey and white matter changes in adult-onset myotonic dystrophy type 1. *NeuroImage Clin*. 2016;12:190-197.
18. Zanigni S, Evangelisti S, Giannoccaro MP, et al. Relationship of white and gray matter abnormalities to clinical and genetic features in myotonic dystrophy type 1. *NeuroImage Clin*. 2016;11:678-685.
19. van Dorst M, Okkersen K, Kessels RPC, et al. Structural white matter networks in myotonic dystrophy type 1. *NeuroImage Clin*. 2019;21:101615.
20. Labayru G, Diez I, Sepulcre J, et al. Regional brain atrophy in gray and white matter is associated with cognitive impairment in Myotonic Dystrophy type 1. *NeuroImage Clin*. 2019;24:102078.
21. Labayru G, Jimenez-Marin A, Fernández E, et al. Neurodegeneration trajectory in pediatric and adult/late DM1: A follow-up MRI study across a decade. *Ann Clin Transl Neurol*. 2020;7(10):1802-1815.
22. Wozniak JR, Mueller BA, Ward EE, Lim KO, Day JW. White matter abnormalities and neurocognitive correlates in children and adolescents with myotonic dystrophy type 1: a diffusion tensor imaging study. *Neuromuscul Disord NMD*. 2011;21(2):89-96.
23. Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet*. 2007;39(1):17-23.
24. Brook JD, McCurrach ME, Harley HG, et al. Molecular basis of myotonic dystrophy: Expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell*. 1992;68(4):799-808.
25. Lee JE, Cooper TA. Pathogenic mechanisms of myotonic dystrophy. *Biochem Soc Trans*. 2009;37(6):1281-1286.
26. Charizanis K, Lee K-Y, Batra R, et al. Muscleblind-like 2-mediated alternative splicing in the developing brain and dysregulation in myotonic dystrophy. *Neuron*. 2012;75(3):437-450.
27. Lueck JD, Lungu C, Mankodi A, et al. Chloride channelopathy in myotonic dystrophy resulting from loss of posttranscriptional regulation for CLCN1. *Am J Physiol-Cell Physiol*. 2007;292(4):C1291-C1297.
28. Santoro M, Masciullo M, Bonvissuto D, Bianchi MLE, Michetti F, Silvestri G. Alternative splicing of human insulin receptor gene (INSR) in type I and type II skeletal muscle fibers of patients with myotonic dystrophy type 1 and type 2. *Mol Cell Biochem*. 2013;380(1-2):259-265.
29. Fugier C, Klein AF, Hammer C, et al. Misregulated alternative splicing of BIN1 is associated with T tubule alterations and muscle weakness in myotonic dystrophy. *Nat Med*. 2011;17(6):720-725.
30. Mankodi A, Logigian E, Callahan L, et al. Myotonic dystrophy in transgenic mice expressing an expanded CUG repeat. *Science*. 2000;289(5485):1769-1773.
31. Sistiaga A, Urreta I, Jodar M, et al. Cognitive/personality pattern and triplet expansion size in adult myotonic dystrophy type 1 (DM1): CTG repeats, cognition and personality in DM1. *Psychol Med*. 2010;40(3):487-495.
32. Winblad S, Lindberg C, Hansen S. Cognitive deficits and CTG repeat expansion size in classical myotonic dystrophy type 1 (DM1). *Behav Brain Funct BBF*. 2006;15(2):16.
33. Carpentier C, Ghanem D, Fernandez-Gomez FJ, et al. Tau exon 2 responsive elements deregulated in myotonic dystrophy type I are proximal to exon 2 and synergistically regulated by MBNL1 and MBNL2. *Biochim Biophys Acta*. 2014;1842(4):654-664.
34. Jiang H, Mankodi A, Swanson MS, Moxley RT, Thornton CA. Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Hum Mol Genet*. 2004;13(24):3079-3088.
35. Annic A, Devos D, Destée A, et al. Early dopasensitive Parkinsonism related to myotonic dystrophy type 2. *Mov Disord Off J Mov Disord Soc*. 2008;23(14):2100-2101.
36. Itoh K, Mitani M, Kawamoto K, et al. Neuropathology does not correlate with regional differences in the extent of expansion of CTG repeats in the brain with myotonic dystrophy type 1. *Acta Histochem Cytochem*. 2010;43(6):149-156.
37. Serra L, Petrucci A, Spanò B, et al. How genetics affects the brain to produce higher-level dysfunctions in myotonic dystrophy type 1. *Funct Neurol*. 2015;30(1):21-31.
38. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*. 2012;489(7416):391-399.
39. Diez I, Sepulcre J. Neurogenetic profiles delineate large-scale connectivity dynamics of the human brain. *Nat Commun*. 2018;9(1):3876.
40. Sepulcre J, Grothe MJ, d'Oleire Uquillas F, et al. Neurogenetic contributions to amyloid beta and tau spreading in the human cortex. *Nat Med*. 2018;24(12):1910-1918.
41. Bueichekú E, Aznárez-Sanado M, Diez I, et al. Central neurogenetic signatures of the visuomotor integration system. *Proc Natl Acad Sci*. 2020;117(12):6836-6843.
42. Mathieu J, Boivin H, Meunier D, Gaudreault M, Begin P. Assessment of a disease-specific muscular impairment rating scale in myotonic dystrophy. *Neurology*. 2001;56(3):336-340.
43. Wechsler D, Seisdedos CN. WAIS-III escala de inteligencia de Wechsler para adultos-III. 1999.
44. Lezak MD, Howieson DB, Loring DW, Hannay HJ, Fischer JS. Neuropsychological assessment, 4th ed. New York, NY, US: Oxford University Press; 2004. xiv, 1016 p. (Neuropsychological assessment, 4th ed.).
45. Casals-Coll M, Sánchez-Benavides G, Quintana M, et al. Estudios normativos españoles en población adulta joven (proyecto NEURONORMA jóvenes): normas para los test de fluencia verbal. *Neurología*. 2013;28(1):33-40.

46. Pena-Casanova J, Quinones-Ubeda S, Gramunt-Fombuena N, et al. Spanish multicenter normative studies (NEURONORMA Project): norms for verbal fluency tests. *Arch Clin Neuropsychol*. 2009;24(4):395-411.
47. Rey A. *FTest de Copia y Reproducción de una Figura Compleja*. S.A.: Tea Ediciones; 2009.
48. Adams N. Contributions to neuropsychological assessment: a clinical manual. *Neurology*. 1995;45(8):1637.
49. Douaud G, Smith S, Jenkinson M, et al. Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. *Brain*. 2007;130(9):2375-2386.
50. De Las HJ, Diez I, Jimenez-Marin A, et al. Brain circuit alterations and cognitive disability in late-onset cobalamin D disorder. *J Clin Med*. 2020;9(4):990.
51. Good CD, Johnsrude IS, Ashburner J, Henson RNA, Friston KJ, Frackowiak RSJ. A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*. 2001;14(1):21-36. <https://doi.org/10.1006/nimg.2001.0786>.
52. Andersson JLR, Jenkinson M, Smith S. Non-linear registration, aka spatial normalisation. FMRIB Technical Report TR07JA2. *Oxf Cent Funct Magn Reson Imaging Brain Dep Clin Neurol Oxf Univ Oxf UK*. 2007.
53. Smith SM, Jenkinson M, Johansen-Berg H, et al. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *NeuroImage*. 2006;31(4):1487-1505.
54. Smith SM, Nichols TE. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *NeuroImage*. 2009;44(1):83-98.
55. Arnatkevic lütè A, Fulcher BD, Fornito A. A practical guide to linking brain-wide gene expression and neuroimaging data. *NeuroImage*. 2019;01(189):353-367.
56. Arloth J, Bader DM, Röh S, Re-Annotator AA. Re-annotator: annotation pipeline for microarray probe sequences. *PLoS One*. 2015;10(10):e0139516
57. Hawrylycz M, Miller JA, Menon V, et al. Canonical genetic signatures of the adult human brain. *Nat Neurosci*. 2015;18(12):1832-1844.
58. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*. 2002;15(1):273-289.
59. Laird AR, Lancaster JL, Fox PT. BrainMap: the social evolution of a human brain mapping database. *Neuroinformatics*. 2005;3(1):65-78.
60. Eickhoff SB, Laird AR, Grefkes C, Wang LE, Zilles K, Fox PT. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp*. 2009;30(9):2907-2926.
61. Burt JB, Helmer M, Shinn M, Anticevic A, Murray JD. Generative modeling of brain maps with spatial autocorrelation. *NeuroImage*. 2020;220:117038.
62. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. *Nat Genet*. 2000;25(1):25-29.
63. Fabregat A, Sidiropoulos K, Viteri G, et al. Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics*. 2017;18(1):142.
64. Ritchie J, Pantazatos SP, French L. Transcriptomic characterization of MRI contrast with focus on the T1-w/T2-w ratio in the cerebral cortex. *NeuroImage*. 2018;174:504-517.
65. Burt JB, Demirtaş M, Eckner WJ, et al. Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nat Neurosci*. 2018;21(9):1251-1259.
66. Fornito A, Arnatkevičiūtė A, Fulcher BD. Bridging the Gap between Connectome and Transcriptome. *Trends Cogn Sci*. 2019;23(1):34-50.
67. Romero-Garcia R, Whitaker KJ, Váša F, et al. Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex. *NeuroImage*. 2018;171:256-267.
68. Romero-Garcia R, Warrier V, Bullmore ET, Baron-Cohen S, Bethlehem RAI. Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism. *Mol Psychiatry*. 2019;24(7):1053-1064.
69. Diez I, Larson AG, Nakhate V, et al. Early-life trauma endophenotypes and brain circuit-gene expression relationships in functional neurological (conversion) disorder. *Mol Psychiatry*. 2020.
70. Tam V, Patel N, Turcotte M, Bossé Y, Paré G, Meyre D. Benefits and limitations of genome-wide association studies. *Nat Rev Genet*. 2019;20(8):467-484.
71. Vermersch P, Sergeant N, Ruchoux MM, et al. Specific tau variants in the brains of patients with myotonic dystrophy. *Neurology*. 1996;47(3):711-717.
72. Sergeant N, Sablonnière B, Schraen-Maschke S, et al. Dysregulation of human brain microtubule-associated tau mRNA maturation in myotonic dystrophy type 1. *Hum Mol Genet*. 2001;10(19):2143-2155.
73. Leroy O, Wang J, Maurage C-A, et al. Brain-specific change in alternative splicing of Tau exon 6 in myotonic dystrophy type 1. *Biochim Biophys Acta*. 2006;1762(4):460-467.
74. Dhaenens CM, Tran H, Frandemiche M-L, et al. Mis-splicing of Tau exon 10 in myotonic dystrophy type 1 is reproduced by overexpression of CELF2 but not by MBNL1 silencing. *Biochim Biophys Acta*. 2011;1812(7):732-742.
75. Ghanem D, Tran H, Dhaenens C-M, et al. Altered splicing of Tau in DM1 is different from the foetal splicing process. *FEBS Lett*. 2009;583(4):675-679.
76. Fragola PV, Luzzi M, Calò L, et al. Cardiac involvement in myotonic dystrophy. *Am J Cardiol*. 1994;74(10):1070-1072.
77. Florek RC, Triffon DW, Mann DE, Ringel SP, Reiter MJ. Electrocardiographic abnormalities in patients with myotonic dystrophy. *West J Med*. 1990;153(1):24-27.
78. Rhodes JD, Lott MC, Russell SL, et al. Activation of the innate immune response and interferon signalling in myotonic dystrophy type 1 and type 2 cataracts. *Hum Mol Genet*. 2012;21(4):852-862.
79. Nakamori M, Kimura T, Kubota T, et al. Aberrantly spliced -dystrobrevin alters -syntrophin binding in myotonic dystrophy type 1. *Neurology*. 2008;70(9):677-685.
80. Thangarajh M, Hendriksen J, McDermott MP, et al. Relationships between DMD mutations and neurodevelopment in dystrophinopathy. *Neurology*. 2019;93(17):e1597-e1604. <https://doi.org/10.1212/WNL.00000000000008363>
81. Kiuchi A, Otsuka N, Namba Y, Nakano I, Tomonaga M. Presenile appearance of abundant Alzheimer's neurofibrillary tangles without senile plaques in the brain in myotonic dystrophy. *Acta Neuropathol (Berl)*. 1991;82(1):1-5.
82. Hernández-Hernández O, Guiraud-Dogan C, Sicot G, et al. Myotonic dystrophy CTG expansion affects synaptic vesicle proteins, neurotransmission and mouse behaviour. *Brain*. 2013;136(3):957-970.
83. Hernández-Hernández O, Sicot G, Dinca DM, et al. Synaptic protein dysregulation in myotonic dystrophy type 1: Disease neuropathogenesis beyond missplicing. *Rare Dis*. 2013;1(1):e25553.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Jimenez-Marin A, Diez I, Labayru G, et al. Transcriptional signatures of synaptic vesicle genes define myotonic dystrophy type I neurodegeneration. *Neuropathol Appl Neurobiol*. 2021;00:1-17. <https://doi.org/10.1111/nan.12725>