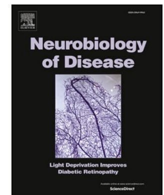
Neurobiology of Disease 162 (2022) 105578

Contents lists available at ScienceDirect

Neurobiology of Disease

journal homepage: www.elsevier.com/locate/ynbdi



# Novel Machado-Joseph disease-modifying genes and pathways identified by whole-exome sequencing

Mafalda Raposo [a,b,*], Conceição Bettencourt [c], Ana Rosa Vieira Melo [b,a], Ana F. Ferreira [b,a], Isabel Alonso [a], Paulo Silva [a], João Vasconcelos [d], Teresa Kay [e], Maria Luiza Saraiva-Pereira [f,n], Marta D. Costa [g,h], Daniela Vilasboas-Campos [g,h], Bruno Filipe Bettencourt [i], Jácome Bruges-Armas [i,j], Henry Houlden [k], Peter Heutink [l], Laura Bannach Jardim [m,n], Jorge Sequeiros [a], Patrícia Maciel [g,h], Manuela Lima [b,a]

[a] *Instituto de Biologia Molecular e Celular (IBMC), Instituto de Investigação e Inovação em Saúde (i3S), Universidade do Porto, Porto, Portugal*
[b] *Faculdade de Ciências e Tecnologia, Universidade dos Açores, Ponta Delgada, Portugal*
[c] *Department of Neurodegenerative Disease and Queen Square Brain Bank for Neurological Disorders, UCL Queen Square Institute of Neurology, London, UK*
[d] *Departamento de Neurologia, Hospital do Divino Espírito Santo, Ponta Delgada, Portugal*
[e] *Departamento de Genética Clínica, Hospital D. Estefânia, Lisboa, Portugal*
[f] *Departamento de Bioquímica, Universidade Federal do Rio Grande do Sul, Brazil*
[g] *Instituto de Investigação em Ciências da Vida e Saúde (ICVS), Escola de Medicina, Universidade do Minho, Braga, Portugal*
[h] *ICVS/3B's - Laboratório Associado, Braga/Guimarães, Portugal*
[i] *Serviço Especializado de Epidemiologia e Biologia Molecular (SEEBMO), Hospital de Santo Espírito da Ilha Terceira (HSEIT), Angra do Heroísmo, Azores, Portugal*
[j] *CHRC - Comprehensive Health Research Centre, Faculdade de Ciências Médicas & CEDOC - Chronic Diseases Research Center, Universidade Nova de Lisboa, Lisboa, Portugal*
[k] *Department of Molecular Neuroscience, Institute of Neurology, University College London and Neurogenetics Unit, National Hospital for Neurology and Neurosurgery, University College London Hospitals, London, United Kingdom, London*
[l] *German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany*
[m] *Departamento de Medicina Interna, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil*
[n] *Serviço de Genética Médica/Centro de Pesquisa Clínica e Experimental, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil*

## ARTICLE INFO

*Keywords:*
MJD
SCA3
Spinocerebellar ataxia
Polyglutamine disease
Age at onset
Genetic modifier

## ABSTRACT

Machado-Joseph disease (MJD/SCA3) is a neurodegenerative polyglutamine disorder exhibiting a wide spectrum of phenotypes. The abnormal size of the (CAG)n at *ATXN3* explains ~55% of the age at onset variance, suggesting the involvement of other factors, namely genetic modifiers, whose identification remains limited. Our aim was to find novel genetic modifiers, analyse their epistatic effects and identify disease-modifying pathways contributing to MJD variable expressivity. We performed whole-exome sequencing in a discovery sample of four age at onset concordant and four discordant first-degree relative pairs of Azorean patients, to identify candidate variants which genotypes differed for each discordant pair but were shared in each concordant pair. Variants identified by this approach were then tested in an independent multi-origin cohort of 282 MJD patients. Whole-exome sequencing identified 233 candidate variants, from which 82 variants in 53 genes were prioritized for downstream analysis. Eighteen disease-modifying pathways were identified; two of the most enriched pathways were relevant for the nervous system, namely the neuregulin signaling and the agrin interactions at neuromuscular junction. Variants at *PARD3, NFKB1, CHD5, ACTG1, CFAP57, DLGAP2, ITGB1, DIDO1* and *CERS4* modulate age at onset in MJD, with those identified in *CFAP57, ACTG1* and *DIDO1* showing consistent effects across cohorts of different geographical origins. Network analyses of the nine novel MJD modifiers highlighted

*Abbreviations:* MJD, Machado-Joseph disease; polyQ, polyglutamine; SCA, spinocerebellar ataxia; CAG, cytosine-adenine-guanine; AO, age at onset; GWAS, genome-wide association studies; WES, whole-exome sequencing; REF, reference sequence; VAR, variant sequence; $CAG_E$, number of CAG repeats in the expanded *ATXN3* allele; CVC, cross-validation consistency; RNAi, RNA interference.

* Corresponding author at: Instituto de Investigação e Inovação em Saúde (i3S), Universidade do Porto, Rua Alfredo Allen, 208, 4200-135 Porto, Portugal.
*E-mail addresses:* msraposo@i3s.up.pt (M. Raposo), c.bettencourt@ucl.ac.uk (C. Bettencourt), ana.fb.ferreira@uac.pt (A.F. Ferreira), paulo.silva@ibmc.up.pt (P. Silva), mlpereira@hcpa.edu.br (M.L. Saraiva-Pereira), martacosta@med.uminho.pt (M.D. Costa), h.houlden@ucl.ac.uk (H. Houlden), Peter.Heutink@dzne.de (P. Heutink), ljardim@hcpa.edu.br (L.B. Jardim), jorge.sequeiros@ibmc.up.pt (J. Sequeiros), pmaciel@med.uminho.pt (P. Maciel), maria.mm.lima@uac.pt (M. Lima).

https://doi.org/10.1016/j.nbd.2021.105578
Received 30 July 2021; Received in revised form 8 November 2021; Accepted 2 December 2021
Available online 3 December 2021
0969-9961/© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

several important molecular interactions, including genes/proteins previously related with MJD pathogenesis, namely between ACTG1/APOE and VCP/ITGB1. We describe novel pathways, modifiers, and their interaction partners, providing a broad molecular portrait of age at onset modulation to be further exploited as new disease-modifying targets for MJD and related diseases.

## 1. Introduction

Variable expressivity, namely in age at onset (AO) of symptoms is a hallmark of polyglutamine (polyQ) disorders, a group of genetically determined late-onset neurodegenerative diseases including Huntington disease and the most common dominant ataxias (spinocerebellar ataxias SCA1, SCA2, SCA3, SCA6, SCA7 and SCA17). At each of these disease *loci*, exonic CAG repeats beyond a critical threshold cause disease (reviewed in Gatchel and Zoghbi (2005)). The number of repeat units in the expanded allele, however, only partially correlates with AO (Gusella et al., 2014; Chen et al., 2018). This incomplete correlation, observed in variable degree in all polyQ disorders, suggests the involvement of modifying factors, namely genetic. Evidence so far seems to indicate that such genetic modifiers operate in a highly complex system of small to moderate effects of multiple modifier genes, which interact together and with environmental factors, resembling a polygenic disease (Gusella et al., 2014; Chen et al., 2018; Génin et al., 2008).

Machado-Joseph disease/spinocerebellar ataxia type 3 (MJD/SCA3) is the most frequent dominant ataxia worldwide; it displays a wide variability in AO (average of ~40 years; range from 4 to 78 years) (Carvalho et al., 2008; Tezenas du Montcel et al., 2014; Raposo et al., 2015), only about 55% of which is explained by size of the (CAG)n tract length (de Mattos et al., 2019). An expanded polyglutamine tract consensually exceeding 51 repeats (reviewed in Bettencourt and Lima (2011)), triggers disease through a cascade of pathological events, leading to neuronal dysfunction and loss (reviewed in Da Silva et al. (2019)). MJD remains untreatable and the discovery of disease-modifying genes may pinpoint novel molecular targets for drug development. Moreover, the identification of such modifiers should allow a better prediction of AO, providing useful information for genetic counselling and patient stratification for clinical trials design. Previous attempts have used mostly candidate gene approaches, proposing as disease modifiers the size of repeats at other (CAG)n disease *loci* (Tezenas du Montcel et al., 2014); Raposo et al., 2015; Jardim et al., 2003, allelic variants at the apolipoprotein E (APOE) *locus* (Bettencourt et al., 2011; Peng et al., 2014) or in DNA repair genes (Bettencourt et al., 2016; Mergener et al., 2020). Difficulties in replication, however, are widely acknowledged and attributed to sample size, study design or population-specific allelic frequencies at the candidate loci (Chen et al., 2018; Raposo et al., 2015). Unbiased genome-wide association studies (GWAS) may circumvent some of these constrain, allowing the simultaneous identification of several modifier variants and their interactions (Petersen et al., 2017), as well as point to modifying pathways. Recently, a large GWAS suggested nine associated *loci* as AO modifiers of MJD, explaining 8% of AO variance (Akçimen et al., 2020). Nevertheless, a considerable part of AO variance remained unanswered and disease-modifying pathways remain to be clarified.

The context of the Azores islands (Portugal), where MJD reaches the highest known prevalence worldwide (de Araújo et al., 2016), provides important advantages for the quest for novel modifiers, including a more homogeneous genetic background is expected, since the islands had a limited number of founders and were subjected to some degree of geographical isolation (Santos et al., 2003; Montiel et al., 2005). Furthermore, extended genealogies of the local MJD families (Lima et al., 1997; Lima et al., 1998), coupled with the regular follow-up of patients for over 20 years, should empower studies using this cohort.

As in other monogenic diseases (e.g., Rahit and Tarailo-Graovac, 2020), major gaps remain in the identification and understanding of genetic modifiers in MJD. Thus, we obtained unbiased whole-exome sequencing (WES) data from pairs of first-degree relative patients, who were highly discordant for AO and then compared with AO-concordant pairs from the same family, proposing a robust criterium to filter putative modifier genes. We further investigated modifier effects of the candidate WES variants using genotype-phenotype correlations in four distinct validation cohorts.

## 2. Materials and methods

### 2.1. Patients and samples

Genomic DNA from blood samples of a total of 282 MJD patients was used. All participants were clinically and molecularly confirmed to be MJD patients; sizing of the CAG tract at *ATXN3 locus* (normal allele - $CAG_N$ and expanded allele - $CAG_E$) was performed for all cohorts at a single laboratory, which routinely performs the molecular test for MJD. Age at onset (AO) was defined as the age at manifestation of the first gait disturbances, reported by the patient or a close relative/caregiver.

#### 2.1.1. Discovery sample

WES was performed in samples from 16 Azorean patients: four AO-discordant and four AO-concordant affected first-degree pairs selected from extended Azorean MJD pedigrees. For each discordant pair, a concordant pair was selected from the same extended kindred. Patients of discordant pairs had a mean AO difference of 9 years (range: 7–11 years), whereas concordant pairs showed a mean AO difference of 2 years (range: 1–4 years). The discovery sample is described in Suppl. Table 1. Each affected pair, both concordant and discordant, had an equal (CAG)n size (±1 repeat) in the expanded allele.

#### 2.1.2. Validation cohorts

for genotype-phenotype correlations, several MJD cohorts were used: 103 Azorean patients, including the 16 patients analysed by WES; 72 patients from mainland Portugal; 78 patients from Brazil; and an additional cohort of 29 patients from the UK (of multi-ethnic background). Characterization of the four validation cohorts is shown in Table 1.

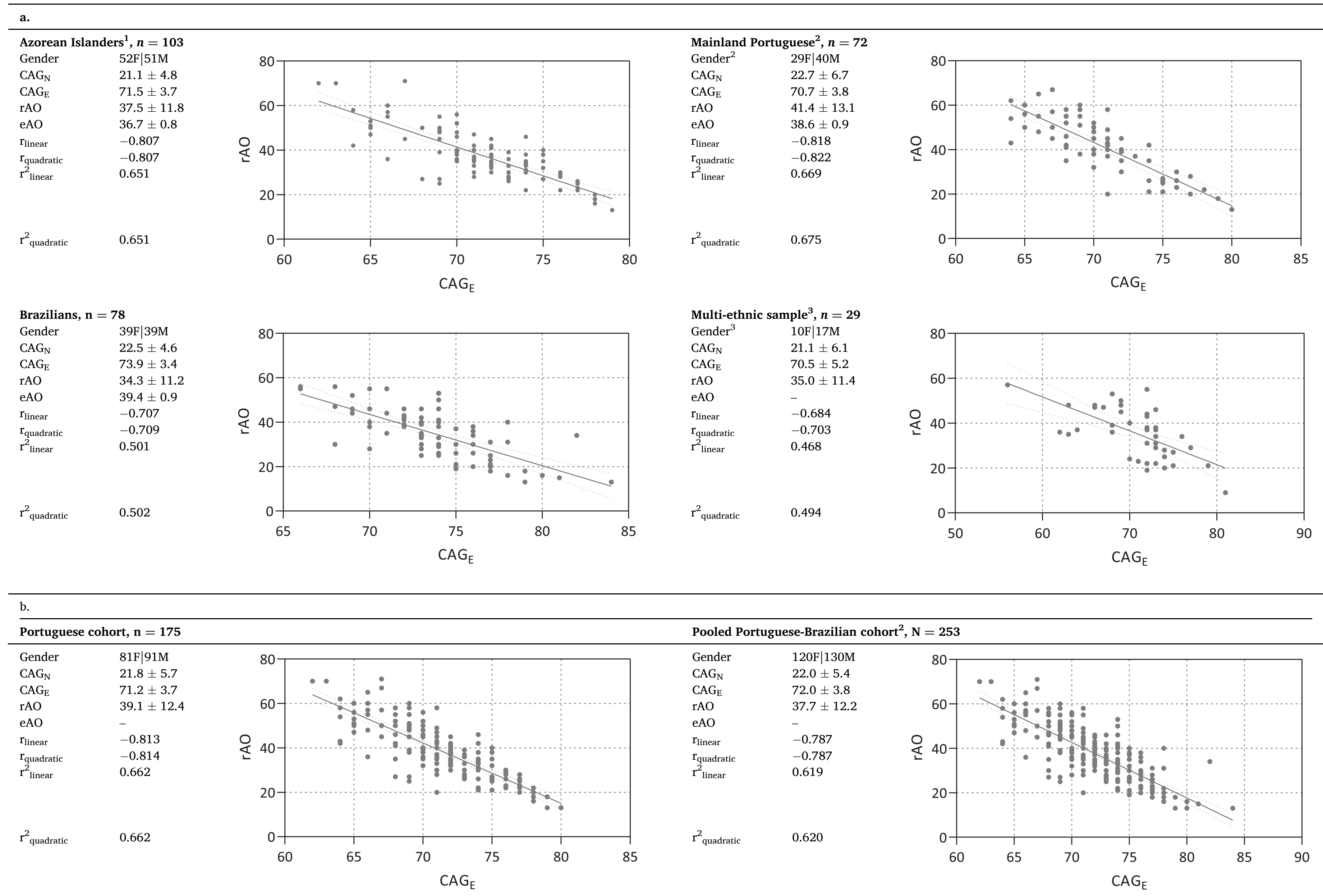
The study design is summarized in Fig. 1.

### 2.2. WES analysis of the discovery sample

WES was performed at Macrogen, Inc. (Suppl. Mat.) for all samples of the discovery sample (Fig. 1A); samples yielded 6.9 to 13.1 GB of high-quality aligned data. Mean target coverage ranged 56.06× to 93.35×, with 91.0%–98.3% being covered at least 10× and 77% being covered at least ≥30× (Suppl. Table 2). A script using R language was used to retrieve, from the total list of variants obtained by WES, those variants whose genotype differed for each discordant pair, but was the same in members of each concordant pair ($n = 233$ variants - list #1, Fig. 1A, Suppl. Table 3). Next, only variants in genes expressed in brain tissues, including cerebellum (expression data were accessed from the GTEx project; 24) were considered for further analysis ($n = 184$ variants – list #2, Fig. 1B, Suppl. Table 3). Exonic variants were sorted by type of mutation, prioritizing nonsense, frameshift, and missense variants. Using information of the Human Splicing Finder (Desmet et al., 2009), intronic and synonymous variants were ranked by potential impact on splicing. The final list (list #3) contained 82 variants, located in 53 genes (Suppl. Table 3 & 4): one frameshift, 18 missense, one splice-acceptor, 43 intronic and 19 synonymous (Fig. 1B), which were successfully

**Table 1**

Demographic, genetic, and clinical features for MJD patients. (a) patients were grouped by geographical origin and (b) pooled Azorean islands and mainland Portuguese patients (=Portuguese cohort), as well as Portuguese (Azorean and mainland Portugal) and Brazilian patients pooled together (=pooled Portuguese-Brazilian cohort).



a.

| **Azorean Islanders[1], $n = 103$** | | **Mainland Portuguese[2], $n = 72$** | |
|---|---|---|---|
| Gender | 52F\|51M | Gender[2] | 29F\|40M |
| $CAG_N$ | 21.1 ± 4.8 | $CAG_N$ | 22.7 ± 6.7 |
| $CAG_E$ | 71.5 ± 3.7 | $CAG_E$ | 70.7 ± 3.8 |
| rAO | 37.5 ± 11.8 | rAO | 41.4 ± 13.1 |
| eAO | 36.7 ± 0.8 | eAO | 38.6 ± 0.9 |
| $r_{linear}$ | −0.807 | $r_{linear}$ | −0.818 |
| $r_{quadratic}$ | −0.807 | $r_{quadratic}$ | −0.822 |
| $r^2_{linear}$ | 0.651 | $r^2_{linear}$ | 0.669 |
| $r^2_{quadratic}$ | 0.651 | $r^2_{quadratic}$ | 0.675 |

| **Brazilians, n = 78** | | **Multi-ethnic sample[3], $n = 29$** | |
|---|---|---|---|
| Gender | 39F\|39M | Gender[3] | 10F\|17M |
| $CAG_N$ | 22.5 ± 4.6 | $CAG_N$ | 21.1 ± 6.1 |
| $CAG_E$ | 73.9 ± 3.4 | $CAG_E$ | 70.5 ± 5.2 |
| rAO | 34.3 ± 11.2 | rAO | 35.0 ± 11.4 |
| eAO | 39.4 ± 0.9 | eAO | – |
| $r_{linear}$ | −0.707 | $r_{linear}$ | −0.684 |
| $r_{quadratic}$ | −0.709 | $r_{quadratic}$ | −0.703 |
| $r^2_{linear}$ | 0.501 | $r^2_{linear}$ | 0.468 |
| $r^2_{quadratic}$ | 0.502 | $r^2_{quadratic}$ | 0.494 |

b.

| **Portuguese cohort, n = 175** | | **Pooled Portuguese-Brazilian cohort[2], N = 253** | |
|---|---|---|---|
| Gender | 81F\|91M | Gender | 120F\|130M |
| $CAG_N$ | 21.8 ± 5.7 | $CAG_N$ | 22.0 ± 5.4 |
| $CAG_E$ | 71.2 ± 3.7 | $CAG_E$ | 72.0 ± 3.8 |
| rAO | 39.1 ± 12.4 | rAO | 37.7 ± 12.2 |
| eAO | – | eAO | – |
| $r_{linear}$ | −0.813 | $r_{linear}$ | −0.787 |
| $r_{quadratic}$ | −0.814 | $r_{quadratic}$ | −0.787 |
| $r^2_{linear}$ | 0.662 | $r^2_{linear}$ | 0.619 |
| $r^2_{quadratic}$ | 0.662 | $r^2_{quadratic}$ | 0.620 |

All continuous variables are shown as mean ± standard deviation (SD), excluding the AOa which is shown as mean ± standard error (SE); [1]includes Azorean patients from discovery cohort; [2] gender for three patients is missing; [3]gender for two patients is missing; $CAG_N$ = number of CAG repeats in the normal allele; $CAG_E$ = number of CAG repeats in the expanded allele; rAO = age at onset (AO) was considered as the age of manifestation of gait disturbances, reported by the patient or a close relative; eAO = adjusted AO calculated using the population of origin as a fixed factor and the covariate appearing in the model are evaluated at $CAG_E = 72$; AO of Brazilian cohort is different from Mainland Portugal cohort and from pooled Portuguese cohort. $CAG_E$ of Brazilian cohort is different from Azorean, mainland Portuguese, multi-ethnic sample and from Portuguese. All tests were performed by an Anova, Tukey HSD $p < 0.05$.

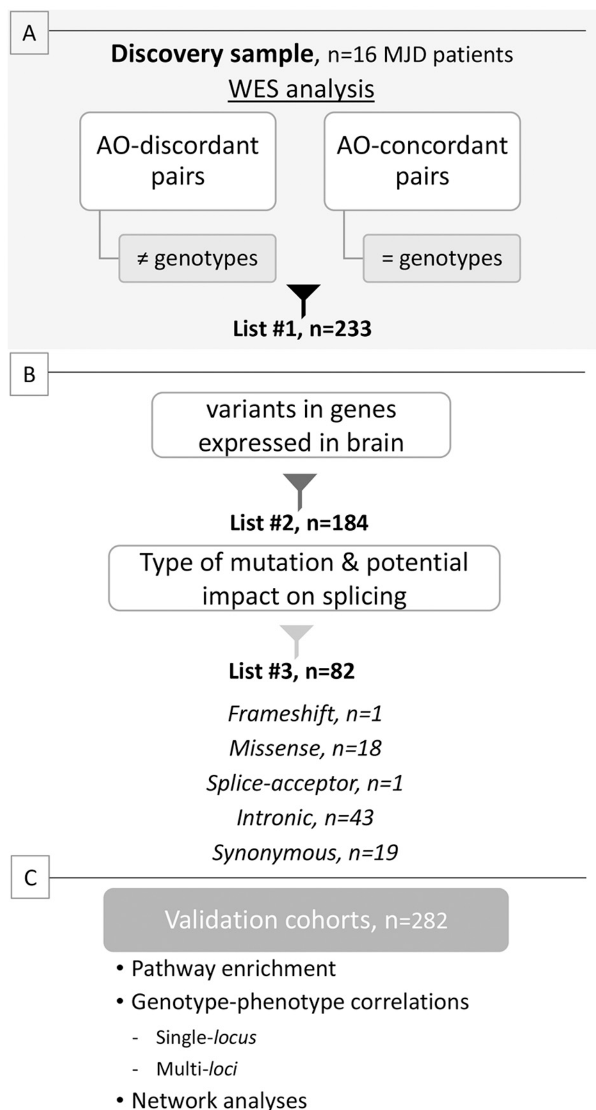



**Fig. 1.** Flow diagram of the study design. (A) Whole-exome sequencing (WES) analyses of AO-discordant and AO-concordant first-degree relative pairs; (B) WES variants were selected using two main criteria (brain expression and potential functional impact) to select the most promising for further analyses; (C) Analysis of modulatory effects of variants/genes. The identification of all variants (List #1, #2 and #3) is available in Suppl. Table 3.

genotyped in the validation cohorts.

### 2.3. *Multiplexed sequencing and data analysis of the discovery sample and validation cohorts*

Candidate variants ($n = 82$) were genotyped (Suppl. Mat.) at Eurofins Genomics (www.eurofinsgenomics.eu), in 282 samples from the validation cohorts (Fig. 1C).

Determination of allelic and genotypic frequencies, conformity with the Hardy-Weinberg equilibrium and linkage disequilibrium were obtained in FSTAT (Goudet, 2002) and GENEPOP (Rousset, 2008). Moreover, allelic and genotypic frequencies in samples from Azores, Mainland Portugal, Brazil and the multi-ethnic cohort were used to perform an exact test of population differentiation (data available upon request).

Independent main effects and gene-gene interactions of allelic variants on MJD AO were tested using parametric (ANCOVA) and nonparametric (Multifactor Dimensionality Reduction) analyses.

An ANCOVA, using the $CAG_E$ as covariate, was performed to calculate and compare adjusted AO (=adjAO) between genotypes. As no differences in the distribution of AO and in $CAG_E$, as well as no differences in genotypic/allelic frequencies of the 82 variants were observed between patients from Azores and mainland Portugal (ANOVA, Tukey HSD $p > 0.05$), these cohorts were pooled together for the analyses (=Portuguese cohort). A pooled Portuguese-Brazilian cohort was also analysed, using size of $CAG_E$ and geographical origin as covariates. To warrant those comparisons were made between genotypic classes with similar number of samples we (1) clustered two genotypic classes, comparing REF|REF versus REF|VAR + VAR|VAR; or (2) excluded the genotypic class with lower number of samples and compared REF|VAR versus VAR|VAR (REF = reference sequence and VAR = variant sequence).

Nonlinear gene-gene interactions were detected by the Multifactor Dimensionality Reduction (MDR) software 3.0.2 (Hahn et al., 2003), using default settings. Cross-validation consistency (CVC), a measure of how often the best model is found across the different tenfold cross-validation interval was calculated; a higher CVC indicates a more consistent result (Hahn et al., 2003). Genotypes for each variant and $CAG_E$ were used as attributes, and AO as the outcome.

Statistical analyses were performed in IBM SPSS Statistics 22 (IBM Corp. Released 2013) and GraphPad Prism 8.0.1. A *P*-value below 0.05 was considered as significant for all tests.

Pathway and interaction-network analyses were generated by IPA (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis; (Krämer et al., 2014)). An interaction network between ATXN3 and candidate modifier genes was explored using the Path Explorer tool, which calculated the "Shortest Path" between 2 molecules. ATXN3 interactors with no links with at least one candidate modifier were removed. All analyses were run under default settings, with exception of the confidence level, which was set to include only experimentally observed interactions. For pathway enrichment analysis, an FDR adjusted P-value <0.05 was considered significant.

## 3. Results

### 3.1. *WES in affected pairs identifies candidate modifier genes of AO in MJD*

We performed WES in a discovery sample of 16 MJD patients, grouped as pairs of AO-concordant and discordant first-degree relatives. An average of 121,105 variants were identified per patient, of which 20% were coding variants (Suppl. Table 2). For 233 variants (list #1, Fig. 1B; Suppl. Table 3), genotypes differed in the AO-discordant pairs and were identical in AO-concordant pairs; 184 variants were prioritized for further investigation (list #2, Fig. 1B, Suppl. Table 3). Characterization of the 82 candidate variants (list #3, Fig. 1B) successfully genotyped is provided in Suppl. Table 4.

Gene enrichment analyses, using the 53 unique genes from list #3 (Suppl. Table 4), resulted in 18 over-represented pathways (Table 2), including two nervous system specific pathways: the neuregulin signaling (4.3%) and the agrin interactions at neuromuscular junction (3.9% of overlapping genes).

### 3.2. *Single-locus analysis in independent cohorts identifies CFAP57, PARD3, CHD5, ACTG1, NFKB1, DLGAP2, ITGB1 and DIDO1 as modifier genes of AO in MJD*

Next, we tested the independent main effects of each of the 82 variants on AO in the Portuguese cohort ($n = 175$). Five variants in five genes showed a significant effect (Fig. 2A): a synonymous variant in *PARD3* (rs11009651), an intronic variant in *NFKB1* (rs4648050), an intronic variant in *CHD5* (rs2273034), a synonymous variant in *ACTG1* (rs1139405) and an intronic variant in *CFAP57* (rs2483688). The

**Table 2**

Eighteen statistically significant canonical pathways were obtained using all genes ($N = 53$) containing variants further analysed in the validation cohorts. The -log (Benjamini-Hochberg *p*-value) greater than 1.3 was considered statistically significant.

| Ingenuity Canonical Pathways | Overlap* | Molecules |
|---|---|---|
| Neuregulin Signaling | 4\|94 (4.3%) | NRG1,ITGB1,MTOR, TGFA |
| ILK Signaling | 5\|199 (2.5%) | ITGB1,DOCK1,MTOR, NFKB1,ACTG1 |
| Rac Signaling (FAK) | 4\|126 (3.2%) | ITGB1,PTK2B,PARD3, NFKB1 |
| Paxillin Signaling | 4\|121 (3.3%) | ITGB1,DOCK1,PTK2B, ACTG1 |
| Calcium-induced T Lymphocyte Apoptosis | 3\|60 (5%) | ITPR3,HLA-DRA, EP300 |
| Signaling by Rho Family GTPases | 5\|256 (2%) | ITGB1,PTK2B,NFKB1, PARD3,ACTG1 |
| TR/RXR Activation | 3\|100 (3%) | MTOR,SREBF2,EP300 |
| Fcγ Receptor-mediated Phagocytosis in Macrophages and Monocytes | 3\|93 (3.2%) | DOCK1,PTK2B,ACTG1 |
| Agrin Interactions at Neuromuscular Junction | 3\|76 (3.9%) | NRG1,ITGB1,ACTG1 |
| Crosstalk between Dendritic Cells and Natural Killer Cells | 3\|89 (3.4%) | HLA-DRA,NFKB1, ACTG1 |
| HER-2 Signaling in Breast Cancer | 3\|99 (3%) | NRG1,ITGB1,PARD3 |
| Tec Kinase Signaling | 4\|176 (2.3%) | ITGB1,PTK2B,NFKB1, ACTG1 |
| B Cell Receptor Signaling | 4\|192 (2.1%) | MTOR,APBB1IP, PTK2B,NFKB1 |
| Death Receptor Signaling | 3\|90 (3.3%) | PARP10,NFKB1, ACTG1 |
| Osteoarthritis Pathway | 4\|206 (1.9%) | ITGB1,MTOR,ANXA2, NFKB1 |
| Calcium Signaling | 4\|198 (2%) | ITPR3,CAMKK1,TPM4, EP300 |
| FAK Signaling | 3\|109 (2.8%) | ITGB1,DOCK1,ACTG1 |
| ErbB Signaling | 3\|108 (2.8%) | NRG1,MTOR,TGFA |

* Number (percentage) of molecules overlapping with the total number of genes belonging to that pathway.

*PARD3* variant was associated with an earlier AO, and the remaining with later AO (Fig. 2A). The *CFAP57* variant explained 4.2% of AO variance, followed by *PARD3* (3.9%), *CHD5* (3%), *ACTG1* (2.6%) and *NFKB1* (2.3%).

In the Brazilian cohort ($n = 78$), two variants in two genes impacted significantly on AO: an intronic variant at *DLGAP2* (rs2293909) and a synonymous variant at *ITGB1* (rs2298141): the *DLGAP2* variant was associated with an earlier onset (explaining 10% of AO variance), whereas the *ITGB1* variant was associated with a later onset (5.3% of variance) (Fig. 2B). In this cohort, *ACTG1* and *CFAP57* maintained the same direction of effect as observed in the Portuguese patients, whereas *PARD3*, *CHD5* and *NFKB1* showed no impact on adjAO. In the Portuguese cohort, *DLGAP2* and *ITGB1* showed an opposite direction of effect, compared to the Brazilian cohort.

To increase statistical power, we pooled the Portuguese and the Brazilian cohorts ($n = 253$); the effects on AO of the two variants at *ACTG1* and *CFAP57* observed in the Portuguese cohort were maintained (Fig. 2C); additionally, two other missense variants were confirmed to modulate AO in the pooled Portuguese-Brazilian cohorts: V1 (rs1883848) and V2 (rs1883847), both at *DIDO1* (Fig. 2C). *DIDO1* variants (V1 and V2), which are not in LD (data not shown) explained an additional 1.6% and 2.5% of AO variance, respectively, beyond geographical origin and size of $CAG_E$.

### *3.3. Multi-locus analysis in independent cohorts confirms NFKB1 and identifies CERS4 as modifiers of AO in MJD*

To detect putative nonlinear interaction effects among all 82 variants, two- and three-*loci* combinations were obtained after performing an exploratory quantitative MDR (qMDR) analysis. In the Portuguese cohort, a two-*loci* interaction model containing $CAG_E$ and *NFKB1* was the best model to predict AO (CVC = 9/10, $p < 0.05$), confirming also the previously described linear independent effect of *NFKB1*. A two-*loci* model containing $CAG_E$ and an intronic variant at *CERS4* (which failed to show an independent effect on AO in previous analyses) was the best model to predict AO (CVC = 8/10, $p < 0.05$) in the Brazilian cohort. In the pooled Portuguese-Brazilian cohorts, the best predictive model of AO was the $CAG_E$**ACTG1* variant combination (CVC = 4/10, $p > 0.05$), although not reaching significance.

### *3.4. The effects of the variants in CHD5, CFAP57, ITGB1 and DIDO1 seem to be maintained in a multi-ethnic cohort*

The variants in *CHD5, CFAP57, ITGB1* and *DIDO1* showed the same independent effect trends in an additional multi-ethnic cohort ($p > 0.05$, Suppl. Fig. 1). None of the two-*loci* combination, namely $CAG_E$**NFKB1* and $CAG_E$**CERS4*, provided a good explanation for AO in this sample. Despite the limited statistical power of this analysis (Suppl. Fig. 2), these results supported an effect of the *CHD5, CFAP57, ITGB1* and *DIDO1* variants on AO across populations of multiple ancestry.

To determine if allelic frequencies for the ten variants in MJD patients were similar to those from control populations (non-MJD), comparisons between our cohort and non-MJD populations were performed; we found that the alternative allele at *PARD3* and *DIDO1-v2* was more frequent in MJD patients then in European or Global populations. Contrarily, the frequency of the alternative alleles at *CERS4* and at *ACTG1* was significantly lower in MJD patients compared to Global and European populations, respectively. Frequencies for the remaining variants were similar between MJD patients and European or Global populations.

### *3.5. Network analyses revealed direct interactions between ATXN3 and NFKB1*

A network analysis was performed by joining ATXN3 and the nine confirmed genetic modifiers provided by genotype-phenotype correlations (eight from the single and one from the multi-*loci* analysis). Several biological relationships were found between ATXN3 and the 9 modifier genes (Fig. 3; Suppl. Table 5), comprising a total of 200 molecules involved in the network. Protein-protein interactions (direct links) were observed between ATXN3 and NFKB1.

We next performed a pathway enrichment analysis, using all molecules from the abovementioned network ($n = 200$) and found the BAG2 (Bcl2-associated athanogene 2) signaling pathway with the strongest enrichment (−log (B-H *p*-value) = 24). Furthermore, protein ubiquitination, Huntington's Disease signaling, and unfolded protein response were also among the more strongly enriched pathways (Suppl. Table 6).

## 4. Discussion

Due to inherent (genetic and non-genetic) confounders, difficulties are acknowledged in the identification of common genetic modifiers between geographically and ethnically distinct cohorts; such difficulties are amplified in rare diseases, such as MJD. In this study, we have used an innovative study design, by applying a combination of exome-wide screening for variants discovery with a candidate gene approach, to identify novel modifiers for MJD. We first performed WES, in a discovery sample of AO-discordant and concordant first-degree relative pairs, selected from the relatively homogeneous and well-documented Azorean cohort (Raposo et al., 2015; Lima et al., 1997), which

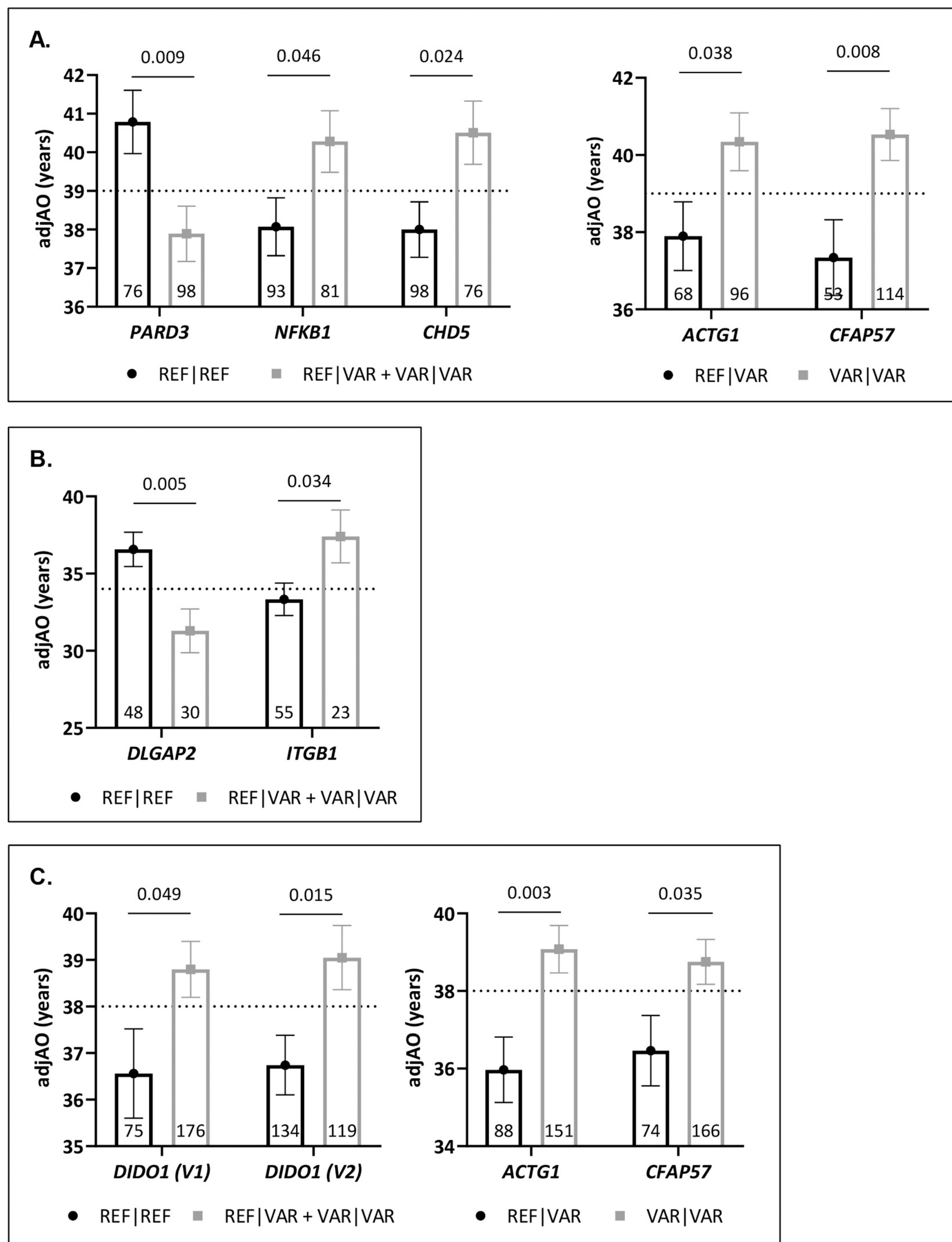



**Fig. 2.** Effect of the candidate variants on adjusted AO (adjAO; mean ± standard error): (A) Portuguese ($n = 175$); (B) Brazilian ($n = 78$); and (C) pooled Portuguese-Brazilian cohorts ($n = 253$). The adjAO corresponds to AO statistically adjusted for mean $CAG_E$ size in each cohort, i.e. (A) 71, (B) 74 and (C) 72. Number of patients by allelic variant/gene are shown inside the bars. Dashed line represents the average AO of each respective cohort. Only variants showing statistically significant differences are displayed ($p < 0.05$).

allowed us to identify variants with potential effects among the high number of variants arising from WES. This discovery step was then complemented by testing the impact on AO of the most promising variants in genotype-phenotype correlations using larger and independent cohorts of MJD patients.

We showed that variants at *PARD3, NFKB1, CHD5, ACTG1, CFAP57, DLGAP2, ITGB1, DIDO1* and *CERS4* modifies AO in MJD, with those in *CFAP57, ACTG1* and *DIDO1* having consistent effects across patients from different geographical origins.

From the 18 statistically enriched pathways found, the neuregulin signaling and agrin interactions at neuromuscular junction were of particular interest, due to their specific function in the nervous system. Neuregulin signaling plays diverse roles in the CNS and peripheral nervous system, namely in synapse formation, radial neuron migration, GABAergic neuron migration, neural crest cells differentiation and migration, neuromuscular junction and Schwann cell development, maturation, and myelination (Kataria et al., 2019). The neuregulin pathway has been linked to processes relevant for neurodegeneration, including oxidative stress and neuroinflammation (Xu et al., 2017; Jiang et al., 2016). The neuromuscular junction is a chemical synapse that is created between motor neurons and skeletal muscles and is covered by Schwann cells (Li et al., 2018). The agrin-LRP4-MuSK signaling drives

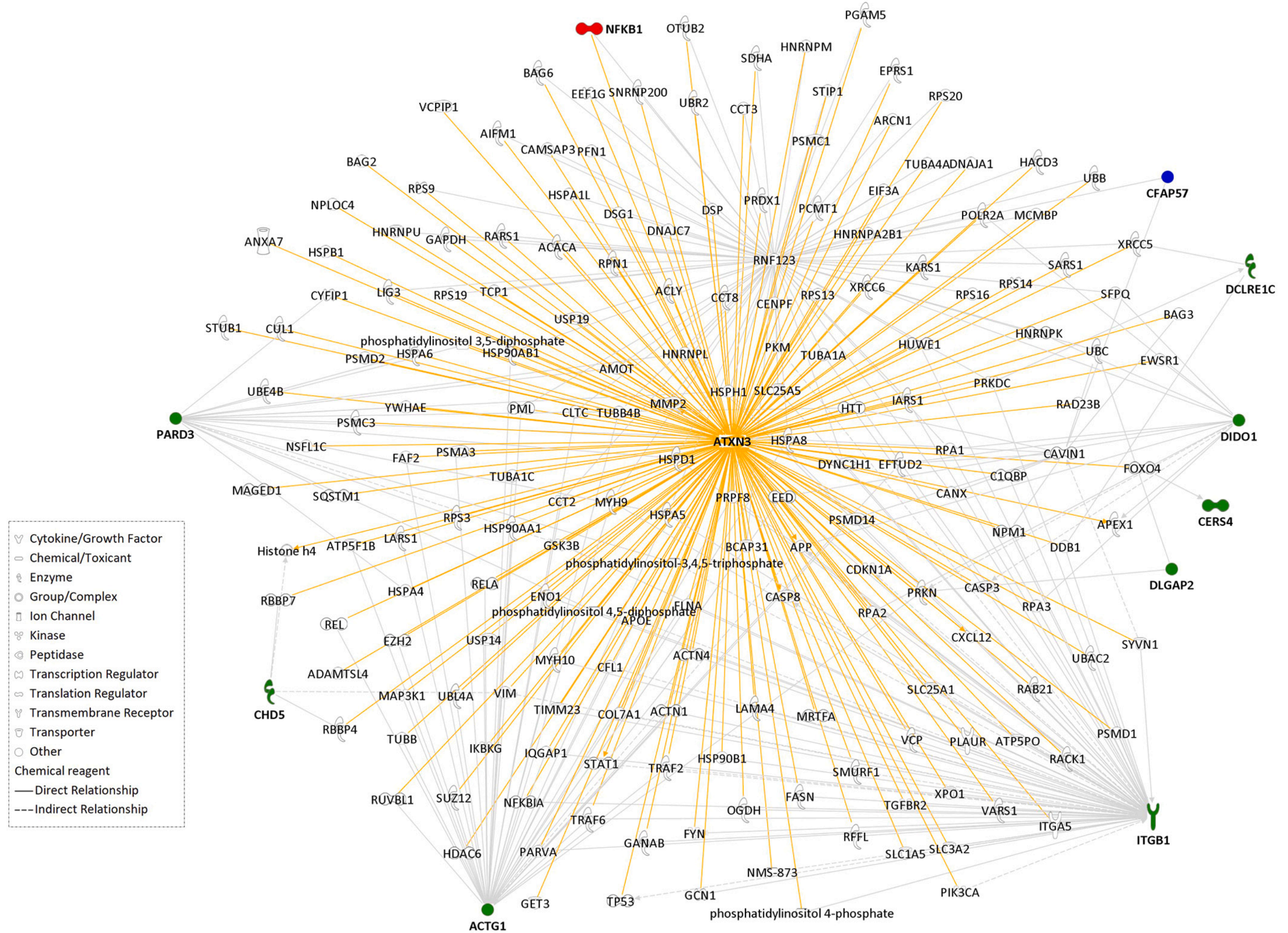



**Fig. 3.** Interaction network showing the relationship between ATXN3 and the candidate modifier genes validated by genotype-phenotype correlations. Network analysis was performed in the IPA software, using the Path Explorer tool. The relationship between the modifier genes and ATXN3 were highlighted, concerning (1) the absence of intermediate molecules (direct connection) – *red,* (2) the presence of a single intermediate molecule – *green* and (3) two intermediate molecules – *blue*. Direct interactors of ATXN3 were highlighted in *orange*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the aggregation of acetylcholine receptors and guarantees efficient signal transduction at the neuromuscular junction (Ohno et al., 2017). This ultimately leads to the initiation of muscle contraction, making the neuromuscular junction essential for mobility. Dysfunction of this signal transduction has been reported in other neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS) and spinal muscular atrophy (SMA) (Li et al., 2018; Ohno et al., 2017). Therefore, we postulate that it might be relevant for the amyotrophy seen in a significant proportion of MJD patients as well as in animal models of the disease (Sara Duarte-Silva, personal communication).

The network analysis considering ATXN3 and the nine genes provided by genotype-phenotype correlations generated a network of 200 genes; pathway-enrichment analysis, using these genes as input dataset, revealed the highest enrichment of the Bcl-2 associated athanogene (BAG) signaling pathway and the protein ubiquitination pathway. The emergence of the protein ubiquitination pathway is consistent with the physiological role of ataxin-3 as deubiquitinating enzyme (Costa, and do C, Paulson HL., 2012). The BAG protein family has been associated with neurodegenerative diseases, including MJD (reviewed in Che et al. (2015); Qin et al. (2016)). Playing a key role in protein degradation, BAG proteins, namely BAG2 and BAG5, modify mutated ataxin-3 toxicity by regulating its stability (Che et al., 2015).

Genotype-phenotype correlations showed that variants in *ACTG1, CFAP57* and *DIDO1* have consistent effects in delaying AO in MJD. *ACTG1* encodes for the actin cytoplasmic 2 protein, an actin monomer (G-actin), which composes the actin cytoskeleton after polymerization (reviewed in Eira et al. (2016)). Alterations in the actin regulatory machinery have been suggested in a *Drosophila* MJD model (Lee et al., 2011; Vu et al., 2018) and actin cytoskeleton disorganization was seen upon ATXN3 silencing or polyQ expansion in cultured cells (Rodrigues et al., 2010; Neves-Carvalho et al., 2015). Data from an RNA-seq experiment using six samples from cerebellum of MJD patients and controls showed that mRNA levels of *ACTG1* were decreased in MJD patients compared to controls (q-value$<$0.05; 43). In healthy population, *ACTG1* levels have been described as highly expressed in cerebellum as well as in other brain areas, not showing significant tissue-specific differences (GTEx Consortium, 2013).

Our network analysis further revealed a direct interaction between ACTG1 and APOE (an ATXN3 interactor), a modifier gene previously described by us and subsequently replicated in a Chinese cohort (Bettencourt et al., 2011; Peng et al., 2014). Another link between ACTG1 and ATXN3 is its direct interaction with PRKN (encoding parkin). Parkin is a E3 ubiquitin ligase, commonly associated with inherited and sporadic Parkinson's disease (reviewed in Dawson and Dawson (2010)).

Parkin is deubiquitinated by ataxin-3 (Durcan et al., 2011); mutant (but not wild-type) ataxin-3 promotes clearance of parkin via autophagy, suggesting that increased turnover of parkin may contribute to the MJD pathogenesis (Durcan et al., 2011; Durcan and Fon, 2013). The synonymous variant in *ACTG1* implied an alteration of the auxiliary splicing sequences; according to in silico predictions its presence should significantly alter the ESE/ESS (exonic splicing enhancer or silencer) motifs ratio (data not shown). The role of splicing defects in human disease is rapidly emerging nowadays; gain or loss of enhancers and silencer sequences can lead to the alternative inclusion of exons, and this type of splicing defect has been associated with neurodegenerative diseases, namely with frontotemporal dementia and Parkinsonism linked to chromosome 17 or spinal muscular atrophy (Daguenet et al., 2015).

*CFAP57* variant also implied a consistent delaying of disease onset; this gene encodes for cilia- and flagella-associated protein 57, a member of the WD40 family of proteins, frequently acting as protein interaction scaffolds to form functional complexes through their WDR domain, in essential cellular processes (Zou et al., 2016). In healthy population, *CFAP57* seems to be similarly expressed in vulnerable and non-vulnerable MJD brain areas (GTEx Consortium, 2013). Moreover, levels of this gene were similar in the cerebellum of MJD patients and controls (Haas et al., n.d.). mRNA levels of *ATG16L2*, a gene which encodes also a protein presenting a WD40 domain, were described to be reduced in blood cells of MJD subjects, including preclinical carriers, when compared to age- and sex-matched controls (Ana F. Ferreira, personal communication). Intriguing, the loss of the WDR domain in LRRK2, another protein of the WD40 family, was block neurotoxicity of multiple LRRK2 Parkinson's disease causative variants (Schapira et al., 2017). Although alterations on splicing signals could be predicted for the intronic variant in *CFAP57,* this should not lead to a significant impact on splicing (data not shown).

Variants in *DIDO1* were also associated to a delaying effect on MJD onset. *DIDO1* encodes for death-inducer obliterator 1 (DIO-1) protein, triggering apoptosis through up-regulation of procaspase 3 and 9 (García-Domingo et al., 2003). In healthy population, *DIDO1* is slightly more expressed in cerebellum compared to other vulnerable (brainstem) and non-vulnerable MJD brain areas (GTEx Consortium, 2013). Interesting, a tendency for an up-regulation of *DIDO1* levels in the cerebellum of MJD patients compared to controls was observed ($p$-value = 0.018 and q-value<0.15; 43). Of note, mutant ataxin-3 has been associated with mitochondrial apoptotic pathway, through activation of caspase-3 and caspase-9 (Chou et al., 2006). DIDO1/CASP3 and ATXN3/CASP3 relationships were also identified in our network analysis. This also highlighted a protein-protein interaction between *DIDO1* and APP (amyloid precursor protein); *APP* mRNA level was shown to be upregulated in CSM14.1 rat cell lines stably expressing human full-length expanded ataxin-3 (Evert et al., 2001). Moreover, several Aβ-immunoreactive deposits were observed in human MJD pons compared to controls (Evert et al., 2001). The impact of the two *DIDO1* missense variants (V1 and V2) on protein structure/function was predicted to be non-pathogenic (benign/tolerant). However, by splicing predictions, the alternative allele of *DIDO1*-V2 implied the activation of a cryptic donor site, and therefore, potentially leading to alterations of splicing (data not shown).

A consistent role of *PARD3, NFKB1, CHD5, DLGAP2, ITGB1* and *CERS4* as genetic modifiers in the different MJD cohorts could not be observed in this study; the statistical power, however, should not be pointed out as the reason to explain this lack of consistency of the variant's effects on AO. For example, in the Brazilian cohort, the *DLGAP2* variant was associated with an earlier onset, explaining 10% of AO variance; a genotype-phenotype correlation study with ~50 patients (25 per group) would be able to detect an effect size of 0.5 in AO (supplementary Fig. 2). This means that the modifier effect of *DLGAP2* on AO should have been found at least in the Portuguese cohort ($N = 175$).

Although previous evidence about the role of genetic modifiers in MJD was grounded on single gene effects, contribution of a single modifier to the phenotype is usually minor, and can be masked by environmental or population effects, as well as by the individual genetic background. Complex interactions between the primary disease-causing gene and a (yet) undetermined number of other modifiers playing a role in a common cell mechanism, as well as a variable combination of multiple gene variants among patients may determine the phenotypic outcome of MJD. As consequence, the effect on AO of variants in *PARD3, CHD5, DLGAP2, ITGB1, NFKB1* and *CERS4* (the last two in interaction with $CAG_E$), whose effects were not found across the several cohorts, may be population-specific but not necessarily without value. Difference in frequency of the *ATXN3* expansion size, in combination with variable allelic frequencies of a variety of modifiers, might be the most important factor to explain non-replication of modifying effects across cohorts. Greene and colleagues showed that alterations in allele frequency at one interacting *locus* may decrease power to detect a main effect at another *locus*: changes in minor allele frequency (MAF) of less than 0.1 may imply a loss of power from 80% to 20% to replicate the main effect (Greene et al., 2009). In this perspective, patient stratification by $CAG_E$ in different cohorts has consistently been reported in MJD (Raposo et al., 2015; de Mattos et al., 2019; Tezenas du Montcel et al., 2014), also observed in this study; Brazilian patients showed larger expansions more frequently than Portuguese patients, or patients from the multi-ethnic cohort (Suppl. Table 8). Of note, the identification of epistatic effects is challenging in MJD, as it is in other diseases caused by dynamic mutations. The large diversity of $CAG_E$ alleles in interaction with multiple modifiers results in a small number of patients in each group to allow further statistical comparisons, which has implications to reach the necessary power, in a rare disease such as MJD.

## 5. Conclusions

In summary, we propose novel MJD-modifying genes and pathways for further investigation as new disease-modifying targets. Unveiling mechanisms of novel MJD-modifying pathways may be a novel attractive avenue to explore pharmacologically than testing single modifier genes, a clear limitation of candidate gene studies. Characterization of the complex contribution of these genes to the MJD phenotype should help improving more individualized disease management and prognosis.

## Ethics approval and consent to participate

The study was approved by local ethics committees; all subjects provided written informed consent.

## Consent for publication

Not applicable.

## Availability of data and materials

Most data are available in this manuscript and in supplementary material. Raw whole exome sequencing data and R script that support the findings of this study will be made available, upon request, to the corresponding author.

## Funding

This work was funded by FEDER - Fundo Europeu de Desenvolvimento Regional funds through the COMPETE 2020 - Operacional Programme for Competitiveness and Internationalisation (POCI), Portugal 2020, and by Portuguese funds through FCT - Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior in the framework of the project PTDC/DTP-PIC/2638/2017 (POCI-01-0145-FEDER-016592); GenomePT (POCI-01-0145-FEDER-022184); ICVS Scientific Microscopy Platform, member of the national infrastructure PPBI - Portuguese Platform of Bioimaging (PPBI-POCI-01-0145-FEDER-

022122; by National funds, through the Foundation for Science and Technology (FCT) - project UIDB/50026/2020 and UIDP/50026/2020; and by the project NORTE-01-0145-FEDER-000013, supported by Norte Portugal Regional Operational Programme (NORTE 2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF). MR is supported by FCT (CEECIND/03018/2018). ARVM (SFRH/BD/129547/2017) and AFF (SFRH/BD/121101/2016) are supported by a PhD grant financed by FCT. CB is supported by the Multiple System Atrophy Trust and Alzheimer's Research UK. MDC received funding from National Ataxia Foundation (NAF) and from FCT (SFRH/BPD/101925/2014); DV-C received a grant from FCT (SFRH/BD/147826/2019).

## Authors contributions

MR and ML conceived, organized and executed the study. MR, CB, ARVM, AFF, IA, JV, TK, MLS-P, BFB, JB-A, HH, LBJ, JS and ML participated in clinical data and blood collection. DV-C, MC and PM performed *C. elegans* experiments. MR, PS, DV-C and MC performed the bioinformatic and/or statistical analyses. MR, CB, DV-C, MC, ARVM, AFF, PH, PM and ML analysed and/or interpreted data. MR, CB, PM, and ML drafted the first version of the manuscript. DV-C, ARVM, AFF, MC, IA, PS, MLS-P, PH, LBJ and JS revised the manuscript. All authors read and approved the final manuscript.

## Declaration of Competing Interest

The authors declare no competing interests.

## Acknowledgements

We gratefully thank Professor Luís Silva (University of the Azores) for the R script development.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nbd.2021.105578.

## References

Akçimen, F., Martins, S., Liao, C., Bourassa, C.V., Catoire, H., Nicholson, G.A., et al., 2020 Mar 23. Genome-wide association study identifies genetic factors that modify age at onset in Machado-Joseph disease. Aging (Albany NY) [Internet]. 12 (6), 4742–4756. Available from: http://www.aging-us.com/article/102825/text.

Bettencourt, C., Lima, M., 2011. Machado-Joseph disease: from first descriptions to new perspectives. Orphanet. J. Rare. Dis. [Internet]. 6, 35. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3123549&tool=pmcentrez&rendertype=abstract.

Bettencourt, C., Raposo, M., Kazachkova, N., Cymbron, T., Santos, C., Kay, T., et al., 2011. The APOE ε 2 allele increases the risk of earlier age at onset in Machado-Joseph disease. Arch. Neurol. 68 (12).

Bettencourt, C., Hensman-Moss, D., Flower, M., Wiethoff, S., Brice, A., Goizet, C., et al., 2016 Jun. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. Ann. Neurol. [Internet]. 79 (6), 983–990. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27044000.

Carvalho, D.R., La Rocque-Ferreira, A., Rizzo, I.M., Imamura, E.U., Speck-Martins, C.E., 2008 Apr. Homozygosity enhances severity in spinocerebellar ataxia type 3. Pediatr. Neurol. [Internet]. 38 (4), 296–299. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0887899407005929.

Che, X.-Q., Tang, B.-S., Wang, H.-F., Yan, X.-X., Jiang, H., Shen, L., et al., 2015 May 4. The BAG2 and BAG5 proteins inhibit the ubiquitination of pathogenic ataxin3-80Q. Int. J. Neurosci. [Internet]. 125 (5), 390–394. Available from: http://www.tandfonline.com/doi/full/10.3109/00207454.2014.940585.

Chen, Z., Sequeiros, J., Tang, B., Jiang, H., 2018 Dec. Genetic modifiers of age-at-onset in polyglutamine diseases. Age. Res. Rev. [Internet]. 48, 99–108. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1568163718300606.

Chou, A.H., Yeh, T.H., Kuo, Y.L., Kao, Y.C., Jou, M.J., Hsu, C.Y., et al., 2006. Polyglutamine-expanded ataxin-3 activates mitochondrial apoptotic pathway by upregulating Bax and downregulating Bcl-xL. Neurobiol. Dis. 21 (2), 333–345.

Costa, M., do C, Paulson HL., 2012 May. Toward understanding Machado-Joseph disease. Prog. Neurobiol. [Internet]. 97 (2), 239–257 [cited 2013 Jan 28]. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22133674.

Da Silva, J.D., Teixeira-Castro, A., Maciel, P., 2019 Oct 5. From pathogenesis to novel therapeutics for spinocerebellar ataxia type 3: evading potholes on the way to translation. Neurotherapeutics [Internet]. 16 (4), 1009–1031. Available from: http://link.springer.com/10.1007/s13311-019-00798-1.

Daguenet, E., Dujardin, G., Valcárcel, J., 2015 Dec 13. The pathogenicity of splicing defects: mechanistic insights into pre- <scp>mRNA</scp> processing inform novel therapeutic approaches. EMBO Rep. [Internet]. 16 (12), 1640–1655. Available from: https://onlinelibrary.wiley.com/doi/10.15252/embr.201541116.

Dawson, T.M., Dawson, V.L., 2010. The role of parkin in familial and sporadic Parkinson's disease. Mov. Disord. [Internet]. 25 (S1), S32–S39. Available from: http://doi.wiley.com/10.1002/mds.22798.

de Araújo, M., Raposo, M., Kazachkova, N., Vasconcelos, J., Kay, T., Lima, M., 2016. Trends in the epidemiology of spinocerebellar ataxia type 3/Machado-Joseph disease in the Azores Islands, Portugal. JSM Brain Sci. 1 (1), 1001.

de Mattos, E.P., Leotti, V.B., Soong, B.-W., Raposo, M., Lima, M., Vasconcelos, J., et al., 2019. Age at onset prediction in spinocerebellar ataxia type 3 changes according to population of origin. Eur. J. Neurol. 26 (1).

Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Béroud, G., Claustres, M., Béroud, C., 2009 May. Human splicing finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. [Internet]. 37 (9), e67. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkp215.

Durcan, T.M., Fon, E.A., 2013. Ataxin-3 and Its E3 Partners: implications for Machado–Joseph Disease. Front. Neurol. [Internet]. 4. Available from: http://journal.frontiersin.org/article/10.3389/fneur.2013.00046/abstract.

Durcan, T.M., Kontogiannea, M., Thorarinsdottir, T., Fallon, L., Williams, A.J., Djarmati, A., et al., 2011 Jan 1. The Machado–Joseph disease-associated mutant form of ataxin-3 regulates parkin ubiquitination and stability. Hum. Mol. Genet. [Internet]. 20 (1), 141–154. Available from: https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddq452.

Eira, J., Silva, C.S., Sousa, M.M., Liz, M.A., 2016 Jun. The cytoskeleton as a novel therapeutic target for old neurodegenerative disorders. Prog. Neurobiol. [Internet]. 141, 61–82. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0301008215300800.

Evert, B., Vogt, I., Kindermann, C., Ozimek, L., de Vos, R., Brunt, E., et al., 2001 Aug 1. Inflammatory genes are upregulated in expanded ataxin-3-expressing cell lines and spinocerebellar ataxia type 3 brains. J. Neurosci. [Internet]. 21 (15), 5389–5396. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11466410.

García-Domingo, D., Ramírez, D., González de Buitrago, G., Martínez-A, C., 2003 May. Death inducer-obliterator 1 triggers apoptosis after nuclear translocation and caspase upregulation. Mol. Cell. Biol. [Internet]. 23 (9), 3216–3225. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12697821.

Gatchel, J.R., Zoghbi, H.Y., 2005 Oct. Diseases of unstable repeat expansion: mechanisms and common principles. Nat. Rev. Genet. [Internet]. 6 (10), 743–755. Available from: http://www.nature.com/articles/nrg1691.

Génin, E., Feingold, J., Clerget-Darpoux, F., 2008. Identifying modifier genes of monogenic disease: strategies and difficulties. Hum. Genet. 124, 357–368.

Goudet, J., 2002. FSTAT [Internet]. Available from: https://www2.unil.ch/popgen/softwares/fstat.htm.

Greene, C.S., Penrod, N.M., Williams, S.M., Moore, J.H., 2009 Jun 2. Failure to replicate a genetic association may provide important clues about genetic architecture. Sorensen TIA, editor. PLoS One [Internet]. 4 (6), e5639. Available from: https://dx.plos.org/10.1371/journal.pone.0005639.

GTEx Consortium, 2013 Jun. The genotype-tissue expression (GTEx) project. Nat. Genet. [Internet]. 45 (6), 580–585. Available from: http://www.ncbi.nlm.nih.gov/pubmed/23715323.

Gusella, J.F., MacDonald, M.E., Lee, J.-M., 2014 Sep 15. Genetic modifiers of Huntington's disease. Mov. Disord. [Internet]. 29 (11), 1359–1365. Available from: http://doi.wiley.com/10.1002/mds.26001.

Haas E, Incebacak RD, Hentrich T, Huridou C, Schmidt T, Casadei N, Maringer Y, Bahl C, Zimmermann F, Mills JD, Aronica E, Riess O, Schulze-Hentrich JM H-SJ. A novel SCA3 Knock-in mouse model mimics the human SCA3 disease phenotype including neuropathological, behavioral, and transcriptional abnormalities especially in oligodendrocytes. Mol. Neurobiol.

Hahn, L.W., Ritchie, M.D., Moore, J.H., 2003 Feb 12. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. Bioinformatics [Internet]. 19 (3), 376–382. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btf869.

Jardim, L., Silveira, I., Pereira, M.L., do Céu Moreira, M., Mendonça, P., Sequeiros, J., et al., 2003. Searching for modulating effects of SCA2, SCA6 and DRPLA CAG tracts on the Machado-Joseph disease (SCA3) phenotype. Acta Neurol. Scand. 107 (3), 211–214.

Jiang, Q., Chen, S., Hu, C., Huang, P., Shen, H., Zhao, W., 2016 Sep. Neuregulin-1 (Nrg1) signaling has a preventive role and is altered in the frontal cortex under the pathological conditions of Alzheimer's disease. Mol. Med. Rep. [Internet]. 14 (3), 2614–2624. Available from: https://www.spandidos-publications.com/10.3892/mmr.2016.5542.

Kataria, H., Alizadeh, A., Karimi-Abdolrezaee, S., 2019 Sep. Neuregulin-1/ErbB network: an emerging modulator of nervous system injury and repair. Prog. Neurobiol. [Internet]. 180, 101643. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0301008219300905.

Krämer, A., Green, J., Pollard, J., Tugendreich, S., 2014 Feb 15. Causal analysis approaches in ingenuity pathway analysis. Bioinformatics [Internet]. 30 (4), 523–530. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24336805.

Lee, S.B., Bagley, J.A., Lee, H.Y., Jan, L.Y., Jan, Y.-N., 2011 Oct 4. Pathogenic polyglutamine proteins cause dendrite defects associated with specific actin cytoskeletal alterations in Drosophila. Proc. Natl. Acad. Sci. U.S.A. [Internet]. 108

(40), 16795–16800 [cited 2012 Jul 18]. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3189048&tool=pmcentrez&rendertype=abstract.
Li, L., Xiong, W.-C., Mei, L., 2018 Feb 10. Neuromuscular junction formation, aging, and disorders. Annu. Rev. Physiol. [Internet]. 80 (1), 159–188. Available from: http://www.annualreviews.org/doi/10.1146/annurev-physiol-022516-034255.
Lima, M., Mayer, F., Coutinho, P., Abade, A., 1997 Jun. Prevalence, geographic distribution, and genealogical investigation of Machado-Joseph disease in the Azores (Portugal). Hum. Biol. [Internet]. 69 (3), 383–391. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9164048.
Lima, M., Mayer, F.M., Coutinho, P., Abade, A., 1998 Dec. Origins of a mutation: population genetics of Machado-Joseph disease in the Azores (Portugal). Hum. Biol. [Internet]. 70 (6), 1011–1023. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9825593.
Mergener, R., Furtado, G.V., de Mattos, E.P., Leotti, V.B., Jardim, L.B., Saraiva-Pereira, M.L., 2020 Mar 5. Variation in DNA repair system gene as an additional modifier of age at onset in spinocerebellar ataxia type 3/Machado–Joseph disease. Neuro. Mol. Med. [Internet]. 22 (1), 133–138. Available from: http://link.springer.com/10.1007/s12017-019-08572-4.
Montiel, R., Bettencourt, C., Silva, C., Santos, C., Prata, M.J., Lima, M., 2005. Analysis of Y-chromosome variability and its comparison with mtDNA variability reveals different demographic histories between islands in the Azores archipelago (Portugal). Ann. Hum. Genet. 69 (2), 135–144.
Neves-Carvalho, A., Logarinho, E., Freitas, A., Duarte-Silva, S., do Costa, M.C., Silva-Fernandes, A., et al., 2015 Jan 1. Dominant negative effect of polyglutamine expansion perturbs normal function of ataxin-3 in neuronal cells. Hum. Mol. Genet. [Internet]. 24 (1), 100–117. Available from: https://academic.oup.com/hmg/article-lookup/doi/10.1093/hmg/ddu422.
Ohno, K., Ohkawara, B., Ito, M., 2017 Oct 3. Agrin-LRP4-MuSK signaling as a therapeutic target for myasthenia gravis and other neuromuscular disorders. Expert Opin. Ther. Targets [Internet]. 21 (10), 949–958. Available from: https://www.tandfonline.com/doi/full/10.1080/14728222.2017.1369960.
Peng, H., Wang, C., Chen, Z., Sun, Z., Jiao, B., Li, K., et al., 2014. APOE ε2 allele may decrease the age at onset in patients with spinocerebellar ataxia type 3 or Machado-Joseph disease from the Chinese Han population. Neurobiol. Aging [Internet]. 35 (9), 2179.e15–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24746364.
Petersen, B.-S., Fredrich, B., Hoeppner, M.P., Ellinghaus, D., Franke, A., 2017 Dec 14. Opportunities and challenges of whole-genome and -exome sequencing. BMC Genet. [Internet]. 18 (1), 14. Available from: http://bmcgenet.biomedcentral.com/articles/10.1186/s12863-017-0479-5.
Qin, L., Guo, J., Zheng, Q., Zhang, H., 2016 Dec 20. BAG2 structure, function and involvement in disease. Cell. Mol. Biol. Lett. [Internet]. 21 (1), 18. Available from: http://cmbl.biomedcentral.com/articles/10.1186/s11658-016-0020-2.
Rahit, K.M.T.H., Tarailo-Graovac, M., 2020 Feb 25. Genetic modifiers and rare mendelian disease. Genes (Basel) [Internet]. 11 (3), 239. Available from: https://www.mdpi.com/2073-4425/11/3/239.
Raposo, M., Ramos, A., Bettencourt, C., Lima, M., 2015. Replicating studies of genetic modifiers in spinocerebellar ataxia type 3: Can homogeneous cohorts aid? Vol. 138. Brain p. e398.
Rodrigues, A.-J., do Carmo Costa, M., Silva, T.-L., Ferreira, D., Bajanca, F., Logarinho, E., et al., 2010 Oct. Absence of ataxin-3 leads to cytoskeletal disorganization and increased cell death. Biochim. Biophys. Acta. - Mol. Cell. Res. [Internet]. 1803 (10), 1154–1163. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0167488910001989.
Rousset, F., 2008. Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol. Ecol. Resour. [Internet]. 8 (1), 103–106. Available from: http://doi.wiley.com/10.1111/j.1471-8286.2007.01931.x.
Santos, C., Lima, M., Montiel, R., Angles, N., Pires, L., Abade, A., et al., 2003. Genetic structure and origin of peopling in the Azores Islands (Portugal): the view from mtDNA. Ann. Hum. Genet. 67 (5), 433–456.
Schapira, M., Tyers, M., Torrent, M., Arrowsmith, C.H., 2017 Nov 13. WD40 repeat domain proteins: a novel target class? Nat. Rev. Drug Discov. [Internet]. 16 (11), 773–786. Available from: http://www.nature.com/articles/nrd.2017.179.
Tezenas du Montcel, S., Durr, A., Bauer, P., Figueroa, K., Ichikawa, Y., Brussino, A., et al., 2014. Modulation of the age at onset in spinocerebellar ataxia by CAG tracts in various genes. Brain [Internet]. 137 (Pt 9), 2444–2455. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24972706.
Vu, A., Humphries, T., Vogel, S., Haberman, A., 2018. Polyglutamine repeat proteins disrupt actin structure in Drosophila photoreceptors. Mol. Cell Neurosci. [Internet]. 93, 10–17. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1044743118301714.
Xu, J., Hu, C., Chen, S., Shen, H., Jiang, Q., Huang, P., et al., 2017 Sep. Neuregulin-1 protects mouse cerebellum against oxidative stress and neuroinflammation. Brain Res. [Internet]. 1670, 32–43. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0006899317302603.
Zou, X.-D., Hu, X.-J., Ma, J., Li, T., Ye, Z.-Q., Wu, Y.-D., 2016 Dec 19. Genome-wide Analysis of WD40 Protein Family in Human. Sci. Rep. [Internet] 6 (1), 39262. Available from: http://www.nature.com/articles/srep39262.