Gene 566 (2015) 158–165

Contents lists available at ScienceDirect

Gene

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

Research paper

Identification of rare protein disulfide isomerase gene variants in amyotrophic lateral sclerosis patients

Paloma Gonzalez-Perez [a,1], Ute Woehlbier [b,c,d,1], Ru-Ju Chian [a], Peter Sapp [a], Guy A. Rouleau [e], Claire S. Leblond [e], Hussein Daoud [e], Patrick A. Dion [e], John E. Landers [a], Claudio Hetz [b,c,f,*], Robert H. Brown [a,**]

[a] Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA
[b] Biomedical Neuroscience Institute, Faculty of Medicine, University of Chile, Santiago, Chile
[c] Institute of Biomedical Sciences, Center for Molecular Studies of the Cell, Program of Cellular and Molecular Biology, University of Chile, Santiago, Chile
[d] Center for Genomics and Bioinformatics, Faculty of Science, Universidad Mayor, Santiago, Chile
[e] Montreal Neurological Institute and Hospital, Department of Neurology and Neurosurgery, McGill University, 3801 rue University, Montreal, QC H3A 2B4, Canada
[f] Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, USA

ARTICLE INFO

Article history:
Received 20 January 2015
Received in revised form 12 April 2015
Accepted 13 April 2015
Available online 22 April 2015

Keywords:
ER stress
Protein disulfide isomerase
PDIA1
ERp57
Amyotrophic lateral sclerosis

ABSTRACT

Disruption of endoplasmic reticulum (ER) proteostasis is a salient feature of amyotrophic lateral sclerosis (ALS). Upregulation of ER foldases of the protein disulfide isomerase (PDI) family has been reported in ALS mouse models and spinal cord tissue and body fluids derived from sporadic ALS cases. Although in vitro studies suggest a neuroprotective role of PDIs in ALS, the possible contribution of genetic mutations of these ER foldases in the disease process remains unknown. Interestingly, intronic variants of the *PDIA1* gene were recently reported as a risk factor for ALS. Here, we initially screened for mutations in two major PDI genes (PDIA1/P4HB and PDIA3/ERp57) in a US cohort of 96 familial and 96 sporadic ALS patients using direct DNA sequencing. Then, 463 familial and 445 sporadic ALS patients from two independent cohorts were also screened for mutations in these two genes using whole exome sequencing. A total of nine PDIA1 missense variants and seven PDIA3 missense variants were identified in 16 ALS patients. We have identified several novel and rare single nucleotide polymorphisms (SNPs) in both genes that are enriched in ALS cases compared with a large group of control subjects showing a frequency of around 1% in ALS cases. The possible biological and structural impact of these ALS-linked PDI variants is also discussed.

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a progressive adult-onset neurodegenerative disease affecting motoneurons in the brain and spinal cord leading to paralysis and death (Pasinelli and Brown, 2006). While most ALS cases are sporadic (sALS), approximately 10% are familial (fALS), caused by rare variants in multiple genes (Leblond et al., 2014). To date, ALS has been associated with 17 genes of high risk and 22 genes of low risk of developing the disease (Leblond et al., 2014). The most common genetic causes of fALS are the recently defined hexanucleotide repeat expansion in the intronic region of *C9orf72* and mutations in the gene encoding cytosolic superoxide dismutase 1 (*SOD1*), which together account for around 50% of fALS cases (Leblond et al., 2014). Mutations in TAR DNA binding protein (*TARDBP*, also known as TDP43) and fused-in-sarcoma/translated-in-liposarcoma (*FUS/TLS*) genes each represent about 5% of fALS cases (Ferraiuolo et al., 2011). Although the mechanisms underlying ALS pathogenesis remain speculative, accumulating evidence indicates that disturbance of proteostasis (Balch et al., 2008) is a common feature of sALS and fALS (Saxena and Caroni, 2011). The accumulation of misfolded proteins is a shared characteristic of many neurodegenerative diseases (Soto, 2012), and is extensively reported in sALS cases and most animal models of the disease (Soto, 2012; Turner et al., 2013). These protein species may result from oxidative damage, endoplasmic reticulum (ER) stress, disturbed calcium homeostasis and/or a global failure due to overload in protein quality control and clearance mechanisms (Hetz and Mollereau, 2014).

The ER is a major compartment involved in protein folding and quality control in the secretory pathway (Walter and Ron, 2011). ER stress has been extensively reported in ALS patients and transgenic animal models of the disease (Matus et al., 2013; Atkin et al., 2014)

Abbreviations: PDI, protein disulfide isomerase; PDIA1/P4HB, protein disulfide isomerase A1/prolyl 4-hydroxylase subunit beta; PDIA3/ERp57, protein disulfide isomerase A3/endoplasmic reticulum resident protein 57; TARDBP, TAR DNA-binding protein 43; SOD1, superoxide dismutase 1; FUS/TLS, fused in sarcoma/translocated in sarcoma; VCP, valosin-containing protein; C9orf72, C9orf72; CNX, calnexin; CRT, calreticulin.

* Correspondence to: C. Hetz, Institute of Biomedical Sciences, University of Chile, Independencia 1027, Santiago, Chile.
** Corresponding author.
E-mail addresses: chetz@hsph.harvard.edu (C. Hetz), robert.brown@umassmed.edu (R.H. Brown).
URL:E-mail addresses:E-mail address: http://www.hetzlab.cl (C. Hetz).
[1] Equal contribution.

http://dx.doi.org/10.1016/j.gene.2015.04.035
0378-1119/

Table 1
ALS-associated *PDI* mutations. Four missense genetic variants in *PDIA1* and *PDIA3* (2 each) were first identified by direct sequencing in the US ALS cohort. Two of these variants were found in patients with familial ALS (fALS) and two in patients with sporadic ALS (sALS). SNP Genotyping Analysis of these variants using TaqMan in >1000 USA controls, as well as, their frequency in 1 K Genome Project and NHLBI Exome variant Server (NHLBI EVS) pointed to a significant association of these variants with ALS.

Gene	Country	Exon	Genomic bp	cDNA	aa change	Seq (UMMS) FALS	Seq (UMMS) SALS	TaqMan Controls	dbSNP	1K Genome Freq	NHLBI EVS EA (8595)	NHLBI EVS AA (4404)	NHLBI EVS Total (12999)	NHLBI EVS Freq
P4HB/PDIA1[a]														
Chr 17	USA	7	79,804,487	c.874G>A	D292N	1/96	0/96	0/1068	rs145209834	0	5	2	7	0.000539
	USA	7	79,804,462	c.899G>A	R300H	0/96	1/96	1/1070	0	0	0	0	0	0
ERp57/PDIA3[b]														
Chr 15	USA	6	44,057,694	c.649G>A	D217N	0/96	1/96	0/1061	0	0	0	0	0	0
	USA	13	44,063,369	c.1441C>A	Q481K	1/96	0/96	0/1074	0	0	0	0	0	0

[a]chr 17: 79,801,034–79,818,544.
[b]chr 15: 44,038,590–44,064,804.

and is one of the earliest events detected in the asymptomatic phase of the disease (Saxena et al., 2009). Genetic and pharmacological manipulation of the unfolded protein response (UPR), an adaptive reaction to cope with ER stress, has been shown to have functional consequences on the progression of experimental ALS reviewed in Hetz et al. (2013). The upregulation of ER chaperones and protein disulfide isomerases (PDIs) has been widely reported in ALS (Andreu et al., 2012). Proteomic analysis of spinal cord tissue of mutant SOD1 transgenic mice revealed as major changes the induction of two PDI family members known as PDIA1 (also referred to as PDI or P4HB) (Atkin et al., 2006) and ERp57 (also referred to as PDIA3 or Grp58) (Atkin et al., 2008), a finding that was confirmed in spinal cord tissue (Ilieva et al., 2007; Hetz et al., 2009) and cerebrospinal fluid (Atkin et al., 2008) of sALS patients. Remarkably, proteomic screening for biomarkers in blood from patients also identified the upregulation of PDIA1 and ERp57 protein levels as one of the best indicators for diagnosis and to monitor disease progression (Nardo et al., 2011). PDIA1 was found to co-localize with protein inclusions containing SOD1, TDP-43, and FUS in tissue from ALS patients, in addition to cellular and mouse models of the disease, possibly indicating a physical association between them (Atkin et al., 2006, 2008; Farg et al., 2012; Honjo et al., 2011). Furthermore, S-nitrosylation of PDIA1 has been observed in sALS, which has a negative effect on neuronal viability (Walker et al., 2010). Finally, intronic variants of *PDIA1* were recently proposed as genetic risk factors for ALS (Kwok et al., 2013). However, the possible contribution of these mutations to ALS pathogenesis has not been directly addressed.

Based on the importance of ER proteostasis disturbances in ALS, we used a candidate gene approach and exome sequencing to screen for possible mutations in the coding region of PDIA1 and ERp57. With this strategy we identified 16 novel missense variants in these two genes. The possible consequences of these substitutions to PDI function and the development of ALS are discussed.

2. Methods

2.1. Direct DNA sequencing

DNA was isolated from venous blood of ALS patients according to standard protocols. The USA cohort included DNA samples from 96 fALS and 96 sALS patients. An additional set of DNA samples in this cohort included >1000 controls subjects that were used for high-throughput SNP genotyping (TaqMan assay). Patients were diagnosed with possible, probable, or definite ALS as per El Escorial criteria (Brooks et al., 2000). No DNA from other family members was available for this study. Whole genome amplification was performed using the Illustra Genomiphi V2 DNA Amplification kit (GE HealthCare cat. No. 25-6600-31). All exons and exon–intron junctions of *PDIA1* and *PDIA3* genes were amplified by PCR with primers designed using Primer 3.0. AmpliTaq Gold PCR Master Mix 2500U (Applied Biosystems cat. No. 4327059) was used to carry out a touchdown PCR in a 30 μl reaction volume. The reaction mixture was incubated at 95 °C for 5 min initially, followed by 30 cycles at 95 °C for 30 s, 65 °C for 30 s; with a −0.5 °C decrement of temperature per cycle, and 72 °C for 1 min. 15 cycles at 95 °C for 30 s, 65 °C for 30 s and 72 °C for 1 min, and a final extension time of 7 min at 72 °C were added. The PCR products were cleaned-up using Exonuclease I 20,000 U (NEB M0293L), S.A. Phosphatase 5000 U (Fisher E70092X) and sequenced bidirectionally by a fluorescently-labeled dideoxy-nucleotide chain termination method. SNPs were confirmed using purified DNA from the patients. High-throughput SNP genotyping was performed using TaqMan assay for each confirmed novel variant in a larger set of unrelated ALS patients and control subjects. The online tools Polyphen-2 and SIFT were used to predict the impact of the amino acid substitutions on the structure and function of PDIA1 and ERp57.

2.2. Exome sequencing

Canadian ALS and control cases were recruited at the following institutes, the "Centre de Recherche du Centre Hospitalier de l'Université de Montréal" (Montreal Qc, Canada) and the "Montreal Neurological Institute and Hospital" (Montreal Qc, Canada). Patients were diagnosed with possible, probable, or definite ALS as per El Escorial criteria (Brooks et al., 2000). Canadian ALS and control cases were studied by whole Exome sequencing using Agilent SureSelectXT Human All Exon V4 for the exome capture, and the Illumina HiSeq 2000 platform from the "McGill University and Génome Québec Innovation Centre" for the high-throughput sequencing. A total of 168 sALS and 100 fALS cases were analyzed. Variants identified in PDI genes were validated by Sanger sequencing using BatchPrimer3 v1.0 for the primer design, AmpliTaq Gold DNA Polymerase (Invitrogen) for the PCR amplification, and Sanger sequencing platform from the "McGill University and Génome Québec Innovation Centre". Additionally, six known SNPs (seen in dbSNP hg19) were identified in our ALS set but their genotype and major allele frequencies were not significantly different from the reported frequency in the general population according to dbSNP (hg19) (data not shown).

Exome sequencing for non-Canadian ALS patients (USA, UK, Italy, Ireland, Spain, Netherlands), here designated as INT for international, was performed as previously described (Smith et al., 2014) at the Keck DNA Sequencing Facility (New Haven, CT). Reads were subsequently aligned to human reference (GRCh37) using BWA (BurrowsWheeler Aligner). PICARD was used to remove duplicate reads and the Genome