Brain Research 1647 (2016) 43–49

Contents lists available at ScienceDirect

Brain Research

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

Review

C9orf72 isoforms in Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration

Shangxi Xiao [a], Laura MacNair [a,b], Jesse McLean [a], Phillip McGoldrick [a], Paul McKeever [a,b], Serena Soleimani [a], Julia Keith [c], Lorne Zinman [c], Ekaterina Rogaeva [a], Janice Robertson [a,b,*]

[a] Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Canada
[b] Department of Laboratory Medicine and Pathobiology, University of Toronto, Canada
[c] Sunnybrook Health Sciences Centre, Canada

ARTICLE INFO

Article history:
Received 22 February 2016
Received in revised form
23 April 2016
Accepted 26 April 2016
Available online 29 April 2016

Keywords:
C9orf72
Hexanucleotide repeat expansion
C9-transcripts
C9-isoforms
ALS/FTLD

ABSTRACT

A hexanucleotide (G_4C_2) repeat expansion in the 5′ non-coding region *C9orf72* is the most common genetic cause of amyotrophic lateral sclerosis and frontotemporal lobar degeneration. Three modes of toxicity have been proposed: gain of function through formation of RNA foci and sequestration of RNA binding proteins; expression of dipeptide repeat proteins generated by repeat-associated non-ATG translation; and loss of function due to C9orf72 haploinsufficiency. Much is known about the proposed gain of function mechanisms, but there is little knowledge of the normal function of C9orf72 and the cellular consequences if its activity is perturbed. Here we will review what is known of C9orf72 at the transcript and protein levels and how changes in C9orf72 expression could contribute to disease pathogenesis.

This article is part of a Special Issue entitled SI:RNA Metabolism in Disease.

Contents

1. Introduction

Amyotrophic lateral sclerosis (ALS) is a devastating adult onset neurodegenerative disease primarily affecting motor neurons of the brain and spinal cord, causing a rapidly progressive paralysis with death typically ensuing 2–5 years from diagnosis. It has been recognized that up to 50% of those affected with ALS demonstrate signs of cognitive and/or behavioral impairment with up to 25% fulfilling the diagnostic criteria of frontotemporal dementia (FTD) (Lomen-Hoerth et al., 2002; Ringholz et al., 2005; Strong et al., 1999; Strong, 2008; Swinnen and Robberecht, 2014). This has led to the development of consensus criteria to diagnose frontotemporal cognitive and behavioral changes as part of the

Abbreviations: HRE, hexanucleotide repeat expansions; DPR, dipeptide repeat protein; RBP, RNA Binding Protein

* Corresponding author at: Tanz CRND, University of Toronto, Krembil Discovery Tower, 60 Leonard Avenue, ON, Canada M5T 2S8.
E-mail address: jan.robertson@utoronto.ca (J. Robertson).

http://dx.doi.org/10.1016/j.brainres.2016.04.062
0006-8993/

diagnostic workup in ALS (Strong et al., 2009). FTD is an umbrella term for a group of disorders affecting the frontal and temporal lobes: semantic variant and non-fluent variant primary progressive aphasia; behavioral variant FTD; corticobasal syndrome and progressive supranuclear palsy (Bang et al., 2015; Irwin et al., 2015). The neuropathological term for these disorders is frontotemporal lobar degeneration (FTLD). The hallmark neuropathology linking ALS and FTLD is the presence of cytoplasmic inclusions of TAR DNA binding protein-43 (TDP-43) in disease-affected neurons and glial cells (Arai et al., 2006; Neumann et al., 2006). TDP-43 is a predominantly nuclear DNA/RNA binding protein involved in many aspects of RNA metabolism (Buratti and Baralle, 2010) and its mislocalization to the cytoplasm of disease affected neurons is a pathological hallmark of ~97% of ALS cases and ~45% of FTLD cases (reviewed in (Ling et al., 2013). The clinical and neuropathological overlap of ALS and FTLD was substantiated by the groundbreaking discovery of mutations comprising of large hexanucleotide $(G_4C_2)_n$ repeat expansions (HRE) in the 5′ non-coding region of *C9orf72* as the most common genetic cause of ALS and FTLD, accounting for ~40% of familial ALS cases, ~25% of familial FTLD cases and remarkably ~7% of sporadic ALS cases, with mutation frequencies of up to 88% in familial ALS/FTLD cases (DeJesus-Hernandez et al., 2011; Gijselinck et al., 2012; Renton et al., 2011, 2014; van der Zee and Van Broeckhoven, 2014). ALS and FTLD are now considered the two extremes of a disease spectrum (DeJesus-Hernandez et al., 2011; Ng et al., 2015; Renton et al., 2011). The precise number of $(G_4C_2)_n$ expansions requisite to cause disease is uncertain and may be influenced by epigenetic factors (Belzil et al., 2013; Belzil et al., 2014; Xi et al., 2013; Xi et al., 2014; Xi et al., 2015b). Nevertheless, common nomenclature in the literature is that >30 repeats are pathogenic, although there are exceptions (Gami et al., 2015; Gijselinck et al., 2015; Xi et al., 2015a).

Research into understanding how HREs cause ALS/FTLD has been aided by the extensive studies already undertaken for other nucleotide repeat diseases, such as Huntington's disease, spinocerebellar ataxias and myotonic dystrophy (reviewed in (Cleary and Ranum, 2014; van Blitterswijk et al., 2012). Three major hypotheses have been promulgated: HRE pre-mRNAs sequester RNA binding proteins (RBPs) forming RNA foci (DeJesus-Hernandez et al., 2011; Renton et al., 2011); HREs undergo repeat-associated non-ATG initiated (RAN) translation generating dipeptide repeat proteins (DPRs) in both sense and antisense directions (Ash et al., 2013; Mori et al., 2013a; Mori et al., 2013b; Zu et al., 2013); finally, a loss of function mechanism due to C9orf72 haploinsufficiency (DeJesus-Hernandez et al., 2011; Mizielinska and Isaacs, 2014; Renton et al., 2011). There is evidence for all three mechanisms in ALS/FTLD, with RNA foci, DPRs and reductions in *C9orf72* transcripts found in disease tissues and patient derived cells (DeJesus-Hernandez et al., 2011; Mizielinska and Isaacs, 2014; Renton et al., 2011). Here we will focus on C9orf72 protein, what little is known, and speculate on how changes in C9orf72 expression could contribute to disease pathogenesis.

2. C9ORF72 transcripts

C9orf72 is highly conserved among metazoans except for *Drosophila melanogaster*. Comparisons of the coding regions of human *C9orf72* orthologues show high sequence similarity both at the nucleotide (nt) and amino acid levels (aa) with mouse showing 91.89% nt and 98.75% aa similarity; rat 89.4% nt, 97.71% aa; chicken 83.3% nt, 93.56% aa; and zebrafish: 68.55% nt, 76.14% aa. This high degree of homology suggests that C9orf72 may fulfill fundamental biological functions. There are three *C9orf72* (C9) transcripts in humans: variant 1 (exons 2-5) encoding a 222 aa isoform called C9-short (C9-S); and variants 2 and 3 (exons 2-11) encoding a 481 aa isoform called C9-long (C9-L) (DeJesus-Hernandez et al., 2011; Renton et al., 2011). The C-terminal sequence of C9-S has an additional lysine residue, encoded by the partially retained nucleotide sequence of intron 5 (Xiao et al., 2015). The HREs are located in the promoter region of variant 2 and the first intron of variants 1 and 3. There have been numerous reports that *C9orf72* transcript levels are reduced by up to 50% either in total or in a transcript specific manner in selected brain regions of C9-linked ALS/FTLD (Belzil et al., 2013; DeJesus-Hernandez et al., 2011; Fratta et al., 2013; Gijselinck et al., 2012; Mizielinska and Isaacs, 2014; Renton et al., 2011). A reduction in *C9orf72* transcripts, in particular variants 1 and 2 was recently confirmed in a large-scale study using frontal cortex and cerebellum from 56 HRE carriers (van Blitterswijk et al., 2015). This study also showed that truncated/aborted transcripts were transcribed in C9-HRE carriers, the presence of which had been reported previously (Haeusler et al., 2014), although this remains controversial (Tran et al., 2015). Moreover higher levels of variant 1 were associated with longer patient survival (van Blitterswijk et al., 2015), suggesting that loss of C9orf72 may contribute to disease pathogenesis. A loss of function pathomechanism has also been supported by the generation of knock-out or deletion of *C9orf72* orthologues in zebrafish and *C. elegans* which resulted in motor dysfunction (Ciura et al., 2013; Therrien et al., 2013). However, reduction of *C9orf72* in mice, either by genetic or treatment with antisense oligonucleotides, did not induce motor neuron deficits, ALS-related pathology, or changes in survival (Koppers et al., 2015; Lagier-Tourenne et al., 2013). More recent studies have shown a role for C9orf72 in immune regulation, with knockout mice developing progressive splenomegaly and lymphadenopathy (Atanasio et al., 2016; O′ Rourke et al., 2016). Interestingly, loss of C9orf72 led to an accumulation of lysosomes in microglia and associated neuroinflammation, changes that were also observed in spinal cord of C9orf72-ALS cases, suggesting that C9orf72 deficiency could contribute to defects of innate immunity (O'Rourke et al., 2016). Collectively, the contribution of C9orf72 protein or lack thereof to the pathogenic mechanism in ALS/FTLD remains uncertain, and is hindered by the lack of a full understanding of C9orf72 function.

3. C9orf72 isoforms and domain structure

Sequence and structural analyses have identified C9orf72 as a DENN (differentially expressed in normal and neoplastic cells) protein, comprised of an N-terminal longin domain, followed by DENN and C-terminal alpha domains (Zhang et al., 2012). DENN domain proteins are GTP-GDP exchange factors (GEFs) for Rab GTPases, which are master regulators of membrane trafficking (Marat et al., 2011; Mizuno-Yamasaki et al., 2012). C9orf72 has been shown to interact with several Rabs and regulate endosomal trafficking in cell culture models, suggesting that haploinsufficiency of C9orf72 could lead to reduced Rab activation causing changes in membrane trafficking and contribute to the pathogenic mechanism causing ALS/FTLD (Farg et al., 2014). Indeed this has been supported by a recent study showing an enrichment of C9orf72 at synapses (Atkinson et al., 2015).

Like human, the mouse orthologue of *C9orf72* (3110043O21Rik) has three transcript variants: V1 (exons 2-11) encoding a 481 aa isoform 1 (Ms-C9-1); V2 (exons 3-11) encoding a 420 aa isoform 2 (Ms-C9-2); and V3 (exons 2-10) encoding a 317 aa isoform 3 (Ms-C9-3) (Atkinson et al., 2015). Mouse isoforms 1 and 3 share the same overall domain structure as C9-L, except Ms-C9-3 lacks 60 aa from the C-terminus (Fig. 1). In comparison, the human isoform C9-S encompasses only the longin domain, whereas Ms-C9-2 contains the DENN-Alpha domains (Fig. 1). Interestingly, sequence