Published in final edited form as:

Trends Neurosci. 2018 May ; 41(5): 247–250. doi:10.1016/j.tins.2018.03.010.

All in the family: repeats and ALS/FTD

Amrutha Pattamatta^{1,2}, John D. Cleary^{1,2}, and Laura P.W. Ranum^{1,2,3,4}

¹Center for Neurogenetics, University of Florida, Gainesville, Florida

²Departments of Molecular Genetics and Neurology, University of Florida, Gainesville, Florida

³McKnight Brain Institute, University of Florida, Gainesville, Florida

⁴Genetics Institute, College of Medicine, University of Florida, Gainesville, Florida

Abstract

In 2011, an intronic (G4C2)•(G2C4) expansion was shown to cause the most common form of amyotrophic lateral sclerosis and frontotemporal dementia. This discovery linked ALS with a clinically distinct form of dementia and a larger group of microsatellite repeat diseases, and catalyzed basic and translational research.

Keywords

C9orf72; amyotrophic lateral sclerosis; ALS; frontotemporal dementia; FTD; GGGGCC; CCGGGG; RNA foci; repeat associated non-ATG translation; RAN translation

Text

Discoveries of novel human disease genes generate scientific interest because they open doors into understanding the causes and mechanisms of disease. It is rare, however, for a gene discovery to captivate scientists the way that the 2011 discovery of the *C9orf72* expansion mutation did. In back-to-back papers published in *Neuron*, two independent groups led by Drs. Radamakers and Traynor reported the discovery of an intronic GGGGCC (or G4C2) expansion mutation as the most common genetic cause of both amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). This discovery captured the attention of scientists all over the world because it identified a common molecular cause for two clinically distinct diseases—ALS and FTD – and because it connected ALS and dementia to a large group of previously described microsatellite repeat expansion disorders.

Historically, ALS and FTD were thought to be distinct disorders. ALS is a fatal neuromuscular disorder that leads to the degeneration of upper and lower motor neurons, resulting in paralysis and eventually death. In contrast, FTD is a neurodegenerative disease

*Correspondence: ranum@ufl.edu (Laura P.W. Ranum).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

affecting primarily the frontal and anterior temporal lobes, which is characterized by behavioral changes, apathy and dementia during the later stages of diseases. A series of separate gene mutations identified in the 1990s and 2000s as causing ALS (e.g. superoxide dismutase *SOD1*) or FTD (e.g. progranulin, *GRN*) was consistent with the distinct clinical features of these disorders ^{1,2}. Although ALS and FTD are clinically distinct, the high frequency of their comorbidity in some families suggested a common underlying genetic mutation(s), at least in some cases. The search for the underlying gene mutation was narrowed in 2006 with the discovery of a genetic locus for familial ALS/FTD on chromosome 9p21 ³. Eventually, in 2011, two groups independently used deep-sequencing of large numbers of independent families to discover a hexanucleotide expansion in the first intron of *C9orf72* gene as the leading cause of familial ALS and FTD ^{4,5}. This surprising discovery created a rich scientific delta by bringing together scientists from the three separate fields of dementia, ALS and microsatellite expansion disorders.

Two decades before the discovery of the *C9orf72* expansion mutation, the microsatellite expansion field was born with the demonstration that CGG and CAG expansion mutations cause Fragile-X syndrome and spinal bulbar muscular atrophy (SBMA), respectively ⁶. These discoveries, and the demonstration that expansion mutations can change in length when transmitted from one generation to the next, provided a molecular explanation for ‘anticipation’, i.e., the earlier onset and more severe disease seen in consecutive generations as observed in many of these disorders. These discoveries also led to an intense hunt for expansion mutations in other neurologic diseases, such as Huntington’s disease, myotonic dystrophy and multiple spinocerebellar ataxias. There are now more than 40 known diseases caused by the expansion of repeats present in the 5’ UTRs, exons, introns or 3’ UTRs of their respective genes. Typically, the molecular mechanisms of these diseases have been classified as protein loss of function, RNA gain of function or protein gain of function.

Early in 2006, linkage analysis in a large Dutch family with autosomal dominant inheritance of both ALS and FTD implicated a 11 Mb region on chromosome 9p13.2-21.3 harboring approximately 103 known genes as encompassing the disease mutation ³. Bryan Traynor’s group at the NIH, used data from genome wide association studies (GWAS) on similar families, and identified a robust unannotated single nucleotide polymorphism (SNP) cluster over the center of chromosome 9p, narrowing the suspect region to a ~200,000bp interval containing three genes: *MOBK2B*, *IFNK* and *C9orf72* (Renton et al., 2011). Traditional sequencing of this region identified no genetic anomaly alluding to the possibility of a repeat expansion or inversion of genes leading to disease. Using meticulous deep sequencing approaches across the ~200,000bp region, Renton et al., observed that there were a large number of SNPs on exon 1a, and that there was a drop in sequence coverage over intron 1a. Manual realignment of these sequences to the reference genome indicated the presence of a hexanucleotide repeat expansion located in the first intron of two isoforms of the *C9orf72* gene and in the promoter region of a third isoform. This mutation was shown to be causative of 46% of familial ALS, 21.1% of sporadic ALS and 29.3% of FTD in a Finnish population analyzed by the authors ⁵. Rosa Rademakers and colleagues, at the Mayo Clinic, also used sequencing to identify the hexanucleotide expansion in a region between exon 1a and 1b on *C9orf72* gene. Using fluorescent fragment length analysis, they identified an apparent abnormal segregation pattern wherein affected patients appeared homozygous and affected

children did not inherit an allele from the affected parent. Suspecting a repeat expansion as the causative mutation, the authors performed repeat primed PCR and Southern blotting and found that all affected individuals carried an expanded hexanucleotide repeat. In the family cohorts analyzed by DeJesus-Hernandez et al., GGGGCC repeats in *C9orf72* were found to cause 12% of familial FTD and 22.5% of familial ALS⁴. These studies were important not only for the discovery of the most common known cause of both familial and sporadic ALS and FTD but also for bringing scientists together, first as a consortium of geneticists to find the gene and second for the scientific fuel it provided to the international research community to study the molecular mechanisms of this disease.

Taking cues from the microsatellite expansion field, both DeJesus-Hernandez et al., and Renton et al., suggested *C9orf72* loss of function as a potential disease mechanism in their initial reports. DeJesus-Hernandez et al. also provided evidence that RNA gain of function effects – similar to what had been established in myotonic dystrophy – could be a possible mechanism for *C9orf72* ALS/FTD. In support of this hypothesis they showed the accumulation of GGGGCC containing nuclear foci in the frontal cortex and spinal cord from affected patients.

By the time of the discovery of the C9-ALS/FTD hexanucleotide repeat, bidirectional transcription and the production of both sense and antisense transcripts, was fast becoming a hallmark of expansion repeat disorders⁷. Additionally, Zu et al., discovered that CAG and CUG expansion mutations can make proteins in all three reading frames without an ATG initiation codon and demonstrated that these repeat-associated non-ATG (RAN) proteins accumulate in both SCA8 and DM1 human and mouse tissues⁸. Taken together, these data demonstrated that a single CAG•CTG expansion mutation can produce eight potentially toxic mutant products: two expansion RNAs and up to six toxic homopolymeric RAN proteins^{8,9}. In 2013, several groups extended these observations to *C9orf72* and demonstrated the accumulation of both sense (polyGA, polyGP, polyGR) and antisense (polyPA, polyPR and polyGP) dipeptide (DPRs) or RAN proteins^{10–12}. Subsequent research has demonstrated that RAN proteins disrupt a number of cellular functions including: (i) proteasome and autophagy function; (ii) liquid-liquid phase separation; (iii) ribosome biogenesis; and (iv) nucleocytoplasmic transport¹. Figure 1 summarizes some of the disease mechanisms that have been described in C9-ALS/FTD.

While there have been groundbreaking advances in understanding the role of the *C9orf72* hexanucleotide expansion mutation, a number of questions remain unanswered. What is the role of RNA vs. RAN protein toxicity in disease? What causes the variable disease presentation found in C9 patients? Why are certain cell types, for instance motor neurons, particularly susceptible to disease? In parallel to efforts to address these important basic research questions, translational and clinical studies have already begun developing and testing a number of therapeutic strategies. Notable amongst these approaches are the iPSC and mouse model screens of antisense oligonucleotides that target the sense repeat containing transcripts, and are nearing clinical trials^{13,14}. Additionally, a separate BAC transgenic mouse model develops key behavioral, neurodegenerative and molecular features of both ALS and FTD and can be used to understand the molecular mechanisms and progression of disease and test therapeutic strategies for their efficacy in preventing motor

neuron and other neurodegenerative changes, paralysis and survival¹⁵. It is remarkable how rapidly the *C9orf72* field has moved from gene discovery and model development to the cusp of therapeutic trials—all in less than 7 years.

Unlike typical genetic mutations, repeat expansions are difficult to detect by both traditional and next-generation sequencing methods. This difficulty is highlighted by the fact that the discovery of a repeat expansion for the most common genetic forms of both ALS and FTD came more than 20 years after the first descriptions of expansion mutations. While advances in sequencing technologies have fueled a lot of the recent progress in gene discovery and genetic diseases, these methodologies are inadequate for the detection of repeat expansion mutations and emphasize the need to develop new strategies to identify repeat expansion mutations. In light of this recent discovery of *C9orf72* expansion mutation being the most common genetic cause of ALS and FTD, an important question arises - how many other unidentified expansion mutations are responsible for common neurodegenerative disorders?

Acknowledgments

The authors thank the National Institutes of Health (RO1 NS098819, RO1 NS040389 and PO1 NS058901), Target ALS, CHDI, the National Ataxia Foundation, the ALS Association, the Packard Center, the Myotonic Dystrophy Foundation, The Marigold Foundation and the Muscular Dystrophy Association for funding.

References

1. Taylor JP Jr, RHB, Cleveland DW. Decoding ALS: from genes to mechanism. *Nature*. 2016; 539:197. [PubMed: 27830784]
2. Lashley T, Rohrer JD, Mead S, Revesz T. Review: An update on clinical, genetic and pathological aspects of frontotemporal lobar degenerations. *Neuropathol Appl Neurobiol*. 2015; 41:858–881. [PubMed: 26041104]
3. Vance C, et al. Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. *Brain J Neurol*. 2006; 129:868–876.
4. DeJesus-Hernandez M, et al. Expanded GGGGCC hexanucleotide repeat in non-coding region of *C9ORF72* causes chromosome 9p-linked frontotemporal dementia and amyotrophic lateral sclerosis. *Neuron*. 2011; 72:245–256. [PubMed: 21944778]
5. Renton AE, et al. A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*. 2011; 72:257–268. [PubMed: 21944779]
6. Pearson CE, Nichol Edamura K, Cleary JD. Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet*. 2005; 6:729–742. [PubMed: 16205713]
7. Batra R, Charizanis K, Swanson MS. Partners in crime: bidirectional transcription in unstable microsatellite disease. *Hum Mol Genet*. 2010; 19:R77–82. [PubMed: 20368264]
8. Zu T, et al. Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci*. 2011; 108:260–265. [PubMed: 21173221]
9. Moseley ML, et al. Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat Genet*. 2006; 38:758–769. [PubMed: 16804541]
10. Ash PEA, et al. Unconventional translation of *C9ORF72* GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron*. 2013; 77:639–646. [PubMed: 23415312]
11. Mori K, et al. The *C9orf72* GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science*. 2013; 339:1335–1338. [PubMed: 23393093]
12. Zu T, et al. RAN proteins and RNA foci from antisense transcripts in *C9ORF72* ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A*. 2013; 110:E4968–4977. [PubMed: 24248382]

13. Jiang J, et al. Gain of toxicity from ALS/FTD-linked repeat expansions in C9ORF72 is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron*. 2016; 90:535–550. [PubMed: 27112497]
14. Lagier-Tourenne C, et al. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci*. 2013; 110:E4530–E4539. [PubMed: 24170860]
15. Liu Y, et al. C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. *Neuron*. 2016; 90:521–534. [PubMed: 27112499]

Figure 1.

Molecular mechanisms of disease in *C9orf72* ALS/FTD. 1) Bidirectional transcription of G₄C₂ repeat expansion leading to sense and antisense RNAs containing G₄C₂ and G₂C₄ repeats. 2) Repeat expansions cause the formation of secondary structures of sense and antisense expansion RNA. 3) Accumulation of sense and antisense RNA into nuclear foci. 4) Export of repeat-containing sense and antisense RNA to the cytoplasm. 5) Repeat associated non-ATG (RAN) translation of sense and antisense RNAs. 6) Sense and antisense RAN proteins are generated and 7) accumulate into perinuclear aggregates. 8) C9-RAN proteins can associate with the nuclear pore complex causing impaired nucleocytoplasmic transport. (reviewed in ¹). Illustration was drawn using images from Servier Medical Art by Servier (<http://www.servier.com/Powerpoint-image-bank>, <https://creativecommons.org/licenses/by/3.0/>).