NDR Whitepaper Issue 4 August 2020

RAN Proteins

The discovery of a repeat expansion[1] in the C9orf72 gene is important because it is a model to explain central pathomechanisms of ALS. Patients with c9ALS have several hundred to several thousand C9orf72G_4C_2 repeats! The normal number of C9orf repeats is less than 30. Pathological mechanisms that are associated with the repeat expansions are 1) a change (loss or gain) of a gene function or 2) toxicity from the buildup of proteins from the expanded repeats.

The consequence of the RNA transcripts that can interact with RNA binding proteins cause problems with nucleocytoplasmic transport. Nucleocytoplasmic transport is important to trafficking proteins to the right cellular compartments. Also, the RNA transcripts can be a template for the synthesis of proteins of repeating dipeptides through RAN (repeat associated non-ATG) translation. These dipeptide proteins can be toxic to cells when they clump. Clumps are called inclusions. The neuronal tissue inclusions seen in ALS are defining or "pathognomonic" to c9ALS/FTD. C9RAN inclusions can be poly(GP), poly(GA), and poly(GR) and in some cases poly(GP), poly(PA), and poly(PR).[2] These are called dipeptide repeats.

Reports that poly(GP) are measurable in blood are interesting; poly(GP) is measurable because it is more soluble than poly(GA) protein. There are validated immunoassay techniques that measure poly(GP) proteins. Although poly(GP) tests aren't commercially available, lab-specific antibodies against dipeptides are used in research studies. Some interesting findings with poly(GP) include *no association* with age at disease onset, site of onset, disease group, or ALSFRS-R score. There was no association between poly(GP) and survival after disease onset. So what is so exciting about this line of research?

An important finding was that immortal lymphoblastoid cell lines made from blood cells from C9orf72 mutation carriers secreted poly(GP)! When the abnormal cells were treated with a drug that targets G_4C_2 RNA the mRNA expression of C9orf72 variants were eliminated. There was a significant reduction in the number of cells bearing foci formed by the poly(GP). An important observation was that poly(GP) production mirrors expression of RAN in the lymphoblastoid cell lines. Importantly, the intracellular protein was decreased with targeted treatment. Scientists had discovered that poly(GP) was stable over time and it may be used as a possible pharmacodynamic marker. That means patient-derived cells can be used in biochemical response assays to identify therapeutics that target the G_4C_2 RNA.

Further experiments showed that neurons that were differentiated from iPSC's (iPSNs)[3] replicated the genetic, transcriptional, and biochemical signatures found in the patient's brain tissue. The secretome[4] from these cells contained poly(GP) and with effective treatment the poly(GP) was reduced in a dose-dependent manner. This supported the idea that the *extracellular* poly(GP) found in the secretome may serve as a surrogate marker for *intracellular* G_4C_2 RNA accumulation.

In C9orf72 mutation carriers the extracellular poly(GP) in CSF fluid may be useful as a pharmacodynamic marker for therapies that target G_4C_2 RNA. Another observation from these works suggested that the

[1] G4C2 is related to C9orf72 gene.
[2] The sense G4C2 and antisense G2C4 produce different sets of proteins.
[3] The iPSC's were derived from a peripheral blood sample.
[4] Secretome is the culture media obtained after growing the cells in the laboratory.

primary source of poly(GP) is released from living cells, not from dying cells. That may indicate detection of poly(GP) has less utility for disease progression but doesn't affect the utility of poly(GP) as a pharmacodynamic marker. This research is good news for ALS patients with C9orf72. What about other gene-associated ALS? How about sporadic ALS (sALS), responsible for the 90% of ALS cases that aren't associated with a gene mutation?

Patients with sALS do not have a dominant gene mutation associated with their disease as do familial cases. More than 100 potential ALS genes have been reported, although the evidence for supporting a causative role varies. Sixteen genes are accepted as being implicated in ALS pathology and these genes fall into three functional categories. ALS-causing mutations are typically missense substitutions, apart from C9orf72, in which ALS is caused by an expansion of an intronic hexanucleotide repeat.

It is thought that sporadic sALS patients have a group of polymorphisms across multiple genes and/or gene families that increase the susceptibility of getting ALS in conjunction with mostly unknown environmental factors. Despite the wide variety of genetic variations that end in ALS, ALS has shown consistencies including vulnerability of cell types, progression, and pathological hallmarks of disease.

Cytoplasmic inclusions are the most prominent pathological feature of ALS and approximately 97% of ALS patients have TDP-43 as the major component of the inclusions. TDP-43 is a nuclear binding protein, and the central pathomechanism in ALS, is an abnormality in nuclear transport resulting in the mis-localization of TDP 43. Nuclear proteins that relocate to the cell cytoplasm become cytoplasmic inclusions. Genetic mutations in the TDP-43 gene only account for 2% of fALS. Surprisingly, other neurodegenerative diseases are associated with mis-localization of TDP-43 as a secondary feature. A common theme in ALS, poly(GP) in C9orf72 patients and TDP-43 in ALS, are cytoplasmic inclusions.

A unifying idea in ALS is that the dysfunction is in nucleocytoplasmic transport through the nucleopore complex (NPC). Some transcripts may deplete RNA binding protein or more compelling (evidence from C9orf cells) cells are toxic from poly-dipeptide produced by unconventional translation of repeat containing transcripts. The c9ALS expansions are a model for a central pathomechanism of ALS and a therapeutic target.