Translational Constraints Control Poly-Dipeptide Repeat Expression and RNA Metabolism from C9ORF72-Derived GGGGCC Repeat Expansions

Solomon Punobyin[1], Paige Metz[2], Katy Ovington[2], Fatima El Azzouzi[1], Emmanuel Sasu[1], Keith Gagnon[1]

[1]Wake Forest University School of Medicine, Winston-Salem, United States, *[2]Southern Illinois University School of Medicine, Carbondale, United States*

Background: The GGGGCC hexanucleotide repeat expansion in *C9ORF72* is the leading genetic driver of amyotrophic lateral sclerosis and frontotemporal dementia (C9FTD/ALS). Translation of this expansion generates toxic dipeptide repeat proteins (DPRs) that accumulate in patient tissues with a characteristic dipeptide species abundance profile, yet the precise mechanism governing this DPR hierarchy remains unclear.

Method: Here, we used inducible reporter constructs with GGGGCC repeats fused to mCherry in human cell lines as well as C9FTD/ALS patient-derived cells. Assays included flow cytometry, FISH, microscopy, RT-qPCR, and Western blot analyses.

Results: We demonstrate that GGGGCC repeat translation strictly requires a functional upstream start codon rather than initiating autonomously through internal ribosome entry. We show that short expansions selectively enhance cognate and near-cognate start codon initiation efficiency, whereas longer expansions progressively reduce total DPR output by driving decline in steady-state repeat-containing mRNA abundance. We find that characteristic DPR hierarchy is intrinsic to mammalian translation and controlled by frame-dependent codon composition. Finally, we uncover an efficient, directionally biased ribosomal frameshifting event during elongation that modulates transcript stability and repeat-containing mRNA localization.

Conclusion: Together, our findings support a multi-layered framework coupling translation elongation fidelity to mRNA surveillance in C9FTD/ALS.