

Nucleic acid acrobatics in repeat expansion disorders

Around forty percent of the human genome is composed of repetitive DNA, yet the expansion of repetitive sequences within ~50 individual genes are associated with devastating human diseases. This phenomenon was first described in 1991, but despite significant advances in disease diagnostics, a compelling molecular understanding of these disorders is still lacking. This represents a major challenge in designing successful therapeutics for these diseases. Importantly, current evidence suggests that these disorders have common underlying molecular themes, making them a prime target for future therapeutic interventions and drug re-purposing.

One common feature of these disorders is the ability of expanded repeats to form non-canonical nucleic acid structures, a concept that was first proposed soon after the genetic basis of these disorders was discovered. In contrast to the stable canonical B-form of DNA adopted by most of the genome, expanded repeats appear to favour non-canonical structures. The list of these structures is continuously growing and includes hairpins, triplexes, cruciforms, slipped structures, R-loops, RNA/DNA hybrids and G-quadruplexes. Interestingly, these structures are also formed in the genome under normal physiological conditions and can be essential for multiple biological processes. The pathological accumulation of these non-canonical structures, however, can interfere with RNA transcription, DNA replication, promote the formation of repressive chromatin marks and even drive repeat instability. These discoveries have sparked promising initial research into therapeutics that aim to target non-canonical structures in repeat expansion diseases. Approaches designed to bind slip-out structures to drive repeat contractions or to interfere with repeat-mediated silencing process by preventing RNA/DNA hybrid formation, are just the tip of the iceberg.

There are still major challenges in studying non-canonical structures, whether in health or disease. One of them is the shortage of robust and specific reagents for their detection *in vivo*. Although some such reagents exist (e.g. S9.6 antibody recognising R-loops), most techniques rely on nucleic acid fragmentation, preventing accurate determination of the true extent and global context of these structures *in vivo*. We also lack a deeper knowledge of the molecular pathways affected by these structures and cellular factors that recognise them. This makes it difficult to understand how they differ from non-pathological structures in the genome. Future work to visualise these structures in living cells and to study their kinetics by single molecule approaches will provide important insight into their role in health and disease.

We are living in exciting times; the field of expansion diseases is poised for new discoveries and ideas that will doubtless unlock the immense potential of nucleic acid therapeutics. A molecular understanding of the 'acrobatics' of non-canonical structures may be the key.