

# Review of: "Linking Dynamic DNA Secondary Structures to Genome Instability"

Albino Bacolla

**Potential competing interests:** The author(s) declared that no potential competing interests exist.

In this study the authors use an exogenous nuclease (S1 nuclease) to cleave single-stranded DNA and generate strand breaks, followed by ligation to adapters and high throughput sequencing to reveal the locations of S1 cleavage (S1-END-seq), with the purpose of detecting whether non-B DNA structures form in human cells in culture. The study is motivated by the growing realization that most genomes contain far more repetitive DNA motifs than expected by chance, and that in vitro most of these repetitive DNA motifs have been shown adopt DNA structures (non-B DNA) that do not conform to the canonical right-handed double-helix originally described by Watson and Crick (Nature, 171, 964, 1953). S1 nuclease, other nucleases and chemical probes have been widely used to characterize non-B DNA structures in vitro, including left-handed Z-DNA, cruciforms, G4 DNA, three-stranded DNA and hairpin-loops. The S1-END-seq technology advances the ability of these tools to establish whether non-B DNA also forms in living cells. The authors present evidence that microsatellite-unstable (MSI) cancer cell lines display an increased formation of cruciform structures by AT-rich repeats, whereas both MSI-positive and MSI-negative cells exhibit asymmetric strand breaks at long (>200 bp) poly(purine.pyrimidine) mirror repeats, known to form intramolecular triplex structures. The authors also use cell-cycle arrest inhibitors and aphidicolin to report that triplex structures are formed on the lagging strand during replication, and that these structures provide a block to replication. In mismatch repair-deficient cells treated with repeating doses of aphidicolin, triplex-forming repeats also displayed increased genetic instability relative to untreated cells. Furthermore, authors conduct S1-END-seq on lymphoblasts from a Friedreich's ataxia patient to show that triplex DNA is formed at the site of GAA expansion but not at a normal unexpanded repeat in a healthy sibling and that, unexpectedly, chemical re-induction of transcription at the Friedreich's ataxia locus further stimulates triplex formation. Finally, the authors conduct a targeted bioinformatics analysis to show that in 10 tumor types S1-END-seq positive triplex-forming repeats are preferred sites of genetic variation. The study is of high technical quality and the narrative is engaging the reader. It remains to be seen how the current findings may be integrated with those from other studies reporting the formation of other non-B DNA structures in cells, such as G4 DNA (Hänsel-Hertsch et al., Nat. Protoc. 13, 551, 2018, doi: 10.1038/nprot.2017.150). If confirmed by more direct methods, these finding would have implications in diseases other than Friedreich's ataxia. Long poly(purine.pyrimidine) mirror repeat sequences capable of forming triplex DNA are frequently found in introns of long human genes highly expressed in multiple areas of the brain (Bacolla et al., Nucleic Acids Res. 34, 2663, 2006, doi: 10.1093/nar/gkl354); these genes

are critical for normal brain function and appear to be susceptible to transcription and replication errors leading to spontaneous strand breaks in toddlers with autistic spectrum disorders (Wang et al., *Cell Stem Cell*, 26, 221, 2020, doi: 10.1016/j.stem.2019.12.013).