Trinucleotide repeats in the human genome: structural characteristics and role in health and disease

Nina Oparina

Engelhardt Institute of Molecular Biology RAS, Moscow, Russia, oparina@gmail.com

Platon Bykadorov

Vavilov Institute of General Genetcs RAS, Moscow, Russia, platon.work@gmail.com

Marina Fridman

Vavilov Institute of General Genetcs RAS, Moscow, Russia, marina-free@mail.ru

Trinucleotide repeats (TNR) expansions are known to be association to more than 40 human disorders, mostly neurodegenerative. Two main classes of these diseases were previously described, according to the expanded repeats location (coding or regulatory region). However, tracks of homotypic monoaminocid tracts are frequent in proteins and numerous cases of regulatory important TNRs and degenerate TNRs were described. We present the comparative study of "stable" and "unstable" trinucleotide repeats based on human population genomics and data on disease associations. Monomer types and number variation, GC-content, flanking sequence motifs, evolutionary patterns etc. were analyzed for all human- and primate-specific trinucleotide repeats for coding, non-coding, intronic, intergenic and regulatory regions. We have shown that several features associated with unstable TNRs surprisingly were enriched also in coding regions. Probable functional implications are discussed.