Statement of research
PROJECT: *Principles of Eukaryotic Genome Organization*

In 2007 we started a new project for the investigation of mid-range nonrandom patterns in genomic sequences. In thousands of genomic regions the composition of A, T, C, or G content or different combinations of these bases exist at extremes far different from the average base composition. We call such compositional extremes genomic mid-range inhomogeneity or MRI if they stretch at least 30 base pairs but less than ten thousand base pairs. To characterize genomic MRI patterns, a public computational resource (*Genomic MRI*) has been created that allows detecting sequence regions with any type of extreme composition. Using this resource it was demonstrated that various MRI regions occupy up to a quarter of the human genome and their existence is maintained via strong fixation bias. For examining mid-range sequence patterns *Genomic MRI* programs do not characterize particular "words" but only the overall compositional content of particular base(s) that we refer to as X (X could be a single nucleotide A, G, C, or T or any of their combinations like A+C, or 2 G+T+C, etc.). *Genomic MRI* allows studying the distribution of X-rich regions in any sequence of interest. These X-rich MRI regions are highly over-represented in mammalian genomes for all kinds X-contexts. We wrote two reviews illuminating that MRI regions are associated with unusual DNA conformations and/or particular DNA properties. In particular, MRI regions have complex arrangements relative to each other and to specific genomic sites such as centromers, telomeres, and promoters, pointing to their important role in genomic functioning and organization.