Investigation of DNA double-strand break repair mechanisms in microsatellite regions using the CRISPR/Cas9 system

Short tandem repeats (STRs), also known as microsatellites, are composed of a 1-6 nucleotide motif repeated 10 -50 times. They are unstable during cell division and show variable lengths within the population as well as somatic heterogeneity. These features are used in population genetics, forensic science and cancer diagnostics. In some cases, the STR length exceeds the normal length and becomes the main cause of many human genetic diseases. Expansions of the (CAG)n motif in functionally unrelated genes is a causative factor in many hereditary neurological and neuromuscular disorders, such as Huntington's disease (HD) and spinocerebellar ataxia type 3 (SCA3).

Recent years have seen tremendous progress in genome editing technology, and these new tools have enabled the creation of better disease models or development of new therapeutic approaches. Genes containing mutant CAG/CTG repeats have also been edited by zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and the clustered regulatory interspaced short palindromic repeats/Cas9 (CRISPR/Cas9) system. The early results have revealed many interesting and unexplained phenomena closely related to the properties of the STR sequences. Activation of DNA repair mechanisms as a result of genome editing in the STR regions led to repeat instability. Interestingly, some of the nucleases induced preferential shortening of the repeats, whereas others caused both expansion and shortening. However, we still do not know how to control these expansions and contractions of the repeats, how DNA breaks within STR flanking sequences may destabilize the repeats and how break type and location, CAG tract length and flanking sequence composition influence STR instability.

The presence of long tracts of repeated sequences in DNA has an impact on the chromatin structure, its stability and the binding of various factors, including the DNA repair proteins. In addition, the presence of a large number of microhomologous regions can prompt the action of noncanonical DNA repair mechanisms that are poorly understood. A better understanding of the mechanisms involved in DNA repair within STR regions will allow us to develop more effective and specific genome editing tools used, e.g., in experimental therapy for diseases caused by the STR expansion. It will also help to predict and control genome editing outcomes.

Therefore, as part of this project, we (i) characterize the targeted endonuclease-induced DNA lesions at the STR region of the *HTT* gene, (ii) analyze the factors affecting the excision/shortening of CAG repeats as a result of DNA repair, (iii) examine the contribution of genomic context to CAG repeat instability, and (iv) analyze the selected proteins and DNA repair mechanisms involved in contraction of the CAG repeats. In contrast to previous studies of STR instability, which were performed mainly in the yeast model (*Saccharomyces cerevisiae*), in this project, we will use human cell lines treated with nucleases targeting the STR region of the *HTT* gene.