# **Computational method for the detection of microsatellite instability in tumor tissue samples**

BudišJaroslav1,2,3, Sýkora Marek4, Kucharík Marcel1,2, Krampl Werner1,2, Pös Ondrej1,2, Styk Jakub1,2,5, Radvánszky Ján2,6,7, Szemes Tomáš1,2,6

*1 Geneton s.r.o., Bratislava 84104, Slovakia*

*2 Science Park, Comenius University in Bratislava, 84104, Slovakia*

*3 Slovak Center of Scientific and Technical Information, Bratislava 81104, Slovakia*

*4 Faculty of Informatics and Information Technologies, Slovak university of technology in Bratislava, 84216, Slovakia*

*5 Faculty of Medicine, Comenius University in Bratislava, 81372, Slovakia*

*6 Department of Molecular Biology, Faculty of Natural Sciences, Comenius University, Bratislava 84215, Slovakia*

*7 Institute of Clinical and Translational Research, Biomedical Research Center, Slovak Academy of Sciences, Bratislava 84505, Slovakia*

Recent advances in sequencing technologies have enabled affordable testing for various genetic diseases. Their application in the field of oncology has great potential for early detection and monitoring of tumor progression. A promising way is the thorough characterization of microsatellites, where the detection of their instable forms can be used as a cancer biomarker.

We compared sequenced DNA fragments isolated from tumour and control tissues. We first mapped them to the reference human genome. Then, thousands of selected microsatellite loci across the genome were thoroughly characterised to detect typical anomalies of unstable forms. The changes were aggregated across the analysed loci to obtain compact features for each sample. Finally, the features were analysed with classification methods to distinguish between tumor and control tissue.

We show that microsatellite instability is readable in our tumour samples. The method has therefore the potential to automate the detection and characterization of ongoing oncology disease from the sequenced genomic data.

The presented work was supported by the Slovak Research and Development Agency (grant ID APVV-18-0319). The presented work was also supported by the OP Integrated Infrastructure within projects with ITMS codes 313011W988, 313011F988, and 313011V578, all co-financed by the European Regional Development Fund.