An atlas of brain regulatory regions and regulatory networks - a novel systems biology approach to pathogenesis of selected neurological disorders

Research project objectives. The Encyclopedia of DNA Elements (ENCODE) project has mapped regions of transcription, transcription factor association, chromatin structure and histone modification and assigned biochemical functions for 80% of the genome to non-coding regions. The studies produced enormous technological advantage and extended our knowledge, but most of the data was collected using established cell lines (mostly tumor cell lines, e.g. neuroblastoma and glioblastoma), which does not reflect "real" regulatory sites in a particular cell/tissue. There is no information on **brain regulatory regions and networks due to inaccessibility of brain tissues**. International efforts such as The Cancer Genome Atlas collected vast amounts of data on genomic alterations, but failed to collect data on maps of the DNA regulatory sites. Hundreds of genome-wide association studies (GWAS) showed that disease- and trait-associated genetic variants concentrate in the non-coding DNA (93%) and their functional meaning is unknown.

Hypothesis. We would like to identify brain regulatory regions and regulatory networks in glial brain tumors (gliomas) and intersect them with ENCODE regulatory DNA maps to make **An atlas of brain regulatory regions and networks**. Due to their diffusive nature, surgically removed gliomas contain tumor tissue and approximately 40% of normal nervous tissue and immune cell infiltrates, which provides a heterogeneous pool of cells. **Our innovative strategy capitalizes on the** recent **advances in the ENCODE project, the completion of the high-resolution maps of the regulatory regions, as well as on new developments in computational data analysis and systems biology.** Based on the assumption that DNA variants associated with specific human diseases or clinical traits concentrate in the regulatory DNA sites revealed by our studies, we will pursuit whether genetic alterations and candidate variants identified by recent GWAS studies in common neuropsychiatric disorders occur within brain specific, regulatory regions.

Research methodology. We shall computationally generate regulatory DNA maps based on the ENCODE data sets and intersect them with experimentally derived brain-specific regulatory sites and open chromatin regions identified in human brain tumors of different grades and representing various neurodevelopmental states. We shall then combine existing proven technologies, including RNA-seq, high resolution DNase I-seq for the analysis of transcription factor occupancy and open chromatin sites as well as ChIP-seq data for the analysis of activating/repressing histone marks in order to generate brain-specific, regulatory DNA maps. We will apply novel and established computational methods and use systems biology approaches to build network models of genetic features in those diseases. In particular, we shall apply new developments in discovering interacting features.

High resolution mapping of DNase I hypersensitive sites across the genome shall be performed on fresh surgical tissue samples using deep sequencing. We shall examine the correspondence between DNase I footprints and known regulatory factor recognition sequences. Comprehensive scans of DNase I hypersensitive regions for transcription factor motifs in the TRANSFAC and JASPAR databases shall be performed to find enrichment of motifs. ChIP-seq data on histone marks in correlation with expression profiles shall be used to confirm active transcription sites and to generate brain-specific, regulatory DNA maps.

In our preliminary studies we have found occurrences of GWAS genetic alterations associated with neuropsychiatric disorders within the regulatory areas defined by ENCODE. Using our newly created **Atlas of brain regulatory regions and regulatory networks** we shall investigate whether genetic alterations and candidate variants identified by GWAS studies in common neuropsychiatric disorders occur within brain specific, regulatory regions.

Research project impact. The results will elucidate the today unknown role role of transcriptional and epigenetic dysfunctions in brain tumors and major neuropsychiatric disorders and likely identify novel predictions that may allow further research on better diagnostics. We may reveal novel targets for the therapy of lethal tumors. The development of new computational approaches and tools will stimulate new research in computing and statistics as well as influence the way other diseases are studied. The results will be published in high impact international journals and be presented at leading domestic and international conferences.