VISIOMICS – Platform supporting an integrated analysis of image and multiOMICs data for biology based advanced tumor diagnostics

F. Kromp[1], E. Bozsaky[1], T. Gerber[1], F. Rifatbegovic[1], Abbasi R[1]., N. Popitsch[1], M. Kauer[1], L. Fischer[2], H. Eghbal-Zadeh[3], K. Bühler[4], S. Alemzadeh[4], K. Matkovic[4], A. Kriegner[5], J. Palme[5], M. Dulovits[5], A. Hanbury[6,], M. Lupu[7], A. Taha[7], A. Bampoulidis[7], C. Ostalecki[8], M. Kampel[9], M. Atanasov[9], M. Berneder[1], A. Ziegler[1], I. M. Ambros[1], S.Taschner-Mandl[1] and P. F. Ambros[1,10]

[1] Children's Cancer Research Institute, St. Anna Kinderkrebsforschung, Vienna, Austria

[2] Software Competence Center Hagenberg (SCCH), Hagenberg, Austria

[3] Institute of Computational Perception, Johannes Kepler University of Linz, Linz, Austria

[4] VRVis Center for Virtual Reality and Visualisation GmbH, Vienna, Austria

[5] Platomics GmbH, Vienna, Austria

[6] Institute of Information Systems Engineering, TU Wien, Vienna, Austria

[7] Research Studios Austria FG, Vienna, Austria

[8] Department of Dermatology, Universitätsklinikum Erlangen, Erlangen, Germany

[9] CogVis Software and Consulting GmbH, Vienna, Austria

[10] Department of Pediatrics, Medical University of Vienna, Vienna, Austria

Background. Biology-based tumor sub-classification, risk stratification and survival prediction are typically based on cellular features, genomics, gene expression or epigenetic data. An integration of these multi-scalar genome-wide (multiOMICs) datasets have demonstrated the potential to obtain new knowledge, and, when combined with clinical data, to improve cancer diagnostics. There is, however, an urgent need for methods and certified platforms supporting multiOMICs data integration.
VISIOMICS aims to develop a platform to systematically collect multiOMICs datasets, fluorescence images and clinical data allowing the application of 1) visual data exploration tools supported by statistical methods and 2) an efficient translation of results and modules into diagnostic analysis workflows.

Methods and results. We use state-of-the-art technologies to develop a database allowing a structured upload of raw data files and metadata information. Based thereon, deep learning and bioinformatics pipelines are established to extract marker expression/gene features from multi-channel fluorescence images, to support (semi-)automatized transcriptomic and genomic data processing including manual expert-guided QC- and annotation steps. Statistical methods for feature selection and correlation analysis in combination with innovative concepts for visual analytics have been tested and allow an integrative, explorative visualization across datasets. Data storage and analysis modules are embedded in a workflow system meeting QC requirements of certified diagnostic labs.

Conclusion. Experts from medicine, biology, bioinformatics, big data analysis, deep learning and visual analytics developed innovative strategies to process, integrate and explore multiOMICs datasets, fluorescence images and clinical data. This will enable novel biological insights and an efficient development of new diagnostic analysis workflows.

Word count: 248