COMMENTARY TO HABILITATION THESIS

Next-Generation Sequencing Bioinformatics for Precision Oncology: From Immunogenetics to Multi-Omics

Mgr. Vojtěch Bystrý, Ph.D.

This habilitation thesis is a commented collection of 19 peer-reviewed publications focused on the development and application of bioinformatics approaches for next-generation sequencing (NGS) data analysis in cancer diagnostics and research. The work spans from foundational contributions to immunogenetics and clonality assessment in hematological malignancies, through the development of bioinformatics tools and platforms, to comprehensive multi-omics integration for precision oncology.

The first part describes the development of the ARResT platform for immunoglobulin and T-cell receptor (IG/TR) repertoire analysis, which became a standard tool within the EuroClonality-NGS consortium for clonality assessment and minimal residual disease (MRD) marker identification. The second part focuses on complementary bioinformatics tools developed to address the growing need for high-sensitivity mutation detection and variant calling optimization, as well as whole transcriptome analysis pipelines for structural variant detection. The final part demonstrates the evolution toward comprehensive cancer genomics—from fusion gene detection through somatic mutation analysis and epigenomics to multi-omics integration and the emerging field of radiogenomics.

The presented body of work represents a progression from specialized immunogenetics applications to broadly applicable multi-omics frameworks, contributing to the translation of NGS technologies into clinical practice for cancer patient management and laying the foundation for European-scale infrastructure development in precision oncology.

This thesis presents a collection of 19 peer-reviewed publications related to NGS data analysis for cancer applications. The publications are organized into three thematic areas, reflecting both the chronological development of the work and the logical progression from specialized to comprehensive approaches:

Part I: Immunogenetics Using NGS in Hematology-Oncology – Development of the ARResT platform for IG/TR repertoire analysis, its adoption as a standard within the EuroClonality-NGS consortium, and clinical applications in MRD detection.

Part II: Bioinformatics Tools and Platforms – Complementary tools for variant calling optimization, standardized sequence assessment, and whole transcriptome analysis pipelines for structural variant detection.

Part III: Cancer Genomics and Multi-omics Integration – Applications spanning genomics (fusion genes, somatic mutations), epigenomics, multi-omics integration in solid tumors, and the emerging field of radiogenomics.

The adaptive immune system generates an enormous diversity of immunoglobulin (IG) and T-cell receptor (TR) molecules through V(D)J recombination, creating unique “molecular fingerprints” for each lymphocyte clone. In lymphoid malignancies, this clonal rearrangement serves as a powerful biomarker for disease detection, classification, and monitoring. The application of NGS to IG/TR analysis has dramatically expanded both the depth and breadth

of possible analyses, enabling simultaneous profiling of millions of rearrangements with unprecedented sensitivity.

My entry into this field began with the development of the ARResT (Antigen Receptor Rearrangement Software for T- and B-cell analysis) platform, which provides a comprehensive suite of tools for IG/TR sequence analysis. The first component, ARResT/AssignSubsets (Publication 1), addressed the challenge of chronic lymphocytic leukemia (CLL) subclassification based on B-cell receptor stereotypy. Building on this foundation, we developed ARResT/Interrogate (Publication 2), an interactive web-based immunoprofiler for comprehensive IG/TR NGS data analysis. Together, these tools became the standard for NGS-based immune repertoire analysis within the European research community, particularly for clonality assessment and MRD detection purposes. The biological power of high-resolution repertoire analysis was demonstrated in collaboration with researchers at Imperial College London, where we showed that the fetal liver is likely the origin of life-long innate B lymphopoiesis in humans (Publication 3).

The translation of NGS-based IG/TR analysis from research to clinical diagnostics requires rigorous standardization and validation. The availability of the ARResT platform enabled the EuroClonality-NGS consortium to establish standardized protocols for clonality assessment and MRD marker identification across leading European laboratories. Publication 4 describes the standardized approach for NGS-based MRD marker identification in acute lymphoblastic leukemia, and Publication 5 provides the complementary quality control and quantification framework essential for clinical implementation. Publication 6 demonstrates the application of TCR repertoire analysis and MRD monitoring in T-cell prolymphocytic leukemia (T-PLL), revealing evidence for graft-versus-leukemia effects following allogeneic stem cell transplantation.

While developing the comprehensive ARResT platform for IG/TR repertoire analysis, the need for complementary analytical tools became apparent. GLASS (Gene Variation Assessment, Publication 7) was developed within the ERIC TP53 Network to standardize the assessment of gene variations from Sanger sequence trace data. ToTem (Tool for Variant Calling Pipeline Optimization, Publication 8) addresses the fundamental challenge of pipeline selection and configuration in NGS data analysis. Publication 9 describes bioinformatics pipelines for whole transcriptome sequencing (WTS) data analysis in leukemia patients with complex structural variants, marking a transition from focused immunogenetics toward broader cancer genomics.

The comprehensive characterization of cancer requires integration of multiple data types. Publication 10 presents an RNA capture NGS strategy for fusion gene assessment in pediatric B-ALL, while Publications 11 and 12 describe somatic mutation detection in cancer-associated genes. Publication 13 demonstrates the integration of epigenomic analysis with established immunogenetic classification schemes through WNT5A methylation profiling in CLL.

A series of collaborative studies with the Medical University of Vienna has contributed to understanding prostate cancer biology through integrated genomic and transcriptomic analysis. Publication 14 describes the role of KMT2C methyltransferase in suppressing prostate cancer metastasis. Publications 15 and 16 elucidate molecular mechanisms of tumor progression. Publication 17 represents a comprehensive multi-omics approach combining machine learning with genomic, transcriptomic, and imaging data. Publication 18 demonstrates the potential of radiogenomic markers for risk stratification in head and neck cancer. Publication 19 evaluates real-world performance of integrative clinical genomics for pediatric precision oncology.

The work presented in this thesis traces my contributions to the development of NGS bioinformatics for cancer applications over the past decade. From the foundational work on immunogenetics and clonality assessment, through the international standardization efforts of EuroClonality-NGS, to comprehensive multi-omics approaches for precision oncology in both hematological malignancies and solid tumors, this body of work reflects the rapid evolution of the field and my own scientific trajectory. Looking forward, my current research focus addresses building infrastructure for high-quality multi-omics analysis and data sharing across Europe, while the emerging field of radiogenomics represents another area of active development where non-invasive imaging can be combined with molecular profiling to provide comprehensive patient characterization.

My Contribution to NGS Bioinformatics in Cancer

Currently, I have co-authored over 38 peer-reviewed journal articles, with the majority focusing on NGS data analysis for hematological malignancies and solid tumors. I have selected 19 publications that best represent my contributions to the development of bioinformatics methods for cancer diagnostics. My contributions to these publications are summarized in the tables below, with attention to experimental bioinformatics work, supervision, manuscript preparation, and research direction.

Part I: Immunogenetics Using NGS in Hematology-Oncology

[1] Bystry, V.; Agathangelidis, A.; Bikos, V.; et al. ARResT/AssignSubsets: A Novel Application for Robust Subclassification of Chronic Lymphocytic Leukemia Based on B Cell Receptor IG Stereotypy. Bioinformatics 2015, 31(23), 3844–3846. (IF=4.98)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
80	–	70	60

[2] Bystry, V.; Reigl, T.; Krejci, A.; et al. ARResT/Interrogate: An Interactive Immunoprofiler for IG/TR NGS Data. Bioinformatics 2017, 33(3), 435–437. (IF=5.61)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
70	30	60	50

[3] Roy, A.; Bystry, V.; et al. High Resolution IgH Repertoire Analysis Reveals Fetal Liver as the Likely Origin of Life-Long, Innate B Lymphopoiesis in Humans. Clinical Immunology 2017, 183, 8–16. (IF=3.99)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
40	–	20	20

[4] Brüggemann, M.; Kotrová, M.; Knecht, H.; Bystry, V.; et al. Standardized Next-Generation Sequencing of Immunoglobulin and T-Cell Receptor Gene Recombinations for MRD Marker Identification in Acute Lymphoblastic Leukaemia; a EuroClonality-NGS Validation Study. Leukemia 2019, 33(9), 2241–2253. (IF=11.53)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
30	–	20	20

[5] Knecht, H.; Reigl, T.; Kotrová, M.; Bystry, V.; et al. Quality Control and Quantification in IG/TR Next-Generation Sequencing Marker Identification: Protocols and Bioinformatic Functionalities by EuroClonality-NGS. Leukemia 2019, 33(9), 2254–2265. (IF=11.53)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
40	–	25	25

[6] Sellner, L.; Brüggemann, M.; Bystry, V.; et al. GvL Effects in T-Prolymphocytic Leukemia: Evidence from MRD Kinetics and TCR Repertoire Analyses. Bone Marrow Transplant 2017, 52(4), 544–551. (IF=4.67)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
30	–	15	15

Part II: Bioinformatics Tools and Platforms

[7] Pal, K.; Bystry, V.; Reigl, T.; et al. GLASS: Assisted and Standardized Assessment of Gene Variations from Sanger Sequence Trace Data. Bioinformatics 2017, 33(23), 3802–3804. (IF=5.61)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
50	30	30	40

[8] Tom, N.; Tom, O.; Bystry, V.; et al. ToTem: A Tool for Variant Calling Pipeline Optimization. BMC Bioinformatics 2018, 19(1), 243. (IF=2.21)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
30	40	20	30

[9] Hynst, J.; Plevova, K.; Radova, L.; Bystry, V.; et al. Bioinformatic Pipelines for Whole Transcriptome Sequencing Data Exploitation in Leukemia Patients with Complex Structural Variants. PeerJ 2019, 7, e7071. (IF=2.38)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
40	50	30	40

Part III: Cancer Genomics and Multi-omics Integration

[10] Grioni, A.; Fazio, G.; Bystry, V.; et al. A Simple RNA Target Capture NGS Strategy for Fusion Genes Assessment in the Diagnostics of Pediatric B-cell Acute Lymphoblastic Leukemia. HemaSphere 2019, 3(3), e250.

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
30	–	20	20

[11] Kubesova, B.; Bystry, V.; et al. Low-Burden TP53 Mutations in Chronic Phase of Myeloproliferative Neoplasms: Association with Age, Hydroxyurea Administration, Disease Type and JAK2 Mutational Status. Leukemia 2018, 32(2), 450–461. (IF=11.53)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
25	–	15	15

[12] Lobello, C.; Tichy, B.; Bystry, V.; et al. STAT3 and TP53 Mutations Associate with Poor Prognosis in Anaplastic Large Cell Lymphoma. Leukemia 2021, 35(5), 1500–1505. (IF=12.53)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
30	–	15	15

[13] Poppova, L.; Bystry, V.; et al. Memory B-Cell like Chronic Lymphocytic Leukaemia Is Associated with Specific Methylation Profile of WNT5A Promoter and Undetectable Expression of WNT5A Gene. Epigenetics 2022, 17(12), 1628–1635. (IF=4.10)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
25	–	15	15

[14] Limberger, T.; Schlederer, M.; Bystry, V.; et al. KMT2C Methyltransferase Domain Regulated INK4A Expression Suppresses Prostate Cancer Metastasis. Mol Cancer 2022, 21(1), 89. (IF=37.30)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
25	–	10	10

[15] Redmer, T.; Raigel, M.; Sternberg, C.; Bystry, V.; et al. JUN Mediates the Senescence Associated Secretory Phenotype and Immune Cell Recruitment to Prevent Prostate Cancer Progression. Mol Cancer 2024, 23(1), 114. (IF=37.30)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
20	–	10	10

[16] Sternberg, C.; Raigel, M.; Limberger, T.; Bystry, V.; et al. Cell-Autonomous IL6ST Activation Suppresses Prostate Cancer Development via STAT3/ARF/P53-Driven Senescence and Confers an Immune-Active Tumor Microenvironment. Mol Cancer 2024, 23(1), 245. (IF=37.30)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
20	–	10	10

[17] Ning, J.; Spielvogel, C.P.; Bystry, V.; et al. A Novel Assessment of Whole-Mount Gleason Grading in Prostate Cancer to Identify Candidates for Radical Prostatectomy: A Machine Learning-Based Multiomics Study. Theranostics 2024, 14(12), 4570–4581. (IF=12.40)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
30	–	15	20

[18] Spielvogel, C.P.; Stoiber, S.; Bystry, V.; et al. Radiogenomic Markers Enable Risk Stratification and Inference of Mutational Pathway States in Head and Neck Cancer. Eur J Nucl Med Mol Imaging 2023, 50(2), 546–558. (IF=9.10)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
30	–	15	20

[19] Pokorna, P.; Palova, H.; Bystry, V.; et al. Real-World Performance of Integrative Clinical Genomics in Pediatric Precision Oncology. Laboratory Investigation 2024, 104(12), 102161. (IF=5.10)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
40	30	25	30

Note: The commentary must correspond to standard expectations in the field and must include a brief characteristic of the investigated matter, objectives of the work, employed methodologies, obtained results and, in case of co-authored works, a passage characterising the applicant's contribution in terms of both quality and content.