Book of Abstracts

22-30 March 2025
Otwock, Poland

Visegrad Fund

NGSchool2025

Sequencing Toolbox for Computational Biologists

Book of Abstracts

22-30 March 2025, Otwock, Poland

NGSchool2025 Team (NGSquad)

Project coordinator:	Karolina Sienkiewicz, Weill Cornell Medicine, Cornell University, US
Organising committee:	Alison Diaz, Theiagen Genomics Akshay Ganesh, Institut d'Investigació Biomèdica de Bellvitge, Barcelona, Spain Szymon Grabia, Medical University of Lodz, PL Marcin Kaszkowiak, Institute of Hematology and Blood Transfusion, PL Gustav Lindwall, Max Planck Institute for Evolutionary Biology, Germany Elene Meshki, Rhenish Friedrich Wilhelm University of Bonn, Germany Soumajit Mukherjee, Department of Infectious Diseases, King's College London, United Kingdom Zuzanna Nowicka, Medical University of Lodz, PL Srishti Patil, DKFZ (German Cancer Research Center), Heidelberg, Germany Karolina Prazanowska, Ajou University School of Medicine, Suwon, Korea Diana Sharysh, Aarhus University, DK Karolina Sienkiewicz, Weill Cornell Medicine, Cornell University, US Urszula Smyczyńska, Medical University of Lodz, PL Wiktoria Wilman, NaturalAntibody S.A., Szczecin, PL
Book of abstracts editing:	Gustav Lindwall
Conference logo designer:	Lauren Emily Fajardo







Preface

Dear attendees,

It is our great pleasure to welcome you all to Otwock for the 7th edition of the NGSchool's flagship event - **NGSchool 2025: Sequencing Toolbox for Computational Biologists**.

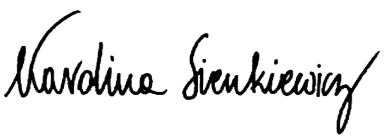
This year's Autumn/Summer School is dedicated to the latest advancements in computational biology and will cover such topics as single-cell data analysis, genomic architecture, spatial transcriptomics, genetic epidemiology and cancer evolution. In our commitment to **fostering knowledge and accessibility of scientific training**, we will be offering all course materials and recorded lectures online, free of charge.

We are proud to present an **outstanding scientific program** that includes fifteen lectures, eleven workshops, participant debates, and five parallel hackathons. Moreover, all NGSchool participants will have a chance to **present their research** and areas of interest during a dedicated session. We trust that you will find the intensive training rewarding and have an opportunity to partake in activities planned as a part of our **extensive social program**. This is an invaluable opportunity for **early-career researchers** to interact with **leading academics and professionals**, both from within and outside the V4 and CEE region.

We would like to extend our heartfelt gratitude to **Visegrad Fund** for their generous sponsorship of this year's NGSchool. Their continued support has been invaluable. Our deepest appreciation also goes out to **all our sponsors and partners for their unwavering support.**

We wish you an enlightening and enjoyable experience at NGSchool2025!

Best regards,



Karolina Sienkiewicz

NGSchool Society President
NGSchool2025 Project Coordinator

Table of contents

NGSchool Society

The previous editions of NGSchool were a big success, and their popularity motivated us to organise this summer school on an annual basis. To regulate formal aspects of our activity and secure flawless organisation of future events, we established NGSchool Society in September 2018.

The goal of the Society is to promote and support science. We do that by organising scientific events (and securing funding for such) and cooperation with scientific institutions and other scientists. While there are many great hands-on courses in bioinformatics, they tend to be expensive, especially for researchers from Central & Eastern European Countries. We decided to make a difference!

Every year we adapt the course programme accordingly to the new trends and developments in sequencing technologies. We try to address new challenges arising in computational biology and high-throughput data analysis. We will keep inviting experts in relevant fields and improve based on feedback from past editions. We are doing our best to secure funding for every course we're organising, making it accessible to everyone.

We are always open to new volunteers! Do you want to shape the upcoming editions of NGSchool? Do you want to help promote science? Do you want to organise excellent and affordable training in computational biology for young, talented scientists? If your answer is yes (or even maybe) to any of the questions - don't hesitate to contact us. We are looking forward to welcoming you to the team!

Sponsors

Visegrad Fund

The International Visegrad Fund is a donor organization, established in 2000 by the governments of the Visegrad Group countries (V4) — Czechia, Hungary, Poland and Slovakia. Its main purpose is to promote the development of closer cooperation among the Visegrad Group (V4) countries by supporting grant projects in the fields of common cultural, scientific and educational projects, youth exchanges, cross-border cooperation and tourism promotion, and by awarding scholarships and artist residency programmes. The Visegrad Fund is the main sponsor of this edition of NGSchool.

Partners

The International Institute of Molecular and Cell Biology in Warsaw (IIMCB), established in 1999, aims to carry out high-quality research in molecular biomedicine, to implement modern biotechnology, as well as to teach and popularize molecular biology and medicine. Research topics at IIMCB cover the wide area of structural biology, bioinformatics, computer modeling, molecular and cell biology, neurobiology, cancer biology, and developmental genomics.

Faculty of Mathematics, Physics and Informatics at Comenius University in Bratislava, founded in 1980, is consistently ranked first in the group of natural sciences in the ranking of faculties in Slovakia. The Faculty offers complex university-level education in all areas of mathematics, physics, and computer science. In collaboration with other units, the institution provides several interdisciplinary study programs (e.g. Biomedical Physics in collaboration with the Faculty of Medicine, and Bioinformatics in collaboration with the Faculty of Natural Sciences).

Semmelweis University in Budapest, has been a leading medical higher education institution, healthcare provider, and center of research excellence in Hungary and Central Europe for over 250 years. Its mission based on the integrity of education, research and development, and patient care has made it a regional center of excellence in the field of health sciences. The University offers academic programs that provide extensive and solid theoretical knowledge as well as competitive

practical skills in medicine, dentistry, pharmaceutical and health sciences, and conductive education.

The Institute of Biology at the Eötvös Loránd University has twelve departments covering wide areas in the life sciences. More than eighty faculty members work in close collaboration with over one hundred researchers who are associated with the institute through various fellowships, grants, and collaborative industrial projects.

ATGenomics is a Mexican company specialized in bioinformatics, computational biology, and biomedical and clinical research. Their mission is to communicate Bioinformatics with good humor, from a more human perspective and based on real experiences.

The Institute of Molecular Biology and Genetics of NASU is represented by 14 Scientific Departments.

Scientific research programme of the Institue is focused on the central trends of molecular biology, genetics and biotechnology.

Core Competencies:

- structural and functional genomics;
- proteomics and protein engineering;
- molecular and cell biotechnologies;

- bioinformatics and computational modeling and design.

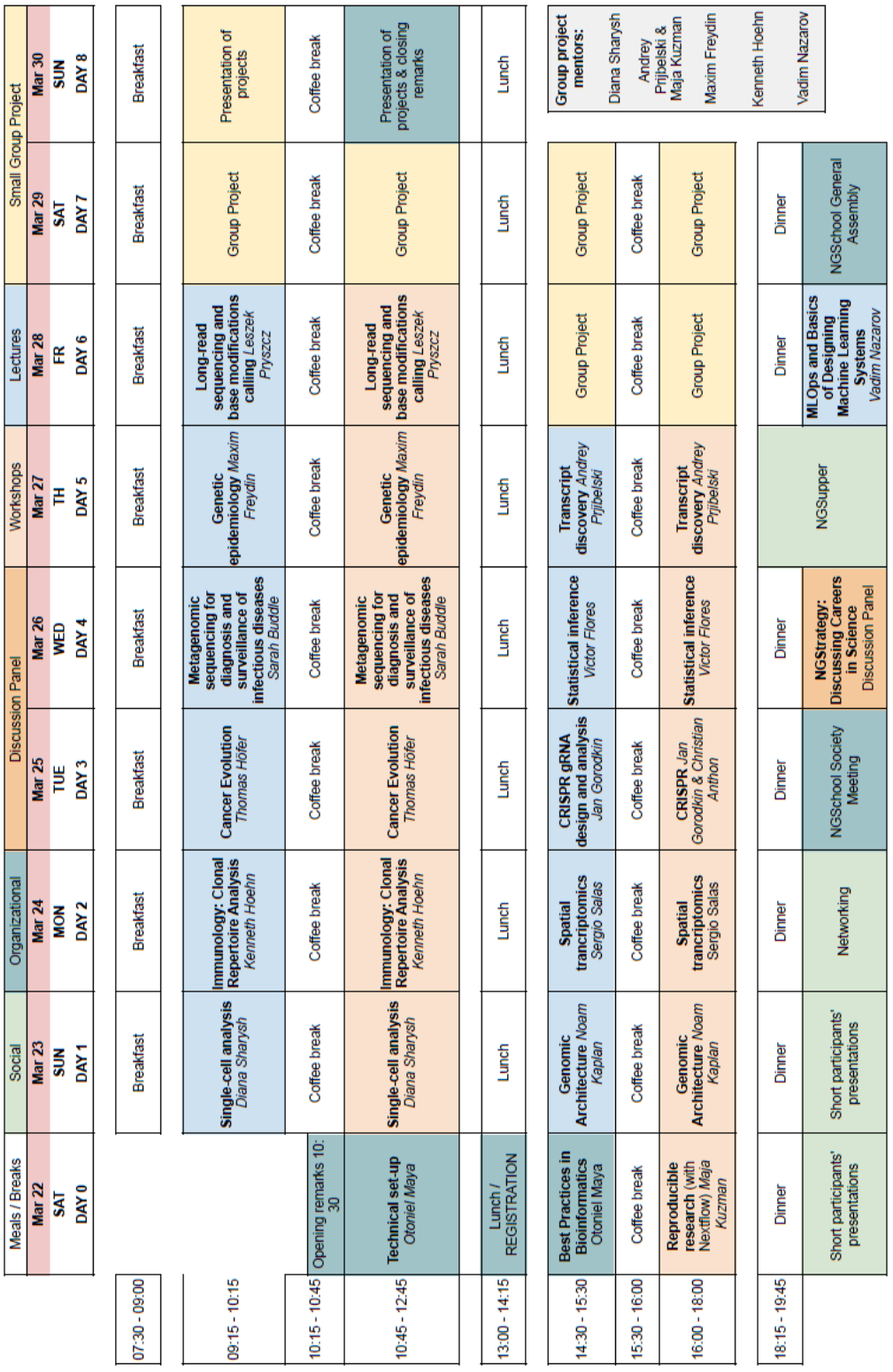
Established in 2019, the Institute of Microbial Biotechnology at the Agricultural University of Georgia is equipped with state-of-the-art scientific research facilities and equipment. It houses laboratories specializing in the bioconversion of plant substrates and the biotechnology of medical fungi.

Recognized as a leading research unit, the institute conducts both fundamental and applied research across a broad spectrum. Its mission extends beyond research to include the training of future scientists and research leaders, ensuring their expertise in both academic and industrial research. The institute provides a conducive environment and resources for career development, supporting master's and doctoral programs in the natural sciences.

ALGATECH – The Centre of Algal Biotechnology evolved from the Laboratory of Algal Research, founded in 1960 in Třeboň. Throughout its history, the Třeboň's site of the **Institute of Microbiology of the CAS** has focused on microscopic algae and their use in

food and feed industries and in human and veterinary medicine. At present, the ALGATECH Centre is an internationally recognised center for basic and applied research of microalgae, cyanobacteria, and photosynthetic bacteria, including the development of algal biotechnology.

Program



	Meals / Breaks	Social	Organizational	Discussion Panel		Workshops	Lectures	Small Group Project	
	Mar 22	**Mar 23**	**Mar 24**	**Mar 25**	**Mar 26**	**Mar 27**	**Mar 28**	**Mar 29**	**Mar 30**
	SAT	**SUN**	**MON**	**TUE**	**WED**	**TH**	**FR**	**SAT**	**SUN**
	DAY 0	**DAY 1**	**DAY 2**	**DAY 3**	**DAY 4**	**DAY 5**	**DAY 6**	**DAY 7**	**DAY 8**
07:30 - 09:00		Breakfast	Breakfast	Breakfast	Breakfast	Breakfast	Breakfast	Breakfast	Breakfast
09:15 - 10:15		**Single-cell analysis** *Diana Sharysh*	**Immunology: Clonal Repertoire Analysis** *Kenneth Hoehn*	**Cancer Evolution** *Thomas Höfer*	**Metagenomic sequencing for diagnosis and surveillance of infectious diseases** *Sarah Buddle*	**Genetic epidemiology** *Maxim Freydin*	**Long-read sequencing and base modifications calling** *Leszek Pryszcz*	Group Project	Presentation of projects
10:15 - 10:45	Opening remarks 10:30	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break
10:45 - 12:45	**Technical set-up** *Otoniel Maya*	**Single-cell analysis** *Diana Sharysh*	**Immunology: Clonal Repertoire Analysis** *Kenneth Hoehn*	**Cancer Evolution** *Thomas Höfer*	**Metagenomic sequencing for diagnosis and surveillance of infectious diseases** *Sarah Buddle*	**Genetic epidemiology** *Maxim Freydin*	**Long-read sequencing and base modifications calling** *Leszek Pryszcz*	Group Project	Presentation of projects & closing remarks
13:00 - 14:15	Lunch / REGISTRATION	Lunch	Lunch	Lunch	Lunch	Lunch	Lunch	Lunch	Lunch
14:30 - 15:30	**Best Practices in Bioinformatics** Otoniel Maya	**Genomic Architecture** *Noam Kaplan*	**Spatial trancriptomics** *Sergio Salas*	**CRISPR gRNA design and analysis** *Jan Gorodkin*	**Statistical inference** *Victor Flores*	**Transcript discovery** *Andrey Prjibelski*	Group Project	Group Project	**Group project mentors:** Diana Sharysh, Andrey Prjibelski & Maja Kuzman, Maxim Freydin, Kenneth Hoehn, Vadim Nazarov
15:30 - 16:00	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	Coffee break	
16:00 - 18:00	**Reproducible research** (with Nextflow) *Maja Kuzman*	**Genomic Architecture** *Noam Kaplan*	**Spatial trancriptomics** Sergio Salas	**CRISPR** *Jan Gorodkin & Christian Anthon*	**Statistical inference** *Victor Flores*	**Transcript discovery** *Andrey Prjibelski*	Group Project	Group Project	
18:15 - 19:45	Dinner	Dinner	Dinner	Dinner	Dinner	NGSupper	Dinner	Dinner	
	Short participants' presentations	Short participants' presentations	Networking	NGSchool Society Meeting	**NGStrategy: Discussing Careers in Science** Discussion Panel		**MLOps and Basics of Designing Machine Learning Systems** *Vadim Nazarov*	NGSchool General Assembly	

Selected speaker's abstracts

Bioinformatics as a Universal Language: From Genomes to Global Impact

Victor Flores
University of Cambridge, UK

Bioinformatics is a fascinating discipline encompassing computer science, biology, mathematics, and many other related areas. But what's more fascinating about bioinformatics is that you can apply basic principles of data analysis to multiple organisms, areas of research, healthcare, and even social sciences. Importantly, we can also have fun doing bioinformatics but ultimately we can also put our knowledge to good use and improve individual's health, develop agriculture programs, help developing environmental policies, and keep deadly viruses at bay with genomic surveillance.

Cancer Evolution

Prof. Thomas Höfer
DKFZ Heidelberg

The vast majority of cancers are linked to gene mutations that provide selective advantage to proliferating cells. At the same time, neutral mutations accumulate in the genomes of somatic cells. I will introduce principal ideas that allow us to infer, from the patterns of neutral variants, clonal selection in somatic tissues. I will then illustrate how to formulate mathematical models to reconstruct the evolutionary dynamics of individual tumors from genome sequencing data. The results of this approach indicate that there is enormous scope for improving early detection of malignant transformation.

Participants' abstracts

Dissecting Tumor Heterogeneity in Brain Cancer Through Single-Cell and Spatial Transcriptomics of Key Mutations

Barnabás Németh
German Cancer Research Center, Heidelberg, Germany.

My research focuses on studying the role of key characteristic mutations in brain tumors and how they influence tumor biology and therapeutic response.
To investigate these mutations, I employ a 10X Genomics-based single-cell transcriptomics workflow on fresh-frozen human brain tumor samples and genetically engineered mouse cell lines with inducible mutations. This approach allows for high-resolution analysis of tumor cell populations and their expression profiles. Sequencing is performed on the Oxford Nanopore platform, enabling simultaneous characterization of different cell types and detection of mutations in our genes of interest. By integrating transcriptomic data with mutational analysis, I aim to gain deeper insights into how these genetic alterations drive treatment resistance.
In parallel, I am developing fluorescent immunohistochemistry (IHC) protocols optimized for formalin-fixed, paraffin-embedded (FFPE) tumor specimens, which are widely available in clinical settings. This allows for the identification and characterization of specific cell types based on protein expression patterns. After successfully identifying these cell populations, I plan to implement the Light-Seq assay on stained sections. Light-Seq, developed by Dr. Sinem K. Saka's group at the European Molecular Biology Laboratory (EMBL), is an advanced spatial transcriptomics method that utilizes light-directed in situ barcoding to spatially index biomolecules in fixed tissues. This enables multiplexed, non-destructive transcriptomic profiling while preserving the integrity of the sample for further downstream analyses. By integrating protein marker data with spatially resolved gene expression, this approach provides a powerful tool for studying the transcriptome of very specific cell populations.
Ultimately, my research aims to assess the effectiveness of current and emerging treatment strategies, particularly those in clinical trials, by characterizing tumor heterogeneity at the molecular and cellular levels. A deeper understanding of these mechanisms could contribute to the development of more precise and targeted therapeutic approaches, potentially improving treatment outcomes for patients with brain tumors.

Immunological Profiling of Resident Kidney Cells: Linking Phagocytic Mechanisms to Inflammatory Protection in Renal Disease

Marlena Typiak
University of Gdansk, Gdansk, Poland.

During my PhD studies I have analyzed genes of receptors for immunoglobulin G, which influenced whether a person encountering the same environmental risks will stay healthy, develop tuberculosis or sarcoidosis. These studies proved that the genetically encoded differences in phagocytosis initiation can be crucial for a disease development. Afterwards, during my postdoctoral fellowship, I conducted research on diabetic nephropathy and an anti-inflammatory Klotho protein, which may protect resident kidney cells from inflammation and damage in the course of diabetes. Now, as an assistant professor at the University of Gdansk, I would like to find a link between the previously analyzed topics and study the immunological features of resident kidney cells, which are able to phagocyte, present foreign antigens, secrete complement components and interleukins. My research is aimed to learn how to modulate the immune response in the kidneys to protect them from damage during diabetes, but also other renal tissue disorders, like lupus nephritis.

Molecular Epidemiology of HIV-1 in Hungary: Ten Years of Transmission Dynamics and Cluster Analysis

Levente Zsichla
Eötvös Loránd University, Budapest, Hungary

HIV-1 causes a lifelong, incurable infection; therefore, monitoring and understanding its transmission patterns across diverse populations and regions is essential for developing and improving prevention strategies. In Hungary, drug resistance genotyping was incorporated into routine clinical practice in 2016, yielding approximately 100 partial pol sequences from ~200 newly diagnosed cases per year over the last decade. Based on these data we aim to present the most extensive molecular epidemiological analysis of the HIV-1 epidemic in Hungary to date.

In total, we analyzed sequences from 1120 patients up until 2024, along with 2199 unique international background sequences. We performed subtyping, drug resistance analysis, maximum likelihood and Bayesian phylogenetic inference, as well as distance-based and phylogenetic clustering analyses to identify and characterize long-term HIV-1 transmission clusters both within Hungary and across international borders.

From the Hungarian sequences, we identified 86 sequence clusters with 44 containing at least five sequences each; 32 of these were predominantly subtype B, and in 38 the MSM risk group was the dominant (>80%). Members of larger clusters (10+ sequences) tended to be younger, more likely to be MSM, and had higher CD4 counts than patients not assigned to large clusters. We identified seven large transmission clusters (with 20+ sequences), of which only two MSM clusters showed substantial growth in 2023 and 2024. When incorporating international sequences (until 2022, final dataset is under analysis), we identified 112 putative transmission clusters, of which 35 contained at least 10 sequences. Of these, 7 (20%) were composed exclusively of Hungarian sequences, 10 (29%) exclusively of international sequences, and 18 (51%) contained a mix of both Hungarian and international sequences. The mixed clusters revealed epidemiological links to several European countries, most frequently to nations within Central Europe.

Multiple independent introductions of the virus occurred in Hungary; however, in recent years, the epidemic has primarily spread within the country's borders. The most intense spread has predominantly affected middle-aged MSM individuals, and our findings indicate a significant separation between MSM and heterosexual risk groups. While we identified several large, long-term clusters, our results suggest that not all are active and, therefore, epidemiologically significant.

Astrocytic Gene Dysregulation in Brodmann Area 25: Glial Signatures and Glutamate Metabolism in Major Depressive Disorder

Aleksandra Herud
Łukasiewicz Research Network - PORT Polish Center for Technology Development, Wrocław, Poland

Major Depressive Disorder (MDD) is one of the main public health issues, affecting approx. 300 million people in all age groups. Disturbed metabolic activity in depression was reported in several brain regions, including subgenual prefrontal cortex. Among those regions is Brodmann area 25 (BA25), a center engaged in stress response, emotion regulation and systemic metabolism. Growing evidence from worldwide research initiatives highlights profound alteration of glial cells across psychiatric symptoms.

In this study, transcriptional analysis of a post-mortem samples from a subpopulation of suicide completers with reported glial abnormalities was performed. Our findings reveal that a substantial share of differentially regulated genes in BA25 were those enriched in glia, with astrocyte-specific genes showing the biggest disregulation, followed by oligodendrocytes and oligodendrocyte precursor cells (OPCs).

In order to investigate potential regulators of altered genes, we performed transcription factor over-representation analysis (ORA) in sets of cell-type specific and unclassified DEGs, using the ChEA 2022 database. With two distinct methods we further examined astrocyte-specific gene networks affected in MDD using GO database. Gene set enrichment analysis (GSEA), applied to all genes identified in astrocyte-enriched nuclei RNA-Seq, highlighted key pathways involved in glutamate metabolism, fatty acid processing, developmental processes, and protein localization at the plasma membrane. Complementary ORA of protein-coding astrocyte-specific DEGs supported these findings and revealed significant role of hormone receptors, with the most prominent being glucocorticoid receptor (GR), encoded by NR3C1 gene, across multiple datasets.

To refine our understanding of astrocyte reprogramming in MDD, we applied a novel protocol for enriching astrocytic nuclei. Based on DGE analysis we again performed ORA and GSEA that identified astrocytes-specific pathway engaged in glutamate uptake and turnover. Key molecular components of this pathway, including the glutamate transporter (SLC1A2) and glutamine synthetase (GLUL), were among the most abundant transcripts but showed significant downregulation in astrocytes from individuals with MDD.

Genomic and Transcriptomic Characterization of RNA Silencing Pathways in Terrestrial Slugs of the *Deroceras* Genus

Kateryna Nemesh
Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic

RNA silencing is a key biological process, yet its evolutionary history and molecular mechanisms remain poorly understood in non-model organisms. Terrestrial slugs, with their distinct ecological adaptations and ease of laboratory maintenance, offer a promising system for such studies. Here, we present Deroceras laeve and Deroceras invadens as emerging models for investigating RNA silencing pathways.

To establish genomic resources, we sequenced and assembled the genomes of D. laeve and D. invadens, generating high-contiguity assemblies (N50 = 3 Mb for D. invadens; chromosomal-level assembly with N50 = 59 Mb for aphallic D. laeve), both with over 90% completeness based on BUSCO analysis. We also sequenced transcriptomes for D. laeve, D. invadens, and the invasive pest Arion vulgaris, identifying key RNA silencing genes, including Dicer, two Argonaute, and two PIWI homologs. Conservation of the DEDH catalytic tetrad in Argonaute and PIWI proteins in Deroceras species suggests functional RNA silencing pathways.

To explore small RNA biology, we generated small RNA sequencing data from all three species. In D. laeve, we annotated 170 primary piRNA clusters, shedding light on their genomic organization.

Somatic Retrotransposition in the Human Brain: Regional Profiles of L1Hs and AluYa5 Activity

Aleksei Kurnosov
Some University, Some City, Some Country

Retroelements constitute nearly 42% of the human genome, making them the most abundant subclass of mobile elements. Although often regarded as “junk” DNA, they play a significant role in genome dynamics. While the vast majority of human retroelements have lost their ability to transpose, a few evolutionarily young families of SINEs (AluYa5) and LINEs (L1Hs), remain active in both somatic and germline cells. Their activity can lead to gene expression alterations and global reshaping of the genome structure.

Our primary objective was to perform a genome-wide search for de novo mobile element insertions in somatic neural cells and compare retrotransposition profiles across different regions of the human brain. To achieve this, we developed a method for retrieving the flanking sequences of active retroelements from genomic DNA using a multi-step selective suppressive PCR approach. This method enriched DNA samples with rare molecules representing somatic insertion events. The resulting samples were sequenced using the Illumina platform, and the sequences were mapped to the human genome. Newly identified insertion coordinates were compared against the coordinates of known retroelements, with insertions unique to a single tissue considered potentially somatic.

We analysed DNA from five tissues – cerebellum, cortex, subventricular zone of the left ventricle, dentate gyrus, and myocardium – of a single individual. Our analysis identified 7,497 somatic L1Hs and 8,990 somatic AluYa5 insertions. The highest number of somatic insertions was found in the dentate gyrus, a region known as a hotspot for adult neurogenesis.

Investigating the Regulatory Role of (p)ppApp in *E. coli*: A Transcriptomic and Phenotypic Perspective

Katarzyna Bryszkowska
University of Gdańsk, Gdańsk, Poland

The stringent response is an important aspect of the global stress response in bacteria, whose main mediators are the nucleotides ppGpp and pppGpp, which are a part of a class of second messenger family of molecules named alarmones.
In Escherichia coli, (p)ppGpp affects transcription mainly in two ways – through direct interaction with RNA polymerase (RNAP) holoenzyme at two distinct sites (site 1 is situated between the β´ and ω subunits, whereas site 2 is situated between the β´and β subunits and is available only in presence of DksA) and indirectly, through a process called σ-factor competition, enabling adaptation to changing environmental conditions.
(p)ppApp, 3′pyrophosphate nucleotide derivatives analogous to (p)ppGpp, have been identified almost 50 years ago in sporulating Bacillus subtilis cells, yet their exact function in bacterial cells is still unknown. Since then, however, a number of enzymes that metabolise (p)ppApp have been characterised, and pppApp synthesis has been confirmed in vivo in E. coli cells.
The objective of my dissertation is to investigate the effects of (p)ppApp on Escherichia coli cells using transcriptomic and phenotypic studies, with the assumption of the following research hypothesis - (p)ppApp is a bacterial signalling nucleotide with global regulatory effects.
In vitro, (p)ppApp affects transcription, specifically the initiation of transcription from the rrnB P1 promoter of E. coli, in the opposite way to (p)ppGpp - promoting it, whereas (p)ppGpp is a known inhibitor of transcription from this promoter. Similar to (p)ppGpp, in vitro (p)ppApp interacts directly with RNAP, but unlike (p)ppGpp it binds close to the active site of RNAP.
The identification of potential cellular (p)ppApp targets may allow a broader understanding of cellular processes and their regulation in E. coli.

Genomic, Transcriptomic, and Proteomic Dissection of Biofilm Formation and Virulence in *Pseudomonas aeruginosa*

Aleksandra Kujałowicz
University of Gdańsk, Gdańsk, Poland

The stringent response is an important aspect of the global stress response in bacteria, whose main mediators are the nucleotides ppGpp and pppGpp, which are a part of a class of second messenger family of molecules named alarmones.
In Escherichia coli, (p)ppGpp affects transcription mainly in two ways – through direct interaction with RNA polymerase (RNAP) holoenzyme at two distinct sites (site 1 is situated between the β´ and ω subunits, whereas site 2 is situated between the β´and β subunits and is available only in presence of DksA) and indirectly, through a process called σ-factor competition, enabling adaptation to changing environmental conditions.
(p)ppApp, 3′pyrophosphate nucleotide derivatives analogous to (p)ppGpp, have been identified almost 50 years ago in sporulating Bacillus subtilis cells, yet their exact function in bacterial cells is still unknown. Since then, however, a number of enzymes that metabolise (p)ppApp have been characterised, and pppApp synthesis has been confirmed in vivo in E. coli cells.
The objective of my dissertation is to investigate the effects of (p)ppApp on Escherichia coli cells using transcriptomic and phenotypic studies, with the assumption of the following research hypothesis - (p)ppApp is a bacterial signalling nucleotide with global regulatory effects.
In vitro, (p)ppApp affects transcription, specifically the initiation of transcription from the rrnB P1 promoter of E. coli, in the opposite way to (p)ppGpp - promoting it, whereas (p)ppGpp is a known inhibitor of transcription from this promoter. Similar to (p)ppGpp, in vitro (p)ppApp interacts directly with RNAP, but unlike (p)ppGpp it binds close to the active site of RNAP.
The identification of potential cellular (p)ppApp targets may allow a broader understanding of cellular processes and their regulation in E. coli.

In Silico Design of a Multivalent Subunit Vaccine Against ASFV and PCV2 Co-Infection in Swine

Lauren Emily Fajardo
DOST-ITDI, Taguig, Philippines

African swine fever virus (ASFV) and porcine circovirus type 2 (PCV2) are two prevalent and economically significant viruses causing high rates of pig mortality and large-scale losses to global pork production. Since there is currently no vaccine simultaneously targeting both viruses, this study aimed to computationally design a safe, stable, and effective multi-epitope based multivalent subunit vaccine against co-infecting ASFV and PCV2. Cytotoxic T-lymphocyte (CTL), helper T-lymphocyte (HTL), and linear B-lymphocyte (LBL) epitopes were screened from sequences of the Rep, Cap, and ORF3 PCV2 proteins. PCV2 epitopes predicted to be antigenic, non-allergenic, and non-toxic were linked to previously screened ASFV epitopes and Vibrio vulnificus FlaB flagellin as an adjuvant to create the final vaccine construct, which underwent physicochemical assessment and structure prediction. The vaccine construct was predicted to be stable, soluble, non-cross-reactive, antigenic, and nonallergenic. An immune simulation demonstrated that the vaccine could elicit robust antibody, T-cell, and B-cell responses. The vaccine construct stably docked to TLR5 and formed significant molecular interactions. A 200-ns molecular dynamics simulation showed that the vaccine-TLR5 complex exhibited stability and compactness throughout the run. These results show that the designed vaccine is safe, stable, and effective and warrants experimental validation.

Genetic Variants Associated with Recurrent Pregnancy Loss in Uzbek Women: A Genome-Wide Association Study

Anastasiya Punko
Center for Advanced Technologies, Tashkent, Uzbekistan

Background: Recurrent pregnancy loss (RPL) is a condition characterised by the loss of two or more consecutive pregnancies before the 20th week of gestation, affecting approximately 15% of all pregnancies. Despite significant advancements in infertility diagnostics and reproductive medicine, the underlying cause remains unknown in 35-60% of RPL cases. This suggests a complex interplay of genetic, epigenetic, and environmental factors contributing to its pathogenesis. In this study, we conducted a genetic analysis to identify variants associated with RPL among women in Uzbekistan, aiming to enhance our understanding of its genetic basis.
Material and Methods: This study included 161 patients diagnosed with RPL and 511 healthy individuals. Genetic screening was performed using the Global Screening Array v.3.0. Hardy-Weinberg equilibrium test and logistic regression analyses were conducted using the SNPassoc software package to identify statistically significant associations between genetic variants and disease risk under an additive genetic model.
Results: Multiple genetic variants demonstrated significant associations with RPL. Notably, rs4871372 (LOC105375725 , p = 3.7 × 10-5, OR = 0.58) and rs2704035 (SULF1 , p = 3.9 × 10-4, OR = 0.63) were identified as protective factors, potentially reducing the risk of RPL. In contrast, rs4916848 (ADGRV1 , p = 1.8 × 10-6, OR = 1.82) and rs478993 (TENM4 , p = 4.5 × 10-4, OR = 1.58) were associated with an increased susceptibility to RPL. These findings highlight the potential role of these genetic variants in influencing pregnancy outcomes and warrant further investigation to elucidate their biological significance.
Conclusion: Our study identified novel genetic associations with RPL in Uzbek women. The IncRNA LOC105375725 (rs4871372) may function as an epigenetic regulator, while rs4916848 likely acts as an enhancer. Additionally, TENM4 has been implicated in mesoderm induction during early embryogenesis. Therefore, these variants can be used as diagnostic and prognostic markers. Further research is needed to understand their functional role of these variants in RPL pathogenesis.

Single-Cell Transcriptomic Analysis Reveals Clonal Evolution and Site-Specific Transformation in Mantle Cell Lymphoma

Dmitrij Kazancev
Institute of Pathological Physiology First Faculty of Medicine, Charles University, Prague, Czechia

Mantle cell lymphoma (MCL) is a chronically relapsing B cell lymphoproliferative malignancy characterized by translocation t(11;14)(q13;q32), cyclin D1 overexpression and significant chromosomal instability. One of the main hypotheses of MCL relapses is a clonal evolution driven by therapy selection of subpopulations carrying adverse genetic and epigenetic aberrations. Our previous whole exome sequencing analysis of 25 MCL tumors showed that the majority of the detected single-nucleotide variants and copy number variants (CNV) were already detectable at diagnosis. To further analyze the subclonal structure of MCL at diagnosis and clonal evolution at relapse, we implemented single cell RNA sequencing analysis of five patients with MCL at diagnosis and at the first clinical relapse after failure of standard immunochemotherapy. All analyzed patients achieved at least partial remission.
Fresh-frozen tumor cells were thawed and sorted to obtain CD45+ population which was processed using droplet-based single cell 3′ protocol (10X Genomics). Resulting libraries were paired-end sequenced and mapped to GRCh38.
Initial clustering of total cell dataset revealed that non-malignant cells have grouped together irrespective of the patient or compartment of origin while the malignant lymphocytes separated into unique clusters. This suggests that genetic background of the MCL cells overrides the impact of the tumor microenvironment.
We then used CNV inference algorithm to analyze subclonal structure of each malignant cell population which showed that the subclones predominantly present at relapse are already detectable as minor subclones at diagnosis. Comparison of these clones at diagnosis with the cells eliminated by therapy revealed common changes in 119 signaling pathways.
Additionally, we have molecularly characterized a late blastoid relapse of originally indolent MCL as long as 7 years after the therapy. We analyzed 3 different compartments at the relapse (blood, bone marrow, and infiltrated intestine) and showed that the blastoid transformation was detectable only in the infiltrated intestine, but not in the bone marrow or blood - these compartments were largely similar to the diagnostic indolent MCL clone detectable in the blood before therapy initiation. These data suggest a late relapse of the indolent MCL with a blastoid transformation in the intestine.
Our data show potential applicability of single cell RNA sequencing for tailored therapy of MCL patients.

Evolution of the Sulfo-EMP Pathway in Bacteria: A Comparative Genomics and Structural Biology Approach

Anna Rybina
Independent researcher, Moscow, Russia

I am a member of a lab that applies comparative genomics to study bacterial evolution, with a focus on carbohydrate metabolism. My research investigates the catabolism of sulfoquinovose, a widespread sulfo-derivative of glucose, and the evolution of the sulfo-Embden–Meyerhof–Parnas (sulfo-EMP) pathway, also known as sulfo-glycolysis, in bacteria. My work integrates phylogenetics, genomic context analysis, structural modeling (including molecular dynamics simulations, protein structure predictions, and docking studies), and gene expression analysis. This research forms the foundation of my PhD dissertation. Beyond my primary work, I am self-studying single-cell data analysis to expand my expertise in cutting-edge genomic technologies. Additionally, I serve as a teaching assistant at the Bioinformatics Institute, where I contribute to educating the next generation of bioinformaticians.

Integrative Transcriptomic Approaches to Uncover Genetic Heterogeneity in Psychiatric Disorders

Georgii Dolgalev
University of Münster, Münster, Germany

Psychiatric disorders such as schizophrenia, bipolar disorder and major depressive disorder are complex and highly polygenic. Although recent genome-wide association studies revealed hundreds of genetic variants associated with these diseases, comprehensive characterization of their functional significance at the molecular, cellular and tissue levels has proven difficult. New methods such as transcriptome-wide association studies promise to advance our understanding of the genetics of psychiatric disorders by linking genetic variants to alterations in gene expression in a tissue-specific manner. CASTom-iGEx, a workflow developed in our laboratory, allows to learn and then predict gene expression in large genotyped cohorts and also perform unsupervised clustering of patients to reveal potential genetic heterogeneity. We have previously applied this workflow to schizophrenia and coronary artery disease, and now we plan to extend the analysis to other psychiatric disorders and neurobiological traits. In parallel, we validate the findings from transcriptome-wide association studies using iPSC-derived neuronal cultures.

Computational Prediction and Design of Cell-Penetrating Peptides Based on Sequence and Experimental Context

Violetta Isakova

The work is devoted to cell-penetrating peptides. A classifier and regressor were written that predict the ability of a peptide to penetrate a cell, taking into account its sequence, chemical structure and experimental conditions. Currently, a generator of cell-penetrating peptides is being written.

Splicing Dynamics of lncRNA PVT1 in Breast Cancer: A Potential Biomarker and Regulator of miRNA-Mediated Gene Expression

Angelos Kozonakis
IMBB-FORTH, Heraklion, Crete, Greece

Long non-coding RNAs (lncRNAs) have been increasingly recognized for their involvement in regulating gene expression and influencing various biological processes. In cancer, they are thought to play important roles by modulating cellular mechanisms such as proliferation and apoptosis. This study investigates the regulatory influence of the long non-coding RNA (lncRNA) PVT1 (Plasmacytoma Variant Translocation 1) in breast cancer, focusing on its splicing dynamics and interactions with microRNAs (miRNAs) and their target genes. Leveraging data from The Cancer Genome Atlas (TCGA) BRCA dataset, encompassing control and tumor samples (de- identified), splicing efficiency was quantified across the splice junctions of the PVT1 locus, with a specific focus on the 3' splice sites that could potentially serve as hotspots for differential splicing. Through supervised and unsupervised machine learning techniques, using 3’ splice site splicing efficiency as features, distinct splice sites were identified that were significantly associated with dysregulated splicing between control and tumor conditions, underscoring PVT1 splicing as a possible interpretable biomarker in cancer. Further analysis examined PVT1’s role as a miRNA sponge and how its splicing dynamics might influence miRNA interactions and downstream gene expression. Linear models identified several genes whose exon-level expression correlated with PVT1’s 3’ splice-site splicing efficiency. Notably, while correlations between PVT1 splicing and expression patterns of miR-200/205 target genes were observed, PVT1’s overall steady-state expression was not predictive, suggesting a regulatory mechanism in trans driven by splicing activity rather than expression levels. These findings underscore the functional importance of PVT1 splicing in gene regulation, presenting it as a potential marker for distinguishing tumor samples and a target for therapeutic intervention. This work lays the foundation for further investigations into PVT1’s role in cancer biology, offering insights into novel biomarkers and strategies for targeted therapy within the PVT1 locus.

Standardizing Glucose Variability Metrics: Rigorous Definitions and Open-Source Implementation for CGM Data Analysis

Jędrzej Chrzanowski
Department of Biostatistics and Translational Medicine, Medical University of Lodz, Lodz, Poland

With the increasing prevalence of diabetes and the introduction of continuous glucose monitoring (CGM) as a standard, there is an urgent need for mathematically rigorous definitions of glucose variability (GV) metrics. The FDA's recent approval of these metrics as clinical trial endpoints underscores this requirement. Although manufacturer-guaranteed reports exist, no public reference dataset standardizes their computation or accuracy, and the existence of several independent open tools and the common practice of ad hoc code creation for GV calculation further widen the gap.

My work aimed to draft rigorous mathematical definitions for standard GV metrics and provide an open-source Python implementation. Additionally, I wanted to identify edge cases for GV metrics and test the proposed Python implementation against manufacturer-provided and open-source platforms.

I began by reviewing the current literature and landscape of GV metrics algorithms employed by open tools. Manufacturers' tools were inaccessible for review due to their closed-code nature. We developed a set of precise mathematical definitions for key GV parameters. These definitions were then implemented in Python. A series of real CGM datasets from most common manufacturers were used to stress-test the implementation, focusing on edge cases such as data gaps and abnormal glucose trajectories. Comparative analyses were performed to benchmark our results against existing manufacturer reports and open-source tools.

We present a comprehensive set of mathematically rigorous definitions for standard GV metrics alongside open-source implementation. The open-source implementation improves GV metrics reproducibility, especially regarding edge cases that previously led to discrepancies. The work standardizes the computation of GV metrics and highlights the importance of addressing edge cases to ensure accuracy.

Coordination of Loop Extrusion, Cohesion, and Transcription on Sister Chromatids in G2 Phase

Dmitry Mylarshchikov
IMBA - Institute of Molecular Biotechnology GmbH, Vienna, Austria

With the increasing prevalence of diabetes and the introduction of continuous glucose monitoring (CGM) as a standard, there is an urgent need for mathematically rigorous definitions of glucose variability (GV) metrics. The FDA's recent approval of these metrics as clinical trial endpoints underscores this requirement. Although manufacturer-guaranteed reports exist, no public reference dataset standardizes their computation or accuracy, and the existence of several independent open tools and the common practice of ad hoc code creation for GV calculation further widen the gap.

My work aimed to draft rigorous mathematical definitions for standard GV metrics and provide an open-source Python implementation. Additionally, I wanted to identify edge cases for GV metrics and test the proposed Python implementation against manufacturer-provided and open-source platforms.

I began by reviewing the current literature and landscape of GV metrics algorithms employed by open tools. Manufacturers' tools were inaccessible for review due to their closed-code nature. We developed a set of precise mathematical definitions for key GV parameters. These definitions were then implemented in Python. A series of real CGM datasets from most common manufacturers were used to stress-test the implementation, focusing on edge cases such as data gaps and abnormal glucose trajectories. Comparative analyses were performed to benchmark our results against existing manufacturer reports and open-source tools.

We present a comprehensive set of mathematically rigorous definitions for standard GV metrics alongside open-source implementation. The open-source implementation improves GV metrics reproducibility, especially regarding edge cases that previously led to discrepancies. The work standardizes the computation of GV metrics and highlights the importance of addressing edge cases to ensure accuracy.