Poster abstracts

1. Multiomic analysis by Teton CytoProfiling for drug screening applications at NGI

Federico Agostini, Karolinska Institutet and Science for Life Laboratories, Sweden

Abstract: The AVITI24 platform, integrated with Teton CytoProfiling chemistry, enables high-throughput, multiomic drug screening by combining cellular imaging and sequencing in a single workflow.

Contact: *federico.agostini@ki.se*

---

2. Agent Lens: LLM-Powered Autonomous Agent for Smart Microscopy

Songtao Cheng1, Hanzhao Zhang1, Hugo Dettner Källander1, Nils Mechtel1, Konrad Olszewski1, Wei Ouyang1:email:. 1 Department of Applied Physics, Science for Life Laboratory, KTH Royal Institute of Technology, Stockholm, Sweden.

Abstract: Smart microscopy has made high-throughput imaging possible, but most systems still rely on fixed, human-designed pipelines that cannot adapt to changing samples, imaging conditions, or analysis needs. This limits error recovery, experiment scalability, and true autonomy. Agent Lens is an AI-powered microscopy framework that uses large language model (LLM) agents on top of the Hypha Agents1 platform to observe images, reason about acquisition goals, and execute microscope actions. By combining LLMs, CLIP-based image embeddings2, segmentation3, and vector search, the system enables interactive navigation of samples and autonomous acquisition of cells of interest.

Contact: *songtao.cheng@scilifelab.se*

---

3. Spatial MetaTranscriptomics reveals a JAK/Stat high activity niche associated with specific microbiota

Yuhao Chi, Hailey Sounart, Maria Lovinelli, Joseph Tintelnot, Johan Thorsson, Ece Kartal, Fabian Springer, Maria Alejandra Ulloa, Ricardo O. Ramirez Flores, Francisca Castillo, Bianca Kern, Gustavo Monasterio, Charlotte Hedin, Georg Zeller, Samuel Huber, Nicola Gagliani,Julio Saez-Rodriguez, Stefania Giacomello & Eduardo J. Villablanca Division of Immunology and Respiratory Medicine, Department of Medicine Solna, Karolinska Institute and University Hospital, Stockholm, Sweden 2 Center of Molecular Medicine, Stockholm, Sweden 3 Heidelberg University, Faculty of Medicine, and Heidelberg University Hospital, Institute for Computational Biomedicine, Bioquant, Heidelberg, Germany 4Universitätsklinikum Hamburg-Eppendorf, I. Medizinische Klinik, 20246, Hamburg, Germany.

Abstract: Inflammatory responses in the colon emerge at the mucosal interface, yet most microbiome measurements average spatially heterogeneous luminal and tissue-associated communities. We combined spatial metatranscriptomics with spatial host transcriptomics to map host–microbe organization along the mouse colon and to link bacterial signals to in situ inflammatory niches. In specific pathogen–free mice, microbial transcripts localized to stool pellets and resolved discrete community states, including a smaller Helicobacter-enriched state consistent with tissue association, whereas adjacent tissue harbored lower diversity than lumen. In gnotobiotic mice colonized with defined eubiotic or dysbiotic communities,

dysbiosis amplified mucosa-proximal belts of JAK-STAT activity and shifted immune microenvironments toward neutrophil enrichment. Within distal JAK-STAT hotspots, association and feature-selection analyses nominated Brachyspira, Eubacterium-G and Malacoplasma as tissue-associated taxa that track inflammatory niches beyond segment identity or luminal carryover. These taxa were detectable and disease-associated across human IBD multi-omics datasets, supporting cross-species relevance of spatially defined host–microbe niches.

Contact: yuhao.chi@ki.se

---

4. SHE: Structural History of Eukarya

Diandra Daumiller, Qiuzhen Li, Patrick Bryant

Abstract: Proteome-wide structural comparison of Eukarya for new evolutionary insights.

Contact: diandra.daumiller@scilifelab.se

---

5. Guiding biochemical experiments with UQSA

Andrei Kramer, KI, Sweden, Federica Milinanni, KTH, Sweden, Pierre Nyquist, Chalmers, Sweden, Jeanette Hellgren Kotaleski, KTH, Sweden, Alexandra Jauhiainen, Astra Zeneca, Sweden, Olivia Eriksson, KTH, Sweden

Abstract: UQSA (Uncertainty Quantification and Sensitivity Analysis) is an R-package for data-driven modeling of dynamical reaction networks. It has primarily been used within the field of biochemical signaling pathway modelling, especially within neuroscience. UQSA has introduced Bayesian parameter estimation and uncertainty quantification to a field which previously relied on uncertain point estimates. It provides confidence intervals to predictions and makes it possible to actually falsify models. UQSA has been used in collaboration with experimental groups e.g. to understand how the protein AKAP79 enables calcineurin to directly suppress protein kinase A activity. UQSA is adopted by EBRAINS – the European Infrastructure for Brain Research and implements efficient and automatic methods for model calibration, validation and analysis on computer clusters. With the latest implementation which included rewriting parts in C and developing the Peano-Bakes Series based algorithm the computational speed was increased two-three orders of magnitude. A prerequisite for UQSA to work is the structured model and data management approach that is used.

Contact: olivia@kth.se

6. Spatial Transcriptomic Differences of Colorectal Liver Metastasis Growth Patterns: Interferon Siganling and Inflamatory Niches

Escriva Conde Maria, Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, Solna, Sweden. Andersson Axel, Department of Diagnostics and Intervention, Surgery, Umeå University, Umeå, Sweden. Vermeulen Peter, Translational Cancer Research Unit, Ziekenhuis aan de Stroom, Campus Augustinus, Antwerp, Belgium. Nilsson Mats, Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, Solna, Sweden. Nyström Hanna, Department of Diagnostics and Intervention, Surgery, Umeå University, Umeå, Sweden.Wallenberg Centre for Molecular Medicine, Umeå University, Umeå, Sweden

Abstract: Introduction: Colorectal cancer liver metastases (CRCLM) display distinct histopathological growth patterns (HGPs), with the replacement (RHGP) and encapsulated (EHGP) patterns linked to markedly different prognoses. Despite their clinical relevance, the underlying microenvironmental dynamics shaping these growth patterns remain incompletely understood.

Materials and Methods: We applied in situ sequencing (ISS) to spatially resolve the transcriptomic landscape of CRCLM tissues, focusing on the tumor–liver interface in RHGP and EHGP lesions. Cellular populations were identified through unsupervised clustering and annotated based on spatial context and marker expression profiles.
All patients provided their informed written consent to participate in the tissue biobank of Region Västerbotten, Sweden.

Results: Our analysis revealed distinct populations of inflamed hepatocytes enriched in RHGP lesions, displaying reduced hepatocyte identity and signs of inflammatory reprogramming. These cells were predominantly located at the tumor-liver interface and may be co-opted by tumor-derived signals to promote progression. In contrast, EHGP lesions were characterized by the presence of well-differentiated, non-inflamed hepatocytes. EHGP lesions also showed spatial enrichment of interferon-stimulated gene expression in stromal and tumor cell populations, supporting immune cell recruitment and activation. Furthermore, fibrotic capsules in EHGP lesions exhibited a zonated structure, with the tumor-facing side influenced by tumor-derived signals and the liver-facing side displaying features of hepatic stellate cell activation and immune involvement. This supports a dynamic, immune–fibrotic niche at the tumor border that may restrict invasion.

Conclusions: Our findings reveal distinct microenvironmental programs associated with RHGP and EHGP in CRCLM, suggesting divergent mechanisms of tumor-host interaction with potential relevance for patient stratification and therapeutic intervention.

Contact: maria.escriva@scilifelab.se

---

7. Fused Gromov-Wasserstein optimal transport for single-cell and spatial transcriptomics data integration

Ferreira Pedro F, KTH Royal Institute of Technology, Sweden; Kviman Oskar, KTH Royal Institute of Technology, Sweden; Lagergren Jens, KTH Royal Institute of Technology, Sweden

Abstract: The spatial organization of evolving cell populations in tumors influences how they progress and respond to treatment. Spatially-resolved transcriptomics data may be used to examine the locations of these diverse cell states in tumor sections. However, these technologies often lack single-cell resolution and instead capture groups of cells in spatial spots, making it difficult to obtain accurate clonal identities and build phylogenetic trees. In contrast, full-length single-cell transcriptomics data are better suited to obtain cell lineages but do not capture their spatial locations in tissue. We present \toolname, an optimal transport and statistical model designed to quantify the concordance between the evolutionary tree and the spatial distribution of cancer cells by integrating single-cell and spatial transcriptomics data from tumors. SpOTr uses Fused Gromov-Wasserstein optimal transport to leverage the tree-based cell distances from one modality and spatial neighborhoods from the other to map single cells to spatial locations. The method learns clone-specific concordances with spatial structure, which can be used to investigate lineage-specific zonation patterns. We validate SpOTr on simulated and real spatial lineage tracing data from a mouse model of breast cancer lung metastasis.

Contact: pedro.ferreira@scilifelab.se

---

8. Unsupervised Ranking of Subtype-Specific Breast Cancer Driver Genes Using Mutual Information Regulatory Networks

Golnaz Taheri, Faezeh Fathi

Abstract: Breast cancer comprises molecular subtypes with distinct transcriptional regulation, but most driver-gene discovery methods analyze it as a single disease and may miss subtype-specific regulatorymechanisms. We propose an unsupervised, network-based framework to identify subtype-specific candidate driver genes from gene expression data. For each subtype, we infer a gene regulatory network using mutual information to capture non-linear dependencies, then derive topological descriptors and expression-derived biological features to represent each gene. We introduce a structure-preserving non-negative self-representation ranking model that exploits feature relationships to prioritize genes without requiring labeled drivers. We apply the framework to TCGA across four subtypes and show that top-ranked genes are enriched for known cancer drivers and subtype-relevant pathways, while also revealing novel subtype-specific candidates supported by survival association and pathway enrichment. These results highlight subtype-dependent regulatory programs and provide testable hypotheses for targeted therapeutic development.

Contact: faezehf@kth.se

---

9. Plasma Protein Biomarkers of Treatment-Induced Brain Injury in Children with Primary Brain Tumors: Early Findings from the BiomarkCBT Study

Dea Gogishvili, Martin Huss Knave, Anders Lundmark, Klara Lahne, Thomas Björk Eriksson, Magnus Borssén, Marie Eliasson-Hofvander, Kleopatra Georgantzi, Lars Karlsson, Per Nyman, Magnus Sabel, Mari Wilhelmsson, Anders Öberg, Jessica, Nordlund, Arja Harila, Christoffer Ehrstedt

Abstract: Introduction: Children treated for primary brain tumors face significant treatment-related side-effects that impact neurocognition and quality of life. Currently, no validated biomarkers exist for treatment-induced brain injury in this population, which limits our ability to identify children at risk, tailor treatments, and guide future interventions to minimise side-effects.

Aim: BiomarkCBT is a national study investigating plasma protein biomarkers of neurotoxicity in pediatric brain tumor patients.

Methods: The first set of samples from 42 patients (22 treatment, 20 controls) treated at Uppsala University Hospital or Skandion Clinic, Uppsala, were analysed using the ultrasensitive Simoa:registered: assay and the Olink:registered: HT Explore panel. Serial plasma samples of 22 patients were collected pre- and post-operatively, and at follow-up (1-7 per patient, 65 in total). Statistical analyses employed limma package in R with repeated measures, accounting for within-patient sample correlation. Additionally, Weighted Protein Co-expression Network Analysis was used to identify co-regulated protein modules and enriched biological processes.

Results: Despite the small sample size, we detected significant proteomic signatures associated with treatment. Protein co-expression analysis identified distinct protein modules correlating with treatment phases and tumor characteristics. Post-operative samples showed elevated immune and inflammatory responses compared to controls. Within the treatment group, post-operative samples showed increased levels of proteins involved in cell-cell adhesion, T-cell activation, and leukocyte migration compared to pre-operative samples. Surgical resection (versus biopsy) was associated with increased wound healing, leukocyte migration, and cell adhesion signatures. During radiotherapy, most significant proteins decreased, though this may reflect post-operative effects. Notably, EDA2R, a known prognostic marker in glioma associated with unfavourable outcomes, was elevated during radiation but decreased over time.

Discussion: Our preliminary data demonstrate detectable proteomic changes in pediatric brain tumor patients throughout treatment, even with limited sample sizes. The predominant signatures reflect immune cell activation and inflammatory responses, particularly following surgical intervention. These early findings support the feasibility of identifying plasma biomarkers for treatment-induced brain injury in children, though larger cohorts with longer follow-up are needed to establish clinically relevant endpoints and distinguish radiation-specific versus surgery-related effects.

Contact: dea.gogishvili@uu.se

---

10. Human in vivo genetic screening for novel CAR T cell design

Florian Albrecht, Laura M. Carroll, Johan Henriksson

Abstract: Problem: Less than 8% of cancer therapies, developed in animal models, are successful in human trials. This calls for a novel method to replace animal models with a more faithful alternative. This project aims to improve CAR T cells by developing the first human in vivo faithful genetic screen in cancer patients, to uncover novel genetic programming opportunities. Screening has been performed in mice, in vivo, using perturb-

seq - a state-of-the-art technology to simultaneously knock out a gene using CRISPR/Cas9 and measuring its impact on the transcriptome. By using single-cell technology and oligo arrays, the genome-wide screening has been performed in a single experiment. However, using this technology on humans would be unethical, as random KO's may harm the patient.

Solution: To overcome this problem, we will exploit that CAR T-cell production naturally causes off-target knockouts (KO) during production, when the CAR-containing virus randomly inserts into a gene. There is thus no need to specifically target genes using the CRISPR/Cas9 technology.

Our technology: This project is enabled by a new type of microfluidics that generate semi-permeable capsules (SPCs), giving increased flexibility to measure multiple aspects of the same cell. Our lab is one of the first labs in Europe to have access to the required equipment.

This project is still in the phase of developing the required new wetlab protocol. This includes (1) the inverse PCR to detect lentivirus insert site and (2) RNA-seq to analyze the associated transcriptome. The data will be analyzed using our novel single-cell computational pipeline, which is the first toolbox tailored for single-cell genome analysis.

Contact: johan.henriksson@umu.se

---

11. AI-assisted integration of H&E morphology and spatial transcriptomics reveals molecular programs of mitotic figures in human tissues

Mengping Long, Taobo Hu, Mats Nilsson; Stockholm University

Abstract: Mitotic figures are key histopathological indicators of cell proliferation and tumor aggressiveness, yet their molecular states and spatial context remain poorly characterized. Traditional assessment relies on visual inspection of H&E-stained sections, providing limited insight into the underlying transcriptional programs that drive mitosis in situ.

Here, we present an AI-assisted framework that integrates deep learning–based analysis of H&E images with high-resolution spatial transcriptomics to systematically characterize mitotic figures at single-cell resolution. Using a convolutional neural network trained to detect and classify mitotic figures from H&E images, we generated cell-centric morphological patches that were spatially aligned with transcriptomic profiles. This enabled direct linkage between mitotic morphology, local tissue architecture, and gene expression states.

Our integrated analysis reveals that mitotic figures are associated with distinct transcriptional programs beyond canonical cell-cycle markers, including differential activation of DNA damage response, chromatin remodeling, and metabolic pathways. Spatially, mitotic cells exhibit non-random distributions, preferentially localizing to specific microenvironmental niches with characteristic cellular compositions and signaling cues. Furthermore, AI-derived morphological features capture heterogeneity among mitotic figures that correlates with divergent molecular signatures, suggesting multiple mitotic states within the same tissue. Together, this work demonstrates that combining AI-driven histopathology with spatial transcriptomics enables a quantitative and molecularly

informed view of mitosis in human tissues. Our framework provides a scalable approach for studying proliferative dynamics in development and disease, and highlights the potential of multimodal integration for advancing spatial systems biology.

Contact: taobo.hu@scilifelab.se

12. Multiomic imprinting of tissue-resident memory T cells in gut and lymph node tissues

Jesse Mihkel, Department of Gene Technology, KTH Royal Institute of Technology, Science for Life Laboratory, Solna, Sweden; Bartolomé-Casado Raquel, Department of Molecular Oncology, Institute for Cancer Research, Oslo University Hospital-Radiumhospitalet, Oslo, Norway and Department of Pathology, Oslo University Hospital-Rikshospitalet, Oslo, Norway; Gu Yisu, Cambridge Stem Cell Institute and Department of Medicine, University of Cambridge, Cambridge, UK; Sarropoulos Ioannis, Cambridge Stem Cell Institute and Department of Medicine, University of Cambridge, Cambridge, UK; Dufva Olli, Cambridge Stem Cell Institute and Department of Medicine, University of Cambridge, Cambridge, UK; Teichmann Sarah, Cambridge Stem Cell Institute and Department of Medicine, University of Cambridge, Cambridge, UK and CIFAR MacMillan Multi-scale Human Programme, CIFAR, Toronto, Ontario, Canada; Koplev Simon, Department of Gene Technology, KTH Royal Institute of Technology, Science for Life Laboratory, Solna, Sweden.

Abstract: Tissue-resident memory T (Trm) cells are a non-circulating subset of T lymphocytes functioning as the first line of adaptive immune defense in epithelial barrier tissues. While their gastrointestinal roles in immune surveillance and tolerance are well established, the precursor cells, gene regulatory networks (GRNs), and tissue-specific ligand-receptor cues remain poorly characterised. Consequently, there is a lack of Trm tissue- and state-specific transcriptional markers and epigenetic signatures that may uncover lymph node imprinting in Trm cells. Here we present a multiomic analysis integrating single-cell gene expression and chromatin accessibility data from the gut and lymph nodes, including mesenteric lymph nodes and pancreaticoduodenal lymph nodes. The underlying workflow consists of a MultiQC approach for preprocessing and quality control, encompassing CellRanger, AMULET, and emptyDropletsMultiome, followed by scVI-based normalisation, cell-type annotation, pseudobulking and SCENIC+ for inferring GRNs and lymph node imprinting. We demonstrate variability of Trm subpopulation-specific epigenetic signatures between different gut regions and lymph nodes, highlighting potential for integrating single-cell gene expression and chromatin accessibility across immunological tissues and cell-types.

Contact: mihkel.jesse@gmail.com

---

13. The pan-disease blood protein profile of autoimmune diseases

Kenrick Josefin, Preger Charlotta, Bueno Álves María, Uhlén Mathias, Nilsson Peter, Edfors Fredrik, Pin Elisa

Abstract: Systemic autoimmune rheumatic diseases (SARDs) are a heterogeneous group of autoimmune conditions characterized by immune system dysregulation leading to chronic inflammation and tissue damage. The overlapping clinical manifestations make differential diagnosis challenging, highlighting the need for novel biomarkers to facilitate

early diagnosis, stratification, and personalized treatment. Five SARDs including idiopathic inflammatory myopathies (n=210), rheumatoid arthritis (n=84), systemic sclerosis (n=100), Sjögren disease (n=99), and systemic lupus erythematosus (n=99), as well as healthy controls (n=400) and controls with acute infectious diseases (n=218) were selected for plasma protein profiling using Olink Explore 1536. Proteins with known association to SARDs as well as novel associations were identified through differential abundance analysis and machine learning. This explorative cross-sectional study demonstrates the importance of a pan-disease approach to biomarker identification within and between the five SARDs.

Contact: josefin.kenrick@scilifelab.se

---

14. Integrated bulk and single-cell transcriptomics identify a high-risk, immune-enriched subset of Luminal A breast cancer

Tatjana Kiseļova, Linn Berglund, Olivia Blomberg1, Stephanie Mobedjina, Alice Åhlund, Arian Lundberg. Scilifelab, KTH Royal Institute of Technology, Sweden

Abstract: Background: Luminal A breast cancer is the most frequent molecular subtype, predominantly associated with favorable prognosis. However, a subset of the patients face recurrence and metastases. The immune landscape of Luminal A subtype remains largely unstudied.

Materials and Methods: We characterized the immune landscape of primary Luminal A breast cancers in Scan-B (n = 4047), and METABRIC (n = 691) cohorts using bulk RNA-seq data and ConsensusTME deconvolution to estimate tumor-infiltrating immune cell population. Tumors were stratified based on immune cell prevalence (immune score). Differential expression analysis was performed using limma, followed by pathway-level Gene Set Enrichment Analysis (GSEA). Survival analysis was performed using Kaplan-Meier curves. Immunotherapy response in immune-hot and immune-cold Luminal A tumors was assessed in I-SPY2 cohort (n = 185).

Results: Higher immune cell infiltration was consistently associated with distant relapse-free survival (p = 0.004). Differential expression analysis identified 1,342 upregulated and 114 downregulated genes in immune-active tumors. GSEA showed significant enrichment of inflammatory response and immune-related pathways. In the I-SPY2 cohort, patients with immune-hot Luminal A tumors exhibited significantly higher rates of pathologic complete response (p = 0.007). Cross-check with single-cell data from Breast Cancer Atlas has shown T-cells (namely CD8+ and CD4+) to be the predominant drivers of immune-active subtype.

Conclusion: Our findings highlight the need to further investigate the role of inflammation in a subset of Luminal A breast cancers traditionally considered immunologically quiescent. These patients may have a distinct risk profile for recurrence and metastatic progression whom could benefit from more tailored therapeutic strategies.

Contact: tatjana.kiselova@scilifelab.se

15. UQSA, an R package for Uncertainty Quantification and Sensitivity Analysis

Andrei Kramer-Miehe, Federica Milinanni, Alexandra Jauhiainen, Olivia Eriksson

Abstract: UQSA is an R package in development (github.com/icpm-kth/uqsa) with the capabilities to solve deterministic initial value problems (as ordinary differential equations, ODEs) or stochastic models using the Gillespie algorithm. The models we work with are from the fields of systems biology and neuroscience, and the package's functions are designed to specifically process reaction network models. Because the package is designed to be fairly modular, it is possible to use an entirely different kind of model, as long as it is of the same class (an ODE). The package can find the unknown parameters of a model using Markov chain Monte Carlo (MCMC) sampling, or approximate Bayesian computation (ABC), for stochastic models. It is also possible to combine the functions in the package and use them together with a standard optimization tool (in R). We use fast solvers, written in C, and offer several different MCMC algorithms.

Contact: andreikr@kth.se

---

16. Delineating genome instability dynamics in breast cancer by time-course bulk and single-cell copy number profiling

Sen Li1,2, Konstantinos L. Georgiadis1,2, Thomas Hatschek1,4, Theodoros Foukakis1,4 & Nicola Crosetto2,3,5 1 Department of Oncology-Pathology, Karolinska Institutet, Stockholm, SE-17165, Sweden 2 Science for Life Laboratory, Tomtebodavägen 23A, Solna, SE-17165, Sweden 3 Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, SE-17165, Sweden 4 Breast Center, Theme Cancer, Karolinska University Hospital, Stockholm, SE-17165 Sweden 5 Human Technopole, Viale Rita Levi-Montalcini 1, 22157 Milan, Italy

Abstract: Estrogen Receptor-positive (ER+) breast cancer represents the most prevalent breast cancer subtype, comprising approximately 70% of cases. Although copy number alterations (CNAs) are a hallmark of aggressive ER+ tumors and have been proposed as prognostic biomarkers, their temporal dynamics during therapy remain insufficiently characterized. In this study, we performed an integrated analysis of CNA evolution using bulk (n = 168 patients) and single-cell (n = 35,328 cells) DNA sequencing data from the PREDIX Luminal B neoadjuvant clinical trial (NCT02603679). Longitudinal profiling of tumors from two treatment arms (armA and armB) uncovered candidate genomic regions potentially linked to therapeutic response and resistance. These included arm-level gains and losses selectively enriched in post-treatment samples from non-responders. Integration with clinical outcomes further revealed treatment-specific CNA trajectories. Single-cell analyses demonstrated that most patients harbored limited subclonal complexity, and therapy predominantly reshaped the tumor landscape through selective expansion or contraction of pre-existing subclones rather than de novo subclone emergence. Our findings nominate recurrent CNAs as candidate predictive biomarkers of neoadjuvant response and suggest that clonal remodeling underlies inter-patient variability in treatment efficacy.

Contact: sen.li@ki.se

17. Predicting multiple sclerosis onset using EBV-related antibodies and machine learning methods in Swedish population

Authors: Yicong Liu(a, b), Neda Sattarnezhad(c, d), Olivia Thomas(a, e), Peggy P. Ho(f), Alison Barret(g), Alexandros Comanescu(g), Tilini Wijerante(g), Lars Alfredsson(a, h), Lawrence Steinman(f), William H. Robinson(c, i), Tomas Olsson(a, b), Tobias V. Lanz(c, g), Ingrid Kockum(a, b), Klementy Shchetynsky(a) Affiliations: a. Neuroimmunology Unit, Department of Clinical Neuroscience, Karolinska Institutet, 171 76 Stockholm, Sweden b. Center for Molecular Medicine, Karolinska University Hospital, 171 76 Stockholm, Sweden c. Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA d. Department of Neurology, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA 94304, USA e. Department of Clinical Neuroscience, Therapeutic Immune Design, Center for Molecular Medicine, Karolinska Institute, 171 77 Stockholm, Sweden f. Department of Neurology and Neurological Sciences, Beckman Center for Molecular Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA g. Institute for Immunity, Transplantation, and Infection, Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA h. Unit of Occupational Medicine, Institute of Environmental Medicine, Karolinska Institutet, 171 77 Stockholm, Sweden i. The Geriatric Research, Education, and Clinical Center (GRECC), Veterans Affairs Palo Alto Health Care System, Palo Alto, CA 94304 USA

Abstract: Background: Epstein-Barr virus (EBV) infection has long been implicated in the pathogenesis of multiple sclerosis (MS). Notably, the EBV nuclear antigen 1 (EBNA1) shares structural similarity with several human central nervous system (CNS) proteins, including ANO2, GlialCAM, and CRYAB. This molecular mimicry may trigger the production of cross-reactive antibodies that mistakenly target human antigens and myelin sheaths, contributing to MS development which also may be used for identifying individuals at risk of developing MS.

Aims and objectives: To assess the diagnostic value of EBV-related antibodies in MS, identify key biomarkers using statistical and machine learning methods, and develop predictive models for early risk detection.

Methods: To explore the diagnostic potential of EBV-related antibodies, we analyzed serum samples from 650 MS patients and 650 healthy controls. IgG reactivity against 33 candidate proteins was measured as median fluorescent intensity (MFI). Predictive models were built using antibody level and logistic regression, LASSO, and random forest with 10-fold cross-validation and bootstrap resampling. Feature selection was performed using significance testing and AIC-based stepwise selection. Model performances were evaluated using accuracy, Cohen's kappa, and area under the receiver operating characteristic curve (AUC).

Results: Initial models using all 33 antibodies achieved an accuracy of 0.65, kappa of 0.30, and AUC of 0.75. LASSO regression selected 21 antibodies and yielded comparable performance (accuracy = 0.65, kappa = 0.31, AUC = 0.74). The random forest model identified 7 stable predictors with slightly improved accuracy (0.66), kappa (0.33), and an AUC of 0.70. Stepwise AIC selection followed by scaling slightly improved the results (accuracy = 0.66, kappa = 0.31, AUC = 0.75). Cross-validation and bootstrap resampling confirmed the stability of these models. These findings suggest that while logistic regression offers some predictive value, complex patterns in antibody reactivity may require more flexible models for improved accuracy.

Conclusion and Future Directions: Antibody-based biomarkers derived from EBV-induced immune responses offer a promising, non-invasive approach for early MS prediction. To enhance model performance, we plan to select independent antibodies based on correlation analysis and explore advanced machine learning methods and deep learning approaches to capture nonlinear patterns. In addition, we aim to define diagnostic cut-off values for antibody titers to support clinical translation. Ultimately, our goal is to integrate antibody-based biomarkers with genetic variants and HLA haplotypes to enhance the predictive accuracy of MS risk models.

Contact: yicong.liu@ki.se

---

18. Computational comparative genomics with FastOMA and Read2Tree

Majidian Sina, Incoming DDLS Assistant Professor, Chalmers University of Technology

Abstract: Genome data keeps piling up, with efforts to sequence 1.5 million eukaryotic species, which could transform our understanding of evolution. However, this requires an overhaul of conventional comparative genomics methods limited to studying tens of genomes. Methods for inferring orthologous genes and phylogenetic relationships are computationally demanding. We have developed FastOMA and Read2Tree which tackle aforementioned challenges. FastOMA is a method for inferring orthology relationships combining k-mer-based placement, species-tree guided subsampling, and highly parallel computing to achieve near-linear performance in the number of input genomes. Read2Tree is a method for inferring phylogenetic trees directly from raw sequencing data bypassing genome assembly, gene annotation, and all-versus-all sequence comparisons. FastOMA and Read2Tree enable comparative genomics at scale.

Contact: sina.majidian@gmail.com

---

19. OrganoFeed: Where Biology Meets Engineering

Munoz Patricia, Dept. of Computer and Systems Sciences Stockholm University, Sweden , Milou Ioanna, Dept. of Computer and Systems Sciences Stockholm University, Sweden , Winkler Thomas, School of Electrical Engineering and Computer Science (EECS), KTH, Sweden

Abstract: Organoid technologies offer powerful models for studying human development and disease, yet their intrinsic variability remains a major limitation for reproducibility and scalability. The OrganoFeed Project aims to overcome this challenge by developing the first AI-driven microrobotic platform capable of monitoring and actively guiding the maturation of vascular organoids in real time. Our system integrates continuous metabolic sensing (glucose and lactate), brightfield imaging, and machine learning models to characterize growth dynamics and predict organoid-specific needs. Based on these predictions, a robotic module performs adaptive media exchange, adjusting feeding interval, volume, and composition for each individual organoid. By transforming organoid culture from passive observation to active, personalized control, OrganoFeed seeks to reduce variability, improve maturation quality, and enable high-throughput and standardized organoid production. This project provides a foundational step toward intelligent biofabrication and closed-loop automation in advanced in vitro systems.

Contacts: patricia.munoz-salvatierra@dsv.su.se

---

20. Characterizing transcriptional mechanisms of human tissue development in the Developmental Genotype-Tissue Expression (dGTEx) project pilot data

Winona Oliveros Diez1, Jose Miguel Ramirez2, Pau Clavell-Revelles2, the dGTEx Consortium, Marta Melé3 and Tuuli Lappalainen4. 1)Science for Life Laboratory, KTH, Solna, Sweden 2) Barcelona Supercomputing Center, Barcelona, Spain 3) Barcelona Supercomputing Center, Barcelona, Spain 4) SciLifeLab & NY Genome Center, Solna, Sweden

Abstract: Background: Cellular phenotypes and gene expression change dramatically from fetal life through adolescence before stabilizing in adulthood. Additionally, genetic and early environmental influences can have long-lasting effects on health. Characterizing transcriptional programs across human development is therefore essential to understanding the origins of many diseases.

Material and Methods: The dGTEx Project extends the GTEx framework to developmental stages in humans and two non-human primates. Here, we characterized the human pediatric transcriptional landscape from dGTEx pilot release of 34 donors and 342 samples across 24 tissues. Results: By quantifying shared gene expression changes across organs, we show that pathways developmentally downregulated with age are more widely shared across tissues than upregulated pathways, suggesting that a common developmental program is progressively silenced across the body. One third of developmentally regulated genes exhibit high genetic constraint across organs, with up- and downregulated genes displaying distinct patterns of selective constraint. On average, 15% of silenced genes during development are genetically associated with developmental diseases. Moreover, developmentally regulated genes that are shared across multiple organs are significantly more likely to be linked to developmental disease. Overall, we show that many disease-associated and drug target genes have dramatic expression changes during pediatric development.

Conclusion: Together, these results indicate that developmentally regulated transcriptional programs are under strong genetic constraint and enriched for developmental disease associations. Many of these genes are downregulated over time and are therefore largely missed in adult cohorts, underscoring the importance of studying gene expression during early human development to understand disease risk.

Contact: winona.oliveros@scilifelab.se

21. Developmental regulation of alternative splicing across human and non-human primate tissues

Philipp Rentzsch, Laura Domènech, Sanna Gudmundsson, Winona Oliveros, Fairlie Reese, Diego Garrido-Martín, Roderic Guigó, Marta Melé, François Aguet, Kristin Ardlie, Tuuli Lappalainen, and the dGTEx Consortium

Abstract: Alternative splicing is a major driver of transcriptomic diversity and is dynamically regulated across tissues and developmental stages. While adult splicing landscapes have been extensively characterized, the developmental trajectory of splicing regulation in humans and its evolutionary conservation across primates remains incompletely understood. Moreover, adult-derived transcript metrics such as proportion expressed across transcripts (pext), widely used for clinical variant interpretation, are limited in their ability to capture early developmental isoform usage.

Here, we leveraged the developmental Genotype-Tissue Expression (dGTEx) resource together with age-matched macaque and marmoset datasets (NHP-dGTEx), including prenatal stages, to characterize developmental splicing programs across humans and non-human primates.

Using splice junction-based cluster modeling across human tissues, we identified hundreds of age-differentially spliced genes (DSGs) per tissue, with the most striking changes observed in testis and muscle. While some of the apparent splicing events can be explained by alternative transcription start site usage, many lead to differences in the expressed coding DNA sequence of protein-coding genes. DSGs are enriched in tissue-relevant biological pathways, such as muscle contraction-related pathways in muscle and small intestine, and are generally overrepresented among genetically constrained genes, highlighting the functional and evolutionary importance of developmental splicing programs.

Orthogonal analyses of transcript isoform usage revealed coordinated isoform redistribution across developmental timepoints. These changes are concordant with shifts in exon-level expression (pext). Analysis of millions of clinically relevant sites (ClinVar variants) using pext revealed significant differences between dGTEx-derived and adult GTEx pext values, underscoring the importance of developmental context for variant interpretation. Cross-species comparisons further demonstrated that many developmental splicing patterns identified in humans are conserved in macaques and marmosets, while also revealing lineage-specific regulatory events.

Together, these results establish a cross-species framework for understanding developmental isoform regulation and its evolutionary dynamics across humans and non-human primates.

Contact: philipp.rentzsch@scilifelab.se

22. Improved biosignature filtering using molecular network structure embedding

Dirk Repsilber, Örebro University, Sweden

Abstract: We have shown that filtering biomarkers can profit from using a priori knowledge from molecular network structure. L-PLS methodology also uses such embedding for improved biosignature predictive or classification modelling. Now, we set out to use molecular network structure for annotation of high-content imaging data, based on these two previous approaches, to use target-gene annotated molecules in the JUMP-MOA dataset for training a toxicity mechanism predictor.

Contact: dirk.repsilber@oru.se

---

23. Modelling multi-modal responses to gene expression dosage in single-cell CRISPR screens with bayesDREAM

Leah U Rosen1, Jasper Panten1 and Tuuli Lappalainen1,2. (1Science for Life Laboratory, Department of Gene Technology, KTH Royal Institute of Technology, Solna, Sweden; 2New York Genome Center, New York, NY)

Abstract: Over 90% of human trait and disease-linked genetic variants have been found in the non-coding genome with many of these variants quantitatively affecting gene expression. Thus, in recent years, understanding how cells respond to changes in gene expression dosage has become a key question in human genetics. Pooled CRISPR screens with single-cell sequencing are emerging as a powerful tool for this, but prior studies have relied on analysis methods that do not fully leverage the single-cell resolution, detect different modalities, or properly account for biases and noisiness of the data.

Here we have developed a Bayesian method to model Dosage Response Effects Across Modalities (bayesDREAM). BayesDREAM leverages the inherent variability of single cell CRISPRi and CRISPRa screens to fit a dosage response function to various trans modalities based on cis gene expression. First we investigate how individual trans genes' expression responds to changes in cis gene dosage and show that cells balance homeostasis at wild type dosage with the ability to react to changes in gene expression. Next, we show that splicing of trans genes quantitatively responds to changes in cis transcription factor dosage, and explore several hypotheses to explain this observation.

BayesDREAM maintains power when applied to transcriptome-wide single cell CRISPR screens with a small number of guides per gene, thus enhancing the analysis of the CRISPR screens that are already being performed. This will have important implications beyond human genetics, such as for the fields of directed differentiation and regenerative medicine as well as in drug development.

Contact: leah.rosen@scilifelab.se

24. Imaging-Based Spatial Transcriptomics at SciLifeLab

Katarina Tiklova, Rapolas Spalinskas, Chika Yokota, SciLifeLab, Department of Biochemistry and Biophysics, Stockholm University, Sweden

Abstract: The In Situ Sequencing (ISS) unit is part of the Spatial Biology platform at SciLifeLab and provides imaging-based spatially resolved single-cell transcriptomics with subcellular resolution in intact tissue sections. The technology enables profiling of hundreds to thousands of RNA targets while preserving spatial context, supporting identification of cell types, cellular microenvironments, and tissue architecture. The ISS service supports projects addressing cell–cell communication, rare cell populations, and complex spatial interactions. Defined panels of proteins can be combined with RNA detection to enable integrated spatial RNA–protein analysis within the same tissue section. The unit offers support in experimental design, sample preparation, data generation, and guidance on data interpretation in close collaboration with users.

Contact: katarina.tiklova@scilifelab.se

---

25. Partial RY-mers increase aDNA alignment specificity

Ivan Tolstoganov, Stockholm University, Sweden; Kristoffer Sahlin, Stockholm University, Sweden

Abstract: One of the key steps in ancient DNA data analysis is the accurate identification of endogenous short reads that originate from the ancient genome. Currently, this step is performed using general-purpose sequence aligners, such as BWA-MEM, with very sensitive settings (i.e., a small seed size). This approach results in increased runtime, even relative to standard DNA alignment workflows, and can reduce mapping specificity.

We present Strobealign-adna, a tool for aDNA alignment that uses damage-aware seed constructs designed to incorporate models of aDNA degradation. We show that Strobealign-adna achieves higher alignment accuracy while also reducing runtime compared to state-of-the-art methods.

Contact: ivan.tolstoganov@math.su.se

---

26. MICA: A Variational Autoencoder Learns Cohort-Invariant Microbiome Signatures of Colorectal Cancer

Wang Xueyao , Saei Amir Ata (Karolinska Institituet, MTC department, Sweden)

Abstract: Colorectal cancer (CRC) has been linked to shifts in the gut microbiome, motivating microbiome based, noninvasive prediction models. However, microbiome cohorts vary substantially across geography, protocols, and sequencing. Models trained on one cohort often fail to generalize to others due to domain shift. We present MICA, a domain-adversarial variational autoencoder that learns a cohort-invariant latent representation from species/strain compositional taxonomic profiles.
MICA compresses microbiome profiles into a low dimensional representation while preserving information for downstream CRC prediction. To better reflect biological

connection among taxa, we represent the input using a phylogeny-aware structure derived from the microbial phylogenetic tree, enabling the model to learn patterns that respect evolutionary relationships rather than treating taxa as independent features. MICA is further trained to minimize cohort specific information in the learned representation, improving robustness to cohort eQects. We evaluate MICA on ~4000 stool metagenomic samples from 18 well controlled CRC cohorts, using nested leave-one-cohort-out cross-validation. Classifiers trained on latent representation achieve comparable CRC prediction performance to strong baselines trained on taxonomic features (including batch correction baselines such as ComBat followed by classification), while showing improved generalization by substantially reducing cohort specific signal (cohort leakage ≤0.40). Linking latent dimensions back to taxa highlights biologically coherent modules, including CRC associated invasive pathogens and health associated commensals. Together, MICA provides a practical framework for robust cross-cohort CRC prediction and interpretable cohort-invariant microbiome signatures across heterogeneous studies.

Contact: xueyao.wang@ki.se

---

28. Uncovering ligand binding determinants in GPCRs using interpretable rule-based model

Yalova Doga, Uppsala Uni, Sweden/Pulinkala Girish, Uppsala Uni, Sweden/ Diaz Alejandro, Uppsala Uni, Sweden/Komorowski Jan, Uppsala Uni, Sweden/Carlsson Jens, Uppsala Uni, Sweden

Abstract: Ligand binding plays a vital role in numerous biological processes, yet the determinants of binding specificity remain incompletely understood. Binding sites frequently lack strong sequence homology or shared motifs, indicating that binding is highly complex and context-dependent. This complexity poses significant challenges in identifying the factors that govern whether a particular ligand will interact with its receptor. In this project, we will investigate applying a predictive and interpretable rule-based model using R.ROSETTA to uncover the determinants of binding specificity for closely related GPCR families. The long-term impact of this work will advance receptor design research and help drive progress in gene therapy.

Contact: dogayalova@gmail.com

---

29. Beyond Chemical Similarity: Modality-Inclusive DDI Prediction with a Knowledge Graph

Yılmaz Ardan, School of Electrical Engineering and Computer Science, SciLifeLab, KTH Royal Institute of Technology, Stockholm, Sweden; Szydlik Szymon, Department of Computer and Systems Sciences, Stockholm University, Stockholm, Sweden; Taheri Golnaz, School of Electrical Engineering and Computer Science, SciLifeLab, KTH Royal Institute of Technology, Stockholm, Sweden*

Abstract: Polypharmacy increases the risk of harmful drug–drug interactions (DDIs), yet exhaustive experimental testing is infeasible. Many computational predictors rely on small-molecule encodings (e.g., SMILES) and chemical similarity, limiting coverage of biologics and obscuring mechanism; moreover, unlabeled pairs are often treated as

non-interacting, biasing models in a positive–unlabeled setting. We introduce a modality inclusive framework, the Drug–Target–Protein Knowledge Graph (DTP-KG), which captures pathway-level context and is combined with MeSH-derived features and a curated set of mechanism-supported negative DDIs. Using MeSH alone, we establish strong and transparent baselines and show that a \emph{Mid} knowledge regime (Low+Mid MeSH depths) yields the most stable performance across interaction categories. Augmenting MeSH with lightweight topological descriptors from 2-hop ego networks (e.g., degree/clustering and path-based proximity) and fusing them via latent-space gating yields statistically significant gains over MeSH-only in five of six categories, with limited headroom for Deep–Deep pairs. Naïve fusion (concatenation or global weighting) does not yield significant improvement. These results indicate that mechanistic context encoded in DTP-KG complements ontology-based features and establish a robust, modality-inclusive baseline for safety-critical DDI prediction spanning small molecules and biologics. The source code is available on GitHub at https://github.com/Golnazthr/DTP-KG

Contact: yilmazardan@gmail.com

---

30. AI-driven prediction of serial killing NK cells

Hanqing Zhang1, Shanshan Xu1, Mårten Björkman2, Björn Önfelt1 1, 1.Department of Applied Physics, KTH Royal Institute of Technology. 2. Division of Robotics, Perception and Learning, KTH Royal Institute of Technology

Abstract: Cellular immunotherapy relies on isolating and expanding immune cells with strong tumor-killing ability, but current selection methods often produce heterogeneous populations and rarely use functional measures such as cytotoxicity. We present an image-based machine learning (ML) approach to detect highly cytotoxic natural killer (NK) cells using data from a single-cell microchip platform in the Önfelt lab, which links imaging-derived behaviors to actual killing behavior. By combining our datasets with advanced ML models, we identify predictive NK cell features associated with effective killing activity. Our solution supports the selection of superior NK cells for therapy and offers a broadly applicable framework for discovering functional immune cell subsets across different cell types.

Contact: hanzha@kth.se