Ales Varabyou

Johns Hopkins University
3100 Wyman Park Drive, Room S216
Baltimore, MD 21211, USA
email: ales.varabyou@jhmi.edu
Website: https://ccb.jhu.edu/people/ales.varabyou
Languages: Belarusian, English, Polish, Russian

Education

2017 – 2023	Ph.D. (Genomics, CS) **Johns Hopkins University** **Advisors**: Dr. Steven Salzberg & Dr. Mihaela Pertea **Thesis**: Computational Study of Transcriptional landscapes from RNA-seq Data
2020	M.Sc.Eng. (Computer Science) **Johns Hopkins University** Advisors: Dr. Steven Salzberg & Dr. Mihaela Pertea
2013-2017	B.A. *cum laude* (Biology & Computer Science) **Luther College** Honors Thesis: *Distributed Adaptive Object Recognition and Tracking* Thesis advisor: Dr. Bradley Miller

Academic Positions

2023-present	Research Scientist at Pertea Lab Developing methods for analyzing translatome expression using RNA-seq data. Building cell-type-specific genome annotations from large RNA-seq datasets. **Department of Biomedical Engineering, Johns Hopkins University**
2023	Supervisor for an undergraduate project (CS). Topic: "Designing a Dynamic and Scalable Web Interface for Genome Annotations"
2023	Supervisor for an undergraduate project (BME). Topic: "Impact of Transcript Clustering on Downstream Expression Analysis"
2020	Teaching Assistant
2017	Research Assistant (Genomics) Salzberg Lab, Institute of Genetic Medicine, **Johns Hopkins School of Medicine**
2015-2017	Research Assistant (Evolutionary Biology) **Luther College**
2016	Teaching Assistant in Human Physiology **Luther College**

Presentations And Talks

Varabyou A. (2023) Topics in Genomics. *Medical School Intensive. Guest Speaker*

Varabyou A. (2023) Precision Gene Catalogs: Segregating Noise and True Signals in RNA-seq Data for Comprehensive Genome Annotation. *Advanced Biomedical Computation (ABC), Seminar Speaker*

Varabyou A. (2022) Investigating Open Reading Frames in reference and novel transcripts using ORFanage. *Banbury Human Gene Annotation Meeting, Speaker*

Varabyou A., Pertea M., Salzberg S. (2019) Annotation-guided alignment of short RNA-seq reads for de-novo transcriptome assembly. Genome Informatics, Cold Spring Harbor, Poster

First Author Publications

Varabyou A., Artamonov M., Bolton D. (Upcoming: 2025). Full-length transcriptome analysis of clonal expansion and proviral integration during acute SIV infection at single-cell resolution. *bioRxiv*

Varabyou A., Artamonov M., Bolton D., Salzberg S., Pertea M. (2025). Comprehensive atlas of HIV-1 and SIV transcriptional diversity. *bioRxiv, 2025-09*

Varabyou A., Erdogdu B., Salzberg S., Pertea M. (2023). Investigating open reading frames in known and novel transcripts using ORFanage. *Nature Computational Science, 10.1038/s43588-023-00496-1*

Varabyou, A., Sommer, M. J., Erdogdu, B., Shinder, I., Minkin, I., ... & Pertea, M. (2023). CHESS 3: an improved, comprehensive catalog of human genes and transcripts based on large-scale expression data, phylogenetic analysis, and protein structure. *Genome biology*, *24*(1), 249.

Varabyou A., Pockrandt C., Salzberg S., Pertea M. (2021). Rapid detection of inter-clade recombination in SARS-CoV-2 with Bolotie, *Genetics, 10.1093/genetics/iyab074*

Varabyou, A., Pertea, G., Pockrandt, C., Pertea, M. (2021). TieBrush: an efficient method for aggregating and summarizing mapped reads across large datasets. *Bioinformatics, 10.1093/bioinformatics/btab342*

Varabyou, A., Salzberg, S. L., & Pertea, M. (2021). Effects of transcriptional noise on estimates of gene and transcript expression in RNA sequencing experiments. *Genome Research*, *31*(2), 301-308.

Liu, R., Yeh, Y. H. J., **Varabyou, A.**, Collora, J. A., Sherrill-Mix, S., Talbot, C. C., ... & Pollack, R. A. (2020). Single-cell transcriptional landscapes reveal HIV-1–driven aberrant host gene transcription as a potential therapeutic target. *Science Translational Medicine*, *12*(543).

Contributing Author Publications

Olivera K., Wei Y., Varabyou A., Ho Y. C. (Upcoming May 2025). Spatial Mapping of SIV-infected Rhesus Macaque Lymph Node Environment During Rebound. *bioRxiv*

Erdogdu, B., **Varabyou, A.**, Hicks, S. C., Salzberg, S. L., & Pertea, M. (2024). Detecting differential transcript usage in complex diseases with SPIT. *Cell Reports Methods*, *4*(3).

Amaral P., Carbonell-Sala S., De La Vega F. M., Faial T., Frankish A., Gingeras T., Guigo R., Harrow J. L., Hatzigeorgiou A. G., Johnson R., Murphy T. D., Pertea M., Pruitt K. D., Pujar S., Takahashi H., Ulitsky I., **Varabyou A.**, Wells C. A., Yandell M., Carninci P., Salzberg S. (2023). The status of the human gene catalogue. *Nature*, *622*(7981), 41-47.

Sommer, M. J, Cha, S., **Varabyou, A.**, Rincon, N., Park, S., Minkin, I., Pertea, M., Steinegger, M., Salzberg, S. (2022). Structure-guided isoform identification for the human transcriptome. *Elife 11*, e82556.

Shifera, A. S, Pockrandt, C., Rincon, N., Ge, Y., Lu, J., **Varabyou, A.**, Jedlicka, A. E., Sun, K., Scott, A. L., Eberhart, C., Thorne, J. E., Salzberg, S.

L. (2021). Identification of microbial agents in tissue specimens of ocular and periocular sarcoidosis using a metagenomics approach. F1000Research, 10(820).

Pertea, M., Shumate, A., Pertea, G., **Varabyou, A.**, Breitwieser, F.P., Chang, Y.C., Madugundu, A.K., Pandey, A. and Salzberg, S.L., (2018). CHESS: a new human gene catalog curated from thousands of large-scale RNA sequencing experiments reveals extensive transcriptional noise. Genome biology, 19(1), pp.1-14.

Leadership and Volunteering

2022-2023	JHU Transportation advisory team. Designed and implemented a real-time tracking system for shuttle service delay reporting and accountability.
2022-2023	GRO Chair for Health and Wellness. Worked with other chairs and student organizations across campuses to create community events and help raise awareness of health and wellness resources on campus. Assisted individual students.
2021-2023	Volunteer as a panelist for academic misconduct hearings. Served on multiple academic hearings alongside faculty and members of school administration. Helped refine school policies and make decisions regarding cases of academic misconduct.
2022	Provost advisory team on healthcare. Served as a student representative of Homewood schools. Helped voice concerns regarding health insurance and advise on upcoming changes.
2022	Translated cybersecurity texts for Ukraine and Belarus during the onset of Russian invasion.
2020-2021	Aided refugees from Belarus. Held bi-weekly online sessions teaching spoken English and relevant computer science concepts. Provided guidance on immigration.
2020	Volunteer for the Emergency COVID-19 JHU hospital supplies team. Coordinated supply delivery and allocation between hospitals. Tracked employees and their shifts. Loaded and delivered supplies within Baltimore City.
2012-2018	Volunteer reviewer and selection committee member for the Belarus National UWC Committee.
2010-2014	Teaching assistant for Belarus Lyceum of Humanities. Supervised high school students on study abroad programs. Assisted teaching Computer Science, English, and Sciences.

Other Positions

2024-2025	Served as an invited external committee member for the MHRP hiring of PhD-level computational biology scientist
2018-2020	Volunteered consulting services to develop variant calling protocol at Aevus diagnostics.
2017-2018	Contributions and management of the MUMmer 4 suite on GitHub (~2,000 citations)
	Reviewer for ISMB/ECCB 2021-2024, Nature, Molecular Biology and Evolution, NAR, OUP Bioinformatics, Genome Biology, Computational and Structural Biotechnology Journal, Springer Nature BMC

2015 - 2017   Workstation Administrator Assistant
Luther College
Supervisor: Matthew Hammen

2014 - 2016   Founder & President of the Luther College Robotics Research Club
Luther College
Academic Advisors: Dr. David Ranum, Dr. Bradley Miller & Dr. Kent Lee

Properties

1. Resilience – very frequently working on exploratory tasks requiring a lot of …
2. Comfortable exploring novel technologies. Very frequently working on topics and techniques previously unexplored by colleagues and others in the field. Help find novel approaches and techniques for my field, such as building complex visualization tools for public use, simplifying deployment of methods to the users with webassembly, building safe and efficient applications in rust and leading efforts to represent transcriptional diversity datasets in more robust sql interfaces addressing challenges of more simple GTF/GFF formats.