ESZTER ARI

Eötvös Lorand University,
Faculty of Science, Institute of Biology

Address: Pázmány Péter sétány 1/c.,
H-1117 Budapest, Hungary

RESEARCH AREA

My research centers on systems-biology questions that can be addressed using bioinformatics methodologies grounded in evolutionary genomics. This includes studying and understanding the evolution of pathogenic bacteria and the emergence and spread of antibiotic resistance and virulence. By analyzing the complete genetic repertoires (genomes) of Escherichia coli and other bacteria, I investigate how specific genetic elements and environmental factors influence evolutionary processes. Using phylogenetic methods, I examine the role of horizontal gene transfer in the dissemination of genes, and from a systems-biology perspective I study how particular pathogens spread across time and space.

Today, biological research has access to an astonishing amount of publicly available data—on the order of petabytes—such as sequenced genomes. However, these data are often not organized in ways that allow them to be readily converted into biological knowledge. My aim is to use bioinformatics tools to derive biologically interpretable insights from large datasets ("big data") and to build databases from them that are accessible to everyone. One such database is TFLink, which focuses on an aspect of gene regulation: transcription factors and their target genes. Yet to reveal and understand gene-expression differences across specific organs and tissue types, it is essential to integrate new, cell-specific sequencing data. To organize these, we also enlist large language models—that is, artificial intelligence. Using a similar approach, we aim to systematize sampling data for hundreds of thousands of bacteria as well, thereby creating a unique opportunity for efficient, big-data-driven epidemiological research.

TECHNIQUES AVAILABLE IN THE LAB

Fundamental bioinformatics methods, phylogenetic and comparative genomics approaches, the Linux command line, the R and Python programming languages, statistical procedures, database use and design, and the application of machine learning and large language models (LLMs).

SELECTED PUBLICATIONS

Ari, E., Vásárhelyi, B. M., Kemenesi, G., Tóth, G. E., Zana, B., Somogyi, B., Lanszki, Z., Röst, G., Jakab, F., Papp, B., & Kintses, B. (2022). A single early introduction governed viral diversity in the second wave of SARS-CoV-2 epidemic in Hungary. **Virus Evol 8(2):** veac069.

Gerber, D., Szeifert, B., Székely, O., Egyed, B., Gyuris, B., Giblin, J. I., Horváth, A., Köhler, K., Kulcsár, G., Kustár, Á., Major, I., Molnár, M., Palcsu, L., Szeverényi, V., Fábián, S., Mende, B. G., Bondár, M., **Ari, E.**, Kiss, V., & Szécsényi-Nagy, A. (2023). Interdisciplinary Analyses of Bronze Age Communities from Western Hungary Reveal Complex Population Histories. **Mol Biol Evol 40(9):** msad182.

Koncz, M., Stirling, T., Hadj Mehdi, H., Méhi, O., Eszenyi, B., Asbóth, A., Apjok, G., Tóth, Á., Orosz, L., Vásárhelyi, B. M., **Ari, E.**, Daruka, L., Polgár, T. F., Schneider, G., Zalokh, S. A., Számel, M., Fekete, G., Bohár, B., Nagy Varga, K., Visnyovszki, Á., … Kintses, B. (2024). Genomic surveillance as a scalable framework for precision phage therapy against antibiotic-resistant pathogens. **Cell 187(21):** 5901–5918.e28.

Liska, O., Bohár, B., Hidas, A., Korcsmáros, T., Papp, B., Fazekas, D., & **Ari, E.** (2022). TFLink: an integrated gateway to access transcription factor-target gene interactions for multiple species. **Database (Oxf.) 2022,** baac083.

Turek, C., Ölbei, M., Stirling, T., Fekete, G., Tasnádi, E., Gul, L., Bohár, B., Papp, B., Jurkowski, W., & **Ari, E.** (2024). Mulea: An R package for enrichment analysis using multiple ontologies and empirical false discovery rate. **BMC Bioinform 25(1):** 334.