

ARI ESZTER



Eötvös Loránd Tudományegyetem,
Természettudományi Kar, Biológiai Intézet

Cím: 1117 Budapest, Pázmány Péter sétány 1/c.

KUTATÁSI TERÜLET BEMUTATÁSA

Kutatásaim középpontjában az evolúciós genomikai megközelítéseket alkalmazó bioinformatikai módszertannal megválaszolható rendszerbiológiai kérdések állnak. Ilyen terület a kórokozó baktériumok evolúciójának, az antibiotikum-rezisztencia és a virulencia kialakulása és terjedésének vizsgálata és megértése. Az *Escherichia coli* és más baktériumok teljes genetikai állományának (genomjának) elemzésével vizsgálom, hogy miként hatnak az egyes genetikai elemek vagy a környezet az evolúciós folyamatokra. Filogenetikai módszertannal kutatom a horizontális géntranszfer szerepét a gének terjedésében, illetve rendszerbiológiai szemlélettel vizsgálom az egyes kórokozók terjedését időben és térben.

A biológiai kutatásokhoz ma már elképesztő mennyiségű (petabájtos nagyságrendű), bárki számára hozzáférhető adathalmaz (pl. megszekvenált genomok) áll a rendelkezésünkre. Azonban ezek az adatok sokszor nem rendezettek megfelelően ahhoz, hogy könnyen biológiai tudássá konvertáljuk őket. Célom, hogy bioinformatikai eszközök segítségével nagy adathalmazokból ("big data") biológiailag értelmezhető következtetéseket vonjunk le, ezekből mindenki számára hozzáférhető adatbázisokat is létrehozak. Ilyen adatbázis a TFLink, ahol a génszabályozás egyik egysége, a transzkripció faktorok és célgénjeik találhatók meg. Azonban ahhoz, az egyes szerv-, szövettípusokban végbemenő génexpressziós különbségeket feltárjuk és megértjük, elengedhetetlen új, sejtspecifikus szekvenálási adatok integrálása. Ezek rendszerezéséhez a nagy nyelvi modelleket, vagyis a mesterséges intelligenciát is segítségül hívjuk. De hasonló eljárással kívánjuk több százezer baktérium mintavételi adatait is rendszerezni, ezzel egyedülálló lehetőséget biztosítva a hatékony, big data epidemiológiai kutatásokhoz.

ELSAJÁTÍTHATÓ TECHNIKÁK

Alapvető bioinformatikai módszerek, filogenetikai és összehasonlító genomikai módszerek, Linux parancssor, R statisztikai programnyelv, Python programnyelv, statisztikai eljárások, adatbázisok használata és felépítése, gépi tanulás és nagy nyelvi modellek használata.

VÁLOGATOTT KÖZLEMÉNYEK

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., Fábán, 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.