

# Szennyezett genomi adatok tisztítására alkalmas szoftvert fejlesztettek a HUN-REN SZBK kutatói

A *Nature Communications* folyóiratban jelent meg a Nagy László által vezetett, a Joint Genome Institute (JGI, USA) és a National Energy Research Scientific Computing Center (NERSC, USA) intézetekkel közös nemzetközi együttműködés új [közleménye](#), mely a szekvencia adatbázisokba tévesen feltöltött, idegen szekvenciákkal szennyezett genom adatok lehetséges következményeivel és a szennyezések kitisztításával foglalkozik.

A közzétett munka fő részeként, a szegedi kutatók elkészítettek egy teljesen új számítógépes szoftvert ([ContScout](#)), amely a korábban más csoportok által közzétett megoldásokhoz képest lényegesen pontosabban azonosítja be és távolítja el a genomi adatokban megbúvó szennyeződések, miközben a horizontálisan géntranszfer révén szerzett géneket érintetlenül hagyja.

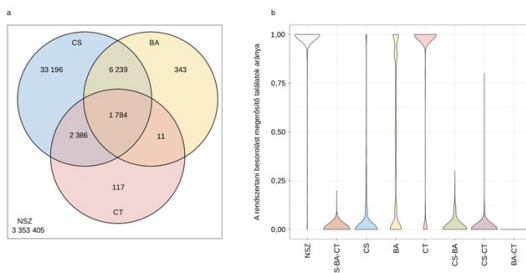
A kutatók a 844 megvizsgált eukarióta genom több mint felében mutattak ki bizonyos mértékű szennyeződést, melyek döntő többsége baktériumokból származott. A szennyeződés néhány esetben akkora mértéket öltött, a közzétett genom adatokból egy második élőlény közel teljes genomja rekonstruálható volt.

A közzétett kézirat részletes példákon keresztül illusztrálja, milyen torzító hatással vannak a szennyező szekvenciák azokra a kísérletekre, amelyek az egyedi géncsaládok, valamint az élőlények evolúciós történetét kívánják feltárni.

Az együttműködésben résztvevő kutatók remélik, hogy a rangos folyóiratban közzétett eredményeikkel sikerül felhívniuk a tudományos közösség figyelmét a genomi szennyeződések eltávolításának fontosságára, valamint, hogy a kifejlesztett szekvencia tisztító szoftver az általánosan elfogadott genomi elemzési munkafolyamatok részévé válhat.

Sajtókapcsolat:

- Hencz Éva, kommunikációs igazgató
- +36 30 155 1803
- [media@hun-ren.hu](mailto:media@hun-ren.hu)



© HUN-REN SZBK

A ContScout program összevetése más kutatócsoportok által korábban közzétett szoftverekkel. A Venn diagram (a) és a violin plot (b) a legszennyezettebb 200 genom adatait tartalmazza. Rövidítések: NSZ nem szennyezés. CS ContScout által azonosított, BA: BASTA által azonosított, CT: Conterminator által azonosított szennyező fehérje.

Eredeti tartalom: HUN-REN Magyar Kutatási Hálózat

Továbbította: Helló Sajtó! Üzleti Sajtószolgálat

Ez a sajtóközlemény a következő linken érhető el:

<https://hellosajto.hu/11233/szenyezett-genomi-adatok-tisztitasara-alkalmas-szoftvert-fejlesztettek-a-hun-ren-szbk-kutatoj/>