

A MAGYAR
TUDOMÁNY
ÜNNEPE



Magyar Tudományos
Akadémia

Bioinformatika 2021

A Magyar Tudományos Akadémia Bioinformatikai Osztályközi Tudományos Bizottsága
és a Magyar Bioinformatikai Társaság tudományos konferenciája

2021. november 12.
Zoom videó-konferencia

Program

1. szekció, Elnök: Pongor Sándor

10:00 – 10:05: **Köszöntő (Pongor Sándor)**

10:05 – 10:25: **Mészáros Bálint**

AlphaFold és fehérje rendezetlenség: a szerkezet becslés új generációja
Structural and Computational Biology Unit, EMBL

10:25 – 10:45: **Csikász-Nagy Attila**

Vírusterjedés egy virtuális városban
PPKE ITBK

10:45 – 11:05: **Gyórfy Balázs**

Új terápiás célpontok azonosítása onkológiában nagy adatbázisok
használatával
SE Bioinformatika Tanszék, TTK Enzimológia Intézet

11:05 – 11:15: **Szünet**

2. szekció – Bioinformatika TDK különdíjat nyert hallgatók előadásai, Elnök: Tusnady Gábor

11:15 – 11:30: **Kis Márton**

Neurális reprezentációk vizsgálata felügyelet nélküli gépi-tanulással
KOKI Biológia számítások kutatócsoport

11:30 – 11:45: **Horváth Márton**

Fajspecifikus elemek felderítése a Candida fajokkal szemben kialakuló
természetes immunválaszban humán szájüregi epitelsejtek esetén
SzTE Mikrobiológia Tanszék

3. szekció – Fiatal Kutatók előadásai, Elnök: Ari Eszter

11:45 – 12:00: **Csabai Luca**

Signalink3: többretegű erőforrás a szövetspecifikus jelátviteli hálózatok
feltárására
ELTE Genetikai Intézet

12:00 – 12:15: **Asbóth András**

A bakteriofág farokrostok kompatibilitásának és diverzitásának feltárása a
Podoviridae és Autographiviridae családokban
SZBK Biokémiai Intézet Szintetikus és Rendszerbiológiai Egység

12:15 – 12:30: **Fekete Zsófia**

Nyúl (*Oryctolagus cuniculus*) cirkuláris RNS-ek azonosítása és jellemzése
betegség modellek létrehozási lehetőségeinek felderítésére
Magyar Agrár- és Élettudományi Egyetem

12:30 – 12:45: **Nagy Gergely**

Az ETS fehérjék DNS-kötésének vizsgálata egér csontvelői makrofágokban
DE ÁOK Biokémiai és Molekuláris Biológiai Intézet

12:45 – 13:00: **Sasvári Péter**

Az ARHGAP25 interakciós hálózatának feltárása humán neutrofil
granulocitákban
SE ÁOK Élettani Intézet

13:00 – 13:30: **Szünet**

4. szekció – Fiatal Kutatók előadásai, Elnök: Gáspári Zoltán

13:30 – 13:45: **Zsidó Balázs Zoltán**

Ligandum dokkolás explicit vizekkel – egy megoldás virális ioncsatornákon
tesztelve
PTE ÁOK Farmakológiai és Farmakoterápiai Intézet Farmakoinformatikai Tanszék

13:45 – 14:00: **Seitz Erik**

Az epheliális-mezenchimális átmenet (EMT) új, többsejtes hálózatos
modelljének kifejlesztése és új, potenciális gyógyszer-célpont megtalálása a
hibrid-EMT-nek a modellbe építésével
SE ÁOK Molekuláris Biológiai Tanszék

14:00 – 14:15: **Koncz Balázs**

A T sejtek pozitív szelekciója vakfoltot eredményez az adaptív
immunfelismerésben
SzTE SZAOK Szent-Györgyi Albert Klinikai Központ Bőrgyógyászati és Allergológiai Klinika

14:15 – 14:30: **Balogh Gergő Mihály**

Láthatóvá tenni a láthatatlant: az Apobec3 mutagenézis hatása a Sars-
Cov-2 adaptív immunfelismerésére
SzTE ÁOK Szent-Györgyi Albert Klinikai Központ Bőrgyógyászati és Allergológiai Klinika

14:30 – 14:45: **Kálmán Zsófia**

PSINDB: új, átfogó és részletes poszt-szinaptikus interakciós adatbázis
PPKE ITBK

14:45 – 15:00: **Miski Marcell**

A poszt-szinaptikus komplexom változékonyságának modellezése a
fehérjegyakoriság függvényében
PPKE ITBK

15:00 – 15:15: **Bohár Balázs**

Sherlock: egy Big Data technológián alapuló adatgyűjtést és elemzést
elősegítő platform
ELTE TTK Genetikai Tanszék

15:15 – 15:30: **Stirling Tamás**

Webchem: Egy R csomag online kémiai adatbázisok elérésére

SzBK Biokémiai Intézet Szintetikus és Rendszerbiológiai Egység

15:30 – 15:40: **Szünet**

5. szekció, Elnök: Hegedűs Tamás

15:40 – 16:00: **Horváth Péter**

Élet a pixelek mögött: mesterséges intelligencia a rákkutatásban és a biológiában

SzBK Biokémiai Intézet Szintetikus és Rendszerbiológiai Egység

16:00 – 16:20: **Kun Ádám**

RNS rátermettségtájképek

ELTE TTK Növényrendszertani, Ökológiai és Elméleti Biológiai Tanszék

16:20 – 16:25: **Zárszó (Pongor Sándor)**

Előadások kivonatai

Vírusterjedés egy virtuális városban

Csikász-Nagy Attila

PPKE ITBK

Pandémia kezelésére különböző intézkedéseket vezethetnek be, mint például a társadalmi távolságtartás, tesztelés/karantén, vakcinázás. A COVID-19 (SARS-CoV-2) világjárványt számos ilyen, u.n. nem-gyógyszeres beavatkozással enyhítik. Nehéz megjósolni, hogy ezek közül a rendelkezések közül melyik a leghatékonyabb egy adott populáció esetében. Egy számítási szempontból hatékony és skálázható, ágens-alapú mikroszimulációs keretrendszert dolgoztunk ki, amely lehetővé teszi, különböző vírusváltozatok terjedése által okozott fertőzési hullámok szimulációját egy város méretű társadalmi környezetben. Modellünkkel megmutattuk, hogy a foglalkozási kockázati csoportokat előtérbe helyező vakcinázási stratégiák minimalizálják a fertőzések számát, de magasabb halálozást tesznek lehetővé, míg a veszélyeztetett csoportok előtérbe helyezése minimalizálja a halálozást, de megnövekedett fertőzési rátával jár. Azt is megállapítottuk, hogy az intenzív vakcinázás a nem gyógyszeres beavatkozásokkal együtt jelentősen visszaszoríthatja a vírus terjedését, míg az alacsony szintű vakcinázás, a korai újraindítása a társadalomnak könnyen visszafordíthatja a járványt a kontrollálatlan állapotba. Elemzésünk rávilágít arra, hogy míg a védőoltás megvédi az időseket a COVID-19-től, a gyermekek nagy százaléka megfertőződhet a vírussal. Ezekon felül, bemutatjuk a különböző karanténos és tesztelési forgatókönyvek előnyeit és korlátait is.

Új terápiás célpontok azonosítása onkológiában nagy adatbázisok használatával

Gyórfy Balázs^{1,2}

¹ TTK Enzimológiai Intézet Onkológiai Biomarker Kutatócsoport

² SE Bioinformatika Tanszék

Onkológiai kezelések során mutációs és génexpressziós változások szolgálhatnak biomarkerként vagy terápiás célpontként. Ezek összekapcsolása egymással illetve nagy adatbázisokkal teszi lehetővé az egyes alcsoportokra vonatkozó elemzések végzését. Az előadás során egy fejlesztés alatt álló rendszert mutatunk be, amely különböző tumortípusok esetén képes egy "multi-omikai" elemzés végzésére.

Élet a pixelek mögött: mesterséges intelligencia a rákkutatásban és a biológiában

Horváth Péter^{1,2}

¹ SzBK Biokémiai Intézet Szintetikus és Rendszerbiológiai Egység

² *Institute for Molecular Medicine, University of Helsinki, Finland,*

Előadásomban áttekintést adok a nagyléptékű fénymikroszkópos kísérletek egysejt szintű feldolgozásának számítástechnikai lépéseiről. Először egy új mikroszkópos képkorrektációs eljárást mutatok be, mely kijavítja a nem egyenletes megvilágításból származó képi hibákat, így támogatva a fényintenzitás alapú mérések helyességét. Ezután új, differenciál geometriára, energia-minimalizációs módszerekre és mesterséges intelligenciára alapuló egysejt szegmentálási módszereket ismertetek. Bemutatom az Advanced Cell Classifier (ACC) (www.cellclassifier.org) gépi tanulási szoftvert, melyet azért fejlesztettünk, hogy a képi jellemzőkből származó információ felhasználásával sejtes fenotípusokat azonosítsunk. Az ACC egy olyan interaktív felületet biztosít, mely segítségével a felhasználók hatékonyan képesek intelligens algoritmusokat sejtek automatikus fenotipizálására tanítani. Azon esetekre, ahol nem lehetséges diszkrét fenotípus kategóriák létrehozása, bemutattunk egy multi-parametrikus regresszió alapuló, eljárást, mely képes biológiai folyamatok elemzésére. A tanulási gyorsaság és a pontosság növelése érdekében egy olyan aktív tanulási sémát dolgoztunk ki, amely kiválasztja a legtöbb információval rendelkező sejt mintákat. A fejlesztett módszerek kombinációit felhasználva különböző egysejt kinyerési stratégiákat mutattunk be. Ismertetem a frissen elért sikeres eredményeinket egysejt DNS és RNS szekvenálás és célzott elektro-fiziológiai elemzések területén.

RNS rátermettségtájképek

Kun Ádám

ELTE TTK Növényrendszertani, Ökológiai és Elméleti Biológiai Tanszék

A genotípus – fenotípus leképezések fontosságát korán észrevették a biológiában, de igazán jó rátermettségtájképek nem igazán ismertek. A modellekben alkalmazottak jelentős része teljesen mesterséges. Egy kivétel talán akad, az RNS szekvenciák (genotípus) és másodlagos szerkezetüket (fenotípus) összekapcsoló tájképek. A genotípus–fenotípus leképezés ebben az esetben az RNS feltekeredésének kísérletesen kimért energetikáján alapul. A valós rátermettséget, azaz egy evolúciós egység növekedési rátájához való hozzájárulást ebből még nem tudjuk meg, de egyre több kísérlet kapcsolja össze a szekvenciát a funkcióval. Előadásomban a genotípus – fenotípus – funkció tájképek alkalmazását mutatom be az élet keletkezésével kapcsolatos kérdések vizsgálatában.

AlphaFold és fehérje rendezetlenség: a szerkezet becslés új generációja

Mészáros Bálint

Structural and Computational Biology Unit, EMBL

A szerkezeti biokémia egyik legnagyobb kihívása a fehérjék térszerkezetének becslése az aminosav sorrend alapján. Ez az úgy nevezett folding probléma az elmúlt közel 60 évben rengeteg módszer fejlesztését motiválta. A legutóbbi, 2020-ban rendezett fehérje szerkezet becslő versenyen (CASP15) a DeepMind által fejlesztett AlphaFold nevű módszer eddig soha nem látott pontosságot ért el – sokak szerint lényegében megoldva a folding problémát. Bár az AlphaFold kétség kívül hatalmas előrelépés a tudomány terület számára, az eredeti célkitűzése szerint az általa célzott probléma csak egy jól körülhatárolt része a szerkezeti biokémiának: a fehérjék monomer, módosítás nélküli szerkezetének meghatározása. Ennek ellenére számos jel mutat arra, hogy az AlphaFold az eredeti céljával együtt számos más kapcsolódó problémát is megoldott, legalábbis részlegesen. Ezek közül az egyik a rendezetlen fehérjék felismerése. Mivel ezek a fehérjék stabil 3 dimenziós szerkezet nélkül látják el a sokszor kritikusan fontos biológiai funkcióikat, definíció szerint az AlphaFold fókuszán kívül vannak. Ennek ellenére az AlphaFold egy meglepően hatékony módszerként használható ezen fehérjék szekvencia-alapú azonosításában és a kölcsönhatásaik modellezésében, amely potenciálisan új fejezetet nyithat a rendezetlen fehérjék szerkezeti tulajdonságainak és funkcióinak vizsgálatában.

Bioinformatika TDK különdíjat nyert hallgatók előadásainak kivonatai

Fajspecifikus elemek felderítése a *Candida* fajokkal szemben kialakuló természetes immunválaszban humán szájüregi epitélisejtek esetén

Horváth Márton

SzTE Mikrobiológiai Tanszék

A *Candida* nemzetség tagjai opportunisták humán patogén kórokozókként ismertek, melyek közül a *C. albicans* és *C. parapsilosis* előfordulásukat tekintve a leggyakoribbak. A *C. parapsilosis*, a *C. albicans*-szal ellentétben az egészséges bőrflóra tagjának tekinthető, továbbá, mindkét faj jelen van egészséges egyének szájüregében. Azonban, míg immunszuppresszált páciensek esetén a szájüregi candidiázis esetek legfőbb kiváltója a *C. albicans*, kifejezetten ezen felszíni fertőzések kiváltása *C. parapsilosis*-ra ritkán vagy egyáltalán nem jellemző. A szájüregben tehát míg a *C. albicans*-ra valódi opportunistáként, addig a *C. parapsilosis*-ra inkább egy kommenzalista fajként tekinthetünk. Ennek oka nagy valószínűséggel az, hogy a két faj fertőzésbiológiája számos szempontból eltér. Ezt a jelenséget a kísérleteink során egészséges eredetű szájüregi epitélisejtekénél is megfigyeltük. Az eredményeink alapján a szájüregi epitélisejtek esetében csak a *C. albicans* váltott ki jelentős gazdasejt károsodást és jelentősebb gyulladási választ is csak ezen faj esetén volt megfigyelhető. Mindezekből adódóan célunk ennek a jelenségnek a feltárása volt, tehát annak megválaszolása, hogy miféle fajspecifikus válaszok jönnek létre egy kommenzalista és egy patogén faj megkülönböztetésekor, illetve a kommenzalista esetén, mi történik magasabb kolonizáltság. A fertőzés utáni transzkriptóm analízis arra utalt, hogy génexpresszió szintjén a legtöbb változás később következik csak be, valamint, hogy a *C. albicans* jóval több gén expresszióját befolyásolja, mint a *C. parapsilosis*. Sikeresen fajra specifikusan megváltozott kifejeződésű géneket beazonosítottunk. További analíziseink arra utaltak, hogy a *C. albicans* jelenléte szinte kizárólag gyulladási útvonalakat és funkciókat aktivált, míg a *C. parapsilosis* fertőzés hatására változatos útvonalak, elsősorban vér- és érképzéssel kapcsolatos és oxidatív stressz választ szabályozó folyamatok aktiválódtak. A vizsgálatok során végeztünk miRNS szekvenálást is, amely elemzése során számos faj specifikusan megjelenő miRNS-t azonosítottunk. Ezek közül néhány miRNS-t ténylegesen a transzkriptóm analízis során beazonosított génexpressziós változásokhoz tudtuk kötni. Összefoglalva, szájüregi epitélisejtek esetén aktív megkülönböztetés zajlik a két faj között, ami feltehetőleg miRNS-ek révén is szabályozott.

Neurális reprezentációk vizsgálata felügyelet nélküli gépi-tanulással.

Kis Márton

KOKI Biológia számítások kutatócsoport

A neurális aktivitást klasszikusan felügyelt tanulós technikákkal analizálták, ahol a vizsgálni kívánt változót, illetve annak pillanatbeli értékét előre rögzíteni kell az analízis során. Ez a keretrendszer nem tesz lehetővé új, eddig ismeretlen vagy kísérletesen nem mérhető változók, funkcionalitások felfedezését melyet egy adott agyterület végez. Kutatásom célja egy felügyelet nélküli mélytanuló módszer, nevezetesen Variációs Autoenkóder (VAE) segítségével felfedni patkány hippocampális aktivitását meghajtó rejtett struktúrát, beazonosítani és dekódolni a navigáció során reprezentált változókat. Ezzel a módszerrel lehetőségünk van kinyerni az adatban rejlő látens struktúrát, egy értelmezhető, potenciálisan alacsonydimenziós reprezentációját az egyébként zajos és magasdimenziós neurális adatnak

Munkám során megmutattam, hogy a hippocampális aktivitás térbeli exploráció során potenciálisan 3 dimenziós, melyből 2 dimenzió a térbeli koordinátákat kódolja. Megmutattam továbbá, hogy a VAE-k sikeresen alkalmazhatóak neurális adaton, azonban, ha a valódi látens faktorok nem Gauss eloszlásúak vagy a leképezés a látensek és a megfigyelések között erősen nemlineáris (mint a felhasznált adatban), a VAE-k nem elég rugalmasak és további nemlineáris dimenzióredukációs technika (pl. Isomap) alkalmazása szükséges.

Fiatal kutatók előadásainak kivonatai

A bakteriofág farokrostok kompatibilitásának és diverzitásának feltárása a Podoviridae és Autographiviridae családokban

Asbóth András¹, Kintses Bálint¹, Ari Eszter^{1,2}, Stirling Tamás¹, Apjok Gábor¹

¹ SzBK Biokémiai Intézet Szintetikus és Rendszerbiológiai Egység

² ELTE TTK Genetikai Tanszék

A bakteriofágok és baktériumok között zajló folyamatos evolúciós fegyverkezési verseny egyik legfontosabb szereplője a baktériumok sejtfalát alkotó poliszacharid alapú kapszid antigének és a fágok fertőzési specificitásáért felelős farokrostok. A bakteriofágok jellemzően erős sejtpusztító kapacitással rendelkeznek a gazdabaktériumokkal szemben, azonban általában kevés különböző törzset ismernek fel, ami korlátozza a terápiás célú alkalmazhatóságukat. Ezért tűztük ki célul, hogy a farokrostok cseréjével kibővítvjük a fertőzési spektrumukat. Célunk egy dinamikusan fejlődő pipeline felépítése, amely képes feltérképezni, hogy mely fágfarkak és fágtestek kompatibilisek egymással, illetve milyen baktériumokat képesek fertőzni. A kompatibilitási problémákat az okozza, hogy a farkak nem tetszőlegesen illeszthetőek a vírusok testébe, ugyanis a fehérje N terminális részén lévő konzervatív "adaptor" régióknak nagy azonosságot kell mutatnia a kicserélt részek között. Az adaptor régió határait még nem térképezték fel, így fontos feladat volt a projekt megvalósításában, hogy annotáljuk ezeket a genomokban. Azonosítottuk a fágfarkak enzimikus doménjeit is, amelyek betekintést adnak a vírus biokémiai aktivitásába, valamint a célpont baktérium kapszid típusának azonosításával meghatározható a specificitásuk is. Eredményünk az adaptor és az enzimikus domének szekvencia azonosság alapján megvalósított klaszterezése, ami predikciós értékű jövőbeli alternatív gazdaspecificitású fágok szintéziséhez. Céljaink elérése új utakat nyithat meg a fágterápiában és kiszélesítheti a funkcionális metagenomika eszköztárát is.

Láthatóvá tenni a láthatatlant: az Apobec3 mutagenezis hatása a Sars-Cov-2 adaptív immunfelismerésére

Balogh Gergő Mihály^{1,2}, Asztalos Leó¹, Pál Csaba³, Kemény Lajos^{1,2} és Manczinger Máté^{1,2,3}

¹ SZTE ÁOK Szent-Györgyi Albert Klinikai Központ Bőrgyógyászati és Allergológiai Klinika

² HCEMM-USZ Skin Research Group

³ SzBK Biokémiai Intézet Szintetikus és Rendszerbiológiai Egység

A SARS-CoV-2 fertőzések hatékony leküzdésének egyik kulcsfolyamata a CD8+ T sejtek általi adaptív immunitás. Ennek előfeltétele a vírusból származó peptidok HLA-I-függő bemutatása a fertőzött sejtek felszínén. A SARS-CoV-2-ben megjelent mutációk egy része szignifikánsan csökkenti bizonyos virális

fehérjeszakaszok affinitását gyakori HLA-I molekulákhoz. De vajon a mutációk csak az immunfelismerés gyengülését eredményezhetik?

Korábbi munkák eredményeivel összhangban a C>U változások jelentős felhalmozódását tapasztaltuk a SARS-CoV-2-ben. E nukleotid-cserék jelentős részéért a veleszületett immunitás elemeiként működő humán APOBEC3 enzimcsalád tagjai a felelősek. Számítógépes predikcióink alapján azonban ezen mutációk nem csupán a leggyakoribbak, de egyúttal nagy arányban vezetnek új, HLA-kötött peptidek kialakulásához. Az APOBEC3 és a HLA-I molekulák funkcionális és evolúciós kapcsolatára utalhat emellett egy közel ezer fős humán kohorton végzett vizsgálatunk eredménye is. Az APOBEC3 által okozott mutációkra érzékenyebbek voltak azon egyének HLA-I molekulái, akikben az APOBEC3G gén magasabb expressziójával összefüggésbe hozott genomi mintázatok voltak jelen.

Eredményeink az APOBEC3 enzimek egy új hatásmechanizmusát vetítik előre: sejtjeink belső mutagén faktoraként „láthatóbbá” teszik a virális peptidek korábban „láthatatlan” részeit az adaptív immunitás számára a HLA-I molekulákon keresztül.

Sherlock: egy Big Data technológián alapuló adatgyűjtést és elemzést elősegítő platform

Bohár Balázs^{1,2}, Fazekas Dávid^{1,2}, Matthew Madgwick^{1,3}, Csabai Luca^{1,2}, Ölbei Márton^{1,3}, Korcsmáros Tamás^{1,3}, Szalay-Bekő Máté¹

¹ Earlham Institute, Norwich Research Park

² ELTE TTK Genetikai tanszék

³ Gut Microbes and Health Programme, Quadram Institute Bioscience, Norwich Research Park

A Big Data technológia korszakában az adatgyűjtés minden eddiginél fontosabb része a biológiai kutatásoknak. Sok esetben ez még időigényesebb folyamat lehet, mint maga az elemzés. Ennek az az oka, hogy adott esetben több különböző nyilvános adatbázist is le kell tölteni, a letöltött fájlok nagy része eltérő adatszerkezettel rendelkezik, a közös formátumra alakítás pedig sok időt és energiát igényel. Erre az egyre gyakrabban felmerülő nehézség kiküszöbölésére készítettünk egy nyílt forráskódú, felhő alapú, Big Data platformot, a Sherlockot (<https://earlham-sherlock.github.io/>). A Sherlock lehetőséget biztosít a különböző biológiai adatok tárolására, konvertálására, lekérdezésére, megosztására valamint generálására, miközben egyszerűsíti és hatékonyabbá teszi az adatkezelést. A platform több különböző Big Data technológián alapuló megoldást használ, mint például a Docker, a PrestoDB, vagy a Hive. A Sherlock célja, hogy az olykor hosszadalmas adatgyűjtési folyamatokat lerövidítse, valamint gyors és hatékony segítséget nyújtson a rendkívül nagyméretű adatfájlok feldolgozásában. Ezenkívül a platform képes több forrásból származó,

különböző adatstruktúrával rendelkező adatot kezelni (interakciós, lokalizációs vagy expressziós), valamint átalakítani egy tömörített és optimalizált formátumra, például az Optimized Row Columnar (ORC) fájlformátumra. A platform github oldalán található különböző online elérhető adatbázisokhoz (interakciós, expressziós, genomikai) megírt, úgynevezett "loader" scriptek segítségével, a felhasználók gyorsan és egyszerűen alakíthatják át erre az egységes formátumra a különböző adatszerkezettel rendelkező fájlokat. A Sherlock platform egy nyílt forráskódú, a github oldalán bárki számára hozzáférhető, ingyenes szoftver, amely lehetővé teszi a gyorsabb és hatékonyabb adatkezelést, adatelemzést, adatintegrációt a modern Big Data technológiák segítségével.

Signalink3: töbrétegű erőforrás a szövetspecifikus jelátviteli hálózatok feltárására

Csabai Luca^{1,2}, Fazekas Dávid^{1,2}, Kadlecsik Tamás², Szalay-Bekő Máté¹, Bohár Balázs^{1,2}, Matthew Madgwick¹, Módos Dezső^{1,3}, Ölbei Márton^{1,3}, Gul Lejla¹, Padhmanand Sudhakar^{1,4}, Kubisch János², Oyebode James Oyeyemi⁵, Liska Orsolya^{6,7,8}, Ari Eszter^{2,6,7}, Bernadette Hotzi², Viktor A. Billes^{2,13}, Molnár Eszter², Földvári-Nagy László⁹, Csályi Kitti², Demeter Amanda¹, Pápai Nóra^{2,10}, Koltai Mihály¹¹, Varga Máté², Lenti Katalin⁹, Farjas J Illés¹⁴, Türei Dénes¹⁵, Csermely Péter¹², Vellai Tibor^{2,13}, Korcsmáros Tamás^{1,3}

¹ Earlham Institute, Norwich, UK

² ELTE TTK Genetikai Tanszék

³ Gut Microbes and Health Programme, Quadram Institute Bioscience, Norwich

⁴ Translational Research in Gastrointestinal Disorders, Leuven, BE

⁵ Zeekay Institute of Advanced and Professional Studies, Lagos, Nigeria

⁶ HCEMM-BRC Metabolic Systems Biology Lab

⁷ SzBK Biokémia Intézet Szintetikus és Rendszerbiológiai Egység

⁸ SzTE Biológia Doktori Iskola

⁹ SE EK Morfológiai és Fiziológiai Tanszék

¹⁰ Institute of Molecular Biotechnology, Vienna, Austria

¹¹ Centre for the Mathematical Modelling of Infectious Diseases (CMMID), London School of Hygiene & Tropical Medicine, London, UK

¹² SE Molekuláris Biológiai Tanszék

¹³ ELKH/MTA-ELTE Genetika Kutató csoport

¹⁴ Citibank Europe plc Hungarian Branch Office, Budapest

¹⁵ Heidelberg University, Faculty of Medicine, and Heidelberg University Hospital, Institute for Computational Biomedicine, Bioquant, Heidelberg, Germany

A jelátviteli hálózatok a sejtek különböző belső vagy külső ingerekre adott válaszát irányító molekuláris mechanizmusokat jelentik. A legtöbb jelenleg rendelkezésre álló jelátviteli adatbázis az egymásba fonódó útvonalak komplex hálózatának csak egy részét tartalmazza, kihagyva a kulcsfontosságú kölcsönhatásokat vagy folyamatokat. Ezért fejlesztettük ki a Signalink3

(<http://signalink.org/>) tudásbázist, amely kézi gyűjtésből származó adatokat szolgáltat a jelátviteli útvonalakról, ezeket integrált adatokkal kiegészítve többféle adatbázisból (kölsönhatás, szabályozás, lokalizáció stb.). A Signalink3 több mint 400 000 újonnan hozzáadott humán fehérje-fehérje kölcsönhatást tartalmaz, mely összesen 700 000 kölcsönhatást jelent *H. sapiens*-re vonatkozóan, így ez az egyik legnagyobb integrált jelátviteli hálózati forrás. A *H. sapiens* mellett a Signalink3 az egyetlen olyan jelenlegi jelátviteli hálózati forrás, amely a *C. elegans* és *D. rerio* modell fajokra vonatkozó szabályozási információkat nyújt, és a legnagyobb forrás *D. melanogaster*-re vonatkozóan. A korábbi verziókhoz képest génexpressziós adatokat, valamint az interaktorok sejtbeli lokalizációját is tartalmazza az erőforrás, így egyedülálló módon lehetővé teszi a szövet-, vagy kompartment-specifikus útvonalak elemzését, elősegítve pontosabb modellek létrehozását. Az adatok szabadon letölthetők széles körben használt formátumokban, többek között CSV, PSI-MI TAB vagy BioPAX formátumban.

Nyúl (*Oryctolagus cuniculus*) cirkuláris RNS-ek azonosítása és jellemzése betegség modellek létrehozási lehetőségeinek felderítésére

Fekete Zsófia¹, Kontra Levente^{1,2}, Barta Endre^{1,2}

¹ Magyar Agrár- és Élettudományi Egyetem

² Debreceni Egyetem

A cirkuláris RNS-ek (circRNS) olyan nem kódoló RNS-ek, amelyek kovalensen körré zárulnak. Általában protein kódoló génekből alakulnak ki a splicing során és főleg exonokat tartalmaznak. Vannak emellett intront is tartalmazó, de kizárólag intront tartalmazó variációk is. Ma már ismert, hogy ezek a molekulák fontos és széleskörű szabályozó szereppel rendelkeznek, akár miRNS sponge-ként, akár cisz-regulátorokként, vagy az epigenetikai mintázat kialakításában. Ismertek fehérjét kódoló circRNS-ek is, bár ezek ritkák.

Úgy véljük hogy, hogy emellett más típusú circRNS-ek is létezhetnek, melyek működési mechanizmusa még nem felderített.

Célul tűztük ki a circRNS-ek nyúlban való általános feltérképezését, amelyhez elsősorban bioinformatikai módszereket alkalmazunk. Publikus, a hosszú nem kódoló RNS-eket tartalmazó (rRNS depléciós) szekvenálási adatokból kiindulva derítjük fel 9 szövettípusban és 7 házinyúl fajtában az azokban megtalálható circRNS-eket. Ezt követően célunk ezen RNS-ek kategorizálása. A rendelkezésre álló, publikált pipeline-ok mellett saját keresési módszert is tervezünk fejleszteni ami a priori módszerrel keres cirkuláris RNS-eket.

Ma már léteznek az eddig felderített circRNS-eket tartalmazó adatbázisok is. A nyúl, annak ellenére, hogy elterjedten alkalmazott állatmodell humán betegségekben, nem szerepel ezekben. Az általunk felderített circRNS-eket

tervezzük összevetni az adatbázisokban más fajokról megtalálható adatokkal, evolúciósan konzervált és nyúl specifikus változatok keresése céljából.

PSINDB: új, átfogó és részletes posztzinaptikus interakciós adatbázis

Kálmán Zsófia¹, Dudola Dániel¹, Mészáros Bálint², Gáspári Zoltán¹, Dobson László^{2,3}

¹ PPKE ITBK

² Structural and Computational Biology Unit, EMBL

³ TTK Enzimológiai Intézet

A posztzinapszis a neuronális jelátvitel felvételi oldala, amelyet több ezer fehérje alkot és amelynek olyan jelentős folyamatokban van elengedhetetlen szerepe, mint a memória és a tanulás. Bár ma már számos különböző fehérje-fehérje interakciós adatbázis létezik, azonban ezek tartalma és felbontása igen különböző lehet az interakciók tekintetében, emellett ritkán tartalmazzák a fehérje-fehérje kapcsolatok kialakulásának részletesebb adatait

Célunk egy átfogó és részletes adatbázis létrehozása volt, amellyel áthidaljuk az előbbiekben felvázolt nehézségeket. Az elkészült PSINDB (PostSynaptic Interaction Data Base) adatbázis az általunk annotált és más adatbázisokban található szinkronizált interakciós adatok mellett olyan szerkezeti jellemzőket is tartalmaz, amelyek a kölcsönhatások szabályozásában töltenek be fontos szerepet. Emellett kiemelő, hogy több tízezer aminosav szintű kötőrégiót, a posztzinaptikus fehérjék izoformáit, mutációs adatokat, valamint a kötőpartnerek Gene Ontology alapú funkcionális fingerprintjét is tartalmazza. Reményeink szerint az adatbázis számos területen használható lesz a rendszerbiológiai modellezéstől a fehérje komplexek vizsgálatára vonatkozó kísérletek tervezéséig.

A T sejtek pozitív szelekciója vakfoltot eredményez az adaptív immunfelismerésben

Koncz Balázs¹, Balogh Gergő Mihály¹, Papp Benjamin Tamás¹, Asztalos Leó¹, Kemény Lajos^{1,2,3}, Manczinger Máté^{1,2,3,4}

¹ SZTE ÁOK Szent-Györgyi Albert Klinikai Központ Bőrgyógyászati és Allergológiai Klinika

² MTA-SZTE Dermatológiai Kutatócsoport

³ HCEMM-UM, Skin Research Group, Szeged

⁴ SzBK Biokémia Intézet Szintetikus és Rendszerbiológiai Egység

Az adaptív immunfelismerés feltétele, hogy a T sejtek megkössék a HLA molekulák által bemutatott peptideket. Ahhoz, hogy kialakuljon egy megfelelően funkcionáló T sejt repertoár, kulcsfontosságú a sejtek pozitív szelekciója a csecsemőmirigyben. Ennek során csak azok a sejtek maradnak életben, amelyek képesek felismerni a kortikális tímusz epitél sejtek felszínén bemutatott humán peptideket. Elméletünk szerint számos idegen peptiddel szemben nem marad fenn specifikus T sejt, mivel a folyamatot saját peptidek mediálják. Vizsgálataink

során a peptideknek azon aminosav-motívumaira fókuszáltunk, amelyek kapcsolatba kerülnek a T sejt receptorokkal. Kevésbé voltak immunogének azok a motívumok, amelyek nagyon ritkán vagy egyáltalán nem fordulnak elő a humán fehérjékben, vagy nem jelennek meg a kortikális tímusz epitél sejtek felszínén. Emellett egészséges egyének naiv T sejt repertoárjában kisebb eséllyel találtunk olyan sejteket, amelyek képesek felismerni ezeket a motívumokat. Továbbá megmutattuk, hogy ezen motívumok HLA-bemutatása növeli a hajlamot fertőző betegségek kialakulására. Eredményeink arra utalnak, hogy a T sejtek pozitív szelekciója egy vakfoltot eredményez az adaptív immunfelismerésben, amely az idegen peptidek egy jelentős részének felismerését akadályozza.

A posztszinaptikus komplexom változékonyságának modellezése a fehérjegyakoriság függvényében

Miski Marcell¹, Keömley-Horváth Bence^{1,2}, Simone Rizetto², Csikász-Nagy Attila^{1,2,3}, Gáspári Zoltán¹

¹ PPKE ITBK

² Cytocast Kft.

³ Randall Centre for Cell and Molecular Biophysics, King's College London, London, UK

A posztszinaptikus denzitás (PSD) egy bonyolult fehérjehálózat a fogadó neuron membránja alatt. Több ezer fehérjéből áll, amelyek számos interakciós doménon és régión keresztül képesek kötni egymást. Jelenlegi ismereteink szerint a PSD dinamikus átrendeződésre képes, és komoly variabilitást mutat a neuron típusától, a szinapszis funkcionális állapotától valamint a cirkadián ritmustól függően. A PSD szerkezetének kísérleti vizsgálatát bonyolultsága és dinamikus jellege nehezíti, de az alkotó fehérjék mennyiségi eloszlására vonatkozóan számos adat áll rendelkezésünkre. Munkánk során hét fő posztszinaptikus fehérje által alkotott komplexek eloszlását modelleztük numerikus szimulációk segítségével az egyedi fehérjék gyakoriságának függvényében. Összesen 524, különböző agyterületekről és egyénektől származó mRNS adatkészletet elemeztünk a Cytocast programban megvalósított ágens-alapú Gillespie algoritmus segítségével. Tapasztalataink szerint a szimulációk robusztusak és reprodukálhatók, a kvázi-egyensúlyi állapot viszonylag korán beáll, és a nagyobb komplexek száma az idő múlásával csak kis mértékben növekszik. Eredményeink arra utalnak, hogy a fehérjegyakoriság és a kialakuló komplexek eloszlása közötti kapcsolat összetett: egyes esetekben a fehérje hozzáférhetőségének viszonylag kis változása a komplexek jelentős átrendeződéséhez vezet. Megállapítottuk, hogy a posztszinaptikus komplexomot főként a PSD-95 állványfehérje és az AMPA receptor gyakorisága befolyásolja egyszerűsített rendszerünkben.

Eredményeink összhangban vannak a posztzinaptikus „nanodomének” jelenlétével a PSD-ben, és kompatibilisek a szinaptom elmélettel is, amely a szinapszis-identitás szerepét hangsúlyozza az agyműködésben. Bár modellünk jelentős egyszerűsítésekkel él, úgy gondoljuk, hogy szimulációink megragadják a posztzinaptikus komplexom egyes alapvető jellemzőit és elősegítik a fehérje turnover szerepének jobb megértését a PSD szerveződésében.

Az ETS fehérjék DNS-kötésének vizsgálata egér csontvelői makrofágokban

Nagy Gergely és Nagy László

DE ÁOK Biokémiai és Molekuláris biológiai Intézet

A makrofágokat meghatározó transzkripciós faktorok (TF-ok) közül a purinban gazdag nukleinsavakat kötő, ún. PU.1 fehérjék találhatóak meg a legtöbb gén szabályozó helyein, beleértve a promótereket és enhanszereket is. A PU.1 az eritroblaszt transzformáció-specifikus (ETS) fehérje szupercsaládba tartozik, és egér csontvelői makrofágokban az interferon szabályozó faktor (IRF) 8-cal működik szorosan együtt – akár közös, ún. kompozit elemeken. A PU.1 mellett egyéb ETS fehérjék is kifejeződnek ezekben a sejtekben, melyek közül három fehérjecsalád képviselőjének is nyilvánosan elérhető a cisztrómja (azaz a kötőhelyek összessége). Ha e cisztrómokban motívumkeresést végzünk, minden alkalommal megtaláljuk az ETS fehérjékre jellemző 5'-GGAA-3' központi szekvenciát, ám ezek szomszédos bázispárjai már közel sem ilyen egyhangúak – mutatnak egyezéseket és specifikus jellegeket is, melyek elkülönítése további elemzéseket igényel.

Munkám során klasztereztem az ETS és IRF8 kötőhelyeket a különböző TF-ok általi lefedettség (ChIP-seq) alapján, és az így elkülönített, egy vagy néhány TF által dominált klaszterekben vizsgáltam az egyes szekvenciák feldúsulását. Az elemzések során az 5' metilálható (és promóter-specifikus) szekvenciák jól elkülönültek az „A/G” és „CA” kezdetű szekvenciáktól, melyeket rendre a GABP/ELF/FLI, a PU.1 és az ELF/FLI/PU.1 fehérjék kötnek a legnagyobb affinitással. Emellett tisztázható volt, hogy a PU.1 és IRF8 csak egyféle hatékony kompozit elemmel (EICE) rendelkezik.

Az ARHGAP25 interakciós hálózatának feltárása humán neutrofil granulocitákban

Sasvári Péter¹, Pettkó-Szandtner Aladár², Wisniewski Éva¹, Czárán Domonkos¹, Csépanyi-Kömi Roland¹

¹ SE Élettani Intézet

² SzBK, Proteomika Labor

Az ARHGAP25 elnevezésű, Rac-specifikus GTPáz aktiváló fehérje egyre nagyobb hangsúlyt kap sejtélettani szempontból. Fontossága az elmúlt években előtérbe került, ugyanis folyamatosan jelennek meg olyan kutatási eredmények,

amelyben az ARHGAP25 expressziójának változása különböző tumorsejtekben kihat ezen sejtek migrációs képességére. Mivel ez a fehérje ilyen sokrétű folyamatokban lényeges szerepet tölt be, felmerül a lehetősége, hogy gyógyszercélpontra váljon, viszont ehhez szükséges pontosan megismerünk működésének részleteit.

Célkitűzés: Az ARHGAP25 partnerfehérjéinek proteomikai analízise humán neutrofil granulocitákban immunoprecipitáció segítségével.

Módszerek: Perifériás vérből nyert neutrofil granulocitákból az immunoprecipitáció elvégzése poliklonális anti-ARHGAP25 antitesttel bevont protein A gyöngyökkel történt. A mintákat tömegspektrometria segítségével analizáltuk. A nyers adatok összehasonlításához a MaxQuant program Label Free Quantification (LFQ) módszerét alkalmaztuk.

Eredmények: Számos fehérjét sikerült azonosítani, ebből 21 darabot detektáltunk minden mintánkban megfelelő "Fold Change" érték felett. A kapott fehérjék között találunk kis G-fehérjéket, illetve az aktin reorganizáció és a membrántranszport folyamatának résztvevőit. A Rac fehérjét, mely az eddig ismert interakciós partnere az ARHGAP25-nek, minden mérés során detektáltuk.

Konklúzió: A kapott fehérjék nagy hányada beleillik az ARHGAP25-ről eddig alkotott képbe, de eredményeink több olyan fehérjét is felfedtek, amelyet eddig nem valószínűsített potenciális partnerfehérjék.

Támogatás: A projekt a Bolyai János Kutatási Ösztöndíj és az NKFIH FK128376 támogatásával készült.

Az epteliális-mezenchimális átmenet (EMT) új, többsejtes hálózatos modelljének kifejlesztése és új, potenciális gyógyszercélpont megtalálása a hibrid-EMT-nek a modellbe építésével

Seitz Erik¹, Nina Kunsic¹, Albert Réka², Csermely Péter¹

¹ SE ÁOK Molekuláris Biológiai Tanszék

² Pennsylvania State University, State College, PA, Egyesült Államok

A rákos metasztázisok (áttétek) kialakulását elősegítő biológiai folyamat az epiteliális-mezenchimális átmenet (EMT). Az átmenet folyamán egy köztes fenotípus fordul elő, a hibrid EMT, amely epiteliális (adhézió) és mezenchimális (migráció) tulajdonságokkal egyaránt jellemezhető. Ez a sejtes állapot képes sejtcsoportosulásokat létrehozni a véráramban, így erősebb életképességet és gyógyszer-toleranciát elérve. E folyamat hálózati modelljének kiépítésével újabb potenciális gyógyszer-célpontok találhatóak, amelyek kezelésével a metasztázisokhoz vezető alakulási folyamat leállítható.

Az EMT-t egy jelátviteli hálózat irányítja. Ennek a folyamatnak a rendszerszintű vizsgálatának céljából két EMT hálózat lett összekötve kilenc jelátviteli útvonalon

keresztül, így létrehozva egy intercelluláris modellt. Ez lehetővé tette a sejt-sejt közti kommunikáció szimulálását. A Boole-modell emberi szövetből származó empirikus adatok alapján lett felépítve, amely 120 nódusból és 425 interakcióból áll. In silico szimulációkkal azonosítani lehetett a modellben a hibrid állapotot, ami be lett építve.

Intercelluláris szimulációk kimutatták, hogy a hálózat egyik nódusának, a COP9 szignalizómának (CSN) gátlását követően a hibrid sejt epiteliális állapotba tranzícionált. Egy ilyen átmenet leállíthatja a metasztázisok terjedését. Ebből adódóan CSN egy jó kiindulási pontja proof-of-concept kísérleteknek egy potenciális rákellenes gyógyszerként.

Webchem: Egy R csomag online kémiai adatbázisok elérésére

Eduard Szöcs¹, Tamás Stirling², Eric R. Scott³, Andreas Scharmüller¹, Ralf B. Schäfer¹

¹ *Universität Koblenz-Landau*

² *SzBK Biokémiai Intézet Szintetikus és Rendszerbiológiai Egység*

³ *Tufts University*

Természettudományos kutatás során gyakran előfordul, hogy online adatbázisokból nagy mennyiségű adathoz kell hozzáférnünk. Az adatok kézi keresése időigényes, hibával járhat, és amennyiben a hozzáférés menete nincsen pontosan dokumentálva, úgy a folyamat nem reprodukálható. Az elmúlt években az adatszolgáltatók jelentős része lehetővé tette, hogy adatbázisaik programozva is elérhetőek legyenek. Ez nagy előrelépés, de a létrehozott felületek használatához programozási ismeretek szükségesek, az adatok felhasználóinak jelentős része viszont nem rendelkezik megfelelő informatikai háttérrel a nagyméretű adatbázisok programozott letöltéséhez.

A webchem R csomagot azért hoztuk létre, hogy megkönnyítsük ezt a folyamatot. Mivel az R programnyelvnek széles felhasználói bázisa van, ezért úgy gondoltuk, hogy egy R csomag ideális megoldás arra, hogy a programozható hozzáférést közelebb hozza az adatok felhasználóihoz. A webchem kis molekulatömegű vegyületekre specializálódott, jelenleg 14 adatbázis érhető el segítségével. A lekérdezések során az adatokat közvetlenül R-be importáljuk, amelyek így könnyen integrálhatók további elemzésekbe. A csomag könnyű és reprodukálható hozzáférést tesz lehetővé, az adatokat ismert struktúrákba rendezi, összességében pedig csökkenti az adattáblák összeállításával eltöltött időt. A csomag folyamatos fejlesztés alatt áll.

Ligandum dokkolás explicit vizekkel – egy megoldás virális ioncsatornákon tesztelve

Zsidó Balázs Zoltán, Börzsei Rita, Szél Viktor, és Hetényi Csaba

PTE ÁOK Farmakológiai és Farmakoterápiás Intézet, Farmakoinformatikai Tanszék

A kísérletesen meghatározott gyógyszer-célpont szerkezetek a számítógépes gyógyszertervezés alapvető elemei, azonban amekkora segítséget nyújtanak, esetekben ugyan olyan félrevezetőek is lehetnek, gyakran nem, vagy rossz pozícióban tartalmaznak szerkezeti vizeket, ezek a jelenségek sokszor eredménytelen gyógyszerfejlesztési projektekhez vezetnek.

Jelen tanulmányban az említett problémára fejlesztettünk egy módszert. A virális ioncsatornák esetében a szerkezeti vízmolekulák kiemelkedő szerepet játszanak a ligandum-kötésben. Ezért, molekula mechanikai, molekuláris dokkolási és molekula dinamikai eljárásokkal vizsgáltuk a virális ioncsatornák gyógyszermolekula kötési mechanizmusát, részben már meglévő komplexek reprodukálásával, részben új gyógyszerjelöltek kötődési pozíciójának újonnan történő meghatározásával, a szerkezeti explicit vízmolekulák figyelembevételével. Mind a régi célpont (influenza A vírus) és az új célpont (SARS-CoV-2) esetében a kísérletes eredményekkel kiváló egyezéseket sikerült elérnünk.

Köszönetnyilvánítás: Eredményeinket az alábbi támogatások felhasználásával értük el és ezért köszönetet mondunk a következő szervezeteknek: Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal (K123836), 2017-1.2.1-NKP-2017-00002 (NAP-2; Chronic Pain Research Group), EFOP-3.6.1-16-2016-0004 and GINOP 2.3.2-15-2016-00050 "PEPSYS". A munka a Magyar Tudományos Akadémia Bolyai János kutatási ösztöndíja támogatásával készült el. Megköszönjük a Kormányzati Informatikai Fejlesztési Ügynökségnek (KIFÜ) a szuperszámítógépes infrastruktúrához való hozzáférést. Az innovációs és technológiai minisztérium ÚNKP-21-3-II és 21-5 kódszámú Új Nemzeti Kiválóság Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült. A projektet támogatta az Európai Unió, és az Európai Szociális Alap. A projekt neve és kódja: Comprehensive Development for Implementing Smart Specialization Strategies at the University of Pécs, EFOP-3.6.1-16-2016-00004.