

A MAGYAR
TUDOMÁNY
ÜNNEPE



Magyar Tudományos
Akadémia

Bioinformatika 2022

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

2022. november 11.

Program

1. szekció, Elnök: Pongor Sándor

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

10:05 - 10:25: **Csabai István**

Öregedési folyamatok molekuláris mechanizmusainak elemzése értelmezhető gépi tanulással

ELTE TTK Fizikai Intézet Komplex Rendszerek Fizikája Tanszék

10:25 - 10:45: **Grolmusz Vince**

A lineáris SVM-ek néhány szép alkalmazása gráfok és sorozatok elemzésében

ELTE Matematikai Intézet PIT Bioinformatikai Csoport

10:45 - 11:05: **Ari Eszter**

Lesz-e az E. coli-ból superbaktérium? – Az antibiotikum-rezisztencia- és virulenciagének evolúciós függőségeinek vizsgálata

HCEMM-BRC Metabolikus Rendszerbiológiai Kutatócsoport; SZBK Biokémiai Intézet Szintetikus és Rendszer-biológiai Egység; ELTE Genetikai Tanszék

11:05 - 11:25: **Lovrics Anna**

Egy rács-alapú, malignus sejtek szimulációjára fejlesztett ágens alapú szoftver bemutatása

TTK Enzimológiai Intézet Gyógyszer Rezisztencia Kutatócsoport

11:25 - 11:35: **Szünet**

2. szekció - Fiatal Kutatók előadásai, Elnök: Csabai István

11:35 - 11:50: **Hajdú Bence**

Az autofágia indukciójában kulcsfontosságú mTORC1-ULK1-PP2A szabályozási háromszög működésének rendszerbiológiai vizsgálata

SE Molekuláris Biológia Tanszék

11:50 - 12:05: **Mentes Anikó**

SARS-CoV-2 PCR primerek célrégióiban megjelenő mutációk vizsgálata

ELTE TTK Fizikai Intézet Komplex Rendszerek Fizikája Tanszék

12:05 - 12:20: **Kalcsevszki Regina**

A SARS-CoV-2 potenciális új variánsai által okozott betegség lefolyásának előrejelzése gépi tanulás segítségével mutációs adatok alapján

PPKE ITBK

12:20 - 12:35: **Koncz Balázs**

A HLA-I és HLA-II variánsok közötti kölcsönhatás befolyásolja a tumorelles immunválasz erősségét és időzítését melanómás betegeknél

SzTE SZAOK Szent-Györgyi Albert Klinikai Központ Bőrgyógyászati és Allergológiai Klinika

12:35 - 12:50: **Munkácsy Gyöngyi**

A TNMplot.com bővítése: új analízisek lefuttatási lehetőségei és új grafikus megjelenítési formák bemutatása

SE Bioinformatika Tanszék

12:50 - 13:20: **Szünet**

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

13:20 - 13:35: **Hoffka Gyula**

Venezuelai ló-láz encephalitis vírus proteáz öninaktivált konformerének vizsgálata molekuladinamikai szimulációkkal

DE ÁOK Biokémiai és Molekuláris Biológiai Intézet Retrovirális Biokémiai Kutató Laboratórium; DE Molekuláris Sejt- és Immunbiológiai Doktori Iskola

13:35 - 13:50: **Kothalawala William Jayasekara**

Vitális funkciók paraméterezése aktivitásmerő okoskarkötőkkel
SE Bioinformatika Tanszék

13:50 - 14:05: **Börzsei Rita Judit**

A protein-arginin-metil-transzferáz-5 enzim foszforilációjának szerkezeti hatásai a H4 hisztonhoz való kötődésére

PTE ÁOK Farmakológiai és Farmakoterápiái Intézet Farmakoinformatikai Tanszék

14:05 - 14:20: **Pándy-Szekeres Gáspár**

GPCRdb, a G protein-kapcsolt receptorok információs adatbázisa

Department of Drug Design and Pharmacology, University of Copenhagen; TTK Szerveskémiai Intézet Gyógyszerkémiai Kutatócsoport

14:20 - 14:35: **Szünet**

4. szekció - Elnök: Simon István

14:35 - 14:55: **Dobson László**

TmAlphaFold adatbázis: AlphaFold2 segítségével becsült transzmembrán fehérje szerkezetek membrán lokalizációja és értékelése

TTK Enzimológiai Intézet Fehérje bioinformatika kutatócsoport; EMBL Structural and Computational Biology Unit

14:55 - 15:15: **Hegedűs Tamás**

Membránfehérjék szerkezete az AlphaFold AI után

SE Biofizikai és Sugárbiológiai Intézet; ELKH-SE Biofizikai Virologia Kutatócsoport

15:15 - 15:35: **Keserű György Miklós**

Fehérjék kötőhelyeinek fragmens alapú azonosítása

TTK Gyógyszerkémiai Kutatócsoport

15:35 - 15:55: **Gyórfy Balázs**

Az immuneterápia hatékonyságának előrejelzése

TTK Enzimológiai Intézet Onkológiai Biomarker Kutatócsoport; SE Bioinformatika Tanszék

15:55 - 16:00: **Zárszó (Gyórfy Balázs)**

Előadások kivonatai

Lesz-e az E. coli-ból szuperbaktérium? - Az antibiotikum-rezisztencia- és virulenciagének evolúciós függőségeinek vizsgálata

Ari Eszter^{1,2,3}, Kiss Enikő², Vásárhelyi Bálint Márk^{2,4}, Fekete Gergely², Stirling Tamás^{4,5}, Ágoston Hunya^{2,3}, Kintsés Bálint^{2,4,5}, Papp Balázs^{1,2,4}

¹HCEMM-BRC Metabolikus Rendszerbiológiai Kutatócsoport

²SZBK Biokémiai Intézet Szintetikus és Rendszer-biológiai Egység

³ELTE Genetikai Tanszék

⁴SZBK Biotechnológiai Nemzeti Laboratórium

⁵HCEMM-BRC Transzlációs Mikrobiológiai Kutatócsoport

Az antibiotikum-rezisztencia- és a virulenciagének leggyakrabban horizontális transzfer útján kerülnek a baktériumokba, ezzel járulva hozzá az új, multidrog-rezisztens patogén variánsok megjelenéséhez. Egyre több bizonyíték utal arra, hogy a donor baktériumban már jelen lévő gének befolyásolják az újabb génnyerések sikerét. Azonban a gének közötti evolúciós függőségeket – vagyis azt, hogy az egyik gén elősegíti-e vagy éppen akadályozza egy második gén nyerését – ezidáig kevéssé tanulmányozták. Ezért több mint 20 000 *Escherichia coli* genom filogenetikai elemzésével elkészítettünk egy nagy felbontású térképet, melynek segítségével vizsgálhatóvá válnak a rezisztencia- és virulenciagének közötti evolúciós hatások. Térképünkéből kiderül, hogy (1) a rezisztencia gének általában elősegítik egymás nyerését; (2) a virulencia gének nem rendelkeznek ilyen általános mintázattal; és (3) – egyes korábbi eredményekkel ellentétben – nincs általános negatív összefüggés a virulencia- és a rezisztenciagének nyerése között, vagyis e két tulajdonság egymástól függetlenül evolválódik az *E. coli* baktériumban. Meglepő módon azt találtuk, hogy az efflux pumpa típusú rezisztenciagének jelenléte a genomban erősen megnöveli az egyéb rezisztenciagének baktériumba jutásának esélyét. Vagyis az efflux pumpák alapján akár előre is jelezhetjük a potenciálisan multidrog-rezisztenssé váló baktérium törzseket a jövőben, esélyt adva ezzel annak, hogy felkészülhessünk ellenük.

Öregedési folyamatok molekuláris mechanizmusainak elemzése értelmezhető gépi tanulással

Aurel Prosz¹, Pipek Orsolya¹, Börcsök Judit¹, Palla Gergely³, Spisák Sándor⁴, Szállási Zoltán¹, Csabai István²

¹Danish Cancer Society Research Center

²ELTE TTK Fizikai Intézet Komplex Rendszerek Fizikája Tanszék

³ELTE TTK Biológiai Fizika Tanszék

⁴TTK Enzimológiai Intézet

A DNS metilációs mintázat meglepő pontossággal tudja becsülni az életkort. Mivel a DNS metiláció epigenetikai vezérlő szerepet játszik biológiai folyamatokban, többen feltételezik, hogy nem csupán korrelációról van szó, hanem okozó tényező is lehet. A becslő módszert felfedező Steve Horvath egy interjúban a becslő konstrukció "eredendő bűnének" nevezi, hogy az egy olyan statisztikai regressziós modellen alapul, amely független a biológiától. Az értelmezhető mesterséges intelligencia (eXplanable AI) módszertanára alapozva kidolgoztuk a XAI-AGE módszert, mely a regressziós modellekkel azonos pontosságú viszont emellett részletes biológiai betekintést enged, hogy mely folyamatokon keresztül vezérelheti a DNS metiláció az öregedés sejtbioológiai folyamatait.

TmAlphaFold adatbázis: AlphaFold2 segítségével becsült transzmembrán fehérje szerkezetek membrán lokalizációja és értékelése

Dobson László^{1,2}, Levente I. Szekeres¹, András Zeke¹, Tusnady E. Gábor¹

¹TTK Enzimológiai Intézet Fehérje Bioinformatika Kutatócsoport

²EMBL Structural and Computational Biology Unit

A mesterséges intelligencia, és azon belül az AlphaFold2 (AF2) új fejezetet nyitott a fehérjeszerkezet becslés területén, megreformálva a szerkezeti biológia számos területét. Mivel a transzmembrán fehérjék hagyományos szerkezet becslési eljárásai bonyolultak és jelentős hibával terheltek voltak, az AF2 nagy segítséget jelent a tudományos közösség számára. A viszonylag csekély számú kísérletes szerkezetet kiegészítve az AF2 több ezer új alfa-helikális membránfehérje 3D becslését biztosítja. A megbízható szerkezeti sablonok/templátok hiánya és az a tény, hogy az AF2-t nem tanították a membrán és vizes közeg fázishatárán lévő transzmembrán fehérjék speciális tulajdonságainak a kezelésére, szükségessé teszi a becsült szerkezetek helyességének értékelését. A munkacsoportunk által fejlesztett Transmembrane AlphaFold adatbázisban (TmAlphaFold adatbázis, <https://tmalphafold.ttk.hu/>) a TMDet segítségével egy egyszerű geometria-alapú módszert alkalmaztunk a fehérjék membránban való elhelyezkedésének a becslésére. A legvalószínűbb membránsík becslésén felül számos módon értékeljük a szerkezeteket, azokat öt megbízhatósági kategóriába rendezve. Eredményeink alapján az AF2 átlagosan nagy pontossággal képes megbecsülni a transzmembrán fehérjék szerkezetét, de sikerült néhány tipikus hibát azonosítani, amely hibák azonosítása előkészíti az utat egy még pontosabb szerkezetbecslő eljárás elkészítéséhez.

A lineáris SVM-ek néhány szép alkalmazása gráfok és sorozatok elemzésében

Grolmusz Vince, Keresztes László, Szögi Evelin és Varga Bálint

ELTE Matematikai Intézet PIT Bioinformatikai Csoport

A mély neurális hálók számos sikere mellett sokszor elfeledkezünk a nagyon "sekély" neurális hálók, azaz a Support Vector Machine-ok (SVM) alkalmazásáról a gépi tanulásban. Néhány meglepő eredményt mutatunk peptidok amiloidképzési hajlandóságának becslésével és az agygráfok egyes tulajdonságok szempontjából fontos éleinek meghatározásával kapcsolatban.

Az immuneterápia hatékonyságának előrejelzése

Gyórfy Balázs^{1,2}

¹ TTK Enzimológiai Intézet Onkológiai Biomarker Kutatócsoport

² SE Bioinformatika Tanszék

Az immun-checkpoint inhibitorok ígéretes hatást mutatnak több daganattípusban. A jelenleg használt biomarkerek, mint például a PD-L1 expresszió és a tumor mutációs teher azonban nem biztosítanak eredményes betegretegezést. Célunk egy olyan adatbázis létrehozása volt, amely génexpressziós és klinikai válaszadatokat egyaránt tartalmaz, és lehetővé teszi az anti-PD-1, anti-PD-L1 és anti-CTLA-4 immunterápiákra adott válasz biomarkereinek azonosítását.

GEO-szűrést hajtottunk végre annak érdekében, hogy azonosítsuk azokat az adatkészleteket, amelyek egyidejűleg rendelkezésre állnak klinikai és transzkriptomikai adatok, függetlenül a rák típusától. A szűrés az anti-PD-1 (nivolumab vagy pembrolizumab), az anti-PD-L1 (atezolizumab vagy

durvalumab) vagy az anti-CTLA-4 (ipilimumab) szerek adásával kapcsolatos vizsgálatokra korlátozódott. ROC elemzést és Mann-Whitney tesztet végeztünk minden génen a terápiás válaszreakcióval kapcsolatos jellemzők azonosítására. Az új biomarker jelöltek további elemzésére és validálására szolgáló elemzési portált hoztunk létre, amely a www.rocplot.com/immune címen érhető el. Összefoglalva, egy adatbázist és egy webes platformot hoztunk létre az immunterápiás válasz biomarkereinek vizsgálatára szolid tumorminták nagy csoportjában. Eredményeink az immunterápiára alkalmas új betegcsoportok azonosításában is segítenek.

Membránfehérjék szerkezete az AlphaFold AI után

Hegedűs Tamás^{1,2}

¹SE Biofizikai és Sugárbiológiai Intézet

²ELKH-SE Biofizikai Virologia Kutatócsoport

Az elmúlt években számos transzmembrán fehérje szerkezetét határozták meg krio-elektronmikroszkópiával. Továbbá az AlphaFold2 (AF2) is nagyszámú, jó minőségű szerkezeti modellt szolgáltatott. Azonban, ha a vizsgálat tárgya egy adott fehérjecsalád, a tagok szerkezeteinek összegyűjtése kihívást jelent a meglévő adatbázisok és szerkezet-összehasonlító algoritmusok ellenére. Ezért megvizsgáljuk ABC (ATP Binding Cassette) fehérjék szerkezetének automatikus azonosításának és közzétételének lehetőségét a 3D-Beacons Hálózaton keresztül (<https://www.ebi.ac.uk/pdbe/pdbe-kb/3dbeacons>). Ezen szupercsalád példáján keresztül mutatjuk be az AF2 membránfehérjéken nyújtott teljesítményét is. Mivel számos ABC fehérje multimerként működik, így az egyláncú fehérjeszerkezeti predikciókat tartalmazó AlphaFold2 adatbázisban nem található meg, ezért az AF-Multimer segítségével meghatároztuk dimerként működő fontos humán ABC fehérjék szerkezetét és más membrán ATPázok fehérje-komplexeit. Predikcióink alapján az AF-Multimer határfoka a vártnál alacsonyabb. Eredményeink felhívják a figyelmet arra is, hogy egyes ABC fehérjékkel végzett kísérleteket szükséges átértelmezni és szerkezetüket kísérletesen meghatározni. A megbízható ABC szerkezeti modelljeinket sikerült a 3D-Beacons Hálózaton keresztül nagy adatbázisokhoz (pl. PDBE-KB) csatolni, ezáltal lehetővé téve széleskörű elérhetőségüket (<https://3dbeacon.hegelab.org>, <https://abc3d.hegelab.org>).

Fehérjék kötőhelyeinek fragmens alapú azonosítása

Keserű György Miklós

TTK Gyógyszerkémiai Kutatócsoport

A ligandumok fehérjekötődéséhez szükséges kölcsönhatástípusok (farmakofór tulajdonságok) és a fehérjék kötőhelyeinek elemzése alapján új protokollt dolgoztunk ki a fehérjék széles köre esetében eredményes fragmens gyűjtemények tervezésére. A SpotXplorer megközelítés során egy olyan kisméretű fragmens könyvtárat hoztunk létre, amely maximalizálja a fehérjék elsődleges kötőhelyein kísérletileg megerősített kölcsönhatási mintázatok lefedettségét. A megközelítés hatékonyságát hagyományos és kihívást jelentő fehérje célpontok ellen vizsgált kísérleti könyvtárral demonstráljuk. A SpotXplorer-szűréssel, mindössze 96 célszerűen megválasztott fragmens kötődésének vizsgálatával, a népszerű célfehérjék (G-fehérje kapcsolt receptorok, proteázok) ismert kölcsönhatási mintázatainak átlagosan 70%-át sikerült azonosítani, és a módszer találatokat adott a közelmúltban azonosított kihívást jelentő célpontok, köztük a SARS-CoV-2 fehérjék ellen.

Egy rács-alapú, malignus sejtek szimulációjára fejlesztett ágens alapú szoftver bemutatása

Kiss Dániel^{1,2}, Pintér Ádám¹, Füredi András², Szakács Gergely^{2,3}, Lovrics Anna²

¹ÓE NJIK Szoftvertervezés és -fejlesztés Intézet

²TTK Enzimológiai Intézet Gyógyszer Rezisztencia Kutatócsoport

³Medical University of Vienna Institute of Cancer Research

Tumorsejtek szimulációjára több különböző szoftver is elérhető, sok közülük nyílt forráskódú. Az egyes szimulációs rendszerek programnyelve, célközönsége, megoldandó feladatai eltérőek. Az előadásban egy általunk fejlesztett új rácsalapú szimulációs rendszert mutatok be, melyet malignus sejtek növekedésének szimulációjára hoztunk létre. Célunk, hogy a szoftver könnyen használható legyen biológiai hipotézisek tesztelésére, laborkísérletek in silico reprodukálására, modellparaméterek illesztésére. Az előadás során publikált számítógépes szimulációk és kísérletek reprodukálása segítségével bemutatom, hogy mire használható a szoftver.

Fiatal kutatók előadásainak kivonatai

A protein-arginin-metil-transzferáz-5 enzim foszforilációjának szerkezeti hatásai a H4 hisztonhoz való kötődésére

Börzsei Rita, Bayarsaikhan Bayartsetseg, Zsidó Balázs Zoltán, Lontay Beáta, Hetényi Csaba

PTE ÁOK Farmakológiai és Farmakoterápiai Intézet Farmakoinformatikai Tanszék

A protein-arginin-metil-transzferáz-5 (PRMT5) enzim különböző fehérjék, köztük a H4 hiszton arginin metilációjáért felelős. Ígéretes gyógyszercélpont, amely szerepet játszik számos betegség, mint hematológiai és szolid tumorok valamint a 2-es típusú diabétesz pathomechanizmusában. Kísérletesen igazolt, hogy a PRMT5 T80 aminosav foszforilációja növeli az enzim metil-transzferáz aktivitását, továbbá a foszforilált enzim emelkedett szintjét mérték több tumor, többek között hepatocelluláris carcinoma esetén is. Munkánk során megalkottuk a natív és a T80-foszforilált PRMT5-metiloszóma protein-50 (MEP 50) H4 hisztonnal alkotott komplex szerkezetét. A komplexek a H4 hiszton teljes egészében tartalmazzák, amelynek a megépítése in situ történt kísérletesen kimért hiszton-fragmens szerkezetek felhasználásával. Molekula dinamikai szimulációk valamint szerkezeti- és energetikai elemzések segítségével összehasonlítottuk a natív és a foszforilált PRMT5-MEP50-H4 enzimkomplexeket valamint az apo szerkezeteket is. Eredményeink atomi felbontású magyarázatot adnak a következő kísérleti tapasztalatokra: (1) Miért nagyobb a T80-foszforilált PRMT5 enzim aktivitása a natív formához képest? (2) A PRMT5 miért a szabad H4 hisztont és nem annak nukleoszómához kötött formáját foszforilálja? Eredményeink szerkezeti információkkal segíthetik új típusú PRMT5 gátlók tervezését és fejlesztését.

Az autofágia indukciójában kulcsfontosságú mTORC1-ULK1-PP2A szabályozási háromszög működésének rendszerbiológiai vizsgálata

Hajdú Bence, Holczer Marianna, Kapuy Orsolya

SE Molekuláris Biológia Tanszék

A sejtek önemésztésért felelős autofágia-függő túlélést fehérje kinázok és foszfatázok egyaránt szabályozzák. A sejtek belső homeosztázisát fenntartó mTORC1 az autofágia kulcsfehérjéjének, az ULK1-nek a foszforilálásával gátolja folyamatot, míg a PP2A foszfatáz képes defoszforilálni azt, ezzel elősegítve az autofágia aktiválódását. Az aktív ULK1 gátolja az mTORC1-et és aktiválja a PP2A-t, továbbá az mTORC1 képes gátolni a PP2A-t is. Elképzelésünk szerint a szabályozási hálózatban így kialakuló pozitív és dupla negatív visszacsatolási hurkok (mTORC1-|PP2A-|mTORC1, mTORC1-|ULK1-|mTORC1, ULK1->PP2A->ULK1) szükségesek az autofágiás és a nem autofágiás állapotok közötti bistabil kapcsolószerű elválasztásához. Vizsgálatunk során rendszerbiológiai megközelítést alkalmaztunk, mely során HEK293T sejteken végzett molekuláris biológiai kísérletek eredményeit matematikai modellek segítségével is vizsgáltuk, hogy jobban megértsük a szabályozási hálózat dinamikai viselkedését. A modell paramétereinek beállítását, illetve a modell validálását a sejtpopulációkon végzett kísérletekből (pl. mTORC1 gátlás, éhezés, PP2A gátlás vagy ULK1 csendesítés) nyert adatokkal végeztük. Megvizsgáltuk az mTORC1-ULK1-PP2A szabályozási háromszög dinamikai jellemzőit, feltételezve, hogy a visszacsatolási hurkok elengedhetetlenek a robusztus sejtválaszhoz különböző celluláris stressz esetén, illetve bizonyítottuk, hogy a PP2A gátlása (okadánsav hozzáadásával) képes autofágia oszcillációt indukálni.

Venezuelai ló-láz encephalitis vírus proteáz öninaktivált konformerének vizsgálata molekuladinamikai szimulációkkal

Hoffka Gyula^{1,2}, George T. Lontos³, Tózsér József¹, Mótyán János András¹

¹DE ÁOK Biokémiai és Molekuláris Biológiai Intézet, Retrovirális Biokémiai Kutató Laboratórium

²DE Molekuláris Sejt- és Immunbiológiai Doktori Iskola

³Basic Science Program, Frederick National Laboratory for Cancer Research

A számítógépes szimulációk lehetővé teszik az enzimek olyan tulajdonságainak vizsgálatát, amelyeket nehéz lenne kísérleti módszerekkel tanulmányozni. A molekuladinamikai módszerek segítségével vizsgálhatjuk a konformációs változásokat, a szubsztrát lehetséges kötődésének módjait és az interakciós hálózatokat is. Az ilyen szimulációkból származó nagy mennyiségű adat vizsgálatához hatékony algoritmusokra van szükség.

A ló-láz encephalitis vírus (VEEV) az emberek és az állatállomány súlyos betegségeinek okozója, ezért a vírus 2-es számú nem-szerkezeti fehérje proteáza (nsP2pro) igazolt gyógyszer-célpont a vírus élelciklusában betöltött fontos szerepe miatt. Együttműködő partnereink meghatározták a VEEV nsP2pro nagy felbontású (1,46 Å) kristályszerkezetét, benne egy nem várt konformációval, melyben az N-terminális köt a szubsztrát helyére. Ilyen öninaktivált konformációt korábban csak az N475A mutáns nsP2pro esetében figyeltek meg.

Az aktív és inaktív konformációk összehasonlításához az Amber szoftverrel végeztünk molekuladinamikai szimulációkat, szabad és szubsztrátot kötő enzimszerkezetek esetében is. A dinamikus hidrogénkötési hálózatok vizsgálatával megtaláltuk az öninaktivált forma kialakulásáért felelős főlánc kölcsönhatásokat, az N475A mutációtól függetlenül. Számításaink fontos különbségeket tártak fel az oligopeptid szubsztrátot kötő vad típusú és N475A mutáns kötési módjai között. Eredményeink hozzájárulhatnak az enzimstabilitás szerkezeti hátterének mélyebb megértéséhez.

A TKP2021-EGA-20 (Biotechnológia) számú projekt a Nemzeti Kutatási Fejlesztési és Innovációs Alapból biztosított támogatással, a TKP2021-EGA pályázati program finanszírozásában valósult meg.

A SARS-CoV-2 potenciális új variánsai által okozott betegség lefolyásának előrejelzése gépi tanulás segítségével mutációs adatok alapján

Kalcsevszki Regina, Pongor Sándor, Ligeti Balázs

PPKE ITBK

Számos tanulmány bizonyította, hogy a COVID-19 kimenetele és a betegségre való fogékonyság statisztikailag összefügg különböző társbetegségekkel, a beteg korával és közvetlenül vagy közvetve a vírus genomjában szereplő mutációkkal.

Munkánk során fejlesztettünk egy mutáció azonosító pipeline-t, valamint gépi tanulás segítségével egy olyan eszközt, ami a betegség kimenetelét képes előre jelezni a genom mintából és a beteg életkorából. A genom mintához — abban eredetileg nem, de a tanulóhalmazban szereplő — mutációkat adva, képesek vagyunk előre jelezni, új potenciális variánsok által okozott betegség kimenetelét.

A GISAID adatbázisából letöltött genom mintákon és a metaadatokon (pl. betegség kimenetele, életkor, ország, mintavétel időpontja) végzett alapos adattisztítás után, 97 ezer mintát dolgoztunk fel, melyekben 1122 gyakori mutációt azonosítottunk.

A végső tanítóhalmaz létrehozásához az adatokat stratifikáltuk a mintavétel időpontjára és országára. A feature halmazban a mutációk mellett a beteg életkora is szerepel. A végső modellt így 2680 enyhe, és 2680 súlyos lefolyású betegséget okozó minta alapján tanítottuk a JADBIO automatikus gépi tanuló eszköz segítségével. A modell előre jelző teljesítménye: ROC AUC 0.855 [0.811, 0.896].

A modell által kiválasztott featureök elemzése során, olyan mutációkat találtunk, melyek összefüggenek a súlyos lefolyással (pl.: NS3a P42L és Q57H, NSP4 A446V, NS8 T11I, NS10 V30L). Ezen mutációkat az omikron variánshoz adva a betegség súlyos lefolyásának predikált értéke 1,6-szorosára nőtt.

A HLA-I és HLA-II variánsok közötti kölcsönhatás befolyásolja a tumorelles immunválasz erősségét és időzítését melanómás betegeknel

Koncz Balázs, Papp Benjamin Tamás, Fülöp Anna Tácia, Balogh Gergő Mihály, Spekhardt Dóra, Manczinger Máté

SZBK Biokémia Intézet Rendszerimmunológiai Munkacsoport

A tumorelles immunválaszhoz alapvetően szükséges, hogy a tumorokban mutációk következtében megjelenő módosult peptideket a HLA-I és HLA-II molekulák bemutassák a T-sejteknek. Ezen HLA molekulák rendkívüli módon diverzek, és a különböző variánsokat eltérő peptidkötő specificitás jellemzi. Ugyan néhány kutatás bizonyítja a HLA diverzitás és bizonyos HLA variánsok hatását a tumorelles immunválaszra, ezek az eredmények ellentmondásosak. Mivel mind a HLA-I, mind a HLA-II molekuláknak alapvető szerepe van a tumoros sejtek felismerésében, a közös hatásukat vizsgáltuk melanómás betegekben. Számos HLA-I és HLA-II allélkombinációt találtunk, melyek pozitív (előnyös kombinációk) vagy negatív (káros kombinációk) összefüggést mutattak áttétes, ICB immunterápiát kapott melanómás betegek túlélésével. Továbbá ezek az allélkombinációk immunterápiát nem kapó, melanómás betegek túlélését is befolyásolták. A tumorminták sejtösszetételét megvizsgálva is nagy különbségeket kaptunk a káros, valamint az előnyös allélkombinációt hordozó betegekben. A káros allélkombinációk primer mintákban a citotoxikus immunválasz jeleit, míg áttétes mintákban az immun elkerülés jeleit, valamint egy hatástalan tumorelles immunitást mutattak. Ezzel szemben az előnyös kombinációt hordozó minták esetében csak az áttétek mutattak aktív citotoxikus immunválaszt. Eredményeink bizonyítják a HLA-I és HLA-II molekulák együttes hatását a tumorelles immunválaszra.

Vitális funkciók paraméterezése aktivitásmerő okoskarkötőkkel

Dr. Kothalawala William Jayasekara, Szigeti Péter, Dr. Gyórfy Balázs

SE Bioinformatikai Tanszék

A Semmelweis Egyetem Bioinformatika Tanszékén egy mesterséges intelligencián alapuló algoritmus fejlesztésébe kezdtünk, mely a későbbiekben alkalmassá válhat a vitális funkciók monitorozására.

E4 okoskarkötők (Empatica Inc.) segítségével gyűjtöttünk adatot egy nagyméretű kontrollcsoporttól. Az okoskarkötővel az alábbi paraméterek mérése lehetséges: pulzushullám görbe és szívfrekvencia-variabilitás (fotopletizmográfiás szenzor, PPG), testhőmérséklet (infravörös szenzor), mozgási adatok (3D accelerométer), elektrodermális aktivitás (EDA szenzor). A mért adatok nyers adat formájában kinyerhetőek az eszközről, mely így normalizálást követően feldolgozható. A mérések során a résztvevők releváns klinikai adatait is rögzítettük (életkor, nem, testsúly, testmagasság, fizikai

aktivitás mértéke, dohányzás, krónikus megbetegedések, vérnyomás). A méréssel egyidejűleg egy AliveCor KardiaMobile 6L hat elvezetéses hordozható EKG-val regisztrátumot készítettünk, melynek görbéje szintén exportálható nyersadatként. Az így mért adatokból különböző mesterséges intelligencián alapuló modellek fejlesztésére kerül sor, melyeket a jövőben betegcsoportokon is tesztelni tervezünk.

A mérési eszközök és a mérések menetének bemutatásán felül az előadást követően szeretnénk meginvitálni az érdeklődőket egy párbeszédre a fejlesztéssel kapcsolatban, illetve a kontrollcsoport bővítésében való részvételre.

SARS-CoV-2 PCR primerek célrégióiban megjelenő mutációk vizsgálata

Mentes Anikó, Papp Krisztián, Visontai Dávid, Stéger József, Csabai István, Medgyes-Horváth Anna, Pipek Orsolya Anna

ELTE TTK Fizikai Intézet Komplex Rendszerek Fizikája Tanszék

A SARS-CoV-2 genomjának folyamatosan növekvő számú mutációja miatt aggályok merültek fel a SARS-CoV-2 diagnosztikájának „arany standardja”, a reverz transzkripció-polimeráz láncreakciónak (RT-PCR) megbízhatóságával kapcsolatban. Kidolgoztunk egy olyan bioinformatikai vizsgálati módszert, amely a WHO által javasolt vagy a tudományos közösség által publikált, nyilvánosan elérhető PCR-tesztek célrégióival átfedő genomi változásokat tárja fel, kifejezetten a PCR tesztek hatékonyságára gyakorolt hatásuk alapján.

Nemzetközi együttműködésben, a VEO (Versatile emerging infectious disease observatory) projekt keretében több, mint 1,2 millió SARS-CoV-2 szekvenciában vizsgáltuk a célrégiókban előforduló mutációkat. Megközelítésünk különbséget tesz a PCR-hatékonyságot esetlegesen károsan befolyásoló és az ilyen értelemben várhatóan semleges mutációk között. A mintákat a jelen lévő mutációk becsült hatása alapján a "téves besorolásra hajlamos" és a "valószínűleg helyesen detektálható" kategóriákba soroljuk egy adott PCR-primer készlettel.

A téves besorolásra hajlamos minták általában napi 2%-os vagy annál alacsonyabb arányban vannak jelen, bár úgy tűnik, hogy bizonyos primer-készletek az Omicron vírustörzs kimutatásában gyengébb teljesítményt nyújtanak. Mivel a SARS-CoV-2 víruspopulációban világszerte átmenetileg különböző variáns törzsek válhatnak dominánssá, egy adott PCR-primer készlet hatékonysága idővel változhat, ezért a primerek célrégióiban található mutációk folyamatos nyomon követése erősen ajánlott.

A TNMplot.com bővítése: új analízisek lefuttatási lehetőségei és új grafikus megjelenítési formák bemutatása

Bartha Áron¹, Munkácsy Gyöngyi¹, Balázs Györfly^{1,2}

¹SE Bioinformatika Tanszék

²TTK Enzimológiai Intézet Lendület Onkológiai Biomarker Kutatócsoport

A tumorban vagy metasztatikus szövetekben magasabb expressziót mutató gének segíthetnek a daganatképződés jobb megértésében, miáltal a progresszió biomarkereiként vagy lehetséges terápiás célpontokként szolgálhatnak. A TNMplot.com oldal háttérében egy általunk összeállított, transzkriptom szintű adatkészletek felhasználásával készült adatbázis áll, amelyben egy online felhasználói felület révén valós időben összehasonlíthatjuk az összes gén normál, daganatos és metasztatikus adatait 22 tumortípusban. A teljes adatbázis 56.938 mintát tartalmaz - ebből 33.520 gén chip és 23.418 RNA-Seq

alapú -, amelyeket az NCBI-GEO, TCGA, TARGET, és GTEx adataiból építettük fel. A 2021-ben publikált verzióhoz képest (A. Bartha, B. Gyórfy, TNMplot.com: A Web Tool for the Comparison of Gene Expression in Normal, Tumor and Metastatic Tissues. *Int. J. Mol. Sci.* 2021, 22(5), 2622;) új analízisekkel és grafikus megjelenítési formákkal bővítettük a webes keresés lehetőségeit: a pan-cancer áttekintést egy heat mappal egészítettük ki, lehetőség van korrelációs mátrix készítésére (Spearman vagy Pearson) és enrichment analízisekre. A statisztikai szignifikancia kiszámítása Mann-Whitney vagy Kruskal-Wallis tesztekkel történik mind az RNA-Seq és gén chip adatok elemzésekor. Az előadásban ezek az újonnan telepített algoritmusok fognak bemutatásra kerülni.

GPCRdb, a G protein-kapcsolt receptorok információs adatbázisa

Pándy-Szekeres Gáspár^{1,2}

¹*Department of Drug Design and Pharmacology, University of Copenhagen*

²*TTK Szerveskémiai Intézet Gyógyszerkémiai Kutatócsoport*

Az FDA által jóváhagyott gyógyszerek ~34%-a hat valamilyen G protein-kapcsolt receptorra (GPCR)¹. Az emberi testben a legkülönbözőbb funkciókért felelnek, így megtalálhatók többek között az ideg-, endokrin-, immun- és keringési rendszerben. Továbbra is népszerű kutatási területnek számítanak, így nagy méretű tudományos közösség foglalkozik ennek a fehérje családnak a legkülönbözőbb aspektusaival. Egy átfogó információs adatbázisnak az igénye már bőven az első GPCR kristályszerkezet előtt felmerült, így a GPCRdb első verziója 1993-ban készült el. Azóta több megújuláson és folyamatos fejlesztésen keresztül ma havi több mint 7000 felhasználóval működik. Szekvencia, szerkezeti, mutációs, funkcionális és egyéb adatok egyesítése mellett, az adatok elemzésére alkalmas online eszközöket fejlesztünk felhasználóink számára. Legfrissebb munkákban aktivációs állapot specifikus AlphaFold modelleket és egy több adatbázist egységesítő GPCR ligandum adattárt hoztunk létre. A GPCR kutatások kiterjedtsége miatt partner molekulákra is kiterjesztettük a GPCRdb-t. 2021-ben publikáltuk az első testvér adatbázisunkat, a GproteinDb-t, valamint jelenleg publikálás alatt van a Biased Signaling Atlas (BSA) és az ArrestinDb.