

A MAGYAR
TUDOMÁNY
ÜNNEPE



Magyar Tudományos
Akadémia

Bioinformatika 2017

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

2017. november 10.

MTA Természettudományi Kutatóközpont
földszinti nagy előadó terem
1117 Budapest, Magyar Tudósok körútja 2.

Program

Délelőtti 1. szekció, elnök: Patthy László

10:00 – 10:05 **Köszöntő**

10:05 – 10:30 **Csabai István**

Eötvös Loránd Tudomány Egyetem, Természettudományi Kar, Fizikai Intézet,
Komplex Rendszerek Fizikája Tanszék

Kollaboratív keretrendszer genomikai adatelemzéshez

10:30 – 10:55 **Barta Endre**

NAIK Mezőgazdasági Biotechnológiai Kutatóintézet,
Mezőgazdasági Genomikai és Bioinformatikai Csoport

Transzkripciós faktorok és kofaktorok topológiájának vizsgálata a DNS-en CHIP-seq
adatok elemzésével

10:55 – 11:20 **Antal Péter**

Budapesti Műszaki és Gazdaságtudományi Egyetem, Villamosmérnöki és Informatikai Kar,
Méréstechnika és Információs Rendszerek Tanszék

*Bayesi rendszerszintű vizsgálatok a közös genetikai háttér felderítésére komplex
fenotípusok, betegségcsoportok és betegségek összessége esetén*

11:20 – 11:40 **Kávészünet**

Délelőtti 2. szekció, elnök: Simon István

11:40 – 12:05 **Fuxreiter Mónika**

Debreceni Egyetem, Biokémiai és Molekuláris Biológiai Intézet, MTA-DE Lendület
Fehérjedinamikai Kutatócsoport

A szerkezeti elmosódottság szerepe a fehérjék szerveződésében

12:05 – 12:30 **Mészáros Bálint**

Eötvös Loránd Tudományegyetem, Természettudományi Kar, Biokémia Tanszék,
MTA-ELTE Lendület Bioinformatika Kutatócsoport

Rendezetlen fehérjék és fehérjerégiók szerepe a rákban

12:30 – 13:10 **Ebédészünet**

Délutáni 1. szekció, elnök: Harrach Balázs

13:10 – 13:35 Pongor Sándor

Pázmány Péter Katolikus Egyetem, Információs Technológiai és Bionikai Kar
Molekuláris jelzések és a mikrobiális közösségek stabilitása

13:35 – 14:00 Gáspári Zoltán

Pázmány Péter Katolikus Egyetem, Információs Technológiai és Bionikai Kar
Térszerkezeti sokaságok szerepe a fehérjék szerkezet-hatás összefüggéseinek megértésében

14:00 – 14:25 Szöllősi Gergely János

Eötvös Loránd Tudományegyetem, Biológiai Fizika Tanszék
MTA-ELTE „Lendület” Evolúciós Genomika Kutatócsoport
Horizontális géntranszfer események mint molekuláris fossziliák

14:25 – 14:45 Kávészünet

Délutáni 2. szekció, elnök: Dosztányi Zsuzsanna

14:45 – 15:00 Bálint Mónika

Pécsi Tudományegyetem, ÁOK, Farmakológiai és Farmakoterápiai Intézet
Hogyan csomagoljunk be egy célpontot?

15:00 – 15:15 Fichó Erzsébet

MTA TTK Fehérjeszerkezet Kutatócsoport
Rendezett rendezetlenség

15:15 – 15:30 Dobson László

MTA TTK Membránfehérje Bioinformatika Kutatócsoport
Rendezetlen fehérjék komplexeinek vizsgálata szekvenciális, szerkezeti és funkcionális szempontból

15:30 – 15:45 Kerepesi Csaba

MTA SZTAKI, ELTE IK Informatika Doktori Iskola
Öregedést befolyásoló fehérjék predikciója gépi tanulással

15:45 – 16:00 Medgyes-Horváth Anna

ELTE-TTK Fizikai Intézet, Komplex Rendszerek Fizikája Tanszék
Humán mitokondrium haplotípus meghatározás és filogenetikai analízis szennyvízmintákból

16:00 – 16:15 Nagy Gergely

MTA-DE Lendület Immungenomikai Kutatócsoport
Makrofág transzkripció faktor kaszkádok vizsgálata ATAC-seq-vel

16:15 – 16:30 Pongor Lőrinc Sándor

Semmelweis Egyetem 2 sz. Gyermekklinika, MTA TTK Onkológiai Biomarker Kutatócsoport
Mintavételezés hatásának vizsgálata ovárium tumoros betegekben újgenerációs szekvenálással

16:30 – 16:45 Pipek Orsolya

ELTE TTK Fizikai Intézet, Komplex Rendszerek Fizikája Tanszék
Egyedi mutációk gyors és megbízható detektálása izogenikus mintákban

Előadások kivonatai

Antal Péter

Budapesti Műszaki és Gazdaságtudományi Egyetem, Villamosmérnöki és Informatikai Kar, Méréstechnika és Információs Rendszerek Tanszék

Bayesi rendszerszintű vizsgálatok a közös genetikai háttér felderítésére komplex fenotípusok, betegségcsoportok és betegségek összessége esetén

A kis hatáserősségű tényezők statisztikai felderítésére az egyik fő trend a mintaszám növelése, amely napjainkra már közel milliós esetszámot eredményezett például gyakori betegségek nagyszámú, gyenge genetikai faktorainak a vizsgálataiban. Egy másik megközelítés a célváltozók számának a növelésén alapul, amely például egy adott betegség esetén a fenotípus és a környezeti leírás részletességének növelését jelenti, de idetartozik a kérdéses betegség vizsgálatának kiterjesztése egy olyan betegségcsoportra, amelyek várhatóan átlapolódó genetikai háttérrel, azaz közös genetikai faktorokkal is rendelkeznek a célbetegséggel. Ezekben az esetekben nem csupán a faktorok interakcióit, hanem a célváltozók kapcsolatrendszerét is szükséges modellezni, amelyhez a Bayes-statisztikai megközelítés és a valószínűségi gráfok modellek egy hatékony keretet biztosítanak.

Az előadásban összefoglalom a keret elméleti háttérét, egy gyakorlati megvalósítását, illetve jelenlegi fejlesztési irányait, amelyek a szemantikailag összekötött, nyílt orvosbiológiai háttértudáshoz, mély tanulási módszerekhez és többcélpontú gépi tanulási módszerekhez kapcsolódnak. A megközelítés alkalmazását több probléma kapcsán is bemutatom, nevezetesen a pszichiátriai betegségek közös genetikai háttérének a vizsgálatán keresztül, gyógyszer-célpont interakció vizsgálatában, az egészséges öregedéssel kapcsolatos, korfüggő betegségekkel szemben általános védettséget jelentő genetikai faktorok kutatásán át, illetve a betegségek átfogó felhasználásán alapuló genetikai kutatásokban.

Barta Endre

NAIK Mezőgazdasági Biotechnológiai Kutatóintézet, Mezőgazdasági Genomikai és Bioinformatikai Csoport

Transzkripciós faktorok és kofaktorok topológiájának vizsgálata a DNS-en ChIP-seq adatok elemzésével

Már régóta ismert, hogy az eukarióta transzkripciós faktorok bizonyos szekvencia motívumokhoz kötődnek a DNS-hez. A transzkripciós faktorokhoz aztán más fehérjék, kofaktorok, represszorok, aktivátorok kapcsolódhatnak akár többszörösen is. Az újgenerációs szekvenálással egybekötött kromatin immunoprecipitáció (ChIP-seq) olyan kísérleti eljárás, amivel meg lehet határozni genom szinten, hogy hol kötődnek ezek a fehérjék a DNS-hez. A ChIP-seq adatok elemzése során csúcs (peak) régiókat határozunk meg. Ezeket tulajdonképpen azok a DNS darabok jellemzik, melyekhez a vizsgált fehérje kötődött. Az a pozíció a DNS-en ahol legtöbb ilyen fragmentet határoztunk meg (summit) az általános konszenzus szerint a DNS-hez kötődő, vagy ahhoz nagyon közel lévő fehérjemolekula által lefedett régió közepe.

A CTCF transzkripciós faktor és a hozzá kötődő kohezin fehérjék ChIP-seq eredményeit vizsgálva azt találtuk, hogy ezeknek a csúcspontjai a CTCF kötőhelyhez és egymáshoz képest meghatározott távolságban helyezkednek el. Hipotézisünk szerint ezek a távolságok az egyes fehérjemolekulák DNS-en való elhelyezkedését, topológiáját tükrözik. Ezek alapján egy új úgynevezett „double embrace” modellt állítottunk fel arra, hogy két CTCF-kohezin komplex hogyan alakít ki egy DNS hurkot.

A továbbiakban azt feltételeztük, hogy más transzkripciós faktoroknál, illetve transzkripciós faktor komplexeknél is fontos információkat nyerhetünk a ChIP-seq adatok elemzésével

ezeknek a fehérjemolekuláknak a topológiájáról a DNS-en. Kifejlesztettünk egy pipeline-t, amivel 1626 reprezentatív transzkripció faktor és 494 reprezentatív kofaktor humán ChIP-seq kísérletnél megállapítottuk, hogy az egyes kísérleteknél a csúcspont (summit) hogyan helyezkedik el a transzkripció faktor kötőhelyekhez képest. Előzetes adataink szerint nem csak a CTCF-kohezin komplex, hanem sok más transzkripció faktor és kofaktor esetében is ezek a vizsgálatok pontos információt szolgáltathatnak a transzkripció faktor és kofaktor komplexek topológiájáról a DNS molekulán. Ez pedig lehetőséget ad, hogy jobban megértsük például azt, hogy a komplex kötőhelyeken hogyan működnek együtt a különböző fehérjemolekulák.

A későbbiekben azt tervezzük, hogy ezeket az eredményeinket egy ChIPsummitdb adatbázisba rendezzük és kereshetővé tesszük a weben keresztül. Az elemzéseket pedig kiterjesztjük az összes elérhető humán és egér ChIP-seq kísérletre, ami lehetővé teszi majd az evolúciós összehasonlításokat is.

Csabai István

Eötvös Loránd Tudomány Egyetem, Természettudományi Kar, Fizikai Intézet, Komplex Rendszerek Fizikája Tanszék

Kollaboratív keretrendszer genomikai adatelemzéshez

Számos tudományterület hihetetlen gyorsasággal alakul át az új műszereknek köszönhetően. Talán mind közül a genomika az ahol leggyorsabban, a Moore törvényt messze meghaladó sebességgel fejlődik a technológia. Az adatbőség a korlátlanul tűnő lehetőségek mellett kihívásokat is hoz: hogyan lehet hatékonyan megosztani, elemezni az adatokat, hogyan tudnak kollaborálni informatikusok és biológusok. Egy fejlesztés alatt álló keretrendszert mutatok be melynek célja az adatmegosztás, a kollaboratív elemzés és a reprodukálható kutatás támogatása.

Fuxreiter Mónika

Debreceni Egyetem, Biokémiai és Molekuláris Biológiai Intézet, MTA-DE Lendület Fehérjedinamikai Kutatócsoport

A szerkezeti elmosódottság szerepe a fehérjék szerveződésében

A fehérjék működése általában egy meghatározott térszerkezethez rendelhető. A fehérjék kölcsönhatásainak alkalmazkodásában azonban lényeges szerepet kapnak az adaptív konformációváltások, melyek dinamikus szerkezeti sokaságokat igényelnek. Ezek átalakulási számos regulációs folyamat kulcspontjai. Csoportunk ezen változások molekuláris mechanizmusaival és kölcsönhatásainak leírásával foglalkozik. Fuzziness-nek, azaz elmosódott szerkezeteken alapuló szabályozásnak nevezzük, amikor a szerkezeti heterogenitás és a biológiai hatás között közvetlen összefüggés található. Létrehoztunk egy elmosódott szerkezetű fehérjekomplexeket tartalmazó adatbázist (FuzDB <http://protdyn-database.org>), amely a bioinformatikai elemzések alapjául szolgál. Kimutattuk, hogy a fehérjék szupramolekuláris szerveződéseinek kialakulásában, működésében és szabályozásában az elmosódottságnak alapvető szerepe van.

Gáspári Zoltán

Pázmány Péter Katolikus Egyetem, Információs Technológiai és Bionikai Kar

Térszerkezeti sokaságok szerepe a fehérjék szerkezet-hatás összefüggéseinek megértésében

A fehérjék szerkezet-hatás összefüggéseinek részletes feltárása elengedhetetlen a gyógyszerfejlesztés és a molekuláris biotechnológiai alkalmazások szempontjából. Az elmúlt évtizedekben egyre inkább világossá vált, hogy ezen összefüggések megértése nem lehetséges a fehérjemolekulák belső mozgásainak atomi szintű leírása, megértése nélkül. Ehhez egy lehetséges megközelítés a különböző időskálájú belső mozgásokra vonatkozó

kísérleti paramétereiket tükröző, és ezáltal a dinamikát is modellező térszerkezeti sokaságok előállítására és jellemzésére. Előadásomban áttekintem az ilyen sokaságok létrehozásának és elemzésének főbb módszereit, és példák segítségével bemutatom azok alkalmazhatóságát.

Mészáros Bálint, Zeke András, Reményi Attila, Dosztányi Zsuzsanna

Eötvös Loránd Tudományegyetem, Természettudományi Kar, Biokémia Tanszék, MTA-ELTE Lendület Bioinformatika Kutatócsoport

Rendezetlen fehérjék és fehérjerégiók szerepe a rákban

Az elmúlt évek kutatásai rávilágítottak, hogy az emberi genom által kódolt fehérjék jelentős része részben vagy teljes egészében rendezetlen. A rendezetlen fehérjék, illetve rendezetlen fehérje régiók (Intrinsically Disordered Regions – IDRs) anélkül tudnak kritikusan fontos biológiai funkciókat ellátni, hogy rendelkeznének egy jól meghatározott 3 dimenziós szerkezettel. A szerkezet hiánya és az ebből fakadó nagyfokú flexibilitás teszi lehetővé az olyan IDR-specifikus funkciók ellátását, mint például a domének flexibilis összekötése, az entrópikus láncok, vagy a membrán nélküli sejtalkotó elemek molekuláris szervezése. Ezek mellett a rendezetlen régiók gyakran tartalmaznak rövid lineáris szekvencia motívumokat, melyeken keresztül specifikusan képesek kölcsönhatni partner fehérjékkel. Az így kialakított kölcsönhatások lehetővé teszik többek között a poszt-transzlációs módosítások, a fehérje lokalizáció, a supramolekuláris komplexek összeszerelésének és a fehérjék lebontásának finoman hangolt irányítását. A motívumokon, mint kompakt funkcionális egységeken keresztül az IDR-ek központi szerepet játszanak a jelátvitelben és a szabályzási folyamatokban.

A rendezetlen régiókat tartalmazó fehérjék biológiai szerepeinek felismerésével párhuzamosan az IDR-ek és különböző betegségek, különösképpen a rák kapcsolata is egyre intenzívebben kutatott terület lett. A fehérje rendezetlenség és a rák pontos kapcsolata azonban nehezen feltérképezhető, mivel mind a szerkezeti biokémiai, mind a tumor szekvenálási kutatások is klasszikusan a rendezett fehérjékre fókuszáltak. Az elmúlt években publikált, nagyskálás genomi szekvenálásból származó eredmények azonban végre lehetőséget kínálnak az IDR-ek tumorigenezisben betöltött szerepének szisztematikus feltérképezésére. A COSMIC adatbázisban található több millió mutáció alapján statisztikai analízisekkel meghatároztuk a humán proteom azon IDR-jeit, melyek szignifikánsan mutálódnak és ezen keresztül járulnak hozzá a rák kialakulásához. Az azonosított IDR-ek mechanisztikus és funkcionális vizsgálata alapján közelebb kerülünk a fehérje rendezetlenség és a rákos megbetegedések közötti összefüggések megértéséhez. Ez a megközelítés egyben arra is rávilágít, hogy a rákkutatásban gyakran használt gén-centrikus képet le kell váltania a (fehérje) funkcionális régió centrikus megközelítésnek. Ez a koncepcionális váltás nem csak a rák teljesebb szintű megértéséhez elengedhetetlen, de új célzott gyógyszerészeti taktikák kifejlesztéséhez is alapvető fontosságú.

Pongor Sándor, Ligeti Balázs, Juhász János és Jády Attila

Pázmány Péter Katolikus Egyetem, Információs Technológiai és Bionikai Kar

Molekuláris jelzések és a mikrobiális közösségek stabilitása.

A sok fajt tartalmazó, multispecies bakteriális közösségek (MBK) sejtjei diffundáló kémiai jelanyagok révén kommunikálnak és kooperálnak egymással, ugyanakkor versengenek egymással a hely és egyéb erőforrások elnyeréséért. Ilyen közösségeket alkotnak az tengerfenék mikrobiális szőnyegei, az állatok bélfloájára vagy akár a talaj mikroba közösségei. A kommunikáció egyik legjobban ismert mechanizmusa az ún. quorum sensing (QS), mellyel a baktériumok populációja szinkronizálja géneik működését és így kollektív viselkedést tudnak kifejteni. Az MBK-k sok esetben stabilisak és ellenállnak a környezeti változásoknak is.

In silico modellek tanúsága szerint a környezetbe kibocsátott kémiai jelek hatására a baktériumok adott helyen fognak összegyűlni, és hogy a jelek kölcsönös megértése hozzájárulhat az közösség stabilitásához. Ugyanakkor, a potenciálisan támadó baktériumfajok jeleinek egyoldalú észlelése, „lehallgatása” révén egy közösség távol tudja tartani a támadót a saját, korlátozott erőforrásaitól. A több faj által közösen értett (detektált) jelzések révén a közösség metabolikus repertoárja is stabilabb lesz, így a közösség jobban tud alkalmazkodni az ökológiai változásokhoz.

Szöllösi Gergely János

Eötvös Loránd Tudományegyetem, Biológiai Fizika Tanszék, MTA-ELTE „Lendület” Evolúciós Genomika Kutatócsoport

Horizontális géntranszfer események, mint molekuláris fossziliák

A genetikai szekvenálás feltűnésével a rokon szekvenciákból génfákat rekonstruáló molekuláris filogenetika zsákutcába került. A megoldatlan kérdések megválaszolása helyett az új genomok megismerése régi vitákat szított fel. Világos, hogy a génfák nem fajfák, hanem mindegyik génfa evolúciós események sorozatának egyedi következménye. Ugyanakkor, ha különbségeiket egy közös fajfa kontextusában modellezzük, a génevolúciót és a fajok diverzifikációját érintő, hihetetlenül gazdag információhoz is hozzáférünk, mely hagyományos módszerekkel nem lenne elérhető. Például horizontális géntranszfer (HGT) kizárólag egyidejűleg létező fajok között fordulhatott elő, tehát a fajképződés időbeli sorrendjéről hordoz információt. Ezekkel a módszerekkel nyitott evolúciós kérdésekre próbálunk választ adni: az egyik ilyen, hogy HGT eseményeket „molekuláris fossziliaként” használva megoldjuk a mikrobiális evolúció időrendiségével kapcsolatos dilemmákat, a Föld történetével összefüggésben is. Utóbbi esetében az elérhető fossziliák száma igen szűkös határok közé szorítja a molekuláris kormeghatározó módszerek használatát.

Fiatal kutatók előadásainak kivonatai

Bálint Mónika, Hetényi Csaba

Pécsi Tudományegyetem, Általános Orvostudományi Kar, Farmakológiai és Farmakoterápiai Intézet

Hogyan csomagoljunk be egy célpontot?

A gyógyszertervezés fontos lépése a célpont-ligandum komplexek atomi felbontású szerkezetének előállítása. Mind a kísérletes technikák, mind az elméleti módszerek korlátokba ütköznek e szerkezetek előállítása során, amennyiben a célpont felszínén a kötőhely hozzátétőleges helyzetét sem ismerjük és/vagy több, esetleg gyengén kötő zsebet keresünk. Ugyanakkor ezek az esetek igen fontosak lehetnek a gyakorlatban. A többszörös, vagy alloszterikus kötődés ugyanis számos betegség pathomechanizmusában kulcsszerepet játszik. Az is gyakran előfordul, hogy csupán a célpont-ligandum kölcsönhatás megléte ismert, de nincs szerkezeti információ róla. Az előadásban egy új, szisztematikus eljárást ismertetünk e kérdések kezelésére.

Dobson László, Fichó Erzsébet, Simon István, Tusnády Gábor és Mészáros Bálint

MTA TTK Membránfehérje Bioinformatika Kutatócsoport, MTA TTK Fehérjeszerkezet Kutatócsoport, MTA – ELTE Lendület

Bioinformatika Kutatócsoport

Rendezetlen fehérjék komplexeinek vizsgálata szekvenciális, szerkezeti és funkcionális szempontból

A rendezetlen fehérjék/szakaszok nem rendelkeznek időátlagban állandó harmadlagos szerkezettel, hanem több alternatív konformáció közötti gyors, folyamatos átalakulásban vannak. Nagyfokú flexibilitásuknak köszönhetően számos funkciót képesek ellátni: molekuláris felismerésben vesznek részt (azáltal hogy más molekulákat vagy ligandumokat képesek megkötni), különböző régiók között flexibilis kapcsolatot biztosítanak vagy lineáris motívumok segítségével hatnak kölcsön specifikus doménekkel stb. Az esetek többségében ezek a fehérjék is képesek rendezett szerkezet kialakítására, amennyiben (pl. egy partnerrel) való kölcsönhatás segíti egy stabilabb állapot létrehozását.

A fehérje-fehérje kölcsönhatásokat osztályozhatjuk az alapján, hogy a natív monomer állapotú partnerek milyen szerkezet kialakítására képesek. Feltéve, hogy két fehérje kölcsönhatásáról beszélünk, a következő kölcsönhatás típusokat különböztethetjük meg: két rendezett fehérje kölcsönhatásakor mindkét fehérje képes önállóan feltekeredni és a kölcsönhatás ezután alakul ki; egy rendezett és egy rendezetlen fehérje kölcsönhatásakor a rendezett fehérje szerkezete templátként szolgál a rendezetlen fehérjének, amely így képes stabilizálódni; két rendezetlen fehérje kapcsolódása esetén a partnerek önállóan nem képesek stabil szerkezetet kialakítani, de a kölcsönhatás segíti egy stabil egység kialakítását.

A fehérjék feltekeredése és más fehérjékkel való kölcsönhatása hasonló biofizikai háttérrel rendelkezik, ugyanazokkal a fő erőhatásokkal, ezért a háromféle kölcsönhatás tárgyalható egységiesen. Ez megmutatja, hogy különböző szinteken (szekvenciális, szerkezeti és funkcionális) hogyan tükröződik a szerkezet kialakulásának háttere (azaz a feltekeredés és a fehérje-fehérje kölcsönhatások viszonya). Ezen tulajdonságok szisztematikus vizsgálata alapján létrehoztuk a különböző típusú kölcsönhatásoknak egy új leírását, külön részletezve a korábban - főleg adatok hiányában - nem vizsgált rendezetlen-rendezetlen kölcsönhatások típusait.

A rendezett-rendezetlen komplexek célzott vizsgálata mára elvezetett a terápiás alkalmazásig, úgy azon folyamatok, amik mögött rendezetlen-rendezetlen komplexek állnak, akkor lesznek a jövőben célozhatóak farmakológiai és biotechnológiai szempontból, ha tisztában vagyunk az ilyen kölcsönhatások alapvető tulajdonságaival. A bemutatott munka jelenti az első lépést ebben az irányban.

Fichó Erzsébet, Simon István, Mészáros Bálint

MTA TTK Fehérjeszerkezet Kutatócsoport

Rendezett rendezetlenség

A rendezetlen fehérjék nem rendelkeznek időben állandó, 3 dimenziós szerkezettel. E tulajdonságuk a rendezett fehérjéktől markánsan különböző molekuláris taktikákon keresztül számos létfontosságú biológiai funkció ellátására teszi őket képessé. Szerepet játszanak többek között a transzkripcióban, a sejtválaszban, különböző jelátviteli és regulációs folyamatokban, mely funkciók betöltése közben gyakran más fehérjékkel hatnak kölcsön. A rendezett fehérjék által kialakított kölcsönhatásokat jól ismerjük, viszont azok a kölcsönhatások, amelyeket rendezetlen fehérjék alakítanak ki egymással, vagy más rendezett partnerekkel, jelenleg kevésbé ismertek. Ugyanakkor az eddig ismert példák is bizonyítják, hogy az ilyen típusú kölcsönhatások rendkívül fontosak, hiszen kapcsolatba hozhatóak a különböző daganatok kialakulásával, a hoszt-patogén kölcsönhatásokkal vagy esszenciális gének transzkripció szabályozásával.

A rendezetlen fehérjék által kialakított kölcsönhatások vizsgálatának egyik hiányzó láncszeme volt a megfelelő méretű és minőségű adatszett. Doktori munkám fókuszpontjában a két vagy több rendezetlen monomerből felépülő fehérjekomplexek állnak, melyek kialakulása során a résztvevő fehérjék a kölcsönhatás során kölcsönösen egymást rendezik (Mutual Synergistic Folding - MSF).

Az általunk létrehozott Mutual Folding Induced by Binding (MFIB) adatbázis több, mint 200 MSF fehérjekomplexet számlál, és tartalmazza ezen fehérjék szerkezeti és funkcionális annotációját is. A rendezetlen fehérjék kölcsönhatásainak elemzéséhez azonban a rendezett és rendezetlen monomerekből felépülő komplexek vizsgálata is szükséges. Így az MFIB-bel összefüggésben elkészítettünk egy másik adatbázist is: a Database of Disordered Binding Sites (DIBS)-t is, mely olyan fehérje komplexeket tartalmaz, ahol csupán egyetlen lánc rendezetlen, mely a rendezett láncokhoz kötődve rendeződik (coupled folding and binding).

Az összegyűjtött adatok immár elegendőek ahhoz, hogy a rendezetlen fehérjék által kialakított kölcsönhatások szekvenciális, szerkezeti és funkcionális elemzése megtörténhessen. Ezen vizsgálatok kiterjesztik a fehérje-fehérje kölcsönhatásokról kialakított képünket, elvezetnek a különböző jelátviteli mechanizmusok mélyebb megértéséhez, és jövőbeli farmakológia vizsgálatokhoz is kiindulópontként szolgálhatnak.

Kerepesi Csaba

MTA SZTAKI, ELTE IK Informatika Doktori Iskola

Öregedést befolyásoló fehérjék predikciója gépi tanulással

Ma körülbelül 300 fehérjéről gondoljuk, hogy befolyásolhatja az emberi öregedési folyamatot. Míg az egyes fehérjék külön-külön nagy figyelmet kapnak, az öregedéssel kapcsolatos fehérjék egésze mint komplex rendszer kevésbé kutatott. Mi gépi tanulási módszereket felhasználva találtunk egy modellt, amely könnyen értelmezhető módon megmondja mi különbözteti meg az öregedéssel kapcsolatos fehérjéket a többi fehérjétől, illetve prediktáltunk olyan új jelölteket, amelyekről eddig nem volt ismert, hogy befolyásolnák az öregedési folyamatot.

Medgyes-Horváth Anna, COMPARE Konzorcium

ELTE-TTK Komplex Rendszerek Fizikája Tanszék

Humán mitokondrium haplotípus meghatározás és filogenetikai analízis szennyvízmintákból

A COMPARE Konzorcium azzal a céllal jött létre 2014-ben, hogy a genomika eszközeivel a fertőző betegségek felismerését felgyorsítsák, ezzel a járványok kitorését előre jósolhassák, káros hatásukat mérsékelhessék. Az újgenerációs szekvenálások eszköztárát felhasználva

számos környezeti mintát vizsgálnak a projekt keretében. Egy alprojekt a konzorciumon belül a világ szennyvízhálózatának felmérését tűzte ki célul (Global Sewage Surveillance Project), a patogén mikroorganizmusok detektálásán keresztül. A projektben jelenleg több mint 60 ország 75 városából származó szennyvízminta DNS-szekvenálása történt meg és további minták gyűjtése és analízise is folyamatban van. Ugyan a pályázat eredeti célja a patogén mikroorganizmusok azonosítása, ezek a szennyvízminták ezen felül is rengeteg információt tartalmaznak. Vizsgálatunk célja a humán mitokondriumok előfordulásának detektálása a szennyvízben, valamint a detektált mitokondriumok haplotípusának meghatározása és evolúciójuknak rekonstruálása volt. A vizsgált minták közül 20 esetben tapasztaltunk megfelelő lefedettséget, ami alapján az adott város mintájához tartozó konszenzus szekvenciát is meg tudtuk határozni, valamint ez alapján a mitokondrium haplotípusát is. A kapott eredmények a legtöbb esetben összhangban voltak az adott környezetre jellemző haplotípussal. Ezen minták alapján egy filogenetikai fát is készítettünk, melyen a kontinensek szétválása jól megfigyelhető. Vizsgálatunk továbbá felhívja arra is a figyelmet, hogy a különböző célokból gyűjtött minták bioinformatikai analizisével számos esetben mennyi értékes további információ is kinyerhető, melyek nem kapcsolódnak szervesen az eredeti célokhoz.

Nagy Gergely

MTA-DE Lendület Immunogenomikai Kutatócsoport

Makrofág transzkripció faktor kaszkádok vizsgálata ATAC-seq-vel

A gének átírását milliányi szabályozó DNS-szakasz befolyásolhatja. Ezek lehetnek serkentő vagy gátló hatással a génexpresszióra, emellett működhetnek horgonyzó elemekként is, meghatározva a kromatin szerkezetét és ezáltal hatást gyakorolva a polimerázok működésére. E régiókban a transzkripció faktorok (TF-ok) a rájuk jellemző rövid DNS-motívumokat kötik, majd további szabályozó fehérjéket toboroznak, melyek komplexe áthidalja a promóter-enhanszer távolságokat; az inzulátorok által kihorgonyzott kohezin gyűrűk pedig segítenek a kromatinhurkok állandósításában. A különféle sejtekben más-más motívumok játszanak szerepet, szoros együttműködésben a sejtre jellemző TF-okkal.

Az ATAC-seq egy olyan, akár egyetlen sejten is alkalmazható újgenerációs szekvenálási eljárás, amellyel megállapítható, mely kromatin régiók férhetőek hozzá a bevitt transzpozázok által. Ezek a régiók nincsenek szorosan nukleoszómákba csomagolva, hanem más kromatin-fehérjék, elsődlegesen TF-ok foglalnak rajtuk helyet, függetlenül attól, hogy promóterek, enhanszerek, silencer-ek vagy inzulátorok. Az enhanszerek és silencer-ek motívumai és az azokat kötő TF-ok nagy változatosságot mutatnak a különböző sejt típusok között, az ATAC-seq ezáltal jól alkalmazható a sejtek differenciációjának vizsgálatára.

Kutatócsoportunk legfőképp a makrofágok transzkripció szabályozásával foglalkozik, és több egyéb modellt is használunk e sejtek különböző tényezőkhöz való alkalmazkodásának tanulmányozására. Az egyik ilyen modell az interleukin-4 hatására történő alternatív makrofág polarizáció; másik fő célunk pedig a regenerálódó izom makrofágjainak cisztrómszintű megfigyelése. E folyamatok ATAC-seq-vel való időbeli követése lehetővé teszi a TF-ok szintjén történő jelátviteli útvonalak elfogulatlan feltérképezését, az eközben érintett célgéncsoportok segítségével pedig megállapítható, hogy a sejtek milyen funkcionális változásokon mennek keresztül. A vizsgált folyamatokban szerepet játszó, eddig kevésbé ismert vagy ismeretlen TF-ok azonosítása hasznos darabokkal szolgálhat ahhoz a mozaikhoz, amelynek kirakása közelebb visz a sejtek működésének és átalakulásának megismeréséhez, és ezáltal a bennük történő esetleges rendellenességek kijavításához is.

Pipek Orsolya

Eötvös Loránd Tudomány Egyetem, Természettudományi Kar, Fizikai Intézet, Komplex Rendszerek Fizikája Tanszék

Egyedi mutációk gyors és megbízható detektálása izogenikus mintákban

A szomatikus mutációk detektálása az új-generációs DNS-szekvenálási eljárások egyik fő célkitűzése. Számos laboratóriumban végzett kísérlet irányul a spontán, illetve a környezeti hatások által okozott mutációs folyamatok pontosabb megismerésére. Az ezek keresésére leggyakrabban használt algoritmusok jelentős része azonban nem több minta együttes tesztelésére lett tervezve, továbbá nem emberi DNS-ek vizsgálata esetén eredményeik kevésbé megbízhatóak. Az irodalomban megjelent kutatások tanúsága szerint a legtöbb kísérlet kiértékelése során az alkalmazott eszközön túl további, egyedileg meghatározott és rosszul dokumentált szűrőkre is szükség van, mely megnehezíti az eredmények reprodukálását, illetve összetettségét. Szerényebb kapacitású számítógépeken pedig bizonyos algoritmusok számítási ideje sok vizsgálni kívánt minta esetén annyira megnő, hogy az a módszert alkalmazhatatlanná teszi.

Ezeknek a problémáknak az áthidalására terveztük az IsoMut nevű eszközt, melynek használata minden olyan kísérleti elrendezésben praktikus, amikor egyedi, a kísérlet során létrejövő mutációkat szeretnénk detektálni több, alapvető genetikai tulajdonságaikban megegyező (izogenikus), de különböző kezeléseknél kitétt mintában.

Az eszköz tesztelésére harminc minta felhasználásával megbízható, ún. mutációs tesztalmozokat állítottunk fel, melyek segítségével a szűrési paraméterek értékeit egy szigorú optimalizációs eljárással határoztuk meg. Megfelelő beállításokkal az IsoMut a hagyományos eszközökhöz képest csökkenteni tudja a fals pozitívok számát, emellett több mint százszor olyan gyorsan lefutatható, mint a napjainkban használt legpontosabb detektáló algoritmus. Az eszköz emellett hatékonyan alkalmazható aneuploid sejtvonalakból származó DNS-mintákon is, mivel a releváns szűrési paramétereket dinamikusan, a lokális ploiditás függvényében változtatja.

Az IsoMutot több, nemrég publikált tanulmány is sikerrel alkalmazta egyedi, kezelés okozta mutációk megbízható detektálására, rendkívül kevés fals pozitív találattal. Ezek a publikációk fontos lépéseket tesznek a különböző vegyszerek, illetve genetikai elváltozások mutagenikus hatásainak feltérképezése felé.

Pongor Lőrinc Sándor, Munkácsy Gyöngyi, Vereczkey Ildikó, Pete Imre, Gyórfy Balázs

MTA TTK Onkológiai Biomarker Kutatócsoport, Semmelweis Egyetem 2 sz. Gyermekklinika, Országos Onkológiai Intézet

Mintavételezés hatásának vizsgálata ovárium tumoros betegekben újgenerációs szekvenálással

Bevezetés. A tumorok növekedése és fejlődése során genetikai változások halmozódnak fel, amelyek hatására új klónok jelenhetnek meg. Több vizsgálat foglalkozott a szolid tumorok térbeli és időbeli heterogenitásával, mely során igazolódott, hogy dinamikusan változik egy tumor genetikai összetétele

Célkitűzés. Célunk a tumorból izolált minta helyének és méretének hatását megvizsgálni újgenerációs szekvenálás során a mutációk detektálására.

Módszerek. Öt ovárium tumoros betegben a daganatos szövet 3 régiójából történt a minták gyűjtése. A szekvenáláshoz a DNS izolálást a következő részekre bontottuk: 1) egy biopszia, 2) biopszia körül több darab, 3) minden tumorrégióból (globális). Valamennyi izolált DNS mintából teljes exom szekvenálást végeztünk.

Eredmények. Az egyes régiókban azonosítható mutációk száma minden esetben lényegesen különbözött. A kópiaszám változások >96%-ban egyeztek meg az egyes régiók között. Négy beteg esetén a tumorminta méretének növekedésével nőtt az azonosítható mutációk száma. Egy beteg esetén az egy biopsziás mintában volt a legtöbb mutáció azonosítható, amely háttérben a tumor hipermutációs fenotípusa állhat. Három beteg esetén azonosítottunk

