

**Funkcionális genomikai vizsgálatok során alkalmazható
kísérletes eszközök fejlesztése és jellemzése**

Ozgyin Lilla

Témavezető: Dr. Bálint Bálint László, MD, PhD



DEBRECENI EGYETEM

MOLEKULÁRIS SEJT- ÉS IMMUNBIOLÓGIA DOKTORI ISKOLA

DEBRECEN, 2019

Funkcionális genomikai vizsgálatok során alkalmazható kísérletes eszközök fejlesztése és jellemzése

Értekezés a doktori (PhD) fokozat megszerzése érdekében
az elméleti orvostudományok tudományágban

Írta: **Ozgyin Lilla**
okleveles molekuláris biológus

Készült a Debreceni Egyetem Molekuláris Sejt- és Immunbiológia doktori iskolája keretében

Témavezető: Dr. Bálint Bálint László, MD, PhD

A doktori szigorlati bizottság:

Elnök: Prof. Dr. Szabó Gábor, az MTA doktora
Tagok: Prof. Dr. Bay Péter, az MTA doktora
Dr. Wiener Zoltán, PhD

A doktori szigorlat helyszíne és időpontja:

Debreceni Egyetem ÁOK, Biokémiai és Molekuláris Biológiai Intézet, In Vitro
Diagnosztikai Tömb 3. emelet, Genomi Medicina és Bioinformatikai
Szolgáltató Laboratórium Tárgyaló szobája
2019. december 9. 11:00

Az értekezés bírálói:

Prof. Dr. Balázs Margit, az MTA doktora
Dr. Bodai László, PhD

A bírálóbizottság:

Elnök: Prof. Dr. Szabó Gábor, az MTA doktora
Tagok: Prof. Dr. Balázs Margit, az MTA doktora
Dr. Bodai László, PhD
Prof. Dr. Bay Péter, az MTA doktora
Dr. Wiener Zoltán, PhD

Az értekezés védésének helyszíne és időpontja:

Debreceni Egyetem ÁOK, Belgyógyászati Intézet „A” épület tanterme
2019. december 9. 13:00

BEVEZETÉS

Az elmúlt évtizedben a funkcionális genomika, valamint új tudományos együttműködési módszerek felemelkedésének lehettünk szemtanúi. Az omikai módszerek gyors ütemű fejlődésével a klinikai kutatások is fokozatosan a nagy áteresztőképességű módszerek felé fordulnak, teljes genom szintű betekintést nyújtva a génkifejeződés szabályozásába patológiás kontextusokban. Az annotált biológiai mintákat illetve nukleinsav szekvenálási adatokat archiváló publikus tárhelyek lehetőséget biztosítanak a meglévő erőforrások hatékonyabb felhasználására. A poszt-genomi kor orvosbiológiai kutatásaival összefüggésben tehát az elterjedőben lévő új modellrendszerek jellemzése, a kulcsfontosságú funkcionális genomikai módszerek továbbfejlesztése, valamint a költségracionalizálás különös jelentőséggel bírnak.

A transzkriptomika és az epigenomika tudományterületeket a nagy áteresztőképességű nukleinsav profilozási módszerek, például az RNS-szekvenálás (RNA-Seq) és a szekvenáláshoz kapcsolt kromatin immunprecipitáció (ChIP-Seq) elterjedése hívta életre. Az RNA-Seq jelentős limitációi közé tartozik a kiindulási minta minőségére való nagyfokú érzékenység, amely többnyire a nem megfelelő mintakezelés, -tárolás vagy -szállítás következménye. Ezzel szemben a kromatin immunprecipitációra jellemző a relatíve alacsony sztenderdizáltság és a hosszú aktív kísérletes idő, amelyek korlátot szabnak a módszer szélesebb körű elterjedésnek.

A humán B-limfoblasztoid sejtvonal (LCL) modellrendszer széles körben használják a genomszekvencia és kvantitatív sejtfenotípusok kapcsolatainak feltárására. Számos genomikai és funkcionális genomikai konzorciális projekt, például a mérföldkőnek számító HapMap Projekt, az 1000 Genom illetve az ENCODE Projekt használt LCL sejtvonalakat. Az LCL-ek genomi asszociációs vizsgálatokban történő használatát támogatja, hogy több ezer LCL és azok annotált genomszekvenciái szabadon elérhetőek a tudományos közösség számára. Mindezek mellett azonban az egyes LCL sejtvonalak közötti genomszekvencia-független funkcionális genomikai variabilitás jelentősen befolyásolhatja ezen vizsgálatok eredményeit.

Tudományos munkánk során a funkcionális genomika tudományterületen használható kísérletes eszközök fejlesztését és jellemzését tűztük ki célul. Egyrészt kifejlesztettünk egy spike-in kontroll rendszert a ChIP protokoll során történő mintavesztés korrigálására. Másrészt egy, a sejtekben található RNS pool stabilizálására használható költséghatékony mintakezelési és tárolási módszer hatékonyságát vizsgáltuk RNS-szekvenáláshoz. Harmadrészt, elsőként vizsgáltuk az LCL modell nem genomi eredetű funkcionális genomikai variabilitásának mértékét és jellegét.

ELMÉLETI HÁTTÉR

Normalizálás kromatin immunprecipitációs kísérletekben

A kromatin immunprecipitáció (ChIP) egy olyan laboratóriumi módszer, amely alkalmas a genomi DNS és különböző fehérjék - például transzkripciós szabályozó fehérjék és poszt-transzkripciósan módosított hisztonok - közötti kapcsolat *ex vivo* vizsgálatára. A módszer mélyszekvenálással összekapcsolva (ChIP-Seq) lehetőséget teremt az epigenetikai jelek teljes genom szintű profilozására. A ChIP-Seq módszert a klinikai kutatásokban is széles körben használják betegség-asszociált epigenetikai mintázatának térképezésére. Azonban több módszertani tényező, például a magas kiindulási sejtszám, a hosszú aktív kísérletes idő és a sztenderdizált protokollok hiánya gátat szab a klinikai körülmények között történő alkalmazás elterjedésének.

A ChIP(-qPCR) protokollok során különböző kontrollok és normalizálók alkalmazására kerül sor, amelyek biztosítják például a változó kiindulási sejtszámból vagy aspecifikus antitest-kötődésből eredő torzító hatások felismerését és kiküszöbölését. Pozitív kontrollként szolgálhat például olyan genomi régiók amplifikációja a célrégiók mellett, amelyekhez a célfehérje nagy bizonyossággal kötődik. A pozitív kontrollokon kívül számos negatív kontroll alkalmazható, elterjedt például az aspecifikus antitesttel történő immunprecipitáció és negatív kontroll genomi régiók amplifikációja is. A ChIP-qPCR-feldúsulásokat szinte kizárólagosan input % formában adjuk meg, amely értéket az IP mintából és az inputból (reprezentatív, nem precipitált, fragmentált kromatin alikvot) származó relatív target mennyiségek elosztásával nyerjük. Az input % megközelítés különösen alkalmas az input kromatin mennyiségi variabilitásának kontrollálására.

Jelenleg nincs egységesen elfogadott belső kontroll ChIP-qPCR kísérletekhez. Ennek legfőbb oka, hogy a megfelelő belső kontroll meghatározására minden kísérleti rendszerben egyénileg lenne szükség. Figyelembe véve, hogy a kísérleti lépések jelentős részére (az input-kontrollált) kromatin fragmentációt követően kerül sor, a kísérletes mintaveszteség jelentős torzító hatásként jelenhet meg megfelelő belső vagy külső kontroll hiányában. Értékes klinikai minták, például biopsziák esetén nem realitás nagyszámú ChIP replika használata, amely esetleg csökkenthetné a fenti torzító hatást. Ilyen esetben ún. spike-in kontrollok jöhetnek számításba, mint posztfragmentációs procedurális kontrollok.

Az elmúlt néhány évben több spike-in kontrollt is leírtak ChIP-Seq kísérletekhez. Minden módszer alapja, hogy valamilyen xenogén anyagot (sejtek vagy fragmentált

kromatin) adnak a kísérletes mintához. A hozzáadott spike anyagot egyazon kísérletben immunprecipitálják pánspecifikus vagy egy külön hozzáadott antitesttel. Az ún. ICeChIP (Internal Standard Calibrated ChIP, belső sztenderd-kalibrált ChIP) lehetővé teszi a hisztónmódosítás denzitásának kalkulációját, ezáltal a kísérletek közvetlen összehasonlítását azáltal, hogy különböző koncentrációkban bárkódolt DNS-t tartalmazó mesterséges nukleoszómákat adnak a kísérleti mintához. Ezen reagensek komoly limitációja, hogy a konzisztens minőségben való előállítás nehezen megoldható. A fenti fejlesztések ellenére a spike-in kontrollok alkalmazása még nem vált elterjedtté.

Bakteriofágok alkalmazhatósága spike-in kontrollként ChIP kísérletekhez

A fágbemutatás fehérjék és egyéb anyagok (pl. más fehérjék, peptidek, DNS vagy ligandok) közötti kölcsönhatások vizsgálatára alkalmas laboratóriumi módszer. A fág-alapú random peptid könyvtárak akár többmilliárd különböző, bakteriofágok felszínéhez kapcsolt („bemutatott”) peptidből állnak. A könyvtárakat fix hosszúságú (pl. 6 vagy 12 aminosavnyi), random szekvenciájú DNS-fragmensek valamely fág (pl. M13) fehérjeburok-génhez (pl. pIV vagy pIII) történő *in-frame* beillesztésével hozzák létre. Az így összeállított fág vektorokat egy *E. coli* törzsbe transzformálják, amely ezt követően ún. fúziós fágokat termel. Az eltérő kísérletes igényeknek megfelelően többféle fág könyvtár van kereskedelmi forgalomban, de speciális igények esetén egyedi könyvtárak is előállíthatóak kitek segítségével.

A fág könyvtárak számos olyan előnyös tulajdonsággal bírnak, amelyek lehetővé tehetik ChIP spike-in kontrollok létrehozását. Egyrészt kifejlészthetőek olyan fágok, amelyek ún. *ChIP-grade* (ChIP kísérletekre alkalmas) antitestekhez kötődnek bemutatott peptidjeiken keresztül. Másrészt mivel a bemutatott fehérje és a fág genom fizikai kapcsolatban állnak egymással, lehetővé válik az affinitás alapján szelektált fágok DNS-alapú relatív kvantifikálása. Ebben az esetben a fágokat a ChIP kísérlet immunprecipitációs lépésében a reakciókhoz adhatjuk, majd a kísérlet végén a kromatin fragmensekkel együtt izolálva kvantifikálhatjuk qPCR segítségével. Az így kapott qPCR-alapú fág visszanyerést reprezentáló számérték poszt-fragmentációs normalizálóként használható. Mindemellett a fágok és bemutatott peptidjeik kevésbé érzékenyek a denaturációra, amely hosszútávú stabilitást biztosíthat a fág-reagensnek. Végül a fágok konzisztens minőségben visszanyerhetőek monoklonális fág reagensek létrehozásával, valamint egy *E. coli* törzs fertőzésével. A filamentózus M13 fág cirkuláris, egyszálú DNS-e újgenerációs szekvenálási könyvtárkészítéskor elvész, így a fágok jelenléte nem interferál a szekvenálás folyamatával.

RNS szekvenálás a klinikai kutatásokban

Az RNS szekvenálás cDNS molekulák nagy átírási képességű szekvenálással történő detektálásán és kvantifikálásán alapuló transzkriptóm-profilozási módszer. A szekvenálási költségek csökkenésének és a bioinformatikai módszerek fejlődésének köszönhetően az RNS szekvenálás vált a nagy átírási képességű RNS biomarker-kutatások vezető módszerévé. Az RNS mennyiségi meghatározására alkalmas módszerek viszonylag érettek, szenzitívek és specifikusak. Mindezekkel összefüggésben a publikált intra- és extracelluláris RNS biomarker-jelöltek száma az elmúlt években folyamatosan emelkedett, amely később RNS-alapú biomarker panelek kifejlesztéséhez vezethet.

A nem egységes RNS szekvenálási módszerek többnyire jelentősen szétválasztottak, így ilyen esetben a sikeres kísérlet szempontjából legkritikusabb tényező a kiindulási minta minősége. A műtétek során gyűjtött klinikai mintákban található RNS molekulák különösen degradálhatók, amelyek fő oka a mintafixálás késlekedése. Míg az RT-(q)PCR-alapú módszerek jellemzően kevésbé érzékenyek a degradációra, az RNS szekvenálás eredményét erősen befolyásolhatja az RNS fragmentáció. Ennek következtében komoly érdeklődés övezi a szövetek/sejtes minták RNS-tartalmának megőrzését magas minőségben későbbi extrakcióhoz.

Szövetbankok és szobahőmérsékleten történő mintatárolás

Mivel a natív biológiai minták érzékenyek számos különböző fizikai és kémiai behatásra, változatos módszereket alkalmaznak a degradáció mértékének csökkentésére és a mérendő tulajdonság megváltozásának minimalizálására. Ezek a módszerek többnyire kémiai inert oldatokban történő tároláson, kémiai fixáláson, (ultra-)mélyfagyasztáson vagy ezek kombinációján alapulnak. A választott módszer függ a minta típusától (pl. teljes szövet vagy molekuláris preparátum), forrásától (pl. biopszia vagy sejtvonal) és a tervezett tárolási időtől (hosszú- vagy rövidtáv). A konzisztens mintaminőség fenntartása biztosítja a hosszú időn keresztül, például ritka betegségekkel összefüggésben, vagy longitudinális vizsgálatokban gyűjtött minták összehasonlíthatóságát.

A szövetbankokban tárolt klinikai minták kiinduló anyagként szolgálnak mind alap-, mind a klinikai kutatásokhoz. A biobankok konvencionálisan ultra-mély fagyasztókat és folyékony nitrogén tartályokat üzemeltetnek a mintakárosodás megelőzésére. Azonban az ultra-mély hőmérsékleten történő tárolás jelentős környezeti terhelést és működtetési költségeket von maga után, amely hatására a szövetbankok fenntarthatóságának kérdése egyre

fontosabb szemponttá válik. Továbbá a szállítás közben, logisztikai akadályok miatt bekövetkező hőmérsékleti ingadozások átmeneti mintaolvadáshoz vezethetnek, fokozva a minták degradációjának kockázatát. A fenti problémákat a szobahőmérsékleten történő tárolás és szállítás orvosolhatná.

Liofilezés

A liofilezés segítségével lehetőség nyílik a sejtek és szövetek szobahőmérsékleten történő tárolására és szállítására. A liofilezés vagy fagyasztva szárítás olyan szárítási módszer, amely vákuum és mérsékelt hőhatás segítségével távolítja el a vízmolekulákat fagyasztott (biológiai) mintákból szublimáción és deszorpción keresztül. A szárítási folyamatot követően a mintákat általában szorosan lezárják és sötétben tárolják fagyáspont feletti hőmérsékleten (pl. szobahőmérsékleten). A következő jellemzők teszik a liofilezést a fagyáspont alatti hőmérsékleten történő tárolás potenciális alternatívájává: (1) alacsony reziduális víztartalom, amely a molekuláris mozgások és a víz által közvetített (bio)kémiai reakciók felfüggesztésével gátolja a molekuláris degradációs útvonalakat; (2) a szárítási folyamat során az alacsony alkalmazott hőmennyiség miatt a hőérzékeny anyagok bomlása megelőzhető; (3) a módszer rövid aktív kísérletes időt igényel, valamint (4) a krioprezervációhoz viszonyítva a tárolási költség csökkenthető, mivel a minták tárolásához nincs szükség erősen kontrollált alacsony hőmérsékletre; továbbá (5) a hőmérsékleti ingadozások kevésbé érinthetik a liofilezett minták integritását. Az egyéb szobahőmérsékleten történő tárolást lehetővé tevő módszerek közül a formalin-fixálás és paraffinba történő beágyazás kiterjedt RNS-degradációt és bázismódosulást eredményez, a nem-keresztelő fixatívok közül például az RNAlater bizonyítottan transzkriptóm-változásokat eredményez.

Az RNS minták integritását szövetliofilezéssel összefüggésben több csoport is vizsgálta változatos liofilezési módszereket és tárolási körülményeket alkalmazva. Általánosságban elmondható, hogy a tanulmányok többségében alacsonyabb RIN értékeket detektáltak, míg az RT-(q)PCR konzisztensek voltak hosszú tárolási időtartamokat követően is. Az RNS stabilitása szempontjából a sötétben, vízelvonószerek jelenlétében történő tárolás kritikusnak mutatkozott, míg a szobahőmérsékleten történő tárolás előnyösebbnek bizonyult a 4°C-on történő tárolással szemben. Továbbá a nukleinsavak stabilitása liofilezett szövetekben összefüggést mutatott a szövet lipidtartalmával és az oxigén-közvetítette peroxidációval. A fentiek ellenére a szövetek liofilezése egy kiaknázatlan módszer szöveti RNS fenntartható tárolásához.

A sejtvonalak fenotípus-instabilitása

A sejt kultúrákban növekedő sejtek genomi eltérései jelentősen hozzájárulhatnak a sejtfenotípus megváltozásához. A sejtvonalak genomi instabilitása a genomi változások normál szervezethez viszonyított magas gyakoriságára utal, beleértve az aneuploidiát, kromoszómális aberrációkat és a kis léptékű mutációkat. A genomi instabilitás különösen jellemző a rákos sejtvonalakra, amelyekben már az *in vitro* tenyésztést megelőzően is jelen lehetnek például sejtciklus-ellenőrzőpontokat vagy DNS-hibajavítást érintő mutációk. Továbbá széles körben elfogadott, hogy a magas passzázsszámú sejtek genomi változásokat szenvednek és mutációkat halmoznak fel, azonban ennek időbelisége és jellege sejtvonalanként eltérhet. Ennek ellenére nincs általános konszenzus a maximálisan megengedhető passzázsszámmal kapcsolatban.

A genomi változások mellett környezeti hatások is hozzájárulnak a tenyésztett sejtek fenotípusának modulálásához. Az egyes sejtek intracelluláris állapota meghatározza a sejtfenotípus megváltoztathatóságát (fogékonyságát) bizonyos stimulusok hatására. Továbbá a sejttenyésztő médiumok legkevésbé kontrollált összetevője, az FCS összetétele a sejtpopulációkat külön evolúciós útvonalakra terelheti. Genetikailag stabil sejtvonalakban (pl. embrionális őssejtek) várhatóan a környezeti hatások játszanak a legjelentősebb szerepet a sejtfenotípus (akár permanens) megváltozásában. Mindezek ismeretében a sejtfenotípus-mérések összehasonlíthatósága érdekében célszerű a tenyésztési körülmények lehető legnagyobb mértékű sztenderdizálása.

Széles körben ismert, hogy a sejtvonalak alapszintű heterogenitása csökken hosszú távú tenyésztés során. A jelenséget gyakran „klonális evolúció” néven emlegetik, amely során a sejttenyésztet egy vagy néhány növekedési előnnyel rendelkező sejt leszármazottai alkotják. Ezen változások kétség kívül hatással annak a fiatalabb és idősebb kultúrákon végzett kísérletek eredményeinek összehasonlíthatóságára. Továbbá óvatosságot igényel a különböző forrásokból beszerzett sejtvonalakkal kapott eredmények összehasonlítása is, ugyanis ezek valószínűleg egymástól függetlenül több „palacknyak-effektuson” mentek keresztül, jelentős geno- és fenotípusos divergenciát eredményezve. A fenti hatások minimalizálására azonos passzázsszámmal rendelkező batcheket tartalmazó sejtbiobankokat hozhatunk létre.

Humán B-limfoblasztoid sejtvonalak mint funkcionális genomikai modellek

A humán B-limfoblasztoid sejtvonalakokat (LCL) nyugvó B-sejtekből Epstein-Barr vírussal (EBV) történő *in vitro* fertőzéssel hozzák létre. Az EBV-fertőzés karakterisztikus fenotípus-változással jár, azonban a genotípus vagy az egyes egyénekből származó LCL-ek egyedi génexpressziós mintázata jellemzően nem változik. Az LCL-ek szolgáltak olyan nagy omikai projektek modelljeiként, mint a HapMap, az 1000 Genom, az ENCODE és a FANTOM5 projektek. Az első kilobázisos felbontású 3D genomi térkép szintén egy limfoblasztoid sejtvonallal készült. Az LCL-ek fő előnyei a következők: 1) az LCL-generálás az egyik legegyszerűbb módszer megújuló mintaforrás létrehozására egészséges egyénekből, 2) az LCL-ek egyszerűen és nagy mennyiségben tenyésztethetők, 3) pre-immortális státuszúak (limitált passzázsszámig osztódóképesek), 4) a genetikai és kariotípusos stabilitásuknak jelentős szakirodalma van, illetve 5) nagyszámú LCL sejtvonallal teljes genom szekvenciája publikusan elérhető. Az utóbbi eredményeképpen az LCL-eket gyakran használják molekuláris- és gyógyszerválasz-QTL-ek meghatározására.

LCL-ek genetikai asszociációs vizsgálatokban

Az elmúlt évtizedben az LCL-ek a humán QTL vizsgálatok egyik fő sejtvonall-modelljévé váltak. A kvantitatív jellegeket meghatározó lokuszok (QTL-ek) olyan genomi helyek, amelyek korrelációt mutatnak kvantitatív fenotípusos jellegekkel, például hiszton módosításokkal (hmQTL), génexpressziós szintekkel (eQTL) vagy gyógyszerválasszal. A heterogén, nagy mintaszámú vizsgálatok esetén különösen előnyös, hogy több mint ezer LCL genotípus-adatuk publikusan elérhetőek. Továbbá online adatbázisokban (pl. GEO és ArrayExpress) releváns kapcsolódó funkcionális genomikai adatok szintén elérhetőek (pl. ENCODE CHIP-Seq adatszettek).

Az LCL sejtvonalakokat egyre gyakrabban használják farmakogenomikai kutatásokban. Ennek fő oka a fentiekén kívül az, hogy nagyszámú hasonló korú, nemű és azonos rasszhoz tartozó egyénekből származó sejtvonallal érhető el, és ezek mentesek az *in vivo* zavaró tényezőktől. Az LCL-eket mind elsődleges, mind validációs modellként alkalmazzák. Az úgynevezett „triangle approach” (háromszög modell) csökkentheti a fals pozitív találatokat azáltal, hogy génexpressziós adatok is bevonásra kerülnek farmakogenomikai vizsgálatokba. Az LCL-eket alkalmazták már például többféle komoterápiás szerrel szembeni érzékenység, valamint lipideszökkentő gyógyszerek és antidepresszánsok hatásának vizsgálatára genomi összefüggésben.

CÉLKITŰZÉSEK

Célkitűzés 1. Egy fágbemutatáson alapuló ChIP-qPCR folyamatkontroll kifejlesztése és jellemzése

A ChIP protokoll folyamatának egy jelentős része nem kontrollált, így a kísérletes mintavesztés miatti torzítás detektálatlan maradhat komparatív ChIP kísérletekben. Vizsgálatunk során a következő célokat tűztük ki:

- Egy olyan fágbemutatáson alapuló ChIP-qPCR spike-in kontroll rendszer kifejlesztése, amely lehetővé teszi a mintavesztés nyomon követését és normalizálását ChIP kísérletekben;
- Antitest által felismerhető androgén receptor (AR) epitópot utánzó fágok szelektálása több körön keresztül történő *in vitro* evolúció során, majd a poliklonális reagens diverzitásának, illetve androgén receptor-ellenes antitesthez való affinitásának ChIP kísérletben történő meghatározása;
- Monoklonális fág reagensok létrehozása a ChIP kísérletben legmagasabb affinitást mutató poliklonális batch-ből, majd a monoklonális reagensok androgén receptor-ellenes antitesthez való affinitásának ChIP kísérletekben történő meghatározása.

Célkitűzés 2. A liofilezés mint alternatív sejtarchiválási módszer RNS-alapú vizsgálatokhoz

A liofilezés bár megjelent, mint potenciálisan költséghatékony biológiai mintatárolási és szállítási alternatíva, széleskörű bevezetése még várat magára. Vizsgálatunk során a következő célokat tűztük ki:

- Az epigallokatekin-gallát lioprotektáns sejtmembrán-stabilizátor hatásának vizsgálata a sejtliofilezést követő RNS- és ChIP-mérésekre;
- RNS-integritás és kitermelés, valamint qPCR-alapú génexpressziós szintek meghatározása 0,1 M trehalóz/PBS-ben liofilezett LCL-sejtekből a liofilezést követően, és kettő illetve nyolc hét szőbahőn történő tárolás után;
- Transzkriptóm-profilozás párosított kontroll és két hétig szobahőmérsékleten tárolt, liofilezett LCL sejteken mRNS-szekvenálással, valamint a könyvtárak minőségének

meghatározása például könyvtárkomplexitás, leolvasás GC-tartalom és mismatch arány vizsgálatával;

- Funkció- és szekvencia-alapú analízisek elvégzése a liofilezett mintákban alulreprezentált transzkriptek jellemzésére.

Célkitűzés 3. A limfoblasztoid sejtvonalak genotípus-független variabilitása génszabályozással és külső stimulusra adott válasszal kapcsolatos sejtfenotípus szinteken.

A genetikai asszociációs vizsgálatokban gyakran használt LCL sejtvonalak nem genetikai eredetű funkcionális genomikai variabilitásának mértéke és jellege eddig nem volt ismert. Vizsgálatunk során izogenikus, azonos egyénből származó LCL-ek használatával a következő célokat tűztük ki:

- Alapvető sejtvonal-karakterizálás elvégzése, beleértve a mikroszatellita-meghatározást, sejtciklus-vizsgálatot és az immunfenotipizálást;
- ChIP-Seq kísérletek elvégzése a H3-as hiszton 27-es lizinjének acetilcsoportja (H3K27ac) ellen termeltetett antitesttel az aktív génszabályozó elemek térképezése és összehasonlítása céljából;
- Az izogenikus LCL-ek transzkriptóm-profilozása mRNS-szekvenálással, majd az érintett gének illetve H3K27ac és RNS szintek összevetése;
- A variábilis gének funkcionális jellemzőinek meghatározása, valamint a variábilisan expresszálódó farmakogének kemoterápiás gyógyszerválaszban betöltött lehetséges szerepének vizsgálata a DPYD gén és az 5-fluorouracil példáján.

ANYAG ÉS MÓDSZER

Sejttenyésztés

A Coriell Cell Repositories-től vásárolt LCL sejtvonalakból (GM22647-GM22651, GM12864, GM12872, GM12873) háromszintű biobankot készítettünk. Kísérleteinkhez minden esetben a 3. szinthez tartozó sejteket (Tier 3) használtunk. A sejteket 15% hőinaktivált FCS-t, 2 mM L-glutamint és 1% penicillin-sztreptomocint tartalmazó RPMI-1640 médiumban tenyésztettük 37°C-on, 5% CO₂ jelenlétében, álló helyzetű T25 vagy T75-ös flaskákban.

Mikroszatellita elemzés

10⁶ sejtet mostunk PBS-sel, majd a genomi DNS-t a Roche cég High Pure PCR Template Preparation Kitjével izoláltuk. Öt mikroszatellita régiót (AMELY/AMELX, D18S51, D8S1179, TH01 és FGA) amplifikáltunk a PowerPlex S5 System segítségével (Promega), majd a termékeket ABI PRISM 3100-Avant Genetic Analyzer műszerrel futtattuk meg. Az eredményeket a GeneMapper ID (v4.1) szoftverrel elemeztük (Laboratóriumi Medicina Intézet – Debreceni Egyetem, Általános Orvostudományi Kar).

Immunfenotipizálás és sejtciklus-elemzés

Reakciónként 10⁶ sejtet inkubáltunk 15 percre, szobahőmérsékleten, sötétben antitestek kombinációival. A mintákat mostuk, majd 1% paraformaldehid/PBS-ben fixáltuk. Az intracelluláris festést az Intrastain (Dako Glostrup) protokollját követve végeztük el. Reakciónként 100.000 eseményt számoltunk le FACS Canto II áramlási citométer segítségével (BD), míg az adatokat FACS Diva (BD) és a Kaluza Software v1.2 (Beckman Coulter) segítségével elemeztük. A használt antitestek forrása: CD5, CD10, CD20, CD22, CD24, CD34, CD38, CD79b, CD81, FMC7, HLA-DR, kappa és lambda markerek (BD); CD19, CD23 és CD43 (Beckman Coulter); CD21 és anti-CD45 (Exbio); nTdT és IgM (Dako). Sejtcikluselemzéshez 2*10⁶ mosott sejtet fixáltunk 70% etanollal 4°C-on, majd a fixált sejteket centrifugáltuk, illetve RNázal és Propidium-jodiddal kezeltük 30 percre, szobahőmérsékleten, sötétben. FACS Calibur II áramlási citométer segítségével (BD) 20.000 eseményt számoltunk le, majd az adatokat ModFit LT (BD) segítségével elemeztük.

In vitro fág evolúció (biopanning)

A Ph.D.TM-7 random heptapeptid könyvtár (NEB) négy egymást követő körben szelektáltuk AR N-20 (humán androgén receptor-ellenes) antitesttel (Santa Cruz

Biotechnology) fedett mágneses gyöngyökkel. TBST/BSA-val blokkolt 10^{11} plakk formáló egységet (PFU-t) immunprecipitáltunk 10 μ g AR antitesttel (10-60 perc, szobahőmérséklet), majd TBST/BSA-val blokkolt protein A:Protein G mágneses gyöngy mixet (1:1) adtunk a reakcióhoz (20 perc, szobahőmérséklet). A gyöngyöket tízszer mostuk TBST/BSA-val, eluáltuk majd neutralizáltuk. Korai log növekedési fázisban lévő F+ ER2738 baktériumokat először fág eluátummal fertőztünk, majd nem szelektív LB-ben növesztettük. A termelt fágokat a bakteriális felülúszóból 1/6 térfogatarányú 20% PEG-8000/2.5 M NaCl-lel precipitáltuk egy éjszakán át, 4°C-on. A fág pelleteket TBS-ben felvettük, majd ismét precipitáltuk. A TBS-ben oldott fágokat használtuk a következő körben végzett *in vitro* evolúcióhoz (10^{10} PFU). A fág koncentrációk méréséhez SmartSpec Plus (Bio-Rad) készüléket használtunk. A fág reagenseket 50% glicerin jelenlétében -20°C-on tároltuk.

Fág szubklónozás

A NEB fágtitrálási és plakk amplifikációs protokollját követve végeztük a fágok szubklónozását. A negyedik körben előállított poliklonális fág reagenst LB-ben hígítottuk, amellyel mid-log fázisban lévő ER2738 sejteket fertőztünk. A fertőzött sejteket 45°C-os “Top Agar”-ban szuszpendáltuk, majd LB/IPTG/X-gal lemezre öntöttük és 37°C-on inkubáltuk egy éjszakán át. A következő napon ER2738 tenyészetet fertőztünk a lemezről származó, egyedi telepekkel. A monoklonális fágokat precipitáltuk és a poliklonális fágokkal megegyező módon kezeltük.

Kromatin immunprecipitáció (ChIP) és ChIP-szekvenálás

Az LCL kísérletek esetén a PBS-sel mosott sejteket 1% metanolmentes formalinnal (Thermo, 10 perc) fixáltuk (izogénenikus sejtek), vagy 2 mM di(N-szukcinimidil) glutaráttal (Sigma; 45 perc) illetve 1 % formalinnal (Thermo, 10 perc) kötöttük keresztbe (trio). A keresztkötési reakciókat 0,125 M glicinnel állítottuk le. A lyo tanulmányhoz kezeletlen, IMT-2-ben liofilezett, IMT-2-ben liofilezett és 1% formalinnal fixált, trehalóz-kezelt és IMT-2-liofilezett, illetve formalinnal előfixált és IMT-2-liofilezett sejteket használtunk. Az EGCG kezelést egy órán át végeztük (2 mM) a kísérlet kezdete előtt. A sejtmagokat izoláltuk, majd szonikáltuk (Bioruptor Plus) vagy MNáz enzimmel kezeltük és enyhén szonikáltuk. Az $5 \cdot 10^6$ sejtnek megfelelő, centrifugálással sejttörmeléktől megtisztított és tízszeresen hígított kromatinokat 2.5 μ g anti-histone H3 (acetyl K27) antitesttel vagy izotípus kontroll antitesttel immunprecipitáltuk egy éjszakán át, 4°C-on. A centrifugálást követően a reakciók felső 90%-át BSA-blokkolt Protein A:Protein G (1:1) mágneses gyöngy mixszel inkubáltuk 6 órán át,

4°C-on. A gyöngyöket összesen hatszor mostuk különböző erősségű pufferekkel, majd az antitest-antigén komplexeket eluáltuk. A keresztkötések felbontását követően a mintákat RNáz A-val és Proteináz K-val kezeltük. A mintákat ezután MinElute PCR Purification Kittel izoláltuk (Qiagen), a DNS-koncentrációt Qubit dsDNA HS Assay Kittel (Thermo) határoztuk meg. A pufferek többségéhez cOmplete Mini proteáz inhibitor tablettát adtuk (Roche). A feldúsítás előtti teljes kromatin minták egy részét "inputként" használtuk ChIP-qPCR kísérletekhez. Negatív kontrollként izotípus kontrollal precipitált mintákat, illetve H3K27ac-negatív genomi régiót használtunk.

A fág ChIP kísérletekhez az IP-Star Automated System-et (Diagenode) használtuk a tisztítási lépésig. A kísérletekhez HEK293T sejtekből származó 1% formalinnal fixált, szonikált HEK293T sejteket használtunk. Az immunprecipitációs 96-lyukú lemez egy-egy sorába az egyes kísérletekhez szükséges reagenseket pipettáztuk: 1 µg AR antitest (vagy izotípus kontroll IgG) immunprecipitációs (IP) pufferben, HEK293T kromatin (~10.000 sejt), 10⁶ fág részecske IP pufferben, 100 µl mosó puffer A, B vagy C, 100 µl TE puffer, 100 µl elúciós puffer, illetve 10 µl Protein A:Protein G mágneses gyöngy mix (1:1). A programozott IP, gyöngykötés, mosás és elúciós lépéseket követően a keresztkötéseket febonítottuk, majd RNázzal és proteázzal kezeltük. A mintákat High Pure PCR Template Preparation Kittel izoláltuk (Roche). A qPCR kísérletekhez M13 univerzális primereket és UPL 48 fluoreszcens próbát (Roche) alkalmaztunk, míg a méréseket LightCycler 480 készüléken (Roche) végeztük.

A szekvenálási könyvtárat 10 ng ChIP DNS-ből készítettük a TruSeq ChIP Sample Preparation Guide 15023092 B (Illumina) protokollt követve. Az elkészült könyvtárat a Genomi Medicina és Bioinformatikai Szolgáltató Laboratóriumban (NextSeq 500 rendszer) és az EMBL Genomics Core Facility-ben (Heidelberg, Germany; HiSeq2000 system) (50-bp, single-end) szekvenáltuk. A BWA 0.7.10 szoftvert használtuk a readok hg19 referencia genomra történő illesztéséhez. A HOMER 4.9.1 szoftver segítségével prediktáltuk a H3K27ac-feldúsulást mutató genomi régiókat, szuper-enhanszereket és a nukleoszómamentes régiókat. Az eltérően acetilált régiókat a pheatmap (R) segítségével klaszterezttük, és a tag denzításokat Java TreeView segítségével vizualizáltuk. A konszenzus régiók meghatározásához, RPKM értékek kalkulálásához és korrelációs hőtérkép-készítéshez DiffBind-ot használtunk. A hőtérképet Plotly 3.0.0 segítségével jelenítettük meg. Az eltérően acetilálódó régiókat two-way ANOVA és Tukey post hoc teszttel (aov(), TukeyHSD())

határoztuk meg (MASS, R; $P < 0,05$; $FC > 2$). A legközelebbi géneket bedtools segítségével határoztuk meg.

MNáz profilozás

A sejtmagokat mostuk, majd 66,6 Gél Egység (GU), 22,2 GU vagy 7,4 GU MNázzal kezeltük 37°C-on 30 percig ($1,5 \cdot 10^6$ millió sejtmag). A leállított reakciókat enyhén szonikáltuk (Bioruptor Plus, Diagenode), majd a sejttörmelékét centrifugálással eltávolítottuk. A felülúszó 80%-át precipitáltuk 3 térfogatnyi abszolút etanollal -20°C-on, egy éjszakán át. A precipitált nukleinsavakat centrifugálással üleptítettük és a pelleteket beszárítottuk. A fragmentált DNS-t a továbbiakban a fenti ChIP protokollnak megfelelően kezeltük. A fragmenseloszlást 1% agaróz gélen vizsgáltuk.

Kromatin konformációs vizsgálat (3C-Seq)

A 3C-szekvenáláshoz egy korábban publikált protokollt követtünk. A sejteket 10% FCS/PBS-ben 1% formalinnal (Thermo) fixáltuk 10 percig szobahőmérsékleten. A reakciókat 0,125 M glicinnel leállítottuk, majd a mosott sejt pelletekből sejtmagot izoláltunk. A mosott magokat 400 U EcoRI-HF (NEB) enzimmel hasítottuk 37°C-on, majd ligáltuk 100 U T4 ligáz (NEB) segítségével (magas hígítás, 16°C). A keresztkötések felbontását, RNáz és proteáz kezelést követően a DNS-t fenol-kloroform-izoamlialkohol (FKI) módszerrel izoláltuk. 10 µg DNS-t 10 U MseI enzimmel (NEB) hasítottunk 37°C-on, az előzőekkel megegyező módon ligáltuk, majd FKI módszerrel izoláltuk. Az így kapott 3C könyvtár mintákat az Expand Long Range, dNTPack (Sigma) kit segítségével, a P3H2 géntesthez kötődő inverz primerekkel amplifikáltuk. A PCR reakciókat a MinElute PCR Purification Kittel (Qiagen) izoláltuk. A 3C-Seq könyvtárakat a TruSeq ChIP Sample Preparation Guide 15023092 B (Illumina) segítségével készítettük, és a mintákat 75 bp hosszúságban szekvenáltuk az EMBL Genomics Core Facility-ben (Heidelberg, HiSeq2000 rendszer).

Liofilezés

Három millió PBS-sel mosott LCL sejtet 0,5 ml 0.1 M D-(+)-trehalóz-dihidrát/PBS-ben vagy IMT-2 oldatban (0.1 M D-(+)-trehalóz dihidrát, 0,945 mg/ml (-)-epigallokatekin-gallát, PBS) liofileztük mikrocentrifugacsövekben. A sejtszuszpenziókat folyékony nitrogénen fagyasztottuk, a csövek szájadékát 1 mm átmérőjű lyukakkal ellátott filmmel lefedtük, majd a mintákat fagyasztva szárító készülékbe helyeztük (CoolSafe 110, ScanVac; kondenzorhőmérséklet: -110°C; Proteomikai Szolgáltató Laboratórium). A liofilezés 6 órán

keresztül, 0,004 mBar nyomáson történt. A ciklus befejeztével a csöveket lezártuk és a liofilezett sejteket vagy azonnal, vagy 2 hét illetve 2 hónap (23–25°C, nem transzparens tárolóban, CaCl₂-dihidrát jelenlétében) tárolást követően processzáltuk.

RNS izolálás

A totál RNS-t trizol reagenssel (UD-Genomed Kft.) izoláltuk. Projekttól függően 2-3*10⁶ kezeletlen vagy liofilezett LCL sejtet PBS-sel mostunk, centrifugálással ülepitettünk, majd 1 ml trizol reagensben vortexeltünk 5 percig. A fázisszeparációt kloroformmal (1:5) és magas sebességű centrifugálással végeztük. Az RNS frakciót a vizes fázisból izopropanollal precipitáltuk (1:1). A pelletet kétszer mostuk 75% fagyos etanollal, majd vákuumcentrifuga segítségével beszárítottuk. A beszárított pelletet nukleázmentes vízben vettük fel, majd a minták tisztaságát NanoDrop 1000 készülékkel (Thermo), míg koncentrációját Qubit fluorométerrel (RNA HS Assay Kit, Thermo) ellenőriztük. Agilent RNA 6000 Nano mikrocsipeken ellenőriztük a fragmenseloszlást és kalkuláltuk a RIN értékeket.

RT-qPCR

A qPCR primereket a Roche UPL Assay Design Center segítségével (UBR2, TRERF1, PTPRJ, SLC6A4, RXRA, TCL1A, SPI1 and CT64 assays) vagy a Primer 3 Plus szoftverrel (DPYD; ACTB; lncRNAs: MALAT1, GAS5, TUG1; eRNAs: eIRF4_-1.9kb, eSPI1_-16kb and eMYC_-170kb) terveztük. A GAPDH primerek a Sigma weboldaláról származnak. A totál RNS mintákat RQ1 DNÁzzal (Promega) kezeltük, majd a reverz transzkripciót SuperScript II enzimmal végeztük (Thermo). Az RT reakciókat qPCR előtt ötszörösére hígítottuk nukleázmentes vízzel. Az EGCG inhibíziós kísérletek során 20 µl RT reakciókat állítottunk össze, amelyben az EGCG végkoncentrációja 1*10⁶ és 1*10⁷ M között volt. A célrégiókat LightCycler 480 SYBR Green I Master (Roche) mestemixszel amplifikáltuk, with 0,375 µM forward és reverz primerrel. A qPCR reakciókat triplikákban mértük és nukleázmentes víz kontrollt használtunk, míg az RT reakció esetén enzim nélküli kontrollt (NAC) alkalmaztunk. Az egyes minták qPCR jeleit ACTB háztartási génre normalizáltuk (2^{-ΔCp}).

RNS szekvenálás

A szekvenálási könyvtárakat 1 µg totál RNS-ből kiindulva az Illumina TruSeq RNA Sample Preparation v2 Guide-ját követve végeztük. A könyvtárakat NextSeq 500 rendszeren szekvenáltuk 50 bp (izogenikus LCL-ek) vagy 75 bp (liofilezés) hosszúságban (egyirányú

leolvasás, single-end). Az izogenikus LCL-ek esetén minden lépést a Genomi Medicina és Bioinformatikai Szolgáltató Laboratóriumban végeztük. A liofilezési tanulmányban a klasztergenerálást, szekvenálást és bázishívást a Semmelweis Egyetem II. sz. Gyermekgyógyászati Klinikáján végezték. A readeket a TopHat v2.0.7 programmal illesztettük a hg19 humán referencia genomra. Az izogenikus LCL-ek esetén a transzkriptek mennyiségi meghatározását edgeR segítségével végeztük, és a CPM>5 géneket expresszáldónak tekintettük; az eltérően expresszáldó géneket szintén edgeR-ral határoztuk meg (FDR = 0,05, FC > 2). A liofilezési tanulmányban a transzkriptek mennyiségi meghatározását Cufflinks-szel végeztük, az RPKM>1 géneket expresszáldónak tekintettük, és a kis RNS-eket kiszűrtük. Az eltérően expresszáldó géneket Cuffdiff segítségével határoztuk meg (FDR = 0,05). A gének funkcionális annotálását a DAVID Bioinformatics Resources 6.8 eszköz segítségével végeztük. A szekvenáló könyvtárakkal kapcsolatos metaadatokat a QoRTs programcsomaggal határoztuk meg. A transzkript információkat a HGNC adatbázisból töltöttük le BioMart segítségével, és az elemzéseket saját bash szkriptekkel végeztük. A statisztikai elemzéseket Mann-Whitney U teszttel végeztük. Az ARE adatokat az ARED-Plus adatbázisból kértük le.

5-FU kezelés és citotoxicitási esszé

96-lyukú U-aljú tenyésztő lemez lyukaiba 20.000 sejtet pipettáztunk „FU médiumban” (indikátormentes RPMI, 15% hőinaktivált FCS, 2 mM L-glutamin, 1 % penicillin-sztreptomicin). A méréseket kvadruplikátumokban végeztük. Az 5-FU stockból (TEVA) tova futó hígítási sort készítettünk „FU médiumban”, majd meghatározott végkoncentrációban hozzáadtuk a sejtekhez. Vivőanyagként (vehicle) ultratiszta vizet használtunk, háttérként sejtek nélküli médiumot használtunk. A sejteket 37°C-on inkubáltuk (5% CO₂) 72 órán át, majd MTT stock (4,5 mg/ml PBS-ben) lyukakhoz történő pipettázását követően a lemezeket was 37°C-on inkubáltuk 6 órán át, sötétben. A sejteket egy órán keresztül lizáltuk, majd a lizátumok abszorbanciáját VICTOR3 Multilabel Plate Reader (PerkinElmer) segítségével mértük 595 nm hullámhosszon.

Kapilláris szekvenálás

A fág genomi DNS-t a High Pure PCR Template Preparation Kittel izoláltuk (Roche). A mintákat -96 gIII primerrel amplifikáltuk, majd a fragmenseket ABI 310 Avant szekvenátorral detektáltuk.

EREDMÉNYEK

Androgén receptor epitópot utánzó, chip spike-in kontrollként használható poliklonális fágok fejlesztése és jellemzése

A NEB cég által forgalmazott Ph.D.TM-7 random heptapeptid könyvtárból kiindulva létrehoztunk androgén receptor (AR) epitópot utánzó poliklonális fág reagenseket. A fágokat ChIP-grade AR-ellenes antitesthez való affinitás alapján négy egymást követő körben szelektáltuk. A négy poliklonális batch-ből származó izolált fág genomokban a 21 variábilis pozíció báziseloszlását kapilláris szekvenálással ellenőriztük és azt tapasztaltuk, hogy az egy nukleobázisítípus által dominált pozíciók száma növekedett, míg a guanidin-dinukleotid háttér csökkent a szelekciós körök számának növekedésével. Ezt követően a poliklonális AR fágok antitest-affinitását klasszikus ChIP kísérleti körülmények között teszteltük: minden reakció keresztkötött és szonikált kromatint (HEK293T), AR-ellenes vagy izotípus kontroll antitestet és 10⁶ fág részecskét. A fág ChIP DNS-t nem variábilis fág genomi DNS szakaszra tervezett qPCR primerekkel amplifikáltuk. Eredményeink azt mutatták, hogy a szimulált ChIP reakciókban az immunprecipitált fágok száma a szelekciós körök számának növekedésével nő. Az izotípus kontrollal végzett kísérletekből nyert qPCR jelek stabilnak mutatkoztak a batch-ek között, amely az aspecifikus (pl. cső falához, mágneses gyöngyöz, antitest konstans régiókhoz) kötődés stabilan alacsony szintjét jelzi. A negyedik körben előállított fág reagens részecskéinek közel 50%-át sikerült visszanyerni a ChIP kísérletben.

Androgén receptor epitópot utánzó, chip spike-in kontrollként használható monoklonális fágok fejlesztése és jellemzése

Következő lépésként egyedi AR fág klónokat szelektáltunk és jellemeztünk ChIP kísérletben mutatott AR-ellenes antitesthez mutatott affinitás alapján. A poliklonális fág keverékek esetén a batchek regenerálásakor (sokszorozásakor) a változó infektivitás és az ER2738 sejtek növekedési különbségei miatti klóneloszlás-eltolódás jelensége a monoklonális reagenseket nem érinti. Továbbá kiküszöbölhető az a jelenség, hogy a reagens egy része a szelekciós környezet egyes komponenseihez kötődik az antitest variábilis régiója helyett. A negyedik szelekciós körben létrehozott poliklonális reagenssel fertőztünk ER2738 sejteket, majd egyedi klónokat termeltettünk és izoláltunk. Félrobotikus ChIP kísérlettel kimutattuk, hogy az egyes klónok Elúciós puffer/IP puffer aránya minden klón esetén 50% felett volt, míg egy klón különösen magas, 90% feletti kötődést mutatott.

Az epigallokatekin-gallát lioprotektáns hatása RT-qPCR és ChIP-qPCR kísérletekre

Korábbi élő emlőssejt liofilezésre vonatkozó vizsgálatokat figyelembe véve első kísérleteinkben az epigallokatekin-gallát (EGCG) tartalmú IMT-2 oldatot választottuk lioprotektánsként. LCL sejteket liofileztünk IMT-2 oldatban, majd a liofilezett sejtekből izolált RNS és DNS mintákat RT-qPCR-ral és ChIP-qPCR-ral vizsgáltuk.

A liofilezett sejtekből kicsapott totál RNS pelletek barnás-szürke színezetűek voltak, amely EGCG kontaminációra utalt. Ennek megfelelően a totál RNS minták $OD_{260/230}$ értéke alacsonyabb volt a kezeletlen kontroll sejtekhez viszonyítva ($P = 0,03$, páros t-próba). A RIN értékek bár magasak voltak a liofilezett minták esetén, azonban alacsonyabbak, mint a párosított kontrollokban ($P = 0,03$, páros t-próba). Emellett a liofilezett mintákban egy magasan expresszáldó háztartási gén, a GAPDH RT-qPCR mérésekor azt tapasztaltuk, hogy a liofilezett mintákhoz tartozó Cp értékek jelentősen magasabbak voltak a kontrollokéhoz viszonyítva, amely különbség már nem jelent meg a totál RNS minták RT reakció előtti 100-szoros hígításakor. EGCG spike kísérlettel igazoltuk az EGCG koncentrációfüggő gátló hatását az RT-qPCR reakcióra.

IMT-2-ben liofilezett sejtek H3K27ac ChIP-qPCR kísérlete során kizárólag a liofilezés előtt fixált (1% FA) sejtek kromatinjai mutattak megfelelő IP hatékonyságot, a nem előfixált sejtekből származó kromatin IP hatékonysága a fragmentáció típusától, a fixálás erősségétől és a trehalózzal történő előzőleges sejteltöltéstől függetlenül alacsony maradt. Az MNáz-profilozás normál, ún. *beads-on-a-string* profilt tárt fel liofilezett sejtek esetén. Az előfixálás nélkül liofilezett sejtekből származó ChIP DNS a totál RNS mintákhoz hasonlóan enyhe barnás elszíneződést mutatott. Az input minták hasonló Cp értékeket mutattak, amely arra utal, hogy a qPCR-gátlás kis valószínűséggel felelős a fenti jelenségért. Annak igazolására, hogy az alacsony IP hatékonyságért az EGCG jelenléte önmagában felelősség tehető, LCL sejteket ChIP kísérlet előtt egy órán át 2 mM EGCG-vel kezeltünk, majd egy pozitív kontroll régió qPCR méréseivel az IP hatékonyság jelentős csökkenését tapasztaltuk. A fentieket figyelembe véve arra a következtetésre jutottunk, hogy az EGCG nem alkalmas lioprotektánsként amennyiben RNS- vagy kromatin-alapú vizsgálatokat tervezünk.

Az izolált RNS minőségének és mennyiségének meghatározása trehalózban történő sejtlíofilezést követően, illetve több hét szobahőmérsékleten történő tárolás után sztenderd módszerekkel

A trehalóz/PBS oldatban líofilezett LCL sejtekből izolált totál RNS minták magas RIN értékkel rendelkeztek (átlag = 9,8), valamint a kitermelés nem különbözött szignifikánsan a kezeletlen sejtekhez viszonyítva. Eltérő mértékben expresszálandó mRNS-ek (extrém alacsony, RPKM < 1; alacsony, RPKM = 1-10; mérsékelt, RPKM = 10-100; magas, RPKM > 100), hosszú nemkódoló RNS-ek és enhanszer-asszociált RNS-ek RT-qPCR-alapú mennyiségi meghatározása szignifikánsan nem eltérő RNS-szinteket mutatott kontroll és líofilezett minták összehasonlításakor. Ezen eredmények arra utalnak, hogy a líofilezett sejtekből nyert RNS minták alkalmasak a génexpresszió RT-qPCR-alapú mérésére.

Trehalózban történő LCL sejt líofilezés hatása a transzkriptómra

Három kontroll és három, líofilezést követően két hétig szobahőmérsékleten tárolt sejtől izolált totál RNS mintát jelöltünk ki mRNS-szekvenálás alapú transzkriptóm-analízisre. A szekvenálási könyvtárak minőségét sztenderd minőségi mutatókkal jellemeztük: az egyedileg térképeződő readok aránya minden mintában 90% fölötti volt, és a duplikált readok aránya nem volt jelentősen eltérő a minták között. Ezt követően a readok GC-tartalmát, kromoszómális eloszlását, biotípus-eloszlását, géntest-lefedettségét és a bázisszintű mismatch rátát vizsgálva szintén arra a következtetésre jutottunk, hogy a kontroll és líofilezett mintákból készült könyvtárak nem mutatnak eltérést. Az eltérő génexpresszió vizsgálatokor magas korrelációt találtunk a két adatszett között ($r^2 = 0.99$), és 28 gén expressziója mutatott csökkenést a líofilezett mintákban (FDR = 0,05; amelyek közül 21 fehérjekódoló gén (FKG), 6 hosszú nemkódoló RNS és 1 pszeudogén). Az alacsonyabb expressziójú gének esetén az eltérés magasabb volt. Génontológiai analízissel a DNS templátú transzkripció ($P = 2,0 \cdot 10^{-5}$) mutatta a legnagyobb feldúsulást, ebbe a csoportba tartozott a POLR2A, az INTS1 és a KMT2D transzkripciók regulátor gének.

A líofilezett mintákban alacsonyabb szinten expresszálandó transzkriptek jellemzői

A következő lépésben célunk az eltérően expresszálandó transzkriptek sajátosságainak felfedése. A géntestek read-lefedettségét vizsgálva kimutattuk, hogy az eltérően expresszálandó gének többsége nem mutat 3' eltolódást ($P > 0,01$; $N = 16$), azonban 8 transzkript esetén szignifikáns pozitív, míg egy esetén szignifikáns negatív (az alacsonyan

expresszálódó *LINC01374*) korreláció mutatkozott a 3' végtől mért távolság és a read-lefedettség között ($P < 0,01$). Az alacsonyabban expresszálódó hosszú nemkódoló RNS-ek és fehérjekódoló RNS-ek (5'UTR+CDS+3'UTR) szignifikánsan hosszabbak voltak és magasabb GC-aránnyal rendelkeztek, mint a megfeleltetett humán transzkriptek ($P < 0,0001$, Mann-Whitney teszt). Emellett a CDS hossz, valamint a CDS és az UTR régiók GC-aránya szignifikánsan magasabb volt eltérően expresszálódó gének esetén ($P < 0,001$; Mann-Whitney teszt). Továbbá 12 fehérjekódoló gén (57%) tartalmazott legalább egy 3'UTR vagy intronikus AU-gazdag régiót (ARE) az ARED-Plus adatbázis alapján.

Izogenikus LCL-ek mikroszatellita-profilja, ploiditása és sejtciklus-progressziója

Elemzéseink eredményeként arra a megállapításra jutottunk, hogy a 26 éves egészséges férfi donorból származó öt izogenikus LCL diploid genotípussal rendelkezik (nem történt genom-multiplikáció), sejtciklus-eloszlása hasonló, valamint mikroszatellita-profiljuk alapján valóban egyetlen férfi donorból származnak kontaminációra utaló jel nélkül.

Az izogenikus sejtvonalak immunfenotipizálása

Áramlási citometriás immunfenotipizálással lehetőségünk nyílt az izogenikus LCL sejtvonalak forrás-sejttípusának azonosítására, a klonalitás meghatározására és a sejtfelszíni marker-szintű heterogenitás vizsgálatára. Kappa-lambda kizáráson alapuló vizsgálatunkban megállapítottuk, hogy egyetlen sejtvonal (sGT_4 sejtvonal; sGT = „azonos genotípus”) lambda+ monoklonális vagy pauciklonális vonal, míg a további sejtvonalak változó arányban mindkét könnyűláncot expresszálták: az sGT_2 és sGT_3 valószínűleg poliklonálisak, míg az sGT_1 és az sGT_5 szintén több klónból származott. A sejtvonalak érett B-sejtekre jellemző sejtfelszíni marker-mintázatot mutattak (CD19+ alacsony oldalszórással, CD20+, CD22+, CD23+, CD45+, HLADR+, dim FMC7+, dim CD21+, dim CD43+, CD5-, CD10-, CD34- és nTdT-). Meglepő módon a CD24 pán-B-sejt marker csupán 0,6-1,9%-ban volt jelen, amely valószínűsíthetően az EBV-fertőzés következménye. Bizonyos markerek nagy variabilitást mutattak az egyes sejtvonalak között – pl. a CD81 54-76%-ban, a CD38 20-87%-ban, míg a cIgM 1-84%-ban volt jelen egy-egy sejtvonalban.

Az izogenikus sejtvonalal genomi szabályozó elem-szintű variabilitása

Annak érdekében, hogy betekintést kapjunk a szabályozóelem aktivitás-szintű LCL variabilitásra azonos genotípus kontextusában, összehasonlítottuk a H3K27ac-jelintenzitásokat (aktív génszabályozó elemek) az izogenikus sejtvonalak között. Arra a megállapításra jutottunk, hogy bár a korrelációs koefficiensek kiemelkedően magasak voltak

($r^2 = 0,9-0,97$), a biológiai replikák klaszterekbe rendeződtek, egyedi epigenetikai profilra utalva. A 42.923 konszenzus régió közel negyede (9.685 hely) variábilis H3K27ac jelet mutatott az izogenikus LCL-ek között (RPKM FC > 2, P < 0,05 legalább két sejtvonal között). Sejtvonal-páronként 1.056-4.174 variábilis szabályozó elemet találtunk. Az intergenikus enhanszerek jelentős érintettségét figyeltük meg, míg az acetilációs szintek relatíve stabilak voltak a promóter elemeken.

Nagy kiterjedésű, koordinált szabályozó elem aktivitás-változás izogenikus LCL-ekben

A szuper-enhanszerek (SE-k) egymástól maximum 12.5 Kb-ra elhelyezkedő, akár több kilobázisnyi genomi szakaszon lineárisan csoportosuló szabályozó elemekből állnak. Az izogenikus LCL-ekből származó H3K27ac ChIP-Seq tagek poolozásával szuper-enhanszereket (SE) prediktáltunk. A vártak megfelelően a prediktált SE-k közül több B-sejt- és immunfunkciókban szerepet játszó gének (pl. PAX5 és IRF2), illetve EBV infekcióhoz köthető gének („EBV szuper-enhanszerek”, pl. BCL2, MIR155 és MYC) közelében helyezkedett el. Az 1.058 putatív SE 49%-a tartalmazott legalább egy variábilis enhanszert, azonban a szuper-enhanszer, mint egység csak 31 (2.9%) esetben volt variábilis (FC > 2, P < 0,05). A legközelebbi gének főként immunfunkciókhoz, például leukocita aktivációhoz (P = $5.5 \cdot 10^{-4}$) és leukocita sejt-sejt adhézióhoz (P = $5.6 \cdot 10^{-4}$) voltak köthetőek. A legnagyobb feldúsulást mutató molekuláris funkciók között találtuk a transzkripció faktor aktivitást (P = $1.2 \cdot 10^{-2}$) és az LPS-kötést (P = $3.2 \cdot 10^{-2}$).

Annak vizsgálatára, hogy a koordinált enhanszer-aktivitás-változás jelensége a SE-k keretein kívül is megvalósul, két módszert alkalmaztunk. Egyrészt a variábilis szabályozó elemeket topológiailag asszociált doménekkel (TAD, LCL HiC adat) vetettük össze, és arra az eredményre jutottunk, hogy nem SE-ként klasszifikált, hosszú genomi régiók mutatnak koordinált H3K27ac szignál-változást. A koordináció több TAD-on keresztül is megvalósulhat, illetve a jel irányultsága változhat egyik TAD-ról a másikra. Másrészt a P3H2 gén egy intragenikus szakaszát csaliként használva multiplex 3C-Seq kísérlettel igazoltuk, hogy a kromatin kontakt frekvencia csökkenése koordinált H3K27ac szignálesökkenéssel jár megabázis hosszúságú genomi régiókn.

A H3K27ac-variabilitás klinikailag releváns útvonalakat érintő transzkriptóm-variabilitást eredményez izogenikus LCL-ekben

A szabályozó elem aktivitás variabilitásának transzkriptómra gyakorolt hatását vizsgálva arra a megállapításra jutottunk, hogy a közeli gének expressziós mintázata bár

követi a H3K27ac aktivitási mintázatot, a gének csupán 4,6%-a (525) mutatott variábilis expressziót (FDR = 0,05, FC > 2). Páronként 25-229 variábilis gént detektáltunk (átlag = 119,8; medián = 107,5). A korábban EBV kópiaszámmal összefüggésbe hozott gének (pl. CXCL16, AGL és ADARB2) egyike sem mutatott variábilis expressziót, arra utalva, hogy a megfigyelt génexpressziós variabilitást nem a variábilis EBV kópiaszám okozza.

A változó gének esetén a biológiai folyamatokhoz tartozó gén ontológiai kifejezések közül a sejt migráció ($P = 2.8 \cdot 10^{-12}$), az intracelluláris jelátvitel ($P = 9.8 \cdot 10^{-9}$) és az apoptózis szabályozás ($P = 3.4 \cdot 10^{-8}$) míg a molekuláris funkciók közül az immun receptor és transzkripció faktor aktivitás mutatta a legerősebb feldúsulást. Meglepő módon 121 variábilis gén tartozott a farmakogének csoportjába (gyógyszerválasszal összefüggésben lévő gének) a Genetic Association Database alapján. Alacsonyabb átlagos génexpressziós szintekhez magasabb CV érték tartozott, az ide tartozó gének jelentős hányada receptor funkcióval bír, sejt felszínen lokalizálódik és jelátvitelben vagy sejt motilitásban van szerepe.

A nem genetikai eredetű variabilitás eltérő gyógyszerválaszhoz vezethet izogenikus LCL-ekben

Ezt követően megvizsgáltuk, hogy gyógyszerválaszban jelentkezik-e különbség izogenikus LCL-ek között, amely nagy jelentőséggel bírhat LCL-alapú farmakogenomikai kutatás esetén. A DPYD gént választottuk mint modell gént, mivel sGT_1 és sGT_2 között szignifikáns expressziós eltérés mutatkozott, illetve a kapcsolt 5-fluorouracil (5-FU) toxicitás MTT-alapú viabilitási esszével mérhető. A génexpressziós különbséget RT-qPCR-ral validáltuk két, RNA-Seq-től független RNS mintán. A sejteket különböző koncentrációban kezeltük 5-FU-val, majd 72 órát követően viabilitást mértünk. A magasabb DPYD expressziót mutató sejtvonal kevésbé mutatkozott érzékenynek az 5-FU kezelésre ($IC_{50} \text{ sGT}_1 = 0,63 \mu\text{M}$, $IC_{50} \text{ sGT}_2 = 1,21 \mu\text{M}$). Eredményeink alapján elmondhatjuk, hogy nem genetikai faktorok szerepet játszhatnak az LCL sejtek gyógyszerválaszára farmakogenomikai tanulmányokban.

DISZKUSSZIÓ

Fág-alapú kromatin immunprecipitációs spike-in kontrollok

A ChIP-qPCR és ChIP-Seq módszerek klinikai alkalmazásának gátat szabnak olyan technikai tényezők, mint az időigényes protokoll, a megvásárolható ChIP kitek szűk választéka, a jelentős változatosságot mutató protokollok és a „jó gyakorlatok” definiálásának hiánya, amely befolyással van a kísérletek reprodukálhatóságára. Míg a kiindulási sejtszámban jelentkező eltérések normalizálhatóak az input % módszerrel, a például pipettázási hibákból származó egyenlőtlen mintavesztés detektálatlan marad. Mivel a mintafeldolgozás jelentős része az input minta kivételét követően történik, az utóbbi komoly megoldatlan problémaként jelentkezik. Replikák használatával valamelyest csökkenthető a fentiek miatti torzító hatás, azonban a limitált mintaszám és magas kiindulási sejtszám-igény miatt nagyszámú replika használata nem minden esetben megoldható. Az RT-qPCR esetén használt háztartási génekhez hasonló belső kontrollok nem alkalmazhatóak, mivel a ChIP jel stabilitása különböző sejttípusok, kezelések és ChIP targetek kontextusában nem ismert. A hiányzó belső kontrollok alternatívájaként ún. spike-in kontrollok alkalmazhatóak.

Egy alkalmas spike-in folyamatkontroll olyan protein-DNS komplex(ek)et tartalmazó reagens, amely az immunprecipitációs lépéstől a ChIP fragmensek izolálásáig a minta része marad, illetve mennyisége a végső mintában egyszerűen meghatározható. Természetesen a spike-in kontroll nem versenghet a kromatin epitópokkal, amely például antitesttel fedett gyöngyök és fág külön reakcióban történő összekapcsolásával és spike-olásával, vagy indifferens epitóp-antitest párral történhet. Ennek segítségével a kísérleten belüli és kísérletek közötti torzítások is kiküszöbölhetőek. Fágbemutató segítségével mind *in vitro* evolúció, mint célzott klónozás segítségével kifejleszthetünk ChIP-grade antitestekhez erősen kötődő fágokat. A fág-alapú spike-in kontrollok legfőbb előnye a korábban leírt módszerekkel szemben az egyszerű sokszorosíthatóság egy baktériumtörzs fertőzését követő egyszerű tisztítási protokollal. Emellett monoklonális fág reagens is létrehozhatóak, amelyek konzisztens minőségben való előállítás tesznek lehetővé.

Az M13KE cirkuláris ssDNS genomja duplaszálú szabad végek hiányában ChIP-szekvenálási könyvtár készítésekor elvész, amely nem jelenthet problémát abban az esetben, ha a választott kvantitálási módszer ChIP-qPCR. Amennyiben a mintát elsősorban szekvenálásra szánjuk, a lineáris dsDNS genommal rendelkező T7 fág megfelelő alternatíva lehet, mint mimotóp-hordozó.

A disszertáció keretein belül bemutattuk egy fágbemutatáson alapuló spike-in folyamatkontroll fejlesztését, amely később a CHIP kísérletek külső kontrolljaként szolgálhat. *In vitro* evolúciós módszer segítségével kifejlesztettünk többféle fág reagenst humán eredetű kromatin-epitópok immunprecipitációjára alkalmas CHIP-grade antitestekkel (AR-ellenes, ER-ellenes (ösztrogén receptor), RXR α -ellenes (retinoid X receptor alfa) és CTCF-ellenes (CCCTC-binding factor) antitestek). Továbbá létrehoztunk monoklonális reagenseket az *in vitro* evolúcióval előállított, kevert fág könyvtárakból annak érdekében, hogy megakadályozzuk a poliklonális reagensek propagálásakor jelentkező, a reprodukálhatóságot befolyásoló fág-arány-változást. Megjegyzendő, hogy ideális esetben az immunizációra használt teljes hosszúságú peptid kódoló szekvenciájának fág genomba történő klónozása is megoldható, ám az immunizáló peptidek szekvenciája kereskedelmi forgalomban lévő CHIP-grade antitestek esetén gyakran nem elérhető. Ebben az esetben kizárólag az *in vitro* evolúciós módszerre szorítkozhatunk. Jelen dolgozat szerzője az AR-fággal kapcsolatos kísérletek tervezéséért, végrehajtásáért és felügyeletéért felelt, így eredményeink az AR-fág példáján kerülnek ismertetésre. Random heptapeptid könyvtár AR-ellenes antitesttel történő többkörös szelekciójával olyan poliklonális stockokat hoztunk létre, amelyek visszanyerése szimulált CHIP kísérletekben a szelekciós körök számával növekedett. Továbbá a legnagyobb affinitást mutató poliklonális stockból generált monoklonális reagensek mindegyike szintén magas affinitást mutatott az AR-ellenes antitesttel CHIP kísérletekben. A módszer fontos limitációja, hogy nem tudtunk előállítani magas affinitású fág stockokat a gyakori CHIP target poszt-transzlációsan módosított hisztonokra. A tárgyalt CHIP kontrollok klinikai kutatásokban, és később diagnosztikai vizsgálatokban történő használatához további kísérletek, például reprodukálhatósági mérések szükségesek.

Humán sejtek liofilezése és szobahőmérsékleten történő tárolása RNS-vizsgálatokhoz

A hagyományos biobankok, különösen az ultra-alacsony hőmérsékletű fagyasztókkal felszerelt facilitások működtetése jelentős és folyamatos anyagi ráfordítással jár, illetve nem elhanyagolható környezeti terhelést von maga után. A nemzetközi kollaborációk számának növekedése és a határokon átívelő mintaszállítás olyan szállítási módszert igényel, amely lehetővé teszi a biológiai minták integritásának megőrzését hőmérsékleti fluktuáció illetve hosszabb, például biztonsági vizsgálatok céljából történő csomagvisszatartás esetén. A liofilezés biológiai minták biztonságos szárítására és a szavatossági idő növelésére javasolt módszer egyes biológiai eredetű gyógyszerek esetén, de folyamatos vizsgálatok tárgya a liofilezés kiterjesztése emlős sejtek és szövetek stabilizálására is. Annak ellenére, hogy

néhány tanulmány RT-qPCR- és gélelektroforézis-alapú vizsgálatokkal kedvező RNS-stabilitásról számolt be liofilezett emlős sejtekben, a sejten belüli RNS stabilizálása céljából végzett liofilezés nem vált általánosan elterjedtté. Vizsgálatunk célja az emlős sejtek liofilezésével kapcsolatos ismereteink bővítése elsősorban RNS-alapú vizsgálatok kontextusában, például alacsony expressziójú gének, enhanszer-RNS-ek és a teljes transzkriptóm vizsgálatával.

Előzetes kísérleteink során trehalózt és EGCG-t tartalmazó IMT-2 oldatot alkalmaztunk lioprotektánsként, amelyet korábban hatékony sejtmembrán-stabilizátorként írtak le, lehetővé téve teljes sejteket igénylő laboratóriumi alkalmazásokat (pl. ChIP). A liofilezést követő rehidrációt követően mikroszkópiusan intakt, ám tripánkékekkel festődő és sejt kultúrában proliferációra nem képes LCL sejteket kaptunk. Az izolált RNS-ek a nem liofilezett kontrollokhoz képest alacsonyabb RIN értékekkel voltak jellemezhetőek, amelynek hátterében a sejtmosások során fellépő RNáz-mediált degradáció vagy a korábban LCL-ekben leírt EGCG-mediált apoptózis-indukció állhatott. Továbbá a totál RNS-sel koprecipitálódó EGCG gátolta a reverz transzkripciót, valamint gátolta egy steady-state állapotban nagy feldúsulást mutató pozitív kontroll régió kimutatását H3K27ac ChIP kísérletekben. Ennek oka lehet valamely kísérleti lépéssel való interferencia, esetleg az EGCG-re adott biológiai válasz. Eredményeink alapján arra a következtetésre jutottunk, hogy a gyors membrán penetrációt mutató, potens génexpressziós modulátor és RT inhibitor EGCG használatát későbbi kísérletekben mellőzzük a kísérletek összehasonlíthatósága érdekében.

A későbbi kísérletekben lioprotektánsként a viszonylag alacsony költségű trehalózt használtunk, amelyről korábban megállapították, hogy képes reaktív oxigénradikálok megkötésére és a sejtek védelmére dehidráció alatt. A költséghatékonyság szem előtt tartásával törekedtünk rövid liofilezési ciklus kialakítására (alacsony műszer energiafogyasztás). A sejtek beszárítása összesen hat órát vett igénybe, amely jelentősen rövidebb a gyógyszeriparban használt ciklushosszakhoz képest. Bár a trehalóz önmagában nem teszi lehetővé mikroszkópiusan intakt sejtek visszanyerését, a liofilezett termék könnyen beoldódott TRIzol reagensben, és izolálást követően magas minőségű (RIN) és mennyiségű RNS izolátumokat kaptunk akár két, vagy nyolc hét szobahőmérsékleten történő tárolást követően is. Különböző expressziós szintű mRNS-ek, illetve lncRNS-ek és eRNS-ek RT-qPCR alapú vizsgálatával megállapítottuk, hogy az expressziós szintek nem térnek el szignifikánsan a párosított kontrolloktól.

Számos tanulmány jutott arra a következtetésre, hogy a PCR-alapú módszerek kevésbé érzékenyek az RNS minta minőségére, így válogatott gének RT-qPCR-alapú módszerrel történő mennyiségi meghatározása nem feltétlenül tükrözi a transzkriptóm-szintű változásokat. Ezt figyelembe véve úgy döntöttünk, hogy RNS-szekvenálással is megvizsgálunk három, a liofilezést követően két hétig szobahőmérsékleten tárolt mintából izolált RNS-mintát párosított kontrolljaikkal. Számos minőségi mutató, például az egyedi leolvasások aránya, a leolvasások duplikációs rátája, GC-aránya, gén-, RNS-biotípus- és kromoszóma-lefedettsége, illetve a könyvtárkomplexitás igazolta a készített szekvenáló könyvtárak magas minőségét. Nem találtunk bázismodifikációra utaló jelet sem, amely a leolvasások illesztését negatívan befolyásolta volna. A 28 alulexpresszáldó gén az összes expresszáldó gén csupán 0,4%-a, alacsony medián fold-change értékkel (kontroll/lio). A génszettben az alacsony expressziójú gének általánosan nagyobb fold-change értékkel voltak jellemezhetőek, amit magasabb degradációs ráta, illetve az alacsony expressziójú gének magasabb expressziós sztochaszticitása és mérési hibája okozhatott. Az eltérően expresszáldó gének általában hosszabbak és magasabb G illetve C aránnyal rendelkeztek, valamint gyakran transzkripciós faktorokat kódoltak. Korábban kimutatták, hogy a hosszabb illetve TF-kódoló transzkriptek kevésbé stabilak mind *in vivo*, mind *in vitro*; azonban a harmadik kodon GC tartalma negatív korrelációt mutatott a degradációs rátával. A fent tapasztaltak és az ARE szekvenciák jelenléte egyes eltérően expresszáldó transzkriptekben arra utalhat, hogy reziduális szabályozott bomlási mechanizmusok okozhatták a 28 transzkript degradációját.

Az egy mintára vonatkoztatott éves költség egy korábbi jelentés szerint liofilezés esetén 3 EUR, -80°C-on történő tárolás esetén 24 EUR, míg folyékony nitrogénben történő tároláskor 31 EUR. A CoolSafe fagyasztva szárító készülék energiafogyasztását, illetve a trehalóz és a folyékony nitrogén árát figyelembe véve saját költségbecslésünk szerint 0,87 USD mintánként, amennyiben csupán egy mintát liofilezünk ciklusonként. Költségcsökkentés céljából célszerű lehet a minták ideiglenes gyűjtése liofilező oldatban -80°C-on, az egy liofilezési ciklusban maximálisan kezelhető mintaszám eléréséig.

Eredményeink alapján elmondhatjuk, hogy a trehalóz oldatban történő liofilezés és szobahőmérsékleten történő tárolás alkalmas módszer lehet RNS-alapú vizsgálatokra szánt humán sejtes minták archiválására.

Az LCL modell genotípus-független sejtfenotípus-variabilitása

A klasszikus és nagy átteresztőképességű molekuláris biológiai technikák kombinálásával lehetőség nyílik a népszerű sejtvonalakkal kapcsolatos általánosan elterjedt nézetek revizálására. Az LCL sejteket használó vizsgálatokban a sejtvonalak közötti különbségekre gyakran – kellően megalapozatlanul – személyek közötti különbségekként hivatkoznak. Néhány tanulmány foglalkozott genotípusfüggő kvantitatív kromatin jellegekkel, például kromatin-hurkolódással összefüggő koordinált kromatin-változásokkal. Azonban a genetikailag eltérő LCL sejtekből álló modelljeik nem voltak alkalmasak a genotípustól független illetve a QTL szignifikancia-határértéket meg nem haladó genotípusfüggő változások közötti különbségtételre.

Vizsgálatunkban azt a célt tűztük ki, hogy megvizsgáljuk az LCL-ek cisz szabályozó elem- és transzkriptóm-szintű variabilitását öt genetikailag azonos LCL sejt felhasználásával. Az izogenikus sejtvonalak előállítása szimulálta a kereskedelmi forgalomban elérhető genetikailag eltérőekét. A vizsgálataink következtében jelentkező variabilitás minimalizálása érdekében minden LCL-t azonos passzázs- és fagyasztás-olvasztási ciklusszámnál vizsgáltunk (háromszintű biobank). A sejtvonalakot párhuzamosan kezeltük, a sejtbegyűjtést a nap azonos szakában végeztük, illetve a random fluktuációk zavaró hatásának megelőzésére biológiai replikákat alkalmaztunk. A modell így az LCL sejtvonalak létrehozása és rövidtávú tenyésztése során kialakuló variabilitást tükrözi.

A sejtvonalak mindegyike euploid és X/Y nemi kromoszómákat tartalmaz, azonos mikroszatellita-profillal és hasonló sejtciklusbeli eloszlással. A sejtvonalak érett B-sejt fenotípussal rendelkeztek és egyetlen sejtvonala kivételével nem mutattak monoklonalitást. Korábbi vizsgálatok szerint a diverzitás szűkülése főként a sejtvonala létrehozásának korai szakaszában jellemző, és kevésbé van rá hatással a későbbi tenyésztés. Feltételezésünk szerint az eredeti B-sejt populáció heterogenitását a nem monoklonális sejtvonalak jobban tükrözik, az egy anyasejtből származó sejtvonala divergenciája mellett. Ezzel szemben a monoklonális sejtvonala vizsgálatunkban nem mutatott kiugró tulajdonságokat.

Reprodukálható H3K27ac mintázatuk alapján egyértelműen elkülöníthető az öt izogenikus sejtvonala. Meglepő módon a vizsgált régiók (9.685 szabályozó elem) közel negyede mutatott variábilis H3K27ac szignált legalább két sejtvonala között. Az intergenikus enhanszerek érintettsége volt a legnagyobb arányú, míg a promóterek aktivitása viszonylag stabilnak mutatkozott. Ez az eredmény párhuzamba állítható korábbi vizsgálatok eredményeivel, amelyek megállapították, hogy a promóterek rezisztensebbek a rövidtávú és

evolúciós skálájú perturbációkkal szemben. Ez a promóter robosztusság részben magyarázható az enhanszer redundancia jelenségével (“árnyék enhanszerek”), azaz egy enhanszer „kikapcsolása” esetén a többi, adott promóterhez kötődő enhanszer továbbra is stabilan tartja a promóter aktivitását és a génexpressziót. Ez a jelenség magyarázhatja azt is, hogy a szuper-enhanszerek egyes alkotó enhanszereinek variabilitása nem okoz szignifikáns génexpressziós eltérést. A promóterek alacsony érintettsége nyomán nem meglepő módon a poly(A)+ RNS-ek esetén alacsony variabilitást találtunk. Ez összhangban van genetikailag eltérő LCL-ekben és élesztőben végzett korábbi tanulmányok eredményeivel.

Egy szövet egyedi sejtjei nagyfokú funkcionális genomikai heterogenitást mutathatnak. A heterogenitás által lehetővé váló fenotípusos plaszticitás hatékonyabb választ tesz lehetővé kiszámíthatatlan külső hatások jelentkezésekor, támogatva a túlélést. Az LCL-ekben megfigyelt variabilitás a klónok egy alcsoportjának szelektív EBV-fertőződéséből, illetve a leszármazott vonalak növekedési különbségéből fakadhat. A magas vagy alacsony citokin-szekréción mutató klónok dominanciája további eltolódást okozhat a populáció fenotípusában. Génontológiai elemzéssel feltártuk, hogy a változó gének között immunfunkciókban szerepet játszó gének és sejtfelszíni receptorok dúsultak fel, amely összhangban van korábbi tanulmányok eredményeivel, amelyek a B-sejt-specifikus sejtfelszíni receptorok kiemelten magas splicing variabilitására mutattak rá.

A farmakogének jelentős érintettsége arra enged következtetni, hogy vizsgálatainknak nem csupán molekuláris QTL-ek, de farmakogenomikai QTL-ek térképezésében is jelentősége lehet. Kísérleteinkben az 5-FU kemoterápiás szerrel szembeni érzékenység korrelációt mutatott a *DPYD* gén expressziójával. Megjegyzendő, hogy korábbi tanulmányok rámutattak több olyan faktorra, például a növekedési rátára, illetve a feltételezhetően örökletes EBV kópiaszámra és ATP szintekre, amelyek befolyásolhatják az LCL-ek gyógyszerre adott válaszát. Tanulmányunk során ezen faktorokat nem vizsgáltuk, így nem zárható ki ezek torzító hatása. Vizsgálataink megerősítették az RNS-expressziót is figyelembe vevő farmakogenomikai „háromszög modell” létjogosultságát. Az elérhető izogenikus sejtvonalak alacsony száma limitálja eredményeink extrapolálhatóságát kiterjedt LCL panelekre.

Vizsgálatunk rávilágít az LCL-ek génszabályozó elem- és génexpressziós szintű variabilitásának jellegére és mértékére, illetve ezek farmakogenomikai vizsgálatokra vonatkozó jelentőségére. A fenti eredmények ellenére véleményünk szerint az LCL sejtvonalak továbbra is a QTL vizsgálatok jelentős modelljei, és a felfedett limitációk racionálisabb kísérlettervezésre adnak lehetőséget.

ÖSSZEFOGLALÁS

A funkcionális genomikai módszerek, például a ChIP-Seq és az RNA-Seq forradalmasították a génszabályozással kapcsolatos vizsgálatokat. Kutatásaink során a funkcionális genomika eszköztárához kapcsolódó vizsgálatokat végeztünk a következő szempontokat szem előtt tartva: kulcsmódszerek fejlesztése, modellrendszerek jellemzése és költségracionalizálás.

A ChIP módszer nem rendelkezik általánosan elfogadott, a kísérletes mintavesztéség normalizálására használható kontrollokkal. Kísérleteink során M13 bakteriofág-alapú peptidkönyvtárból kiindulva sikeresen feldúsítottunk kromatin epitópotkat utánzó bakteriofágotkat *in vitro* evolúció segítségével. A kromatin mintákhoz kevert fág részecskék nagy affinitással kötődtek a ChIP-minőségű antitestekhez, és kvantitálhatóak voltak ChIP eluátumokból qPCR segítségével. Mindezek alapján a bemutatott módszer kiindulópontként szolgálhat spike-in kontrollok előállításához különböző ChIP-minőségű antitestekhez.

Az RNS-szekvenálás egyre jelentősebb szerepet játszik az orvosbiológiai kutatásokban, ám a fagyasztott szövetek tárolásból eredő magas működtetési költségek miatt a szobahőmérsékleten történő mintatárolásra alkalmas módszerek fejlesztése nagy jelentőséggel bírhat. Kísérleteink során sejtek liofilezéssel történő stabilizálásának hatását vizsgáltuk RNS-alapú mérésekre, mint potenciálisan költségcsökkentő alternatíva. Míg az epigallokatekin-gallát nem bizonyult alkalmas lioprotektánsnak RNS-alapú vizsgálatokhoz, a trehalóz megfelelő védelmet nyújtott liofilezés és többhetes szobahőmérsékleten történő tárolás során, magas kitermelést és kiváló RNS-minőséget eredményezve mind alacsony-, mind magas áteresztőképességű vizsgálatokhoz.

A molekuláris- és gyógyszer-QTL-ek térképezésére gyakran használt humán LCL sejtvonalak intrinzik epigenomi és transzkriptóm-szintű variabilitása korábban ismeretlen volt. Öt, azonos személyből származó LCL sejtvonal vizsgálata során kimutattuk, hogy az aktív (H3K27ac-jelölt) génszabályozó elemek közel negyede mutat variábilis acetilációt, amely enyhébb transzkriptóm-variabilitással párosult. Továbbá a génexpressziós eltérések kemoterápiás szerekkel szembeni eltérő sejtválaszt okozhatnak. Emiatt javasolt az RNS-szintek beépítése az LCL-alapú QTL-vizsgálati tervekbe.

Összegezve, eredményeink kiindulási pontként szolgálhatnak költséghatékonyabb és racionálisabb kísérlettervezéshez a funkcionális genomika keretein belül.

KÖSZÖNETNYILVÁNÍTÁS

Köszönetemet szeretném kifejezni témavezetőmnek, Dr. Bálint Bálint Lászlónak töretlen szakmai és személyes támogatásáért B.Sc, M.Sc. és PhD képzésem során. Külön szeretném kifejezni hálámat azért, hogy lehetővé tette számomra a legváltozatosabb kísérletes és bioinformatikai módszerek elsajátítását, ösztöndíjak megpályázásában és konferencia részvételekben nyújtott támogatását, valamint azt, hogy támogatta angol-magyar orvos- és egészségtudományi szakfordítói képzésemet a PhD képzés alatt.

Szeretném megköszönni Dr. Horváth Attilának személyes támogatását és azt, hogy aktívan közreműködött bioinformatikai ismereteim fejlesztésében és elsőszerzős publikációim létrejöttében.

Hálával tartozom hallgatótársaimnak, Bojsuk Dórának, Erdős Edinának, Csumita Máriának, Czipa Eriknek és Halász Lászlónak inspiráló és gondolatébresztő szakmai beszélgetéseinkért és a sok együtt töltött vidám pillanatért.

Köszönettel tartozom Dr. Eric Solernek, a Laboratory of Molecular Hematopoiesis (Inserm UMR 967, Fontenay-aux-Roses, Franciaország) vezetőjének és Tharshana Stephennek, hogy lehetőséget kaptam gyakorlatot szerezni a multiplexált 3C-Seq módszerben Campus Hungary-ösztöndíjas időszakom alatt.

Továbbá köszönettel tartozom szakmai támogatásukért a Genomi Medicina és Bioinformatikai Szolgáltató Laboratórium munkatársainak: Dr. Póliska Szilárdnak, Mátyás Erzsébetnek, Pallér Ádámnak, Nagy Évának, Kerekes Tamásnak, a Nagy Laboratórium munkatársainak: Dr. Ixchelt Cuaranta-Monroy-nak és Fürtös Ibolyának, valamint a Proteomikai Szolgáltató Laboratórium munkatársainak: Dr. Csősz Évának, Sólyom Kamillának és Darai Zoltánnának, illetve a Laboratóriumi Medicina intézet munkatársainak: Dr. Hevessy Zsuzsannának és Kálmáncheyné Gombos Évának.

Hálás vagyok Dr. Katarzyna Blaszczyk és Dr. Daina Skiriutė vendégkutatóknak, hogy egymást kiegészítő munkatársakká és barátokká válhattunk.

Végül végtelen hálával tartozom imádott családomnak, különösen édesanyámnak, édesapámnak és nővéremnek feltétel nélküli támogatásukért és azért a stabil háttérért, amely lehetővé tette a kitűzött céljaim elérését.



**DEBRECENI
EGYETEM**

**DEBRECENI EGYETEM
EGYETEMI ÉS NEMZETI KÖNYVTÁR**

H-4002 Debrecen, Egyetem tér 1, Pf.: 400
Tel.: 52/410-443, e-mail: publikaciok@lib.unideb.hu

Nyilvántartási szám: DEENK/289/2019.PL
Tárgy: PhD Publikációs Lista

Jelölt: Ozgyin Lilla
Neptun kód: RPI5U5
Doktori Iskola: Molekuláris Sejt- és Immunbiológia Doktori Iskola

A PhD értekezés alapjául szolgáló közlemények

1. Keresztessy, Z., Erdős, E., **Ozgyin, L.**, Kádas, J., Horváth, J., Zahuczky, G., Bálint, B. L.:
Development of an antibody control system using phage display.
J. Biotechnol. 300, 63-69, 2019.
DOI: <http://dx.doi.org/10.1016/j.jbiotec.2019.05.009>
IF: 3.163 (2018)
2. **Ozgyin, L.**, Horváth, A., Hevessy, Z., Bálint, B. L.: Extensive epigenetic and transcriptomic
variability between genetically identical human B-lymphoblastoid cells with implications in
pharmacogenomics research.
Sci Rep. 9, 1-16, 2019.
DOI: <http://dx.doi.org/10.1038/s41598-019-40897-9>
IF: 4.011 (2018)
3. **Ozgyin, L.**, Horváth, A., Bálint, B. L.: Lyophilized human cells stored at room temperature
preserve multiple RNA species at excellent quality for RNA sequencing.
Oncotarget. 9 (59), 31312-31329, 2018.
DOI: <http://dx.doi.org/10.18632/oncotarget.25764>





**DEBRECENI
EGYETEM**

**DEBRECENI EGYETEM
EGYETEMI ÉS NEMZETI KÖNYVTÁR**

H-4002 Debrecen, Egyetem tér 1, Pf.: 400
Tel.: 52/410-443, e-mail: publikaciok@lib.unideb.hu

További közlemények

4. Horváth, A., Dániel, B., Széles, L., Cuaranta-Monroy, I., Czimmerer, Z., **Ozgyin, L.**, Steiner, L., Kiss, M., Simándi, Z., Póliska, S., Giannakis, N., Raineri, E., Gut, I. G., Nagy, B., Nagy, L.: Labelled regulatory elements are pervasive features of the macrophage genome and are dynamically utilized by classical and alternative polarization signals.
Nucleic Acids Res. 47 (6), 2778-2792, 2019.
DOI: <http://dx.doi.org/10.1093/nar/gkz118>
IF: 11.147 (2018)
5. **Ozgyin, L.**, Erdős, E., Bojcsuk, D., Bálint, B. L.: Nuclear receptors in transgenerational epigenetic inheritance.
Prog. Biophys. Mol. Biol. 118 (1-2), 34-43, 2015.
DOI: <http://dx.doi.org/10.1016/j.pbiomolbio.2015.02.012>
IF: 2.581
6. Blaszczyk, K., Olejnik, A., Nowicka, H., **Ozgyin, L.**, Chen, Y. L., Chmielewski, S., Kostyrko, K., Wesoly, J., Bálint, B. L., Lee, C. K., Bluysen, H. A.: STAT2/IRF9 directs a prolonged ISGF3-like transcriptional response and antiviral activity in the absence of STAT1.
Biochem. J. 466 (3), 511-524, 2015.
DOI: <http://dx.doi.org/10.1042/BJ20140644>
IF: 3.562
7. Franyó, D., Boros Oláh, B., **Ozgyin, L.**, Bálint, B. L.: Befolyásolja-e az életmód génjeink működését?: az epigenetikai kutatások irányvonalai és eredményei.
LAM KID. 2 (1), 37-42, 2012.

A közlő folyóiratok összesített impakt faktora: 24,464

A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre): 7,174

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudományometriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.



Debrecen, 2019.07.30.