

dc_1977_21

MTA doktori értekezés tézisei

Új dimenziók a genomszerkezet kutatásban

Dr. Székvölgyi Lóránt



Biokémiai és Molekuláris Biológiai Intézet

Általános Orvostudományi Kar

Debreceni Egyetem

2022

Tartalom

BEVEZETÉS	4
A GENOM FELÉPÍTÉSÉNEK SZERKEZETI ALAPVONÁSAI.....	5
A NUKLEOSZÓMA	6
AZ R-HUROK	8
A HOMOLÓG REKOMBINÁCIÓ MECHANIZMUSA.....	9
TÉRBELI GENOMSZERKEZET ÉS HOMOLÓG REKOMBINÁCIÓ	10
CÉLKITŰZÉSEK.....	12
ANYAGOK ÉS MÓDSZEREK.....	13
Élesztő mutánsok.....	13
<i>Arabidopsis</i> minták.....	13
Humán sejt kultúra.....	14
Meiózis és sporuláció élesztőben	14
Southern blot és DSB analízis	14
Élesztő 2-hibrid screen.....	14
Rekombináns fehérje expresszió és GST-pull down	15
Kromatin immunprecipitáció és kompetíciós kromatin immunprecipitáció (ChIP és c-ChIP) szekvenálás.....	15
Kromatin turnover ráta becslése c-ChIP adatokból	16
Meiotikus transzkriptom microarray	16
RNS-DNS hibrid immunprecipitáció (DRIP).....	17
A DRIP és ChIP peak-ek genomi annotációja.....	18
KisRNS szekvenálás (sRNS-seq)	18
<i>In situ</i> Hi-C.....	18
<i>In silico</i> restrikciós enzim emésztés	19
Konfokális lézerpásztázó mikroszkópia (CLSM)	19
Fluoreszcencia-visszatérés fotoelhalványodás után (FRAP).....	19
Fluoreszcencia-korrelációs spektroszkópia (FCS).....	20
Farmakogenomikai elemzések.....	20
EREDMÉNYEK ÉS MEGBESZÉLÉSÜK	21
A Set1 komplex szerepe a meiotikus DNS törések kialakulásában.....	21
Az Spp1 kromatinkötődési dinamikájának vizsgálata a meiózis során.....	23

A hiszton H3K56 acetiláció szerepe a meiotikus homológ rekombináció iniciációjában.....	25
Kromoszómális R-hurkok és R-hurok regulátorok molekuláris vizsgálata.....	26
ÚJ MEGÁLLAPÍTÁSOK.....	33
KÖSZÖNETNYILVÁNÍTÁS.....	35
A DISSZERTÁCIÓ ALAPJÁT KÉPEZŐ SAJÁT KÖZLEMÉNYEK.....	36

BEVEZETÉS

A természet máig megoldatlan talánya, hogy vajon hogyan tekeredik fel egy közel két méter hosszú makromolekula, az emberi DNS, egy körülbelül 5 μm átmérőjű sejtmag parányi dimenziójában? Az eukarióta sejtek örökítő egységei, a kromoszómák, nem csupán kétszálú dezoxiribonukleinsav láncok, hanem DNS-ből, hiszton+nem-hiszton fehérjékből és ribonukleoproteinekből álló makromolekuláris komplexek, amelyek számos nyitott (vagyis rugalmas) és zárt (rigid) kromatin kompartmentet alakítanak ki az interfázisos sejtmagban. Előbbi régiókat klasszikusan eukromatinnak nevezzük, míg az utóbbi képlet neve heterokromatin. Ezen felül, a kromatin szerkezeti alapegysége, a nukleoszóma, nagyszámú és kémiaiag rendkívül változatos poszttranszlációs módosításokon megy keresztül, amelyek kombinatív módon, a különféle sejtbiokémiai funkcióknak megfelelően finomhangolják a nyitottabb és zártabb kromatin régiók szerkezetét és működését, szűkítve a lehetséges feltekeredési útvonalak számát. A kromatin emiatt rendkívül dinamikus és adaptív, vagyis belső és külső ingerekre képes drasztikusan megváltoztatni térbeli konformációját.

A genom flexibilitását növelő fontos tényező a DNS molekula egyik vagy mindkét szálát megszakító folytonossághiányok (SSB - szimplaszál törés; DSB - duplaszál törés) és speciális, ún. "nem-B DNS" struktúrák (pl. R-hurkok) jelenléte. DNS hibák, törések, és R-hurkok gyakorlatilag folyamatosan keletkeznek a sejtekben a különféle sejtélettani és/vagy patológiás folyamatok révén, ám azokat hatékony hibakorrekciós rendszerek és replikatív hiba-elkerülő útvonalak javítják ki. Megjegyzendő, hogy a DNS törésekre általában, mint valamilyen patológiás állapotra gondolunk (részben jogosan, hiszen azok növelik a genom instabilitását, amely pl. a rosszindulatú sejtek jellemzője), viszont az utóbbi évek kutatásai számos olyan fiziológiás folyamatot tártak fel (Niehrs and Luke 2020), amelyek genetikailag programozott kromoszómatörések vagy R-hurkok létrejötte nélkül nem játszódhatnának le (Santos-pereira and Aguilera 2015; Szekvolgyi and Nicolas 2010). A genom szerkezetével kapcsolatban azonban számos megválaszolatlan (vagy részben feltárt) kérdéssel szembesülünk: mennyire specifikus a kromatin feltekeredése eltérő sejt típusokban; mekkora az egyedi sejtek közötti variabilitás; hogyan változik a 3D kromatin mintázat a szöveti differenciáció, az öregedés, és különféle patológiás állapotok során; van-e különbség a diploid testi sejtek és a haploid (meiotikus) ivarsejtek kromatinszerkezeti sajátosságai között; milyen összefüggés van a magasabbrendű kromatinszerkezet, a DNS léziók és R-hurok struktúrák kialakulása között; melyek a térbeli genomszerkezet főbb molekuláris komponensei, különös tekintettel a hiszton módosító

enzimek, hiszton modifikációk és nem-kódoló RNS-ek szerepére? A következő pontokban ezekre a kérdésekre keressük a választ.

A GENOM FELÉPÍTÉSÉNEK SZERKEZETI ALAPVONÁSAI

A virális genom igen szorosan pakolódik: a kapszid térfogata majdnem megegyezik a benne lévő nukleinsav térfogatával (Krebs et al. 2018). Ezzel szemben a baktériumok genetikai állománya a sejtterefogat töredékét foglalja el. *E. coli* sejteket lizálva tömegesen izolálhatunk a széttöredezett sejtburkokhoz kihorgonyozott, fehérjékkel-asszociált, hurokszerű DNS rostokat. Etídium-bromid (EtBr) titrálásos mérések alapján az izolált DNS rostok ún. zárt-cirkuláris negatív szuperhélix formában vannak, amely a DNS molekula természetes alapállapota (Calladine 2004). Az izolálás során véletlenszerű egyszáltörések (ún. nick-ek) keletkeznek a nukleoid DNS-ben, amelyek *elméletileg* megszüntetik az EtBr által létrehozott energetikailag kedvezőtlen pozitív szuperhelikális állapotot. Ám a valóságban nem ez történik: a nukleoid nem relaxálódik, hanem megtartja az EtBr okozta pozitív szupertekercs állapotát, ami azt jelenti, hogy a bakteriális genom számos egymástól független doménből áll. A genomban körülbelül 100 ilyen egység van, melyek mindegyike ~43 kb (13 μ m) nagyságú autonóm hurkot alkot (Krebs et al. 2018). E topológiai domének egymástól függetlenek, mivel a szuperhelicitás változása nem propagálódik egyik hurokról a másikra; így eltérő szuperhélix állapot tartható fenn a baktérium genom különböző részein, amely alapvetően befolyásolja a különféle gének/géncsoportok koordinált expresszióját vagy csendesítését.

Ha eukarióta sejtekből sóextrakciós eljárással izolált genomon végezzük el a fenti EtBr titrálást, hasonló következtetésre jutunk: egy fehérje-mátrixhoz kötődő illetve abból kibomló kromatin hurkok tömege figyelhető meg (Berezney and Coffey 1974), melyben a negatív szuperhélix egységek relaxációjához ~85 kb-onként 1 nick szükséges (Calladine 2004). A "nettó" alultekert állapot energetikailag megkönnyíti a DNS szálak szétválását például a transzkripció iniciációja vagy a pre-replikációs komplexek összeszerelése során (Naughton et al. 2013). A fehérje-mátrixot az alkalmazott extrakciós eljárástól függően scaffold-nak, nukleoszkeletonnak, vagy "halo"-nak hívják, jellegzetes fehérje komponensei evolúciósan konzerváltak prokariótáktól az emberig (Sureka and Mishra 2021). A mátrix fehérjék közül a topoizomeráz II enzim kitüntetett szerepet játszik a DNS szuperhelicitásának szabályzásában és a transzkripció/replikáció/rekombináció során felgyülemelő torziós-feszüléssel járó többletenergia eloszlásában (Pommier et al. 2016; Razin et al. 2014; Zhang et al. 2014; Gasser et al. 1986; Adolph et al. 1977). Fontos megjegyeznünk, hogy a transzkripció és DNS replikáció önmagában indukálja a szuperhelikális domének kialakulását, amelyek szabályozása

- a topoizomerázokon kívül - a DNS-nukleoszóma interakciók kontrollált megváltoztatásával lehetséges. Nukleoszóma beépülés hatására a DNS feszített állapotba kerül, amelyet energetikailag a hiszton-DNS kölcsönhatások tartanak fenn. A nukleoszóma disszociációja viszont újrendezi a kialakult szuperhelikális struktúrát, melynek eredményeképpen a DNS alultekertté válik.

A NUKLEOSZÓMA

A nukleoszóma két-két kópia H2A, H2B, H3 és H4 hisztonból és ezen oktamer köré tekeredő 145-147 bp szuperhelikális DNS-ből áll (Olins and Olins 1974). A nukleoszóma központi ("core") részét és a vele kölcsönható DNS-t 1997-ben röntgenkristallográfias módszerrel karakterizálták (Luger et al.). A tanulmány részletesen ismerteti a DNS kettős hélix kis árok és a hiszton-fold domén közötti kölcsönhatásokat, többek között elektrosztatikus kölcsönhatásokat, hidrogénkötéseket és apoláris kölcsönhatásokat.

Proteomikai és genomikai tulajdonságaik alapján a nukleoszóma hisztonfehérjék két csoportra oszthatók: általános hisztonok (H2A, H2B, H3, H4) és hiszton-variánsok (H2A.X, H2A.Z, H3.3, CENP-A, stb.). Eukariótákban az általános hisztonokat kódoló gének esszenciálisak (deléciójuk letális) és tandem ismétlődő tömbökben csoportosulnak a genom különböző pontjain. Mivel számosságuk meghaladhatja a több ezres nagyságrendet, funkcionális vizsgálatuk rendkívül komplikált (ez alól kivételt képez néhány egyszerűbb eukarióta modellszervezet (pl. pékéllesztő), amelyben speciális genetikai eljárásokkal megoldható a hiszton klaszterek kiütése és azok rekombináns hiszton génekkel történő helyettesítése.

A nukleoszóma további jellemzői a flexibilis hisztonvégek, melyek áthaladnak a DNS szuperhélix felett és között. A nukleoszóma ezen N- és C-terminusai szabályozó szerepet töltenek be a különféle poszttranszlációs hiszton módosítások révén, ám mechanikai szerepük is van a kromatindinamika és a nukleoszóma-konformáció szabályozásában. A főbb hiszton módosítások közé tartozik a lizin acetiláció/metiláció/ubiquitináció/szumoiláció, arginin metiláció, és szerin foszforiláció, de ismertek egyéb, "különleges" módosítások is, például az arginin citrullináció (deimináció) (Zhu et al. 2021) vagy glutamin szerotoniláció (Farrelly et al. 2019). Ezen kémiai csoportok nem csupán a hisztonok túlnyúló N- és C-terminális oldalláncaira kerülhetnek, hanem a DNS-sel közvetlen kapcsolatban álló nukleoszóma felszínekre is, amelyre példa a DNS be- és kilépési pontjához (ún. "dyad axis"-hoz) közel eső H3 lizin 56 acetiláció (Yang et al. 2008).

A hiszton módosítások közül az egyik legabundánsabb és leginkább vizsgált aktiváló modifikáció a H3K4 trimetiláció (H3K4me3), amely tipikusan a gének 5'-végén helyezkedik el, és rendkívül konzervatív, élesztőktől az emberig az aktív génátírást stimuláló epigenetikai jel. *Saccharomyces*

Cerevisiae-ben a H3K4 metilációt egyetlen lizin-specifikus metiltranszferáz végzi, amelynek neve Set1C / COMPASS (Complex of Proteins Associated with Set1) (Miller et al. 2001; Dehe and Geli 2006). Az Set1C önmagában katalitikusan inaktív; mono- di- és trimetiláz aktivitásaihoz további fehérjék kapcsolódása szükséges, amelyek élesztőben az Spp1, Bre2, Swd1, Swd2, Swd3, Sdc1 és Shg1. A Swd1 és Swd3 elengedhetetlen a komplex stabilitásához és a Set1 katalitikus alegység működéséhez, ezek hiányában a Set1C destabilizálódik és a H3K4 metiláció nem megy végbe. Az Swd2 alegység a di- és trimetilációhoz szükséges (a monometilációhoz nem), amelynek viszont előfeltétele egy másik hiszton módosítás, a Rad6/Bre1 E3 ubiquitin ligáz által katalizált H2B lizin 123 mono-ubiquitináció (H2BK123ub) (Vitaliano-Prunier et al. 2008; Kim et al. 2009; Jeon et al. 2018). (Ez a mechanizmus izgalmas például szolgál a különféle hisztonmodifikációk és hisztonmódosító enzimek funkcionális interakciójára.) A Set1 és Swd2 alegységek N-terminális régiói kooperatív kölcsönhatást alakítanak ki az RNS polimeráz II C-terminális doménjével (CTD), amely a Set1C/COMPASS működését a transzkripció elongációjához kapcsolja (Bae et al. 2020). Az Swd2 ezenkívül elősegíti a transzkripció terminációjában résztvevő komplexhasítási és poliadenilációs faktor (CPF) működését, tovább erősítve a Set1C komponensek szerepét a génexpresszió szabályozásában (Cheng et al. 2004). A Bre2, Sdc1, és Spp1 alegységek a H3K4 trimetilációhoz szükségesek, lényegében ismeretlen mechanizmussal.

Az Spp1 a katalitikus (Set1) domén teljesen-aktív konformációjának kialakulásához szükséges; különlegessége, hogy H3K4 di/trimetil olvasó funkcióján kívül a H3 arginin 2 oldallánc aszimmetrikus dimetilációját (H3R2me2a) is felismeri és megköti a PHD-finger doménjén keresztül (Santos-Rosa et al. 2009; Nakanishi et al. 2008). Az Spp1-H3R2me2a interakció azonban kizárja a H3K4 metilációt (Guccione et al. 2007; Iberg et al. 2008), annak előbb le kell kerülnie a kromatinról a Set1C katalitikus aktivációjához. H3R2me2a-mentes nukleoszómákon az Spp1 stabil kölcsönhatásokat tud kialakítani a mono- di- és trimetilált H3K4 oldallánccal, amely így egyrészt védi a 4-es lizint a demetilázok hatásától, másrészt elfedi az R2 oldalláncot az arginin-metilázok elől. Nem meglepő, hogy a két hisztonmódosítás ellentétes genomi eloszlást mutat, ugyanis a H3R2me2a a heterokromatikus régiókra, míg a H3K4me az eukromatinra jellemző. Tovább árnyalja a képet, hogy a szimmetrikus arginin dimetiláció (H3R2me2s) viszont elősegíti a H3K4me olvasó egységek bekötődését a di/trimetilált H3K4 hisztonokhoz (Migliori et al. 2012) és ezzel serkenti a transzkripció aktivációját (a szabályozás részletei azonban kevésbé ismertek). A Set1C egyéb funkcióiról és az Spp1 további interakciós partnereiről a meiotikus homológ rekombinációról szóló fejezetben lesz szó részletesebben.

A represszív hisztonmodifikációk közül kiemelendő a heterokromatin keletkezésben és epigenetikai továbbörökítésében kulcsszerepet játszó H3 lizin 27 metiláció (H3K27me), valamint a DNS metiláz rendszerrel szorosan együttműködő H3 lizin 9 metiláció (H3K9me) (Allshire and Madhani 2017). E módosítások részt vesznek a heterokromatin alapvető biológiai funkcióiban: i) a genom védelmében a kontrollálatlan transzpozonokkal és endogén retrovírusokkal szemben, ii) a pluripotencia és sejtdifferenciálódás szabályozásában a transzkripció faktorok által közvetített kromatinszerkezeti változások és sejt-átprogramozás akadályozásával, iii) a mitotikus sejtosztódás szabályozásában a centromerikus kinetokor komplexek összeszerelődésén keresztül.

AZ R-HUROK

Az R-hurkok háromszálú nukleinsav struktúrák a genomban, amelyek egy RNS-DNS hibridből és egy kiszorított, szimplaszálú DNS-ből (ssDNS) állnak). Mára nyilvánvalóvá vált, hogy az R-hurkok az életfa minden szintjén előforduló epigenetikai elemek, élesztőtől az emberig (Sollier and Cimprich 2015). Míg korábban csupán transzkripció melléktermékként tekintettek rájuk, az utóbbi évtizedben több fiziológiás, illetve patológiás folyamattal is kapcsolatba hozták őket (Niehrs and Luke 2020).

Normál körülmények között a promóterek közelében lévő CpG szigetek metilációjának gátlásával vagy nem-kódoló RNS-eket toborozásával szabályozzák a génexpressziót (Arab et al. 2019; Ginno et al. 2012; Boque-Sastre et al. 2015; Sun et al. 2013), intermediert biztosítanak a T4 bakteriofág, az *Escherichia coli* plazmid, illetve a cirkuláris mitokondriális DNS replikációjához (Aguilera and García-Muse 2012) és a kromoszóma telomerek dinamikájához (Balk et al. 2013), továbbá B limfocitákban az R-hurkok ssDNS szála enzimikus hatásoknak illetve mutációknak kitéve hozzájárul az immunglobulin nehézlánc izotípus váltásához (Yu et al. 2003). Patológiás körülmények között azonban az R-hurkok felhalmozódása súlyos veszélyt jelent a kromoszómák integritására, amely genom instabilitáshoz és rákos folyamatokat elindulásához vezethet (Crossley et al. 2019; Wells et al. 2019; Boque-Sastre et al. 2015). Az R-hurkokat különféle neurodegeneratív betegségekkel is összefüggésbe hozták, úgymint az Amiotrófiás Lateral Sclerosis (ALS), Aicardi-Goutieres szindróma, Friedreich ataxia, és a Fragilis X szindróma. Jelenlegi tudásunk szerint az alábbi mechanizmusok képesek szabályozni az R-hurkok mennyiségét a genomban: (1) RNázH enzimek (RNázH1 és RNázH2), amelyek szekvencia-független módon elhasítják az RNS-t az RNS-DNS hibridben (Lockhart et al. 2019; Skourti-Stathaki and Proudfoot 2014). (2) RNS-DNS hibrid helikázok, mint például a Senataxin, Aquarius, Pif1, Srs1/BLM (Yun et al. 2017), UPF1, DHX9, DDX1, DDX19, DDX21, DHX30, amelyek részt vesznek az R-hurkok feloldásában, ezáltal pedig az R-hurok-mediált DNS károsodás és

genomi instabilitás megelőzésében (Mischo et al. 2011). A Senataxin ezenkívül transzkripció terminációs faktorként működik a kódoló és nem-kódoló géneknél egyaránt, továbbá részt vesz a transzkripció és replikáció koordinációjában a replikációs villa integritásának fenntartásán keresztül (Alzu et al. 2012). (3) A topoizomerázok képesek megváltoztatni a DNS supercoiling-ot, amely egyébiránt folyamatos R-hurok keletkezéshez vezetne a transzkripció és replikáció során (Tuduri et al. 2009). A Topoizomeráz I (Top1) szerepéről tudunk többet, annak DNS relaxációs aktivitása hozzájárul a kotranszkripció R-hurkok eltávolításához, amely a transzkripció leállítását (“pausing”) és a DNS repliszómával történő “frontális” ütközést okozná (Manzo et al. 2018; Cristini et al. 2019). A transzkripció és replikáció közötti konfliktus a genom instabilitás egyik fő forrása, amely végsősoron R-hurok képződésre vezethető vissza. (4) RNS processzálo és biogenezis faktorok, például az ASF/SF2 (Li and Manley 2005) vagy a THO/TREX (TRanscription and EXport) komplex (Domínguez-Sánchez et al. 2011), amelyek az mRNS érési/szállítási folyamatokban vesznek részt. (5) Rekombinázok és anti-rekombinázok, amelyek elősegítik vagy gátolják az R-hurkok keletkezését. A homológ rekombináció során a száláthelyezést katalizáló Rad51 rekombináz (és bakteriális homológja, a RecA) például segíti az R-hurok képződését *in vitro* és *in vivo* (Kasahara et al. 2000; Wahba et al. 2013). Az antirekombináz Srs2 viszont a Rad51-el ellentétes hatást fejt ki, vagyis akadályozza az R-hurkok kialakulását (Wahba et al. 2013; Yun et al. 2017). Ezenkívül az Srs2/BLM és UPF1 helikázok (Yun et al. 2017; Ngo et al. 2021), a Rad52 (Yasuhara et al. 2018), BRCA1 (Cristiano et al. 2019; Bhatia et al. 2014; Renaudin et al. 2021), BRCA2 (Hatchi et al. 2015, 2021; Zhang et al. 2017), FANCM (Liang et al. 2019; Schwab et al. 2015) rekombinációs modulátor fehérjék szintén dokumentált szerepet játszanak az R-hurkok felismerésében és a duplaszál DNS törések (DSB-k) javításában.

Az R-hurok szerkezetek keletkezésének és turnover-ének szabályozása tehát szorosan összefügg a genom integritását befolyásoló tényezőkkel, mint például a mutagenézis és homológ rekombináció. Az R-hurkok rekombinogén hatására további bizonyíték, hogy elősegítik a száláthelyezés folyamatát a DSB hibajavítás során. A DSB-k processzállását végző MRX komplex (Mre11, Rad50, Xrn2/Nbs1) kölcsönhatása az RNS polimerázzal egy R-loop szintézisét eredményezi a visszavágott DNS-vég közvetlen közelében (Ohle et al. 2016; Yasuhara et al. 2018), amely szükséges és elégséges feltétele a nem-mutagén homológia-vezérelt hibajavításnak (HDD).

A HOMOLÓG REKOMBINÁCIÓ MECHANIZMUSA

Az evolúció során a különféle DNS sérülésekre különböző hibajavító útvonalak specializálódtak, amelyek felismerik/átalakítják a léziókat és helyreállítják a genom integritását. A sejtek túlélését

leginkább veszélyeztető folytonossághiányok a DNS duplaszál törések (DSB-k), amelyek szekvencia-homológia alapján történő (hibamentes) korrekcióját a homológ rekombináció (HR) útvonala végzi. A hibajavítás alapját képező nem-sérült templát mitotikusan osztódó sejtekben az S fázisban szintetizálódó testvérkromatidokról származik, míg meiotikus sejtekben (ősvarsejtekben) kizárólag az apai vagy anyai homológ kromoszómákról. Ez egy döntő különbség a mitotikus és meiotikus HR között, a jelenséget interhomológ torzításnak (“interhomologue bias”) nevezzük (Schwacha and Kleckner 1997). A másik döntő különbség a DNS törések keletkezési módjából és gyakoriságából adódik. Mitotikus DSB-k főleg a replikáció vagy transzkripció megakadása során, illetve a két rendszer összeütközésekor keletkeznek, amely viszonylag ritka esemény, a DSB-k eloszlása pedig sztochasztikus. A meiotikus sejtek viszont speciális DSB-, rekombinoszóma-, és szinaptonémális komplex fehérjéket expresszálnak, amelyek egyszerre nagyszámú törést generálnak térben és időben kontrollált módon (becsült számuk: 90-140 DSB / sejt; (Buhler et al. 2007), >1000-szeresére emelve a HR gyakoriságát (Székvölgyi et al. 2015; Székvölgyi and Nicolas 2010). A meiotikus DSB-k pozíciója konkrét “konszenzus” DNS szekvenciához nem köthető, ám közvetlen környezetükben dúsulást mutatnak bizonyos “aktiváló” hiszton modifikációk (pl. H3K4me3, H3K36me3, H3K9ac) illetve bizonyos hiszton módosító enzimek deléciója (pl. Gcn5, Rpd3, Sir2, Set1, Set2, Dot1, Rad6, Him-17) néhány hotspot régió esetén szignifikánsan csökkentette a DSB keletkezés gyakoriságát (Székvölgyi et al. 2015; Székvölgyi and Nicolas 2010). Ezek közül a H3K4 trimetiláció szerepe igazolódott vissza a teljes genomon élesztőtől az emberig, mivel mind a Set1 hiányos élesztő sejtek (Sollier et al. 2004; Borde et al. 2009), mind a Prdm9 (meiózis-specifikus H3K4 hiszton metiláz (Mihola et al. 2009; Berg et al. 2011) mutáns egerek, emberszabású majmok, és emberek súlyos meiotikus defektust mutattak.

TÉRBELI GENOMSZERKEZET ÉS HOMOLÓG REKOMBINÁCIÓ

A kromatin funkciója szorosan kapcsolódik a genom háromdimenziós szerkezetéhez, amely természetesen érinti a homológ rekombináció folyamatát is. Emlősökben néhány strukturális fehérjét már jellemeztek, melyek fontos szerepet játszanak a kromatin hurkok kialakulásában és mutációjuk gyakran kapcsolódik fejlődési rendellenességekhez és más betegségekhez. Ilyen fehérje a CCCTC-kötő faktor (CTCF) (Wit et al. 2015; Vietri Rudan et al. 2015; Guo et al. 2015; Nagy et al. 2016; Nichols and Corces 2015; Tang et al. 2015), a YY1 (Weintraub et al. 2017), és a kohezin komplex tagjai (Haarhuis et al. 2017; Fudenberg et al. 2016; Mizuguchi et al. 2014; Sofueva et al. 2013; Costantino et al. 2020; Rao et al. 2017; Lazar-Stefanita et al. 2017; Busslinger et al. 2017), amelyek

kötődését a helyi hiszton környezet, másodlagos DNS struktúrák és a DNS szekvencia szabályozza. Meiotikus sejtekben ezidáig a CTCF-paralóg BORIS szerepét igazolták (Rivero-Hinojosa et al. 2021), illetve azt is tudjuk, hogy a mitotikus kohézinek meiózis-specifikus kohézinekre cserélődnek (Rec8, Rad21L), amelyek a kromatin hurkok szervezéséhez szükségesek és ellenállnak a szeparáz hasításnak, amely a testvérkromatid kohézió fennmaradásához szükséges a meiózis I anafázisában (Watanabe 2012). A meiotikus kohézinek ún. tengelyfehérjékkel együtt (pl. Red1, Hop1, Zip1, Zip3) hozzák létre a homológ kromoszómák központi tengelyét ("axis"), amely az ezután kialakuló szinaptonémális komplex laterális elemeként funkcionálva proximitásba hozza az anyai és apai homológokat a rekombináció iniciációja során. A tengelyfehérjék citológiai és genomikai térképezése igazolta, hogy azok kolokalizálnak az esszenciális DSB-fehérjékkel (élesztőben: Spo11, Mer2, Mre11, Xrn2, Rad50, Ski8, Mei4, Rec102, Rec104) (Blat et al. 2002; Arora et al. 2004; Maleki et al. 2007; Panizza et al. 2011; Murakami et al. 2020; Karányi et al. 2018), vagyis a rekombináció biokémiai lépései (DNS törés, száláthelyezés, hibajavítás, rekombináns termékek keletkezése) a kromoszóma tengelyhez kötötten játszódnak le. A DSB hotspot-ok viszont meglepő módon nem ezekre a szakaszokra esnek, hanem a tengellyel "ellentétes" kromatin hurok régiókra (Blat et al. 2002). Ezt a szerkezeti paradoxont a Nancy Kleckner által javasolt "hurok-tengely" modell bevezetése oldotta fel ("tethered loop-axis", (Blat et al. 2002), amely szerint a kromatin hurkok (ahol a DSB forrópontok vannak) és a kromoszómák tengelye (ahol a töréseket okozó enzimek és rekombinációs fehérjék vannak) fizikai kölcsönhatásba kell lépjenek egymással. A feltételezett hurok-tengely kölcsönhatás molekuláris mechanizmusa egészen a közelmúltig ismeretlen maradt.

CÉLKITŰZÉSEK

1) Kutatómunkám során céлом volt a meiotikus homológ rekombináció iniciációjának molekuláris vizsgálata, a DSB képződés epigenetikai szabályozásának mélyebb megértése, a főbb molekuláris komponensek és fehérje-fehérje interakciós partnerek azonosítása. Kiemelt figyelmet fordítottam a Set1C/COMPASS hiszton metiláz komplex szerepére, mivel a hiszton H3K4 trimetiláció és a rekombinációs forrópontok statisztikai korrelációja már korábban ismert volt (Borde et al. 2009), azonban az ok-okozati kapcsolat bizonyítása nem sikerült. További céлом volt a “hurok-tengely komplex” hipotézis alátámasztása vagy elvetése a kapott adatok alapján. Céлом volt megvizsgálni más hiszton modifikációk szerepét is a fenti folyamatokban, amelyek nem az N-terminális hiszton részekre koncentrálnak, hanem a nukleosómák központi globuláris doménjére esnek (pl. H3K56 acetiláció).

2) Az R-hurok azonosítása és kromoszómális térképezése kulcsfontosságú a genom stabilitását szabályozó folyamatok megértéséhez és a repair faktorok és DNS léziók kromatinszerkezeti aspektusainak feltárásához. Céлом volt egy olyan univerzális genomikai térképezési módszer kidolgozása, amely az “alap” DRIP módszer inkonzisztenciáit kikerülve alkalmas az R-hurok nagy pontosságú azonosítására bármely modellorganizmusban. További céлом volt, hogy összefüggést találjak a kromatin hurok letapadási pontok, a DNS törések és az R-hurok sejtmagi lokalizációja között. A térbeli genomszerkezet vizsgálatok kapcsán kiemelt céлом volt a Hi-C módszer implementálása a laboratóriumomban. A fenti eredményekre és metodikákra alapozva célul tűztem ki új R-hurok regulátor funkcióval rendelkező fehérjék azonosítását és genomikai karakterizálását. A lúdfű (*Arabidopsis thaliana*) kicsiny genomja (~130 Mb) és genetikai manipulálhatósága miatt tökéletes alanya e komplex vizsgálatoknak, elsőként ebben a modellorganizmusban tanulmányoztam a *Nodulin homeobox* (NDX) transzkripciós faktor szerepét az R-hurok képződés, DNS metiláció, és a magasabbrendű 3D kromatinszerkezet összefüggésében. Humánspecifikus R-hurok szabályozó fehérjékkel pedig farmakogenomikai elemzéseket végeztem, hogy összefüggést találjak az R-hurok regulátorok expressziója és különféle rákos megbetegedésekben mutatott túlélés-asszociációk és kemoterápiás szerekekkel szembeni érzékenység/rezisztencia kialakulása között. Utóbbi vizsgálatok, az alapkutatói jelentőségükön túl, utat nyitottak az R-hurok és R-hurok szabályozó fehérjék klinikai kiaknázásához diagnosztikai markerként vagy terápiás célpontként.

ANYAGOK ÉS MÓDSZEREK

Élesztő mutánsok

A kísérletekben használt *Saccharomyces cerevisiae* törzsek az SK1 genomi háttérből származtak. Az összes géndeléciót az EUROFAN II géndeléciós könyvtárból származó PCR-amplifikált KanMX4 vagy HphMX vagy NatMX deléciós kazettákkal való transzformációval és keresztezéssel vittük be a szülői sejtekbe. A plazmid shuffle mutánsokat izogén *hht1Δ::HphMX hht2Δ::KanMX* pARS-CEN(*HHT2*, *HHF2-URA3*) és *hht1Δ::HphMX hht2Δ::KanMX dmc1Δ::LEU2* pARS-CEN(*HHT2*, *HHF2-URA3*) törzsek transzformálásával hoztunk létre, majd a haploid szülők keresztezésével és tetradelemzéssel homozigóta diploidokat készítettünk. A Set1C alegységek az UAS_{GAL} szekvenciához történő célzott irányításához használt konstrukciókat PCR-rel állítottuk elő, a géneket NcoI/Pst1 vagy BamHI/PstI fragmentumok formájában vittük be a pASIN plazmidba a Gal4 DNS-kötő domént (GBD) kódoló szekvenciától 3'-irányban. A GBD-Set1C fúziós fehérjéket a konstitutív pADH1 promóter szabályozása alatt fejeztük ki az endogén *trp1-1* lókuszbba integrálva. Minden GBD-törzsből delécióval inaktiváltuk a megfelelő vad típusú Set1C allélt, hogy elkerüljük a fehérje kompetíciót a GBD- és endogén allélok között az UAS_{GAL} szekvenciához való kötődésért. A co-IP kísérletekhez az Spp1 fehérjét 3xHA epitóppal fuzionáltattuk, amelyet az *SPP1* gén C-terminális részébe vittünk be a pFA6a-3HA-KanMX6 plazmidról. Az SPP1-3xHA fehérje expressziója a 2 mikronos pBD827 plazmidról történt. Az SPP1 PHD doménjének delécióját két PCR amplikon fuzionálásával végeztük, amelyek -569 pozíciótól (ATG-hez képest) +3-ig és +235-től +1062-ig terjedtek. Ezt a PCR terméket a Yiplac211 *URA3* integratív plazmidba klónoztuk, amelyet NruI-el linearizáltunk. A kromoszómába integrált plazmidot hordozó sejteket 5-fluoroorotsav (FOA) lemezekre oltottuk, hogy kisselektálhassuk az *spp1Δ4-234* mutáns allélt hordozó, *spp1ΔPHD*-nek elnevezett rekombinánsokat. Az Spp1 cink-ujj (CXXC) motívumának (C263GYC266) delécióját úgy végeztük el, hogy a GBD-SPP1-et hordozó integratív vektorból (pASIN-SPP1) 12 bp-t kitöröltünk a Quickchange site-directed mutagenesis PCR segítségével. A GBD-SPP1Δ789-801 mutáns a GBD-SPP1ΔCXXC nevet kapta. A pASIN-SPP1ΔCXXC plazmidot XbaI-vel linearizáltuk, majd az endogén *trp1-1* lókuszbba integráltuk.

Arabidopsis minták

Kísérleteinkben a következő *Arabidopsis thaliana* (lúdfű) ökotípusokat használtuk: Col-0, *ndx1-4*, flag-NDX (genotípus: flag-NDX / *ndx1-1*(FRI) / *flc-2*) és NDX-GFP (genotípus: NDX- GFP / *ndx1-1*(FRI) / *flc-2*). A magokat sterilizáltuk és agar-MS lemezekre vetettük, majd 4 °C-on, sötétben

tartottuk 2 napig. A lemezeket ezután 21 °C-os, hosszú nappalos (LD, 16 óra világos, 8 óra sötét) környezetbe helyeztük 10 napra. Minden méréshez 10 napos csíranövényeket használtunk.

Humán sejtkultúra

Az R-hurkokkal kapcsolatos kísérletekben használt Jurkat sejteket (humán lymphoblastoid leukémia) 10% (v/v) magzati borjúsérummal (FBS), 2 mM glutaminnal, penicillinnel és sztreptomocinnel kiegészített RPMI-1640 tápfolyadékban (Sigma, R5886) tenyésztettük 5% széndioxid koncentráció mellett.

Meiózis és sporuláció élesztőben

A szinkronizált sporuláció elindításához a diploid törzseket gazdag tápközegben (YPD) növesztettük 24 órán át, majd pre-sporulációs tápoldatba (SPS) oltottuk át őket át és éjszakán át növesztettük $\sim 4 \times 10^7$ sejt/ml sűrűségig. Az SPS tenyészeteket centrifugálással összegyűjtöttük, majd 1% kálium-acetáttal mostuk és előmelegített sporulációs tápoldatban szuszpendáltuk 2×10^7 sejt/ml sűrűségig (SPM; nem-fermentálható szénforrásként 1% kálium-acetátot tartalmaz, aminosavakkal és nukleotidokkal kiegészítve az auxotróf markereknek megfelelően). A meiotikus DNS-replikáció és szinkronitás követése áramlási citometriás méréssel történt. A fertilitást (spóra életképesség) tetrad elemzéssel határoztuk meg az életképes spórák arányából.

Southern blot és DSB analízis

A meiotikus DSB-k kimutatása és kvantifikálása *sae2Δ* mutánsokban történt, amely megállítja a DNS törések processzállását. A genomi DNS-t fenol-kloroformmal extraháltuk, majd megfelelő restrikciós enzimekkel emésztettük és Southern blot-tal analizáltuk. A PCR-által generált próbákat ^{32}P -dCTP-vel jelöltük Ready Prime kit (GE Healthcare) segítségével, és a radioaktív jeleket Storm Phosphorimager (Molecular Dynamics) segítségével detektáltuk. A DSB frekvenciákat az ImageQuant szoftverrel (Molecular Dynamics) határoztuk meg.

Élesztő 2-hibrid screen

Az *Saccharomyces cerevisiae* teljes-genom könyvtár ("ORFome") szűrését yeast 2-hibrid screening rendszerrel végeztük a Hybrigenics SA (Párizs) cég segítségével. Csaliként pB27 vektorba klónozott *SPP1* ORF-et használtunk (lexA, C-terminális fúzió).

Rekombináns fehérje expresszió és GST-pull down

A GST-t és a GST-Mer2 fúziós fehérjét a bakteriális lizátumokból Glutathione Sepharose 4B-n tisztítottuk az ajánlásoknak megfelelően (GE Healthcare). A teljes hosszúságú Spp1 fehérje és trunkált Spp1-fragmentumok *in vitro* expressziójához a plazmid DNS-eket TNT-kapcsolt *in vitro* transzlációs rendszerrel (Promega) írtuk át mRNS-é és fehérjévé. A tisztított GST-t vagy a GST-Mer2 fúziós fehérjét az *in vitro* transzlált [³⁵S]-metionin-jelölt fehérjékkel inkubáltunk. A kötőpuffer 1 ml-es aliquotját 15 µl glutation-sepharose-val (GE Healthcare) inkubáltuk, majd a gyantát ülepítettük és háromszor mostuk 1 ml 0,01% NP-40-et tartalmazó PBS-sel. A fehérjéket 40 µl NuPAGE minta pufferrel eluáltuk és 4-12%-os Bis-Tris NuPAGE gélen (Invitrogen) megfuttattuk. A géleket 30 percig fixáló oldatban (10% ecetsav, 40% metanol) fixáltuk, majd 30 percig 1 M nátrium-szalicilát oldattal kezeltük és vákuumban 3M whatman papíron szárítottuk. A szárított géleket Fujifilm filmen 24-72 órán keresztül exponáltuk és BAS-1800II képelemző készüléken szkenneltük.

Kromatin immunprecipitáció és kompetíciós kromatin immunprecipitáció (ChIP és c-ChIP) szekvenálás

A meiotikus élesztősejteket (4×10^7 sejt/ml) 50 ml térfogatban gyűjtöttük össze az ábrákon jelzett időpontokban és 1% formaldehiddel 20 percig szobahőmérsékleten fixáltuk. A reakciót 125 mM glicinnel, szobahőmérsékleten történő inkubálással állítottuk le 5 perc alatt, majd a sejteket háromszor jéghideg 1x TBS, pH 7,5 (20 mM Tris-HCl, pH 7,5, 150 mM NaCl) oldatban mostuk. *Arabidopsis* csíranövények (GFP-NDX) esetén magát a növényi mintát fixáltuk a fenti módon, majd folyékony nitrogén-tartalmú dörzsmozsárban mechanikusan homogenizáltuk a sejteket. Ezt követően a sejteket 500 µl lízis pufferben (50 mM Hepes, KOH pH 7,5, 140 mM NaCl, 1 mM EDTA, 1 % Triton X-100, 0,1 % Na-deoxikolát, 1 tabletta teljes protein inhibitor koktél (Roche)), savval mosott üvegyöngyökkel lizáltuk 10 percig FastPrep bead beater készülék segítségével. A kromatin mintákat ultrahanggal (Bioruptor, Diagenode) fragmentáltuk átlagosan 300 bp méretűre. A teljes sejtextraktumból kivettünk 50 µl immunprecipitáció előtti mintát (input), a maradék mintát 20 percig centrifugáltuk (12000 fordulat/perc, 4 °C) a sejtörmelék elválasztásához. Az immunoprecipitációt (IP) úgy hajtottuk végre, hogy 450 µl extraktumot adtunk a mágneses protein G dynabeads (Dyna) pellethez, amelyből 50 µl 2×10^7 gyöngynek felelt meg. A gyöngyöket előinkubáltuk a megfelelő antitestekkel, amely "hagyományos" ChIP esetén 9E11 monoklonális egér anti-myc antitest volt (ab56, Abcam), c-ChIP esetén pedig 9E11 anti-myc antitest (ab56, Abcam) a konstitutív allél ellen és anti-GFP poliklonális nyúl antitest (ab290, Abcam) a kompetíciós (indukált) allél ellen, éjszakán át 4 °C-on. Egyéb yeast ChIP kísérletekben anti-GAL4BD (Euromedex) és anti-

H3K4me3 (Abcam) antitesteket is használtunk. Az *Arabidopsis* CHIP mérésekhez anti-flag (#2044, New England Biolabs) és anti-GFP (#ab290 Abcam) antitesteket használtunk. Az IP mintákat kétszer mostuk lízis pufferrel, majd újabb kétszer 360 mM NaCl-dal kiegészített lízispufferrel, kétszer mosó pufferrel, és végül egyszer 1x TE-vel, pH 7,5. A keresztkötéseket TE/1% SDS oldatban oldottuk fel, éjszakán át, 65 °C-on. Ezután a fehérjéket proteináz K-val emésztettük 3 órán át 65 °C-on. A nukleinsavat PCR clean-up oszlopon tisztítottuk, majd RNS-emésztést (10 ug RNáz) végeztünk 1 órán át 37 °C-on. A CHIP DNS-t végül 50 µl nukleáz-mentes dH2O-ban oldottuk fel. Az NGS könyvtárakat az Illumina TruSeq CHIP mintaelőkészítési protokollja szerint állítottuk elő. A könyvtárakat 150 bázispáros paired-end read-ekre szekvenáltuk Illumina HiSeq 2500 szekvenátorral (EMBL Genomics Core Facility, Heidelberg, Németország). A nyers fastq file-ok feldolgozása és a peak meghatározás a DRIP-seq módszerhez hasonlóan történt.

Kromatin turnover ráta becslése c-CHIP adatokból

Az Spp1 és Set1 genomi kötőhelyek turnover rátáját a GFP-Spp1 és a Spp1-myc kompetíciós ChIP-seq adatok alapján határoztuk meg. Kiszámítottuk a GFP-Spp1 / MYC-Spp1 "okkupancia" arányokat minden egyes időpontban és minden genomi kötőhelyre (peak-re), majd exponenciális modellt illesztettünk az adatokra: $\frac{GFP}{myc} = (1 - e^{-\lambda t})$, ahol a *GFP/myc* az "okkupancia" arány, λ a turnover ráta, t pedig a kompetíciós allél indukciójától számított idő. A modell illesztése után kiszámítottuk a becslések standard hibáját és t-próbát végeztünk a modell illeszkedésének jóságának becslésére. Összesen 977 c-CHIP kötőhelyet és turnover rátát írtunk le a modellel.

Meiotikus transzkriptom microarray

20 ml sporuláló (SPM) élesztő tenyészetekből Trizol reagenssel totál RNS-t izoláltunk a sporuláció kezdőpontjában (0h) és az azt követő időpontokban (2h, 4h, 6h). A cDNS szintézis során aminoallyl-dUTP-t (Fermentas) építettünk be a mintákba. A 0h mintát referenciának tekintettük és Cy3 festékkel jelöltük (Amersham), míg a 2-6h mintákat Cy5 festékhez kapcsoltuk. A Cy3-mal jelölt referencia cDNS-t azonos mennyiségben hibridizáltuk a Cy5-jelzett mintákkal 8x15K transzkriptom microarray-khez (Agilent). A mosásokat, a szkennelést, az adatok normalizálását és az elemzést a gyártó protokollja szerint végeztük. Az eredmény file a 0h időponthoz képest mért relatív mRNS expressziós változásokat mutatja.

RNS-DNS hibrid immunprecipitáció (DRIP)

Humán sejtek keresztkötéséhez 1%-os paraformaldehid oldatot használtunk 10 percig, aminek a hatását 2.5 M glicinnel szüntettük meg 5 perc alatt szobahőmérsékleten. *Arabidopsis* csíranövények (Col-0 és *ndx1-4* minták) esetén magát a növényi mintát fixáltuk a fenti módon, majd folyékony nitrogén-tartalmú dörzsmozsárban mechanikusan homogenizáltuk a mintát. Az előkészített preparátumokat 1 ml lízis pufferben (500 µl 2x lízis és 500 µl TE) lizáltuk. Humán minták esetén a sejtlízist két különböző hőmérsékleten is elvégeztük: 65 °C 7 óra vagy 37 °C, éjszakán át. A nukleinsav izolálást NucleoSpin Tissue Kit segítségével végeztük és 100 µl elúciós pufferben eluáltuk a mintákat. A tisztított nukleinsav preparátumot 300 µl Tris-HCl (pH 8.5) oldatban fragmentáltuk szonikálással kétszer 5 percig, hogy átlagosan 500 bázispár méretű DNS fragmenteket kapjunk. A fragmentum analízist 1%-os agaróz gélelektroforézissel végeztük el. Amennyiben szükséges volt, további szonikálást alkalmaztunk. A szonikált DNS mintákat NucleoSpin Gél és PCR tisztító kittel tisztítottuk, majd 100 µl elúciós pufferben eluáltuk. 12 µg DNS-t 100 µl térfogatra hígítottunk 5 mM Tris-HCl (pH 8.5) oldattal. A minták 2%-át input DNS-ként eltettük. A minták felét 80 µl végtérfogatban 8 µl *E. coli* RNase H-val kezeltük (NEB) 37 °C-on, éjszakán át. Az immunszelekcióhoz Dynabeads Protein A mágneses gyöngyöket PBS/EDTA tartalmú 0.5%-os BSA-val blokkoltunk. Az S9.6 antitest immobilizációhoz 50 µl blokkolt gyöngyöt 10 µg S9.6 antitesttel inkubáltunk folyamatos forgatás közben IP pufferben 4 °C-on 4 órán át. 6 µg fragmentált genomi DNS-t adtunk a keverékhez és gyengéden forgattuk egy éjszakán át 4 °C-on. A gyöngyöket ülepítettük és mostuk egymás után 1 ml lízis pufferrel (alacsony só cc. mellett), 1 ml lízis pufferrel (magas só cc. mellett), 1 ml mosó pufferrel és 1 ml TE-vel 4 °C-on, kétszer. Az elúció 100 µl elúciós pufferrel történt 15 percen keresztül 65 °C-on. NucleoSpin Gél tisztítás és PCR Clean-up kit után, a nukleinsavakat 55 µl elúciós pufferben eluáltuk. A kinyert DNS-t kvantitatív valós-idejű PCR készülékkel (LightCycler 480, SYBR Green I Master) mértük le és QuantStudio 12K Flex valós-idejű PCR rendszerrel elemeztük ki. Az adatokat a relatív dCT módszerrel analizáltuk. Az RNS-DNS hibrid feldúsulást az IP/Input arányból számoltuk ki. Az NGS könyvtárakat az Illumina TruSeq ChIP protokoll alapján készítettük el. A leolvasott szekvenciákat az emberi (hg19) vagy *Arabidopsis thaliana* (tair10) referencia genomra illesztettük BWA-MEM algoritmussal alapbeállítások mellett. Kiszűrtük azokat a leolvasott szekvenciákat, amelyek alacsony térképeződési értékkel rendelkeznek, illetve PCR duplikátumok voltak vagy feketelistás régiókon lokalizálódtak. A replikákat összevontuk és MACS2-t használtunk a szignifikáns genomi régiók azonosításához input-normalizálás mellett. A feldolgozott és összevont illesztésekből bamCoverage alkalmazásával szignál fájlokat generáltunk. Az RPKM értékeket 20 bázispáros

ablakokban határoztuk meg minden egyes mintára egy 60 bázispáros csúszóablak alkalmazásával. Az elkészített szignál fájlokat R-ben ábrázoltuk ggplot2 és ggbio csomagok segítségével.

A DRIP és CHIP peak-ek genomi annotációja

A CHIP-seq és DRIP-seq adatok elemzése során az NGS szekvenálási adatok feldolgozásából kapott .bed fájlokkal dolgoztunk. A DRIP és CHIP peak-ek genomi eloszlásának a meghatározásához a GenomicRanges R csomagot alkalmaztuk, amellyel a DRIP/CHIP régiók és az annotációs kategóriák átfedő területeit meghatároztuk. Az átfedő területeket számítógép-generált véletlen átfedésekhez hasonlítottuk. A Bedtools szoftver segítségével a fájlokban tárolt kötőhely pozíciókat véletlenszerűen megváltoztattuk a kromoszómákon belül, ezzel az eredeti fájlokkal megegyező peak számot és méretet, de eltérő pozíciókat kaptunk. A szignifikáns dúsulást mutató helyeket a Jbrowse genomböngésző programmal vizualizáltuk.

KisRNS szekvenálás (sRNS-seq)

Arabidopsis Col-0 és *ndx1-4* palántákból totál RNS-t tisztítottuk a northern blot-nál leírt módszerrel. Az RNS-minták minőségét Agilent bioanalizátorral ellenőriztük (RIN > 9). Az NGS könyvtárakat az Illumina NEBNext® Multiplex Small RNA Library Prep protokollja szerint készítettük és Illumina NextSeq500 szekvenátorral szekvenáltuk 1x50 bp leolvasással. Az eredményeket az sRNAnalyzer pipeline segítségével elemeztük (Wu et al. 2017) az alábbiak szerint: az Illumina adaptereket Cutadapt segítségével levágtuk, majd a leolvasások méretét a 19-25 nt tartományba korlátoztuk. Az sRNAnalyzer segítségével a kiválasztott méretű read-eket a miRbase adatbázishoz (<https://www.mirbase.org/>) és egy nemrégiben közzétett átfogó kis RNS-lókus-adatbázishoz (Hardcastle et al. 2018) annotáltuk. A Col-0 és *ndx1-4* minták közötti differenciális sRNS expressziót a Deseq2 algoritmussal határoztuk meg ($p < 0,05$, $\text{abs}(\log_2(\text{fc})) > \log_2(1,5)$) és az eredményeket R-ben ábrázoltuk. Az sRNS-seq adatokat sRNS rt-qPCR-rel validáltuk.

In situ Hi-C

A Hi-C-t 10 napos *Arabidopsis thaliana* csíranövényeken végeztük Col-0 és *ndx1-4* mutáns genomi háttérben. A növényi mintát 1%-os formaldehiddel fixáltuk, majd folyékony nitrogén-tartalmú dörzsmozsárban mechanikusan homogenizáltuk. A sejtmag izolálásához Miracloth szűrőt használtunk az Arima Genomics protokollja alapján. A Hi-C kísérlet kivitelezéséhez Arima-HiC Kit-et használtunk (Arima Genomics Ltd.). A kromatin feldarabolását DpnII és Hinfl restrikciós enzimekkel végeztük, a végreparálást és biotininilálást a kit reagensével végeztük a gyártó ajánlása szerint. A keresztkötött/biotinilált fragmentumokat a permeabilizált sejtmagokon belül (*in situ*) ligáltuk, majd

szonikálás után NGS könyvtárat készítettünk az Illumina TrueSeq protokoll alapján. A paired-end NGS szekvenálás során mintánként közel 200 millió read-et kaptunk, amely a jelenlegi legjobb felbontás az *Arabidopsis thaliana* 130 megabázis méretű genomjáról. A leolvasásokat ún. "kiméra read"-ként kaptuk vissza, mivel a read-ek eltérő genomi szakaszokról származtak. Az interakciós fragmentumokat a TAIR10 referencia genomra térképeztünk a Juicer pipeline-al (Durand et al. 2016). A Juicer először a nyers .fastq adatokat transzformálta Hi-C interakciós adatokká, majd ezeket normalizálta és interakciós mátrixokká alakította és annotálta. A Juicer ezenkívül annotálta a DNS hurkokat, loop anchor motívumokat, és kontakt doméneket, melyeket a HiCCUPS algoritmus (Rao et al. 2014) segítségével azonosított. A .hic fájlokból származó térbeli interakciókat .bed fájlokba konvertáltuk 1 kilobázis (kB), 5kb, 10kb és 25kb nagyságú genomi felbontásokkal.

In silico restriktív enzim emésztés

Adott restriktív enzim-kombinációk ("RE koktélok") által generált elméleti fragmenthossz eloszlás kiszámításához az emberi, *Arabidopsis*, és élesztő genomot *in silico* feldaraboltuk a DECIPHER R csomag segítségével. A hasítóhely pozíciók alapján kiszámoltuk a restriktív fragmentumok hosszát és eloszlását. A keletkezett fragmenthossz-eloszlások statisztikai összehasonlításához 300 véletlenszerűen kiválasztott értéket hasonlítottunk össze Wilcoxon Rank Sum teszttel, 100 alkalommal. A p-értékeket Benjamini & Hochberg módszerrel korrigáltuk.

Konfokális lézerpasztázó mikroszkópia (CLSM)

A humán sejteket minden mérés előtt 1%-os formaldehiddel fixáltuk. A CLSM képeket Olympus FluoView 1000 konfokális mikroszkóppal vettük fel, amelyhez 60x olajimmerziós objektívet (NA 1,35) használtunk. A gerjesztési és emissziós szűrők a következők voltak: EGFP, 488 nm gerjesztés, 500-540 nm detektálás. Tíz, 0,7-1,1 µm vastagságú optikai szeletet gyűjtöttünk minden egyes mag esetében, Kalman szűrőmodot alkalmazva a zaj és alternatív gerjesztés csökkentése érdekében és a cross-talk kizárására. A kolokalizáció mértékét az ImageJ program JACoP plugin-jével számoltuk ki.

Fluoreszcencia-visszatérés fotoelhalványodás után (FRAP)

A FRAP méréseket GFP-Spp1 és GFP-Set1 élesztő sejteken végeztük Olympus FluoView 1000 IX-81 invertált konfokális mikroszkóppal, UPlanAPO 60x (NA 1.2) vízimmerziós objektívvel. Az GFP-t 488 nm-en gerjesztettük argon-ion lézerrel és a fluoreszcenciát egy 500-550 nm-es sávszűrőn detektáltuk. A FRAP mérésekben kioltási régióknak négyzet alakú területeket választottunk a magon belül. A sporuláló élesztősejtekben elvégzett FRAP kísérletekhez (0-6 óra SPM-ben) óránként vettünk mintákat és a méréseket 1% kálium-acetáttal bevont mikroszkópos tárgylemezeken végeztük el. A

Set1-GFP vagy Spp1-GFP fehérjéket expresszáló sejteket véletlenszerűen választottuk ki 100 μM CuSO_4 indukciót követően, majd öt előképet (256 \times 256 képpontos terület, 15x zoom, $\sim 9 \mu\text{W}$ lézer teljesítmény) készítettünk és 500 ms-os kioltási idő mellett, 100 %-os lézer teljesítménnyel (900 μW) végeztük el a sejtmag fotokioltását. A képeket másodpercenként gyűjtöttük 1 percen át. Az *Arabidopsis thaliana*-ban végzett FRAP mérések esetén az NDX-GFP jelet 10 napos csíranövények gyökércsúcsában mértük a fenti paraméterekkel azonos beállításokkal, 25 másodpercig.

Fluoreszcencia-korrelációs spektroszkópia (FCS)

A GFP-Set1 és GFP-Spp1 FCS kísérletekhez Olympus FluoView1000 konfokális mikroszkópot használtunk. Meiotikus élesztő sejteket mintavételeztünk minden órában folyékony sporulációs tenyészetekből, majd a sejteket 1% kálium-acetát agarral (SPM) bevont mikroszkópos tárgylemezre cseppentettük. Mérés előtt a GFP-Set1 és GFP-Spp1 expressziót 100 μM CuSO_4 oldattal indukáltuk. Az autokorrelációs görbéket ALV-5000E korrelációs kártyával számoltuk ki az egyes sejtmagokban, három véletlenszerűen kiválasztott pontban, 10 \times 8 másodperces mérést követően. Gazdag táptalajon (YPD) növesztett mitotikus tenyészetek esetében a sejteket $A_{260} = 1,0$ optikai denzitás értékig növesztettük, majd YPD agarral bevont tárgylemezen, 100 μM CuSO_4 indukciót követően mértük. Minden mérés szobahőmérsékleten (22 $^\circ\text{C}$) történt. Az autokorrelációs függvényeket és a származtatott FCS paramétereket a QuickFit 3.0 szoftverrel számítottuk ki. Az *Arabidopsis thaliana*-ban végzett FCS mérések esetén az NDX-GFP jelet 10 napos csíranövények gyökércsúcsában mértük a fenti paraméterekkel azonos beállításokkal.

Farmakogenomikai elemzések

Összesen 36 R-hurok regulátor fehérjét vizsgáltunk 33 primer tumor típusban, amelyek génexpressziós és túlélés adatait a TCGA adatbázisból töltöttük le (The Cancer Genome Atlas Research Network et al. 2013). A túléléselemzés Kaplan-Meier görbéit a "survival" szoftverrel generáltuk R-ben. Az R-hurok regulátorok génexpresszió-függő gyógyszer-szenzitivitás asszociációit 276 FDA által engedélyezett szerre számítottuk ki a GDSC adatbázisból letölthető tumorsejtvonalak adatai alapján (Yang et al. 2013). Az asszociációkat az IC_{50} értékek (félhalálos dózis) és az adott R-hurok regulátor gén medián mRNS-seq értékek Spearman rangkorrelációja alapján ábrázoltunk. A statisztikai számításoknál a p-értékeket a Benjamini-Hochberg módszerrel korrigáltuk.

EREDMÉNYEK ÉS MEGBESZÉLÉSÜK

A Set1 komplex szerepe a meiotikus DNS törések kialakulásában

A H3 lizin 4 trimetiláció és a Set1 komplex (Set1C) “megjelöli” a rekombinációs forrópontokat élesztőtől az emlősökig (Sollier et al. 2004; Borde et al. 2009; Parvanov et al. 2010), azonban azt nem ismertük, hogy ez a rendkívül abundáns hisztonmódosítás hogyan lép kapcsolatba a DSB hotspot régiókkal és a DNS duplaszáltöréseket (DSB-eket) okozó Spo11 komplexxel. Ennek tisztázására olyan Set1C-mutánsokkal végeztünk funkcionális vizsgálatokat, amelyekben egyesével inaktiváltuk a komplex összes alegységét és megvizsgáltuk, hogy bizonyos jól karakterizált rekombinációs forrópontokban megváltozik-e a DNS törések gyakorisága és eloszlása. Southern blot kísérleteink szerint minden Set1C-mutánsban (Set1, Sdc1, Bre2, Swd1, Swd3, Spp1) szignifikánsan csökkent a meiotikus DSB-k mennyisége (az Shg1 alegység kivételével), amely a módosítható lizin 4 oldallánctól és így a H3K4 metilációtól függött. A H3K4 trimetiláció és a DNS törések mechanikai kapcsoltságának bizonyításához egy olyan kísérleti rendszert terveztünk, amelyben a kromoszómák előre meghatározott pontjain (célzottan) elhelyeztünk hiszton H3K4me3 módosításokat, majd ugyanitt detektáltuk a DNS töréseket. Ez a funkcionális megközelítés a CRISPR korszak előtt kis túlzással “forradalminak” számított, amellyel azt kívántuk bizonyítani, hogy amikor elhelyezzük a kromoszómán a hiszton-jelet, akkor valóban a várt pozícióban törik el a DNS és nem máshol, ok-okozati kapcsolatot igazolva a két változó között a korábbi statisztikai összefüggés helyett. Ehhez egy olyan géntechnológiai eszközt fejlesztettünk, amely egyrészt felelős a vizsgált H3K4me3 kialakulásáért, másrészt szelektíven odairányítható a DNS molekula előre megadott szakaszaihoz egy Gal4 DNS kötő domént (GBD) tartalmazó „célzó” egység segítségével. A vizsgált genomi régiók jól karakterizált “hideg” (coldspot) szakaszok voltak, amelyek normál körülmények között nem vesznek részt rekombinációs eseményekben, viszont *UAS_{GAL4}* motívumokat tartalmaztak a konstrukt megkötéséhez. A GBD-Set1 és GBD-Spp1 törzsek minden tekintetben úgy viselkedtek, mint a vad típusú sejtek, kivéve, hogy a *GAL2* coldspot régióban, ahol vad típusú sejtekben egyáltalán nincs rekombináció, a fúziós fehérjék DNS töréseket váltottak ki, így bizonyítva az ok-okozati összefüggést a H3K4me3 jel és a rekombináció iniciációja és között. A GBD-targeting kísérletek váratlan és elsőre megmagyarázhatatlan eredménye volt, hogy amíg a Set1 katalitikus egység által kiváltott meiotikus DSB-k képződése a módosítható H3K4 oldallánctól és az Spp1 jelenlététől függött, addig az Spp1 hatása független volt a H3K4 oldallánc és a Set1 jelenlététől. A fenti eredmények legegyszerűbb magyarázata, hogy az Spp1 valamilyen módon fizikai kontaktusba lép a meiotikus DSB-komplex egyik (vagy több) tagjával, esetleg magával a Spo11-el, amelyet odatoboroz a rekombinációs coldspot

régiókhoz, ahol emiatt DNS törések keletkeznek. Hipotézisünket élesztő 2-hibrid (Y2H) screen kísérlettel igazoltuk *in vitro*, amelyben a teljes (~6000) élesztő fehérjekódoló génekre vonatkozóan ("ORFome") elvégeztünk egy szűrést oly módon, hogy csaliként az Spp1 LexA-fúziós fehérjét használtuk. A screen eredményeként a Mer2 rekombinoszóma fehérjét azonosítottuk az Spp1 fehérje-fehérje interakciós partnereként, amely a Spo11 közvetlen partnere és aktivátora (Arora et al. 2004; Zhang et al. 2020; Rousova et al. 2021) és a kromoszóma tengelyre lokalizálódik (Panizza et al. 2011). A kölcsönhatást glutation S-transzferáz / Mer2 pull-down kísérletekkel validáltuk különböző hosszúságú (trunkált) *in vitro* transzlált Spp1 polipeptidek felhasználásával, amelyben a Mer2-vel való kölcsönhatáshoz szükséges domént az Spp1 131 aminosavból álló C-terminális régiójára térképeztük. A két fehérje fizikai kölcsönhatását *in vivo* co-IP és ChIP-qPCR kísérletekben is megerősítettük.

Az Spp1 fehérje Mer2-interakciós doménjének közvetlen szerepét további funkcionális kísérletekben igazoltuk. Ehhez az Spp1-ből kitöröltük a fehérje korábban leírt cinkujjszerű (CXXC) doménjét (Murton et al. 2010), amely egybeesett az általunk felfedezett Mer2-interakciós doménnel, majd megvizsgáltuk, hogy a deléción befolyásolja-e a DSB képződést és a Mer2 kötődését a *GAL10* coldspot régióban. A GBD-*Spp1_{CXXCA}* mutánsban mind a DSB-képződés mind a Mer2-kötés szignifikánsan csökkent a GBD-Spp1-hez képest (**21/A-B ábra**), amely azt jelzi, hogy az Spp1-Mer2 interakciós motívum szükséges a Mer2 kromatin kötéséhez és a DSB keletkezéshez.

A hiszton H3K4 trimetilációnak van-e közvetlen mechanikai szerepe az Spp1-Mer2 interakció létrejöttében és a meiotikus DSB-k keletkezésében? Ez egy kulcsfontosságú kérdés, amely megválaszolásához az Spp1 fehérjéből kitöröltük a H3K4me3 jel kiolvasásáért felelős PHD domént (*spp1_{PHDA}* mutáns) és megvizsgáltuk a DSB-keletkezés gyakoriságát. Az *spp1_{PHDA}* szignifikánsan csökkentette a meiotikus DSB-k számát a vizsgált szakaszokon. Mindebből az következik, hogy az Spp1 H3K4me3-olvasó PHD doménje specifikus szerepet játszik a meiotikus DSB-k kialakulásában. A fenti eredményeket összegezve, a meiotikus homológ rekombináció iniciációjához szükség van H3K4me3 modifikációra, amelyet a Set1C/COMPASS katalizál, szükség van Mer2 DSB-fehérjére, amely a Spo11-komplexet rögzíti a kromoszóma tengelyhez (Panizza et al. 2011), és szükség van Spp1-re, amely a PHD-finger és Mer2-ID doménjén keresztül fizikai kapcsolatot teremt az epigenetikai jel és a DNS törések között. Kísérleteink alapján egy új rekombinációs modellt vezettünk be, amely egyrészt kiterjesztette a régebbi hurok-tengely modellt (Blat et al. 2002), másrészt az Spp1-

Mer2 kölcsönhatás alapján molekuláris szinten megmagyarázta a meiotikus DNS-törések kapcsolatát a Set1C-katalizált H3K4me3 módosítással.

Az Spp1 kromatinkötődési dinamikájának vizsgálata a meiózis során

A fenti hurok-tengely modell alapján szeretnénk volna részleteiben megérteni az Spp1 speciális DSB-segítő funkcióját a meiózis során. Például azt, hogy az Spp1 a Set1C/COMPASS hisztonmetiláz komplex részeként fejt-e ki rekombináció-potencírozó hatását, vagy esetleg a Set1C-ről leválva, attól függetlenül működik. A kérdés megválaszolásához 9xmyc epitóppal jelölt Spp1 és Bre2 fehérjéket expresszáltunk meiotikusan differenciálódó élesztő sejtekben (Spp1-9xmyc, Bre2-9xmyc) majd ChIP-szekvenálással meghatároztuk a genomi eloszlásukat (cisztrum) a meiózis profázisa során (0-6h SPM). A Bre2 fehérje a Set1 komplex egyik stabil alegysége, ezért genomi eloszlása markerként használható a Set1C pozíciójának jelölésére. Az azonosított kromatinkötő helyek átfedésének az elemzése azt mutatta, hogy az Spp1 csúcsok mintegy 46%-a egybeesik a Bre2/Set1C kötőhelyekkel a meiózis alatt, míg 54%-uk attól független genomi eloszlást mutat. A különbség és metszet halmazok annotációja szerint a Set1C-független Spp1 frakció ("Spp1-only" kötőhelyek; n=2447) jelentős dúsulást mutattak a Mer2 és Red1 kromoszómatengely kötőhelyek felett, míg ugyanitt a "Bre2-only" (n=283) és a közös Spp1/Bre2 csúcsok ("common", n=2077) nem dúsultak. Fontos megjegyezni, hogy az Spp1 kromatinkötése progresszív növekedést mutatott a Mer2 kötőhelyeken a meiózis profázisa alatt, míg ugyanitt a Bre2/Set1C peak-ek egyáltalán nem mutattak dúsulást. Eszerint, a meiotikus differenciáció során az Spp1 molekulák mintegy fele leválik a Set1 komplexről és "átkerül" a Mer2 kötőhelyekre, s mindez a Set1C/COMPASS alapfunkciójától függetlenül történik.

Az Spp1 kromatin dinamikájának mélyebb megértéséhez időfüggő klaszteranalízist végeztünk az azonosított ChIP-seq kötőhelyeken, így az Spp1 csúcsokat osztályokba tudtuk sorolni a hasonlóságuk alapján. Két fő kinetikus csoportot különítettünk el az Spp1 kötőhelyekre jellemző ChIP-seq értékek időbeli változása alapján: 1. dinamikus Spp1 kötőhelyek, amelyek a meiózis előrehaladtával fokozatosan megjelentek (ún. megjelenő alcsoport) vagy eltűntek (ún. eltűnő alcsoport), 2. statikus helyek, amelyek időben állandó asszociációt mutattak az Spp1-el. A dinamikus és statikus kötőhelyek funkcionális annotációja alapján i) a megjelenő Spp1 csúcsok erősen dúsultak a kromoszóma tengely mentén (a Mer2 és Red2 kötőhelyeken); ii) az eltűnő Spp1 kötőhelyek ugyanitt depléciót mutattak, ellenben a riboszómális fehérje (RPG) és snoRNS géneken felülreprezentáltak voltak; iii) a statikus Spp1 peak-ek erősen kötődtek a ncRNS génekhez.

A továbbiakban funkcionális ChIP-seq elemzést végeztünk különféle Spp1 funkcióvesztéses mutánsokban, amelyekben az Spp1 nem tud kötődni a H3K4me3 módosításhoz (*Spp1^{PHDΔ}* mutáns), vagy nem kapcsolódik a Mer2 fehérjéhez (*Spp1^{CXXCΔ}* mutáns), vagy nincs jelen a H3K4me3 modifikáció (H3K4R és H3R2A hiszton mutánsok), amelyet az Spp1 a Mer2-höz horhonyozhatna a PHD doménjén keresztül. Mindhárom esetben a feltételezett hurok-tengely kölcsönhatás megszűnését és a meiotikus DSB-k szintjének jelentős csökkenését várhatjuk. A vad típusú sejtekben azonosított kötőhelyekhez viszonyítva a mutánsokban detektált Spp1 kötőhelyek kb. 50%-a eltűnt, vagyis az Spp1 fehérje PHD doménje, CXXC doménje, illetve a H3K4me3 módosítás szükséges az Spp1 stabil kromatinkötéséhez. Mindebből arra következtettünk, hogy az Spp1 H3K4me3-olvasó PHD doménje, a Mer2-interakciós CXXC doménje, illetve a H3K4me3 módosítás jelenléte történő kromatin-kölcsönhatáshoz. Összehasonlítóképpen kielemeztük szükséges a Mer2-vel a Bre2 ChIP-seq kötőhelyek Mer2-vel való asszociációját, amely a várakozásnak megfelelően egyáltalán nem mutatott Mer2-dúsulást.

A továbbiakban az Spp1 kromatin-dinamikáját kvantitatívan jellemeztük különféle genomikai és mikroszkópos módszerek segítségével (c-ChIP, FRAP, FCS). A kompetíciós ChIP (c-ChIP) eljárásban (Lickwar et al. 2012, 2013) egy konstitutív és egy indukálható Spp1 izoformát "versenyeztettünk" egymással a genomi kötőhelyekhez történő kötődésért, amely alapján ún. kicserélődési rátát számoltunk (turnover rate), amely az időegység alatt történő Spp1 kicserélődések számát mutatja az egyes Spp1 kötőhelyekre vonatkozóan (Spp1 kicserélődés / perc / kötőhely; mértékegysége: [1/perc]). Mérésünkben 9xmyc tag-gel jelöltük a konstitutív Spp1 allélt amelyet endogén promóteréről expresszáltunk (Spp1-9xmyc), míg a kompetitor Spp1 allélt GFP-vel fúzionáltattuk egy réz-indukálható promóter (*pCUP1*) szabályozása alatt (GFP-Spp1). A megfelelő időbeli felbontóképesség elérése érdekében a réz-indukciót követően 30 percenként vettünk mintát a meiotikus tenyészetekből, majd western blot elemzéssel ellenőriztük a kompetitor allél expressziójának exponenciális növekedését. Ezt követően kromatin immunprecipitációt (ChIP) végeztünk anti-myc és anti-GFP antitestekkel, majd genomi kötőhelyeket azonosítottunk a "hagyományos" ChIP-seq elemzéshez használt bioinformatikai módszerekkel. Kicserélődési rátát azokra a peak-ekre definiáltunk, amelyek illeszthetőek voltak az elméleti exponenciális modell alapján ($\frac{GFP}{myc} = (1 - e^{-\lambda t})$, ahol λ a turnover ráta, t pedig a kompetitor allél indukciójától számított idő. A modell illesztése után kiszámítottuk a becslések standard hibáját (SE), amely figyelembe vételével összesen 977 c-ChIP kötőhelyet és Spp1 turnover rátát definiáltunk. Elemzésünkéből kiderült, hogy

az Spp1 c-ChIP kötőhelyek markánsan szétválaszthatók turnover rátájuk alapján, amelyek nagyban átfednek az Spp1 kötőhelyek kinetikus viselkedésük szerinti klasztereivel (megjelenő, eltűnő, konstant Spp1 peak-ek). Megállapítottuk, hogy minél erősebb egy adott kromatin régió Mer2-asszociációja, annál kisebb az Spp1 kicserélődési rátája (**27/E ábra**), ami azt jelenti, hogy a Mer2-gazdag területek stabilan kötik az Spp1-et, illetve az Spp1 Mer2-axiális helyekkel való kapcsolata lecsökkenti annak mobilitását. Összefoglalva, az Spp1 kötőhelyek differenciális turnover értékei kétféle Spp1 pool jelenlétét bizonyítják a meiózis során, amelyek markánsan eltérő asszociációt mutatnak a H3K4me3 modifikációval (Set1C-függő frakció) és a meiotikus DSB-komplexxel (Set1C-független frakció). A c-ChIP eredményeket kvantitatív mikroszkópos módszerekkel (FRAP, FCS) validáltuk élő (nem-fixált) meiotikus sejtekben, amely megerősítette, hogy az Spp1 a meiózis során Set1C-független működésre vált, amely dinamikai tulajdonságainak megváltozásában és kromoszómális relokalizációjában tettenérhető.

A hiszton H3K56 acetiláció szerepe a meiotikus homológ rekombináció iniciációjában

A hiszton H3K4 trimetiláció a genom számos pontján beindíthatja a homológ rekombinációt, azonban más hisztonmódosítások is szükségesek a folyamat flexibilitásához és a rekombinációs haplotípusok sokféleségéhez (Székvölgyi and Nicolas 2010). További hisztonmodifikációk és alternatív epigenetikai útvonalak azonosítása tehát segíthet megérteni a meiotikus crossover mintázatok kialakulását és azok evolúciós plaszticitását. Kísérleteinkben ezért a H3 hiszton globuláris doménjén elhelyezkedő lizin 56 acetiláció szerepére összpontosítottunk (H3K56ac), amely a DNS be- és kilépési pontjainál destabilizálhatja a hiszton/DNS kölcsönhatásokat (Buning and Van Noort 2010; Simon et al. 2011). Kísérleteinkben ezért funkcionális vizsgálatokat végeztünk egy sor élesztő mutánsban, amelyekben lecsökkentettük vagy megnöveltük a H3K56ac szintjét a meiotikus differenciáció során. A H3K56 acetiláció szinte teljesen eltűnt az *asf1Δ* és *rtt109Δ* mutánsokban, ami összhangban van az Asf1 (hiszton chaperon) és Rtt109 (K56-specifikus acetiltranszferáz) ismert biokémiai szerepével a H3K56 acetiláció kromatinra írásával kapcsolatban (Tsubota et al. 2007). A K56ac szignál jelentősen megnőtt a *hst3/4Δ* kettős-mutánsban, amely megakadályozza a H3K56ac deacetilációját (Celic et al. 2006). Ezután a *HIS4-LEU2* "mesterséges" rekombinációs hotspot régióban Southern blot módszerrel térképeztük és kvantifikáltuk a meiotikus DSB-eket és a crossover rekombinációk mennyiségét. Méréseink szerint az összes mutáns fiziológiás DSB-szintet produkált (amely gyakorisága a vad típusú mintákkal összevethető volt; 6-8% a görbék globális maximuma alapján), azonban az *asf1Δ* és *hst3/4Δ* mutánsokban a DSB-k módosult kinetikával jelentek meg és tűntek el,

kb. 60 perces időbeli késéssel. Fontos kiemelni, hogy a mutánsokban azonosított meiotikus DSB-k erősen rekombinogének voltak (lásd az R1/R2 rekombinánsokat), vagyis a *HIS4-LEU2* szakaszon keletkező DNS törések a “természetes” repair útvonalakon processzálódtak.

A modifikálható lizin 56 aminosav oldallánc (K56) közvetlen szerepének teszteléséhez egy funkcionális kísérletet végeztünk, amelyben egy nem-módosítható H3K56A mutáns hiszton molekulát expresszáltunk a H3 hiszton fehérje egyedüli forrásaként a sejtekben. A DSB-eket Southern blot-tal térképezve háromszoros csökkenését figyeltünk meg a H3K56A mutánsban, amit ezután független RPA ChIP-qPCR módszerrel is validáltunk. A K56 acetiláció és a DSB hotspotok kapcsolatát H3K56ac ChIP-Chip és RPA ChIP-Chip módszerrel vizsgáltuk a teljes genomban, amelyekben anti-H3K56ac és anti-H3 antitestek felhasználásával immunprecipitáltuk a hiszton fehérjéket, majd meghatároztuk a H3K56ac nukleosómák dúsulását az össz H3-hoz képest, illetve ugyanezt elvégeztük anti-Rfa1 antitestekkel kontroll és H3K56A mutáns sejtekben, amellyel a meiotikus DSB-eket jelöltük meg. Eredményeink szerint a H3K56ac hisztonok preferenciálisan dúsultak az RPA ChIP peak-ek közvetlen közelében (a DSB-k mellett) a kontroll sejtekben, míg nem mutattak asszociációt a H3K56A mutánsban azonosított (“megmaradó”) RPA ChIP peak-ekkel. Az RPA ChIP peak-ek száma szignifikánsan lecsökkent a H3K56A mutánsban, mivel a mutáció a kontroll sejtekben detektált 1004 RPA peak 65,8%-át eliminálta (661 peak), 343 RPA-helyet nem érintett (metszet halmaz), míg 182 esetben “ektópikus” RPA peak-ek keletkeztek, amelyek nem voltak jelen a vad típusú sejtekben. Fontos kiemelni, hogy az RPA szignál K56A mutánsban tapasztalt csökkenését nem detektáltuk kromoszóma tengely (Mer2) kötőhelyeken és véletlenszerűen kiválasztott (random) genomi pozíciókban, amely az RPA kötőhelyeken kapott különbség specificitását mutatja. Az RPA peak-ek számának jelentős genomi csökkenése összecseng a Southern blot eredményekkel, amely összességében H3K56 oldallánc közvetlen szerepét bizonyítja a rekombináció iniciációjában. A H3K56-acetilezett hisztonok tehát fiziológias szintre állítják meiotikus DNS törések mennyiségét, mintegy finomhangolva a rekombinációs források működését.

Kromoszómális R-hurok és R-hurok regulátorok molekuláris vizsgálata

Az R-hurok struktúrák rekombinogén és mutagén hatása régóta ismert, azonban a DNS törésekkel való közvetlen fizikai kapcsolatukat nem sikerült bizonyítani. Funkcionális kísérletek elvégzéséhez egy megbízható R-hurok térképezési módszer szükséges, amely hatékonyan képes az R-hurok struktúrákat azonosítani a genomban, azonban időközben bevezetett DRIP módszer (Ginno et al. 2012) rendre inkonzisztens adatokat generált, amely a publikált R-loop eredmények jelentős

variabilitásában tettenérhető. Ezért célul tűztük ki egy univerzális genomikai térképezési módszer kidolgozását, amely az “alap” DRIP módszer torzításait kikerülve alkalmas az R-hurkok nagy pontosságú azonosítására bármely modellorganizmusban.

A publikált protokollokat áttanulmányozva meghatároztuk a legfontosabb kísérleti változókat (sejtfixálás, lízis hőmérséklete, nukleinsav izolálás, szabad RNS eltávolítása, genom-fragmentáció), amelyek kombinálásával 40 kísérleti sémát (osztályozó változót) terveztünk. Az osztályozó változók segítségével felmértük, hogy azok hogyan rangsorolnak különböző teszt lókuszokat R-hurok státuszuk alapján (“tréning szett”), amelyet publikus adatokból és saját mérésekből előzetesen megismertünk. Saját DRIP-seq kísérletünket kétféle humán mintán végeztük el (Jurkat T-sejt leukémia sejtvonal és CD4+ T limfociták), amelyben 88,830 és 99,337 R-hurok régiót azonosítottunk.

Az azonosított DRIP kötőhelyekből konszenzus R-hurok készletet állítottunk elő és valós-pozitív / valós-negatív teszt szakaszokat választottunk tréning szettként. A kiválasztott lókuszokon RNázH-kontroll mellett 5 független ismétlésben elvégeztük a 40-féle DRIP kísérletet, amely ~4000 DRIP-qPCR értéket (hozamot) eredményezett. A DRIP-qPCR adatokat ezután ROC elemzés független változójaként felhasználva meghatároztuk a 40 függő változónk (osztályozó változók) szenzitivitását és specificitását. A ROC görbék összehasonlítása során tíz osztályozó változó esetén (5, 6, 13, 15, 17, 18, 19, 21 és 24) magas AUC értékeket kaptunk (>0,7), vagyis ezek a DRIP sémák nagy szenzitivitással és specificitással képesek az R-hurkokat azonosítani. Négy DRIP sémában 0,5 körüli AUC-t kaptunk (2, 1, 11, 16), ami arra utal, hogy ezek a kísérletek véletlenszerű válaszokat adnak, ezért alkalmazásuk feltétlenül kerülendő. A legjobban teljesítő két DRIP séma a következő volt: 5 és 13, 79.1-79.5%-os AUC-val, 70-75%-os érzékenységgel, és 72-80% specificitással. Fontos megfigyelésünk volt, hogy a DRIP kísérletekben rutinszerűen alkalmazott restriktív enzim (RE) felismerési szekvenciák eloszlása nem véletlenszerű a genomban, a vágási helyek sűrűsége szignifikánsan magasabb volt az intergénikus régiókban, mint a fehérjekódoló ORF-ekben. Különösen az első exonokban alulreprezentáltak az RE felismerő motívumok, amely torzított mintavételt okoz. Emiatt a restriktív enzimekkel történő genom-fragmentálás és DRIP szekvenálás hosszú, restriktív fragmentum méretű R-hurkokat azonosít a fehérjekódoló gének közelében, amely a kódoló ORF-ek első exonjánál okozza a legnagyobb torzítást. Szonikált mintáknál ellenben az R-hurkok genomi pozíciója pontosan behatárolható, mivel az ultrahang véletlenszerűen darabolja a DNS-t. A nukleinsav fragmentáció tehát jelentősen befolyásolja az R-hurkok méreteloszlását és genomi térképezhetőségét, és ezáltal a biológiai funkciójuk pontos megítélését. A fenti torzítás, amelyet “first-exon bias”-ként írtunk le, előzetes *in silico* restriktív elemzéssel tesztelhető és kiküszöbölhető.

A fenti metodikai tapasztalatokra alapozva célul tűztük ki a kromoszómális R-hurok struktúrák nagy pontosságú annotációját és új R-hurok regulátor fehérjék azonosítását különféle modell szervezetekben. Első kísérletünket lúdfűben (*Arabidopsis thaliana*) végeztük, mivel ebben a modellorganizmusban elsőként azonosítottak egy R-hurok-asszociált fehérjét, amely a virágzást szabályozó *FLOWERING LOCUS C (FLC)* gén területén specifikusan megkötött egy R-hurok struktúrát (Sun et al. 2013). Ez a fehérje a *Nodulin homeobox (NDX)* transzkripció faktor volt, az egyetlen ismert makromolekula kísérleteink kezdetéig, amely közvetlen szerepet játszott egy R-hurok felismerésében. A lúdfű kicsiny, teljesen annotált genomja és viszonylag egyszerű genetikai manipulálhatósága lehetőséget kínált az NDX-el kapcsolatos hipotéziseink gyors és hatékony tesztelésére. Kísérleteinkben párhuzamosan vizsgáltuk az NDX és R-hurok kromoszómális eloszlását ChIP és DRIP szekvenálással. Az NDX genomi kötőhelyeit független transzgenikus vonalakban expresszált N-terminálisan jelölt flag-NDX és C-terminálisan jelölt NDX-GFP fúziós fehérjék kromatin immunprecipitációjával és NGS vizsgálattal határoztuk meg. Az NDX kötőhelyek jellemzően a kromoszómák centromerikus és pericentrikus régióiban halmozódtak fel, ahol sűrűségük (84 peak/Mb) szignifikánsan nagyobb volt, mint a kromoszóma karok mentén (16 peak/Mb). Előbbi régiók funkcionálisan az ún. perifériás heterokromatinnak felelnek meg, amely többnyire erősen kondenzált, génekben szegény és funkcionálisan teljesen vagy részlegesen inaktív, míg az utóbbi régiók eukromatikusak, tehát laza szerkezetűek és elsősorban aktívan átíródó géneket tartalmaznak. Ezt követően DRIP-seq kísérletekben elemeztük az *Arabidopsis thaliana* kromoszómáin elhelyezkedő R-hurok számát és elhelyezkedését. Genomi annotációjukat az NDX kötőhelyekkel összevetve megállapítottuk, hogy az R-hurok teljesen komplementer pozíciót foglalnak el az NDX kötőhelyekhez képest, 97%-uk ugyanis semmilyen átfedést nem mutat az NDX peak-ekkel, és fordítva, az NDX peak-ek 94%-a független az R-hurokuktól.

Az R-hurok jellemzően a gének 5' UTR régiójában dúsultak és mérsékelt együttálást mutattak bizonyos aktiváló hiszton modifikációkkal (H3K4me2, H3K4me3). Az NDX kötőhelyek ezzel szemben a transzpozonokon / transzpozon géneken dúsultak és represszív hiszton módosításokkal (H3K9me2, H3K27me1) mutattak erős korrelációt. Az *Arabidopsis*-ban azonosított kilenc funkcionális kromatinállapot közül (Sequeira-Mendes et al. 2014) az NDX a heterokromatikus state 8-9-el mutatott szignifikáns átfedést, míg az R-hurok az eukromatikus state 1-2-re voltak jellemzőek. A legmarkánsabb együttállást a közelmúltban azonosított nem-kódoló kisRNS lókuszek (Hardcastle et al. 2018) egy csoportja, a 'klaszter 6-9' sRNS-ek, és az NDX kötőhelyek között figyeltük

meg, melyek a pericentrikus heterokromatin-asszociált siRNS-eket (het-siRNS) kódolják. Az R-hurkok ebben az esetben is antagonisztikus eloszlást követtek az NDX kötőhelyekhez képest, mivel a 'klaszter 1-2' sRNS lókuszokon mutattak szignifikáns kötődést.

A továbbiakban egy T-DNS inzerciós mutánsban (*ndx1-4*) teszteltük az NDX fehérje het-siRNS asszociációjának funkcionális következményeit. Kontroll (Col-0) és *ndx1-4* mutáns növényekben kisRNS szekvenálást (sRNS-seq) végeztünk, és megállapítottuk, hogy az NDX hiánymutánsban szignifikánsan több funkcionális kisRNS (21nt/24nt) szintetizálódik, mint NDX jelenlétében. A kilenc sRNS osztály közül a 4-9. csoport mutatott megnövekedett expressziós szintet az *ndx1-4* mutánsban, míg az 1-3. csoport nem különbözött a vad típustól. Ez a transzkripció heterogenitás összhangban van az sRNS osztályok elsődleges funkcionális felosztásával, amely pontosan az 1-3. és 4-9. csoportok között történik (előbbi a fehérjekódoló génekkel és a promóterekkel mutat szoros asszociációt, míg utóbbi az epigenetikailag aktivált siRNS-ekkel (easiRNS) és az RNS-függő DNS metilációs (RdDM) útvonallal (Matzke and Mosher 2014).

Mivel a növényekben a het-siRNS-ek szabályozzák a *de novo* DNS metilációt (másnéven RNS-függő DNS metilációt; RdDM) és a géncsendesítést (Matzke and Mosher 2014; Hardcastle et al. 2018), DNS-metilom vizsgálatokat végeztünk Col-0 és *ndx1-4* növényekben biszulfid-szekvenálással (BS-seq), amely a nem-metilált citozin bázisok timinné konvertálásán alapul. BS-seq kísérletünkben mindhárom szekvencia kontextusban (CG/CHG/CHH) meghatároztuk a metilált citozinok arányát Col-0 és *ndx1-4* növényekben, majd differenciálisan metilált régiókat (DMR) azonosítottunk, amelyek szignifikánsan alacsonyabb vagy magasabb metilációs szintet mutattak a mutánsban (előbbieket hipo-DMR-eknek, utóbbiakat hiper-DMR-eknek hívjuk). Elemzésünk 2449 hipometilált és 1597 hipermetilált régiót mutatott ki a CHH és CHG szekvenciáknál, amelyek jellemzően a pericentromérius régiókban (hipo-DMR-ek) és a kromoszóma karokon (hiper-DMR-ek) dúsultak. A DMR-ek közül a hipo-CHH/CHG DMR-ek szignifikáns kolokalizálót mutattak az NDX kötőhelyekkel (amelyeket a vad típusú növényben azonosítottunk), amely az NDX direkt hatására utal. A fenti siRNS és DNS-metilációs változások kromatinszerkezeti hátterének vizsgálatához Hi-C kísérleteket végeztünk Col-0 és *ndx1-4* mutáns növényekben, hogy kiderítsük, a megfigyelt különbségek összefüggésbe hozhatók-e a genomszerkezet változásaival. Mindkét mintából ~200 millió paired-end read-et szekvenáltunk, amelyekből nagyfelbontású Hi-C interakciós mátrixokat generáltunk. A mátrixok kvantitatív elemzésével szignifikáns különbségeket azonosítottunk a Col-0 és *ndx1-4* minták között, amely számos genomi régióban megnövekedett (piros) vagy csökkent (kék) Hi-C értéket tárt fel. E különbségek a kromatinszerkezet globális megváltozását jelzik az NDX fehérje

hiányában. Az egyik legmarkánsabb változást az ún. KNOT (csomó) struktúrában figyeltük meg, amely a vad típusú növényekre jellemző tíz interkromoszómális régió helyett (Grob et al. 2014) új *de novo* elemekkel egészült ki az *ndx1-4* mutánsban. A másik markáns Hi-C változás a centromerekre és pericentromerikus régiókra jellemző, amely szignifikánsan csökkent intrakromoszómális kapcsolatokat és felerősödő interkromoszómális interakciókat mutat az *ndx1-4* mutánsban. Az eltűnő intracentromerikus kölcsönhatások a kromatin fellazulására utalnak NDX hiányában (dekompakció), amely egybeesik az NDX jellemző genomi kötőhelyeivel és erősen korrelál az *ndx1-4* mutánsban megfigyelhető sRNS-expressziós és CHH/CHG-metilációs változásokkal. A fentieket összefoglalva, különféle genomikai és sejtbiológiai módszerekkel az NDX fehérjéről bebizonyítottuk, hogy egy alapvető heterokromatin regulátor, amely a korábbi elképzelésekkel szemben nem az R-hurok-gazdag eukromatikus kromoszóma karokon, hanem a heterokromatikus centromer és pericentromerikus régiókban működik, ahol a het-siRNS-ek transzkripcióját és a represszív CHH/CHG metilációt kontrollálja.

További vizsgálatainkban figyelmünket a humánspecifikus R-hurok regulátor fehérjék felé fordítottuk, amelyeket egyáltalán nem ismertünk a 2018-as évig. Ekkor jelent meg egy tanulmány (Wang et al. 2018), amelyben tömegspektrometriás vizsgálattal és immunprecipitációval több száz potenciális R-hurok-kötő fehérjét azonosítottak *in vitro*, amely megnyitotta az utat az R-hurok regulátor gének vizsgálatához különféle fiziológias és patológias folyamatokban. A rákos megbetegedések kapcsán például az elmúlt 10-12 év kutatásában világossá vált, hogy az R-hurok kontrollálatlan keletkezése és a tumorokra jellemző genom instabilitás között van valamilyen mechanikai kapcsolat, azonban a résztvevő faktorokat és funkcionális kölcsönhatásokat nem tudták azonosítani. A fenti összefüggések mélyebb megértéséhez farmakogenomikai vizsgálatokat végeztünk a szakirodalomban azonosított (és kísérletesen validált) R-hurok regulátor fehérjékkel, amely azon az elképzelésen alapult, hogy a daganatok permanensen magas R-hurok szintje kihasználható a rákos sejtek kemoterápiás kezelésekkal szembeni érzékennyé tételére. Vizsgálatainkhoz olyan R-hurok regulátor géneket választottunk, amelyek egyértelmű R-hurok asszociációval rendelkeznek a kísérletes adatok alapján, különféle molekuláris útvonalakon működnek (pl. DNS hibajavítás, splicing, hiszton modifikáció, stb.), és bizonyítottan szerepet játszanak a rák kialakulásában. Ezen R-hurok regulátorok mRNS szignatúráit 33 primer tumorban vizsgáltuk a Cancer Genome Atlas (TCGA) adatbázis mRNS-seq adatai alapján, amelyet egészséges kontroll szövetekhez viszonyítva számos tumor-specifikus génexpressziós különbségeket azonosítottunk. Ezt követően a primer tumorokban mért génexpressziós értékeket nagyszámú rákos

betegcsoport túlélési adataival kapcsoltuk össze, majd megvizsgáltuk, hogy az R-hurok regulátorok transzkripció szintje korrelál-e a betegek túlélési arányával. Elemzésünkben 12,862 Kaplan-Meier túlélési görbét állítottunk elő és 179 olyan tumor típust azonosítottunk, amely szignifikáns R-hurok regulátor-függő túlélés asszociációt mutatott. Az R-hurok regulátorokat magasan expresszáló és alacsonyan expresszáló betegcsoportok túlélési esélyei szignifikánsan és egyértelműen elváltak egymástól, és az esetek 70%-ában (n=123) az R-hurok regulátorok alacsony expressziós szintje volt előnyös a betegség lefolyása szempontjából. További megfigyelés, hogy bizonyos gének, pl. az RNASEH2A, BLM, és BRCA1 (top 3), számos ráktípusban mutattak szignifikáns túlélés-asszociációt, amely ezen fehérjék általános szerepére utal, így potenciális biomarkerként tekinthetünk rájuk. Ezzel szemben más gének, pl. az ATXN2, DDX19A és U2AF1 (utolsó 3), hatása csak bizonyos tumor típusokra korlátozódott, amely nagyfokú szövetspecifitásukra utal és csökkenti diagnosztikai vagy terápiás jelentőségüket.

A következő lépésben szinergisztikus gyógyszer kölcsönhatásokat kerestünk a túlélés asszociációt mutató primer tumorokkal szövettanilag és genetikailag egyező tumorsejtvonalakban. Vizsgálatunkban 267 FDA (US Food and Drug Administration) által engedélyezett terápiás készítmény érzékenységi interakcióit határoztuk meg 1065 tumorsejtvonal farmakogenomikai adatai alapján (GDSC, (Yang et al. 2013)). Számításainkban azonosítottuk azokat a tumor típusokat, amelyek érzékenyebben reagálnak (vagy rezisztenssé válnak) bizonyos terápiákra egy adott R-hurok regulátor expressziójának függvényében. Például, a CX-5461 nevű kemoterápiás szerrel (RNS polimeráz I inhibitor, G4 ligand) kezelt sarcoma sejtvonalban a THOC2 R-hurok gén magas expressziója magas IC₅₀ értékekkel (félhalálos dózis) korrelált, vagyis a CX-5461 nem hatékony THOC2-pozitív sarcoma-k kezelésére, használata kerülendő. Ezzel szemben a CX-5461-el kezelt melanoma és endometriális tumorsejtekben az ATXN1 és TREX1 R-hurok gének magas expressziója alacsony IC₅₀ értékekkel korrelált, vagyis a ATXN1-pozitív és TREX1-pozitív tumorok érzékenyen reagálnak a CX-5461 terápiára, használatuk javasolt. A vizsgált terápiás szerek számos biológiai útvonalat felöleltek (a protein kináz és MAPK szignalizációtól a citoskeletonon át a DNS hibajavításig), amelyek közül az PI3K/mTOR, MAPK/sejtciklus és DNS replikáció útvonalakra ható szerek mutatták a legtöbb génkölcsönhatást. Összességében 508 olyan R-hurok regulátor-függő gyógyszer kölcsönhatást detektáltunk, amely egyben szignifikáns túlélés asszociációt is mutatott a megfelelő primer tumoros populációban.

Az R-hurok regulátorok expressziós státusza tehát jelentősen befolyásolja a betegek túlélési esélyeit valamint a gyógyszeres kezelésekre adott választ is módosíthatja. Ezen farmakogenomikai

összefüggések klinikai felhasználása még várat magára, azonban megfelelő preklinikai vizsgálatok elvégzése után alkalmas lehet a betegcsoportok prioritizálására. A rákos sejtekben zajló R-hurok képződési folyamatok és az azokat kontrolláló fehérjék tehát potenciális biomarkerek és/vagy terápiás célpontok, amelyek alkalmasak lehetnek egyes daganattípusok kemoterápiás szerekekkel szembeni érzékenyítésére. Eredményeink, az alapkutatási jelentőségükön túl, új utat nyithatnak az epigenetikai terápiák továbbfejlesztéséhez, amely a daganatok R-hurok szintjének módosításán alapul.

ÚJ MEGÁLLAPÍTÁSOK

A meiotikus homológ rekombinációval kapcsolatos munkánkban megmutattuk, hogy a Set1C hiszton metiláz komplex által katalizált H3K4me3 okozati szerepet játszik a DNS duplaszáltörések kialakulásában. Igazoltuk, hogy a DSB-k kialakulása függ a módosítható lizin 4 oldallánc (H3K4) jelenlététől, a Set1 alegység katalitikus aktivitásától, és az Spp1 alegység jelenlététől. Bizonyítottuk, hogy az Spp1 a DSB képződés során i) fizikai kölcsönhatásba lép a DSB-komplex egyik tagjával (Mer2), amely egy cinkujyszerű Mer2-interakciós doménon keresztül történik (CXXC Mer2-ID), ii) felismeri és megköti a Set1-által katalizált H3K4me3 szignált, amely a PHD-finger doménje által történik. Bizonyítottuk, hogy a Mer2-ID és PHD-domének deléciója a DSB-szint jelentős csökkenését okozza. Egy új rekombinációs modellt vezettünk be, amely egyrészt kiterjesztette a régebbi hurok-tengely modellt (Blat et al. 2002), másrészt az Spp1-Mer2 kölcsönhatás alapján molekuláris szinten megmagyarázta a meiotikus DNS-törések kapcsolatát a Set1C-katalizált H3K4me3 módosítással. Megmutattuk, hogy az Spp1 diffúziós tulajdonságai és dinamikus turnover rátája alapvető fontosságú a rekombinációt megelőző kromatinváltozások kialakulásához. Ennek kapcsán bizonyítottuk, hogy az Spp1 Set1C/COMPASS-független működésre vált a meiózis során, amely elősegíti a rekombinációs iniciációs helyek aktiválását.

A Set1C mellett egyéb epigenetikai modifikációk szerepét is bizonyítottuk a meiotikus rekombináció során. Megmutattuk, hogy a központi hiszton-fold doménre koncentrálnak H3 lizin 56 acetiláció (H3K56ac) fiziológias szintre állítja be meiotikus DSB-k mennyiségét és ezzel finomhangolja a rekombinációs forrópontok működését.

Az R-hurok struktúrákkal kapcsolatos munkánk során kidolgoztunk egy nagy pontosságú analitikai eljárást, amely az R-hurok azonosítására és genomi térképezésére alkalmas bármely modellorganizmusban. Felismertük az ún. "első-exon torzítás" jelenségét, amely bizonyos genomfragmentációs eljárásoknál torzítja az R-hurok méreteloszlását és genomi térképezhetőségét, és ezáltal a biológiai funkciójuk megítélését. E metodikára alapozva R-hurok térképezést végeztünk *Arabidopsis thaliana*-ban, hogy az addig egyetlen ismert R-hurok-kötő fehérje, a *Nodulin homeobox* (NDX) működését molekuláris szinten megértsük. Az NDX fehérjéről korábban azt feltételezték, hogy egy általános R-hurok regulátor, amely kötődésével stabilizálja e struktúrákat és ezzel bizonyos gének aktivitását befolyásolja (Sun et al. 2013). Ezzel szemben mi megmutattuk, hogy az NDX egy alapvető heterokromatin regulátor, amely a korábbi elképzelésekkel szemben nem az R-hurok-gazdag eukromatikus kromoszóma karokon, hanem a heterokromatikus régiókban működik. Bizonyítottuk, hogy az NDX kromoszómális eloszlása ellentétes az R-hurokokhoz képest és az RNS-függő DNS metilációs rendszerrel (RdDM) együttműködve a heterokromatin homeosztázisát

szabályozza a centromerikus és pericentrometikus régiókban. Az NDX hatásmechanizmusával összefüggő genomszerkezet változások detektálásához genomkonformáció leképezés alapú módszereket implementáltunk. NDX hiánymutánsban végzett Hi-C vizsgálatainkban jelentős genomszerkezet változásokat detektáltunk az NDX kötőhelyekben gazdag genomi régiókban (pericentromerek), amely erősen korrelált a megfigyelt het-siRNS expresszió és DNS metilációs változásokkal. E molekuláris változások fenotípusos következményeit további funkcionális vizsgálatok elvégzésével próbáljuk feltérképezni.

Humánspecifikus R-hurok regulátor fehérjékkal végzett farmakogenomikai vizsgálatainkban megmutattuk, hogy R-hurok regulátorok expressziós szintje kihasználható a rákos sejtek kemoterápiás kezelésekkal szembeni érzékenyvé tételére. Emellett számos olyan tumor típust azonosítottunk, amely szignifikáns R-hurok regulátor-függő túlélés asszociációt mutatott, vagyis az R-hurok regulátorokat magasan expresszáló és az alacsonyan expresszáló betegcsoportok túlélési esélyei szignifikánsan elváltak egymástól. A fenti összefüggések miatt az R-hurok regulátorokat mindenképp érdemes potenciális diagnosztikai markerként vagy terápiás célpontként tekintetbe venni.

KÖSZÖNETNYILVÁNÍTÁS

Az értekezésben bemutatott eredmények a Genomszerkezet és Rekombináció Kutatócsoport munkájából származnak, amelyért elsősorban Nekik tartozom köszönettel: Karányi Zsolt, Miskei Márton, Feró Orsolya, Nagy Éva, Horváth Adrienn, Boros-Oláh Beáta, Fillér Csaba, Hornyák Lilla, Fazekas-Bálint Ágnes, Varga Dóra, Mosolygó-Lukács Ágnes, Nagy Dénes, Sipos Éva, Hetey Szabolcs, Halász László, Fürtös Ibolya, Petrucz Anita.

A kutatócsoport nem jöhetett volna létre a Magyar Tudományos Akadémia Lendület programjának támogatása nélkül, amelyért köszönettel tartozom az MTA döntéshozóinak és a bírálóknak.

Hálás vagyok a Debreceni Egyetem ÁOK Biokémiai és Molekuláris Biológiai Intézetnek, hogy 2015-ben befogadta a csoportomat, lehetőséget és teret biztosítva a kutatásainknak.

Végül, köszönettel tartozom családomnak. Hálás vagyok végtelen türelmetekért, bátorításotokért, és kitartásotokért, amely nélkül bizonyosan nem juthattam volna el idáig.

A DISSZERTÁCIÓ ALAPJÁT KÉPEZŐ SAJÁT KÖZLEMÉNYEK

1. Karányi Zs. Mosolygó-L Á, Feró O, Horváth A, Boros-Oláh B, Nagy É, Hetey Sz, Holb I, Szaker H, Miskei M, Csorba T, **Székvölgyi L**: NODULIN HOMEBOX is required for heterochromatin homeostasis in Arabidopsis. *Nature Communications*. 13 5058, 2022. (D1, IF: 17.694)
2. Karányi, Z., Hornyák, L., **Székvölgyi, L.**: Histone H3 lysine 56 acetylation is required for formation of normal levels of meiotic DNA breaks in *S. cerevisiae*. *Front. Cell. Dev. Biol.* 7 1-18, 2020. (Q1, IF: 6.684)
3. Boros-Oláh, B., Dobos, N., Hornyák, L., Szabó, Z., Karányi, Z., Halmos, G., Roszik, J., **Székvölgyi, L.**: Drugging the R-loop interactome: RNA-DNA hybrid binding proteins as targets for cancer therapy. *DNA Repair*. 84 1-10, 2019. (Q1, IF: 3.339)
4. Hegedűs, É., Kókai, E., Nánási, P., Imre, L., Halász, L., Jossé, R., Antunovics, Z., Webb, M., El Hage, A., Pommier, Y., **Székvölgyi, L.**, Dombrádi, V., Szabó, G.: Endogenous single-strand DNA breaks at RNA polymerase II promoters in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 46 (20), 10649-10668, 2018. (D1, IF: 11.147)
5. Karányi, Z., Halász, L., Acquaviva, L., Jonás, D., Hetey, S., Boros-Oláh, B., Peng, F., Chen, D., Klein, F., Géli, V., **Székvölgyi, L.**: Nuclear dynamics of the Set1C subunit Spp1 prepares meiotic recombination sites for break formation. *J. Cell Biol.* 217 (10), 3398-3415, 2018. (D1, IF: 8.891)
6. Hornyák, L., Dobos, N., Koncz, G., Karányi, Z., Páll, D., Szabó, Z., Halmos, G., **Székvölgyi, L.**: The role of indoleamine-2,3-dioxygenase (IDO) in cancer development, diagnostics, and therapy. *Front. Immunol.* 9 (151), 1-8, 2018. (Q1, IF: 4.716)
7. Roszik, J., Fenyőfalvi, G., Halász, L., Karányi, Z., **Székvölgyi, L.**: In Silico Restriction Enzyme Digests To Minimize Mapping Bias In Genomic Sequencing. *Mol. Ther. Methods. Clin. Dev.* 6 66-67, 2017. (Q2, IF: 3.681)
8. Halász, L., Karányi, Z., Boros-Oláh, B., Kuik-Rózsa, T., Sipos, É., Nagy, É., Mosolygó, Á., Türk-Mázló, A., Rajnavölgyi, É., Halmos, G., **Székvölgyi, L.**: RNA-DNA hybrid (R-loop) immunoprecipitation mapping: an analytical workflow to evaluate inherent biases. *Genome Res.* 27 1063-1073, 2017. (D1, IF: 10.101)
9. **Székvölgyi, L.***, Ohta, K., Nicolas, A*.: Initiation of meiotic homologous recombination: flexibility, impact of histone modifications, and chromatin remodeling. * levelező szerzők *Cold Spring Harbor Perspect. Biol.* 7 (5), 1-16, 2015. (D1, IF: 9.173) * megosztott levelező szerzők
10. Acquaviva, L*, **Székvölgyi, L.***, Dichtl, B., Dichtl, B., Saint André, C., Nicolas, A., Géli, V.: The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. *Science*. 339 (6116), 215-218, 2013. (D1, IF: 31.477) * megosztott első szerzők

11. **Székvölgyi, L.**, Nicolas, A.: From meiosis to postmeiotic events: Homologous recombination is obligatory but flexible. *FEBS J.* 277 (3), 571-589, 2010. (Q1, IF: 3.129)
12. **Székvölgyi, L.**, Rákossy, Z., Bálint, B., Kókai, E., Imre, L., Vereb, G., Bacsó, Z., Goda, K., Varga, S., Balázs, M., Dombrádi, V., Nagy, L., Szabó, G.: Ribonucleoprotein-masked nicks at 50-kbp intervals in the eukaryotic genomic DNA. *Proc. Natl. Acad. Sci. U.S.A.* 104 (38), 14964-14969, 2007. (D1, IF: 9.598)

EGYÉB ELSŐ ÉS UTOLSÓSZERZŐS KÖZLEMÉNYEK

13. Miskei, M., Horváth, A., Viola, L., Varga, L., Nagy, É., Feró, O., Karányi, Z., Roszik, J., Miskey, C., Ivics, Z., **Székvölgyi, L.**: Genome-wide mapping of binding sites of the transposase-derived SETMAR protein in the human genome. *Computational and Structural Biotechnology Journal.* 19 4032-4041, 2021. (D1, IF: 6.155)
14. Szabó, Z., Hornyák, L., Miskei, M., **Székvölgyi, L.**: Two targets, one hit: new anticancer therapeutics to prevent tumorigenesis without cardiotoxicity. *Front. Pharmacol.* 11 1-5, 2021. (Q1, IF: 5.988)
15. Hetey, S., Boros-Oláh, B., Kuik-Rózsa, T., Li, Q., Karányi, Z., Szabó, Z., Roszik, J., Szalóki, N., Vámosi, G., Tóth, K., **Székvölgyi, L.**: Biophysical characterization of histone H3.3 K27M point mutation. *Biochem. Biophys. Res. Commun.* 490 (3), 868-875, 2017. (Q1, IF: 2.559)
16. **Székvölgyi, L.**, Imre, L., Doan-Xuan, Q., Hegedűs, É., Bacsó, Z., Szabó, G.: Flow Cytometric and Laser Scanning Microscopic Approaches in Epigenetics Research. *Methods Mol. Biol.* 567 99-111, 2009. (Q3)
17. **Székvölgyi, L.**, Bálint, B., Imre, L., Goda, K., Szabó, M., Nagy, L., Szabó, G.: Chip-on-beads: flow-cytometric evaluation of chromatin immunoprecipitation. *Cytometry A.* 69A (10), 1086-1091, 2006. (Q1, IF: 3.293)
18. **Székvölgyi, L.**, Hegedűs, É., Molnár, M., Bacsó, Z., Szarka, K., Beck, Z., Dombrádi, V., Austin, C., Szabó, G.: Nick-forming sequences may be involved in the organization of eukaryotic chromatin into ~50 kbp loops. *Histochem. Cell Biol.* 125 (1-2), 63-73, 2006. (D1, IF: 3.22)