



Budapesti Műszaki és Gazdaságtudományi Egyetem
Vegyészmérnöki és Biomérnöki Kar
Biomérnök MSc képzés

**Genomi uracilmintázatok jellemzése: szenzorfejlesztés és funkcionális elemzés
embrionális fejlődési modellben**

OTDK dolgozat

Szajkó Milda Blanka

Témavezető: **Dr. Békési Angéla**, adjunktus

BME VBK, Alkalmazott Biotechnológia és Élelmiszertudományi Tanszék

HUN-REN TTK, Molekuláris Élettudományi Intézet

Konzulens: **Holub Eszter**, doktorandusz

ELTE TTK, Biológia Doktori Iskola

HUN-REN TTK, Molekuláris Élettudományi Intézet

Alkalmazott Biotechnológia és Élelmiszertudományi Tanszék

Genom Metabolizmus és Biostruct Kutatócsoport

Lezárás dátuma: 2024. 12. 30.



Tartalomjegyzék

Köszönetnyilvánítás	4
Rövidítések jegyzéke.....	5
1. Bevezetés.....	7
2. Szakirodalmi áttekintés	8
2.1. Uracil megjelenése a DNS-ben	8
2.1.1. Citozin-dezamináció.....	8
2.1.2. Timint helyettesítő beépülés.....	9
2.2. A DNS-beli uracil javítása	10
2.2.1. Uracil-DNS-glikozilázok mint szenzorok	11
2.3. Kromatinszerveződés és szabályozó elemek.....	12
2.3.1. Hisztonmódosítás	13
2.3.2. DNS-metiláció.....	14
2.4. DNS-szekvenálási technikák az epigenom vizsgálatára	14
2.4.1. Genom-hozzáférhetőség térképezésére használt technikák	15
2.4.2. Antitestalapú technikák	16
2.4.3. Uracilosodási mintázat meghatározása	17
2.5. A zebraahal mint modellorganizmus	18
2.5.1. A zebraahal genom.....	18
2.5.2. A zebraahal korai embrionális fejlődése	18
3. Célkitűzés	21
4. Anyagok és módszerek.....	22
4.1. A funkcionális genomikai adatok bioinformatikai elemzése	22
4.1.1. A zebraahal genomi uracilosodási mintázatának meghatározása	22
4.1.2. Adatbázisból származó fájlok és feldolgozásuk.....	22
4.1.3. Genomszegmentációs analízis.....	24
4.1.4. Genomi mintázatok hasonlósági keresése a GIGGLE programcsomaggal.....	26
4.2. A szenzorfejlesztéshez használt kiindulási konstruktok	27
4.3. Konstruktok előállítása molekuláris klónozással	28
4.3.1. Polimeráz láncreakció	28
4.3.2. KLD reakció	29
4.3.3. Transzformálás XL1Blue törzsbe.....	29
4.4. Kolónia PCR és a klónok felszaporítása	30
4.5. Fehérjetermelés	30
4.5.1. Transzformálás BL21 sejtekbe	30

4.5.2.	Fehérjeexpresszió	31
4.6.	Fehérjetisztítás.....	31
4.6.1.	Sejtek feltárása	31
4.6.2.	Biotinilálás	32
4.6.3.	Tisztítás oszlopon	32
4.6.4.	Dialízis	34
4.6.5.	Töményítés	34
4.7.	Gélelektroforézis	34
4.7.1.	Agaróz gélelektroforézis	35
4.7.2.	SDS-PAGE.....	35
4.7.3.	Elektroforetikus mobilitás eltolódás esszé	36
4.8.	Western blot	37
4.9.	U-DNS-t detektáló csepp blot	38
5.	Eredmények és értékelésük	39
5.1.	Genomi uracilmintázat bioinformatikai elemzése korai zebrahal embriókban.....	39
5.1.1.	Szatelliták átfedése az uracilos régiókkal.....	39
5.1.2.	Uracileloszlási és funkcionális genomikai profilok genomszegmentációs analízise	41
5.1.3.	Az uracilosodott régiókhoz hasonló genomi mintázatok keresése GIGGLE programmal	42
5.1.4.	Az összehasonlító bioinformatikai elemzés eredményeinek értelmezése	43
5.2.	Uracil-DNS-t felismerő szenzorok fejlesztése	46
5.2.1.	A kiindulási konstruktok C'-terminálisának bővítése AviTaggel.....	46
5.2.2.	A konstruktok próbatisztítása és <i>in vitro</i> biotinilálhatóságának tesztelése	47
5.2.3.	A His-3xFLAG-ΔUNG-Avi konstruktt termelése és tisztítása	49
5.2.4.	A His-1xFLAG-UdgX-Avi konstruktt termelése és tisztítása.....	52
5.2.5.	A szenzorfehérjék alkalmazhatóságának tesztelése csepp bloton	57
5.2.6.	A szenzorfehérjék előállításának és jellemzésének tanulságai.....	60
6.	Összegzés	62
	Irodalomjegyzék.....	64
	Függelék	69

Köszönetnyilvánítás

Köszönettel tartozom Prof. Dr. Vértessy G. Beátának, hogy csatlakozhattam a kutatócsoportjához, és így egy összetartó, inspiráló közösség tagja lehetek.

Köszönöm a témavezetőmnek, Dr. Békési Angélának, hogy a kezdetektől fogva hitt bennem, motivált a fejlődésre, és mindig számíthattam rá. Hálás vagyok a türelméért, szakmai és emberi példamutatásáért.

Köszönöm a konzulensemnek, Holub Eszternek a rengeteg segítséget, lelki és szakmai támogatást.

Köszönöm az egész Genom Metabolizmus és Biostruct Kutatócsoportnak, hogy befogadtak, figyeltek rám, és bárkihez fordulhattam a kérdéseimmel.

Köszönöm a családomnak, hogy érdeklődéssel és támogatással fordultak a tanulmányaim és a kutatásom felé, és köszönöm a barátaimnak, hogy kitartottak mellettem.

A jelen munka részét képezi az Egyetemi Kutatói Ösztöndíj Program keretében megvalósuló kutatásomnak (EKÖP-24-2-BME-216), melyet a Kulturális és Innovációs Minisztérium, és a Nemzeti Kutatási, Fejlesztési és Innovációs Alap támogat.

Rövidítések jegyzéke

1. táblázat: Az alkalmazott rövidítések jegyzéke.

Rövidítés	Jelentés
5FdUR	5-fluoro-2'-dezoxiuridin
AID	aktiváció indukált citozin-dezamináz
APOBEC	apolipoprotein B mRNS-t editáló katalitikus polipeptid
ATAC-seq	transzpozázhoz hozzáférhető kromatin vizsgálata szekvenálással
BA	benzamidin
BER	báziskivágásos javítómechanizmus
CA	klóramfenikol
CAPS	N-ciklohexil-3-amino-propánszulfonsav
CAR	karbenicillin
ChIP-seq	kromatin immunprecipitációhoz kapcsolt szekvenálás
DNS	dezoxiribonukleinsav
dTTP	dezoxi-timidin-trifoszfát
dUTP	dezoxi-uridin-trifoszfát
EDTA	etilén-diamin-tetraecetsav
GEO	Gene Expression Omnibus adatbázis
HEPES	2-(4-(2-hidroxietil) -1-piperazinil)-etánszulfonsav
hpf	megtermékenyítés után eltelt órák (hours post fertilization)
HRP	tormaperoxidáz
IP	immunprecipitáció
IPTG	izopropil- β -d-1-tiogalaktopiranozid
LB tápoldat	Luria-Bertani tápoldat
MeDIP-seq	Metilált DNS immunprecipitációs szekvenálás
MMR	hibaspár-javítás
MNáz	mikrokokkális nukleáz
MNase-seq	mikrokokkális nukleáz emésztés mély szekvenálással
NER	nukleotidkivágó javítás

Ni-NTA	nikkel-nitrilo-triecetsav
NT	nem kezelt (non-treated)
OD	optikai denzitás
ON	éjszakán át tartó (overnight)
PBS	izotóniás foszfát puffer
PCR	polimeráz láncreakció
PMSF	fenil-metil-szulfonil-fluorid
SDS	nátrium-dodecil-szulfát
SMUG1	egyszálú DNS-specifikus egyfunkciós U-DNS-glikoziláz
TBS	TRIS-pufferolt sóoldat
TC	tetraciklin
TRIS	trisz-(hidroximetil)-amino-metán
TS	timidilát-szintáz
Tween-20	polioxietilén(20)-szorbitán-laurát
U	uracil
UDG	uracil-DNS-glikoziláz
U-DNA-Seq	U-DNS-immunprecipitációhoz kapcsolt szekvenálás
U-DNS	DNS-beli uracil
UNG	uracil-DNS-N-glikoziláz
β-ME	β-merkaptoetanol

1. Bevezetés

Az uracil a ribonukleinsavat (RNS) felépítő nukleotidok négyféle bázisának egyike. A dezoxiribonukleinsavban (DNS) uracil helyett timin bázisok képeznek specifikus párt az adeninekkel. Az uracil bázis megjelenhet a DNS-molekulában is, azonban előfordulása erősen korlátozott. A DNS-ben az uracilt a sejtek alapvetően hibaként értelmezik, és báziskivágásos mechanizmussal hatékonyan javítják, melynek első enzimei az uracil-DNS-glikozilázok [1]. A sejtek nukleotid-anyagcseréjének zavarása, különösen a timidilát-bioszintézis gátlása régóta alkalmazott rákterápiás stratégia. A timidilát-bioszintézis gátlásával megnövekszik sejtbeli dezoxi-uridin-trifoszfátnak (dUTP) a dezoxi-timidin-trifoszfáthoz (dTTP) viszonyított aránya, amely nagymértékű timint helyettesítő uracilbeépülést okoz, mely erősen citosztatikus hatású.

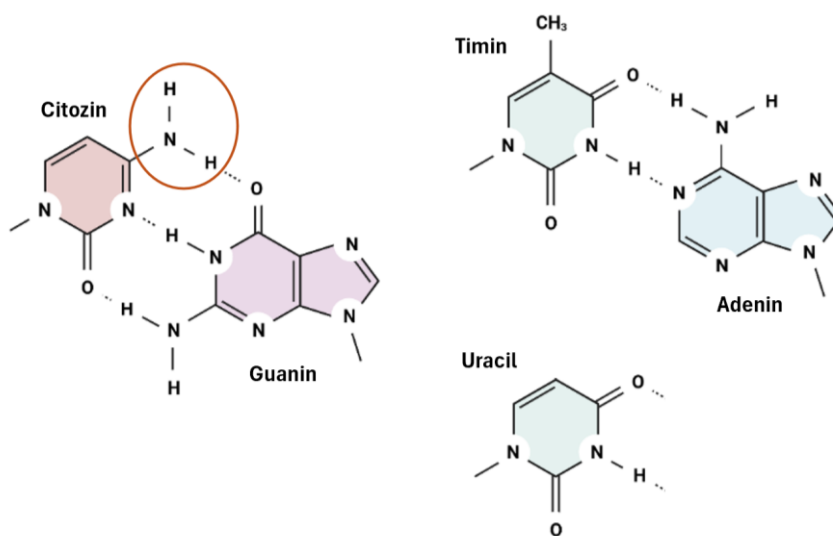
Ugyanakkor számos kutatás alátámasztja, hogy a DNS-beli uracilnak lehet biológiai szerepe, akár epigenetikai tényezőnek is tekinthető bizonyos helyzetekben. Például olyan különböző biológiai folyamatokban is szerepet játszik, mint az immunitás [2] vagy a *Drosophila* metamorfózisa során zajló programozott sejthalál [3]. Emellett a kutatócsoportunk (Genom Metabolizmus és Biostruct Kutatócsoport) legfrissebb eredményei szerint a zebrahal korai embrionális stádiumaiban szintén magas a dUTP/dTTP arány, és emelkedett a DNS-beli uracil szintje is. Ennek funkcionális jelentőségét alátámasztja, hogy amennyiben ezt dUTP-t bontó enzim mikroinjektálásával megzavarjuk, az komoly fejlődési zavarokat okoz. A csoport egy saját fejlesztésű uracil-DNS-immunprecipitációhoz kapcsolt szekvenálással (U-DNA-seq, [4]) meghatározta a genomi uracilosodási mintázatot is 2,5 órás zebrahal embriókban. Ennek funkcionális genomikai adatokkal való összehasonlítása révén közelebb kerülhetünk az uracil egyedfejlődésben betöltött szerepének megértéséhez, hozzájárulva annak biológiai jelentőségének feltárásához.

A DNS-ben található uracil vizsgálatához elengedhetetlen az azt minél hatékonyabban felismerő szenzorok alkalmazása. Laborunkban jelenleg is használnak uracil-DNS-glikoziláz alapú szenzorokat, azonban ezek érzékenysége és stabilitása – különösen a nehezen izolálható, komplex mátrixból kinyert és alacsony uraciltartalmú minták esetében – nem mindig megfelelő. Így ezeknek a továbbfejlesztése fontos és szükséges törekvés afelé, hogy olyan eszközök álljanak rendelkezésünkre, amelyek alkalmasak a kicsi, ám sok esetben igen fontos különbségek kimutatására.

2. Szakirodalmi áttekintés

2.1. Uracil megjelenése a DNS-ben

A DNS-t felépítő nukleotidokban a négy kanonikus bázis (a purinvázis adenin és guanin, valamint a pirimidinvázis citozin és timin) mellett előfordulhatnak nem kanonikus bázisok is. Ezek közül az uracil az egyik leggyakoribb [5], amely a citozinhoz és a timinhez hasonlóan pirimidinvázis, utóbbtól csak egy metilcsoportban különbözik (1. ábra).



1. ábra: A DNS-ben előforduló kanonikus bázisok és az uracil szerkezeti képlete (forrás: BioRender). Az ábrán a négy kanonikus bázis látható a bázispárosodás szabályainak megfelelően. Az uracil a timintől csak egy metilcsoportban különbözik, emiatt a timinhez hasonlóan az adeninnel párosodik. A citozin oxidatív dezaminációjával (az aminoscsoport narancssárga körrel jelölve) uracil keletkezik.

Az uracil megjelenhet a DNS-ben a citozin dezaminációjának eredményeként, vagy beépülhet a szintézis során a timin helyett [6].

2.1.1. Citozin-dezamináció

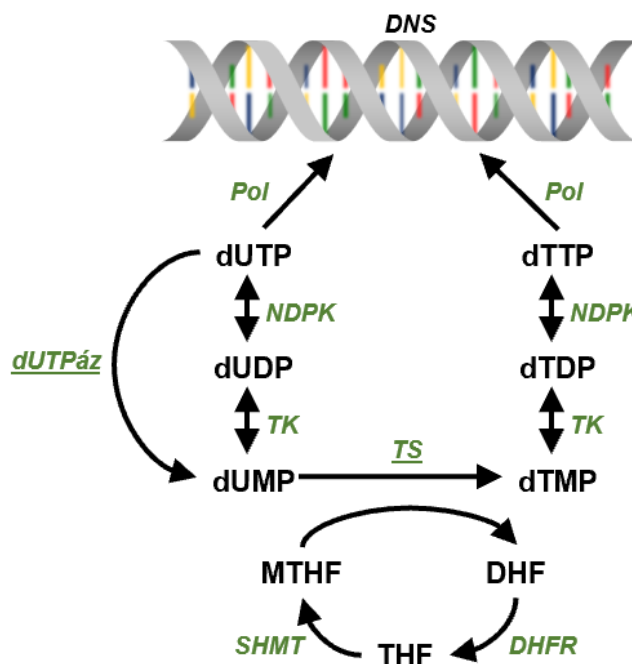
A citozin dezaminációjának során keletkező uracil javítás nélkül mutációt eredményezhet. Ha a kódoló szálon történik a dezamináció, akkor az a következő replikáció során C > T mutációt okoz, ha a nem kódoló szálon keletkezik az uracil, akkor az a kódoló szálon G > A átmenetként jelenik meg [7].

A citozin dezaminációjának történhet spontán módon, melyet a sejtekben levő víz [8], vagy bizonyos vegyszerek, például a biszulfid [9] iniciálhatnak. Emellett a folyamatot katalizálhatják enzimek, a DNS-citozin-dezaminázok is. Ilyen enzimek az AID/APOBEC család tagjai, melyek közül az APOBEC3 enzimek fontos szerepet játszanak a veleszületett immunitásban. A vírusokfertőzésekre válaszul citokinek termelődnek, melyek hatására az APOBEC3 enzimek

indukálódnak, és a vírusok genetikai anyagát megváltoztatva vesznek részt a vírus replikációjának és terjedésének gátlásában [10]. A család másik tagja, az AID, mutációkat okoz az antitesteket kódoló gének variábilis és kapcsoló régióiban, így a szomatikus hipermutációs és osztályváltó rekombinációs folyamatokon keresztül jelentős szerepet játszik az antitestek diverzifikációjában [2].

2.1.2. Timint helyettesítő beépülés

A citozin-dezamináció mellett a másik lehetőség az uracil megjelenésére a DNS-ben a timin helyetti beépülés, ami az előbbivel ellentétben nem vezet mutációhoz, azonban több szempontból is megzavarhatja a DNS-metabolizmust [11]. A legtöbb polimeráz nem képes különbséget tenni a dezoxiuridin és a dezoxtimidin között, ezért a sejtek a celluláris dUTP/dTTP arány alacsonyan tartásával igyekeznek elkerülni az uracil beépítését a DNS-be [12]. Ezért elsősorban a timidilát-bioszintézis (2. ábra) egyik kulcsenzime, a dUTPáz felelős, amely katalizálja a dUTP bomlását dezoxi-uridin-monofoszfátra (dUMP) és pirofoszfátra. Eztáltal egyrészt szabályozza a dUTP-szintet, másrészt pedig a keletkező dUMP prekursora a dezoxi-timidin-monofoszfát (dTMP) szintézisének, melyet a timidilát-szintáz katalizál [13]. A dTMP-t több lépésben kinázok alakítják dTTP-vé.



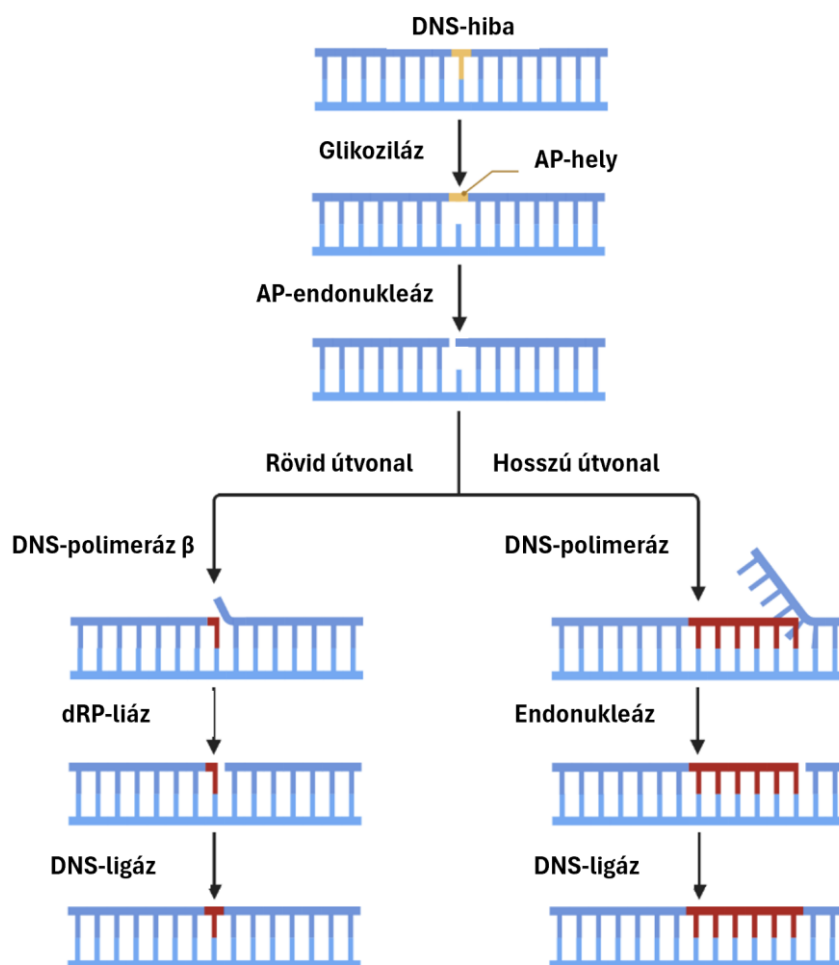
2. ábra: Timidilát-bioszintézis (az ábrát Holub Eszter készítette). A nyilakon az adott folyamatot katalizáló enzimek rövidítései találhatóak, a két kulcsenzim aláhúzással jelölt. Pol: DNS-polimeráz, NDPK: nukleozid-difoszfát-kináz, TK: timidin-kináz, dUTPáz: dezoxiuridin-trifoszfátáz, TS: timidilát-szintáz, DHFR: dihidrofolát-reduktáz, SHMT: szerin-hidroxitimetiltranszferáz. A nyilak végén az enzimek által katalizált folyamatok kiindulási és termék vegyületei láthatók. dUTP: dezoxiuridin-trifoszfát, dUDP: dezoxiuridin-difoszfát, dUMP: dezoxiuridin-monofoszfát, dTMP: dezoxtimidin-monofoszfát, dTDP: dezoxtimidin-difoszfát, dTTP: dezoxtimidin-trifoszfát, MTHF: 5,10-metiléntetrahidrofolát, DHF: dihidrofolát, THF: tetrahidrofolát.

Amennyiben a timidilát-bioszintézis enzimeit nem működnek megfelelően a sejtbeli dUTP/dTTP arány megnő, így több uracil épül be a DNS-be. Ennek hatására megváltozhat a DNS-kötő fehérjék, pl. transzkripciós faktorok affinitása a DNS-hez [11] [14], valamint az uracilt eltávolító hibajavító mechanizmusok túlhajtott működése sejtciklus-feltartóztatást vagy sejthalált eredményezhet [15]. Emiatt a timidilát-bioszintézis enzimeit gátló gyógyszereket régóta alkalmazzák rákterápiára, hiszen a sejtosztódáshoz szükséges DNS-szintézis során fejtik ki hatásukat. A rákos sejtek intenzíven és folyamatosan osztódnak, ezért különösen érzékenyek az ilyen jellegű gátlásra [16].

2.2. A DNS-beli uracil javítása

A DNS-beli uracil javítása elsősorban báziskivágásos javítómechanizmussal (base excision repair, BER) történik (3. ábra) [1]. A folyamatot az uracil-DNS-glikozilázok (UDG) indítják el, amelyek az uracil és a dezoxiribóz közötti N-glikozidos kötést hasítják el, így abázisos helyet (AP-hely) eredményezve a DNS-szálon [17]. Ezt követően az AP-endonukleáz hasítja a DNS-t az AP-hely 5'-végénél, majd a DNS-polimeráz újraszintetizálja a hiányzó szakaszt, végül pedig a ligáz ismét összeköti a DNS-szálat. A ligálási lépés előtt a túllógó DNS-szakaszt egy nukleáz távolítja el [18]. A BER kétféle útvonalon haladhat: rövid útvonal (short path, SP) és hosszú útvonal (long patch, LP) [17] [18]. A rövid útvonalban a β -polimeráz egyetlen nukleotidot épít be, a hosszú útvonal összetettebb, és több fehérje együtműködésével egy 2-10 oligonukleotid hosszú szakasz szintetizálódik [19].

A BER mellett a citozinból keletkezett uracil eltávolítása elvileg lehetséges hibáspárjavító-mechanizmussal (mismatch repair, MMR) is [20]. Ebben az útvonalban a nem kanonikus bázispárokat ismeri fel egy MutS homológ enzim, amely a MutL homológ közreműködésével megindítja a hibás nukleotid cseréjét egy kivágó javítási mechanizmus segítségével [21]. Az MMR során a „hibás nukleotid” cseréje hosszabb DNS-szakaszok lebontásával és újraszintetizálásával jár, amire jellemzően a sejtciklus S-fázisában, a DNS-szintézishez kapcsolódóan van lehetőség a nagy processzivitású DNS-polimerázok jelenlétében [22].



3. ábra: A báziskivágásos javítómechanizmus sematikus ábrája (a BioRender módosított ábrája). A hibás bázist a glikoziláz eltávolítja abázisos helyet eredményezve. Az abázisos helyen az AP-endonukleáz hasítja a DNS gerincét, majd a DNS-polimeráz újraszintetizálja a hiányzó szakaszt. Egy nukleáz eltávolítja a túllógó DNS-szakaszt, majd a ligáz összeköti a DNS-szálat. A rövid útvonalon csak egy, a hosszú útvonalon több nukleotid is beépül.

2.2.1. Uracil-DNS-glikozilázok mint szenzorok

Amint ezt a csoportunkban felismerték, az uracil-DNS-glikozilázok természetes vagy módosított változatai szenzorként is alkalmazhatók az uracil detektálására a DNS-ben. A szenzorként használt UDG fehérjék képesek nagy specificitással megkötni az uracilos DNS-t, lehetővé téve a DNS-beli uracil vizsgálatát többféle kísérleti elrendezésben is: globális mennyiségi meghatározás – csepp blot, in situ lokalizáció – immuncitokémia, genomi mintázatok – DNS-immunprecipitációhoz kapcsolt szekvenálás. Az ilyen szenzorok fontos eszközök az uracil jelenlétének, keletkezésének és biológiai szerepének kutatásában.

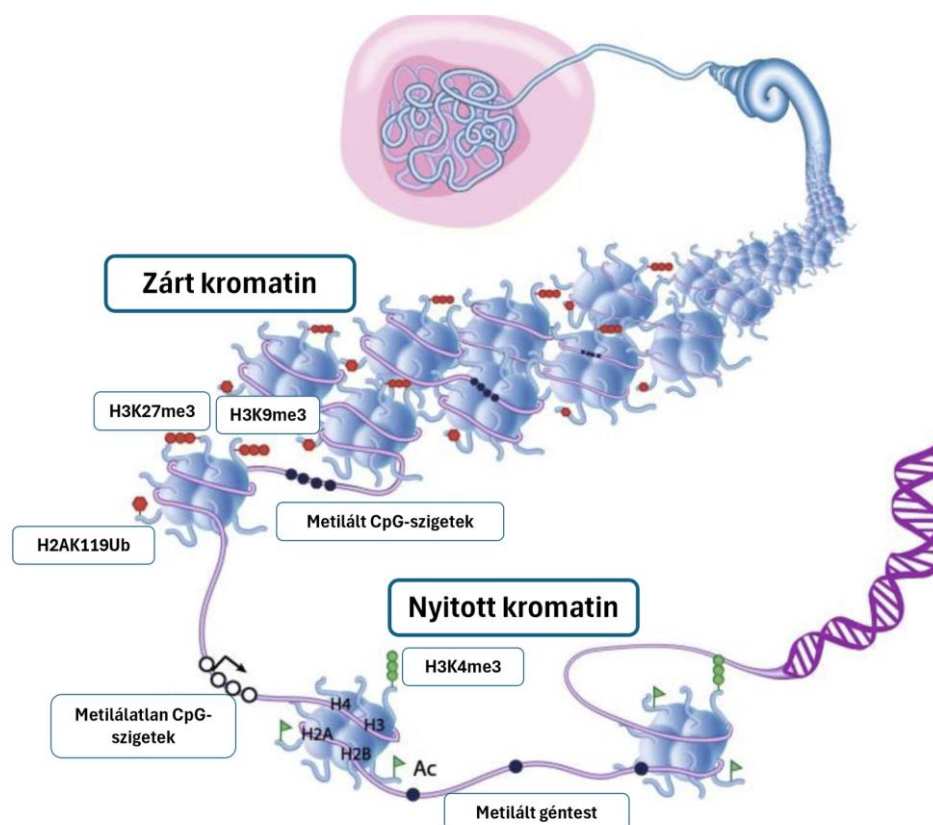
Az egyik ilyen szenzorként alkalmazott fehérje a *Mycobacterium smegmatis*-ből származó UdgX, amely denaturáló körülmények között is kötődik az uracilt tartalmazó DNS-hez [23]. Az UdgX specifikusan felismeri az uracilt, kivágja azt, majd a H109 hisztidinen keresztül kovalensen kötődik az abázisos helyhez, miközben az uracil az aktív hely zsebében marad [24].

A kutatások azt mutatják, hogy az UdgX előnyben részesíti az egyszálú DNS-ben található uracilt a kettős szálú DNS-hez képest, és az egyszálú, uraciltartalmú oligomereket képes 1:1 arányban kvantitatívan kovalens komplexben kötni. Az imidazol viszont inhíbeálhatja a fehérjét: 250 mM imidazolt tartalmazó kezelés már jelentős aktivitáscsökkenést eredményez. Az UdgX képes lokalizációs szígnál nélkül bejutni a sejtmagba, így akár *in vivo* uraciljelölő reagensként is alkalmazható [6].

Az UNG (uracil-N-glikoziláz) enzimek jól konzervált uracil-DNS-glikozilázok, amelyeket számos organizmusban azonosítottak, beleértve például az embert és az *Escherichia coli*-t is [25]. Az UNG a BER egyik iniciáló DNS-glikozilázaként specifikusan felismeri a DNS-beli uracilt és kivágja. Az UNG aktivitásában kulcsszerepet játszó aminosavak mutációjával létrehozható olyan katalitikusan inaktív fehérje, amely felismeri az uracil-DNS-t, de nem képes kivágni. Az emberben ismert UDG-k közül az egyik legfontosabb az UNG magi izoformája (UNG2), melynek a D154N és H277N pontmutációi [26] ilyen katalitikusan inaktív fehérjét eredményeznek, amit sikeren alkalmaztak az uracil csepp blot alapú *in vitro*, valamint immuncitokémiai megközelítésű *in situ* vizsgálatára is. A pontmutációkon kívül ezekben a kísérletekben használt ΔUNG szenzorban az eredeti fehérje N'-terminálisától számított első 84 aminosavból álló peptidszakaszt eltávolították, csökkentve ezáltal más fehérjékkel, többek között a replikációs fehérjékkel való kölcsönhatást [27].

2.3. Kromatinszerveződés és szabályozó elemek

Az eukarióták sejtmagjában található DNS-t és fehérjét tartalmazó struktúra a kromatin, amely segít a genetikai információ rendezésében és szabályozásában. A kromatin nukleoszómak tömbjéből áll, amelyet körülbelül 146 bázispár kettős szálú DNS alkot hisztonfehérjék oktamereire tekeredve [28]. A kromatin több különböző állapotban létezhet, amelyeket különböző epigenetikai módosítások kombinációi határoznak meg. A kromatin epigenetikai állapota számos mechanizmussal szabályozható, beleértve a DNS és a hisztonfehérjék kémiai módosítását. Az eukromatinban a génátírás aktív, míg a heterokromatinban – a nehezen hozzáférhető szerkezete miatt – génexpresszió tipikusan nem zajlik. A fakultatív heterokromatinban az eukromatikus és heterokromatikus állapot közötti átmenet lehetséges, melyet a kromatin szerveződésében szerepet játszó hisztonfehérjék aktuális módosításai és a nukleoszómak átrendeződése szabályoz (4. ábra) [29][30].



4. ábra: Az ábra a kromatin szerveződését és a hozzá kapcsolódó epigenetikai módosításokat mutatja be, amelyek befolyásolják a génexpresszió szabályozását. Két fő kromatin állapot látható: zárt kromatin és nyitott kromatin. A zárt kromatin sűrű szerkezetű, ahol a H3K27me3, H3K9me3 és promóterekben található metilált CpG-szigetek gátolják a génexpressziót. A nyitott kromatin lazább, a H3K4me3 és a hiszton-acetiláció a génaktivitást segítik elő [31].

2.3.1. Hisztonmódosítás

A maghisztonok (H2A, H2B, H3 és H4) központi globuláris doménje a többi hisztonnal való kölcsönhatásban és a DNS nukleoszóma köré tekerésében vesz részt, a rendezetlen N'-terminális farokdomén pedig poszttranszlációs módosításokra alkalmas aminosavakat, például lizint, arginint és szerint tartalmaz nagy mennyiségben. Ezeknek az aminosavaknak a specifikus pozíciójában végbemenő módosítások a sejtek számára fontos jelként szolgálnak, mivel befolyásolják a kromatin szerkezetét és az adott régió transzkripciós aktivitását.

Az egyik típusú hisztonmódosítás az acetiláció, amit számos sejttypusban összefüggésbe hoztak a transzkripcionálisan aktív kromatinnal. Az acetiláció a farokdomén specifikus pozícióban található lizinjein mehet végbe. Semlegesíti a lizin pozitív töltését, ami hozzájárulhat a transzkripciós aktiváláshoz a kromatin szerkezetének relaxációja révén.

A hisztonok az acetiláción túl módosíthatók számos más módon is, például a lizin és arginin metilációjával vagy a szerin foszforilációjával is, amelyek szintén a farokdomének specifikus aminosavain mennek végbe, és összefüggésben állnak a transzkripciós aktivitás változásaival. Amellett, hogy ezek a módosítások befolyásolják a kromatinszerkezet lazaságát, kötőhelyet is

jelenthetnek bizonyos transzkripciót támogató vagy éppen represszálo fehérjéknek. Például a H3 hiszton egyes aminosavainak metilálása (pl. H3K4me3 = H3 4. lizinjének trimetilált formája, H3R17me), acetilációja (pl. H3K9ac, H3K14ac, H3K27ac) és foszforilációja (pl. H3S10p) mind hozzájárulnak a transzkripció aktivációjához. Ugyanakkor a H3K9me3 és H3K27me3 módosítások represszióval és kromatinkondenzációval társulnak. A promóterek és enhanszer régiók környékén megtalálható hisztonmódosítások különböznek. Például a promóterekre a H3K4me3, míg az enhanszerekre a monometilált forma (H3K4me1) jellemző [30].

2.3.2. DNS-metiláció

Nemcsak a hisztonok, hanem a DNS is metilálható, és ennek szintén szerepe van a kromatinsztruktúra és a transzkripció szabályozásában, valamint részt vesz a rekombinációban, a replikációban, a transzpozonok szabályozásában, az imprinting létrehozásában és az X-kromoszóma inaktivációjában is. Eukarióta DNS-ben a leggyakoribb metilált bázismódosulat az 5-metilcitozin. Emellett ritkábban előfordul a 6-metiladenin is, amit eredetileg csak prokariótákra jellemző módosításnak gondoltak. A citozin metiláltsága leginkább a CpG-dinukleotidokra jellemző. A CpG-dinukleotidok globálisan viszonylag ritkák, a gének promóter és regulátor régióiban található CpG-szigeteken kívül leginkább az intronokban vagy intergenikus régiókban, gyakran repetitív szakaszokban, transzpozonokban találhatók [32]. A fiatal sejtek genomjában található CpG-k többsége citozin-metilációval rendelkezik. A promóterekben levő CpG-k metilációja transzkripció repressziójához vezet, ennek megfelelően erősen expresszáldó gének promóterei nem tartalmazzak DNS-metilációt. A fejlődés során a DNS-metilációnak fontos szerepe van bizonyos gének csendesítésében, amelyek kifejeződésére már nincs szükség [33]. Emellett a DNS mutációs rátájának megváltoztatása révén evolúciós jelentőséggel is bír. Az 5-metilcitozin könnyebben dezaminálódik, mint a normál citozin. Ráadásul a metilált citozinból nem uracil, hanem timin keletkezik, aminek a pontos kijavítása kevésbé valószínű. Emiatt a csírvonalban erősen metilált szekvenciák CpG-hiányossá válnak, miközben a TpG-k és CpA-k száma növekszik [34].

2.4. DNS-szekvenálási technikák az epigenom vizsgálatára

Az epigenom vizsgálatára többféle módszer létezik, melyek specifikussága eltérő. A legspecifikusabb eljárások egy adott hiszton- vagy DNS-módosítás, valamint transzkripció faktor elhelyezkedését és gyakoriságát térképezik fel antitestalapú módszerekkel. Egyes esetekben hasznosabb egy átfogóbb képet kapni a génszabályozási környezetről, különösen ha

egy megfigyelt jelenség mögött álló epigenetikai változások pontos jellege ismeretlen. Napjainkban már számos alternatív módszer áll rendelkezésre, amelyek a génszabályozási környezet szélesebb vizsgálatát teszik lehetővé [29]. Ezeknek a lényeges elemei az új generációs szekvenálási (NGS) technikák. Az NGS módszerek megjelenése jelentősen felgyorsította a genomikai kutatásokat, mivel lehetővé teszik a gyors és költséghatékony szekvenálást, akár párhuzamosan több millió DNS-fragmens esetében is, ami nagyban elősegíti a genetikai variációk, mutációk és betegségek molekuláris alapjainak azonosítását. Az NGS különböző alkalmazásai lehetővé teszik a genomszintű profilozást, ami a hagyományos technikákkal nehezen kivitelezhető lenne. Így epigenetikai elemzésekre is kiválóan alkalmazhatók, például a DNS-metilációs mintázatok és a kromatin átalakulások vizsgálatára, amelyek kulcsfontosságúak a génexpresszió szabályozásában. A legújabb valós idejű és egymolekulás NGS módszerek elviekben alkalmasak lehetnek a különböző bázismódosulatok közvetlen detektálására is, de ezek az eljárások még kidolgozás alatt vannak. Ugyanakkor a sokkal olcsóbb és elterjedtebb rövid olvasatos technológiákra alapuló eljárások széles választéka rendelkezésre áll, amelyekből a továbbiakban csak a munkám során előfordulókat fogom bemutatni.

2.4.1. Genom-hozzáférhetőség térképezésére használt technikák

A kromatinon belüli szabályozó elemek elsősorban a genom hozzáférhető régióiban találhatók. A nyitott kromatin kimutatása segíthet a DNS-sel kölcsönhatásba lépő fehérjék kötőhelyeinek azonosításában is [35].

A nukleoszómák elhelyezkedése a DNS mentén alapvetően befolyásolja a génexpressziót. A nukleoszómák ellenállóak a mikrokokkális nukleáz (MNáz) végzett emésztéssel szemben. Ezt használja ki az MNase-seq technika, mely során MNáz enzimes kezelést követően a megmaradt DNS-t izolálják és szekvenálják (5. ábra, c). Mivel a legtöbb modellszervezet genomi szekvenciája már ismert, így a nukleoszómák helyzetének meghatározásához elegendő a nukleoszómális DNS-fragmens két végének rövid szekvenálása, amire a rövid olvasatos, páros végű NGS technikák (Illumina paired-end sequencing) kiválóan alkalmasak [36].

Szintén a kromatin átfogó profilozására alkalmas módszer a transzpozáz-érzékeny kromatin szekvenálás (Assay for Transposase-Accessible Chromatin using sequencing, ATAC-seq), mely az MNase-seq által készített genomszintű szabályozási térképekhez nagyon hasonló végeredményt szolgáltat, azonban gyorsabb és egyszerűbb vizsgálatot tesz lehetővé, és kevesebb kiindulási mintát igényel [29]. Az eljárás során egy módosított, hiperaktív Tn5 transzpozázt használnak, ami a megnövekedett hozzáférhetőségű régiókban egyidejűleg vágja

el a DNS-t és kapcsolja hozzá a szekvenáláshoz szükséges adaptereket (5. ábra, d). Az adaptertől adapterig PCR-ezett fragmenseket páros-végű rövid olvasatos szekvenálásnak vetik alá. Így a szekvenálási olvasatok alapján azonosíthatók a kromatin hozzáférhető régiói és a fragmenshosszak periodicitása alapján térképezhetők a nukleoszómák pozíciói [37].

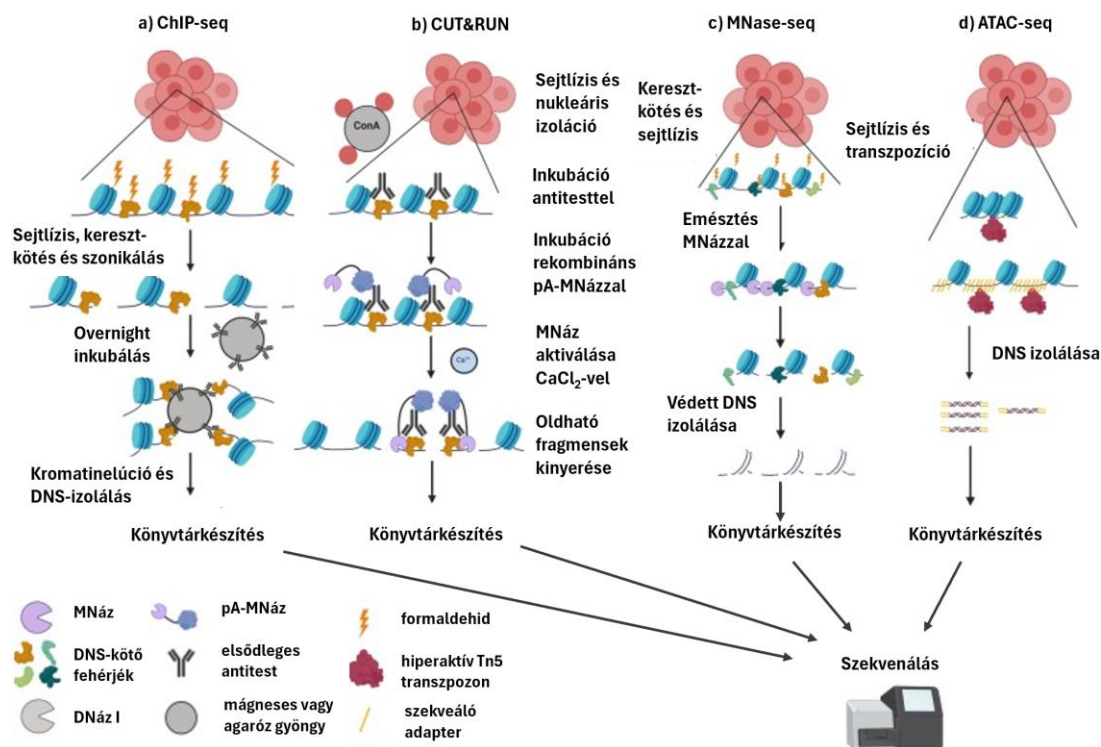
2.4.2. Antitestalapú technikák

A hisztonfehérjék és más DNS-kötő fehérjék aktuális genomiális elhelyezkedése nem határozható meg csupán a DNS-szekvencia és/vagy a hozzáférhetőség alapján, ezért egyedi fehérjekötődés-profilozási technológiákat alkalmaznak [35].

A kromatin immunprecipitációhoz kapcsolt szekvenálás (ChIP-seq) segítségével DNS-hez kötődő fehérjék kötőhelyeit lehet azonosítani. A módszer lényege, hogy a DNS-fehérje komplexeket keresztkötés segítségével stabilizálják, majd fragmentálják a kromatint. A kívánt fehérjére specifikus antitest hozzáadásával immunprecipitáció segítségével feldúsítják a fehérje által kötött fragmenseket. Ezután szétválasztják a fehérjéket a DNS-fragmensektől, és a tisztított DNS-szakaszokat szekvenálják NGS technikákkal (5. ábra, a) [38].

A DNS metilációs mintázatának a feltérképezésére alkalmas módszer a metilált DNS immunprecipitációján alapuló szekvenálás (MeDIP-seq). Az eljárás során a metilált citozinra specifikus antitestet használnak, és a ChIP-seq-hez hasonlóan immunprecipitáció segítségével dúsítják fel a metilált DNS-szakaszokat [39]. A különbség pusztán annyi, hogy nem kromatint, hanem DNS-t izolálnak – mivel nem fehérje-DNS kölcsönhatást vizsgál, keresztkötésre sincs szükség.

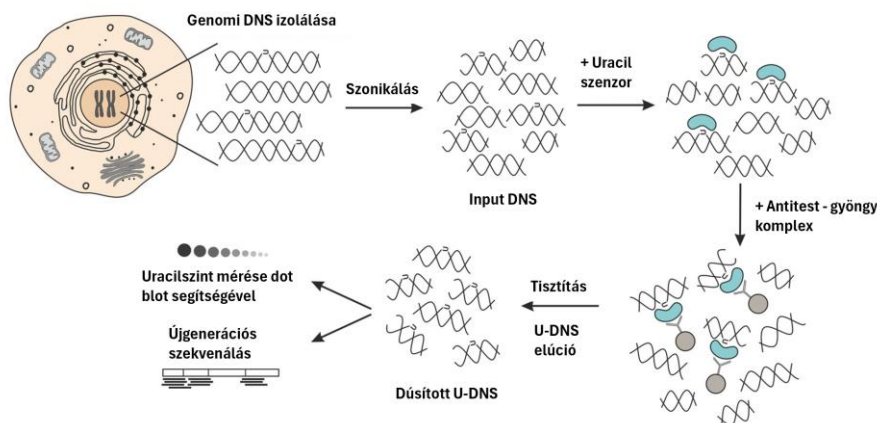
A CUT&RUN szekvenálás natív fehérje-DNS kölcsönhatások magas felbontású feltérképezésére alkalmazható. Az eljárás során a ChIP-seq-hez hasonlóan specifikus antitesttel ismerik fel a vizsgálandó fehérjét, azonban nem szükséges keresztkötés és szonikálás. Ehelyett egy rekombináns protein A-MNáz fúziót használnak. A protein A (pA) hozzáköt az antitesttel jelölt fehérjéhez, és az MNáz megemésztí a DNS-t a fehérje közvetlen környezetében. Az így kihasított DNS-fragmentumokat extrakció után szekvenálják (5. ábra, b). A CUT&RUN technika előnye, hogy a háttérjel gyakran alacsonyabb, mint a ChIP-seq eljárás során, mivel a pA-MNáz kromatinhoz való nem specifikus kötődése megfelelő mosás után viszonylag alacsony. Emiatt kisebb leolvasási lefedettséget igényel, és kevesebb kiindulási sejtből is elvégezhető [40].



5. ábra: DNS-szekvenálási technikák ([35] módosított ábrája). a) Kromatin immunprecipitációhoz kapcsolt szekvenálás (ChIP-seq), b) Antitesttel kijelölt célfehérjék mellett nukleázos emésztéssel kapcsolt szekvenálás (CUT&RUN), c) Mikrokokkális nukleáz emésztéssel kapcsolt szekvenálás (MNáz-seq), d) transzpozáz-érzékeny kromatin szekvenálás (ATAC-seq).

2.4.3. Uracilosodási mintázat meghatározása

A kutatócsoportunk által fejlesztett U-DNS-immunprecipitációhoz kapcsolt szekvenálással (U-DNA-Seq) meghatározható a genomi uracilosodási mintázat. A módszer lényege, hogy a tisztított és szonikálással fragmentált DNS-mintához katalitikusan inaktív ΔUNG szenzort (ld. 2.2.1. fejezet) adnak, amely a DNS-beli uracilt felismeri és hozzáköt. A szenzorra specifikus antitest-gyöngy komplex hozzáadásával immunprecipitáció segítségével dúsítják az uraciltartalmú DNS-fragmenseket, majd ezeket szekvenálják NGS technikával (6. ábra) [4].



6. ábra: Az U-DNA-seq technológia lépései [4]. Az uracil-DNS felismerésére szolgáló ΔUNG szenzort a világoskék alakzat jelöli. Az antitest-gyöngy komplex szűrőként jelölt.

2.5. A zebrahal mint modellorganizmus

A zebrahal (*Danio rerio*) egy édesvízi hal, mely népszerű modellszervezet a gerincesek génfunkcióinak tanulmányozásában. Előnyös tulajdonsága, hogy fenntartása költséghatékony, magas megtermékenyítési aránnyal rendelkezik, és külső megtermékenyítéssel szaporodik, ami egyszerűvé teszi az *in vitro* megtermékenyítést is, emellett viszonylag rövid a generációs ideje. A zebrahal embriók ellenállóak, és képesek túlélni különböző eljárásokat közvetlenül a megtermékenyítés után, például a genetikai manipulációt vagy mikroinjektálást. Az embriók áttetszősége lehetőséget nyújt a belső struktúrák fejlődésének vizsgálatára, valamint a jelölt részecskék (sejtek, fehérjék, nanorészecskék stb.) mozgásának és biodisztribúciójának nyomon követésére valós időben [41].

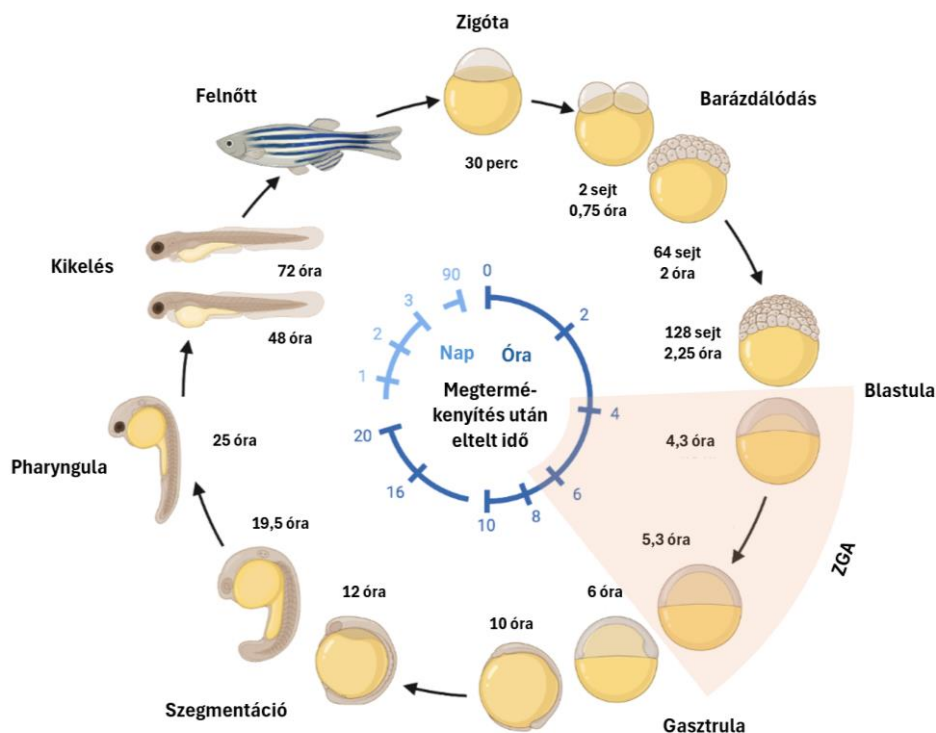
2.5.1. A zebrahal genom

Az emberi genetikai betegségek hatékony modellezéséhez elengedhetetlen a zebrahal gének és génstruktúrák megismerése, azok az emberitől való eltérésének azonosítása. A humán referenciagenommal való összehasonlítás azt mutatja, hogy az emberi gének körülbelül 70%-ának, és a betegségekkel összefüggő gének 82%-ának van zebrahalban megtalálható ortológja. A zebrahal genomjában kiemelkedően magas az ismétlődő szekvenciák aránya (kb. 50%) [42]. A szatelliták ismétlődő szekvenciaszakaszok, amelyek a legtöbb eukarióta kromoszóma heterokromatikus régióiban találhatók, és hozzájárulnak a kromoszómák fontos struktúráihoz, a heterokromatin kialakulásához és a genom stabilitásához [43]. A zebrahalban a szatellita régiók a jelentős részét a MOSAT típusú teszi ki, ami egyenletesen oszlik el az összes kromoszómán, de funkciójáról nincs konkrét tudásunk. Ezzel szemben a BRSATI néhány klaszterben található meg kromoszómánként, és a legkiterjedtebb klaszter a feltételezett centromerikus pozíció közelében helyezkedik el [42].

2.5.2. A zebrahal korai embrionális fejlődése

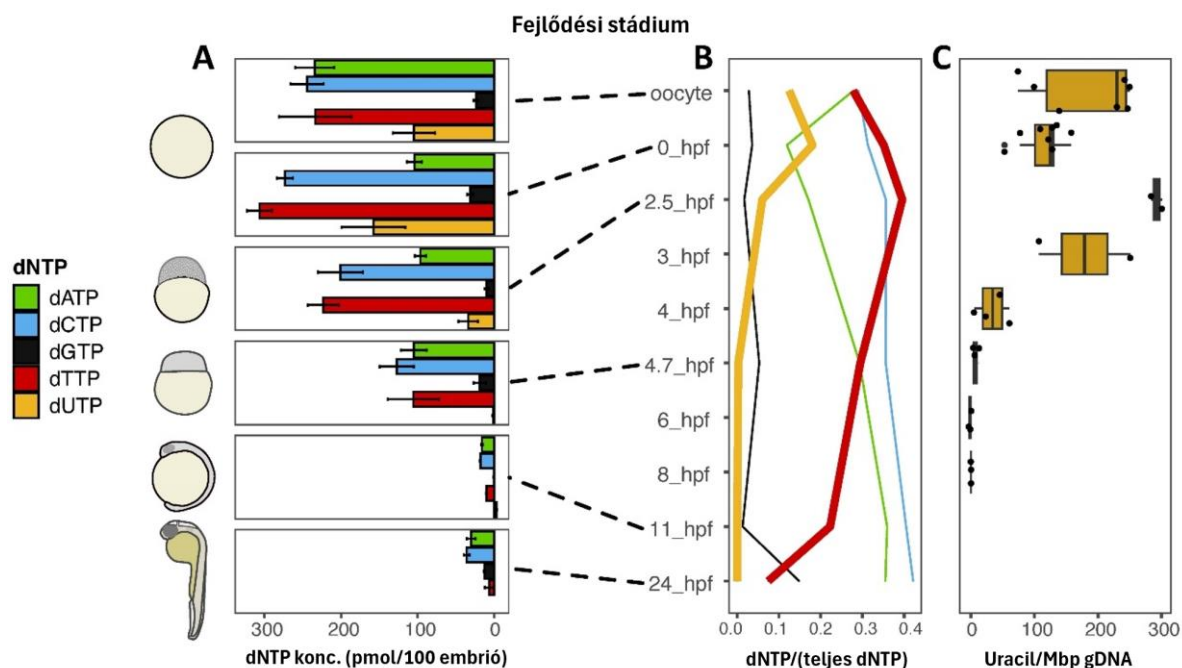
A zebrahal embrionális fejlődésének főbb fázisait az 7. ábra szemlélteti. Az első sejtosztódás a megtermékenyítést követően 45 perccel történik, ezután az osztódások szinkronban zajlanak 15 percenként az első 3 órában. Az első hasadás előtt eltelt viszonylag hosszú idő kísérleti szempontból igen előnyös, mivel lehetőséget teremt a különféle genetikai manipulációkra és kezelésekre. A megtermékenyítést követően 2,25 órával kezdődik a blastula periódus, ebben a fázisban 3 óra után megkezdődnek az aszinkron sejtosztódások és a zigotikus génaktiváció (ZGA) is kezdetét veszi. Ezt a szakaszt anyai-zigotikus átmenetnek (Maternal to Zygotic Transition, MZT) is nevezik. A ZGA előtt az embrió az anyától származó fehérjék és mRNS-

ek segítségével működteti életfolyamatait. Az MZT során az zigotikus gének transzkripciója fokozódik, az embrionális pluripotens sejtek elindulnak a differenciálódás felé [44].



7. ábra: A zebrahal embrionális fejlődése (forrás: BioRender). A külső körben a zebrahal főbb egyedfejlődési stádiumainak sematikus ábrái láthatók. Ezek nagyobb szakaszokra oszthatók, amiknek az elnevezéseit a körön kívül olvasható feliratok jelölik. A kör belsejében a megtermékenyítés után eltelt idő szerepel az egyes stádiumokhoz rendelve. A zigotikus génaktiváció (ZGA) időszaka barack színnel jelölve látható.

A kutatócsoportunkban mért korábbi (jelenleg publikálás alatt lévő) eredmények alapján a zebrahal embrionális fejlődésének legelső szakaszán magas dUTP- és genomi uracilszint mérhető, ami a ZGA-t követően jelentősen lecsökken (8. ábra). Azt is megfigyelték, hogy amennyiben dUTP-t bontó enzimet (aktív dUTPáz) mikroinjektálnak az embriókba, a fejlődés során komoly rendellenességek lépnek fel. Ez alapján feltételezhető, hogy a kezdeti magas uracilszintnek, valamint az uracil későbbi processzállódásának jelentős szerepe lehet az embrionális fejlődésben.



8. ábra: A dUTP és genomi uracilszint korai zebrahál embrióban (az ábrát Dr. Varga Máté készítette). A: A dNTP-készlet mennyiségi meghatározása különböző stádiumú embriókból. Az 5féle dNTP (színek bal oldalon) mennyiségét 100-100 embrióból határozták meg a csoportban fejlesztett módszerrel [45]. A hibaszálók a 3 biológiai párhuzamos mérés eredményének standard szórását jelölik. Az embrionális stádiumok sémái a megfelelő diagramok mellett, a kapcsolódó életkorral (B panel függőleges tengelye) összekötve jelennek meg. B: A relatív dNTP szintek alakulása a korai embrionális fejlődés során. A teljes dNTP mennyiségre vonatkoztatott relatív értékek a megtermékenyítés után eltelt idő (hpf) függvényében. C: A genomi uracil tartalom változása. A box diagramon az medián és a kvartilisek szerepelnek, a B panel skálájához illesztve.

3. Célkitűzés

Kutatásom alapvető célja közelebb kerülni a zebraahal embriókban megfigyelt megemelkedett dUTP- és genomi uracilszint egyedfejlődésbeli jelentőségének megértéséhez. Ehhez első megközelítésben bioinformatikai elemzéseket terveztem végezni, összehasonlítva az adatbázisokból legyűjtött funkcionális genomikai adatokat a kutatócsoport által meghatározott genomi uracilosodási mintázattal. A kutatásom második felében a kutatócsoportunkban uracilszint mérésére használt szenzorokat kívántam tovább fejleszteni. Ennek célja két olyan fehérjeszenzor létrehozása molekuláris klónozással, amelyben a fehérje – korábbi tapasztalatok alapján stabil – N'-terminálisa komibálható egy C'-terminálisra tervezett AviTaggel, ami az enzimatis biotinilálást lehetővé teszi, így a specifikus és erős biotin-avidin kölcsönhatást kihasználva javíthatjuk az eddig is sokoldalúan alkalmazott szenzorok felhasználhatóságát.

A specifikus céljaim a következők voltak:

1. Az embrionális genom uracilban dúsult régiói és különböző szatellita régiók átfedésének mennyiségi jellemzése.
2. A funkcionális genomikai jellemzőkkel való összehasonlításhoz releváns adatok legyűjtése és rendszerezése: 0-8 hpf (fertilizáció után eltelt órák) stádiumokra vonatkozó genom-hozzáférhetőségi, illetve DNS-kötő fehérjék és hisztonmódosítások genomi eloszlását leíró szekvenálási adatokból készült genomi profilok (bigwig fájlok) és régiók (bed fájlok).
3. A 2,5 órás zebraahal embriókban mért uracilosodási mintázat és szintén 2,5 hpf körüli stádiumra vonatkozó genomi profilok genomszegmentációs elemzése.
4. Az uracilosodási mintázattal átfedő kromatin jellemzők keresése az embrionális fejlődés szélesebb tartományában (0-8 hpf).
5. A csoportban jelenleg használt uracilfelismerő szenzorok (His-3xFLAG-ΔUNG és a His-1xFLAG-UdgX) C'-terminálisának bővítése AviTaggel.
6. Az AviTaggel ellátott konstruktok termelésének és tisztításának finomítása.
7. Az előállított szenzorfehérjék kötődésének jellemzése TBE-PAGE-n és alkalmazhatóságának tesztelése csepp bloton.

4. Anyagok és módszerek

4.1. A funkcionális genomikai adatok bioinformatikai elemzése

Az általam legyűjtött funkcionális genomikai adatok bioinformatikai elemzését témavezetőmmel, Dr. Békési Angélával közösen végeztem.

4.1.1. A zebrahal genomi uracilosodási mintázatának meghatározása

A zebrahal uracilosodási mintázatát a kutatócsoportom munkatársai korábban 2,5 órás embriókban immunprecipitációval kapcsolt szekvenálással (U-DNA-seq [4]) meghatározták. A mintázatokat az egyértelműen illesztett olvasatokkal kapott lefedettségek alapján, az immunprecipitációval kapott dúsított (IP) és az input lefedettségek hányadosának kettes alapú logaritmusával jellemezték. Az uracilban dúsult régiókat három küszöbérték ($\log_2(\text{IP}/\text{input})$) értékek $> 0,25$ /megengedő/; $0,5$ /közepes/; 1 /szigorú/ alkalmazásával a $\log_2$ görbék alapján határozták meg.

4.1.2. Adatbázisból származó fájlok és feldolgozások

Munkám során a Gene Expression Omnibus (GEO) adatbázisból gyűjtöttem funkcionális genomikai adatokat (ld. Függelék 16. táblázat). Ezek olyan szekvenálási eredmények, amelyek a DNS hozzáférhetőségét (ATAC-seq, MNase-seq), metiláltságát (MeDIP-seq), valamint a DNS-hez kötődő fehérjék előfordulását (ChIP-seq, CUT&RUN) jellemzik. Mivel a U-DNA-Seq technikával kapott uracilosodási mintázatunk 2,5 órás embriókból származik, ezért az adatgyűjtés során ehhez viszonylag közeli fejlődési szakaszokra fókuszáltam. A letöltött fájlok a megtermékenyítéstől számítva maximum 8 órás embriókra vonatkoznak (0-8 hpf), valamint spermiumból és oocitából kapott adatokat is használtam. Alapvetően kétféle fájltypust gyűjtöttem: intervallumokat definiáló „bed” fájlokat, valamint genomi szintű görbéket, melyeket lefedettségi adatokból származtattak (bigwig fájlok). A gyűjtött adatokat az összehasonlító elemzés előtt formai és minőségi szempontból szűrtem.

4.1.2.1. Integrative Genomics Viewer

A fájlok vizuális megjelenítéséhez az Integrative Genomics Viewer (IGV) programot használtuk [46]. Ezzel a programmal az intervallumfájlok és a folyamatos lefedettségi görbék egyaránt szemléltethetők.

4.1.2.2. Linux parancsok és awk

Az analízist Linux operációs rendszerű számítógépen végeztük, felhasználva a Linux parancsokat. Ezek közül az alábbiakban a számunkra legfontosabbakat mutatom be egy-egy példán keresztül.

wc: Megszámolja egy fájl vagy standard bemenet (például egy parancs kimenete) sorainak (-l), szavainak (-w) vagy karaktereinek (-m) számát.

`wc -l NAME.bdg` → Megszámolja a NAME.bdg fájl sorait.

sort: Szöveges dokumentumok rendezésére szolgál. Alapértelmezés szerint betűrendbe vagy numerikus sorrendbe rendezi a szöveges adatokat, de számos más opcióval módosítható.

`sort -k1,1 -k2,2` → Először az első oszlop szerint rendezi sorrendbe az adatokat, majd a második oszlopot is figyelembe veszi.

grep: Szöveges fájlokban vagy bemenetekben kereshetünk vele megadott elemeket. Számos kapcsolóval módosítható, így akár bizonyos részek eltávolítására vagy beillesztésére is alkalmas.

`grep -vw "track" NAME1.bed > NAME2.bed` → A NAME1.bed fájlból kiszűri azokat a sorokat, amelyek önálló szóként nem tartalmazzák a "track" kifejezést, majd ezeket egy új fájlba, a NAME2.bed-be írja.

ls: Kilistázza a könyvtárak tartalmát. Számos opcióval módosítható a fájlok és könyvtárak megjelenítése, például rejtett fájlok listázása, részletes információk, méret vagy idő szerinti rendezés.

`ls LIBRARY/*bigwig > list_of_bigwigs.txt` → LIBRARY könyvtárban található bigwig kiterjesztésű fájlok nevét kilistázza, majd beleírja a list_of_bigwigs.txt. nevű új fájlba.

cat: Kapcsoló nélkül megjeleníti a megadott fájl tartalmát. A megfelelő beállításokkal alkalmazható fájlok másolására és összefűzésére is.

`cat NAME2.bed >> NAME1.bed` → A NAME1.bed eddigi tartalma után írja a NAME2.bed tartalmát.

Az awk egy Linux parancs és programozási nyelv, amely lehetőséget biztosít az adatok feldolgozására, kezelésére, valamint formázott kimenetek előállítására. Az eszközzel számos szövegfeldolgozási művelet elvégezhető, és megkönnyíti az összetett adatkijelölések

végrehajtását. Az elemzések során több alkalommal is használtuk ezt a parancsot, erre egy példa az alapvető kromoszómákra való szűrés.

```
awk '($1=="chr1") || ($1=="chr2") || ($1=="chr3") || ($1=="chr4") ||  
($1=="chr5") || ($1=="chr6") || ($1=="chr7") || ($1=="chr8") ||  
($1=="chr9") || ($1=="chr10") || ($1=="chr11") || ($1=="chr12") ||  
($1=="chr13") || ($1=="chr14") || ($1=="chr15") || ($1=="chr16") ||  
($1=="chr17") || ($1=="chr18") || ($1=="chr19") || ($1=="chr20") ||  
($1=="chr21") || ($1=="chr22") || ($1=="chr23") || ($1=="chr24") ||  
($1=="chr25") {print $0}' NAME.bdg > NAME.filtered.bdg
```

4.1.3. Genomszegmentációs analízis

A Segway genomszegmentációs felügyelet nélküli tanulóalgoritmus a folyamatos lefedettségű görbék összehasonlító elemzésére szolgál. A módszer lényege, hogy a program a genomot kis ablakokra osztja, és a betáplált görbéket ezekben tanulja. Majd ezeket hasonlóság alapján megadott számú szegmenscsoportba osztja és felcímkézi. Az azonos címkével ellátott régiók tehát bizonyos tulajdonságokkal rendelkeznek a megfigyelt adatokban. A szegmentálási feladat magában foglalja a szegmenshatárok megtalálását, miközben egyidejűleg címkéket rendel a szegmensekhez [47].

A Segway nagy számítási kapacitást igényel, ezért leszűkítettük a fájlokat 0-4,5 hpf embriókból, valamint oocitából és spermiumból származó adatokra. Bemenetként minden esetben az IP/input lefedettségű arány görbéket használtuk BedGraph (bdg) formátumban. A bdg tulajdonképpen nagyon hasonló a wig fájlhoz, ugyanúgy egymást követő, rövid genomi szakaszokhoz rendel értékeket és tömörítetlen szöveges fájlformátum.

A fájlokat awk segítségével szűrtük az alapvető kromoszómákra.

Több ChIP-seq adat esetében a lefedettségű görbék nem voltak osztva az input minta adataival, ezért az egységesítés érdekében a szükséges fájlokon elvégeztük ezt a műveletet.

```
bigwigCompare -b1 ChIPfile.bw -b2 input.bw -o ChIPfile.ratio.bdg -of  
bedgraph --binSize 100 -v -p 32 --skipZeroOverZero --skipNAs --  
operation ratio
```

- ➔ -b1: bemenet (az immunprecipitált mintából származó fájl); -b2: bemenet (az input fájl); -o: kimenet (a két bemenet aránya); -of: a kimeneti fájl formátumát adja meg; -binSize: megadja, hogy hány bázispár hosszú szakaszokban számoljon; -v: a feldolgozottságról üzenetet küld; -p: megadja, hogy hány szálon fut; --skipZeroOverZero: 0/0 arány esetén az adatot kihagyja; --skipNAs: ha valamelyik bemeneti fájlban hiányzó érték van, kihagyja a számításból; --operation ratio:

meghatározza a két bemeneti fájl közötti műveletet, jelen esetben a hányados kiszámítását

Az elemzéshez a bigwigben adott fájlokat BedGraph formátumba kell konvertálni, amihez BigWigToWig eszközt használtuk [48].

```
bigWigToWig NAME.bw NAME.bdg
```

Problémát jelentett az is, hogy nem minden fájlban azonos referenciagenom-verzióhoz voltak rendelve az adatok, ezért liftOver paranccsal egységesítettük a GRCz11/danRer11 referenciagenomhoz.

```
liftOver INPUT.bed chain OUTPUT.bed unMapped
```

- ➔ A chain fájl meghatározza, hogyan kell a koordinátákat az egyik referenciagenomról a másikra átvinni, az unMapped pedig azokat a szakaszokat tartalmazza, amiket nem tudott átfordítani.

A Segway első lépése az adatok elnevezése és betöltése egy megfelelő adatstruktúrába.

```
genomedata-load -s danRer11.core.fa -t SAMPLE1=NAME1.bdg -t  
SAMPLE2=NAME2.bdg ... GENOMELOAD
```

- ➔ -s: a referencia genom (zebrahal danRer11 genomi adatai); -t {SAMPLE1}={FILE1}: itt adhatók meg az egyes analizálandó fájlok és azok elnevezései, GENOMELOAD: a betöltött adatokat tartalmazó kimeneti fájl

A következő lépés a genomikus adatok tanulása, ahol a Segway train elindítja a Segway tanuló folyamatát: a program a minták közti különböző jellemző mintázatokat tanulja meg a genomi ablakokban.

```
SEGWAY_CLUSTER=local segway train GENOMELOAD--resolution=100 --num-  
labels=12 -- minibatch-fraction=0.01 train12_ZebrafishE2p5_UDNAseq
```

```
SEGWAY_CLUSTER=local segway train GENOMELOAD--resolution=100 --num-  
labels=25 -- minibatch-fraction=0.01 train25_ZebrafishE2p5_UDNAseq
```

- ➔ SEGWAY_CLUSTER=local: az elemzés helyét adja meg (helyi); GENOMELOAD: a betöltött genomikus adatokat tartalmazó fájl; --resolution=100: meghatározza a szegmensek hosszát (100 bázispár); --num-labels=12: a szegmensek számát adja meg; --minibatch-fraction=0.01: megadja, hogy az adatoknak mekkora részét használja fel egy-egy iterációban; train12_ZebrafishE2p5_UDNAseq: a kimeneti fájl vagy mappa neve

A Segway annotate azonosítja a genomikus régiók mintázatait, és a korábban megtanult minták alapján szegmentálja és címkézi a genomot.

```
SEGWAY_CLUSTER=local          segway          annotate          GENOMELOAD
train12_ZebrafishE2p5_UDNaseq_1 Segway_annot_ZebrafishE2p5_UDNaseq
```

➔ SEGWAY_CLUSTER=local: az elemzés helyét adja meg (helyi); segway annotate: végrehajtja a Segway annotációs folyamatát; GENOMELOAD: a betöltött genomikus adatokat tartalmazó fájl; train12_ZebrafishE2p5_UDNaseq: a korábbi tanítási eredményeket tartalmazó fájl; Segway_annot_ZebrafishE2p5_UDNaseq: az annotáció kimeneti fájlja

A Segtools segítségével számítottuk ki, hogy a jelintenzitások hogyan oszlanak el az egyes szegmensekben.

```
segtools-signal-distribution segway.bed.gz GENOMELOAD
```

➔ segway.bed.gz: a Segway által létrehozott annotált fájl; GENOMELOAD: a betöltött genomikus adatokat tartalmazó fájl

4.1.4. Genomi mintázatok hasonlósági keresése a GIGGLE programcsomaggal

A GIGGLE intervallumfájlok (bed, narrowPeak, broadPeak kiterjesztésű fájlok) összehasonlítására alkalmas keresőmotor, amely gyors és nagymértékben skálázható. Az elemzés eredményeként kapott GIGGLE pontszám (combo score) arról ad információt, hogy az adott fájl mennyire hasonlít az általunk megadott fájlhoz. Az elemzés Fisher's Exact kétoldali tesztet tartalmaz, ami az egyes intervallumok kiterjedésbeli különbségeiből eredő hibák kiküszöbölésére szolgál [49].

Az intervallumfájlokat is átfordítottuk liftOverrel a GRCz11 referenciagenomra. Ezután rendszereztük és tömörítettük a fájlokat a GIGGLE csomaghoz tartozó szkripttel, mivel a GIGGLE használatához bed.gz tömörített formátumban kell lenniük.

A keresés előtt indexelni kell a könyvtárunkat, hogy az adatok struktúrált formában legyenek, lehetővé téve a gyors keresést.

```
giggle index -i "bed_sorted/*gz" -o bed_sorted_b -f -s
```

➔ -i: bemeneti fájl (bed_sorted könyvtárból a tömörített formátumú fájlok); -o: az indexelt fájlokat tartalmazó könyvtár; -f: a már létező indexeket felülírja; -s: statisztika számítása

A GIGGLE Search segítségével lehet keresni az indexelt állományban.

```
giggle search -i bed_sorted_b -q REFERENCE.bed.gz -s -g 1340613534 >
OUTPUT.csv
```

➔ -i: indexelt fájlok könyvtárát adja meg (jelen esetben ez a bed_sorted_b); -q: lekérdezési fájl, ami alapján átfedéseket keres a program; -s: statisztika számítása; -g: a referencia genom mérete; OUTPUT.csv: az eredményeket tartalmazó fájl

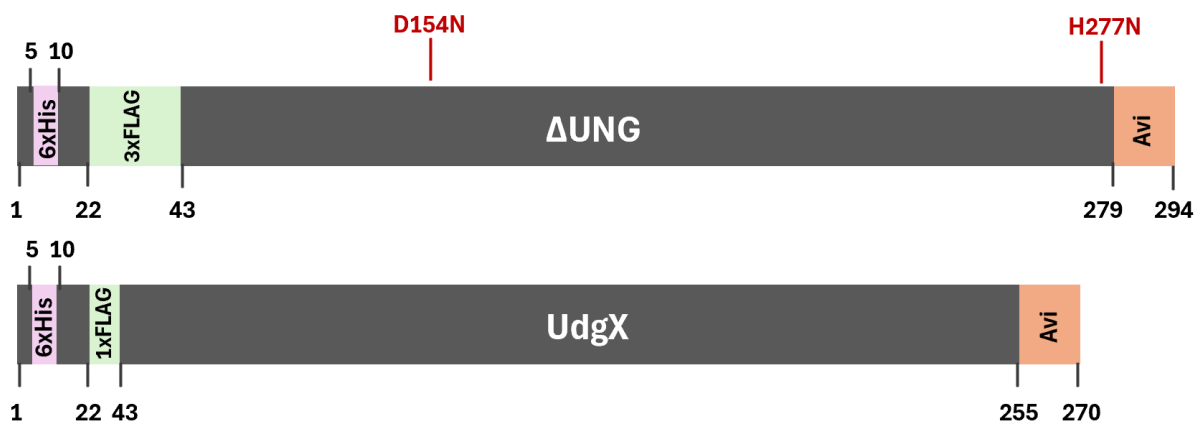
4.2. A szenzorfejlesztéshez használt kiindulási konstruktok

Kutatómunkám során kétféle szenzor továbbfejlesztésén dolgoztam: Δ UNG és UdgX alapú szenzorok (ld. Szakirodalmi áttekintés 2.2.1. fejezet)

Az általam kiindulásként használt szenzorok N'-terminálisán FLAG-címke található (a Δ UNG-nál 3xFLAG, az UdgX esetében pedig 1xFLAG). Ezt a konstrukciót mindenképpen meg szeretnénk volna tartani, mivel egyrészt nem akartuk elveszíteni az anti-FLAG antitest festés lehetőségét, másrészt pedig a korábbi tapasztalatok alapján az N'-terminális FLAG-címkét tartalmazó szerzorfehérjék termelése bizonyult a legstabilabbnak.

A konstruktok His-címkét is tartalmaznak, így Ni^{2+} - alapú affinitás kromatográfiával tisztíthatók.

A rendelkezésre álló His-3xFLAG- Δ UNG (pET15b plazmidban) és a His-1xFLAG-UdgX (pET45b plazmidban) konstruktok C'-terminálisára AviTaget terveztem (9. ábra), ami enzimatikusan biotinilálható. Ennek előnye, hogy valamilyen avidinszármazék használatával közvetlenül festhető, nincs szükség másodlagos ellenanyagra. Ez kisebb méretű komplexet és eggyel kevesebb disszociációs lehetőséget jelent. További előny, hogy a biotin-sztreptavidin kölcsönhatás igen erős ($K_d \approx 10^{-15}$ M), emiatt a szenzor biotinilálása nagyobb érzékenységet és pontosságot jelenthet. A szenzorfehérjék szekvenciája a Függelékben található (37. ábra).



9. ábra: A tervezett szenzorok. A Δ UNG N'-terminális 6xHis-címkével és 3xFLAG-címkével rendelkezik. Az eredeti UNG fehérjéhez képest két pontmutációt tartalmaz, amik bordó színnel jelöltek. Az UdgX N'-terminálisán 6xHis- és 1xFLAG-címke található. Mindkét szenzor C'-terminálisa egy 15 aminosav hosszú AviTaggel bővült.

4.3. Konstruktok előállítása molekuláris klónozással

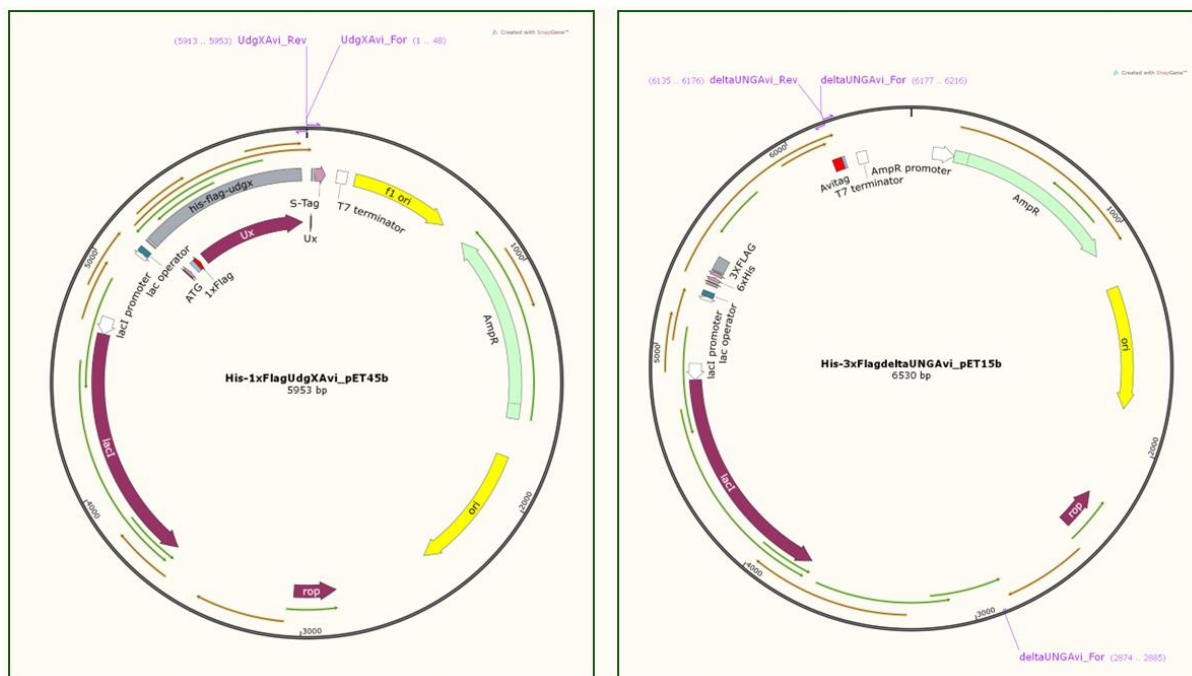
4.3.1. Polimeráz láncreakció

Az AviTag szekvenciáját a reverse és forward primerekbe terveztem, és a teljes plazmidot sokszorosítottam fel PCR reakció segítségével. A primerek szekvenciáját és a hozzájuk tartozó olvadási, valamint javasolt anellációs hőmérsékletet a 2. táblázat tartalmazza.

2. táblázat: A PCR reakció primerei és anellációs hőmérsékletek Δ UNG és az UdgX alapú szenzorokhoz. A szekvenciákban a kis betűvel jelölt szakaszok az AviTaget kódoló, a nagy betűvel jelölt részek pedig a plazmidhoz kitapadó szakaszt jelentik. T_m az oligonukleotidok olvadási hőmérséklete, T_a az anellációs hőmérséklet.

	Oligonukleotidok szekvenciája	T_m	T_a
Δ UNG for.	tcagaaaatcgaatggcacgaaTAACTCGAGGATCCGGCT	65°C	63°C
Δ UNG rev.	gcctcgaagatgtcgttcaggccGATGTACCTGTAGGTGTCC	62°C	
UdgX for.	tcagaaaatcgaatggcacgaaTGATAACTCGAGTCTGGTAAAGAAAC	64°C	65°C
UdgX rev.	gcctcgaagatgtcgttcaggccTGGCCTGACATCTGCTGC	68°C	

A PCR reakciót SnapGene program segítségével szimuláltam, a kapott His-3xFLAG- Δ UNG-Avi_pET15b és His-1xFLAG-UdgX-Avi_pET45b plazmidokat a 10. ábra szemlélteti. A His-3xFLAG- Δ UNG-Avi_pET15b plazmid esetében a forward primer egy másik helyre is képes kitapadni, viszont itt jelentősen alacsonyabb az anellációs hőmérséklet ($T_m=55^\circ\text{C}$).



10. ábra: His-3xFLAG- Δ UNG-Avi_pET15b és His-1xFLAG-UdgX-Avi_pET45b plazmidok SnapGene-ben ábrázolva

A PCR reakcióhoz Q5 High-Fidelity DNS-polimerázt használtam, a komponenseit a 3. táblázatnak megfelelően jégen összemértem, majd két részletre szedtem. A 10-10 μl térfogatú reakcióelegyekkel a Függelék 6. táblázatában leírt körülmények között elvégeztem reakciót.

Mind a Δ UNG, mind az UdgX alapú szenzorral elvégeztem a reakciót 63 °C és 65 °C anellációs hőmérséklettel is.

3. táblázat: Az AviTag PCR reakció komponensei.

Komponens	20 μ l reakcióelegy	Végző koncentráció
Q5 High-Fidelity 2X Master	10 μ l	1X
10 μM forward primer	1 μ l	0.5 μ M
10 μM reverse primer	1 μ l	0.5 μ M
plazmid DNS	1 μ l	változó
nukleázmentes víz	7 μ l	-

A PCR reakció termékeit agaróz gélen megfuttattam (ld. 4.7.1. fejezet).

4.3.2. KLD reakció

A KLD enzimmix három enzimet tartalmaz: a kinázt, ligázt és DpnI-t. A kináz a DNS 5'-végeit foszforilálja, így a ligáz képes a DNS két végét összekapcsolni. A DpnI enzim a mutagenézis során fontos, mivel hasítja a metilált adenin bázisokat tartalmazó szekvenciákat, így képes a metilált szülői DNS specifikus eltávolítására, érintetlenül hagyva az újonnan szintetizált mutációt tartalmazó szálat. A PCR termékből 1 μ l-t hozzáadtam 5 μ l KLD reakciópuffer és 3 μ l nukleázmentes víz elegyéhez, majd hozzáadtam 1 μ l KLD enzimmixet. Az így kapott 10 μ l reakcióelegyet 1 órán át szobahőmérsékleten inkubáltam.

4.3.3. Transzformálás XL1Blue törzsbe

A KLD terméket minden további kezelés vagy tisztítás nélkül XL1Blue kompetens sejtekbe transzformáltam az alábbiak szerint. Az XL1Blue törzs tetraciklin (TC) rezisztenciával rendelkezik, az ampicillin/karbenicillin (CAR) rezisztencia pedig a plazmidokban kódolt.

- 50 μ l XL1Blue sejthez hozzáadtam 5 μ l KLD terméket.
- 35 percig jégen inkubáltam, majd 2 percig 42 °C-ra tettem, ezt követően pedig 3 percig ismét jégre tettem.
- Hozzáadtam 200 μ l LB tápoldatot, majd 37 °C-on 1 órát rázattam.
- Asztali centrifugán 1 perc alatt lepörgettem a mintákat, majd a felülúszóból 120 μ l-t leszívtam, és a maradékban visszaszuszpendáltam a pelletet.
- TC, CAR antibiotikumot tartalmazó Luria-Bertani (LB) agarlemezre szélesztettem.

A telepek éjszakán át (overnight) nőttek 37 °C-on.

4.4. Kolónia PCR és a klónok felszaporítása

Az éjszakán át tartó inkubálás után felnőtt telepekből kiválasztottam a Δ UNG és az UdgX alapú konstruktokból 12-12-t. A kiválasztott telepekből pipettával mintát vettem, előre feljelölt és számozott agarlemezre szélesztettem, majd a pipettán maradt kis mennyiséget a PCR reakcióelegybe (4. táblázat) szuszpendáltam. A PCR reakcióhoz MyTaq Mix-et és T7 primereket használtam. A reakció körülményeit a Függelék 7. táblázata tartalmazza.

4. táblázat: Kolónia PCR reakció komponensei.

Komponens	10 μ l reakcióelegy	Végő koncentráció
MyTaq Mix 2X	5 μ l	1X
10 μ M forward primer	0,5 μ l	0.5 μ M
10 μ M reverse primer	0,5 μ l	0.5 μ M
nukleázmentes víz	4 μ l	-

A kolónia PCR termékeket agaróz gélen megfuttattam. Kontrollként a kiindulási His-3xFLAG- Δ UNG_pET15b és His-1xFLAG-UdgX_pET45b plazmidok PCR termékeit is futtattam a mintákkal. Az eredmények alapján kiválasztott klónokat felszaporítottam. Ehhez a megfelelő klónokból pipettával mintát vettem, és 5 ml TC és CAR tartalmú LB tápoldatba szuszpendáltam. A kultúrákat éjszakán át rázattam 37 °C-on.

4.5. Fehérjetermelés

A kolónia PCR alapján kiválasztott és felszaporított klónokból izoláltam a plazmidokat a NucleoSpin Plasmid Kittel (Macherey-Nagel) a kit leírása szerint. A módszer lényege, hogy lúgos lízis segítségével szabadítják fel a plazmid DNS-t az *E. coli* sejtekből. A semlegesítést követően a fehérjék, a nagyméretű genomikus DNS és sejtörmelék kicsapódik, és centrifugálással eltávolításra kerül. A tiszta oldatból a plazmid DNS-t az oszlop szilikagél membránjára kötik. A szennyeződések mosási lépésekkel távolítják el, és a tiszta plazmid DNS-t enyhén lúgos pufferrel vagy vízzel eluálják.

A plazmidok koncentrációját NanoDrop™ 2000 (ThermoFischer Scientific) UV-spektrofotométer segítségével meghatároztam, és Sanger-szekvenálásra küldtem a mintákat (Microsynth).

4.5.1. Transzformálás BL21 sejtekbe

A konstruktot tartalmazó plazmidokat BL21 *ung-* pLysS fehérjetermelő törzsbe transzformáltam. Az *ung-* sejtvonalt az esetleges endogén UNG szennyezés elkerülése

érdekében fontos. Az UNG enzim jelenléte ugyanis a szenzorfehérje használata közben az uracil kivágását eredményezhetné a DNS-ből, ami torzíthatná a mérési eredményeket. A pLysS plazmid egyrészt az indukció előtti expresszió erősebb kontrollját, másrészt a sejtek könnyebb lízisét biztosítja.

A transzformálást a 4.3.3. pontban leírtakkal megegyező módon végeztem, annyi különbséggel, hogy a táptalaj a karbenicillin mellett klóramfenikolt (CA) tartalmazott, mivel a pLysS plazmid klóramfenikol rezisztenciát okozó gént tartalmaz. A sejtek egy részét agarlemezre szélesztettem, a másiktól pedig 5 ml-es előkultúrát indítottam.

A sejtek éjszakán át nőttek 37 °C-on.

4.5.2. Fehérjeexpresszió

A próbaexpresszióhoz a felnőtt kultúrákból 500 µl-t oltottam 50 ml CAR- és CA-tartalmú, folyékony LB tápoldatba, majd a megfelelő optikai denzitás (OD) (5. táblázat) elérésekor izopropil-β-D-1-tiogalaktopiranoziddal (IPTG) indukáltam. A ΔUNG termelésére a kutatócsoportban volt már tapasztalat, az UdgX esetében viszont nehézségek adódtak, ezért erre vonatkozóan több körülményt is kipróbáltam. A próbaexpressziók körülményeit az 5. táblázat foglalja össze.

5. táblázat: Próbaexpressziók körülményei.

	ΔUNG	UdgX ₁	UdgX ₂	UdgX ₃
OD	0,6	1,5	0,6	0,6
IPTG koncentráció	0,6 mM	1 mM		
Expressziós hőmérséklet	18 °C	37 °C	18 °C	37 °C
Expressziós idő	overnight	2 óra	overnight	4 óra
Rázató fordulatszáma	180 rpm			

A nagyléptékű expressziót a próbaexpresszió alapján kiválasztott körülményekkel végeztem többször 500 ml LB tápoldatban, melyet 5 ml előkultúrával oltottam be. Ez mindkét szenzornál OD=0,6-nál történő indukálást jelentett, a ΔUNG esetében 18 °C-on éjszakán át, az UdgX-nél 37 °C-on 4 órán át tartó expresszióval.

4.6. Fehérjetisztítás

4.6.1. Sejtek feltárása

Az expresszióra szánt idő leteltével a sejteket átöntöttem centrifugacsövekbe, és 30 percig centrifugáltam 3500 rpm-en, 4 °C-on. A centrifugálás után minden műveletet jégen végeztem.

A felülúszót leöntöttem, a pelletet pedig izotóniás fosztátpufferben (PBS) újraszuszpendálva mostam, majd ismét centrifugáltam. A pelletet ebben az állapotban folyékony nitrogénben történő fagyasztást követően -80°C-on tárolhatjuk a későbbi feltárásig. A feltáráshoz a pelletet felsuszpendáltam lízispufferben (Függelék 13. táblázat). A szuszpenziót potterbe öntöttem, és homogenizáltam. A szuszpenziót 5 × 3 percig jégen hűtve szonikáltam fél másodperces szünetekkel, közepes szonikáló fejjel és erősséggel (Bandelin Sonoplus, HD 2070). Centrifugálást követően (20 000 rpm, 4 °C, 40 perc) a felülúszóból (extraktum) a konstruktot Ni-NTA (nikkel-nitrilo-trieetsav) gyantán tisztítottam, vagy előbb biotiniláltam és monomer-avidin agaróz oszlopon választottam el. A feltárás és tisztítás lépései során minden esetben mintát vettem, amit SDS-gélelektroforézissel (ld. 4.7.2) analizáltam.

4.6.2. Biotinilálás

Az *in vitro* biotinilálást BirA enzim segítségével oldatban (extraktum) vagy gyantához kötött fehérje konstruktokkal csőforgatón inkubálva végeztem. A biotiniláláshoz szükséges reakcióelegy: 16 mM MgCl₂, 8 mM ATP, 80 μM bioitin, 40 μl BirA [50]. A kit leírásából kiderül, hogy a BirA aktivitását több szokásos pufferkomponens is negatívan befolyásolja: így kerülendő a glicerín, a NaCl és a β-merkaptóetanol. Ezért a pufferekbe eleve KCl-t használtam NaCl helyett, illetve a gyantához kötött állapotban történő biotinilálás előtt puffert is cseréltem. A biotiniláló puffer összetétele a Függelékben (12. táblázat) található.

4.6.3. Tisztítás oszlopon

4.6.3.1. Ni²⁺-affinitás kromatográfia

Az elválasztás előtt előkészítettem a Ni-NTA gyantát: átmostam desztillált vízzel, majd 30 mM imidazolt tartalmazó lízispufferrel. A gyantára rátettem az extraktumot, és inkubáltam 4 °C-on 1 órát folyamatos forgatás közben, majd betöltöttem az oszlopba.

Az áteső leszedése után a ΔUNG alapú szenzort LS (low salt, alacsony sótartalmú) pufferrel, majd HS (high salt, magas sótartalmú) pufferrel mostam, hogy az aspecifikusan kikötött fehérjéket eltávolítsam. Abban az esetben, amikor az oszlopon biotiniláltam a fehérjét, a HS pufferes mosás után egy plusz mosási lépéssel csökkentettem az ionerősséget és a β-merkaptóetanol-koncentrációt, amelyek a BirA aktivitását gátolnák. Ezek után adtam hozzá a biotiniláló reagenseket, és csőforgatón éjszakán át inkubáltam 4°C-on. Ezt követően újra oszlopba töltve a BirA tartalmú átesőt leengedtem, LS, majd VHS (very high salt, nagyon magas sótartalmú) pufferrel, végül a mosó pufferrel mostam a gyantát (a pufferek összetétele a Függelék 12. táblázatában található). A mosási lépésekhez 10x oszloptérfogatnyi puffert

használtam. A fehérjét 350 mM imidazollal (LS-pufferben), valamint, hogy az elúció teljességéről megbizonyosodjak, második lépésben 1 M imidazollal eluáltam.

Az UdgX alapú szenzor feltárását a saját lízispufferében (Függelék 13. táblázat) a Δ UNG konstrukttal megegyező módon végeztem. Az UdgX esetében kritikus, hogy a fehérje szerkezetében fontos szerepet játszó vas-kén klasztert megőrizzük, ezért a kelátorok és az oxidációs állapotot jelentősen befolyásoló pufferkomponensek alkalmazását minimalizálni igyekeztem. Ugyanakkor a korábbi tapasztalatok alapján a magas ionerősség szükséges a fehérje oldatban tartásához. Ezért a Ni^{2+} -affinitás kromatográfia során is kerültem a változatos pufferes mosásokat, és az alap puffer A-val mostam hosszabban. Hasonlóan, az elúciónál az imidazolkoncentrációt is minimalizálni igyekeztem: puffer A-ban oldott 250 mM, és második lépésként 500 mM imidazollal próbálkoztam.

A fehérjék elúcióját minden esetben Bradford-reagenssel követtem nyomon: parafilmre cseppentve 1 μ l eluátumhoz 20 μ l 5x híg Bradford-reagenst tettem. A kromatográfiás frakciókból, valamint az elúció után visszamaradt gyantából is 20-20 μ l mintát vettem SDS-gélre. Az eluált fehérjét dialízist követően töményítettem.

4.6.3.2. Monomer avidin affinitás kromatográfia

A már biotinilált fehérje tisztítása monomer avidin agaróz oszlopon is végezhető. Az avidin tetramer formája (sztreptavidin) irreverzibilisen köti a biotint, viszont monomerként reverzibilisen. Ennek megfelelően a monomer avidin oszlopról biotint tartalmazó pufferrel eluálható a biotinilált fehérje. A használat előtt az esetlegesen jelenlevő irreverzibilis kötőhelyeket blokkolni kell biotinnal való feltöltést követő alacsony pH-jú regenerálással.

A monomer avidin-agaróz 50%-os szuszpenzióból 1,5 ml-t oszlopba töltöttem, kétszeres oszloptérfogatnyi puffer A-val, majd 2 mM biotint tartalmazó oldattal recirkuláltatva (10×1 ml), utána kétszeres oszloptérfogatnyi glicines pufferrel (0,1 M; $p=2,8$) mostam, végül pedig négyszeres oszloptérfogatnyi puffer A-val hoztam egyensúlyba. A biotinilált extraktot dializáltam (ld. 4.6.4.) a szabad biotin eltávolítása érdekében, mivel ez nagyban rontaná az oszlop kapacitását. A biotinilált és dializált extraktot (300 ml kultúrából származó, 45 ml oldat) steril szűrtem az esetleg keletkezett aggregátumok eltávolítása érdekében, majd a monomer avidin agarózzal inkubáltam 4 °C-on, egy órán keresztül.

A szuszpenziót beletöltöttem üvegszűrővel ellátott oszlopba 10 ml puffer A-val (Függelék 11. táblázat) mostam az oszlopot. A mosás után 5 mM puffer A-ban oldott biotinnal eluáltam a fehérjét szakaszosan: 1-1 lépésben annyi eluenst adtam hozzá, ami éppen ellepi az oszlopot,

majd elzárva 15 percig 4 °C-on inkubáltam, és leengedtem. A frakciókart egybe gyűjtöttem, összesen 9 ml-t. Ezt dialízist követően töményítettem.

4.6.4. Dialízis

A dialízis membránt (12-14 kDa cutoff, Merck Sigma) kimostam desztillált vízzel kívül-belül, majd 5 percig forraltam 10 mM Na-bikarbonát és 10 mM EDTA elegyében (pH=8). Ezután ismét alaposan átmostam desztillált vízzel. Hideg desztillált vízzel is megmostam, aztán 20% etanolban tároltam. Felhasználás előtt a membránt alaposan átmostam desztillált vízzel. A fehérjeoldatot a dialízismembránba töltöttem, és gondosan lezártam. Belehelyeztem a dialízispufferbe (Függelék 13. táblázat), és legalább $2 \times 1,5$ órán keresztül dializáltam.

4.6.5. Töményítés

Töményítő csövet (10 kDa cutoff, Merck Millipore) kimostam (feltöltöttem és centrifugáltam) desztillált vízzel és dialízis pufferrel, majd feltöltöttem a dializált fehérjeoldattal és centrifugáltam (4 °C, 4000 rpm, 5-10 perc). A centrifugálási lépések között megkevertem az oldatot és feltöltöttem újra a csövet. Ilyenkor mindig ellenőriztem, hogy nem kezd-e látható csapadék megjelenni, és időnként koncentrációt is mértem. Végül a töményítő oszlopot kimostam desztillált vízzel és 20% etanolban tároltam.

A töményítés után NanoDrop™ 2000 (ThermoFischer Scientific) UV-spektrofotométer segítségével megmértem a fehérje koncentrációját.

4.7. Gélelektroforézis

A kutatómunkám során számos alkalommal végeztem gélelektroforézist. A gélben a minták makromolekulái az egyenfeszültség hatására vándorolnak a méretüktől, alakjuktól és töltésüktől függően eltérő sebességgel, így elválaszthatók egymástól. A különböző mérettartományokba eső makromolekulák elválasztására két különböző rendszert használunk. A nagyméretű és lineáris nukleinsavak gélelektroforézise a nagyobb pórusméretű agaróz gélen egy horizontális rendszerben történik. A fehérje molekulák és a kisebb méretű oligonukleotidok nagyobb felbontású elválasztását pedig poliakrilamid gélben, egy vertikális rendszerben végezzük. A gélek elkészítéséhez összemértem a szükséges komponenseket, majd a megfelelő gélöntő apparátusba öntöttem. A gél megszilárdulása után behelyeztem a futtató kádba, és feltöltöttem a megfelelő pufferrel. Felvittem a mintákat, majd beállítottam a paramétereket, és elindítottam a futtatást. A gél ezek után fehérje vagy nukleinsav festékkel festettem. A gélek

beolvasásához Chemidoc XR+ (Bio-Rad) gépet, a kiértékeléshez ImageLab (Bio-Rad) programot használtam.

4.7.1. Agaróz gélelektroforézis

Az nukleinsavak méret szerinti elválasztását egyszerűbbé teszi, hogy töltés és alak tekintetében nagymértékben hasonlóak. Ezért általában nem használnak denaturáló körülményeket. A méret szerinti elválasztást az RNS-ek másodlagos struktúrái vagy például a plazmidok „super coil” szerkezete nehezítheti.

Jelen munkában a KLD reakcióhoz készült PCR és a kolónia PCR termékeinek vizsgálatához használtam. A futtatáshoz TAE pufferrel készített 1%-os agaróz gélét öntöttem, amihez az oldódást segítő melegítést követően 0,01% GelRed (Biotium) festéket tettem. A mintákat mintakoktélban (Gel Loading Dye Purple (6X), NEB) vittem fel, futtatáshoz szintén TAE puffert használtam, aminek összetétele a Függelék 9. táblázatában található. Méretjelzőként a Thermo Fisher GeneRuler 1 kb Plus DNS létráját használtam. A futtatást 100 V-on, 40 percig végeztem.

4.7.2. SDS-PAGE

A fehérjék gélelektroforézise során meg kell oldani a töltés és alak változatosságából eredő problémát. A nátrium-dodecil-szulfát (SDS) detergens tulajdonsága révén kitekeri a fehérjéket és egységes negatív töltéssel látja el őket. Továbbá a kisebb és – az akrilamid:biszakrilamid arány variálásával – jobban szabályozható pórusméretű poliakrilamid gélek használata nagyobb felbontást biztosít. Ennek megfelelően az SDS-sel végzett poliakrilamid (37,5:1) gélelektroforézis (SDS-PAGE) alkalmas a fehérjéink méret szerinti elválasztására és elemzésére.

Az SDS-gél két részből áll: tömörítő- és szeparálógél (összetétel ld. Függelék 14. táblázat). A tömörítő gél (5% akrilamid) biztosítja, hogy az elválasztás során a fehérjék egy pontból induljanak, a szeparálógélben (12% akrilamid) történik a fehérjék elválasztása. A gél polimerizációjának megindításához TEMED és 1% APS oldat hozzáadása szükséges. Először a szeparálógélt öntöttem meg, majd a megszilárdulása után rápipettáztam a tömörítő gélre, amibe belenyomtam a zsebek kialakítására alkalmas fésűt, és megvártam amíg megköt.

A mintákhoz a felvitel előtt SDS-mintakoktélt (Függelék 10. táblázat) adtam olyan mennyiségben, hogy a mintakoktél ötszörösére híguljon. 95 °C-on 5 percig forraltam, majd bemértem a mintákat és a fehérjelétrát (Grisp GRS Protein Marker MultiColour) az SDS-gél zsebeibe.

Az elektroforézis futtatópuffer összetétele a Függelék 9. táblázatában található. A futtatást 200 V egyenfeszültségen 45 percig végeztem.

A megfuttatott gélt desztillált vízzel többször megmostam, majd Coomassie alapú PageBlue festékben (ThermoFisher Scientific) áztattam 40 percet. A festéket ezután leöntöttem a gélről, és desztillált vízzel mostam a megfelelő differenciálódásig, amíg a fehérjespecifikus, kékre festődő sávok kiemelkedtek a háttérből.

4.7.3. Elektroforetikus mobilitás eltolódás esszé

A TBE-PAGE (TRIS-borát-EDTA-PAGE) alapvetően nukleinsavak elválasztására használható módszer, mely során szintén akrilamid gélt használunk, de az SDS-PAGE-val szemben itt az akrilamid mellett a keresztkötő bisz akrilamid aránya nagyobb (19:1), és nem öntünk tömörítő gélt. Munkám során denaturáló és nem denaturáló TBE-PAGE-val is dolgoztam, előbbi 7 M ureát is tartalmazott. A futtatáshoz TBE puffert használtam (Függelék 9. táblázat). A TBE puffer EDTA-t tartalmaz, ami megköti a kétértékű fémionokat, így gátolja a nukleázok működését.

A TBE-PAGE-t kutatásom során az elektroforézis mobilitás eltolódási vizsgálatához (EMSA) használtam, ami a termelt szenzorfehérjék kötő aktivitásának ellenőrzésére szolgált. A fehérjékhez olyan 5'-végen Cy3 fluoroforral jelölt 30-mer oligonukleotidot adtam, ami középen egyetlen uracilt tartalmaz (5'-AGGAATTGGTATUATGATTAGAGTGGAGTAA-3'). Abban az esetben, ha a fehérje aktív, felismeri az uracilt, komplexet képez az oligonukleotiddal, akkor a gélen ez eltolódást eredményez. A specificitás ellenőrzése érdekében olyan oligonukleotiddal is elvégeztem a kísérletet, ami az uracil helyett citozint tartalmaz. Kontrollként UNG kezelt oligonukleotidokat alkalmaztam, amiket 0,1 M NaOH hozzáadása után 5 percig 95 °C-on melegítettem. Az UNG kivágja az uracilt, a lúgos melegítés hatására pedig az abázisos hely mentén hidrolizál a cukor-foszfát gerinc, így az uraciltartalmú oligomerekből rövidebb szakaszok keletkeznek. A géltre minden esetben 4 pmol oligomert és különböző mértékű moláris feleslegben lévő szenzor elegyét vittem fel.

A denaturáló gél esetében a minták futtatása előtt előfuttattam a gélt 300 V-on 1 órán át, ami a denaturációt, és a rendszer egyensúlyba hozását segíti. Ezt követően a zsebeket Hamilton-fecskendővel átmostam a kidiffundáló urea eltávolítása érdekében, és betöltöttem a mintákat, majd elindítottam a futtatásukat 300 V-on. A futtatást addig végeztem, amíg a futási front el nem érte a gél ¾-ét.

4.8. Western blot

A Western blot az SDS-gélen elválasztott fehérjék specifikus detektálására alkalmas módszer. Az elválasztott fehérjéket a gélből egy hidrofób membránhoz kötjük, majd valamilyen módon jelöljük. Ez lehetséges az adott fehérjére specifikus antitesttel (elsődleges és másodlagos antitest), vagy más specifikus kölcsönhatást (pl. biotin-sztreptavidin) kihasználva is.

A megfuttatott SDS-gélt desztillált vízzel mostam a futási front eltűnéséig, majd CAPS pufferbe (Függelék 8. táblázat) áztattam. A gélt először metanolban, majd CAPS pufferben áztatott hidrofób PVDF membránra (Immobilon-P Transfer Membrane, Merck) helyeztem, amit CAPS pufferrel átitatott szűrőpapírokkal és szivacsokkal együtt az erre alkalmas présbe fogtam. Ezt a blottoló készülékbe helyeztem, és feltöltöttem CAPS pufferrel. A behelyezésnél figyeltem arra, hogy membrán a pozitív oldal felé legyen, mivel a fehérjék a lúgos CAPS ($\text{pH} > 11$) oldatban többnyire deprotonáltak, így negatív töltésűek, ezért az egyenáram hatására pozitív irányba haladnak. A készülékbe jégakkut tettem, valamint az egész apparátust jeges fürdőbe helyeztem és kevertetem, hogy a folyamat során ne melegedjen túl. A blottolás 350 mA áramerősségen, 3 órán át zajlott. Ezután a membránt egy 50 ml-es Falcon-csőbe hajtottam és 5 ml 5% tejport tartalmazó TBS-T-ben (Függelék 8. táblázat) inkubáltam szobahőmérsékleten 1 óráig (vagy éjszakán át 4 °C-on), csőforgatóban biztosítva az egyenletes folyadékfedettséget.

Ezt követően a festési eljárás következett, melynek lényege, hogy a membránra rögzített fehérjéket specifikus jelöléssel detektálhatóvá tegyük. Ehhez az „ellenanyagokat” (anti-FLAG antitest, tormaperoxidázal (HRP) konjugált másodlagos antitest vagy sztreptavidin-HRP) 5% tejport tartalmazó TBS-T oldatban vittem fel 10 000x hígításban. A meg nem kötött fehérjék eltávolításához minden esetben 3×10 másodperc, majd 2×20 perc TBS-T oldatos mosást végeztem.

Először a membránra felvittem az elsődleges „ellenanyagot” (egérben termelt anti-FLAG antitest vagy sztreptavidin-HRP), és 1 óráig inkubáltam folyamatos forgatás közben. Az inkubáció után TBS-T-vel mostam. Ezután az anti-FLAG-es festés esetében hozzáadtam a másodlagos ellenanyagot (HRP-konjugált anti-egér antitest), amit szintén 1 óráig inkubáltam, majd TBS-T-vel mostam. Az előhívást mindkét esetben megelőzte 3×10 másodperc, majd 10 perc TBS-sel történő mosás a Tween eltávolítása érdekében. A sztreptavidinhez és a másodlagos ellenanyaghoz konjugált, az előhíváshoz szükséges HRP fényérzékeny, ezért miután ezeket a membránra tettem, gondosan befedtem a csövet alufóliával.

Az előhíváshoz ECL (Enhanced Chemiluminescence, Millipore) oldatot (a reagenseket 1:1 arányban alkalmazva) használtam. A membránra pipettáztam az előhívó oldatot, és az 1 perces reakcióidő elteltével Bio-Rad ChemiDoc™ MP készülékkel detektáltam a kemilumineszcens jelet és a Bio-Rad ImageLab szoftverrel elemeztem.

4.9. U-DNS-t detektáló csepp blot

A szenzorok teszteléséhez két genomi DNS-mintát használtam. Ezeket egyaránt uracil-DNS-glikoziláz inhibitort (UGI) stabilan expresszáló HCT116 sejtvonalból izolálták, ahol az uracil javítása gátolt. Az egyik mintához (a továbbiakban 5FdUR minta) a sejtvonalat timidilát-szintáz gátló gyógyszerrel, 5-fluoro-2'-deoxiuridinnal (5FdUR) kezelték, aminek hatására a sejtbeli dUTP/dTTP arány megnő, és az UGI által gátolt hibajavításnak köszönhetően magasabb genomi uracilszintet (600-800 U/millió bp) eredményez. Emellett a másik minta (a továbbiakban NT minta) gyógyszeres kezelés nélküli sejtvonalból származik, ennek megfelelően alacsony a genomi uracilszintje (1-2 U/millió bp). A mintákat a blotra felvittem intakt formában és UNG kezelés után is. Az UNG kezelés eltávolítja az uracilt a DNS-ből, így vizsgálható a szenzorok specifitása. A kalibrációhoz dUTPáz és UNG kiűtött (*dut-/ung-*) *Escherichia coli* genomi DNS-t használtam, melyek uraciltartalma kb. 5900/millió bp [27].

A csepp blotot pozitív töltésű nylon membránon (GE Healthcare Amersham Hybond™-N⁺) végeztem, amit használat előtt 10 percig TE pufferben (Függelék 8. táblázat) áztattam. A mintákból négy pontos, a kalibrációs standardból pedig hét pontos felező hígítást készítettem TE pufferrel. Az NT mintából 0,5 µg, az 5FdUR mintából 0,05 µg, a standardból pedig 0,005 µg mennyiséget vittem a hígítási sor első pontjába. A mintákat 5 percig 95 °C-on denaturáltam, majd jégre helyeztem és lepörgettem asztali centrifugán. A beáztatott membránt a készülékbe helyeztem, majd összeszereltem a csepp blot apparátust. A mintákból 40 µl-t pipettáztam a lyukakba, majd vákuum segítségével rászívtam a membránra. A membránt üveglapra tettem száradni 10 percig, majd szárítószekrénybe helyeztem 2 órára 80 °C-on a keresztkötések kialakulása érdekében. Az így keresztkötött nukleinsavakat biztonságosan lehet festeni a membránon. Ehhez a blokkoló pufferben hígított szenzorral 2 óráig szobahőmérsékleten vagy éjszakán át 4 °C-on inkubáltam. Az oldat 15 µg/ml szenzorfehérjét, 5% tejport, 50 µg/ml lazac DNS-t, és a ΔUNG esetén 10 mM, az UdgX esetén 2 mM β-merkaptóetanolt tartalmazott ETBS-T-ben. A festés és az előhívás a 4.8. pontban ismertetett Western blotos eljárással megegyezett, azonban TBS és TBS-T helyett ETBS-t, valamint ETBS-T-t használtam (Függelék 8. táblázat). Ez azért szükséges, mivel a csepp blot esetében DNS-mintával dolgozunk, és az EDTA a kétértékű fémionok megkötésével gátolja a nukleázok működését.

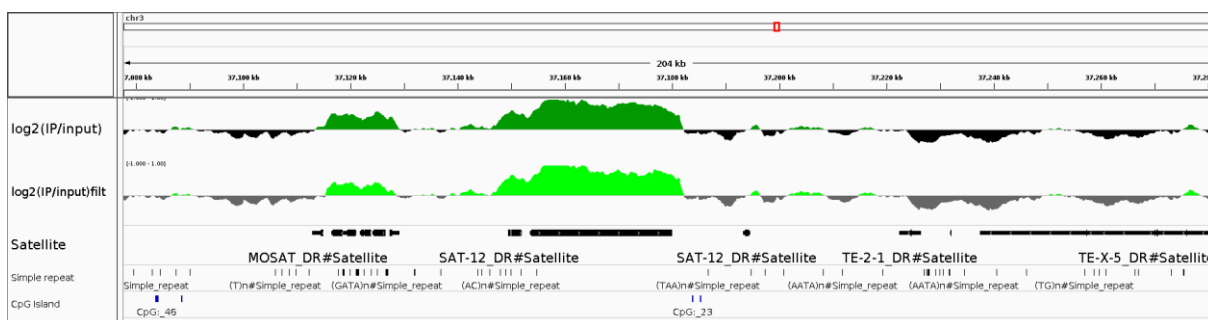
5. Eredmények és értékelésük

5.1. Genomi uracilmintázat bioinformatikai elemzése korai zebrahal embriókban

A bioinformatikai elemzések során a 2,5 órás zebrahal embriókból U-DNA-Seq segítségével kapott uracilosodási mintázatokat hasonlítottuk össze a szatelliták elhelyezkedésével és funkcionális genomikai adatokkal.

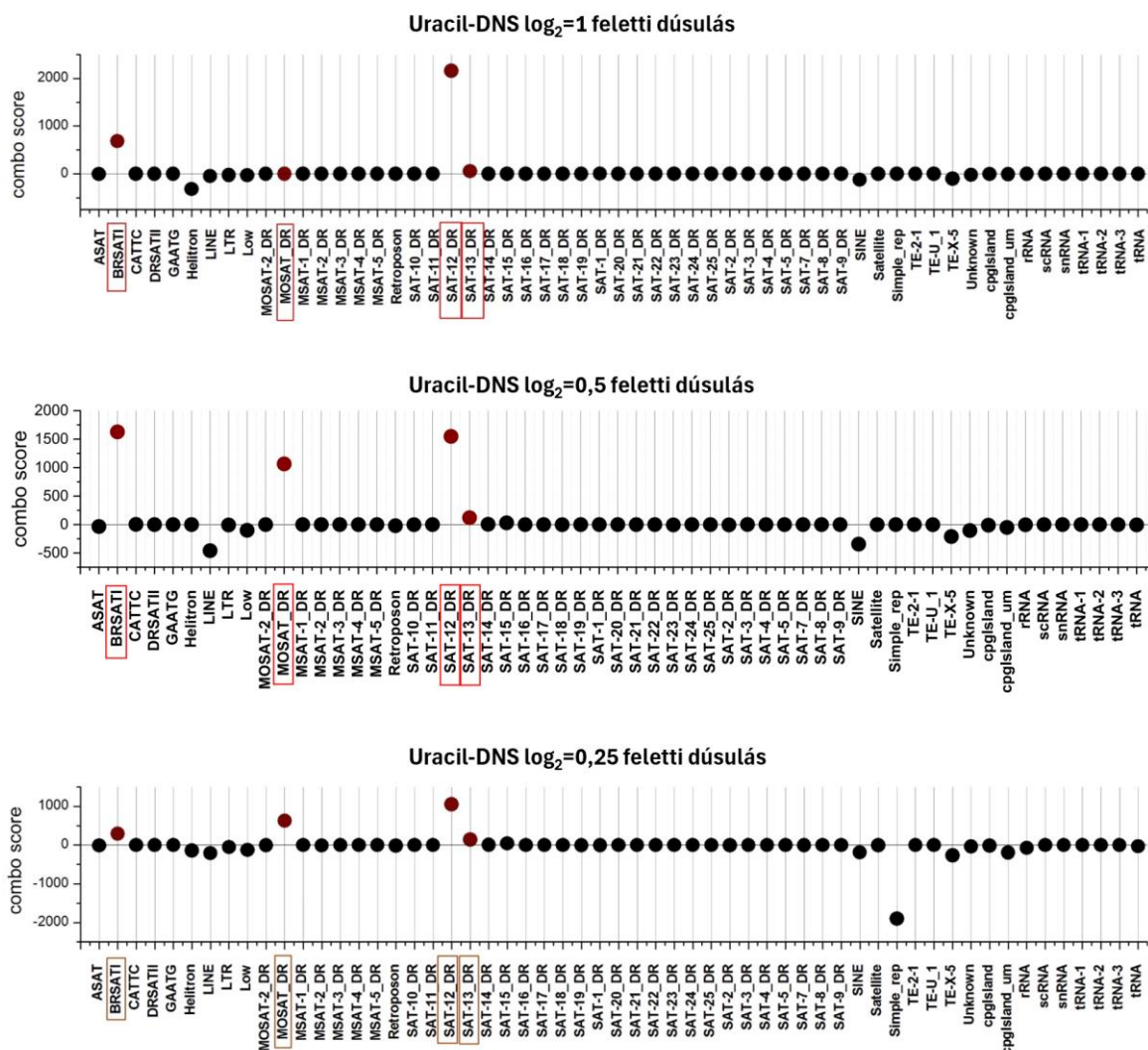
5.1.1. Szatelliták átfedése az uracilos régiókkal

A processzált U-DNA-Seq adatokat IGV-ben vizualizálva szembeötlő volt, hogy a zebrahal genom szatellitái átfedést mutatnak a 2,5 órás embrióból származó, uracilban dúsult DNS-szakaszokkal (11. ábra).



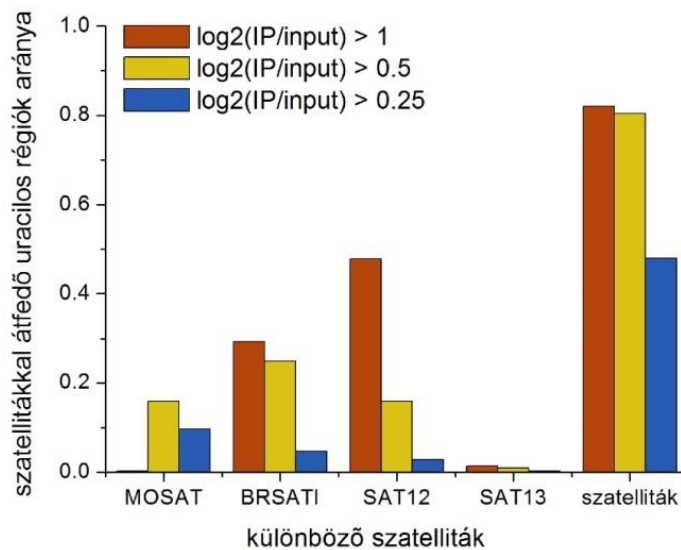
11. ábra: A 3. kromoszóma 37078-37282 bp közti 204 kb-os részlete, IGV-ben ábrázolva: az uracilos genomi régiók jól láthatóan átfednek a MOSAT és SAT-12 szatellitákkal. Az U-DNA-Seq adatokat a $\log_2(\text{IP}/\text{input})$ értékek reprezentálják, ahol a pozitív, dúsult régiókat zölddel emeltük ki. A nyers adaton (sötét zöld) és a nem egyértelműen térképezhető (MAPQ=0) olvasatok kizárásával szűrt adaton (világos zöld) számolt görbék egyaránt mutatjuk. A repetitív szakaszokra és a CpG szigetekre vonatkozó annotációt a UCSC Table Browserből töltöttük le (<http://genome.ucsc.edu/cgi-bin/hgTables>).

Az átfedés mértékének genomi szintű meghatározásához a $\log_2(\text{IP}/\text{input})$ görbék alapján uracilosodott régiókat definiáltunk három küszöbérték ($\log_2(\text{IP}/\text{input})$ értékek $> 0,25$; $0,5$; 1) alkalmazásával. Az összehasonlító elemzéseket mindhárom vonatkozásában a GIGGLE program segítségével végeztük el. A UCSC Table Browserből letölthető „repeatmasker” által definiált szatellitákon végzett GIGGLE elemzés eredményéből (12. ábra) látható, hogy a leguracilosabb régiók ($\log_2(\text{IP}/\text{input})=1$ feletti dúsulás) elsősorban a SAT-12 és a BRSATI szatellitákkal mutatnak átfedést, valamint kisebb mértékű hasonlóság a SAT-13 esetében is megfigyelhető. Megengedőbb küszöbértéknél ($\log_2(\text{IP}/\text{input})=0,25$ és $\log_2(\text{IP}/\text{input})=0,5$ feletti dúsulás) az előbbieket mellett a MOSAT is magasabb pontszámot mutat.



12. ábra: Szatellitákra végzett GIGGLE elemzés eredményei. A különböző szatelliták a UCSC Table Browserből letölthető „repeatmasker” által definiáltak. A combo score a GIGGLE által számított hasonlósági pontszám, ami az intervallumok átfedésének mértékéről és szignifikanciájáról ad információt (ld. Anyagok és módszerek 4.1.4. fejezet). A három diagramon különböző küszöbértékekkel meghatározott uracilosodási mintázatokhoz hasonlítottuk a szatelliták elhelyezkedését.

A szatellitákkal átfedő uracilos régiók arányát témavezetőm konkrétan is meghatározta (13. ábra). A nagyon és közepesen uracilos régiók 80%-a valamilyen szatellitával átfed. Ha a kevésbé uracilos részeket is figyelembe vesszük, akkor is 50%-os átfedés mutatkozik. A GIGGLE elemzésben magasabb pontszámú MOSAT, BRSATI, SAT-12 és SAT-13 szatellitákra külön lebontva látható, hogy ez a négy valóban ki is adja az összes szatellitára vonatkozó átfedést. A leguracilosabb régiók nagy része a SAT-12-vel, illetve a BRSATI-gyel fed át, míg a megengedőbb határral megállapított uracilosodással a MOSAT típusú szatelliták is jelentős átfedést mutatnak. A SAT-13 szatelliták alacsony számban vannak jelen a genomban, emiatt az ezekkel való átfedés mennyiségileg kevésbé számottevő.



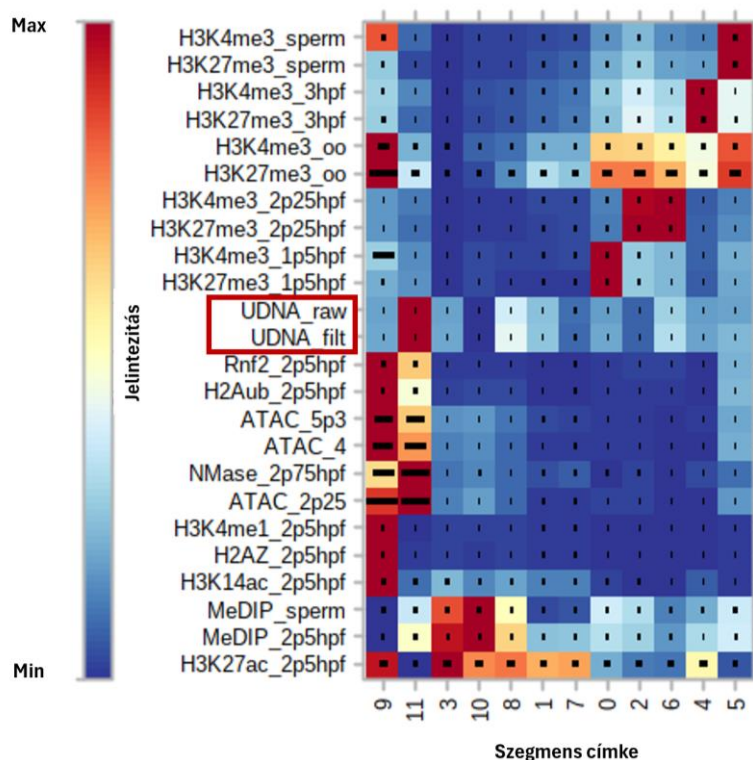
13. ábra: A szatellitákkal átfedő uracilos régiók aránya. Az uracilos régiók különböző küszöbértékekkel meghatározottak, ezeket külön színek jelölik. A jobb szélső blokkban az összes szatellitára vonatkozó eredmények szerepelnek: a nagyon és közepesen uracilos régiók 80%-ban átfednek valamilyen szatellitával. Ettől balra az uracilos régiókkal jelentősebb mértékben átfedő szatelliták láthatók.

A zebraahalban található szatellitákról kevés funkcionális adat ismert. A BRSATI AT-gazdag szatellita, kromoszómánként jellemzően csak néhány (1-3) klaszterben található, és a feltételezett centromerikus pozíciókkal áll együtt [42]. A centromérák lokalizációját a referencia genom legújabb, GRCz11 verziójában sem tudjuk pontosan, habár a centroméra marker szekvenciát már 2000-ben meghatározták [51]. A MOSAT a legkiterjedtebb szatellita, a Zv9 referencia alapján a genom 0,59%-át teszi ki, és egyenletesen oszlik el a kromoszómákon [42]. A funkciójáról annyit lehet tudni, hogy a CATGCTAA ismétlődő szakaszán a korai embrionális stádiumokban érintett ún. „nem CpG” metilációval, mely a ZGA-ig drasztikusan lecseng [52].

5.1.2. Uracileloszlási és funkcionális genomikai profilok genomszegmentációs analízise

A szatellitákra vonatkozó érdekes eredmények után felmerült a kérdés, hogy más kromatinszervező epigenetikai markerek és a tapasztalt uracilos régiók között találunk-e hasonlóságot. Ennek vizsgálatához először széleskörű adatgyűjtést végeztem (ld. Anyagok és módszerek 4.1.2. fejezet, Függelék 16. táblázat). A folyamatos lefedettségi görbéket tartalmazó fájlokon genomszegmentációs analízist csináltunk a Segway programcsomag segítségével. Az eredményként kapott hőtérképen láthatók az algoritmus által meghatározott szegmensek, és

azokon belül az egyes szekvenálási adatokból kapott átlagos jelintenzitás. Minél jobban hasonlítanak a fájlok, annál inkább megegyezik a térképen kapott színezetük. A leguracilosabb régiók a 2,5 hpf körüli stádiumból származó ATAC-seq és az MNase-seq adatokkal mutatnak korrelációt, amelyek a hozzáférhető DNS-szakaszokat térképezik. A szórás jellemzően nagy és az idősebb embriók esetében a hasonlóság lecsökken. Az elemzés során vizsgált stádiumban (1,5 – 3 hpf) epigenetikai markerekre végzett ChIP-seq adatokkal nem volt jelentős korreláció.

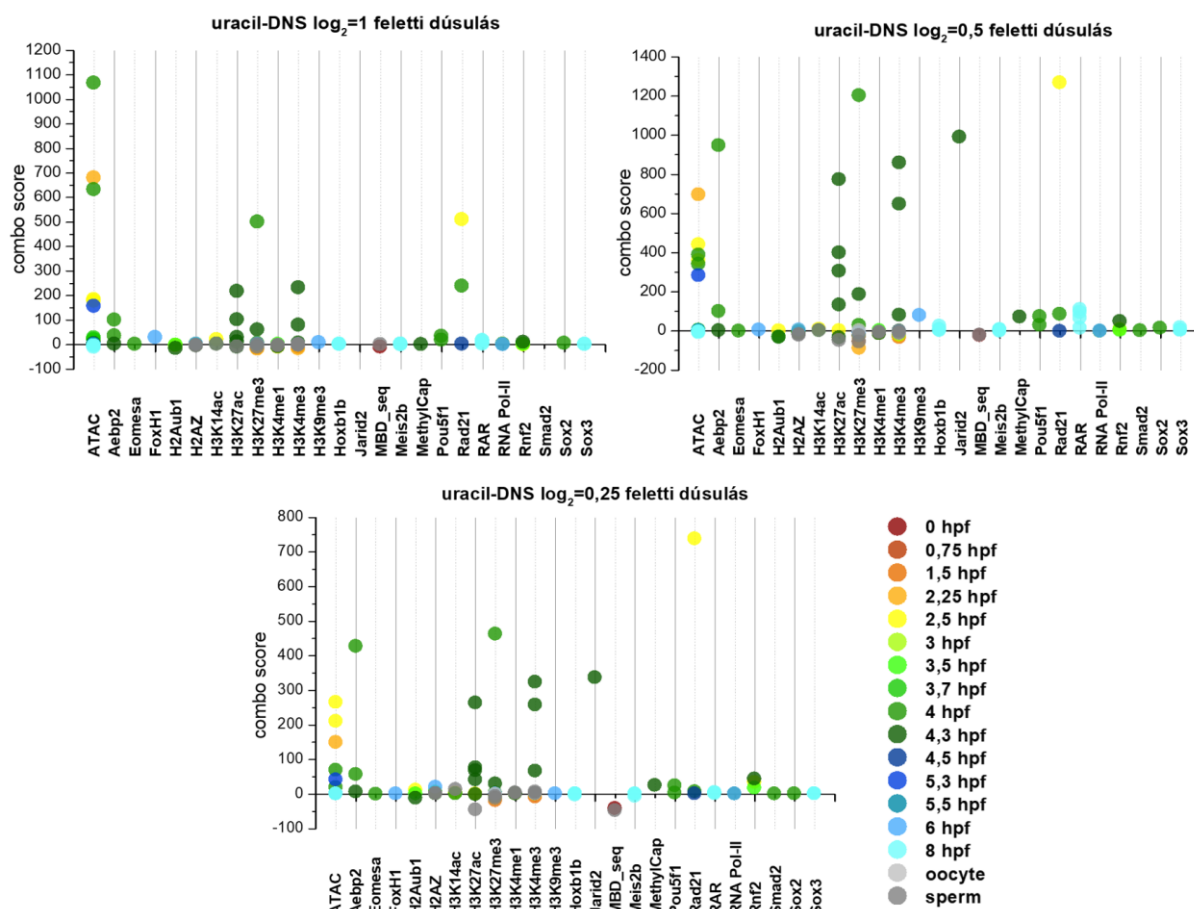


14. ábra: Genomszegmentációs analízis eredményeként kapott hő térkép. Az U-DNA-seq adatok bordó kerettel jelöltek, a „filt” a MAPQ=0 olvasatokat nem tartalmazó, szűrt adatokat jelöli, a „raw” pedig a szűrés nélküli adatfájlt. A 11. címkével ellátott szegmensbe tartoznak az uracilban leginkább dúsult régiók. A színskála (bal oldalon) az IP/input lefedettség arány görbékből számított jelintenzitás soronként normált értékét mutatja. A mezőkben lévő fekete sáv szélessége a jelintenzitás szórását jelzi a szegmensben belül.

5.1.3. Az uracilosodott régiókhoz hasonló genomi mintázatok keresése GIGGLE programmal

A GIGGLE elemzést kiterjedtebb adatkészleten, 0-8 hpf embrionális stádiumból származó adatokon végeztük (15. ábra). Ahogy a genomszegmentációs analízis eredményeinél, úgy itt is látszik, hogy a genomi uracilmintázat és a jobban hozzáférhető DNS-szakaszok között megfigyelhető hasonlóság. Ez elsősorban a 2,25-5,3 hpf stádiumokban jellemző. A 2,5 órás embrióból származó adatok közül egyedül a Rad21 ChIP-seq szerepel magas pontszámmal, viszont az idősebb embrióból származó adatoknál ez már nem figyelhető meg. Az elemzések során több, 4 óra körüli ChIP-seq adattal is jelentős hasonlóságot kaptunk: az Aebp2 és Jarid2

fehérjék, valamint több módosított hiszton (H3K27ac, H3K27me3, H3K4me3) is jellemzően magas pontszámmal szerepel.



15. ábra: A GIGGLE elemzés eredményei funkcionális genomikai adatokra. A három diagramon különböző küszöbértékekkel 2,5 hpf-nél meghatározott uracilosodási mintázatokhoz hasonlítottuk a fejlődési stádiumok szélesebb körére vonatkozó kromatin-hozzáférhetőségi és ChIP-seq jellegű kísérletekből származó genomi intervallum adatokat. Az egyes adatfájloknak a genomi uracilosodási mintázattal való átfedését a combo score jellemzi. A grafikonon az adatokat epigenetikai faktorokként csoportosítottuk, a különböző fejlődési stádiumhoz tartozó adatokat eltérő színnel jelöltük (ld. jobb oldalon).

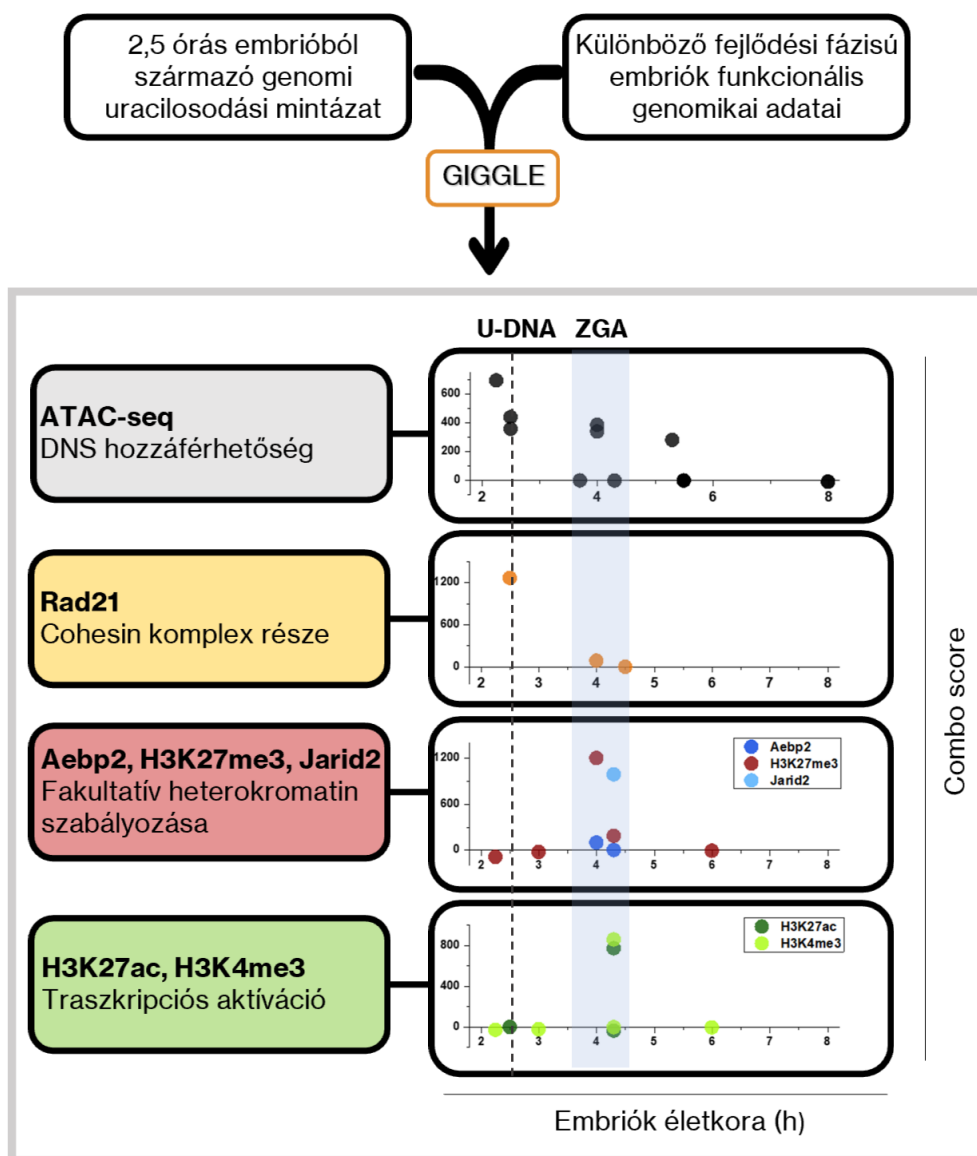
5.1.4. Az összehasonlító bioinformatikai elemzés eredményeinek értelmezése

A 16. ábra a GIGGLE elemzés alapján kiemelkedő adatok időfüggését szemlélteti. A Rad21 fehérje a kohéziós komplex részeként hozzájárul a kromatinszerkezet és a génszabályozás fenntartásához. A ZGA előtti szakaszban jellemzően a BRSATI és a SAT-1 szatellitákkal együtt a pericentromerikus régiókban lokalizálódik. Kimutatták, hogy a Rad21 csendesítése késlelteti a ZGA-t, anélkül, hogy a sejtciklust befolyásolná [53]. A Rad21-re vonatkozó ChIP-seq adatok a 2,5 óránál mutatnak nagyobb átfedést az uracilosodási mintázattal, ez a ZGA környékén már nem jellemző.

Az Aebp2 és a Jarid2 fehérjék a Polycomb repressziós komplex 2 (PRC2) kofaktorai, ezáltal szerepet játszanak a génaktiváció szabályozásában. A PRC2 metilálja a H3 hiszton 27-es lizinjét, ezáltal géncsendesítést eredményezve, ami elengedhetetlen a sejtek

differentiálódásához [54]. Az ezekre vonatkozó, ZGA környéki ChIP-seq adatok mutatnak hasonlóságot a 2,5 órás embrióból származó genomi uracilosodási mintázattal. A H3K27ac és a H3K4me3 a transzkripció aktivációval összefüggésbe hozható epigenetikai markerek [30], melyek közül a 4,3 órás embrióból származó adatok mutatnak korrelációt az uracilosodással.

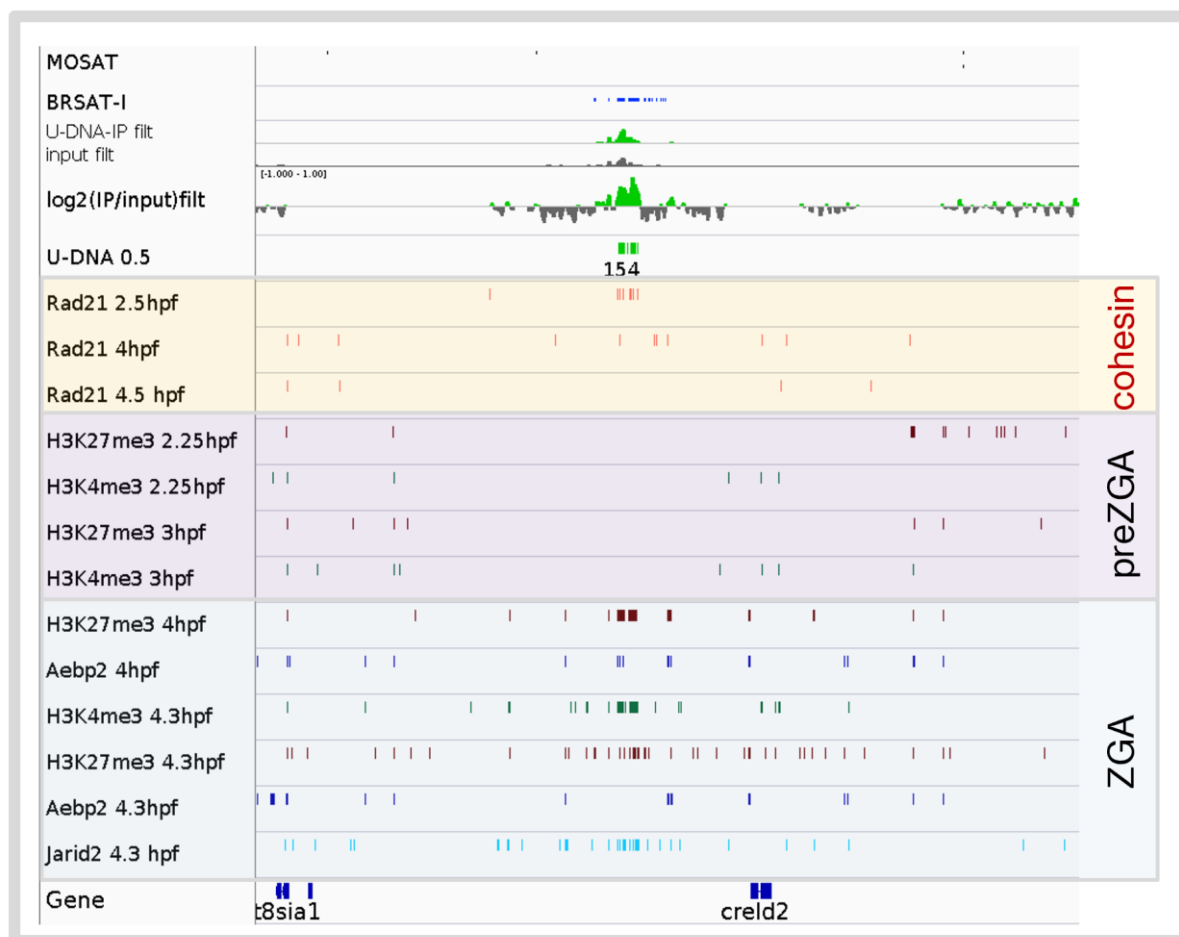
A DNS-hozzáférhetőséget térképező ATAC-seq esetén 2,5 óra környékén a legmagasabb a pontszám, a későbbi szakaszokban lecsökken a hasonlóság.



16. ábra: A GIGGLE elemzés során kiemelkedő pontszámot kapott funkcionális genomikai jellemzők időfüggése. A hasonlósági pontszámok faktoronként színezve és funkcionális csoportokra bontva szerepelnek. Az időtengelyen az U-DNA-Seq adatok (2,5 hpf, szaggatott vonal) és a ZGA (kék mező) embrionális stádiumát kiemeltük.

Ezek az összefüggések szemmel is jól követhetők egyes BSATI szatellitával átfedő erősen uracilos régiók példáján (17. ábra). Ebben a genomi szakaszban a 2,5 órás embrióból származó Rad21 adatok besűrűsödnek, ugyanakkor a transzkripció szabályozásban fontos epigenetikai

markerek és szabályozó fehérjék hiányoznak a zigotikus génaktiváció előtt (preZGA). Ugyanebben a régióban a ZGA környékén ezek az epigenetikai markerek feldúsulnak, a Rad21 fehérje lokalizációja viszont már nem jellemző.



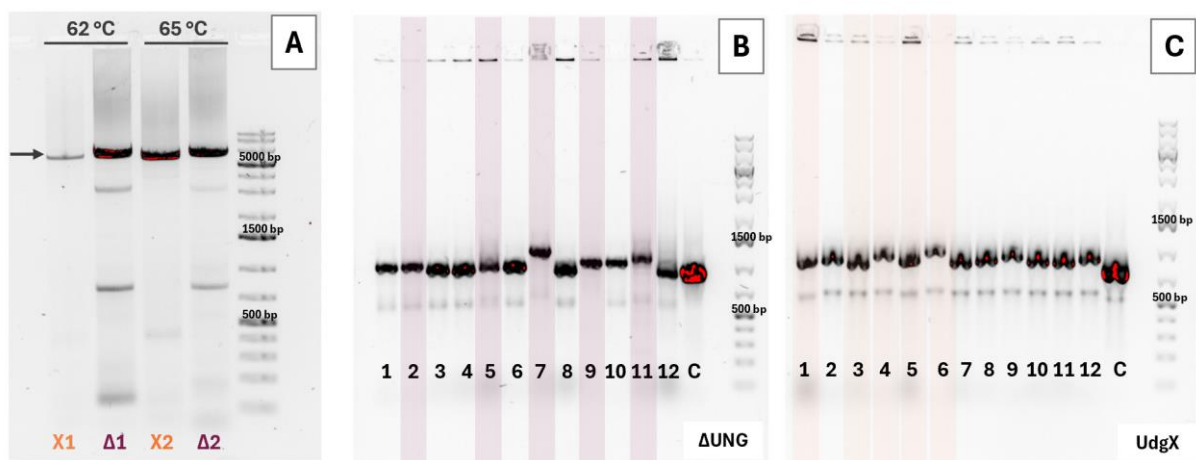
17. ábra: Epigenetikai faktorok mintázatának időbeli változása egy 2,5 hpf stádiumban erősen uracilos genomi régióban. A 4. kromoszóma 26 258 540-27 622 905 nt közti szakaszán IGV-ben megjelenítettük a MOSAT és BRSATI szatellitákat (kék), a genomi uracilosodási profilokat (IP: zöld, input: szürke, log₂(IP/input): zöld-szürke) és régiókat (0.5 log₂(IP/input) küszöbérték, zöld), valamint 3 időponthoz tartozó Rad21 ChIP-seq adatokat (cohesin, sárga mező). Ezek alatt a transzkripció szabályozásban fontos hiszton markerek és faktorok mintázatai láthatók a preZGA és ZGA időszakokra bontva.

Az eredmények alapján elképzelhető, hogy az uracilosodási mintázat valamilyen módon szerepet játszik a zigotikus génaktiváció folyamataiban. A Rad21-ről kimutatták, hogy a korai pericentromérás lokalizációjából egyes korai gének közelébe való átrendeződésével hozzájárul a ZGA-hoz [53]. A genomi uraciltartalom ezzel párhuzamosan, feltehetőleg a BER révén cseng le, ami a száltöréses intermediereken keresztül elősegítheti a kromatínátrendeződéseket.

5.2. Uracil-DNS-t felismerő szenzorok fejlesztése

5.2.1. A kiindulási konstruktok C'-terminálisának bővítése AviTaggel

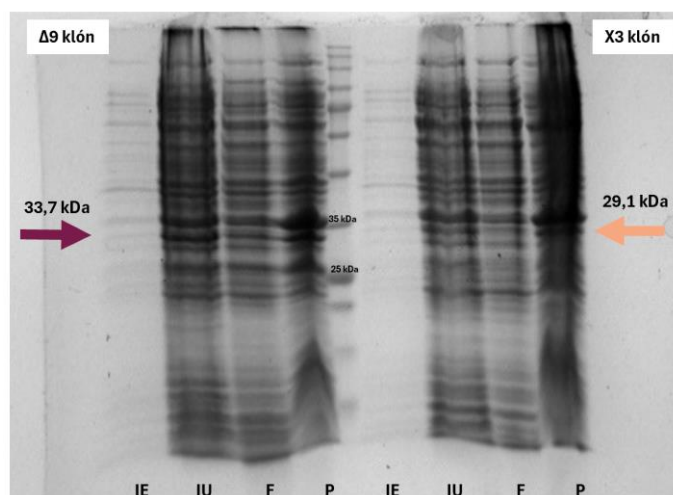
A His-3xFLAG- Δ UNG és His-1xFLAG-UdgX kiindulási konstruktok C'-terminálisához PCR-reakció segítségével illesztettem hozzá az AviTaget úgy, hogy az AviTaget kódoló szekvenciát a teljes plazmidot felszaporító primerekbe terveztem (ld. Anyagok és módszerek 4.3. fejezet). A reakciót két különböző anellációs hőmérséklettel (62 °C és 65 °C) is elvégeztem, amelyből mindkét konstruktnál a 65 °C bizonyult jobbnak (18. ábra A panel). Az UdgX esetén ezen a hőmérsékleten több specifikus termék keletkezett, a Δ UNG szenzornál pedig a primer nem specifikus kitapadása kevésbé érvényesült.



18. ábra: Új AviTaggel bővített konstruktok klónozása. A: Az AviTaget is tartalmazó PCR-termékek agaróz gélképe. Az X az UdgX alapú, a Δ pedig a Δ UNG alapú szenzort jelöli. Az 1-es index a 62 °C-hoz, a 2-es index a 65 °C-os anellációs hőmérsékletre tartozik. B-C: Kolónia PCR eredménye agaróz gélen. A Δ UNG alapú (B) és az UdgX alapú (C) szenzor egyes klónjait számok jelölik. Az AviTaget nem tartalmazó plazmidon készült PCR a „C”-vel jelölt negatív kontroll. Mindkét szenzor esetében 5-5 klónt választottam ki, ezek színes sávokkal jelöltek. A méretjelzőként a Thermo Fisher GeneRuler 1 kb Plus DNS létráját használtam (jobb oldalon).

A PCR-termékeken végzett *in vitro* KLD reakció elegyeivel transzformált sejtekből nőtt telepeket kolónia PCR-reakcióval szűrtem (18. ábra, B és C panelek). Ennek eredménye alapján mindkét szenzortípusból 5-5 klónt választottam ki. Ezeket szekvenálásra küldtem, melyekből 4-4 klón tartalmazta hibátlanul a kívánt konstruktok szekvenciáját.

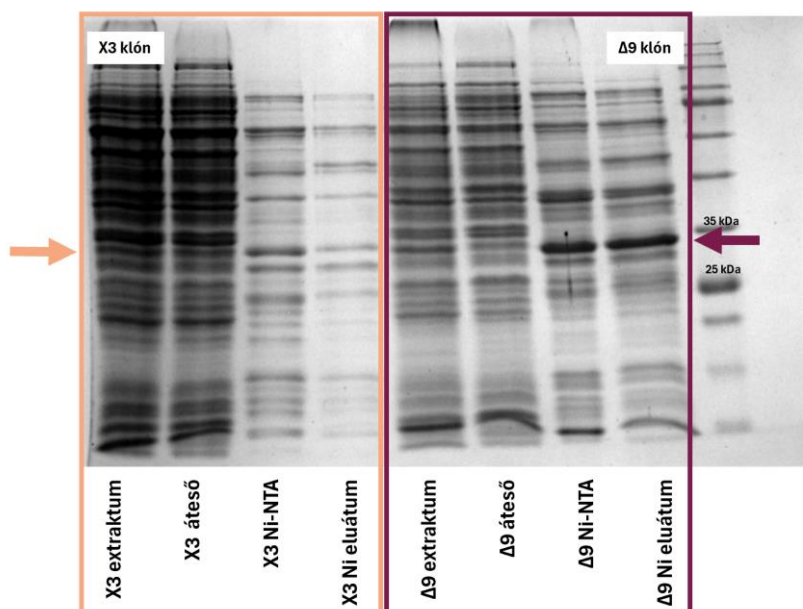
A kiválasztott klónokkal próbaexpressziót végeztem, amit SDS-PAGE-n elemeztem. Ennek eredménye alapján a további lépésekhez a Δ 9 és X3 klónokat választottam ki (19. ábra), ugyan a többi a minta esetében is hasonló eredményeket kaptam. A fehérje expressziója nem volt túl erős, de a termelt mennyiség jelentős része megjelent a felülúszóban, tehát a fehérje oldható.



19. ábra: A $\Delta 9$ és X3 klónok próbaexpressziójának eredménye SDS-PAGE-n. A rövidítések feloldása: IE=indukálás előtt, IU=indukálás utáni expresszió végén, F=felülúszóból, P=pelletből vett minta. A bordó színű nyíl a Δ UNG alapú szenzor, a barackszínű nyíl az UdgX alapú szenzor futási vonalát jelöli, felette a molekulatömegük szerepel. A molekulatömeg-markernek a Grisp GRS Protein Marker MultiColour létrát használtam.

5.2.2. A konstruktok próbatisztítása és *in vitro* biotinilálhatóságának tesztelése

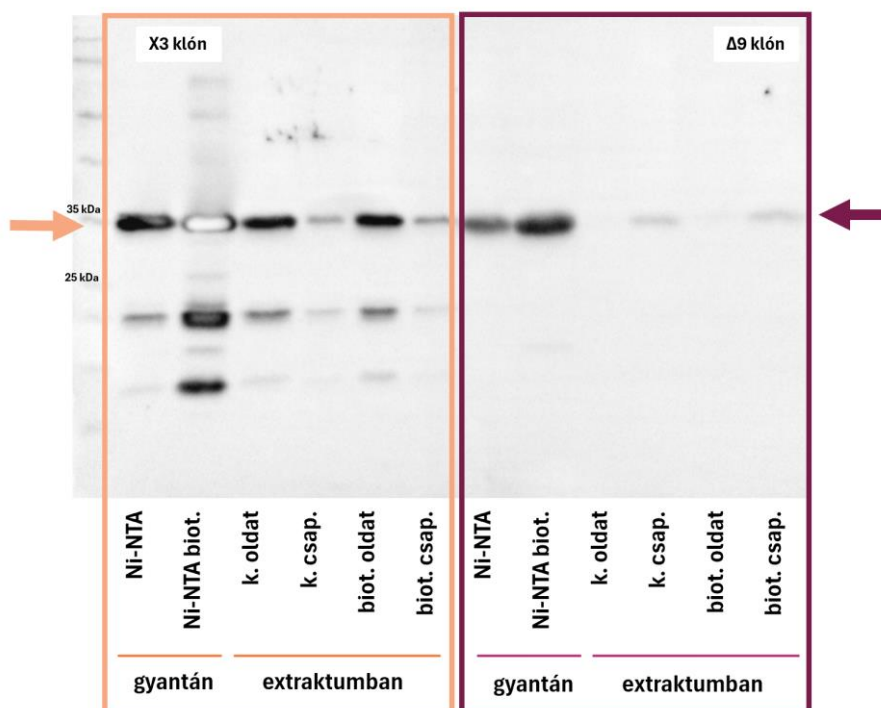
A kiválasztott klónok próbaexpresszióiból egyúttal próbatisztítást is végeztem Ni-NTA oszlopon. Az eredmények (20. ábra) alapján a Δ UNG expressziója megfelelőnek bizonyult: a tisztítás után az eluátumban látványosan erősebb intenzitással volt jelen a kívánt szenzor a többi fehérjéhez képest. Ezzel szemben az UdgX alacsonyabb expressziót mutatott, ezért ennek a termelését a későbbiekben több körülmény kipróbálásával igyekeztem optimalni.



20. ábra: A próbaexpresszióból végzett tisztítás eredménye SDS-PAGE-n. Az egyes szenzorok esetében a nyílak jelölik a várt futási magasságot. Az extraktum sejtlízist követő centrifugálás utáni felülúszót jelenti. A Ni-NTA átesőjéből, az imidazolos eluátumból és a gyanta maradékából („Ni-NTA”) vett mintákat analizáltam.

A próbaexpresszióból készített 5-5 ml extraktummal teszteltem, hogy az *in vitro* biotinilálást milyen körülmények között tudom hatékonyabban elvégezni. Ehhez tisztított BirA fehérjét használtam, melyet a laborban előzőleg Dr. Nyíri Kinga állított elő. A biotinilálás elvileg elvégezhető extraktumban inkubálva, aminek előnye, hogy a fehérje számára természetesebb közeg. Hátránya viszont, hogy az extraktumban levő proteázok is aktívak lehetnek, valamint a nagyobb térfogat miatt több reagenst igényel, ráadásul a hozzáadott ATP-t más enzimek is kontrollálatlanul fogyasztathatják. Másik lehetőség, hogy a Ni-oszlopon feldúsított fehérjét még a gyantához kötött állapotában biotiniláljuk. Ebben az esetben a kisebb térfogat és nagyobb tisztaság előnyt jelent, ahogy az is, hogy a BirA feleslegtől is könnyen megszabadulhatunk. Mindemellett problémát okozhat, hogy sztérikus akadály léphet fel. A harmadik lehetőség, hogy a fehérjét először megtisztítjuk, majd csak ezt követően biotiniláljuk. Ennek hátránya, hogy egy újabb tisztítási lépést igényel, mellyel a BirA eltávolítható.

Az első tesztek alapján úgy tűnt, hogy az *E. coli* sejtekben eleve biotinilálódnak a fehérjék – az UdgX jobban, a Δ UNG kevésbé –, mivel a biotinilálás előtti minták is pozitívak voltak (21. ábra).

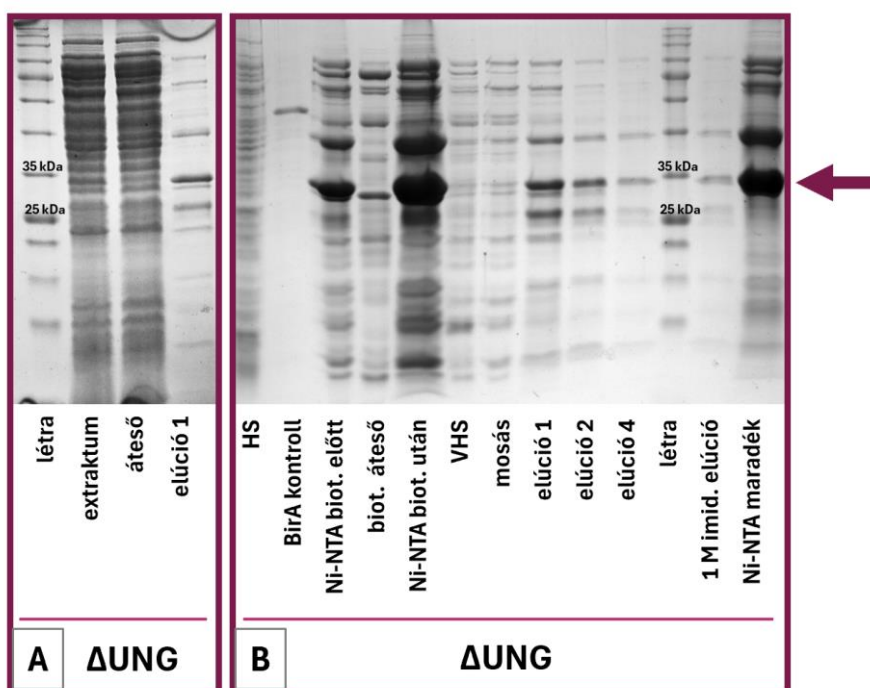


21. ábra: A próbiabiotinilálás eredménye sztreptavidin-HRP-vel előhívott Western bloton. Az oszlopon (gyantán) történő biotinilálás, a biotinilálás előtt (Ni-NTA) és után (Ni-NTA biot.) a gyanta szuszpenzióból vett minták alapján sikeres lehetett. Az extraktumban próbált biotinilálás, a kontroll (k. oldat) és a biotinilált (biot. oldat) minta alapján nem adott hozzá az alap biotiniláltsághoz. Továbbá az inkubálás végére megjelent csapadékok mintái alapján a Δ UNG konstrukt nagymértékben kicsapódott.

Az extraktumban próbált biotinilálás nem volt eredményes, mivel a biotinilált és a kontroll minta között nem mutatkozott különbség. Továbbá az inkubálás során a Δ UNG nagy része csapadékba került. A gyantán próbált biotinilálás mindkét esetben sikeres lehetett, bár a heterogén szuszpenzióból vett minták mennyiségi összehasonlítása nem megbízható.

5.2.3. A His-3xFLAG- Δ UNG-Avi konstrukt termelése és tisztítása

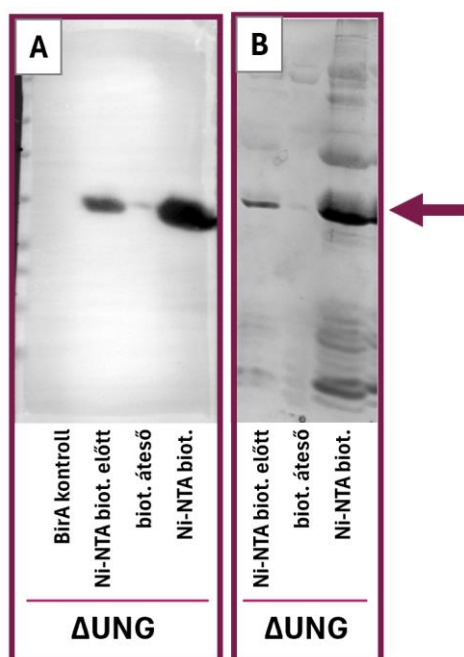
Az összesen 4 liter BL21 *ung-* kultúrában termelt Δ UNG alapú szenzort Ni^{2+} -affinitás kromatográfiával tisztítottam, és a próbaexpresszióból végzett próbabiotinilálás eredményei alapján az oszlopon biotiniláltam. A folyamat során vett mintákat SDS-PAGE-n futtattam (22. ábra). A szenzorfehérje megkötődött az Ni-NTA gyantán, mivel az átesőben elhalványodott a csík az adott pozícióban az extraktumhoz képest, az elúcióban viszont erősödve jelent meg (22. ábra, A panel). Az oszlopon való biotinilálás során nem degradálódott látható mértékben. Az oszlopról 350 mM imidazolt tartalmazó pufferrel csak lassan eluálódott, az 1 M imidazollal eluált frakcióban is jelentős mennyiségű fehérje volt, sőt az elúció után is sok maradt a gyantán (22. ábra, B panel).



22. ábra: A Δ UNG tisztítása és biotinilálása során vett minták SDS-PAGE-n. A Δ UNG megkötődött a gyantán (A panel), az oszlopról lassan eluálódott 350 mM imidazolt tartalmazó pufferrel (elúció 1,2,4), és az 1 M imidazolt tartalmazó frakcióban is jelentős mennyiségű fehérje jelent meg. Az elúció után a gyantából vett mintában (Ni-NTA maradék) kifejezetten sok fehérje maradt (B panel). A HS a magas sótartalmú pufferes mosás, a VHS a nagyon magas sótartalmú pufferes mosás során vett mintákat jelöli.

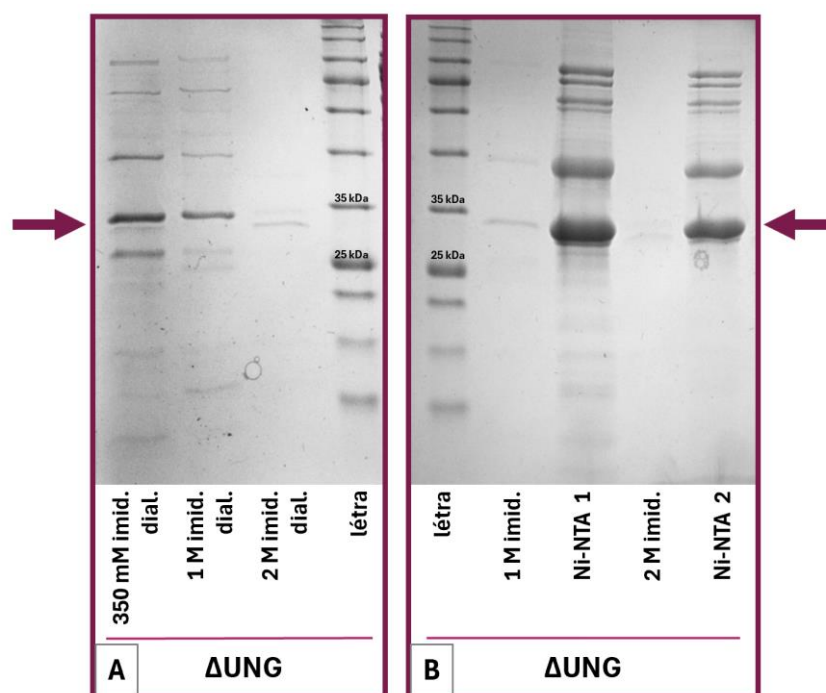
A biotinilálás sikerességének ellenőrzéséhez Western blotot készítettem (23. ábra, A panel). A fehérjét kötő gyantából biotinilálás előtt vett minta is mutat jelet a sztreptavidines festés után,

ami azt jelenti, hogy a Δ UNG már a sejten belül is biotinilálódott. Ehhez képest az oszlopon biotinilálás után erősebbnek tűnik a jel. Figyelembe véve, hogy a heterogén rendszer miatt a gélre felvitt minták mennyisége nehezen kiszámítható, a blottolás után visszamaradt gél is megfestettem (23. ábra, B panel). Látható, hogy gyantából a biotinilálás után vett minta esetében nagyobb mennyiségű fehérje került a gélre, ezért is lehet erősebb a jel a Western bloton, így a biotinilálás sikeressége továbbra is kérdéses.



23. ábra: A Δ UNG szenzor gyantán való biotinilálásának ellenőrzése. A: Sztreptavidin-HRP-vel festett Western blot. A bloton a fehérjét kötő gyantából biotinilálás előtt és után vett mintákat, valamint az átesőt és egy BirA kontrollt futtattam. B: A blottolás után visszamaradt gél PageBlue-val festve.

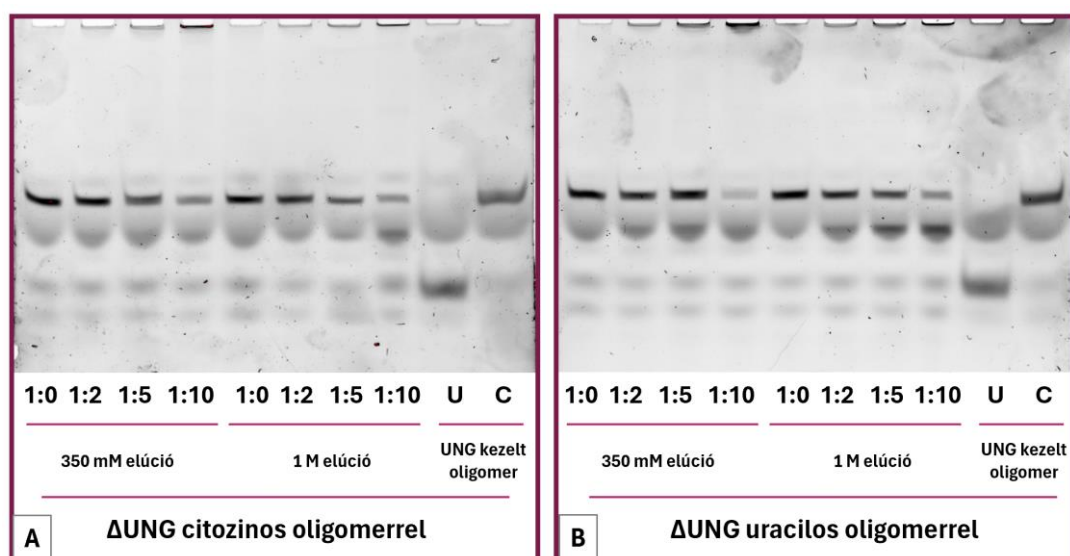
A tisztítás és dializálás után vett minták a 350 mM és 1 M imidazolos elúció esetében is a korábbi tapasztalatokhoz hasonló eredményt adtak: a Δ UNG-on kívül más fehérjék is megjelentek a gélen, azonban láthatóan kisebb mennyiségben (24. ábra, A panel). Mivel az 1 M imidazolos elúció során jelentős mennyiségű fehérje maradt a gyantán, ezért 2 M imidazolt tartalmazó oldattal is megpróbáltam eltávolítani az oszlopról. Ekkor az eluátumban már nem jelent meg a Δ UNG, viszont a gyantából vett mintában továbbra is nagy mennyiségben látható (24. ábra, B panel). Ezt a gyantára csapódott fehérjét sem SDS-sel, sem guanidin-hidrokloriddal nem sikerült eltávolítani.



24. ábra: A Δ UNG tisztítása során vett minták SDS-PAGE-n. A: A tisztított és dializált fehérjeoldatok. A legtöbb fehérje a 350 mM imidazolt tartalmazó pufferrel eluálódott, de az 1 M-os frakcióban is jelentős mennyiségű fehérje volt. A 2 M-os frakcióban már nem jelent meg intakt Δ UNG szenzor. B: Az 1 M és 2 M imidazolos eluátumok, valamint az elúció után a gyantából vett minták. A fehérje elúciója nem javult az imidazol koncentráció emelésével, viszont a gyantán 2 M imidazol után is sok maradt.

A termelt Δ UNG alapú szenzorfehérje töményíthető és stabil, az előző konstruktokhoz hasonló tisztaságú. A 350 mM imidazollal eluált frakcióból összesen 2 ml 3,6 mg/ml koncentrációjú fehérjét sikerült előállítani, azaz körülbelül 7 mg lett a kihozatal. Az 1 M imidazolos elúció pedig további 2 mg (600 μ l; 3,4 mg/ml) fehérjét eredményezett. Ez körülbelül tízszeres mennyiség a csoportban eddig előállított N'-terminális AviTaget tartalmazó Δ UNG szenzorhoz képest. Ugyanakkor megközelítőleg tizede a His-3xFLAG- Δ UNG nem biotinilálható konstruktnak, aminek elsődleges oka lehet, hogy a fehérje az oszlopon biotinilálás közben részben kicsapódott. Meg kell még jegyezni, hogy a jelentősebb szennyezőfehérje-mennyiség miatt az UV spektrummal mért fehérjekoncentráció felülbecsli a valódi értéket.

A szenzor uracil-DNS-t kötő aktivitását EMSA-val teszteltem natív TBE-PAGE-n (25. ábra). Az eredmények a citozint és az uracilt tartalmazó oligonukleotidok estében nagyon hasonlóak. Komplexnek megfelelő eltolódott sáv a gélen nem figyelhető meg, csupán a fehérje nagyobb moláris feleslegénél a zsebekben maradt festődés látható. Azonban ez jel kisebb mértékben a citozinos oligomernél is jelentkezik. Ez alapján az alkalmazott kísérleti elrendezésben specifikus Δ UNG:U-DNS komplex nem mutatható ki egyértelműen.

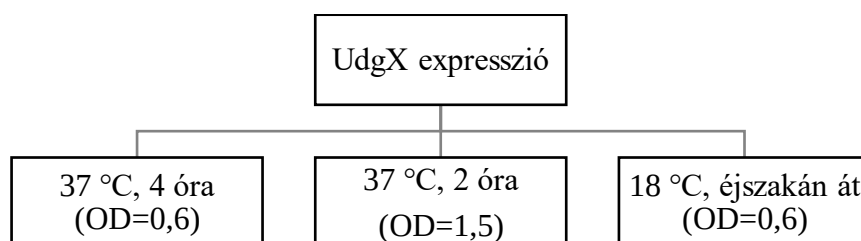


25. ábra: A Δ UNG EMSA tesztelése TBE-PAGE-n. Az alkalmazott oligomerek közül az egyik uracilt, a másik citozint tartalmazott egy meghatározott pozícióban. Kontrollként UNG kezelést végeztem az oligomereken (U: uracilt tartalmazó UNG kezelt oligomer és C: citozint tartalmazó UNG kezelt oligomer). A szenzorfehérjéket különböző moláris feleslegben adtam az oligomerekhez. A: A citozinos oligomerrel végzett teszt. B: Az uracilos oligomerrel tesztelt minták. Komplexnek megfelelő eltolódott sáv egyik esetben sem látható.

5.2.4. A His-1xFLAG-UdgX-Avi konstrukt termelése és tisztítása

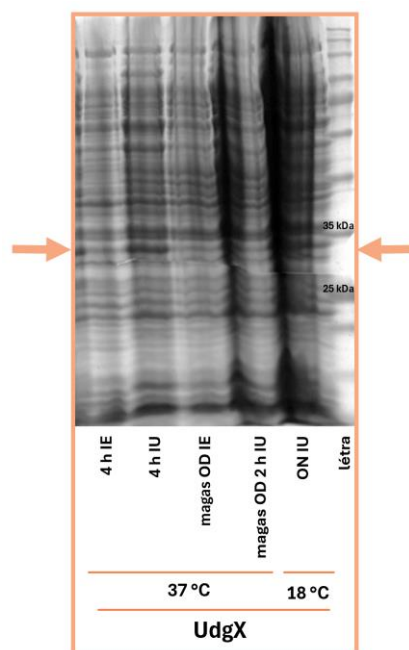
5.2.4.1. Termelés optimalizálása

Az UdgX esetében a kezdeti próbaexpresszió alacsony hozama miatt igyekeztem javítani a termelést más körülmények kipróbálásával, melyek a 26. ábrán láthatók.



26. ábra: Az UdgX expressziójának javításához kipróbált különböző körülmények.

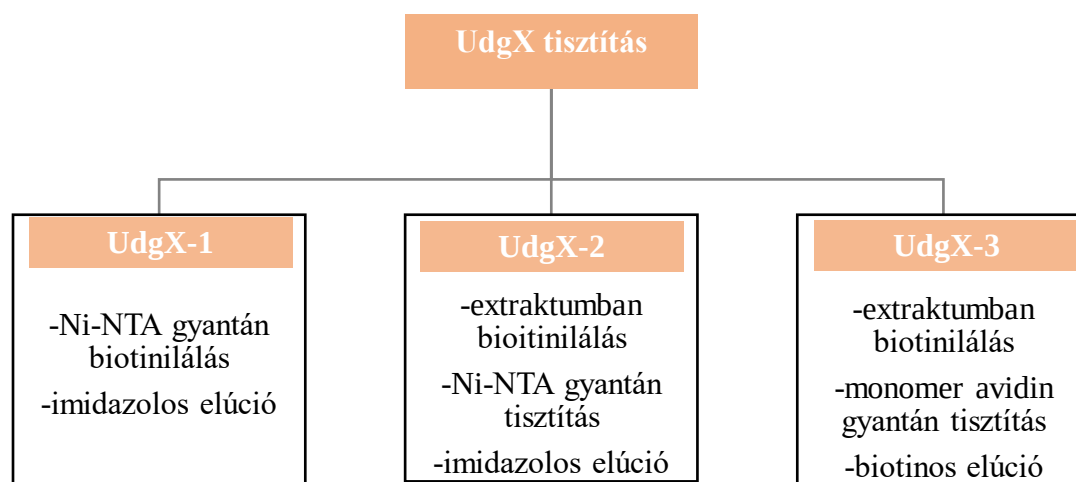
Az eredmények (27. ábra) alapján a 37 °C-on, 4 órás expresszió bizonyult a legerősebbnek, ezért a nagyléptékű termelést ilyen körülmények között végeztem.



27. ábra: Az UdGx próbaexpressziókból vett minták SDS-PAGE-n. Az IE az indukálás előtt, az IU az indukálás után, az expresszió végén vett mintákat jelöli. Az egyes minták esetében feltüntetett idő az expresszió hosszát fejezi ki órában. A „magas OD” 1,5-ös optikai denzitást jelent. Az ON az éjszakán át (overnight) tartó expresszióra utal. A legerősebb expresszió a 37 °C, 4 óra inkubálás esetében látható.

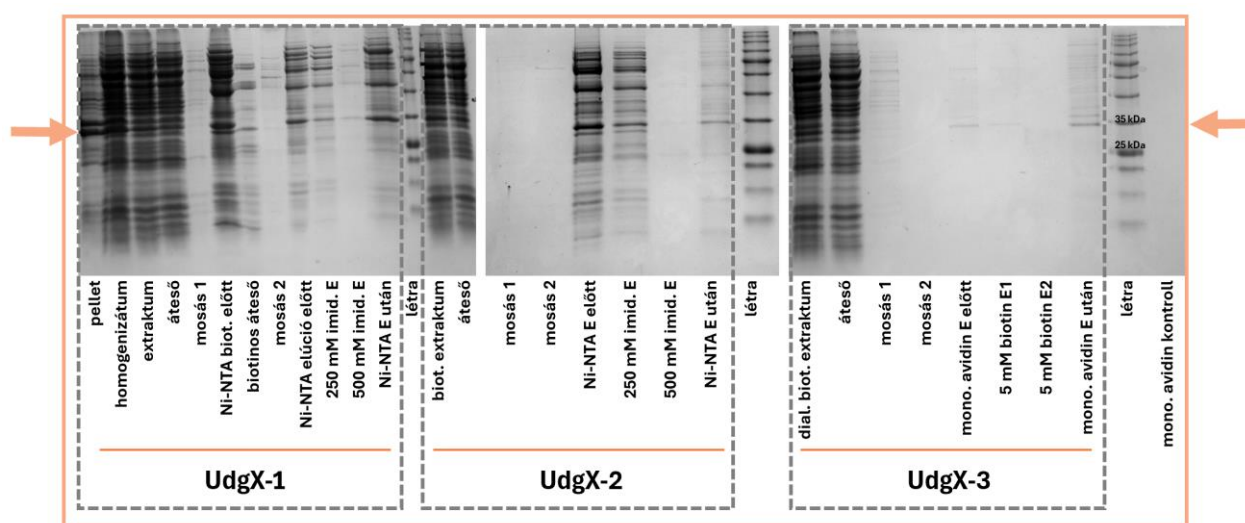
5.2.4.2. Nagyléptékű expresszió és tisztítás

A nagyléptékű (összesen 1 liter kultúra) UdGx expressziót követően a sejtsapadékot feltártam, és az extraktumból a fehérjét három különböző módon tisztítottam (28. ábra). Az UdGx-1 frakciót a Ni-NTA gyantán biotiniláltam, majd imidazolos pufferrel eluáltam. Az UdGx-2 és UdGx-3 esetében a biotinilálást az extraktumban végeztem, majd az UdGx-2 frakciót Ni-NTA gyantán, az UdGx-3-at monomer avidin gyantán tisztítottam. Az utóbbi esetében egy dializálási lépés megelőzte a gyantára töltést a szabad biotin eltávolítása érdekében.



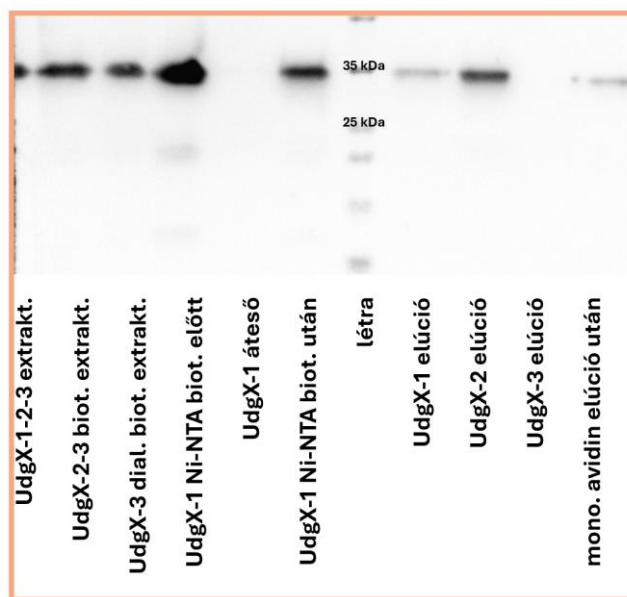
28. ábra: Az UdGx háromféle tisztítási módja.

Az expresszió és tisztítás során vett minták SDS-gélen megfuttatott mintázatai a 29. ábrán láthatók. Az UdgX-1 és UdgX-2 frakciók esetében az UdgX kikötődött a Ni^{2+} -alapú gyantára, majd a 250 mM imidazolos elúcióval a fehérje egy részét sikerült visszanyerni. Az 500 mM imidazolt tartalmazó elúcióval már nem távozott jelentős mennyiségű fehérje, de a gyantán még ezt követően is sok maradt. Ez az UdgX-1 és UdgX-2 esetében is hasonló mértékben figyelhető meg. Az UdgX-3 frakciónál a monomer avidines tisztítás kevésbé volt sikeres: az UdgX nagy része meg sem kötődött a gyantán, hanem az átesőben jelent meg. Az eluátumban alig látható a szenzorfehérje, viszont nagyobb mennyiség maradt a gyantán is.



29. ábra: Az UdgX tisztítása során vett minták SDS-PAGE-n. Az UdgX-1 Ni-NTA gyantán biotinilált és imidazollal eluált frakció. Az UdgX-2 és UdgX-3 extraktumban biotinilált, az UdgX-2 Ni-NTA gyantán, az UdgX-3 monomer avidin agarózon tisztított. Az „E” rövidítés az elúciót jelöli.

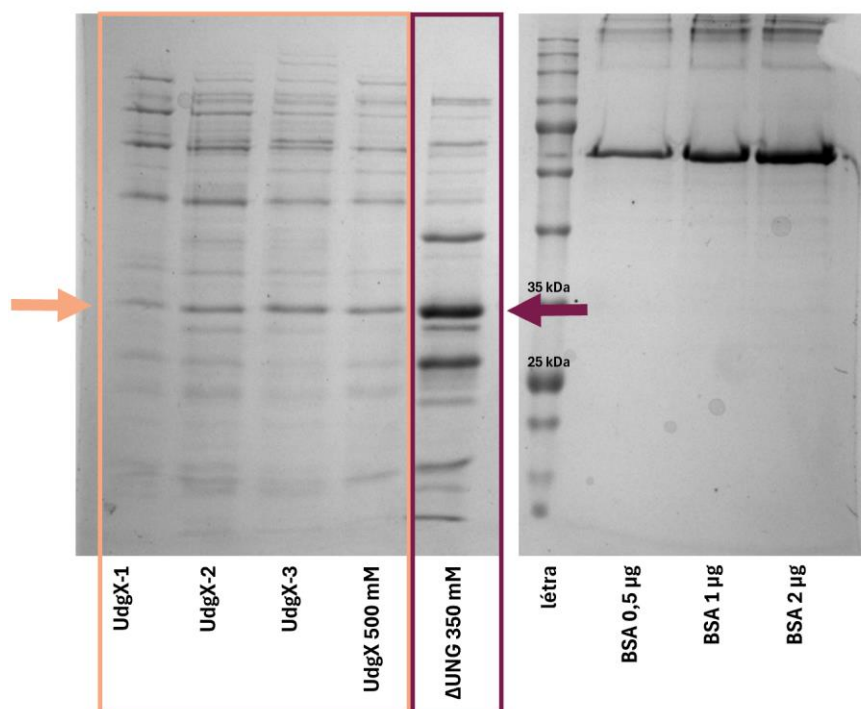
A biotiniláltság vizsgálatához Western blotot készítettem (30. ábra). Az UdgX már a sejten belül is biotinilálódott, és az extraktumban biotinilálás után vett minta sem mutat erősebb jelet egyik frakció esetében sem. Az oszlopon biotinilálás nem bizonyult hatékonynak, bár a heterogén szuszpenzió pipettázási nehézsége miatt a mennyiségi következtetések pontatlanok.



30. ábra: Az UdgX bionitinálásának ellenőrzésére Western bloton. Az UdgX-1 Ni-NTA gyantán biotinilált és imidazollal eluált frakció. Az UdgX-2 és UdgX-3 extraktumban biotinilált, az UdgX-2 Ni-NTA gyantán, az UdgX-3 monomer avidin agarózon tisztított.

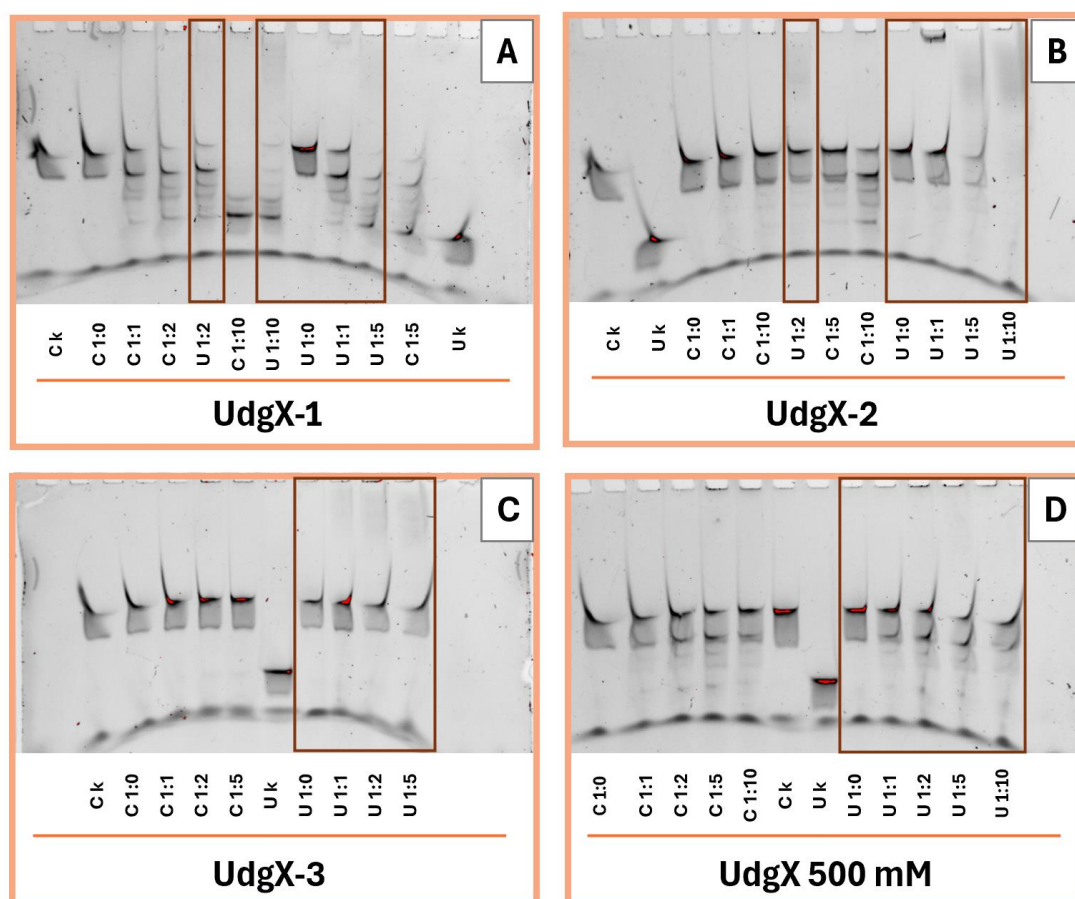
A dializált és töményített mintákat SDS-PAGE-n futtattam (31. ábra). Az UdgX az egyes frakciókban jól azonosíthatóan megjelent, azonban nem sokkal erősebb intenzitással, mint a szennyező fehérjék. Az UdgX-1 frakció a szenzorfehérjéhez viszonyítva több szennyezőt tartalmaz, mint a másik kettő. A Δ UNG-gal összehasonlítva látszik, hogy az UdgX mindhárom preparátuma esetén kevésbé sikerült feldúsítani a szenzorfehérjét. A szenzorfehérje preparátumok töményíthetők voltak és stabilak maradtak.

Az oszlopon biotinilált UdgX-1 összesen 700 μ l 5,3 mg/ml, az extraktumban biotinilált és Ni^{2+} -affinitás kromatográfiával tisztított UdgX-2 1 ml 10 mg/ml, a monomer avidin gyantán tisztított UdgX-3 pedig csupán 700 μ l 0,27 mg/ml névleges koncentrációjú fehérjét eredményezett. Az 500 mM imidazollal eluált nagyon híg frakciókat egyesítettem, és töményítés után 1 ml 0,71 mg/ml fehérjét nyertem ki. Az UV spektrumok A_{260}/A_{280} aránya alapján nukleinsav szennyezést nem tartalmaztak, ugyanakkor a valódi koncentrációk a szennyező fehérjék miatt a mértnél jóval alacsonyabbak lehetnek. Hogy ezt megbecsüljem, a töményített minták mellett ismert mennyiségű BSA hígítási sort is futtattam a gélen (31. ábra). Ez alapján az látható, hogy a Δ UNG szenzor a névleges koncentrációjának mintegy fele, míg az UdgX preparátumokban csupán nyolcada, sőt az UdgX-1 esetén ennél is rosszabb az arány.



31. ábra: A dializált és töményített minták SDS-PAGE-n futtatva. A szenzor preparátumokból az UV spektrummal meghatározott koncentrációk alapján 2-2 μg-ot vittem fel a géltre. A valódi mennyiségek becslése érdekében ismert mennyiségű BSA fehérjét is futtattam.

Az UdGX aktivitásának és az uracilos DNS-hez való kovalens kötésének teszteléséhez denaturáló TBE-PAGE-n EMSA-t futtattam. Az UdGX-1 frakció (32. ábra, A panel) esetében az figyelhető meg, hogy ahogy a hozzáadott szenzorfehérje aránya növekszik, úgy egyre több degradált termék jelenik meg a gélen. Ezt valószínűleg valamilyen nukleáz szennyező okozta. Az uracilos oligomereknél ezek a degradációs termékek magasabban futnak, ez utalhat valamilyen mértékű UdGX kötődésre és/vagy a degradációval szembeni részleges védelemre is. Az UdGX-2 esetén (32. ábra, B panel) csupán a citozinos kontroll oligomer degradálódik a nagy feleslegben adott fehérjés mintában, ugyanilyen aránynál az uracilos oligomer teljesen a magasabb mérettartományokba tolódik, ahol egy definiálatlanabb foltban jelentkezik. Ez az eltolódás koncentrációfüggő és minden bizonnyal a UdGX specifikus és kovalens kötődésére utal. Az UdGX-3 mintánál (32. ábra, C panel) is hasonló, csak gyengébb jelet kapunk, ami az alacsonyabb fehérje koncentrációval magyarázható. Ezesetben nem látunk aspecifikus degradálódást egyik oligomernél sem. Az 500 mM imidazollal eluált UdGX (32. ábra, D panel) minta már nem mutat aktivitást, ugyanakkor némi aspecifikus degradálódás látható mindkét oligomer esetén. Ez összecseng az irodalmi adattal [6], miszerint a nagyobb koncentrációjú imidazol hatására az UdGX U-DNS-t kötő aktivitása drasztikusan lecsökken.



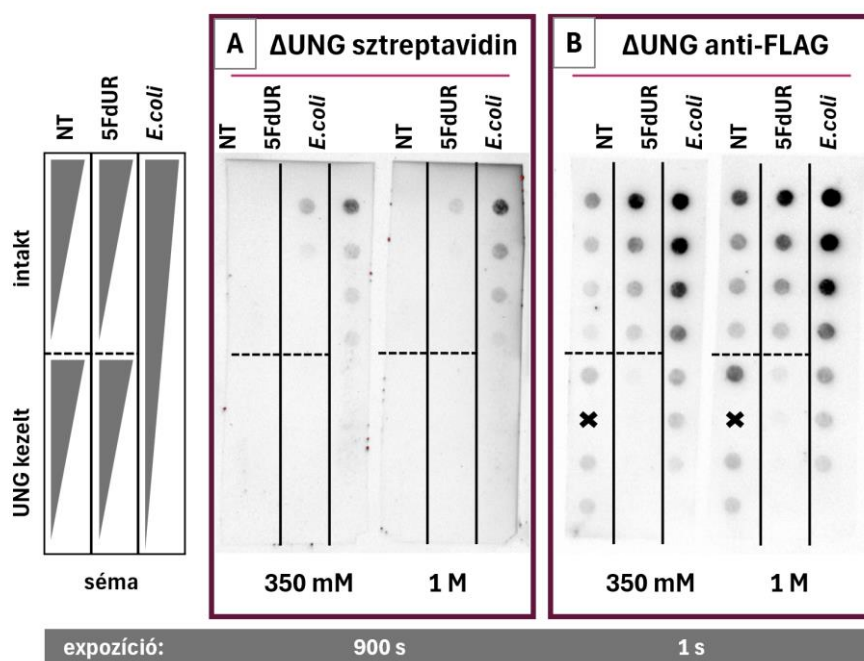
32. ábra: Az UdgX EMSA tesztelése TBE-PAGE-n. Az alkalmazott Cy3 fluoroforral 5'-végen jelölt 30-mer oligonukleotidok közül az egyik uracilt (U), a másik citozint tartalmazott (C) egy meghatározott pozícióban. Kontrollként UNG kezelést végeztem az oligomereken („U k” és „C k” jelölés). A szenzorfehérjéket növekvő moláris feleslegben adtam az oligomerekhez. A: Az UdgX-1 Ni-NTA gyantán biotinilált és imidazollal eluált frakció. B: Az UdgX-2 extraktumban biotinilált, Ni-NTA gyantán tisztított preparátum. C: Az UdgX-3 extraktumban biotinilált, monomer avidin agarózon tisztított preparátum. D: Az UdgX 500 mM az oszlopon maradt fehérjéknek 500 mM imidazolt tartalmazó elúciójából származó frakció.

5.2.5. A szenzorfehérjék alkalmazhatóságának tesztelése csepp bloton

A szenzorfehérjék alkalmazhatóságának vizsgálatához csepp blot méréseket végeztem. Minden blotnál két HCT116 sejtvonalból származó genomi DNS-mintát használtam: egy gyógyszeres kezelés nélküli, alacsony uraciltartartalmút (NT minta), valamint egy 5FdUR kezelt, magas uraciltartalmút (5FdUR minta). A mintákat intakt formában és *in vitro* UNG kezelés után is felvittem a blotra, ezzel ellenőrizve a szenzorok specificitását az uracilos DNS-re.

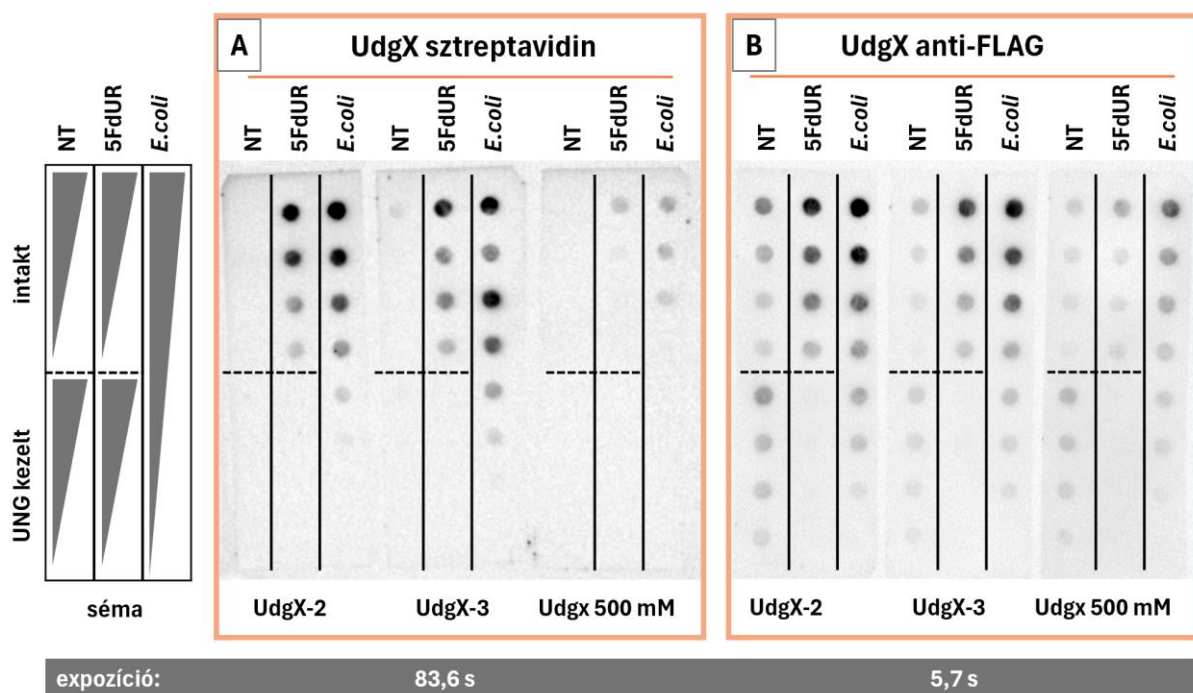
A Δ UNG esetében (33. ábra) a 350 mM és 1 M imidazollal eluált szenzorok nagyon hasonló eredményt mutatnak. Szembetűnő, hogy a sztreptavidines festésnél még 900 másodperces exponálás után is jóval gyengébb a jelintenzitás, mint ami az anti-FLAG antitest esetében már 1 másodpercnél tapasztalható. Az anti-FLAG antitest festésnél látható, hogy a tízszeres mennyiségben felvitt alacsony uraciltartalmú NT minta UNG kezelés után is ad jelet. Ebből arra lehet következtetni, hogy az anti-FLAG antitest vagy a HRP-konjugált másodlagos

ellenanyag aspecifikus kötődést is mutat a nagyobb mennyiségben felvitt DNS-cseppekhez. A sztreptavidines festésnél ilyen jellegű probléma nem jelentkezik, bár a gyenge jel miatt nem kizárható. Az, hogy ugyanazon szenzorral fejlesztett blot a FLAG-címkén (1 másodperces expozíciós idővel), illetve a biotin-jelölésen keresztül (900 másodperes expozícióval) detektálva ilyen mértékben különböző jelet ad, a szenzor alacsony fokú biotinizáltságára is utalhat.



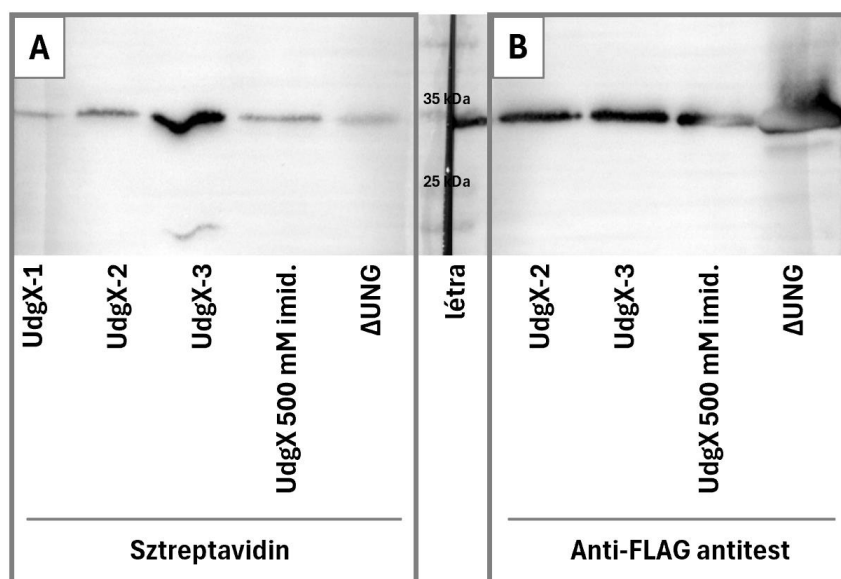
33. ábra: A Δ UNG szenzor alkalmazhatóságának tesztelése csepp bloton. Minden blot esetében a séma (bal oldalon) szerint az első oszlopban a nem gyógyszerkezelt HCT116 sejtvonalból (NT minta), a második oszlopban az 5FdUR kezelt HCT116 UGI sejtvonalból (5FdUR minta) származó gDNS szerepel. A mintákból 4 pontos, felező hígítást készítettem, az NT esetén 0,5 μ g, az 5FdUR esetén 0,05 μ g kezdő mennyiséggel. A szaggatott vonal felett az intakt, alatta pedig az *in vitro* UNG kezelést követően felvitt minták szerepelnek. Az utolsó oszlopban szintén felező hígítással a *dut-ung-* *E. coli* DNS-ből készített kalibrációs sor látható. A: A sztreptavidin-HRP-vel 900 s expozíciós idővel előhívott blotok. B: Az anti-FLAG elsődleges és HRP-konjugált anti-egér Ig másodlagos antitesttel 1 s expozíciós idővel előhívott blotok. Mindkét esetben a 350 mM, valamint az 1 M imidazollal eluált szenzorfehérjéket teszteltem. A fekete x technikai hibából eredő hiányzó mérési pontot jelöl.

Az UdgX alapú szenzor esetében a TBE-PAGE-n végzett kötődési esszék tanulságai alapján az UdgX-2, UdgX-3 és az UdgX-500 preparátumokból származó szenzorokat teszteltem csepp bloton is (34. ábra). Az UdgX-2 és -3 hasonló jelet adnak, azonban az 500 mM imidazollal eluált fehérjénél már jelentős jelcsökkenés figyelhető meg. Ahogy a Δ UNG-nál, úgy az UdgX-nél is látszik, hogy a sztreptavidines festés hosszabb expozíciót igényel, viszont ebben az esetben sokkal kisebb a különbség a két típusú előhívással kapott intenzitás között. Szintén megfigyelhető az anti-FLAG antitesttes festésnél az aspecifikus jel (az UNG-gal emésztett minta jele nem tűnik el), ami a sztreptavidines festés esetében nem jelentkezik. Ez alátámasztja, hogy a nagy DNS-mennyiségeknél tapasztalható aspecifikus jelet nem a szenzorfehérje, hanem valószínűleg valamelyik ellenanyag okozza.



34. ábra: Az UdgX szenzor alkalmazhatóságának tesztelése csepp bloton. A membránra felvitt minták és elrendezésük a 33. ábrán feltüntetettellel megegyezik. A: A sztreptavidin-HRP-vel 83,6 s expozíciós idővel előhívott blotok. B: Az anti-FLAG antitesttel 5,7 s expozíciós idővel előhívott blotok. Mindkét esetben az UdgX-2 (extraktumban biotinilált, Ni-NTA gyantán tisztított), UdgX-3 (extraktumban biotinilált, monomer avidin agarózon tisztított) frakciókat és az 500 mM imidazollal eluált szenzorfehérjéket is teszteltem.

A Δ UNG sztreptavidines festése után jelentősen nagyobb expozíciós időre volt szükség, mint az UdgX esetében. Erre adhat magyarázatot a 35. ábra, amely jól szemlélteti, hogy a Δ UNG kevésbé biotinilálódott, mint az UdgX preparátumok. Az anti-FLAG antitesttel festett Western blot alapján látszik, hogy a Δ UNG-ból a blotra felvitt fehérjemennyiség nagyobb, a sztreptavidinnel való festés esetében mégis kisebb a jelintenzitás az UdgX frakciókhoz képest. A legnagyobb biotiniláltságot az UdgX-3, azaz a monomer avidinen tisztított frakció mutatja, ami várható volt, hiszen ez az oszlop a biotiniláltságot kihasználva dúsítja a fehérjét. Ez az erősebb biotiniláltság kompenzálhatja a TBE-PAGE-n tapasztalt gyengébb látszólagos kötődést, így a csepp blotos eredmények esetén az UdgX-3 végül nem mutatkozik gyengébbnek.



35. ábra: A különböző szenzorfehérje preparátumok Western bloton. A: sztreptavidin-HRP-vel, B: anti-FLAG antitesttel előhívott blot. A két membránra egyforma mennyiséget vittem fel az egyes fehérjékből, mely az UV spektrum alapján számítva mind az öt preparátumra 1 µg.

5.2.6. A szenzorfehérjék előállításának és jellemzésének tanulságai

A két tervezett szenzor klónozását és bakteriális expresszióját megvalósítottam. A tisztításuk a korábbi jó stabilitású konstrukciókhoz hasonló mértékben sikerült. Az *in vitro* biotinizálásukra több elrendezést is kipróbáltam tisztított BirA fehérje alkalmazásával, ami nem vezetett igazán sikerre. Itt mindkét szenzor esetében a jövőben frissen termelt rekombináns BirA-t tartalmazó extraktumok összekeverésével fogok próbálkozni, amiről a csoportban jó tapasztalatok vannak. A funkcionális jellemzésből az látható, hogy az UdgX preparátumom valóban képes specifikus kovalens komplexet képezni a modellként használt egyszálú uracilos nukleinsavval (egyetlen uracilt tartalmazó 30-mer oligonukleotid), míg a reverzibilisen kötődő ΔUNG szenzor U-DNS komplexeit natív EMSA-val sem tudtam megbízhatóan detektálni. Ez alapján minden okunk megvan arra, hogy az UdgX-ben egy szenzitívebb és robusztusabb U-DNS szenzorfehérjét reméljünk, amennyiben a biotinizáltságát tovább tudjuk optimalizálni. A csepp blot tesztekben az is megerősítést nyert, hogy a nagy mennyiségben felvitt, alacsony uraciltartalmú minták specifikus mérésére az egyszerűbb biotin-sztreptavidin rendszer alkalmasabb, mint FLAG-címke két antitestből álló, nagyobb komplex révén való detektálása.

		ΔUNG	UdgX		
Expresszió		OD=0,6; 18 °C, éjszakán át	Próbaexpresszió körülményei		
			OD=1,5; 37 °C, 2 óra OD=0,6; 18 °C, éjszakán át OD=0,6; 37 °C, 4 óra → nagyléptékű expresszió		
Feltárás					
Biotinilálás		Ni-NTA gyantán Sikeressége kérdéses, kicsapódott a fehérje.	UdgX-1	UdgX-2	UdgX-3
			Ni-NTA gyantán Nem hatékony, kicsapódott a fehérje.	Extraktumban Nem tett hozzá a sejten belüli biotiniláláshoz.	
Elválasztás		Ni-NTA oszlopon Megkötődött a fehérje.	Ni-NTA oszlopon Megkötődött a fehérje.	Monomer avidin agarózon A fehérje nagy része nem kötődött meg.	
Elúció		350 mM imidazol Lassú elúció. 1 M imidazol Jelentős fehérjemennyiség és sok maradt a gyantán is.	250 mM imidazol Fehérje egy része eluálódott. 500 mM imidazol Ebben már kevés fehérje távozott, viszont így is nagy mennyiség a gyantán maradt.	5 mM biotin Az eluátumban alig volt fehérje, nagyobb mennyiség a gyantán maradt.	
Dialízis és töményítés					
EMSA		→ Specifikus ΔUNG:U-DNS komplex nem mutatható ki egyértelműen.		→ Van kovalens komplex, ami részben véd a degradáció ellen is.	
Csepp blot		→ Az alacsony biotiniláltság miatt a szteptavidinnel való festés gyenge jelet ad. → A biotin-szteptavidin rendszer specifikusabbnak tűnik, mint az anti-FLAG alapú.		→ A ΔUNG-hoz képest jobb a biotiniláltsága. → A biotin-szteptavidin rendszer specifikusabb jelet ad, mint az anti-FLAG alapú. → Az 500 mM imidazzal eluált frakcióban jelentős a kötési aktivitás csökkenése.	

36. ábra: A két szenzorfehérje előállításának és jellemzésének összegzése.

6. Összegzés

A kutatásom alapvető célja az volt, hogy hozzájáruljak a DNS-ben található uracil biológiai szerepének megértéséhez. Ehhez zebrahal embriókból származó adatokkal bioinformatikai elemzéseket végeztem. A zebrahal mint modellorganizmus különösen érdekes, mivel a kutatócsoportunkban végzett kísérletek alapján a korai embrió magas genomi uracilszintje valamilyen módon szerepet játszhat az egyedfejlődésének folyamatiban. Elemzéseim során funkcionális genomikai adatokat vettem össze a csoportunk által meghatározott genomi uracilosodási mintázattal. Emellett a kutatásom második felében egy érzékeny uracilfelismerő szenzor fejlesztésén dolgoztam, amely a jövőben lehetőséget nyújthat a szenzorok hatékonyabb és szélesebb körű alkalmazására. Többek között alkalmas eszköz lehet a zebrahal embrió modellben részletesebb *in vitro* és *in vivo* kísérletekhez is.

A bioinformatikai elemzések alapján az uracilban dúsult régiók jelentős átfedést mutatnak bizonyos szatellitákkal, amelyek közül kiemelkedő a feltételezett centromérikus pozíciókkal kolokalizáló BRSATI [42], valamint a korai embrionális fejlődés során a „nem CpG” metilációban érintett MOSAT [52]. Ugyan ezen szatelliták pontos szerepét nem ismerjük, a ZGA során feltehetőleg kromatinszerveződési tényezőként jelennek meg. A funkcionális genomikai adatok közül a 2,5 órás embrionális stádiumban – amelyből a genomi uracilosodási mintázatunk is származik – a DNS hozzáférhetőségét térképező ATAC-seq és a kohéziós komplex részét képező Rad21 fehérje adataival volt jelentős hasonlóság, ami az idősebb embriókból származó adatok esetében már kisebb mértékben, vagy egyáltalán nem figyelhető meg. A Rad21 korai pericentromérás lokalizációjából bizonyos gének közelébe való átrendeződésével hozzájárul a zigotikus génaktivációhoz [53]. A genomi uraciltartalom ezzel párhuzamosan cseng le, ami az uracil kivágása révén keletkező száltöréses intermediereken keresztül elősegítheti a kromatinátrendeződéseket. A ZGA környékén (~4 órás embrió) több transzkripció szabályozásban is szerepet játszó fehérjével és epigenetikai markerrel is mutatkozott korreláció. Ezek alapján feltehető, hogy az uracilosodási mintázat valamilyen módon szerepet játszik a zigotikus génaktiváció folyamataiban. Annak érdekében, hogy pontosabb képet kapjunk az uracilosodási mintázat és a funkcionális genomikai adatok közötti átfedésekről, az eddigi eredmények alapján érdemes lenne a nagyobb számítási igényű genomszegmentációs analízist kiterjeszteni a későbbi stádiumokból származó releváns ChIP-seq adatokra is, ami a jövőbeni kutatási terveim részét képezi.

A kutatásom második felében a csoportunkban eddig használt uracilfelismerő szenzorokat fejlesztettem tovább. Sikeresen előállítottam olyan Δ UNG és UdgX alapú szenzorfehérjéket, amelyek az eddigi konstrukciókhoz képest egy C'-terminális AviTaggel bővültek. A termelés kihozatala jelentősen meghaladta a laborunkban eddig előállított biotininálható szenzorokét, azonban a hasonló, de AviTaget nem tartalmazó konstrukcióhoz képest alulmaradt. A töményíthetőség és stabilitás tekintetében viszont megfeleltek az elvárásunknak. Denaturáló gélen végzett EMSA alapján, az UdgX konstrukció az uracilos oligomerrel egyértelműen kovalens komplexet képzett. Az AviTag elviekben lehetőséget teremt az *in vitro* biotininálásra, ami egyrészt lehetővé tehetné a monomer avidinen való tisztítást, másrészt pedig a biotininált fehérje sztreptavidinnel festhető, amivel a szenzor érzékenysége javítható. A kísérletek során egyelőre a monomer avidin agarózszalagos tisztítás, és a szenzorfehérjék biotininálása nem bizonyult elég hatékonynak. A Ni^{2+} -affinitás kromatográfiával tisztított, részben biotininált szenzorokkal viszont sikerült csepp bloton szisztematikusan összehasonlítani a két szenzor, illetve a két detektálási módozat specificitását és érzékenységét. Ezek alapján a sztreptavidines festés specifikusabb jelet ad az uracilos DNS-re, mint a korábban alkalmazott anti-FLAG antitesttel végzett, ami különösen az alacsony uraciltartalmú minták esetén kritikus. Ezért a jövőre vonatkozóan a szenzorok hatékonyabb biotininálási eljárásának kifejlesztését tűztük ki célul.

Irodalomjegyzék

- [1] H. E. Krokan and M. Bjørås, “Base excision repair,” *Cold Spring Harb. Perspect. Biol.*, vol. 5, no. 4, p. a012583, Apr. 2013, doi: 10.1101/cshperspect.a012583.
- [2] S. U. Siriwardena, K. Chen, and A. S. Bhagwat, “Functions and Malfunctions of Mammalian DNA-Cytosine Deaminases,” *Chem. Rev.*, vol. 116, no. 20, pp. 12688–12710, Oct. 2016, doi: 10.1021/acs.chemrev.6b00296.
- [3] V. Muha *et al.*, “Uracil-containing DNA in *Drosophila*: stability, stage-specific accumulation, and developmental involvement,” *PLoS Genet.*, vol. 8, no. 6, p. e1002738, Jun. 2012, doi: 10.1371/journal.pgen.1002738.
- [4] H. L. Pálincás *et al.*, “Genome-wide alterations of uracil distribution patterns in human DNA upon chemotherapeutic treatments,” *eLife*, vol. 9, Sep. 2020, doi: 10.7554/eLife.60498.
- [5] B. G. Vértessy, “Search and destroy: versatile proteins offer unique structural solutions against uracil in DNA,” in *Amino acids, peptides and proteins: volume 43*, vol. 43, M. Ryadnov and F. Hudecz, Eds. Cambridge: Royal Society of Chemistry, 2019, pp. 1–13.
- [6] J. A. Stewart, G. Schauer, and A. S. Bhagwat, “Visualization of uracils created by APOBEC3A using UdGX shows colocalization with RPA at stalled replication forks,” *Nucleic Acids Res.*, vol. 48, no. 20, p. e118, Nov. 2020, doi: 10.1093/nar/gkaa845.
- [7] G. Chen, S. Mosier, C. D. Gocke, M.-T. Lin, and J. R. Eshleman, “Cytosine deamination is a major cause of baseline noise in next-generation sequencing,” *Mol. Diagn. Ther.*, vol. 18, no. 5, pp. 587–593, Oct. 2014, doi: 10.1007/s40291-014-0115-2.
- [8] A. S. Bhagwat, “DNA-cytosine deaminases: from antibody maturation to antiviral defense,” *DNA Repair (Amst)*, vol. 3, no. 1, pp. 85–89, Jan. 2004, doi: 10.1016/j.dnarep.2003.09.008.
- [9] H. Hayatsu, “Bisulfite modification of nucleic acids and their constituents,” *Prog. Nucleic Acid Res. Mol. Biol.*, vol. 16, pp. 75–124, 1976.
- [10] S. Stavrou and S. R. Ross, “APOBEC3 proteins in viral immunity,” *J. Immunol.*, vol. 195, no. 10, pp. 4565–4570, Nov. 2015, doi: 10.4049/jimmunol.1501504.
- [11] A. Beke, “Uracil a DNS-ben: hiba vagy jel? A dUTPáz szabályozása és egy újonnan azonosított U-DNS nukleáz a *Drosophila melanogaster* fejlődésében,” Doctoral dissertation, 2007.
- [12] J. Wardle *et al.*, “Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya,” *Nucleic Acids Res.*, vol. 36, no. 3, pp. 705–711, Feb. 2008, doi: 10.1093/nar/gkm1023.

- [13] R. Then, “Thymidylate Synthase,” in *xPharm: The Comprehensive Pharmacology Reference*, Elsevier, 2007, pp. 1–10.
- [14] D. K. Rogstad, P. Liu, A. Burdzy, S. S. Lin, and L. C. Sowers, “Endogenous DNA lesions can inhibit the binding of the AP-1 (c-Jun) transcription factor.,” *Biochemistry*, vol. 41, no. 25, pp. 8093–8102, Jun. 2002, doi: 10.1021/bi012180a.
- [15] T. Seno, D. Ayusawa, K. Shimizu, H. Koyama, K. Takeishi, and T. Hori, “Thymineless death and genetic events in mammalian cells.,” *Basic Life Sci.*, vol. 31, pp. 241–263, 1985, doi: 10.1007/978-1-4613-2449-2_15.
- [16] R. D. Ladner, “The role of dUTPase and uracil-DNA repair in cancer chemotherapy.,” *Curr. Protein Pept. Sci.*, vol. 2, no. 4, pp. 361–370, Dec. 2001, doi: 10.2174/1389203013380991.
- [17] T. Visnes, M. Akbari, L. Hagen, G. Slupphaug, and H. E. Krokan, “The rate of base excision repair of uracil is controlled by the initiating glycosylase.,” *DNA Repair (Amst)*, vol. 7, no. 11, pp. 1869–1881, Nov. 2008, doi: 10.1016/j.dnarep.2008.07.012.
- [18] Y. Kubota, R. A. Nash, A. Klungland, P. Schär, D. E. Barnes, and T. Lindahl, “Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase beta and the XRCC1 protein.,” *EMBO J.*, vol. 15, no. 23, pp. 6662–6670, Dec. 1996, doi: 10.1002/j.1460-2075.1996.tb01056.x.
- [19] J. Woodrick *et al.*, “A new sub-pathway of long-patch base excision repair involving 5’ gap formation.,” *EMBO J.*, vol. 36, no. 11, pp. 1605–1622, Jun. 2017, doi: 10.15252/embj.201694920.
- [20] J. L. Perkins and L. Zhao, “The N-terminal domain of uracil-DNA glycosylase: Roles for disordered regions.,” *DNA Repair (Amst)*, vol. 101, p. 103077, May 2021, doi: 10.1016/j.dnarep.2021.103077.
- [21] R. R. Iyer, A. Pluciennik, V. Burdett, and P. L. Modrich, “DNA mismatch repair: functions and mechanisms.,” *Chem. Rev.*, vol. 106, no. 2, pp. 302–323, Feb. 2006, doi: 10.1021/cr0404794.
- [22] J. Jiricny, “Postreplicative mismatch repair.,” *Cold Spring Harb. Perspect. Biol.*, vol. 5, no. 4, p. a012633, Apr. 2013, doi: 10.1101/cshperspect.a012633.
- [23] P. B. Sang, T. Srinath, A. G. Patil, E.-J. Woo, and U. Varshney, “A unique uracil-DNA binding protein of the uracil DNA glycosylase superfamily.,” *Nucleic Acids Res.*, vol. 43, no. 17, pp. 8452–8463, Sep. 2015, doi: 10.1093/nar/gkv854.
- [24] W.-C. Ahn *et al.*, “Covalent binding of uracil DNA glycosylase UdgX to abasic DNA upon uracil excision.,” *Nat. Chem. Biol.*, vol. 15, no. 6, pp. 607–614, Jun. 2019, doi: 10.1038/s41589-019-0289-3.
- [25] L. Aravind and E. V. Koonin, “The alpha/beta fold uracil DNA glycosylases: a common origin with diverse fates.,” *Genome Biol.*, vol. 1, no. 4, p. RESEARCH0007, Oct. 2000, doi: 10.1186/gb-2000-1-4-research0007.

- [26] C. D. Mol *et al.*, “Crystal structure and mutational analysis of human uracil-DNA glycosylase: structural basis for specificity and catalysis.,” *Cell*, vol. 80, no. 6, pp. 869–878, Mar. 1995, doi: 10.1016/0092-8674(95)90290-2.
- [27] G. Róna *et al.*, “Detection of uracil within DNA using a sensitive labeling method for in vitro and cellular applications.,” *Nucleic Acids Res.*, vol. 44, no. 3, p. e28, Feb. 2016, doi: 10.1093/nar/gkv977.
- [28] D. E. Schones *et al.*, “Dynamic regulation of nucleosome positioning in the human genome.,” *Cell*, vol. 132, no. 5, pp. 887–898, Mar. 2008, doi: 10.1016/j.cell.2008.02.022.
- [29] F. C. Grandi, H. Modi, L. Kampman, and M. R. Corces, “Chromatin accessibility profiling by ATAC-seq.,” *Nat. Protoc.*, vol. 17, no. 6, pp. 1518–1552, Jun. 2022, doi: 10.1038/s41596-022-00692-9.
- [30] G. M. Cooper, Ed., *The Cell: A Molecular Approach*, 8th ed. 2019.
- [31] H. Holliday, L. A. Baker, S. R. Junankar, S. J. Clark, and A. Swarbrick, “Epigenomics of mammary gland development.,” *Breast Cancer Res.*, vol. 20, no. 1, p. 100, Sep. 2018, doi: 10.1186/s13058-018-1031-x.
- [32] K. A. Szigeti *et al.*, “[Role and alterations of DNA methylation during the aging and cancer].,” *Orv. Hetil.*, vol. 159, no. 1, pp. 3–15, Jan. 2018, doi: 10.1556/650.2018.30927.
- [33] S. Pal and J. K. Tyler, “Epigenetics and aging.,” *Sci. Adv.*, vol. 2, no. 7, p. e1600584, Jul. 2016, doi: 10.1126/sciadv.1600584.
- [34] G. B. Dixon, L. K. Bay, and M. V. Matz, “Evolutionary consequences of DNA methylation in a basal metazoan.,” *Mol. Biol. Evol.*, vol. 33, no. 9, pp. 2285–2293, Sep. 2016, doi: 10.1093/molbev/msw100.
- [35] D. C. Klein and S. J. Hainer, “Genomic methods in profiling DNA accessibility and factor localization.,” *Chromosome Res.*, vol. 28, no. 1, pp. 69–85, Mar. 2020, doi: 10.1007/s10577-019-09619-9.
- [36] K. Cui and K. Zhao, “Genome-wide approaches to determining nucleosome occupancy in metazoans using MNase-Seq.,” *Methods Mol. Biol.*, vol. 833, pp. 413–419, 2012, doi: 10.1007/978-1-61779-477-3_24.
- [37] J. D. Buenrostro, B. Wu, H. Y. Chang, and W. J. Greenleaf, “ATAC-seq: a method for assaying chromatin accessibility genome-wide.,” *Curr. Protoc. Mol. Biol.*, vol. 109, p. 21.29.1-21.29.9, Jan. 2015, doi: 10.1002/0471142727.mb2129s109.
- [38] P. J. Park, “ChIP-seq: advantages and challenges of a maturing technology.,” *Nat. Rev. Genet.*, vol. 10, no. 10, pp. 669–680, Oct. 2009, doi: 10.1038/nrg2641.

- [39] M. Weber *et al.*, “Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells.,” *Nat. Genet.*, vol. 37, no. 8, pp. 853–862, Aug. 2005, doi: 10.1038/ng1598.
- [40] S. J. Hainer and T. G. Fazzio, “High-Resolution Chromatin Profiling Using CUT&RUN.,” *Curr. Protoc. Mol. Biol.*, vol. 126, no. 1, p. e85, Apr. 2019, doi: 10.1002/cpmb.85.
- [41] C. Gutiérrez-Lovera, A. J. Vázquez-Ríos, J. Guerra-Varela, L. Sánchez, and M. de la Fuente, “The potential of zebrafish as a model organism for improving the translation of genetic anticancer nanomedicines.,” *Genes*, vol. 8, no. 12, Nov. 2017, doi: 10.3390/genes8120349.
- [42] K. Howe *et al.*, “The zebrafish reference genome sequence and its relationship to the human genome.,” *Nature*, vol. 496, no. 7446, pp. 498–503, Apr. 2013, doi: 10.1038/nature12111.
- [43] D. Ugarkovic, “Functional elements residing within satellite DNAs.,” *EMBO Rep.*, vol. 6, no. 11, pp. 1035–1039, Nov. 2005, doi: 10.1038/sj.embor.7400558.
- [44] N. L. Smith and D. Kimelman, “Establishing the body plan,” in *The zebrafish in biomedical research*, Elsevier, 2020, pp. 81–88.
- [45] É. V. Surányi *et al.*, “Using Selective Enzymes to Measure Noncanonical DNA Building Blocks: dUTP, 5-Methyl-dCTP, and 5-Hydroxymethyl-dCTP.,” *Biomolecules*, vol. 13, no. 12, Dec. 2023, doi: 10.3390/biom13121801.
- [46] H. Thorvaldsdóttir, J. T. Robinson, and J. P. Mesirov, “Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.,” *Brief. Bioinformatics*, vol. 14, no. 2, pp. 178–192, Mar. 2013, doi: 10.1093/bib/bbs017.
- [47] M. M. Hoffman, O. J. Buske, J. Wang, Z. Weng, J. A. Bilmes, and W. S. Noble, “Unsupervised pattern discovery in human chromatin structure through genomic segmentation.,” *Nat. Methods*, vol. 9, no. 5, pp. 473–476, Mar. 2012, doi: 10.1038/nmeth.1937.
- [48] R. M. Kuhn, D. Haussler, and W. J. Kent, “The UCSC genome browser and associated tools.,” *Brief. Bioinformatics*, vol. 14, no. 2, pp. 144–161, Mar. 2013, doi: 10.1093/bib/bbs038.
- [49] R. M. Layer, B. S. Pedersen, T. DiSera, G. T. Marth, J. Gertz, and A. R. Quinlan, “GIGGLE: a search engine for large-scale integrated genome analysis.,” *Nat. Methods*, vol. 15, no. 2, pp. 123–126, Feb. 2018, doi: 10.1038/nmeth.4556.
- [50] Avidity, “Reaction conditions for BirA biotin ligase.”
- [51] R. B. Phillips and K. M. Reed, “Localization of repetitive DNAs to zebrafish (*Danio rerio*) chromosomes by fluorescence in situ hybridization (FISH).,” *Chromosome Res.*, vol. 8, no. 1, pp. 27–35, 2000, doi: 10.1023/a:1009271017998.

- [52] S. E. Ross, A. Angeloni, F.-S. Geng, A. de Mendoza, and O. Bogdanovic, “Developmental remodelling of non-CG methylation at satellite DNA repeats.,” *Nucleic Acids Res.*, vol. 48, no. 22, pp. 12675–12688, Dec. 2020, doi: 10.1093/nar/gkaa1135.
- [53] M. Meier *et al.*, “Cohesin facilitates zygotic genome activation in zebrafish.,” *Development*, vol. 145, no. 1, Jan. 2018, doi: 10.1242/dev.156521.
- [54] V. Kasinath *et al.*, “Structures of human PRC2 with its cofactors AEBP2 and JARID2.,” *Science*, vol. 359, no. 6378, pp. 940–944, Feb. 2018, doi: 10.1126/science.aar5700.

Függelék

His-3xFLAG-ΔUNG-Avi:

MGSSHHHHHHSSGLVPRGSHMDYKDHDGDYKDHDIDYKDDDDKPLAARNVPVGFGESWKKHLSGEFGKPYFIKLM
GFVAEERKHYTVYPPPHQVFTWTQMCDIKDVKVILGQNPYHGPNAHGLCFSVQRPVPPPSLENIYKELSTDIEDFV
HPGHGDLGWAKQGVLLNAVLTVRAHQANSHKERGWEQFTDAVSWLNQNSNGLVFLWGSYAQKKGSAIDRKRH
HVLQTANPSPLSVYRGFFGCRHFSKTNELLQKSGKKPIDWKELDTYRYIGLNDIFEAQKIEWHE

His-1xFLAG-UdgX-Avi:

MAHHHHHHHVGTSNDDDDKSPDPNWELDYKDDDDKPSSGLVPRGSHMAGAQDFVPHADLAELAAAAGECRGCG
LYRDATQAVFGAGGRSARIMMIGECPGDKEDLAGLPFVGAGRLLDRALEAADIDRDALYVTNAVKHFKFTRAAGGKRRI
HKTPSRTEVVACRPWLIAEMTSVEPDVVLLGATAAKALLGNDFRVTQHRGEVLHVDDVPGDPALVATVHPSSLLRGPK
ERESAFAGLVDDLVAADVRLGLNDIFEAQKIEWHE

6xHis-tag
 FLAG címke
 AviTag
 pontmutáció

37. ábra: A His-3xFLAG-ΔUNG-Avi és His-1xFLAG-UdgX-Avi szenzorok aminosav-szekvenciái.

6. táblázat: Az AviTag PCR reakció körülményei.

Lépés	Hőmérséklet	Idő
Kezdeti denaturáció	95 °C	1 perc
35 ciklus	95 °C	20 másodperc
	63 vagy 65 °C	30 másodperc
	72 °C	3,5 perc
Végső meghosszabbítás	72°C	10 perc
Tartás	4 °C	-

7. táblázat: Kolónia PCR reakciókörülményei.

Lépés	Hőmérséklet	Idő
Kezdeti denaturáció	95 °C	5 perc
35 ciklus	95 °C	30 másodperc
	58 °C	30 másodperc
	72 °C	30 másodperc
Végső meghosszabbítás	72°C	10 perc
Tartás	4 °C	-

8. táblázat: Western blothoz és csepp blothoz használt pufferek összetétele. A komponensek rövidítései: etilén-diamin-tetraecetsav (EDTA), trisz-(hidroximetil)-amino-metán (TRIS), polioxietilén(20)-szorbitán-laurát (Tween-20).

	TBS	TBS-T	ETBS	ETBS-T
EDTA	-	-	1 mM	1 mM
TRIS-HCl (pH=8)	20 mM	20 mM	25 mM	25 mM
NaCl	140 mM	140 mM	137 mM	137 mM
KCl	-	-	2,7 mM	2,7 mM
Tween-20	-	0,1 V/V%	-	0,05 V/V%
pH	7,5	7,5	7,5	7,5

9. táblázat: Az elektroforézisek során használt futtatópufferek összetétele.

	TAE	TBE	SDS-PAGE-puffer
EDTA	20 mM	2 mM	-
TRIS-HCl (pH=8)	1 M	90 mM	25 mM
ecetsav	20 mM	-	-
bórsav		88 mM	-
glicin	-	-	190 mM
SDS	-	-	0,1 V/V%

10. táblázat: Az SDS-mintakoktól összetétele.

	SDS-mintakoktól (5X)
TRIS-HCl (pH=6,8)	250 mM
SDS	10%
DTT	10%
glicerol	50 V/V%
brómfenolkék	0,05 V/V%

11. táblázat: A puffer A és a Ni-NTA puffer összetétele.

	Puffer A	Ni-NTA puffer
TRIS-HCl	20 mM	50 mM
KCl	500 mM	500 mM
β-merkaptóetanol	-	2 mM
glicerol	10 V/V%	-
pH	7,5	8,0

12. táblázat: ΔUNG tisztításhoz és a biotiniláláshoz használt pufferek.

	LS	HS	VHS	Mosó	Elúciós	*biotiniláló
TRIS-HCl	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM
KCl	30 mM	300 mM	500 mM	30 mM	30 mM	200 mM
β-ME	5 mM	5 mM	5 mM	10 mM	10 mM	2 mM
imidazol	-	-	40 mM	50 mM	350 mM	-
pH	8,0	8,0	8,0	8,0	8,0	8,0

13. táblázat: A szenzorok tisztítása során használt lízis- és dialízispufferek összetétele.

	Lízispuffer		Dialízispuffer	
	ΔUNG	UdgX	ΔUNG	UdgX
TRIS-HCl (pH=8)	50 mM	20 mM	30 mM	20 mM
KCl	300 mM	500 mM	-	500 mM
NaCl	-	-	140 mM	
EDTA	0,5 mM	-	1 mM	-
Triton X-100	0,1 V/V%	-	-	-
Glicerín	-	10 V/V%	-	10 V/V%
Tween-20	-	-	0,01 V/V%	-
PMSF	1 mM	0,5 mM	-	0,1 mM
benzamidin	5 mM	5 mM	-	1mM
β-ME	10 mM	2 mM	15 mM	2 mM
lizozim	0,1 mg/ml	0,1 mg/ml	-	-
DNase	0,05 mg/ml	0,02 mg/ml	-	-
RNase	0,01 mg/ml	0,02 mg/ml	-	-
proteáz inhibitor	1 db tabletta/50 ml puffer		-	-
pH	8,0	8,0	7,4	8,0

14. táblázat: Az SDS-PAGE gélósszetétele. A táblázatban szereplő mennyiségek 4 gél megöntéséhez elegendőek.

	Tömörítő gél (4%-os)	Elválasztó gél (12%-os)
40%-os 37,5:1 akrilamid-biszakrilamid	1,95 ml	6 ml
10 %-os SDS-oldat	200 µl	200 µl

1,5 M TRIS-HCl (pH=8,8)	-	5 ml
0,5 M TRIS-HCl (pH=6,8)	5 ml	-
10 % APS	100 µl	100 µl
TEMED	10 µl	10 µl
víz	12,85 ml	8,8 ml

15. táblázat: A TBE-PAGE gélösszetétele. A táblázatban szereplő mennyiségek 6 gél megöntéséhez elegendőek.

	Natív	Denaturáló
19:1 akrilamid-biszakrilamid	4,5 g	4,5 g
TBE (10x tömény)	3 ml	3 ml
10 % APS	240 µl	240 µl
TEMED	24 µl	24 µl
urea	-	12,6 g
víz	11,1 ml	11,1 ml

16. táblázat: Letöltött funkcionális genomikai adatokra vonatkozó információk.

Genomszegmentációs analízishez használt genomi profilok				
térképezett tulajdonság/faktor	adattípus	embrió kora (hpf)	GEO szériaszám	eredeti referenciagenom
hozzáférhetőség	ATAC-seq	2,25	GSE168361	Zv10
hozzáférhetőség	MNase-seq	2,75	GSE109410	GRCz11
hozzáférhetőség	ATAC-seq	4	GSE168361	Zv10
hozzáférhetőség	ATAC-seq	5,3	GSE168361	Zv10
H2Aub	ChIP-seq	2,5	GSE168360	Zv10
H2AZ	ChIP-seq	2,5	GSE95030	Zv10
H3K14ac	ChIP-seq	2,5	GSE95030	Zv10
H3K27ac	ChIP-seq	2,5	GSE95030	Zv10
H3K27me3	ChIP-seq	1,5	GSE110663	Zv9
H3K27me3	ChIP-seq	2,25	GSE110663	Zv9
H3K27me3	ChIP-seq	3	GSE110663	Zv9
H3K27me3	ChIP-seq	oocita	GSE110663	Zv9
H3K27me3	ChIP-seq	spermium	GSE110663	Zv9
H3K4me1	ChIP-seq	2,5	GSE95030	Zv10

H3K4me3	ChIP-seq	1,5	GSE110663	Zv9
H3K4me3	ChIP-seq	2,25	GSE110663	Zv9
H3K4me3	ChIP-seq	3	GSE110663	Zv9
H3K4me3	ChIP-seq	oocita	GSE110663	Zv9
H3K4me3	ChIP-seq	spermium	GSE110663	Zv9
metiláció	MeDIP-seq	2,5	GSE52703	Zv9
metiláció	MeDIP-seq	spermium	GSE52703	Zv9
Rnf2	ChIP-seq	2,5	GSE168360	Zv10
GIGGLE elemzéshez használt intervallumfájlok				
térképezett tulajdonság/faktor	adattípus	embrió kora (hpf)	GEO szériaszám	eredeti referenciagenom
hozzáférhetőség	ATAC-seq	2,25	GSE168361	Zv10
hozzáférhetőség	ATAC-seq	2,5	GSE100286	Zv9
hozzáférhetőség	ATAC-seq	2,5	GSE100286	Zv9
hozzáférhetőség	ATAC-seq	4	GSE168361	Zv10
hozzáférhetőség	ATAC-seq	4	GSE152743	Zv10
hozzáférhetőség	ATAC-seq	5,3	GSE168361	Zv10
hozzáférhetőség	ATAC-seq	8	GSE233691	GRCz10
hozzáférhetőség	ATAC-seq	8	GSE233691	GRCz10
Aebp2	ChIP-seq	4	GSE168360	Zv10
Aebp2	ChIP-seq	4	GSE168360	Zv10
Aebp2	ChIP-seq	4,3	GSE168360	Zv10
Eomesa	ChIP-seq	4	GSE51891	Zv9
FoxH1	ChIP-seq	6	GSE133990	Zv9
H2Aub1	ChIP-seq	2,5	GSE168360	Zv10
H2Aub1	ChIP-seq	3,5	GSE168360	Zv10
H2Aub1	ChIP-seq	4,3	GSE168360	Zv10
H2AZ	ChIP-seq	2,5	GSE95030	Zv10
H2Az	ChIP-seq	3,5	GSE168360	Zv10
H2AZ	ChIP-seq	4	GSE95030	Zv10
H2AZ	CUT&Tag	6	GSE178559	GRCz11
H2AZ	ChIP-seq	spermium	GSE95030	Zv10
H3K14ac	ChIP-seq	2,5	GSE95030	Zv10

H3K14ac	ChIP-seq	4	GSE95030	Zv10
H3K14ac	ChIP-seq	spermium	GSE95030	Zv10
H3K27ac	ChIP-seq	2,5	GSE95030	Zv10
H3K27ac	ChIP-seq	4,3	GSE143306	GRCz11
H3K27ac	ChIP-seq	4,3	GSE32483	Zv8
H3K27ac	ChIP-seq	4,3	GSE32483	Zv8
H3K27ac	ChIP-seq	4,3	GSE32483	Zv9
H3K27ac	ChIP-seq	spermium	GSE95030	Zv10
H3K27me3	ChIP-seq	0,75	GSE110663	Zv9
H3K27me3	ChIP-seq	1,5	GSE110663	Zv9
H3K27me3	ChIP-seq	2,25	GSE110663	Zv9
H3K27me3	ChIP-seq	3	GSE110663	Zv9
H3K27me3	ChIP-seq	4	GSE168360	Zv10
H3K27me3	ChIP-seq	4	GSE168360	Zv10
H3K27me3	ChIP-seq	4,3	GSE70847	Zv9
H3K27me3	CUT&RUN	6	GSE178343	GRCz11
H3K27me3	ChIP-seq	oocita	GSE110663	Zv9
H3K27me3	ChIP-seq	spermium	GSE110663	Zv9
H3K27me3	ChIP-seq	spermium	GSE95030	Zv10
H3K4me1	ChIP-seq	2,5	GSE95030	Zv10
H3K4me1	ChIP-seq	3,5	GSE168360	Zv10
H3K4me1	ChIP-seq	4,3	GSE32483	Zv8
H3K4me1	ChIP-seq	4,3	GSE32483	Zv8
H3K4me1	ChIP-seq	4,3	GSE32483	Zv9
H3K4me1	ChIP-seq	spermium	GSE95030	Zv10
H3K4me3	ChIP-seq	0,75	GSE110663	Zv9
H3K4me3	ChIP-seq	1,5	GSE110663	Zv9
H3K4me3	ChIP-seq	2,25	GSE110663	Zv9
H3K4me3	ChIP-seq	3	GSE110663	Zv9
H3K4me3	ChIP-seq	4,3	GSE143306	GRCz11
H3K4me3	ChIP-seq	4,3	GSE32483	Zv8
H3K4me3	ChIP-seq	4,3	GSE32483	Zv8
H3K4me3	ChIP-seq	4,3	GSE32483	Zv9
H3K4me3	CUT&RUN	6	GSE178343	GRCz11
H3K4me3	ChIP-seq	oocita	GSE110663	Zv9

H3K4me3	ChIP-seq	spermium	GSE110663	Zv9
H3K4me3	ChIP-seq	spermium	GSE95030	Zv10
metiláció	MBD-Seq	0	GSE52110	Zv9
Jarid2	ChIP-seq	4,3	GSE168360	zv10
metiláció	MethylCap-seq	4,3	GSE70847	Zv9
H3K9me3	CUT&RUN	6	GSE178343	GRCz11
Hoxb1b	ChIP-seq	8	GSE233695	GRCz10
Hoxb1b	ChIP-seq	8	GSE233695	GRCz10
Meis2b	ChIP-seq	8	GSE233695	GRCz10
Meis2b	ChIP-seq	8	GSE233695	GRCz10
metiláció	MBD-Seq	spermium	GSE52110	Zv9
Pou5f2	ChIP-seq	4	GSE39780	Zv9
Rad21	ChIP-seq	2,5	GSE84600	Zv9
Rad21	ChIP-seq	4	GSE167125	Zv10
Rad21	ChIP-seq	4,5	GSE84600	Zv9
RARaa	ChIP-seq	8	GSE233695	GRCz10
RARaa	ChIP-seq	8	GSE233695	GRCz10
RNA pol II	ChIP-seq	4	GSE152742	Zv10
RNA pol II	CUT&RUN	6	GSE178343	GRCz11
Rnf2	ChIP-seq	2,5	GSE168360	Zv10
Rnf2	ChIP-seq	3,5	GSE168360	Zv10
Rnf2	ChIP-seq	4,3	GSE168360	Zv10
Smad2	ChIP-seq	4	GSE51891	Zv9
Sox2	ChIP-seq	4	GSE39780	Zv9
Sox3	ChIP-seq	8	GSE233695	GRCz10
Sox4	ChIP-seq	8	GSE233695	GRCz10