

EGYETEMI DOKTORI (PhD) ÉRTEKEZÉS

**Az ösztrogén receptor alfa-vezérelt
szuper-enhanszerek kialakulásának vizsgálata**

Bojcsuk Dóra

Témavezető: Dr. Bálint Bálint László



DEBRECENI EGYETEM

MOLEKULÁRIS SEJT- ÉS IMMUNBIOLÓGIA DOKTORI ISKOLA

Debrecen, 2019

Tartalomjegyzék

1. Bevezetés	3
1.1. A Humán Genom Projekt hagyatéka	3
1.2. Az eukarióta genom szerveződése	4
1.3. A transzkripció szabályozásának szintjei eukariótákban	8
1.4. <i>Cisz</i> - és <i>transz</i> szabályozó elemek	10
1.4.1. Promóter elemek	10
1.4.2. Enhanszer és silencer elemek	10
1.4.3. Transzkripciós faktorok	12
1.5. Hiszton módosítások	13
1.6. ChIP-seq: a transzkripciós faktor kötőhelyek genom-szintű azonosítása	15
1.7. Szuper-enhanszerek	16
1.7.1. Szuper-enhanszerek azonosítása	19
1.8. A magreceptorok	21
1.8.1. A magreceptorok csoportosítása	23
1.9. Az ösztrogén receptor alfa	25
1.9.1. MCF-7 sejtvonala: modell az ER α tanulmányozásához	26
1.9.2. Az ösztrogén receptor alfa általi génszabályozás lehetséges mechanizmusai	27
1.9.3. Az ösztrogén receptor alfa partnerei a transzkripciós szabályozásban	30
2. Célkitűzés	33
3. Anyagok és módszerek	35
3.1. Adatgyűjtés	35
3.1.1. Szekvenálási adatok az ligand-indukálható szuper-enhanszerek karakterizálásához	35
3.1.2. Szekvenálási adatok a különböző típusú ER α kötőhelyek karakterizálásához	37
3.2. Adatfeldolgozás	38
3.2.1. ChIP-seq adatok elemzése	38
3.2.1.1. Lefedettségi érték (RPKM) meghatározása	39
3.2.1.2. Szuper-enhanszerek azonosítása	40
3.2.1.3. De novo motívumkeresés	40
3.2.1.4. Hőtérkép készítése	41
3.2.1.5. Hisztogram készítése	41
3.2.1.6. Dobozdiagram készítése	41

3.2.1.7. Statisztikai módszerek.....	41
3.2.1.8. Adatábrázolás	42
3.2.2. RNS szekvenálási adatok elemzése	43
4. Eredmények.....	44
4.1. Az ösztrogén receptor alfa-vezérelt szuper-enhanszerek azonosítása	44
4.2. Anya és leány enhanszerek azonosítása	45
4.2.1. A kromatin transzkripció aktivitásának vizsgálata.....	50
4.2.2. Specifikus szekvenciamotívumok azonosítása	53
4.3. További ligand-indukálható szuper-enhanszerek vizsgálata	56
4.3.1. A vizsgálatba bevont transzkripció faktorok.....	56
4.3.2. Anya és leány enhanszerek dinamikájának és motívumainak vizsgálata	56
4.4. Az ER α , FoxA1 és AP2 γ jelenlétének vizsgálata egymás szuper-enhanszer régióiban.....	64
4.5. ER α -vezérelt szuper-enhanszerek: áttekintés	66
4.6. A nem szuper-enhanszerek részét képező ER α kötőhelyek jellemzése	68
4.6.1. A kötőhelyek csoportosítása az indukálhatóság és az ERE megléte alapján	68
4.6.2. Az ösztradiol hatása az ER α kötési gyakoriságára a különböző csoportokban	70
4.6.3. A különböző csoportokban feldúsult szekvenciamotívumok	71
4.6.4. A kromatin hozzáférhetőségének vizsgálata a különböző csoportokban	73
4.6.5. A FoxA1, AP2 γ és GATA3 fehérjék megléte a különböző csoportokban	74
4.7. Az ER α célgének expressziója	77
4.7.1. A szuper-enhanszerek által szabályozott gének.....	78
5. Diskusszió	80
6. Összefoglalás.....	88
7. Summary.....	90
8. Kulcsszavak	92
9. Rövidítések jegyzéke.....	93
10. Függelék.....	97
10.1. A magreceptor szupercsalád tagjai és ligandjaik.....	97
10.2. Az értekezés alapjául szolgáló közlemények	99
11. Köszönetnyilvánítás.....	101
12. Hivatkozások.....	102

1. Bevezetés

1.1. A Humán Genom Projekt hagyatéka

Azokra az izgalmas kérdésekre, hogy két fajt vagy két embert mi tesz egyedivé, mitől függ a szemünk és a hajunk színe, vagy mi okozza azt, hogy bizonyos betegségekre fogékonyabbak vagyunk másoknál, a választ a genetikai állományunkban, tehát a dezoxiribonukleinsav (DNS) bázissorrendünkben (szekvenciánkban) találjuk. A DNS egy olyan biopolimer, amely alapvetően négyféle dezoxiribonukleotid egységből épül fel. Minden egyes nukleotidot egy foszfátcsoport, egy dezoxiribóz (cukor) és egy nitrogén tartalmú heterociklusos bázis alkot, mely lehet adenin (A), guanin (G), citozin (C) vagy timin (T). A nukleotidok cukor- és foszfátelemei 3' – 5' irányban foszfodiészter kötéssel képesek összekapcsolódni és polinukleotid láncot létrehozni; ezen túlmenően pedig két, egymással szemközti, de ellentétes lefutású (antiparallel), hidrogénhidak (H-hidak) segítségével kapcsolódó polinukleotid lánc létrehozza a duplaszálú DNS-t, melyet a kialakult csavart szerkezete alapján kettős hélixnek is szokás nevezni. A két lánc közötti távolság minden pozícióban azonos, ami annak a komplementaritásnak köszönhető, hogy a purinvázis adenin minden esetben a pirimidinvázis timinnel szemben helyezkedik el és alkot egy bázispárt, míg a szintén purinvázis citozin a pirimidinvázis guaninhoz kapcsolódik (Alberts et al., 2002).

Az emberi génállomány (genom) megközelítőleg 3,2 milliárd ($3,2 \times 10^9$) bázispár jól meghatározott sorrendiségével írható le, melynek első változatát 2001-ben, a Humán Genom Projekt (HGP) során közzétették (Lander et al., 2001). Ennek a 3,2 milliárd nukleotidnak az egymásutánisága azonban csak általánosságban érvényes. Két nem rokon személynek a genetikai állománya közötti különbség 0,1%-ra becsülhető, és ez eredményezi – a fent említett – eltérő tulajdonságokat és egyedi „genetikai ujjlenyomatot”.

A HGP révén megismerhettük a gének elhelyezkedését a genomban, illetve a fehérjekódoló génkészletünket is – melynek mérete 20-25000 génre becsülhető – (International Human Genome Sequencing Consortium, 2004). Az úgynevezett (ún.) háztartási (*housekeeping*) génektől eltekintve, amelyek a sejtek alapvető folyamataiért felelősek, a közel 25000 génnek legfeljebb a fele mutat bármiféle aktivitást az egyes sejtekben; a különböző sejtek pedig a génkészlet más-más része által szabályozódnak. Ez a sejt/szövet-specifikus génexpressziós mintázat, illetve a gének megfelelő időben történő ki- és bekapcsolása határozza meg a szöveti fejlődés molekuláris mechanizmusát, a sejtek sorsát, anyagsere állapotát, összességében tehát a szerteágazó biológiai folyamatok transzkripciós szabályozását. A génexpressziós mintázat mindemellett számos betegség indikátora is lehet (Niehrs & Pollet, 1999; Song, Ahn, Suh, Davis, & Lee, 2013).

A HGP magával hozta a DNS szekvenálás és a bioinformatika rohamos fejlődését is, melyek segítségével ma már néhány nap alatt megismerhetjük akár a saját genomunk nukleotid sorrendjét; a HapMap Projekt által azonosított egynukleotidos eltérésekkel összevetve pedig akár az azokban rejlő „hibákat” is (International HapMap Consortium, 2003).

1.2. Az eukarióta genom szerveződése

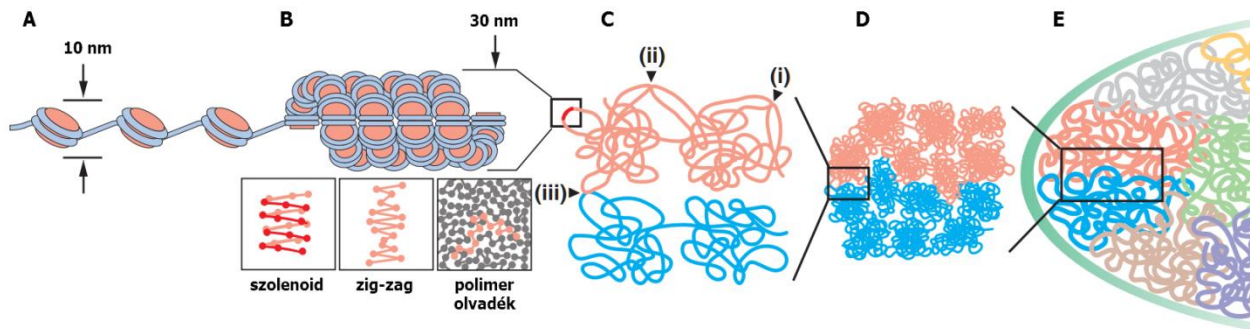
A sejtmaggal rendelkező – ún. eukarióta – sejtek és genomjaik jelentősen nagyobbak és összetettebbek a prokariótákéhoz képest. A sejtmag, melynek tartalmát a maghártya választja el a citoplazmától, tartalmazza a 2×3200 megabázispárnyi DNS-t. Hogyan képes ez az összesen megközelítőleg 2 méter hosszúságú fonál elférni a 6 μm átmérőjű sejtmagban? A válasz a kromatin egy magasabb rendű, jól strukturált, kompaktabb szerveződésében rejlik.

Ennek elérése érdekében a DNS először nukleoszóma szerkezetet vesz fel, melynek során hisztonokra, egészen pontosan a H2A, H2B, H3 és H4 hisztonok két-két kópiájára tekeredik fel. Ez nukleoszómánként ~147 bázispárnyi feltekeredett DNS-t jelent, a mononukleoszómák pedig 10-80 bázispárnyi összekötő, ún. internukleoszómális DNS-sel kapcsolódnak egymáshoz, így kialakítva egy gyöngyfűzérre hasonlító struktúrát, melynek átmérője 10 nm. Ugyanakkor ez a feltekeredés még csupán a DNS 5-10-szeres tömörödését teszi lehetővé. Ebben a formában a hisztonokra való feltekeredés ellenére a DNS még hozzáférhető lehet nem-hiszton fehérjék számára, a hisztonok aminoterminális végei pedig általában elérhetőek maradnak a módosító enzimek számára, lehetővé téve ezáltal a nukleoszóma struktúra lokális fellazítását, letekeredését.

A következő lépésben a H1/H5 hiszton stabilizálja a szomszédos nukleoszómákat, ami elősegíti egy még kompaktabb, ~50-szeres tömörödést eredményező forma kialakulását. Az így formálódó, 30 nm átmérőjű kromatin rostok felvehetnek tekercsszerű, ún. szolenoid fonalra hasonlító szerkezetet, mely révén minden nukleoszóma képes kapcsolódni az ötödik vagy hatodik szomszédos nukleoszómával, vagy olyan cikk-cakk (zig-zag) szerkezetet, mely minden második nukleoszóma kapcsolódását teszi lehetővé (Dorigo et al., 2004; Robinson, Fairall, Huynh, & Rhodes, 2006; Schalch, Duda, Sargent, & Richmond, 2005). Habár a 30 nm-es rost szerkezetét *in vitro* széles körben tanulmányozzák, az *in vivo* létezésére szolgáló bizonyítékok korlátozottak. A jelenleg elfogadott nézet szerint ez a fajta szerkezet az alacsony kromatin-koncentrációjú és alacsony ionerősségű kromatinkészítményekben preferált, ahol az intramolekuláris kölcsönhatások részesülnek előnyben az intermolekuláris kölcsönhatásokkal szemben (Bian & Belmont, 2012; Felsenfeld & Groudine, 2003; Hübner, Eckersley-Maslin, & Spector, 2013; Maeshima, Hihara, & Eltsov, 2010) **(1. ábra)**.

A legmagasabb szerveződési szint azonban a kromoszóma szerkezet révén valósul meg. Ennek elérése érdekében a 30 nm-es rost először Kondenzin II gyűrűk által a tövüknél összefogott nagyméretű hurkokat formál. E hurkok töve lényegében egy Kondenzin II-ben gazdag tengelyt képez, melyből sugárirányban állnak ki a különböző méretű hurkok. Ezeket végül a Kondenzin I gyűrűk tagolják tovább a tömör kromoszóma létrehozásához (Gibcus et al., 2018). A kromatin kromoszómákba szerveződése ugyanakkor csak a sejtek osztódása során figyelhető meg; a sejtek életének nagy részében a DNS-nek közel a fele többé-kevésbé hozzáférhető formában van jelen a sejtmagban található nem-hisztón fehérjék számára, lehetővé téve ezáltal a gének DNS-kötő fehérjék általi szabályozását.

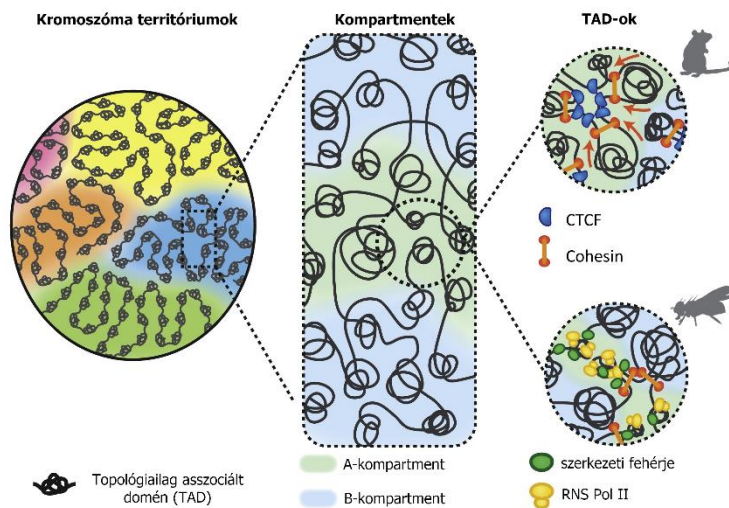
Az újgenerációs szekvenálás mellett a kromatin háromdimenziós szerkezetét vizsgáló módszerek és a bioinformatika rohamos fejlődésével az utóbbi 10-15 évben egyre részletesebb kép tárult elénk a genomot illetően. Ismertté vált, hogy a kromatin rostok olyan komplex geometriai alakzatot vesznek fel, melyek ún. fraktál globulusokhoz hasonlítanak (Lieberman-Aiden et al., 2009). A felvett alakzat révén a globulusok különféle kölcsönhatást képesek kialakítani önmagukkal és egymással egyaránt: a leggyakrabban az azonos kromatin részlegen belüli interakciók figyelhetők meg, de előfordulhatnak más kromatin részleggel történő, de még egyazon kromoszóma területen, ún. territóriumon belüli kölcsönhatások is. Nagyon ritkán, de létrejöhet interakció a szomszédos territóriumok között is (**1. ábra**).



1. ábra. Az eukarióta genom szerveződése. A sematikus ábra a kromatin szerveződésének legfontosabb szintjeit mutatja be. **(A)** A DNS duplahélix, melynek átmérője eredendően 2 nm-re tehető, a nukleoszómák által kialakított gyöngyfüzér („beads on a string”) struktúrát veszi fel elsődlegesen (10 nm). **(B)** A kromatin rostok megközelítőleg 30 nm átmérőjűek, melyek felvehetik mind a szolenoid, mind a zig-zag (ún. cikk-cakk) szerkezetet. **(C)** Ezt követően a kromatin rostok olyan komplex geometriai alakzatot vesznek fel, melyek az ún. fraktál globulusokhoz hasonlítanak. A globulusokba való szerveződés révén a rostok kölcsönhatása megfigyelhető egyrészt ugyanazon elkülönt kromatin részlegen belül (i), vagy más-más kromatin részleggel, de egyazon kromoszóma territóriumon belül (ii), másrészt megfigyelhető interakció a szomszédos territóriumok között is (iii). A globulusok egymással való szoros összefonódása **(D)** révén rajzolódnak ki a kromoszóma territóriumok **(E)** (Hübner et al., 2013).

A kromoszóma territóriumokon belül meghatároztak aktív és inaktív kromatin szerkezeteket (az ún. „A”, ill. „B” kompartmenteket). Leginkább az aktív (A) kompartmenteken belül találhatóak meg az ún. topológiaiilag asszociált domének (TAD-ok), melyek olyan 10-100 kilobázis nagyságú kromatin egységek, melyeken belül nagy számban figyelhetőek meg kromatin-kromatin (pl. promóter-enhanszer) interakciók. Ezek a kapcsolatok a DNS hurkolódása révén valósulnak meg, lehetővé téve két – lineárisan mérve – távoli genomi régió kölcsönhatását **(2. ábra)** (Dixon et al., 2012; Fortin & Hansen, 2015; Lieberman-Aiden et al., 2009; Nora et al., 2012). A TAD-ok révén a genom kisebb-nagyobb szakaszokra különül el, melyeken belül a DNS szabályozó szakaszai és a hozzájuk kötődő fehérjék nagyobb valószínűséggel kerülnek fizikai kölcsönhatásba, mint két TAD között. A TAD határok, melyeket gyakran a CCTC-kötő faktor

(CTCF) jelenléte jellemez, nagymértékű megőrzöttséget mutatnak a különböző fajok között. A bennük fellépő mutációk a kromatin hurkolódását megváltoztatva a transzkripció hibás szabályozását is eredményezhetik, végső soron pedig akár fejlődési rendellenességhez vagy daganat kialakulásához is vezethetnek (Gong et al., 2018).



2. ábra. A kromatin magasabb rendű, háromdimenziós szerkezete. A kromoszóma territóriumok az interfázisos kromoszómák alapvető szerkezeti egységei. A kromatin aktív (A) és inaktív (B) kompartmentekre különíthető el. Ezeket topológiailag asszociált domének (TAD-ok = hurkok) építik fel, melyek végeit gyakran a CTCF fehérje és a Cohesin komplex tartja össze (egérben és

emberben egyaránt), míg az ecetmuslicában a CTCF-en kívül egyéb szerkezeti/”inzulátor” fehérjék is képesek ezt a feladatot ellátni (Hug & Vaquerizas, 2018).

1.3. A transzkripció szabályozásának szintjei eukariótákban

A gének szabályozása alapvető biológiai folyamat a sejtek túlélése és működése szempontjából, melyet tökéletesen kifinomult szabályozási mechanizmusok koordinálnak. Az első génszabályozó modellt 1960-ban két, később Nobel-díjjal jutalmazott francia biokémikus, *François Jacob* és *Jacques Monod* alkotta meg, akik azt találták, hogy az *E. coli* laktóz-metabolizmusában szerepet játszó *lac operon* bekapcsolása és szabályozása során a *lacZ*, *lacY* és *lacA* gének termékei (RNS, fehérje) képesek visszacsatolás révén szabályozni a gének expresszióját (Jacob & Monod, 1961). Ez a koncepció képezte alapját a kortárs génexpressziós modelleknek.

A jelenlegi tudásunk szerint a DNS-kötő, génszabályozó fehérjék, avagy transzkripciósfaktorok (TF-ok) ún. *transz* szabályozó elemként fejtik ki serkentő vagy gátló hatásukat, mégpedig olyan *cisz* szabályozó elemekkel (CRE; *cis*-regulatory element) való kölcsönhatás révén, mint a promóterek, enhanszerek, silencer-ek és inzulátorok, vagy olyan *cisz* szabályozó modulokon (CRM; *cis*-regulatory module) keresztül, melyek esetenként több, mint 10 CRE-t is magukban foglalhatnak (Lelli, Slattey, & Mann, 2012; Niu, Tabari, Ni, & Su, 2018). Az eukarióta egyedfejlődés során a gének megfelelő szabályozásáért – részben ilyen CRM-eken keresztül – a transzkripciósfaktorok jól meghatározott csoportjainak az együttműködése a felelős. E csoportok összetétele kombinatorikus és eltérő a fejlődés különböző fázisaiban, lehetővé téve ezáltal a gének sejttípus- és környezetfüggő szabályozását (Istrail & Davidson, 2005; Jeziorska, Jordan, & Vance, 2009).

Bár *Johann Gregor Mendel*t tartják a „genetika atyjának”, maga a *genetika* kifejezés 40 évvel korábbra, *gróf Festetics Imre* nevéhez köthető. Festetics azáltal, hogy feljegyezte a beltenyészett vonalak keresztezése során megfigyelhető örökletes különbségeket, ugyan nem írta le a faktoriális genetikát, irányt mutatott követőinek (Szabó, 2010). Az *epigenetika* kifejezés *Conrad Waddington* nevéhez fűződik az 1940-es évekből, amit később a biológia egy olyan ágaként definiált, amely a gének és a termékeik közötti azon kölcsönhatásokat vizsgálja, amelyek a fenotípus létrehozásáért felelősek (Waddington, 1968, 2012). Az évek során e kifejezés jelentése fokozatosan szűkült; az epigenetika jelenleg azon meiotikusan és/vagy mitotikusan öröklődő változásokat jelenti a gének működésében, amelyek eredendően nem járnak a DNS szekvenciájának a megváltoztatásával (Wu & Morris, 2001). A DNS szabályozó elemei mellett ez – a kromatin szintjén való epigenetikai szabályozás – az eukarióta transzkripciósfolyamatok másik fontos szintje, amely magában foglalja a nukleoszómába csomagolt „cél” DNS szakasz

hozzáférhetőségének, valamint a DNS metiláltsági fokának a szabályozását (Dupont, Armant, & Brenner, 2009).

1.4. Cisz- és transz szabályozó elemek

1.4.1. Promóter elemek

A *cisz* szabályozó elemek egyik legjobban tanulmányozott típusai a promóter elemek. Ezek, bár elengedhetetlenek a transzkripció végbemeneteléhez, önmagukban csupán egy alapszintű mRNS mennyiség képződéséhez elegendőek.

Az egyik legismertebb promóter elem a TATA-box, amely a szekvenciájáról (TATAAT) kapta a nevét, és mind prokariótákban, mind eukariótákban az egyik alapvető transzkripció iniciációs helyet jelöli a gének promóter régiójában (Pribnow, 1975). A TATA-box mellett az eukarióta promóterekben előforduló elemek még a CAAT-box és GC/GT-box. Bár bizonyos gének esetében több promóter is azonosítható – melyek feltételezhetően más-más körülmények között lépnek működésbe –, a legtöbb eukarióta gén egy promóterrel rendelkezik, mely rendszerint a transzkripció start hely (TSS; transcription start site) közelében, a géntől közvetlenül 5' irányban (*upstream*) található (Brown & Feder, 2005).

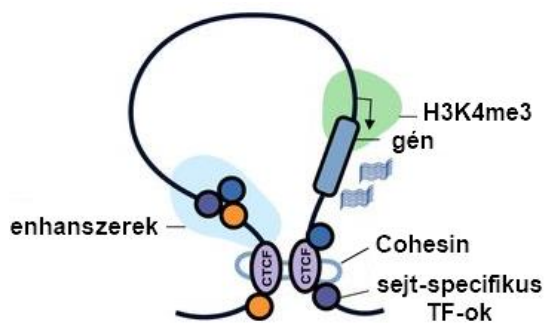
1.4.2. Enhanszer és silencer elemek

A promóterekhez viszonyítva az enhanszerek, mint a transzkripciót fokozó szekvenciák, sokkal nagyobb variabilitást mutatnak, és ők felelnek a génexpresszió valódi finomhangolásáért (Banerji, Rusconi, & Schaffner, 1981; Wittkopp & Kalay, 2012; Wray, 2007). A DNS ezen szakaszai – hasonlóan a promóterekhez – tipikusan néhány száz bázispár hosszúságúak és megfelelő felszint biztosítanak a transzkripció faktorok rövid, specifikus DNS szakaszokhoz

(motívumokhoz) való kötődéséhez (Spitz & Furlong, 2012). A génhez képest mind 3' (*upstream*), mind 5' (*downstream*) irányban elhelyezkedhetnek, de megtalálhatóak intronokban, több ezer kilobázisra (kb) a szabályozott géntől, egy másik génen, vagy akár másik kromoszómán is (Kleinjan & van Heyningen, 2005). Ebből következik, hogy hatásukat nem feltétlenül a legközelebbi gén promóterén fejtik ki; sőt, bizonyos esetekben több gént is szabályozhatnak.

Az emberi genomban azonosított enhanszerek száma 400000 – 1400000-re tehető, melyből körülbelül néhány tízezer mutat csak aktivitást egy-egy sejttípusban (ENCODE Project Consortium, 2012; Heintzman et al., 2009; Thurman et al., 2012; Visel et al., 2009; Whyte et al., 2013). Az enhanszerekkel ellentétben a silencer-ek olyan DNS szekvenciák, melyek gátló fehérjéket képesek kihorgonyozni – azok visszavetik a promótereken kialakult transzkripció „gépezetek” működését, így végsősoron a célgén csendesítését végzik (Ogbourne & Antalis, 1998).

A genom két távoli régiójában fellelhető promóter és enhanszer/silencer elemek, illetve a hozzájuk kapcsolódó transzkripció faktorok közötti fizikai kapcsolat létrejöttét a kromatin hurkolódása (looping) teszi lehetővé, melyet az ún. inzulátor elemek specifikus kötése is stabilizálhat. Az inzulátor elemek képesek megakadályozni a szomszédos kromatin domének (TAD-ok) közötti nem megfelelő kölcsönhatások létrejöttét (Gaszner & Felsenfeld, 2006; Sun & Elgin, 1999). Ezen elemeket gerincesekben elsődlegesen a CTCF köti, ugyanakkor a CTCF kötőhelyek csak kis hányada működik valóban inzulátor elemként (**3. ábra**) (Bell, West, & Felsenfeld, 1999; M. Liu et al., 2015).



3. ábra. Promóter-enhanszer interakció létrejötte a kromatin CTCF-mediált hurkolódása által. Két CTCF molekula egy Cohesin gyűrű segítségével képes a kromatin hurkolódását előidézni és fizikai interakciót létesíteni a hurkon belül elhelyezkedő gén promótere és az enhanszerek, illetve a hozzájuk kapcsolódó transzkripciós faktorok (TF-ok) között (Arzate-Mejía, Recillas-Targa, & Corces, 2018).

1.4.3. Transzkripciós faktorok

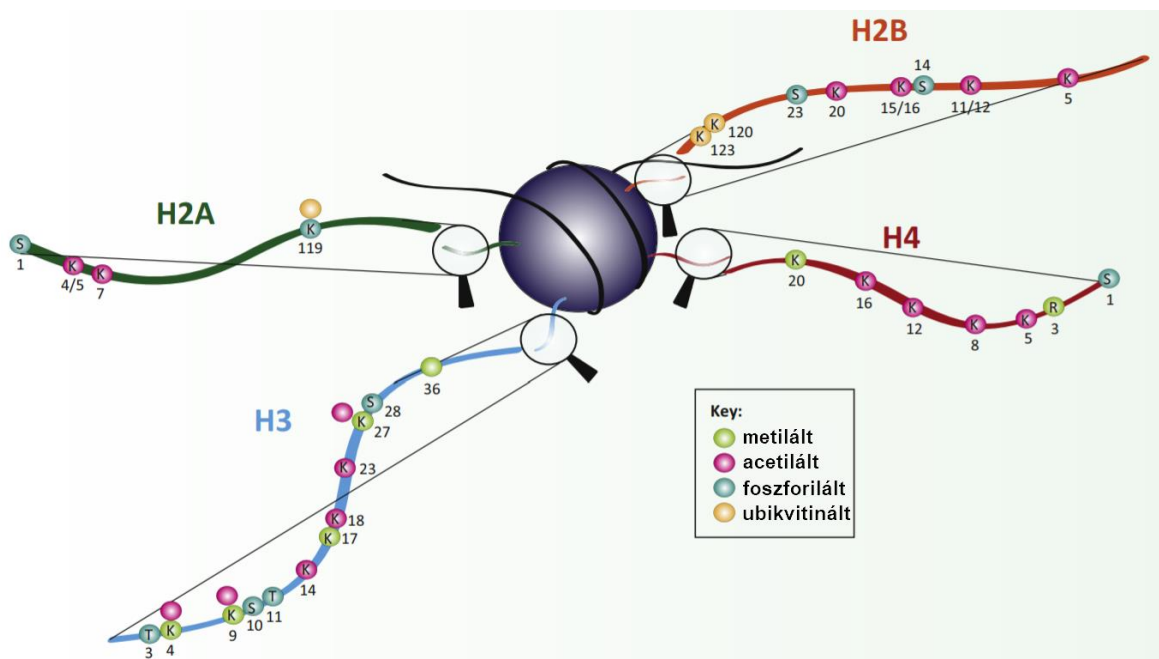
A transzkripciós faktorok azok a CRE-ek specifikus DNS szekvenciáit (szekvencia motívumait) felismerő fehérjék, melyek a kromatin és a transzkripció szabályozása révén gyakorolnak pozitív vagy negatív hatást a gének expressziójára. Egy random szekvenciához képest a saját felismerő szekvencia motívumaikhoz ezerszeres vagy akár még nagyobb affinitással képesek kötődni, és ez az elsődleges szelekció teszi lehetővé azt, hogy hatásukat csak a célgéneken fejthessék ki (Vaquerizas, Kummerfeld, Teichmann, & Luscombe, 2009). Amennyiben a kromatin hozzáférhető a transzkripciós faktor számára, és az elfoglalja a motívumát, a génekre gyakorolt hatását további ko-faktorok (pl. a TFII complex, a Mediátor komplex vagy az acetil-transzferáz aktivitással rendelkező P300 fehérje) kötésével, valamint poszt-transzlációs (pl. hiszton) módosítások létrehozása által fejtheti ki. A felépült fehérje komplex már képes az RNS polimeráz II enzimet kihorgonyozni a „core” promóterhez és aktiválni azt (Reiter, Wienerroither, & Stark, 2017).

A transzkripció szabályozása az egyes sejttípusokban tipikusan különböző transzkripciós faktorok kombinációja révén valósul meg, melyek közül néhánynak az együttműködése is elegendő ahhoz, hogy egy adott sejttípusnak egyedi jelleget kölcsönözzenek. Érdekes, hogy a

transzkripció faktorok elsősorban a nyitott kromatint preferálják, ugyanakkor maga a nyitott kromatint is részben transzkripció faktorok szabályozzák. A humán transzkripció faktorok katalógizálása során ezidáig összesen 1639 transzkripció faktort azonosítottak, melyek közel háromnegyede rendelkezett saját motívummal (Fulton et al., 2009; Lambert et al., 2018).

1.5. Hiszton módosítások

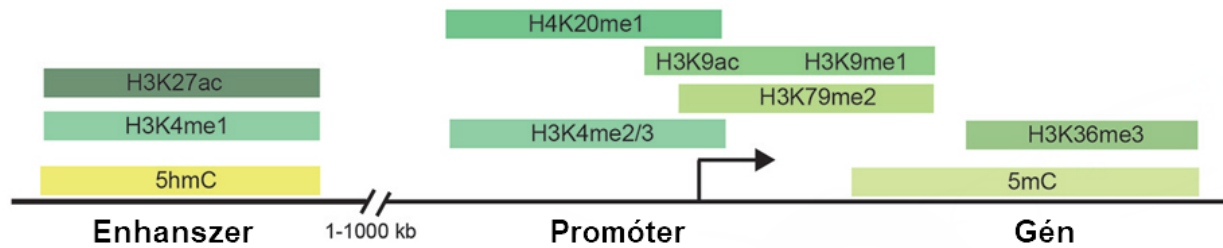
A hiszton fehérjék aminoterminális végei a kromatinkondenzáció során is mindvégig elérhetőek maradnak az enzimek számára. Ezek az ún. hiszton farkak különféle poszt-transzlációs módosításokon (PTM; post-translational modification) eshetnek át, melyek szerepet játszanak mind a kromatin kompaktságának és a nukleoszóma konformációjának modulálásában, mind a transzkripció szabályozásában. Ezek a módosítások, melyek között fellelhető az acetiláció, metiláció, foszforiláció, ubikvitináció, szumoiláció, ADP riboziláció és a deamináció, sokféleképpen befolyásolhatják a kromatin hozzáférhetőségét, és jelenlétük korrelálhat a gének transzkripció állapotával is, hiszen a nukleoszómák előnytelen pozícionálása korlátozhatja a transzkripció faktor motívumához való hozzáférést **(4. ábra)** (Kouzarides, 2007; Lawrence, Daujat, & Schneider, 2016).



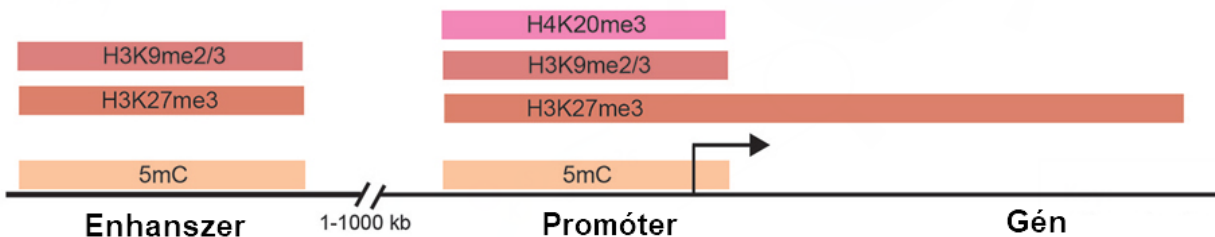
4. ábra. A hiszton végeken előforduló poszt-transzlációs módosítások sematikus ábrája. A H2A, H2B, H3 és H4 hisztonok N-terminális végeit érintő módosítások pozícióit fekete számok jelölik. Az adott pozícióban metilált, acetilált, foszforilált és ubikvitinált aminosavakat eltérő színű körök jelzik, míg a bennük található karakter magát az érintett aminosavat jelöli (K = lizin, R = arginin, S = szerin, T = treonin) (Lawrence et al., 2016).

Egyes hiszton módosítások genom-szintű feltérképezése alkalmas az enhanszerek és promóterek katalogizálására, továbbá azok transzkripcióra gyakorolt hatásának előrejelzésére. A H3 hiszton 27-es pozíciójában található lizin oldallánc acetilációja (H3K27ac) egyaránt jellemző az aktív enhanszerekre és aktív promóterekre, míg ugyanezen oldallánc metilációja (pl. H3K27me3) inaktív kromatint eredményez. A transzkripciósan aktív és represszált gének egyes szabályozó régióira jellemző – leggyakoribb – hiszton módosításait az **5. ábra** taglalja.

Aktív gén



Represszált gén



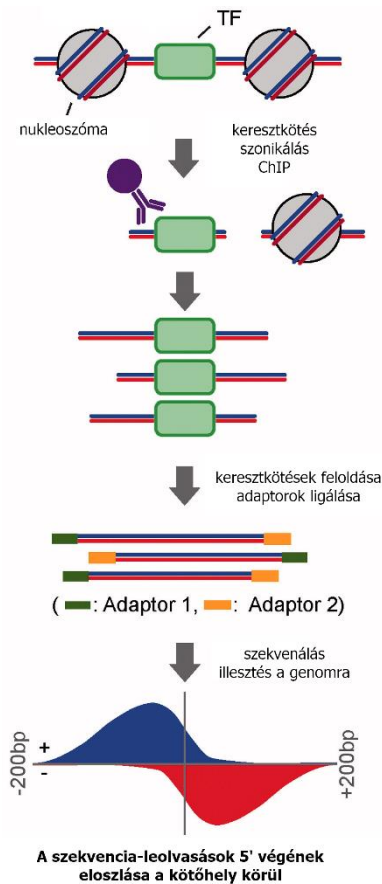
5. ábra. Az aktív és represszált gének enhanszer és promóter régióira, illetve a génre eső hisztonmódosítások.

Az aktívan átíródó gének promóterét a H3K27ac és a H3K4me1 mellett az 5hmC (5-hidroxi-metilcitozin) jellemzi, mindemellett a számos egyéb, magára a génre eső hisztonmódosítás segít facilitálni az iniciációt és az elongációt. A represszált gének esetében az 5mC (5-metilcitozin) mellett megfigyelhető, hogy mind az enhanszer, mind a promóter régiókon, illetve magán a génen is azonos módosítások fejtik ki gátló hatásukat. Fontos megjegyezni, hogy míg a H3K27me3 ideiglenes, a H3K9me3 szignál egy állandó jelzője a heterokromatikus állapotnak (Layman & Zuo, 2015).

1.6. ChIP-seq: a transzkripció faktor kötőhelyek genom-szintű azonosítása

Az újgenerációs szekvenálással kapcsolt kromatin immunprecipitációs (ChIP-seq; chromatin immunoprecipitation followed by sequencing) módszer alkalmas különböző kromatin-asszociált fehérjék, mint pl. a transzkripció faktorok vagy hiszton módosító enzimek teljes genom szinten történő detektálására, amennyiben azok a DNS-sel valamilyen interakcióban vannak (Johnson, Mortazavi, Myers, & Wold, 2007; Mardis, 2007). Ez az interakció jelenthet direkt DNS kötést, vagy fehérje-fehérje interakción keresztüli indirekt kapcsolódást is.

A módszer kivitelezése tipikusan 1-10 millió sejtet igényel, melynek során a vizsgálni kívánt fehérje jelenléte az ellene termeltetett antitest alkalmazásával detektálható. A ChIP-seq főbb lépéseit a **6. ábra** taglalja.



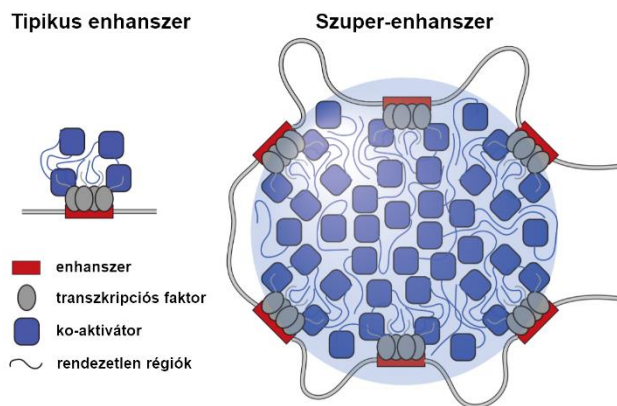
6. ábra. A ChIP-seq módszer főbb lépései.

A ChIP során első lépésben a fehérjét (beleértve a transzkripciós faktorokat) kovalensen keresztköti a DNS-sel, tipikusan formaldehiddel történő kezelés során. Ezt követően a kromatint rövid szakaszokra feldarabolják (pl. szonikálással), majd a keletkezett fragmenteket a vizsgálni kívánt fehérjére specifikus antitesttel immunprecipitálják. Az antitest által ki nem fogott, nem specifikus fehérje-DNS komplexek eltávolításra kerülnek, míg a specifikusak esetében a keresztkötéseket feloldják, és a fehérje által kötött DNS darabokat tisztítással visszanyerik. Az így kapott DNS-ből könyvtárat készítenek, és minden egyes molekulának meghatározzák a bázissorrendjét. A szekvencia leolvasásokat ezt követően a referencia genomra illesztik, majd a leolvasások – háttérhez képest történő – feldúsulása alapján meghatározzák a kötőhelyeket (Mahony & Pugh, 2015; Mardis, 2007).

1.7. Szuper-enhanszerek

A szuper-enhanszerek (SE-ek) olyan szabályozó egységei a genomnak, ahol az enhanszerek viszonylag kis régió belül nagy számban csoportosulnak, és mindez szokatlanul magas ko-faktor, különösképp Mediátor kötődéssel társul (Hnisz et al., 2013; Lovén et al., 2013; Whyte et al., 2013). A szuper-enhanszereken a Mediátor komplex 1-es alegysége (MED1) mellett számos más, megnövekedett enhanszer aktivitást kiváltó faktor együttes jelenléte is megfigyelhető; ilyenek pl. a hiszton acetiltranszferáz aktivitással rendelkező P300, a hiszton H3K27ac és

H3K4me1/2, mint aktív enhanszer- és promóter markerek, az acetilált hisztonokat felismerő BRD4 (Bromodomain 4) fehérje vagy maga a kromatin hozzáférhetősége (Di Micco et al., 2014; Ogryzko, Schiltz, Russanova, Howard, & Nakatani, 1996; Sengupta et al., 2015). A legújabb kutatások azt is kimutatták, hogy a szuper-enhanszerek ún. kondenzátumokba rendeződött, fázisszeperált multimolekuláris egységekben vannak jelen, lehetővé téve a biokémiai reakciók sejten belüli elkülönülését (7. ábra) (Sabari et al., 2018).



7. ábra. A szuper-enhanszerek fázisszeperációja. Az ábra a szuper-enhanszerek és ko-aktivátoraik által formált transzkripció apparátust reprezentálja. A tipikus enhanszerekénél csupán a klasszikus összeszerelődés figyelhető meg. A fehérjék közötti kapcsolatot mindkét esetben alacsony affinitású interakciók vagy rendezetlen régiók tartják fent (Sabari et al., 2018).

Létezik egy modell, mely szerint bizonyos transzkripciósan aktív gének ún. „transzkripciós gyárakban” helyezkednek el. Ezekben a gyárakban a kromatin hurkolódásának hasonlóan fontos szerepe lehet, mint a szuper-enhanszerek multimolekuláris egységeiben. Bennük az egymástól távol – akár másik kromoszómán – elhelyezkedő gének fizikailag közel kerülnek egymáshoz, a szintén a „gyáron” belül elhelyezkedő nagy mennyiségű fehérjék – esetenként szuper-enhanszerek – révén pedig egyidejűleg írodnak át. Mindez lehetővé teszi, hogy egy biológiai reakció során a kapcsolt funkcióval rendelkező génekről válaszként keletkezett fehérjetermékek koordinált időben és mihamarabb a citoplazmába jussanak a sejtmagból (Feuerborn & Cook, 2015; Osborne et al., 2004).

A szuper-enhanszerek a tipikus enhanszerekhez képest fontosságuk abban rejlik, hogy olyan gének szabályozásában vesznek részt, melyek a vizsgált sejt jellegét lényegesen meghatározzák (Hnisz et al., 2013). Dagantos sejtekben a szuper-enhanszereket olyan onkogén funkcióhoz társítható gének közelében azonosították, mint a c-Myc vagy a TAL1 (Mansour et al., 2014; Tolani, Gopalakrishnan, Punj, Matta, & Chaudhary, 2014). Egy 2014-es tanulmányban neuroblasztóma sejtekben gátolták a ciklin-dependens kináz 7 (CDK7; cyclin-dependent kinase 7) fehérje működését, amely közvetve nemcsak a Myc fehérjecsalád egyik tagjának, a MYCN-nek az expresszióját és a tumor növekedésének visszaesését modulálta, hanem olyan szuper-enhanszerek aktivitását is, melyek további fontos onkogéneket szabályoztak (Chipumuro et al., 2014; Kwiatkowski et al., 2014). Előzetes tanulmányok azt is kimutatták, hogy egy-egy szuper-enhanszer meglétének azonosítása alkalmas lehet annak megítélésére, hogy bizonyos gyógyszerek alkalmazása előnyös-e, vagy sem egy betegség, mint pl. az akut mieloid leukémia kezelése során (Ledford, 2018).

A szuper-enhanszereket számos tanulmányban a lókuszt kontroll régiókkal (LCR; locus control region) és a „stretch” enhanszerekkel azonosították (Q. Li, Peterson, Fang, & Stamatoyannopoulos, 2002; Parker et al., 2013). Mivel mindhárom fogalom esetében enhanszer klaszterekről beszélhetünk – melyek egy hányada át is fed egymással –, a szuper-enhanszerek egyedi szerepét nagyon sokan vitatják. Míg a szuper-enhanszereket az aktív kromatin és hiszton markerek jelenléte, sejt-specifikus transzkripciós faktorok és azok motívumai jellemzik, a „stretch” enhanszerek a TSS-től rendszerint távol elhelyezkedő, legalább 3 kb méretű szabályozó régiók.

A „stretch enhanszerek a szuper-enhanszerekhez képest kevésbé konzerváltak és alapvetően nem mutatnak aktivitást, de megfelelő szignál révén aktiválhatóak („poised” enhanszerek). Az általuk elfoglalt régiók mérete a genomban a szuper-enhanszerek által elfoglalt

régiók kétszeresére tehető, így a szuper-enhanszerek 85%-a a „stretch” enhanszerek csupán kis hányadával (13%-ával) mutat átfedést (Khan, Mathelier, & Zhang, 2018). A lókuszt kontroll régiók általában több gént szabályoznak, melyek a fejlődés vagy a differenciáció során különböző mértékben expresszálódnak. Bár a LCR, mint fogalom, arra enged következtetni, hogy LCR-ek a szuper- vagy „stretch” enhanszerekhez hasonlóan egy enhanszer klaszterként funkcionálnak, több tanulmány szerint az őket felépítő egyedi enhanszerek egymástól távol eső, intronikus és intergenikus eloszlást mutathatnak (Q. Li et al., 2002; Parker et al., 2013).

Egy másik, sokak által vitatott kérdés, hogy vajon az egyedi enhanszerek aktivitása függ-e a szomszédos enhanszerek aktivitásától; hatásuk egy klaszteren belül a gének expressziójára nézve additív-e, vagy egymást erősítve, szinergista módon értelmezhető (Hay et al., 2016; Shin et al., 2016).

Ma már a szuper-enhanszereknek saját adatbázisaik is vannak, ahol több száz sejtvonal adatai között válogathatunk (Jiang et al., 2019; Khan & Zhang, 2016; Wei et al., 2016). Bár azonosításuk egy vagy akár több faktor együttes jelenléte alapján történhet, a szuper-enhanszerek azonosításának alapja minden esetben azonos: azon enhanszercsoportok kiszűrése a cél, melyek a sok fehérje jelenléte miatt a szekvenálási minták alapján a legnagyobb csúcsokat mutatják.

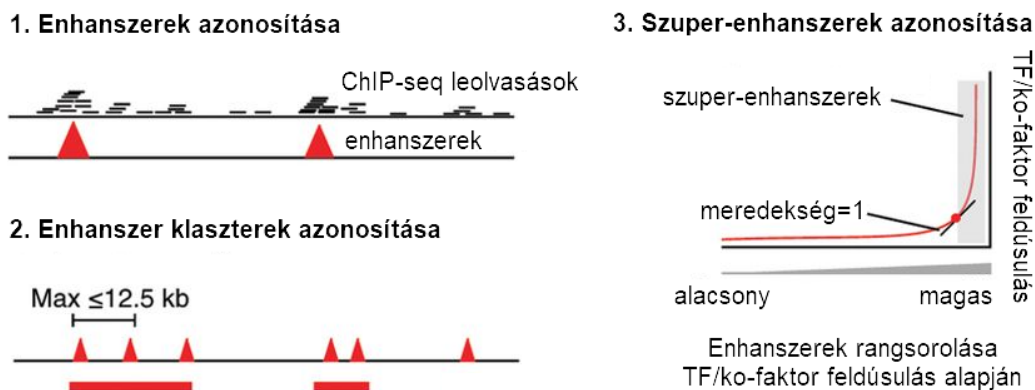
1.7.1. Szuper-enhanszerek azonosítása

A két legszélesebb körben alkalmazott szuper-enhanszereket azonosító program a Rank Ordering of Super-Enhancers (ROSE), illetve a Hypergeometric Optimization of Motif EnRichment (HOMER) (Heinz et al., 2010; Whyte et al., 2013). Az azonosításhoz kiindulhatunk a ChIP-seq kísérletből származó (mester) transzkripciós faktor vagy ko-faktor (MED1, BRD4) kötőhelyekből, vagy alapul vehetjük a H3K27ac-jelölt nukleosómákkal fedett régiókat is –

utóbbi a legáltalánosabban elfogadott módszer **(8. ábra)**. Első lépésben az egyedi enhanszer elemeket a fent említett valamelyik program segítségével csoportokba rendezzük, úgy, hogy a két enhanszer közötti távolság ne haladja meg a 12,5 kb-t. Második lépésben ezeket az enhanszer csoportokat a szekvencia leolvasásaik sűrűsége (denzitása) alapján sorbarendezzük, majd azon pont alatt, ahol a meredekség nem haladja meg az 1-et, az enhanszereket tipikus enhanszereknek tekintjük, a fölötte levő enhanszer klasztereket pedig szuper-enhanszereknek **(8. ábra)** (Pott & Lieb, 2015).

A legáltalánosabban használt faktorok a szuper-enhanszerek azonosítása során:

H3K27ac, MED1, BRD4, P300, vagy a vizsgált sejt mester transzkripció faktor



8. ábra. A szuper-enhanszerek azonosításának lépései. Az enhanszerek elsődleges azonosítása ChIP-seq kísérletből származó transzkripció faktor (TF), ko-faktor vagy hiszton csúcsok (régiók) alapján történik. Az enhanszer klaszterek meghatározásánál az egyetlen kitétel az, hogy a két szomszédos enhanszer közötti távolság ne legyen nagyobb, mint 12,5 kb. A szekvenálási leolvasások alapján számolt lefedettség értékek alapján sorbarendezt enhanszerek azon hányadát, melyek meredeksége meghaladja az 1-et, a program szuper-enhanszerként azonosítja (Pott & Lieb, 2015).

Összességében néhány száz szuper-enhanszer azonosítható egy-egy kísérletből, melyek száma vagy kezdő- és végpontja a két meghatározási módszerrel – az eltérő matematikai megközelítés végett – részben eltérhet.

1.8. A magreceptorok

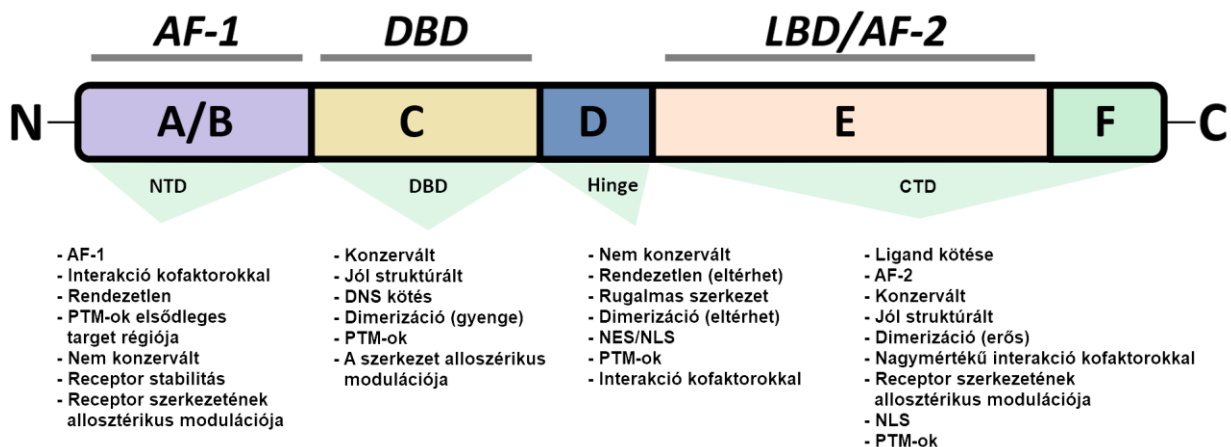
A 19. század végén *Sir George Beatson* elsőként számolt be arról, hogy azon nők körében, akiknek a petefészkük eltávolításra került, sokkal kisebb valószínűséggel alakul ki emlőrák a hátralévő életük során. Ezzel egyidejűleg rámutatott arra, hogy egy szerv képes egy másik és egyben különálló szerv szekrécióját irányítani, továbbá elsőként feltételezte az emlőrák és a hormonok közötti kapcsolatot (Beatson, 1896). Az 1920-as évek végén bár izolálták az ösztrogén három formáját is, az, hogy ezek a hormonok miképp fejtik ki hatásukat, egy időre még megválaszolatlan maradt (Doisy, Veler, & Thayer, 1929). Néhány évtizeddel később *Elwood V. Jensen* és *Jack Gorski* azonosítottak egy hormon-kötő komponenst, amely képes volt az ösztradiollal interakcióba lépni – ekkorra tehető a magreceptorok első tagjának, az ösztrogén receptornak (ER) a felfedezése (Gorski, Toft, Shyamala, Smith, & Notides, 1968; E. V. Jensen et al., 1968).

A magreceptor szupercsalád jelenleg emberben ismert 48 tagja számos fiziológiai folyamat szabályozásában vesz részt. Ilyenek például a metabolizmus, sejtproliferáció, gyulladás, vagy épp a cirkadián ritmus, de fontos szereppel bírnak olyan patológiai folyamatokban is, mint a metabolikus szindróma, kardiovaszkuláris megbetegedések, vagy a legtöbb daganatos megbetegedés (Gerhart-Hines et al., 2013; Glass & Ogawa, 2006; Oyekan, 2011; Ranhotra, 2013; Yang, Lamia, & Evans, 2007). Központi szerepüket az a tény is alátámasztja, hogy a farmakológiai gyógyszerek megközelítőleg 13%-ának célpontját a magreceptorok teszik ki (Overington, Al-Lazikani, & Hopkins, 2006).

A magreceptorok többsége ligand-kötő doménjén (LBD; ligand-binding domain) keresztül kisméretű, zsíroldékony jelátviteli molekulákkal (pl. szteroid hormonokkal) képes közvetlen kapcsolatot kialakítani. Ezek a ligandok – zsíroldékony tulajdonságuk révén – a lipid kettősréteg permeabilitását kihasználva azon könnyen átjutnak. A létrejött hormon-receptor

komplex számos ko-regulátor (aktivátor vagy represszor) fehérjével kiegészülve a sejtmagban a DNS egy specifikus szekvenciájához (a receptor saját válaszadó eleméhez) közvetlenül képes kötődni és szabályozni a célgén expresszióját (Kushiro, Nambara, & McCourt, 2003). Néhány magreceptor azonban nem rendelkezik ligand-kötő árokkal, ezért szabályozásuk valamilyen más mechanizmus által valósul meg.

A magreceptorok a LBD mellett általában négy további doménnel rendelkeznek. Egy konzervált DNS-kötő domén (DBD; DNA-binding domain) révén a receptor közvetlenül képes kötődni a DNS egy specifikus szekvenciájához; emellett jelen van még egy N-terminális, egy C-terminális, valamint egy összekötő (Hinge) régió is (9. ábra).



9. ábra. A magreceptorok általános felépítésének vázlata. Az N-terminális domén (NTD), a DNS-kötő domén (DBD), a Hinge régió és a C-terminális domén (CTD) mind eltérő módon járulnak hozzá a receptor működéséhez. Funkcióik és fontosabb tulajdonságaik az egyes domének alatt pontokba szedve olvashatóak. Fontos megjegyezni, hogy nem minden funkció vonatkozik minden magreceptorra. NLS: magi lokalizációs szignál (nuclear localization signal), NES: magi export szignál (nuclear export signal), PTM: poszt-transzlációs módosítás (post-translational modification) (Aagaard, Siersbæk, & Mandrup, 2011; Cui, Shen, & Li, 2013).

Az N-terminális régió szekvenciáját tekintve variábilis és tartalmazza az AF-1 (activation function 1) transzaktivációs régiót, mely lehetővé teszi az interakciót más transzkripció

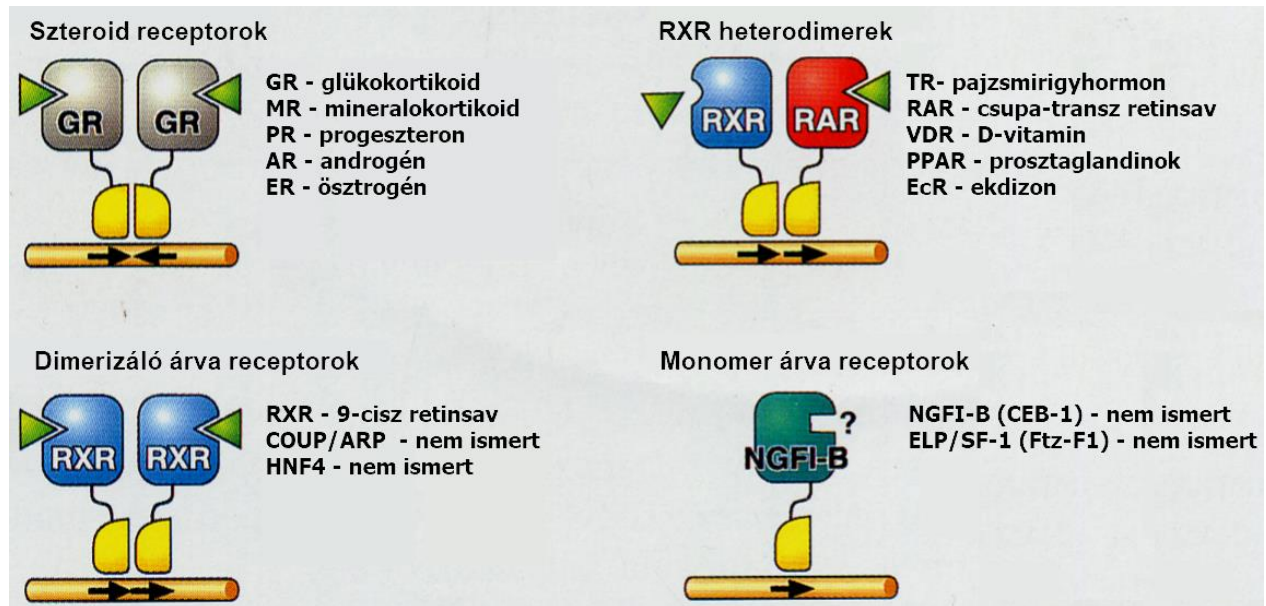
faktorokkal, hatása pedig független a ligand jelenlététől. A C-terminális régió mindamellett, hogy mediálja a ligand kötődését, a receptor dimerizációját és a magi transzlokációt, magában foglalja a ligand-függő aktivációs domént (AF-2) is, amely a transzkripció aktiválásában vesz részt. A Hinge elsődlegesen ún. kapocs szerepet tölt be a DBD és LBD között (Giguère, Hollenberg, Rosenfeld, & Evans, 1986; Mangelsdorf et al., 1995).

Habár a magreceptorok a doménszekvenciájukat tekintve nagymértékű homológiát mutatnak (pl. az ösztrogén receptor alfa és béta ($ER\alpha$ és $ER\beta$) szekvenciája között 96%-os a hasonlóság), a kisebb-nagyobb eltérések teszik lehetővé, hogy eltérő ligandokat ismerhessenek fel és más-más koregulátorokkal lépjenek kölcsönhatásba. Mindez nagyban elősegíti a különböző célgének idő- és szövetspecifikus módon történő szabályozását és a transzkripciós folyamatok finomhangolását.

1.8.1. A magreceptorok csoportosítása

A magreceptor szupercsalád tagjainak osztályozása történhetne akár filogenetikai megközelítés alapján is, de a szakirodalom négy különböző osztályt (I-IV.) különít el a ligand-specificitásuk, dimerizáló és DNS-kötő képességeik alapján (Mangelsdorf et al., 1995; Nuclear Receptors Nomenclature Committee, 1999). Az I. osztály azon szteroid hormon receptorokat foglalja magába, melyek – javarészt – természetes ligandjaik általi aktiválás következtében homodimerként kötik a DNS-t, mégpedig két fordítottan ismétlődő (inverted repeat; IR) magreceptor félhelyen keresztül. Ebbe az osztályba tartoznak a glükokortikoid (GR), mineralokortikoid (MR), progeszteron (PR), androgén (AR) és ösztrogén (ER) receptorok. A II. osztály tagjait azok a további ligand-függő receptorok adják, melyek a Retinoid X Receptorral (RXR-val) heterodimerizálni képesek. Ide tartoznak a tiroid hormon (TR), retinsav (RAR), D-

vitamin (VDR), ekdizon (EcR) és a peroxiszóma proliferátor aktivált (PPAR) receptorok. Jellemző módon azonos irányban ismétlődő (direct repeat) félhelyeken keresztül kötik a DNS-t. A III. és IV. osztályba olyan árva (ismert természetes ligand nélküli) receptorok tartoznak, melyek vagy homodimerként képesek a direct repeat elemekhez kötődni, vagy monomerként egyetlen félhelyhez kötődnek (**10. ábra**) (Mangelsdorf et al., 1995; Olefsky, 2001).



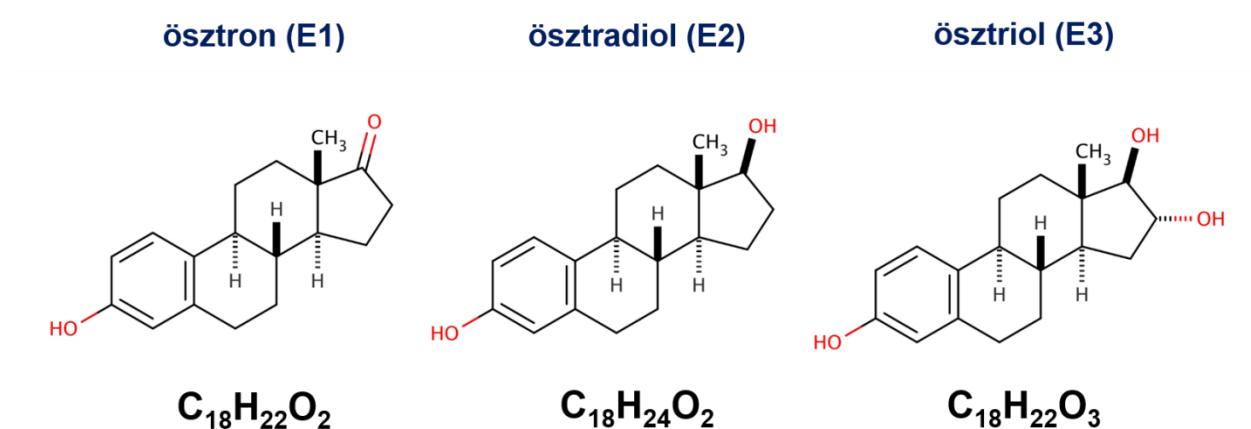
10. ábra. A magreceptor szupercsalád tagjainak négy fő osztályba való besorolása. Az ábra reprezentálja a szteroid hormon receptorok (I. osztály), az RXR heterodimerek (II. osztály), a dimerizáló árva receptorok (III. osztály) és a monomer árva receptorok (IV. osztály) legfontosabb jellemzőit. Az egyes osztályok mellett a legismertebb és/vagy legfontosabb receptorok vannak feltüntetve, továbbá azok természetes vagy elsődleges ligandjai. A nyilak a magreceptor félhelyek irányultságát jelölik (Mangelsdorf et al., 1995). A *magreceptor szupercsalád 48 tagját és azok ligandjait a Függelék tartalmazza.*

Az árva receptorok a magreceptor szupercsalád többi tagjával bár nagy szerkezeti hasonlóságot mutatnak, természetes ligandjuk nem ismert, vagy legalábbis a felfedezésük idején még nem volt az (Evans, 1988). Ezeknek a receptoroknak vagy hiányzik a ligand-kötő zsebük, vagy valamilyen más mechanizmus révén, pl. poszt-transzlációs módosítás által szabályozódnak,

de ismert, hogy néhány közülük a környezetükben lévő xenobiotikumokra reagál – így megteremtve a genom és a környezet közötti kapcsolatot (Wang et al., 2003; Xie & Evans, 2001). Fontos megjegyezni, hogy a magreceptor félhelyek túlságosan gyakran előforduló motívumok a genomban ahhoz, hogy önmagukban specifikusan működhessenek; a legtöbb monomer árva receptort az AGGTCA szekvencián felüli bázisok jelenléte teszi specifikussá (Johnston et al., 1997; Wilson, Paulsen, Padgett, & Milbrandt, 1992).

1.9. Az ösztrogén receptor alfa

Az 1920-as évek végén *Edward A. Doisy*, *Clement D. Veler* és *Sidney A. Thayer* az elsők között izolálták az ösztrogének közül az ösztront (E1) és az ösztriolt (E3), majd később Doisy az ösztradiolt (E2), melyek alapvetően hidroxilcsoportjuk számában és a receptor iránti affinitásukban mutatnak eltérést (**11. ábra**). Jóllehet az E1 és E3 is nagy affinitással rendelkező ligandok, az ösztrogén receptoroknak (ER) sokkal gyengébb agonistái az E2-höz képest (Kuiper et al., 1997).



11. ábra. A legfontosabb ösztrogén vegyületek képletei. Az ösztrogének három fő fajtája az ösztrol (E1), ösztradiol (E2) és az ösztriol (E3), melyek a hidroxilcsoportjaik számában és a receptor iránti affinitásukban mutatnak különbséget.

Az E2 (17 β -ösztradiol) fiziológiai funkcióinak szabályozásáért alapvetően a két ösztrogén receptor, az ösztrogén receptor alfa (ER α) és béta (ER β) felelnek. Míg utóbbi a prosztata epitél sejtjeiben, a húgyhólyagban, a petefészek granulóza sejtjeiben, a vastagbélben és a zsírszövetben fejeződik ki leginkább, és markáns hatást képes kifejteni a központi idegrendszerben és az immunrendszerben, a magreceptor szupercsalád talán egyik legintenzívebben tanulmányozott tagja, az ER α leginkább az emlőmirigyben, méhben, a petefészek theca sejtjeiben, a csontban, májban, zsírszövetben, továbbá a férfi reproduktív szervekben és a prosztata sztróma állományában van jelen. Kiemelt szerepe ugyanakkor annak tudható be, hogy az emlőrákos nők háromnegyedében kulcsfontosságú transzkripciós faktor, és a széles körben alkalmazott anti-ösztrogén hatású készítmények, mint például a tamoxifen, csak a hormonreceptor-pozitív daganatok esetében képesek hatásukat kifejteni (Davies & Hiscox, 2011; Enmark et al., 1997; Gustafsson, 2003; Harris, 2007; Heldring et al., 2007; Elwood V Jensen & Jordan, 2003; Writing Group for the Women's Health Initiative Investigators et al., 2002). Az ER α -nak három izoformája ismert az emberi szervezetben, melyek közül a leghosszabb – amely rendelkezik az NTD-nel is – 595 aminosavból áll (Green et al., 1986). Mérete 67 kDa, és a 6-os kromoszóma rövid karján (6q25.1) elhelyezkedő *ESR1* (Estrogen Receptor 1) gén kódolja.

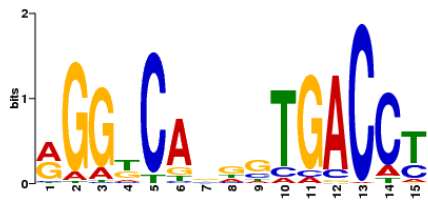
1.9.1. MCF-7 sejtvonalt: modell az ER α tanulmányozásához

A legáltalánosabban használt sejtvonalt az ER α általi génszabályozás *in vitro* tanulmányozásához az emlő eredetű MCF-7 (Michigan Cancer Foundation-7) sejtvonalt, amely egy 69 éves metasztatikus emlőrákkal diagnosztizált nő, *Helen Marion* pleurális effúziójából (mellkasi folyadékgyüleméből) származik. 1973-ban *Herbert Soule* volt az, aki stabilizálta a sejtvonaltat, de annak ellenére, hogy tisztában voltak a szöveti eredetével, csupán a sejtek metabolizmusának és

transzportmechanizmusainak vizsgálatára használták. Később *Kathryn Horwitz* volt az első, aki szorgalmazta, hogy az MCF-7 kiváló *in vitro* modellként szolgálhat a emlődaganat endokrin terápiára adott válaszának, illetve a ER α E2 általi szabályozásának vizsgálatában is (Horwitz, Costlow, & McGuire, 1975; Levenson & Jordan, 1997; Soule, Vazquez, Long, Albert, & Brennan, 1973). Az MCF-7 a mai napig számos genomikai tanulmány alapját szolgálja; az első ChIA-PET kísérletet – mely a genom-szintű ER-vezérelt DNS hurkolódást vizsgálta – is MCF-7 sejtekben végezték el (Fullwood et al., 2009).

1.9.2. Az ösztrogén receptor alfa általi génszabályozás lehetséges mechanizmusai

Az ER α ligand hiányában, bár túlnyomórészt a sejtmagban lokalizálódik, valójában a sejtmag és a citoplazma között mozog (Dauvois, White, & Parker, 1993). Az ösztradiol általi aktivációja során (mely $K_d=10^{-10}$ M affinitással kötődik a receptor ligand-kötő zsebéhez) a receptor elsőként disszociálódik a hsp90 hősokkfehérjéről, amely mindaddig meggátolta hatásának kifejezésében, majd a LBD konformációváltozása révén a receptor a citoplazmából a sejtmagba transzlokálódik, ahol közvetlenül képes kötődni a DNS bizonyos szabályozó régióihoz, az ösztrogén válaszadó elemekhez (ERE, estrogen response element) (Abderrahman & Jordan, 2018). Az ERE-t elsőként 1986-ban a *Xenopus* vitellogenin A2 génjének promóter régiójában azonosították, majd 1993-ban *John W. Schwabe* volt az, aki meghatározta az ER α DBD-je és a DNS közötti komplex teljes kristályszerkezetét, mely egyértelműsítette, hogy az ER α szimmetrikus dimerként köt egy olyan palindrom kötőhelyhez, amely két magreceptor félhelyet tartalmaz, közöttük három véletlenszerű bázispárral (**12. ábra**) (Klein-Hitpass et al., 1986; Klinge, 2001; Schwabe et al., 1993; Truong et al., 1973).



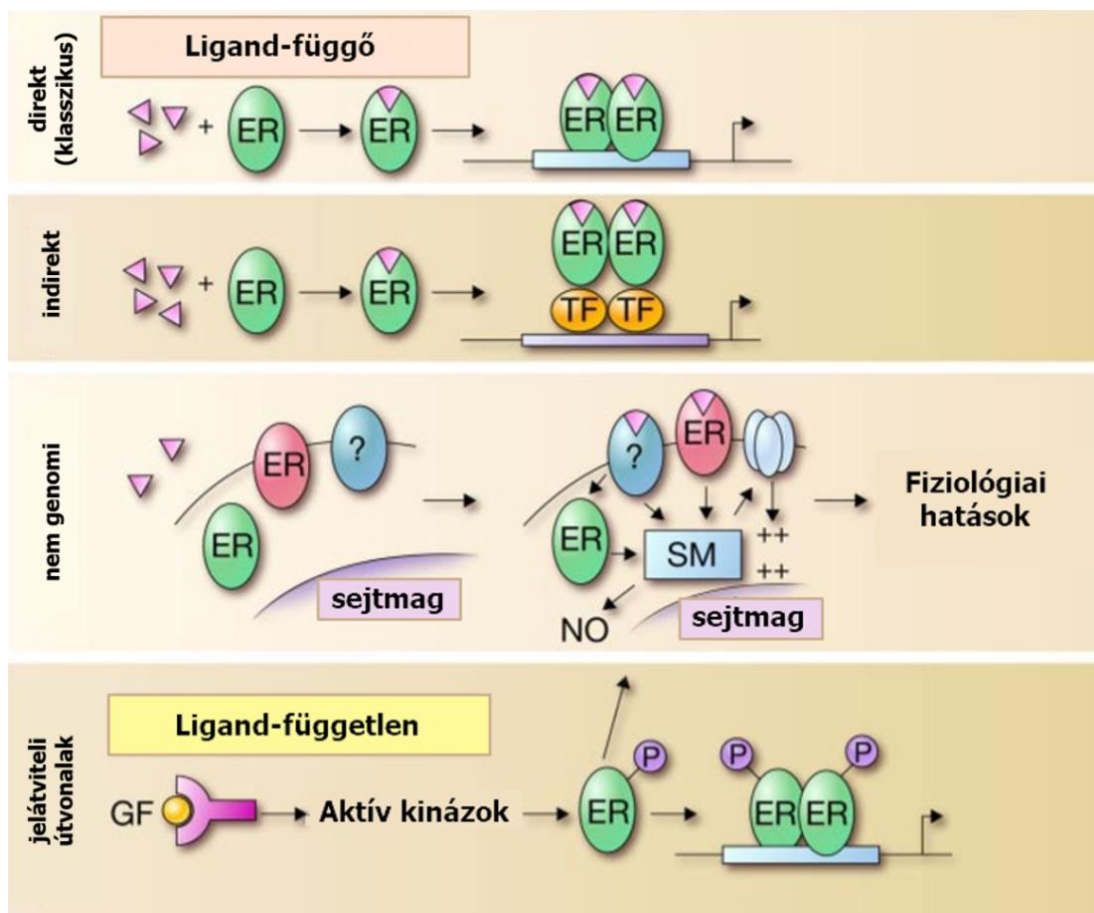
12. ábra. Az ösztrogén válaszadó elem (ERE) sematikus ábrázolása.

Az ERE-t felépítő AGGTCA_{nnn}TGACCT palindrom konszenzus szekvencia logója. Az egyes karakterek magassága a nukleotidok információtartalmával arányosan súlyozott (*forrás: factorbook.org*).

Az emberi genomban megközelítőleg egy millió ERE található, de az ER α egyidejűleg csak egy nagyon kis hányadukon keresztül fejt ki hatását (Vega et al., 2006). Ennek az elsődleges magyarázata az lehet, hogy az ERE-k egy része a fehérje számára nem elérhető, zárt szerkezetű kromatin régióin belül helyezkedik el, de magyarázható a jelenség az ERE-k szekvenciájában lévő eltérésekkel, vagy azzal is, hogy a különböző sejtek eltérő, sejt-specifikus enhanszerkészlet révén szabályozzák a működésükhöz szükséges géneket, és más-más ko-faktorok jelenléte szükséges a kromatin hozzáférhetőségének elősegítéséhez.

Az ER α számos esetben képes olyan géneket is szabályozni, amelyek szabályozó régióiban nincs jelen az ERE. Ebben az esetben hatását más transzkripciós faktor komplexeken keresztül, ún. fehérje-fehérje interakció révén, a DNS-hez indirekt módon kötődve fejt ki, amiben segítségére lehetnek olyan ismert fehérjék, mint például az AP-1 vagy Sp-1 (Safe, Kim, & Kim, 2008). Mindemellett a receptor ligand általi aktivációja elősegítheti számos további ko-regulátor, illetve az RNS polimeráz kötődését is.

Habár az ER α elsődlegesen ligand-függő DNS-kötő transzkripciós faktorként ismert, ligand nélküli formában történő kötődését a promóterhez már nagyon korán, 1990-ben kimutatták (Murdoch, Meier, Furlow, Grunwald, & Gorski, 1990). Ligand-független aktivitása számos intracelluláris jelátviteli útvonalon keresztül, a receptor protein kinázok általi foszforilációja révén valósulhat meg (**13. ábra**).



13. ábra. Az ER α általi génszabályozás lehetséges mechanizmusai. A klasszikus (direkt), ligand-függő szabályozás során a hormon-receptor komplex kialakulását követően a receptor közvetlenül kötődik az ösztrogén válaszadó elemhez (ERE). Az indirekt, vagy kihorgonyozást igénylő szabályozás más fehérjék elsődleges kötésén keresztül jön létre. A nem genomi szabályozás során a ligand bár aktiválja a receptort, az vagy a membránhoz kötött formában van jelen, vagy a klasszikus receptor helyett egy izoformáját aktiválja. Az is előfordulhat, hogy a természetes ligandja helyett egy teljesen más jel fogja a receptort a citoplazmában aktiválni, továbbá a jelátviteli kaszkádot másodlagos hírvivő molekulák (SM; second messengers) fogják beindítani, melyek vagy ioncsatornákat modulálnak, vagy megnövelik a nitrogén-oxid (NO) szintjét a citoplazmában, így produkálva gyors fiziológias választ anélkül, hogy mindez befolyásolná a célgén expresszióját. A ligand-független mechanizmus létrejöhet más jelátviteli útvonalak, például a növekedési faktor receptor jelátvitelére is. Ezen folyamat során az aktivált kinázok foszforilálják a receptorokat, így elősegítve a dimerizációjukat és a DNS-hez való kötődésüket a célgén szabályozásának érdekében (Heldring et al., 2007).

1.9.3. Az ösztrogén receptor alfa partnerei a transzkripció szabályozásban

Az ER α mellett, hogy működését különféle ko-faktorok vagy ko-faktor komplexek vezérlik (melyek egyaránt lehetnek ko-aktivátor vagy ko-represszor fehérjék/fehérje komplexek), képes számos DNS-kötő transzkripció faktorral is interakciót kialakítani a gének megfelelő működésének szabályozása érdekében. Ezek a transzkripció faktorok működhetnek az ER α kollaboratív faktoraként, megkönnyíthetik annak kötődését a DNS-hez való kihorgonyozása révén, de lehetnek előkészítő, ún. „*pioneer*” faktorai is. Az ER α -val interakciót kialakítani képes transzkripció faktorok közé tartoznak a Forkhead box 1 (FoxA1), Activator Protein 2 gamma (AP2 γ), GATA-binding protein 3 (GATA3), Retinsav receptor alpha/gamma (RAR α/γ), Activator protein 1 (AP-1), vagy a Signal transducer and activator of transcription 1 (STAT1) fehérjék (Candelaria, Liu, & Lin, 2013; Z. Liu et al., 2014).

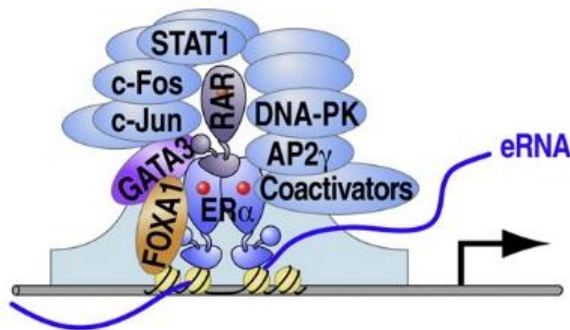
A legfontosabb közülük is a FoxA1, amely a „szárnyas hélix” DNS-kötő doménjén keresztül (amely nevét a H₃ felismerő hélixet szárnyként (*wing*) ölelő W₁ és W₂ hurkok konformációjáról kapta) képes a zárt kromatinhoz hozzáférni és imitálni a linker hisztonokat – így működik *pioneer* faktoraként az ER α -nak (L A Cirillo et al., 1998; Lisa Ann Cirillo et al., 2002; Clark, Halay, Lai, & Burley, 1993). A FoxA1 többnyire elengedhetetlen az ER α -kromatin kölcsönhatás létrejöttéhez és az ER α célgének szabályozásához. Az ER α transzkripció faktor kötőhelyek 50%-ának közelében megtalálható a FoxA1 motívuma, de hatását távoli interakció révén is képes kifejteni. Képes továbbá előrehaladott betegségekből az ER α kötődésének átprogramozására is (Carroll et al., 2005, 2006; Hurtado, Holmes, Ross-Innes, Schmidt, & Carroll, 2011; Ross-Innes et al., 2012). Fontos megjegyezni, hogy a FoxA1 gén expressziója pozitív prognosztikai markerként is működik az ER⁺ emlődaganatos betegek körében és expressziós szintje pozitívan korrelál az endokrin terápia iránti érzékenységgel (Badve et al., 2007).

A FoxA1-hez hasonlóan az AP2 γ és GATA3 fehérjék motívuma is jelentős feldúsulást mutat az ER α által kötött genomi régiókban (Lin et al., 2007). ChIA-PET adatok alapján az E2 kezelést követően távoli interakciót mutató ER α kötőhelyek 88%-ánál kimutatható volt az ER α , FoxA1 és GATA3 fehérjék együttes előfordulása, mely arra engedett következtetni, hogy a három transzkripciós faktor egy komplexet alkotva vesz részt az E2 kezelésre adott transzkripciós válaszok finomhangolásában (Fullwood et al., 2009; Kong, Li, Loh, Sung, & Liu, 2011).

Tan és munkatársai ugyanakkor kimutatták, hogy az AP2 γ a FoxA1-gyel közösen képes elfoglalni az ER α kötőhelyek jelentős részét már a ligandkezelést megelőzően is, és nélkülözhetetlenek az ER α -mediált intrakromoszómális interakciók jelentős hányadában. Mindemellett elsődlegesen az AP2 γ jelenléte szükséges ehhez a ko-lokalizációhoz, felvetve annak lehetőségét, hogy a két transzkripciós faktor közösen, egy pioneer komplexként képes facilitálni az ER α kötődését (Tan et al., 2011).

A FoxA1, GATA3 és AP2 γ fehérjék önálló és/vagy együttes szabályozó szerepét további transzkripciós faktorok is képesek fokozni. Az ERE-n keresztül létrejövő legerősebb ER α -vezérelt enhanszerekre jellemző például, hogy az E2 kezelést követően képesek egy 1-2 MDa méretű, DNS-kötő transzkripciós faktorokból álló komplexet kialakítani. Ezen enhanszereket „mega transcription factor-bound in *trans*”, azaz MegaTrans enhanszereknek neveték el. A MegaTrans komplex részét képezi többek közt az AP-1 transzkripciós faktor család két tagja, a c-Fos és a c-Jun, a STAT1 és az RAR α vagy RAR γ , mely utóbbiaknak a géncsendesítése egyenesen megszünteti a komplex további komponenseinek kötődését, és gátolja a komplex célgénre gyakorolt hatását is. A komplex kialakulása nem csupán az enhanszer RNS-ek (eRNS) átíródásához szükséges, de lehetővé teszi különféle ko-aktivátorok és enzimek, mint például a P300 és a MED1 fehérjék vagy protein kinázok kombinatorikus kötődését is a

komplexhez, és a számos „összeszerelő” forma tovább képes fokozni az enhanszer aktivitását és a transzkripciót (14. ábra) (Z. Liu et al., 2014).



14. ábra. A ERα-vezérelt MegaTrans enhanszer felépítése. Az ERE-hez kötődő ERα és a közelében található FoxA1 fehérje, mely az ERα kötését facilitálja, további DNS-kötő fehérjéket (GATA3, AP2γ, RARα vagy RARγ, c-Jun, c-Fos, STAT1) és ko-aktivátorokat rekrutálnak egyazon genomi régióhoz és létrehozzák a MegaTrans enhanszert (Z.

Liu et al., 2014).

Míg a szuper-enhanszerek egyedi enhanszereket foglalnak magukba és főként a sejt fejlődéséhez nélkülözhetetlen, illetve betegség-asszociált gének szabályozásában vesznek részt, a MegaTrans komplex azon egyedi enhanszerek kijelölésében játszik szerepet, melyek az adott sejtípus vonatkozásában kulcsfontosságú szereppel bírnak.

2. Célkitűzés

I. A ligand-indukálható szuper-enhanszerek működésének vizsgálata

Publikusan elérhető ER α transzkripciós faktor ChIP-seq adatok (újra)elemzése során megfigyeltük, hogy a specifikus ösztadiol kezelést követően aktiválódó ER α -vezérelt szuper-enhanszerek már a ligand általi stimulust megelőzően is jellemezhetőek egy, maximum néhány olyan kötőhellyel, amelyet ER α foglal el.

Hipotézis: az elsőként aktiválódó kötőhelyek alapját képezik a későbbiekben a ligand hatására kialakuló szuper-enhanszereknek.

Hogy megbizonyosodjunk, szerepet játszanak-e az elsődleges enhanszerek a szuper-enhanszer jellegű együttműködés inicializálásában, megvizsgáltuk, van-e különbség:

1. az elsődleges (anya vagy „mother”) és másodlagos (leány vagy „daughter”) enhanszerek ligand kezelésre adott válaszában;
2. a kötődésük dinamikájában (időbeli követéssel);
3. az általuk elfoglalt DNS szekvencia motívumokban;
4. ko-faktoraik és ko-regulátoraik között; valamint hogy
5. megfigyelhető-e ez a jelenség más, ligand-indukálható transzkripciós faktorok esetében is?

Mivel az MCF-7 sejtvonal egy nagyon fontos modell, sok kutatócsoport végzett benne az ER α transzkripciós faktorra és annak ko-faktoraira irányuló genom-szintű vizsgálatokat, melyek nyers adatai bárki számára elérhetőek a különféle adatbázisokban. Eredményeink teljes körű bemutatására ezért mi is az ER α -t választottuk elsődleges modellünknek.

II. Az ER α cisztróm karakterizálása az ösztrogén válaszadó elem (ERE) jelenlétének és a kötőhelyek aktiválhatóságának függvényében

Közel 80000 ER α kötőhelyet osztályoztunk azon tulajdonságaik alapján, hogy rendelkeztek-e ERE-vel, ill. kötötte-e őket ER α ösztradiol hiányában is vagy csak annak hozzáadását követően. Ennek során további 6535 olyan kötőhely meglétét azonosítottuk a genomban, melyek az anya enhanszerekhez hasonlóan rendelkeztek ERE-vel és detektálható volt rajtuk az ER α már a specifikus ligand kezelést megelőzően is.

Célunk volt annak meghatározása, hogy:

1. van-e különbség az anya enhanszerek és a hozzájuk hasonló, nem szuper-enhanszerek részét képező ER α kötőhelyek génexpresszióra gyakorolt hatásában; továbbá
2. van-e különbség a különböző besorolású kötőhelyek által preferált DNS szekvencia motívumokban, társfaktoraikban, továbbá a génexpresszióra gyakorolt hatásukban.

3. Anyagok és módszerek

3.1. Adatgyűjtés

3.1.1 Szekvenálási adatok az ligand-indukálható szuper-enhanszerek karakterizálásához

A ligand-indukálható szuper-enhanszerek karakterizálásához a Gene Expression Omnibus (GEO) és a Sequence Read Archive (SRA) adatbázisokból összesen 8, humán és egér sejtekből és sejtvonalakból származó ChIP-seq, továbbá a humán MCF-7 sejtvonalból származó ChIP- és DNáz-seq minta kezeletlen, oldószerrel (*etanol (EtOH)* vagy *dimetil-szulfoxid (DMSO)*), természetes liganddal (*ösztadiol (E2)*, *dihidrotesztozteron (DHT)*, *lipopoliszaharid (LPS)*, *csupa-transz retinsav (ATRA)* vagy *kalcitriol (D3)*) vagy szintetikus ER α agonistával (*tamoxifen (Tam)* vagy *fulvestrant (Fulv)*) kezelt adatait töltöttünk le és dolgoztunk fel, melyek listáját az **1.** és a **2. táblázat** tartalmazza.

Transzkripciósfaktor	Sejt/Sejtvonal	Kezelés	Faj	Azonosító (oldószeres)	Azonosító (kezelt)	Referencia
ER α	MCF-7	E2	humán	GSM614611	GSM614610	(Schmidt et al., 2010)
ER α	MCF-7	E2	humán	GSM365925	GSM365926	(Welboren et al., 2009)
FoxA1	MCF-7	E2	humán	GSM588929	GSM588930	(Tan et al., 2011)
AP2 γ	MCF-7	E2	humán	GSM1469997	GSM1469998	(Z. Liu et al., 2014)
AR	LNCaP	DHT	humán	GSM1527822	GSM1527834	(Takayama et al., 2015)
JUNB	BMDM	LPS	egér	GSM1022318	GSM1022319	(Ostuni et al., 2013)
RAR	F9	ATRA	egér	GSM1370730	GSM1370736	(Chatagnon et al., 2015)
VDR	Intestinal	D3	egér	GSM1694859	GSM1694862	(Lee et al., 2015)

1. táblázat. A vizsgált ligand-indukálható transzkripciósfaktorok listája. A táblázat azokról a ChIP-seq mintákról tartalmaz információkat, melyek segítségével az anya és leány enhanszerek általános dinamikája azonosítva lett. Az oszlopok tartalmazzák a vizsgált transzkripciósfaktorok neveit, a sejtet/sejtvonalat, amellyel a kísérletet végezték, a ligandkezelés típusát, a fajt, a minták GEO azonosítóit és a mintákhoz tartozó referenciákat.

(Transzkripció) faktor	Azonosító (oldószeres)	Azonosító (E2 kezelt)	Azonosító (kezeletlen)	Azonosító (Tam/Fulv)	Referencia
MED1	GSM1469999	GSM1470000	-	-	(Z. Liu et al., 2014)
P300	GSM1470013	GSM1470014	-	-	
DNáz I	GSM822389	GSM822390	-	-	(He et al., 2012)
H3K27ac	GSM1382472	GSM1382482	-	-	(Brunelle et al., 2015)
BRD4	GSM1348516	GSM1348519	-	-	(Z. Liu et al., 2014)
ERα	GSM589236	-	-	-	(Joseph et al., 2010)
ERα	GSM1534720	-	-	-	(Franco, Nagari, & Kraus, 2015)
ERα	GSM1534721	-	-	-	
ERα	-	-	GSM1858620	-	(Swinstead et al., 2016)
ERα	-	-	GSM1858621	-	
ERα (0 min)	-	-	GSM1325246	-	(Guertin, Zhang, Coonrod, & Hager, 2014)
ERα (2 min)	-	GSM1325247	-	-	
ERα (5 min)	-	GSM1325248	-	-	
ERα (10 min)	-	GSM1325249	-	-	
ERα (40 min)	-	GSM1325250	-	-	
ERα (160 min)	-	GSM1325251	-	-	
ERα (siCTL) (FoxA1 géncsendesített)	GSM631465	-	-	-	(Hurtado et al., 2011)
ERα (siFoxA1) (FoxA1 géncsendesített)	-	GSM631467	-	--	
ERα (antagonista kísérlet)	GSM365925	-	-	-	(Welboren et al., 2009)
ERα (antagonista kísérlet)	-	GSM365926	-	-	
ERα (tamoxifen kezelt; antagonista kísérlet)	-	-	-	GSM365927	
ERα (fulvestrant kezelt; antagonista kísérlet)	-	-	-	GSM365928	

2. táblázat. A szuper-enhanszerek karakterizálásához használt transzkripció faktorok listája. A táblázat azokról a humán MCF-7 sejtvonalban végzett ChIP-seq mintákról tartalmaz információkat, melyek segítségével az ER α -vezérelt szuper-enhanszereket jellemeztük. Az oszlopok tartalmazzák a vizsgált (transzkripció) faktorok neveit, a minták GEO azonosítóit – aszerint, hogy milyen kezelést kaptak –, és a mintákhoz tartozó referenciákat.

3.1.2. Szekvenálási adatok a különböző típusú ER α kötőhelyek karakterizálásához

Az ösztradiollal kezelt ER α cisztróm ~80000 transzkripciós faktor kötőhelyének jellemzéséhez és csoportosításához a Gene Expression Omnibus (GEO) és a Sequence Read Archive (SRA) adatbázisokból letöltött és feldolgozott, a humán MCF-7 sejtvonalban végzett ChIP-, DNáz- és RNA-seq adatok listáját és azok legfontosabb jellemzőit a **3. és a 4. táblázat** tartalmazza.

Kísérlet típusa	(Transzkripciós) faktor	Azonosító (oldószeres)	Azonosító (E2 kezelt)	Referencia
ChIP-seq	ER α	GSM614611	GSM614610	(Schmidt et al., 2010)
ChIP-seq	FoxA1	GSM588929	GSM588930	(Tan et al., 2011)
ChIP-seq	AP2 γ	GSM1469997	GSM1469998	(Z. Liu et al., 2014)
ChIP-seq	GATA3	GSM1469992	GSM1469993	
ChIP-seq	ER α siCTL	-	GSM631465	(Hurtado et al., 2011)
ChIP-seq	ER α siFoxA1	-	GSM631467	
ChIP-seq	H3K27ac	GSM1382472	GSM1382482	(Brunelle et al., 2015)
ChIP-seq	P300	GSM1470013	GSM1470014	(Z. Liu et al., 2014)
DNáz-seq	DNáz I	GSM822389	GSM822390	(He et al., 2012)
ChIP-seq	HDAC2	GSM1010825	-	(Gertz et al., 2013)
ChIP-seq	SIN3	GSM1010862	-	

3. táblázat. A különböző típusú ER α kötőhelyek karakterizálásához használt adatok listája. A táblázat azokról az MCF-7 sejtvonalban végzett ChIP- és DNáz-seq mintákról tartalmaz információkat, melyek segítségével a ~80000 ER α transzkripciós faktor kötőhelyek négy, általunk meghatározott csoportját jellemeztük. Az oszlopok tartalmazzák a kísérletek típusát, a vizsgált (transzkripciós) faktorok neveit, az oldószerrel és E2-vel kezelt minták GEO azonosítóit és a mintákhoz tartozó referenciákat.

Kísérlet típusa	E2 kezelés tartama (min)	Azonosító	Referencia
RNA-seq	kezeletlen	GSM1533414	(Honkela et al., 2015)
RNA-seq	5 min	GSM1533415	
RNA-seq	10 min	GSM1533416	
RNA-seq	20 min	GSM1533417	
RNA-seq	40 min	GSM1533418	
RNA-seq	80 min	GSM1533419	
RNA-seq	160 min	GSM1533420	
RNA-seq	320 min	GSM1533421	
RNA-seq	640 min	GSM1533422	
RNA-seq	1280 min	GSM1533423	

4. táblázat. A génexpressziós profilok meghatározásához használt RNA-seq adatok listája. A táblázat azokról az MCF-7 sejtvonalban végzett RNA-seq mintákról tartalmaz információkat, melyek segítségével jellemeztük a különböző típusú ER α kötőhelyek által szabályozott gének expressziós profilját. Az oszlopok tartalmazzák a kísérletek típusát, az ösztradiollal történő kezelés idejét, valamint a minták GEO azonosítóit és a mintákhoz tartozó referenciát.

3.2. Adatfeldolgozás

3.2.1. ChIP-seq adatok elemzése

A nyers szekvenciaadatok a GEO és SRA adatbázisokból való letöltés után a következő módon lettek elemezve: a szekvencialeolvasások (ún. *read*-ek) illesztése a humán hg19 vagy az egér (mm9) referenciagenomhoz a BWA (Burrows-Wheeler Aligner) szoftverrel (v7.10), a BAM (Binary Alignment Map) fájlok generálása pedig a SAMtools-szal (v0.1.19) történt (H. Li et al., 2009; H. Li & Durbin, 2009). A transzkripció faktor kötőhelyek meghatározásához a MACS2 (Model-based Analysis of ChIP-Seq 2) szoftvert (v2.0.10) használtuk, a műtermékeket pedig az ENCODE (Encyclopedia of DNA Elements) „feketelistás” genomi régióinak gyűjteménye

alapján szűrtük ki a BEDtools parancsaival (Dunham et al., 2012; Quinlan & Hall, 2010; Zhang et al., 2008). A HOMER (Hypergeometric Optimization of Motif EnRichment) (v4.2) *makeTagDirectory* programjának segítségével – a BAM fájlokat alapul véve – a HOMER számára is megfelelő szűrt bemenetet, ún. könyvtárat (*tag directory*) készítettünk a leolvasásokból. A fragmenthosszt egységesen 150 nukleotidra állítottuk be. A bedgraph-ok, melyek a genom szekvencia leolvasások általi lefedettségét reprezentálják, a fent említett HOMER könyvtárakból készültek a *makeUCSCfile* parancs segítségével (Heinz et al., 2010). A fent leírtak érvényesek a DNáz-seq adat elemzésére is, ugyanakkor a szélesebb H3K27ac régiók azonosításához a *--broad* kapcsolót használtuk MACS2-höz.

3.2.1.1. Lefedettség érték (RPKM) meghatározása

Az RPKM (Reads Per Kilobase per Million mapped reads) az a lefedettség érték, amely megadja a leolvasások számát 1000 bázisonként, egy millió térképezett leolvasásra vetítve, azaz azt mutatja meg, hogy arányosan hány leolvasás lenne az egységesen 1 kb hosszú régiókon, ha összesen 1 millió térképezett leolvasással dolgoznánk:

$$RPKM = R(p) \times \frac{10^3}{L(p)} \times \frac{10^6}{TR} ,$$

ahol $R(p)$ a vizsgált kötőhelyre (peakre) eső leolvasások (read-ek) számával, $L(p)$ a vizsgált kötőhely hosszával, TR az összes térképezett leolvasás számával egyenlő. Az RPKM értékek a BAM fájlból, a transzkripció faktor kötőhelyek középpontjától számított ± 50 bp-nyi régióon lettek számolva a BEDtools (v2.23.0) *coverageBed* parancsával (Quinlan & Hall, 2010).

3.2.1.2. Szuper-enhanszerek azonosítása

A szuper-enhanszerek minden esetben a ligand kezelt ChIP-seq mintákból lettek meghatározva a HOMER *findPeaks* parancsával, a *-style super* kapcsolót használva. A szuper-enhanszerek egyedi kötőhelyeinek középső 50 bp-nyi régióján RPKM értéket számoltunk az oldószerrel kezelt (kontrollnak tekinthető) mintából, majd szuper-enhanszerenként a legnagyobb intenzitású kötőhelyet *anya enhanszernek* neveztük el. A – többnyire – csak ligandkezelés hatására, másodlagosan elfoglalt kötőhelyek a *leány enhanszer* nevet kapták (Bojcsuk, Nagy, & Balint, 2017).

3.2.1.3. De novo motívumkeresés

A *de novo* motívumfeldúsulásokat a transzkripciós faktor kötőhelyek középpontjától számított ± 50 bp-nyi régióon vizsgáltuk, melyhez a HOMER *findMotifsGenome* programját használtuk (Heinz et al., 2010). A célzottan keresett motívumok 10, 12, 14 és 16 bp hosszúságúak voltak. Az elemzés során kapott *P*-értékek a vizsgált régiókon belüli feldúsulás és a HOMER által véletlenszerűen generált háttérsekvenciákon belüli feldúsulás összehasonlítása révén lettek meghatározva. A HOMER által alkalmazott statisztikai módszert a 3.2.1.7. fejezet ismerteti.

A motívumok eloszlását reprezentáló hőtérképek létrehozásához a feldúsult motívumok mátrixai lettek visszatérképezve a HOMER *annotatePeaks* parancsával, melyhez az *-mbed* kapcsolót használtuk. A térképezés a transzkripciós faktor kötőhelyek középpontjától számított 2 kb-nyi régióon történt, 20 bp-nyi blokkokban (ún. *bin*-ekben).

3.2.1.4. Hőterkép készítése

A leolvasásokat megjelenítő hőterképek (ún. *read distribution heat map*) elkészítéséhez a HOMER *tag* könyvtárait és *annotatePeaks* parancsát használtuk. Ehhez továbbá a *-ghist*, *-size 2000*, *-hist 50* kapcsolók segítségével definiáltuk azt, hogy a lefedettség értékeket a HOMER a vizsgált régió középpontjától számított ± 1000 bp-nyi régióban, 50 bp-nyi blokkokban számolja.

3.2.1.5. Hisztogram készítése

Az átlagos lefedettség értékeket reprezentáló hisztogramok a HOMER *tag* könyvtáraiból lettek számolva, szintén a HOMER *annotatePeaks* parancsával. A lefedettség értékek a vizsgált régiók középpontjától számított ± 1000 bp-nyi régiókon lettek számolva, melyből a középső ± 250 bp lett ábrázolva, 1 bp felbontással.

3.2.1.6. Dobozdiagram készítése

A dobozdiagramokkal (ún. box plotokkal) a kötőhelyek középpontjától számított ± 50 bp-nyi régióban számolt RPKM értékeket vizualizáltuk. A box-ok az első- és harmadik kvartiliseket reprezentálják, a vízszintes vonal a medián RPKM értéket jelöli, a szélsőértékeket pedig, melyek nem tartalmazzák az alsó és felső decilist, a dobozból kiinduló vonalak jelölik. Az alkalmazott statisztikai módszereket a 3.2.1.7. fejezet ismerteti.

3.2.1.7. Statisztikai módszerek

A 3.2.1.6. fejezethez kapcsolódóan: a dobozdiagramok vizualizálása során alkalmazott statisztika párosított (egymintás) t-próba, vagy párosítatlan (kétmintás) t-próba volt, annak megfelelően, hogy az összehasonlított csoportok egyazon régióban számolt értékekből származtak-e vagy

egymástól független, nem egyenlő elemszámú populáció átlagát tesztelték. Szignifikancia szintek: ns: nem szignifikáns; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

A 3.2.1.3. fejezethez kapcsolódóan: a szekvencia motívumok *de novo* feldúsulását a HOMER program az alapértelmezés szerint binomiális valószínűség-eloszlás révén határozza meg. Ezen statisztikai módszer segítségével feltételezi, hogy a vizsgálni kívánt DNS szekvenciák (mind az ismeretlen, mind a háttérként szolgáló) osztályozása független a bennük megtalálható motívumok előfordulásától. A motívumok háttérhez viszonyított feldúsulásának meghatározása során ez az eljárás figyelembe veszi az ismeretlen és a háttérszekvenciák elemszámát, továbbá azt, hogy az egyes típusok közül hány tartalmaz visszatérő motívumot. Ezekből az értékekből meghatározható annak a valószínűsége, hogy egy adott esemény (jelen esetben a motívum megléte) hányszor következett be az ismeretlen, ún. targetált szekvenciákban a háttérhez viszonyítva, feltételezve azt, hogy a target szekvencia és a motívum között nincs kapcsolat. Végeredményében a motívumok az így generált P-értékeik alapján lesznek rangsorolva.

3.2.1.8. Adatábrázolás

A lefedettségértékeket és a motívumok eloszlását reprezentáló hőtérképeket a Java TreeView (v1.1.6r4) program segítségével ábrázoltuk (Saldanha, 2004). A dobozdiagramok (ún. *box plotok*), a hisztogramok és vonaldiagramok a GraphPad Prism 6-os verziójával lettek ábrázolva. A Venn diagram a BioVenn webes applikáció segítségével készült (Hulsen, de Vlieg, & Alkema, 2008). A bedgraph fájlok az Integrative Genomics Viewer-ben (IGV) lettek megjelenítve (Thorvaldsdottir, Robinson, & Mesirov, 2013).

3.2.2. RNS szekvenálási adatok elemzése

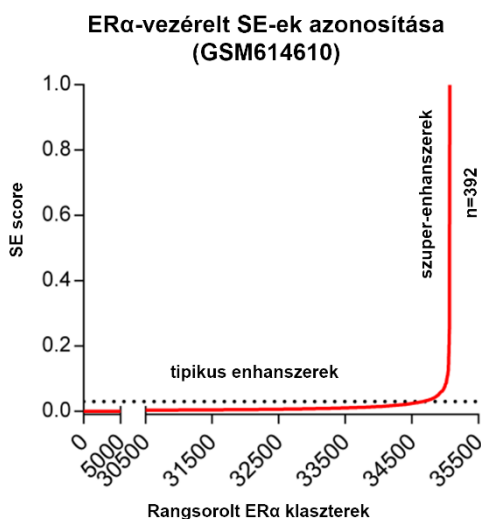
A nyers szekvenciaadatok a GEO adatbázisból lettek letöltve, majd a leolvasások a humán hg19 referencia genomhoz lettek illesztve a TopHat program (v2.0.7) segítségével. Az FPKM (Fragments Per Kilobase of transcripts per Million mapped reads) értékek, melyek megadják a fragmentek számát 1000 bázisonként, egy millió térképezett leolvasásra vetítve, a Cufflinks (v2.0.2) program alapértelmezett paramétereivel lettek számolva (Trapnell et al., 2010). A transzkripció faktor kötőhelyek a legközelebbi fehérjekódoló génekhez lettek hozzárendelve (azaz annotálva) a PeakAnnotator program segítségével (Salmon-Divon, Dvinge, Tammoja, & Bertone, 2010).

4. Eredmények

4.1. Az ösztrogén receptor alfa-vezérelt szuper-enhanszerek azonosítása

A szuper-enhanszerek olyan sejtípus-függő génexpressziós folyamatok szabályozásában vesznek részt, melyek nagyban hozzájárulnak a sejtazonosság kialakulásához (Hnisz et al., 2013; Whyte et al., 2013). A transzkripcióban betöltött kimagasló szerepük miatt egy genom-szintű vizsgálat során azonosításuk és az általuk szabályozott gének megismerése indokolt, különösképp betegségekben, vagy valamilyen megváltozott állapotban.

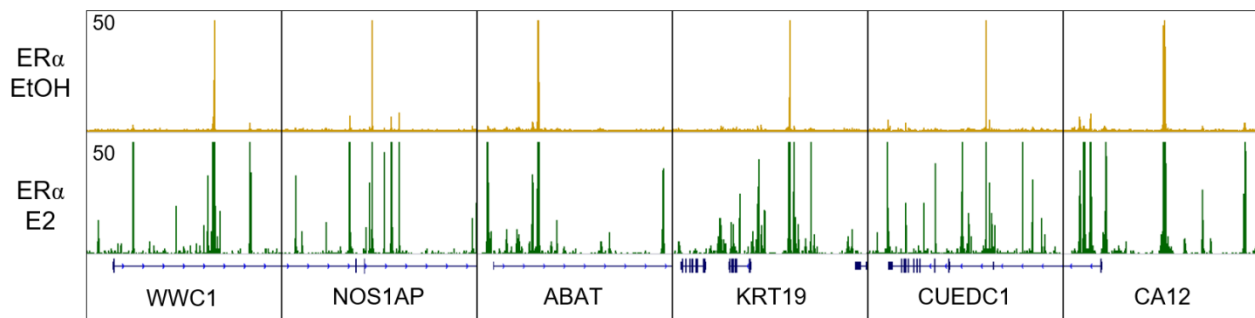
Az ösztrogén receptor alfa (ER α) transzkripciós faktor általi génszabályozás vizsgálata során –mely munkacsoportunk elsődleges fókuszában állt az elmúlt évek során – az emlő eredetű MCF-7 sejtvonalban összesen 392 szuper-enhanszer régiót azonosítottunk, melyek 4.042 egyedi enhanszert foglaltak magukba (**15. ábra**). Az azonosításhoz az ösztradiollal (E2) indukált ER α transzkripciós faktor kötőhelyeit vettük alapul, melyek a GSM614610 GEO azonosítójú ChIP-seq mintából lettek prediktálva (Schmidt et al., 2010).



15. ábra. Az ER α -vezérelt SE-ek azonosítása. A GSM614610 E2 kezelt ER α ChIP-seq minta felhasználásával a SE-ek a HOMER programmal lettek azonosítva (a 3.2.1.2. pontban leírtak szerint). Pontozott vonal metszi azt a pontját a görbének, ahol a meredekség értéke 1. A denzitásuk alapján sorbarendezett ER α kötőhelyek klaszterei közül 392 esett 1-es meredekség fölé, és ezek, mint SE-ek képezték alapját a további elemzéseinknek.

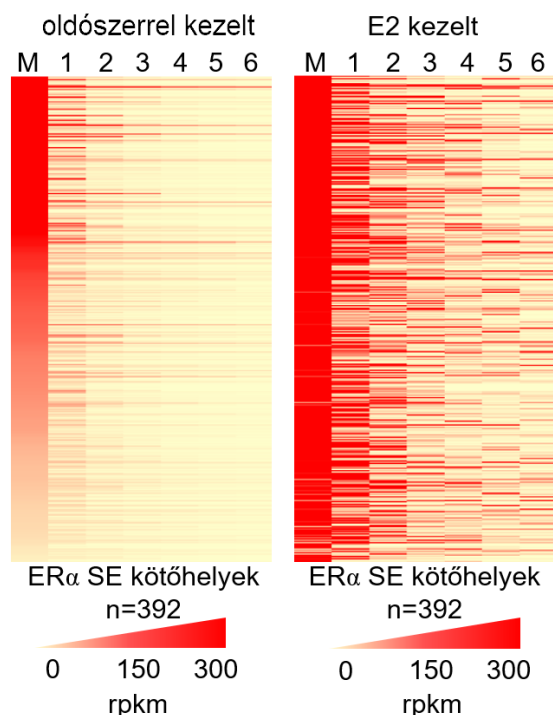
4.2. Anya és leány enhanszerek azonosítása

A szuper-enhanszer régiókat az IGV genomböngészőben megtekintve arra lettünk figyelmesek, hogy az oldószerrel (etanollal; EtOH) kezelt mintában is megfigyelhető egy-egy elfoglalt ER α kötőhely minden egyes szuper-enhanszer esetében. Feltételezve, hogy az elsőként megjelenő kötőhelyek alapját képezhetik a későbbiekben a ligand hatására megjelenő szuper-enhanszereknek, a szuper-enhanszerek ezen kötőhelyeit „anya” enhanszereknek (n=392; minden szuper-enhanszer esetében a legnagyobb intenzitású csúcs) neveztük el, míg a másodlagos, csak ligandkezelést követően aktivált kötőhelyek a „leány” enhanszer (n=3650) nevet kapták. A továbbiakban eszerint hivatkozunk rájuk. Néhány példát a **16. ábra** reprezentál.



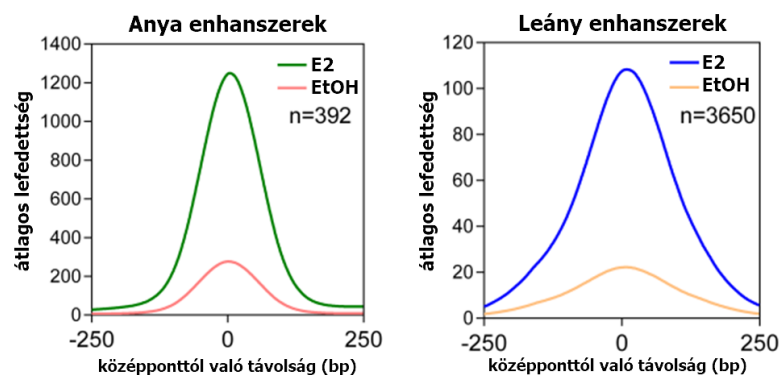
16. ábra. Az ER α -vezérelt SE-eket egy, vagy néhány ER α kötőhely használata jellemezi ligandkezelés hiányában is. Az Integrative Genomics Viewer (IGV) genomböngészőben az oldószerrel (GSM614611) és E2-vel kezelt (GSM614610) ChIP-seq mintákat megjelenítve jól látható az anya enhanszerek jelenléte. A genomböngészőben a skála mindkét minta esetében 50.

Az anya és leány enhanszerek 392 szuper-enhanszeren belüli egyidejű ábrázolásához először sorbarendeztük az anya enhanszereket az oldószerrel kezelt ChIP-seq mintából számolt RPKM értékeik alapján, és ezeket az értékeket hőtérképen ábráztuk. Ezt követően a további 6 legnagyobb lefedettséggel rendelkező leány enhanszert egy sorban jelenítettük meg a hozzá tartozó anya enhanszerrel (**17. ábra**).



17. ábra. A 392 ERα-vezérelt SE anya és leány enhanszerei. A hő térkép a 392 ERα SE anya (M, mint „mother”) és a top 6 (1-6) leány enhanszereinek lefedettségértékét (RPKM) reprezentálja az oldószerrel és az E2-vel kezelt ERα ChIP-seq mintából számolva. Minden sor egy SE-t jelenít meg; az első oszlop minden esetben az anya enhanszer, ugyanazon sorban a további oszlopok pedig a top 6 leány enhanszert jelölik. A két hő térképen mind a sorok, mind az oszlopok sorrendjei megegyeznek.

Látható, hogy az oldószerrel kezelt mintában az anya enhanszerek nagyon jól elkülönülnek. A leány enhanszerek csak nagyon kevés hányada van jelen és azok is alacsony lefedettséget mutatnak. Kezelés hatására viszont nemcsak a leány enhanszerek jelennek meg, de az anya enhanszerek is tovább indukálódnak (**17. ábra**). Ennek az indukálódásnak a mértéke a **18. ábrán** látható.

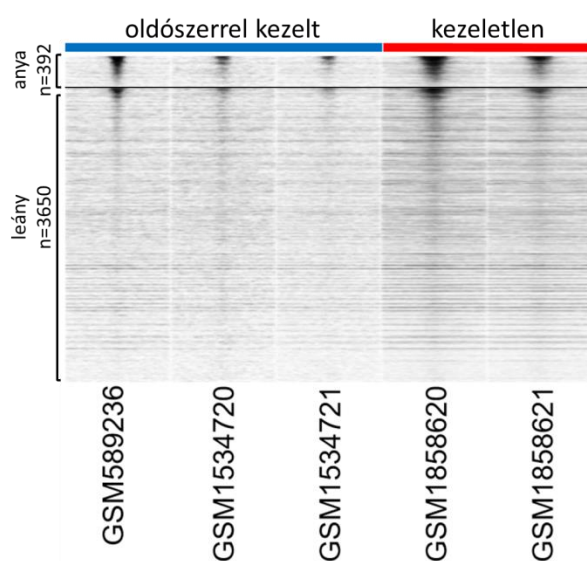


18. ábra. Az anya és leány enhanszerek átlagos lefedettsége E2 kezelés előtt és után. A hisztogram a 392 anya és 3650 leány enhanszer átlagos lefedettségét (RPKM) ábrázolja E2 kezelés hiányában (EtOH) és E2 kezelést követően (E2). A lefedettség

értékek a kötőhelyek középpontjától mért ± 250 bázispárnyi (bp) távolságon, 1 bp felbontással lettek számolva és ábrázolva.

Az E2 kezelés hatására az anya enhanszerek átlagos lefedettsége ~300-ról ~1200-ra nőtt, míg a leány enhanszereké E2 hatására is csupán átlagosan ~110-re tehető. Ez a tízszeres különbség a két enhanszercsoport kötési gyakoriságában bár valóban arra utal, hogy az anya enhanszerek kivételesek lehetnek, felmerül a kérdés, okozhatja-e a DNS-hez való „idő előtti” kötődésüket az etanol, mint alkalmazott oldószer.

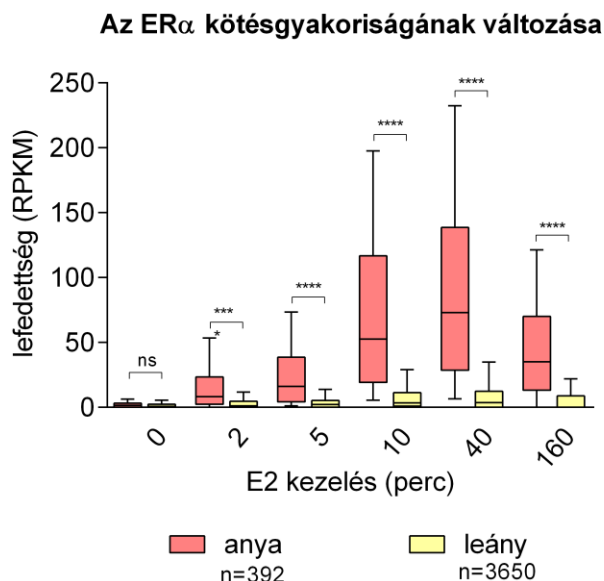
Ennek megválaszolására megvizsgáltuk az adatbázisokban fellelhető MCF-7 sejtvonalban végzett és oldószerrel kezelt (etanollal vagy dimetil-szulfoxiddal, azaz DMSO-val), illetve kezeletlen ER α ChIP-seq mintákat. Azt találtuk, hogy az anya enhanszerek jelenléte mind az oldószerrel kezelt, mind a kezeletlen minták esetében egyértelműen detektálható volt, szemben a leány enhanszerekkel. Mindez összhangban van az eddigi eredményekkel, és arra utal, hogy az oldószer jelenléte egyik enhanszercsoport kötését sem változtatja meg **(19. ábra)**.



19. ábra. Az anya enhanszerek meglétének vizsgálata oldószerrel kezelt és kezeletlen mintákban. A szekvencia leolvasásokat ábrázoló hőtérkép 3 oldószerrel kezelt és 2 kezeletlen mintán reprezentálja a 392 anya és 3.650 leány enhanszer kötését. Az ábrán minden sor egy ER α kötőhelynek felel meg. Minden oszlop az E2-kezelt ER α minta kötőhelyeinek középpontjához viszonyított ± 1 kilobázispárnyi (kb) régiót mutatja. A minták nevei a GEO adatbázisban található azonosítóikkal egyeznek meg.

Egy, az időfüggést vizsgáló, ún. *time-course* kísérlet, amely az ER α kötődésének gyakoriságát reprezentálja eltérő időpontokban történő E2 kezelést követően, szintén arra utal,

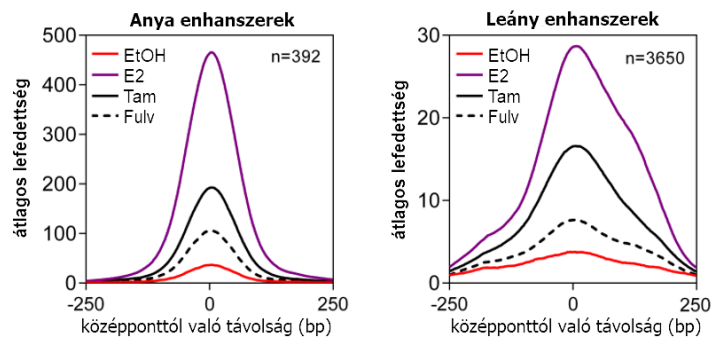
hogy az ER α az anya enhanszereket a leány enhanszerek előtt foglalja el, mindemellett kötődése a későbbi időpontokban, pl. 160 perc elteltével is detektálható (20. ábra).



20. ábra. Az anya és leány enhanszerek kötődésének dinamikája az idő függvényében. Az ábra a 392 anya és a 3650 leány enhanszer kötődését reprezentálja a 0. percben, továbbá az E2 kezelést követő 2, 5, 10, 40 és 160. percben. A boxok az anya- és leány enhanszerek középső 200 bp-nyi régiójában számolt lefedettségértéket (RPKM) reprezentálják. Párosítatlan (kétmintás) t-próba, ahol ns = nem szignifikáns; **** P < 0.0001.

Bár a 0. percben végzett kísérlet ellentmond a 19. ábrán bemutatott eredménynek – mely szerint az anya enhanszerek a kezeletlen mintákban is detektálhatóak –, a szóban forgó minta esetében az immunprecipitáció hatékonysága a többi mintához képest alacsony, 1,17 volt. A minta gyengébb minősége miatt nem összehasonlítható a kísérlet többi tagjával, melyek esetében az immunprecipitáció átlagos hatékonysága 4,71-re tehető.

Meglepő módon a parciális agonistaként működő tamoxifennel és a tiszta ER α agonistaként működő fulvestranttal történő kezelések bár lecsökkentették az ER α kötődésének mértékét, az anya enhanszereken detektálható ER α mennyisége továbbra is jelentős maradt (21. ábra).



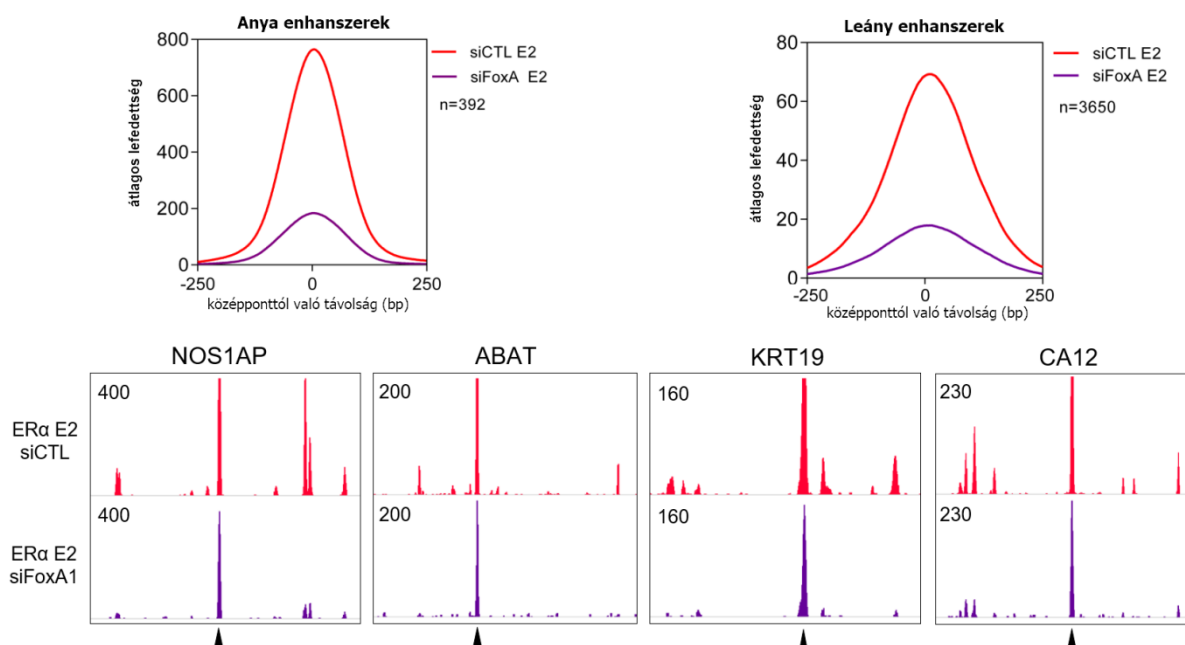
a kötőhelyek középpontjától mért ± 250 bázispárnyi (bp) távolságon, 1 bp felbontással lettek számolva és ábrázolva.

21. ábra. Antagonisták hatása az ER α

kötésére. A hisztogramok az ER α kötés átlagos lefedettségét reprezentálják az anya és leány enhanszereken oldószeres (EtOH; etanol), E2, tamoxifen (Tam) és fulvestrant (Fulv) kezelést követően. A lefedettség értékek

Az antagonista hatású készítményeken felül megvizsgáltuk azt is, hogy a FoxA1 siRNS-mediált géncsendesítése hogyan befolyásolja az ER α kötődését az anya és leány enhanszereken. A FoxA1-ről bár ismert, hogy jelenléte, legyen az közvetlen vagy közvetett, az emlőrákos sejtekben szinte kivétel nélkül minden ER α kötéshez szükséges; hiánya – a tamoxifennel és fulvestranttal történő kezeléshez hasonlóan – bár globálisan csökkentette a kötés mértékét, azt szintén nem volt képes teljes mértékben meggátolni (22. ábra) (Hurtado et al., 2011).

A fenti megfigyeléseink alapján elmondható, hogy minden ER α -vezérelt szuper-enhanszerre jellemző egy rendkívül aktív, erős kötődést biztosító régió, amely ún. előfutárként foglalja el a DNS-t és feltehetően szerepet játszik a szuper-enhanszerek ligandkezelést követő kialakulásában. Az általunk anya enhanszerként elnevezett elsődleges enhanszerek nem csupán megelőzik társaikat, a leány enhanszereket, de időben is egy sokkal stabilabb fehérje-DNS interakció jellemzi őket.

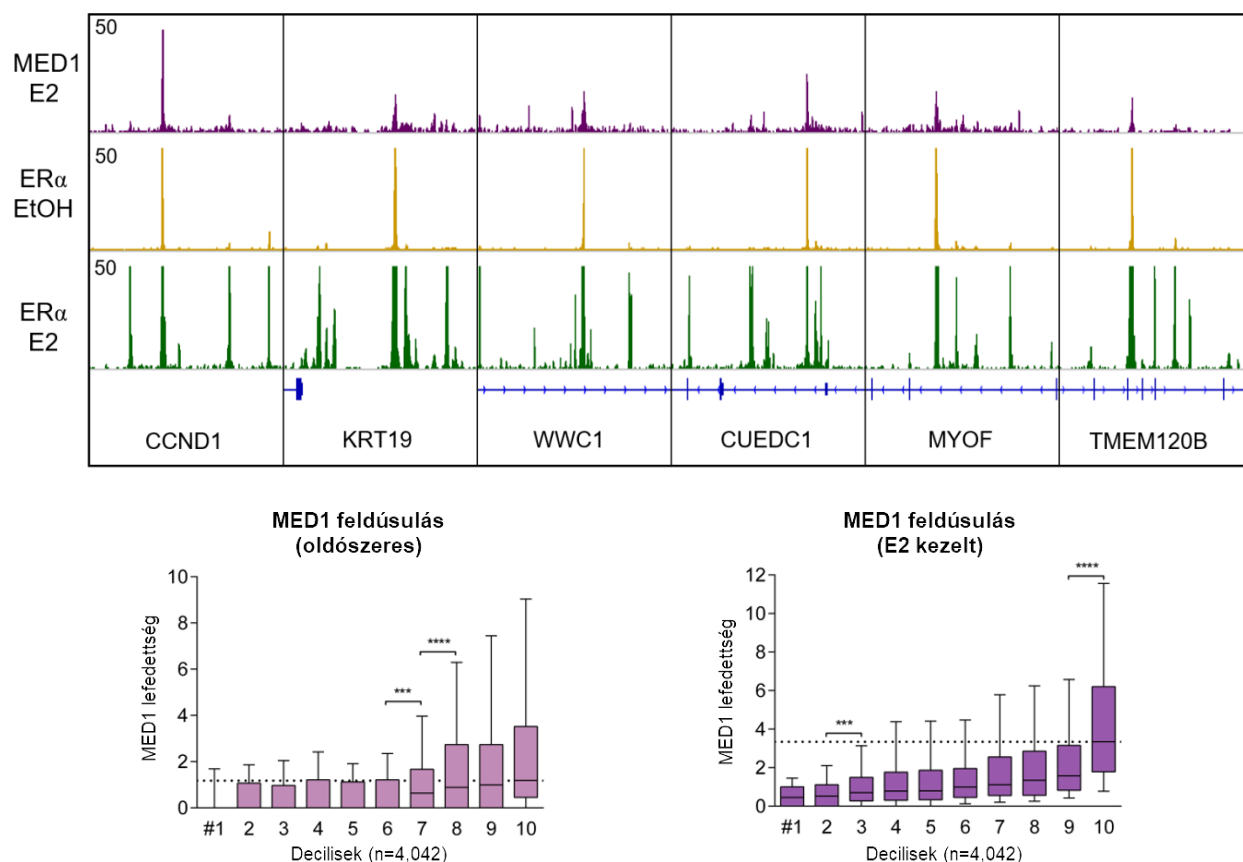


22. ábra. A FoxA1 hiányának a hatása az ER α kötésére. A hisztogramok az ER α kötés átlagos lefedettségét reprezentálják a NOS1AP, ABAT, KRT19 és CA12 géneket szabályozó anya és leány enhanszereken, kontroll (siCTL) és FoxA1 siRNS-mediált géncsökkentett (siFoxA1) MCF-7 sejtekben, E2 kezelést követően. A lefedettség értékek a kötőhelyek középpontjától mért ± 250 bázispárnyi (bp) távolságon, 1 bp felbontással lettek számolva és ábrázolva. Az alsó panel az IGV genomböngészőben megjelenített néhány példán keresztül mutatja be a ChIP-seq lefedettség változását a kontroll és az siRNS-mediált FoxA1 géncsökkentést követően. A skála minden esetben a szuper-enhanszeren belüli legnagyobb csúcs méretéhez, az anya-enhanszerhez lett beállítva, melyeket fekete nyilak jelölnek.

4.2.1. A kromatin transzkripció aktivitásának vizsgálata

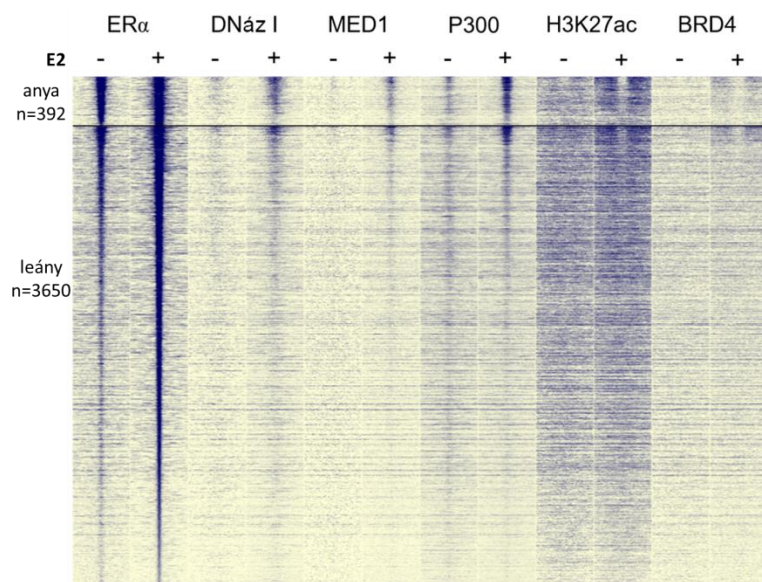
Ahogy a bevezetőben is szó esett róla, a szuper-enhanszerek szokatlanul magas ko-faktor, különösképp Mediátor kötődéssel társulnak. A Mediátor komplex több mint 30 fehérjét foglal magába, melyek közül az 1-es alegysége, a MED1 markáns jelenlétét azonosították a szuper-enhanszereken. A Mediátor szerepe abban rejlik, hogy hidat képez a(z) (szuper-)enhanszerek és a szabályozott gének promótere között (Hnisz et al., 2013; Lovén et al., 2013; Whyte et al., 2013).

Az ER α -vezérelt szuper-enhanszerek esetében azt találtuk, hogy a MED1 E2 kezelést követően ugyanazon genomi régióknál helyezkedik el, mint az anya-enhanszerek, továbbá kötőgyakorisága követi a vizsgált szuper-enhanszereket alkotó 4042 egyedi kötőhely ER α általi lefedettségét (23. ábra). A MED1 kizárólagos jelenléte az anya enhanszereknél tovább erősíti azok fontos szerepét a szuper-enhanszereken belül.



23. ábra. A MED1 jelenléte az ER α szuper-enhanszereken belül. Az IGV genomböngészőben (fent) megjelenített adatok az E2-kezelt MED1, és a kezeletlen és E2-kezelt ER α ChIP-seq minták lefedettségét reprezentálják a CCND1, KRT19, WWC1, CUEDC1, MYOF és TMEM120B géneket szabályozó szuper-enhanszer régiókon. A skála minden minta esetében 50. A box plot-ok (dobozdiagramok) az oldószerrel (bal) és E2-vel (jobb) kezelt MED1 lefedettségértékeket mutatják 10 decilisre osztva. A decilisekbe tartozó kötőhelyek, melyeken a MED1 denzitás lett számolva, az ER α lefedettség értékek növekvő sorrendje alapján követik egymást. Egymintás (párosított) t-próba, ahol * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

A MED1 mellett megvizsgáltuk a szuper-enhanszerek egyedi kötőhelyeinek a DNáz I hiperszenzitivitását, a P300 és BRD4 fehérjék és a H2K27ac hisztonmódosítás jelenlétét is. A DNáz I hiperszenzitivitás vizsgálatának fontossága abban rejlik, hogy az átmenetileg vagy tartósan nukleoszóma hiányos kromatin régiók fokozott érzékenységet mutatnak a DNáz I enzimmel történő emésztésre, jelezve ezáltal a transzkripciósan aktív helyeket (Zhong et al., 2016). Hőtérképen ábrázolva ezeknek a (ko-)faktoroknak a jelenlétét az anya és leány enhanszereken azt találtuk, hogy az anya enhanszerek a fehérjék számára sokkal nyitottabb kromatin régióban helyezkednek el, továbbá E2 kezelést követően jelentősen megnő a MED1, P300, H3K27ac és a BRD4 jelenléte ezeken a helyeken. Ezzel ellentétben a leány enhanszereken, bár látható némi növekedés a vizsgált fehérjék mennyiségében (a BRD4 kivételével), összességében sokkal kisebb aktivitás jellemzi őket **(24. ábra)**.



24. ábra. Az aktív enhanszereket és nyitott kromatin régiókat jelölő faktorok jelenléte. A hőtérkép az oldószerrel (-) és E2-vel (+) kezelt ERα, DNáz I, MED1, P300, H3K27ac és BRD4 faktorok lefedettségét mutatja az ERα kötőhelyek alapján definiált anya (n=392) és leány (n=3650) enhanszerek középpontjához viszonyított ± 1 kbp-nyi régióban.

4.2.2. Specifikus szekvenciamotívumok azonosítása

Az eddigi eredményeink alapján azt feltételeztük, hogy az anya és leány enhanszerek aktivitásában megfigyelhető különbségekért más ko-faktorok lehetnek a felelősek, melyek eltérő affinitással járulhatnak hozzá az ER α DNS-hez történő kihorgonyozásához. Ennek megválaszolására megvizsgáltuk a *de novo* motívumfeldúsulásokat a két enhanszer csoport kötőhelyein (25. ábra).

Anya enhanszerek					Leány enhanszerek				
	P-érték	Target %	Bg %	Motívum		P-érték	Target %	Bg %	Motívum
	1e-200	62.2 %	3.37 %	ERE		1e-236	59.6 %	29.94 %	NR félhely
	1e-61	42.99 %	8.65 %	NR félhely		1e-44	8.96 %	3.31 %	FoxA1
	1e-43	29.57 %	5.37 %	DR -1		1e-24	7.64 %	3.52 %	AP-1
	1e-20	20.12 %	5.27 %	MARE					

25. ábra. Motívumfeldúsulások. Az anya-és leány enhanszerek kötőhelyein a háttérhez képest statisztikailag feldúsult motívumok. (Target %: megadja, hogy a vizsgált régiók (392 anya és 3650 leány) hány százalékában található meg az adott motívum; Bg%: megadja, hogy a véletlenszerűen generált, háttérként szolgáló szekvenciákban az adott motívum hány százalékban van jelen).

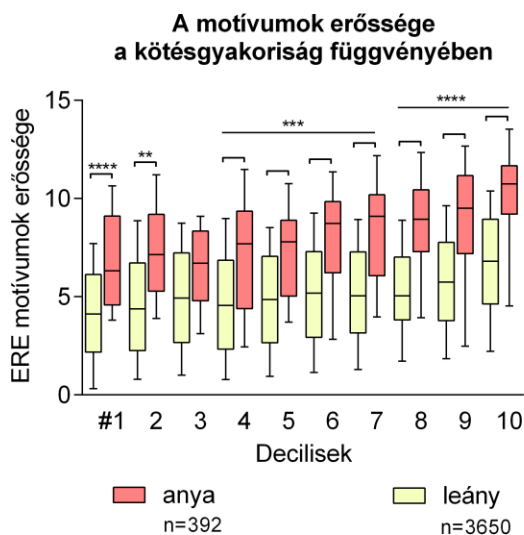
A 392 anya enhanszer 62,2%-át az erős, nagy affinitású, ún. kanonikus ERE motívum jelenléte jellemezte, amely a 3,37%-os háttérfeldúsuláshoz képest egy nagyon markáns feldúsulást jelent. Emellett 42,99%-ban azonosítható volt az AGGTCA magreceptor félhely (NR félhely), továbbá 29,57%-ban egy olyan direkt módon ismétlődő (direct repeat; DR), valószínűleg szintén a magreceptorokra jellemző motívum, amelyben az első félhely utolsó nukleotidja képezte a második félhely első nukleotidját. Korábban ehhez hasonló motívumfeldúsulást az ER α vonatkozásában nem publikáltak. A MARE motívum, melyet egy

MAF/BACH vagy MAF/NFE2 dimer képes kötni, 20,12%-os feldúsulást mutatott, bár ebben a motívumban is fellelhetőek az NR félhely nyomai.

A 3650 leány enhanszerre a teljes ERE nem, csupán a magreceptor félhely 59,6%-os feldúsulása volt jellemző, viszont ugyanez a motívum a kontrollként szolgáló random háttérszekvenciák 29,94%-ában ugyancsak megtalálható volt. Emellett jóval kisebb százalékban, de megjelentek a FoxA1 (8,96%) és az AP-1 (7,64%) fehérjék motívumai is **(25. ábra)**.

A viszonylag kisszámú szekvencia (n=392) ellenére, ha figyelembe vesszük az erős *P*-értéket (1e-200) és az ERE feldúsulásának mértékét (62,2% vs. 3,37%), az eredmény arra utal, hogy a nagy mennyiségű ER α jelenléte az anya enhanszereken (ligand hiányában is) elsősorban ezeknek a kanonikus válaszadó elemeknek köszönhető. Ezzel szemben az ER α a leány enhanszereket többnyire fehérje-fehérje interakció révén, félhelyeken és a FoxA1 és AP-1 fehérjéken keresztül foglalja el helyét a DNS-en. Mindez nem zárja ki az ERE jelenlétét a leány enhanszereken, csupán azok statisztikai feldúsulása nem volt azonosítható.

A következő lépésben megvizsgáltuk, hogy a motívumok erőssége hogyan függ össze a fehérjekötés gyakoriságával **(26. ábra)**. Ezt úgy valósítottuk meg, hogy megkerestük a *de novo* feldúsult ERE motívumhoz legjobban hasonlító DNS szekvenciát minden egyes (4042) ER α kötőhely középpontjától számított 200 bp-nyi régió belül. A HOMER program hasonlóság alapján adott motívumerősség értékeit (*score*) így minden kötőhelyhez hozzá tudtuk rendelni **(26. ábra)**.



26. ábra. Az ERE motívumok és az ER α kötések gyakorisága közötti összefüggés. A boxok a leány enhanszerek esetében az E2-kezelt, míg az anya enhanszerek esetében az oldószerrel kezelt ER α lefedettség értékek növekvő sorrendjében, 10 csoportban (decilisben) reprezentálják az anya és leány enhanszerek középső 200 bp-nyi régiójában azonosított ERE motívumok erősségét. Egymintás (párosított) t-próba, ahol ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$.

Ez az elemzés tovább erősítette azt, hogy az anya enhanszerek minden esetben erősebb ERE-vel rendelkeznek (a 8-10. decilisben a különbség az anya és leány enhanszerek között már nagyon jelentős, $P < 0.0001$), továbbá az is kirajzolódott, hogy az ER α kötésének gyakorisága követi a motívum erősségét.

Összefoglalásként elmondható, hogy az elsődleges kötőhelyeket, melyeknek a DNS-sel való interakciója nem igényli a ligand általi stimulációt, szignifikánsan erősebb válaszadó elemek vezérlik és ezek feltehetően nem igénylik más transzkripció faktor(ok) jelenlétét sem.

Elsődleges megfigyeléseinket az ER α -n kívül 4 további egér és humán sejtvonalon, összesen 6 másik transzkripció faktor vonatkozásában validáltuk.

4.3. További ligand-indukálható szuper-enhanszerek vizsgálata

4.3.1. A vizsgálatba bevont transzkripciós faktorok

Megfigyelésünket, mely szerint az egyik legszélesebb körben vizsgált ligand-indukálható transzkripciós faktor, az ER α kötőhelyei által elfoglalt szuper-enhanszerek egy jól körülírható mechanizmus révén épülnek fel, 6 további ligand által indukálható transzkripciós faktor ChIP-seq adatait felhasználva próbáltuk meg alátámasztani. Ezen faktorok közé tartozott 3 másik magreceptor, név szerint az egér bél epitélisejtekben vizsgált D-vitamin receptor (VDR), az F9 egér embrionális karcinóma sejtvonalban vizsgált retinsav receptor (RAR), és a humán LNCaP prosztataráksejtben vizsgált androgén receptor (AR), továbbá a C57BL/6 egértörzs primer csontvelő eredetű makrofágjaiban vizsgált JUNB, és az MCF-7 sejtvonalban vizsgált, az ER α – talán – legfontosabb társfaktoraiként ismert FoxA1 és AP2 γ faktorok. A FoxA1 E2 általi indukálása a korábbi tanulmányok alapján, bár nem okoz szignifikáns változást a FoxA1 kötésében, bevonását számos további összehasonlítás indokolta (Hurtado et al., 2011). Emellett további egy ER α mintán is megismételtük a szuper-enhanszereket karakterizáló alapvető elemzéseket.

4.3.2. Anya és leány enhanszerek dinamikájának és motívumainak vizsgálata

Amikor a VDR, RAR, AR, JUNB, FoxA1, AP2 γ és ER α faktorok anya és a top 6 leány enhanszereit hőtérképeken egyidejűleg ábrázoltunk egymás mellett oldószeres/kezeletlen és ligand által indukált állapotban, a FoxA1 kivételével, melyről ismert, hogy az E2 általi indulálása nem okoz szignifikáns változást a FoxA1 kötésében, minden minta esetében ugyanazt a jelenséget tapasztaltuk, mint az eredendően vizsgált ER α -nál **(lásd 17. ábra)**: az anya enhanszerek elsődleges jelenléte a vizsgált faktorok mindegyikénél megfigyelhető volt már a

specifikus ligand kezelést megelőzően, ligand általi aktivációt követően a kötési gyakoriság pedig a leány enhanszereken is megnőtt **(27. ábra)**.

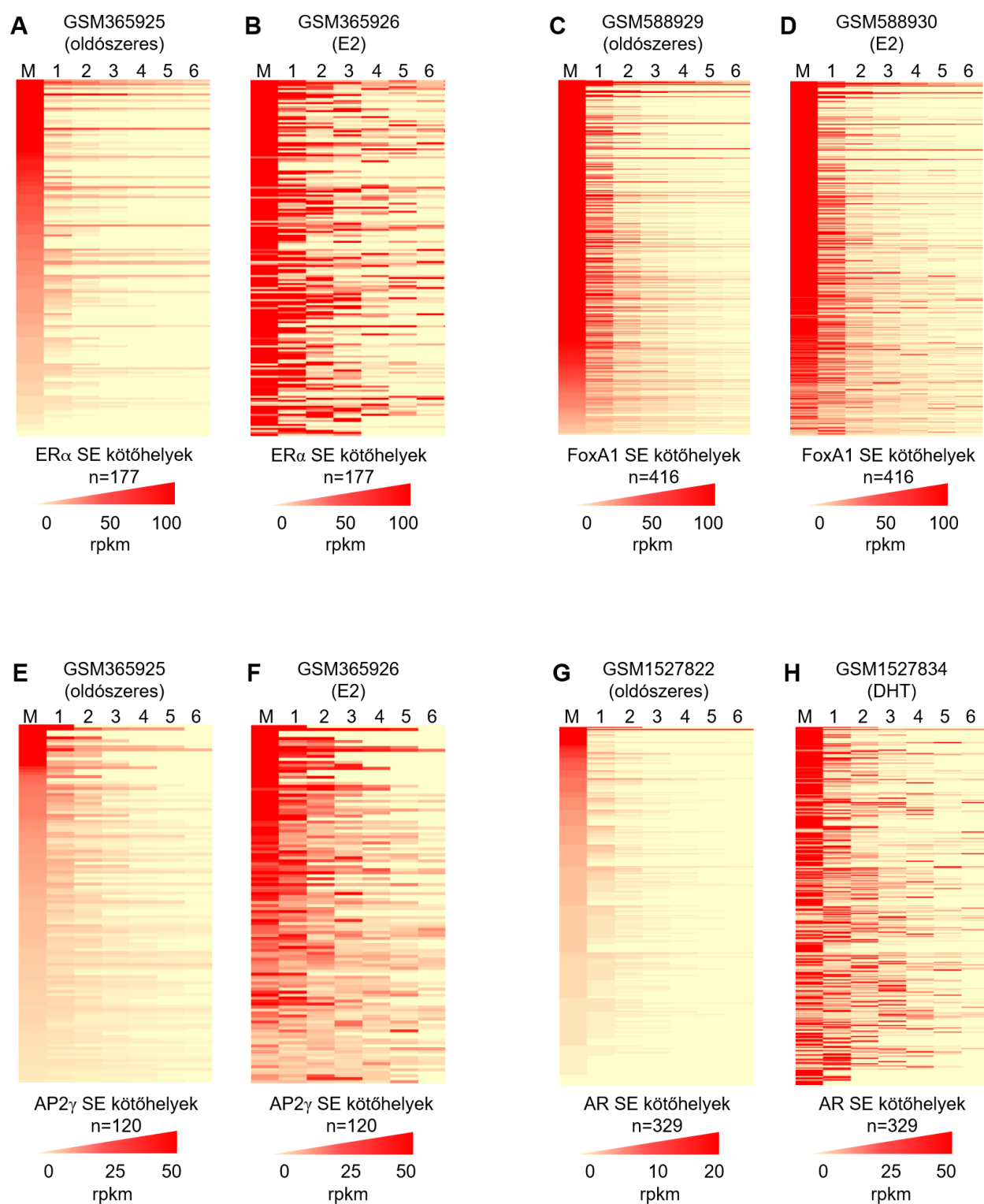
A hisztogramok, melyek a vizsgált faktorok anya- és leány enhanszereinek átlagos lefedettségét reprezentálják a kötőhelyek közppontjától számított ± 250 bp-nyi régióban, minden vizsgált transzkripciós faktor (kivéve a FoxA1) esetében ténylegesen számszerűsítették, hogy az anya-enhanszerek már ligand hiányában nagyobb mennyiségű fehérjét képesek toborozni, mint a leány enhanszerek a ligand hozzáadását követően **(28. ábra)**.

A hisztogramok, melyek az átlagos lefedettséget számszerűsítik, minden vizsgált transzkripciós faktor (kivéve FoxA1) esetében azt mutatták, hogy az anya-enhanszerek már ligand hiányában képesek nagyobb mennyiségű fehérjét toborozni, mint a leány enhanszerek a ligand hozzáadását követően **(28. ábra)**.

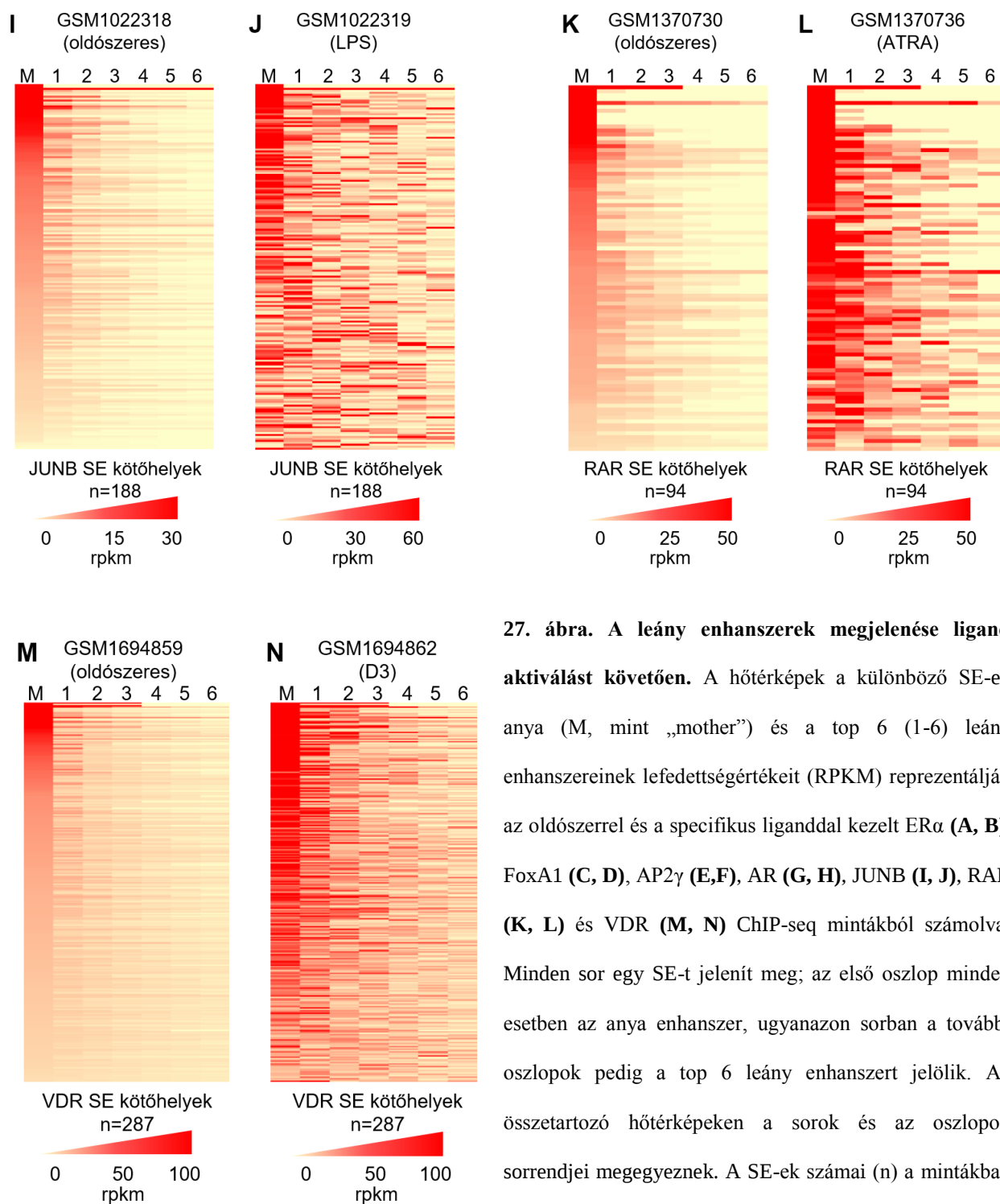
Ezen felül az anya enhanszerek motívumfeldúsulásai is lényegesen jobb, erősebb, specifikusabb motívumokat eredményeztek a másodlagosan megjelenő leány enhanszerekhez képest **(29. ábra)**. Az AP2 γ leány enhanszereit érdekes módon a fő ko-faktorainak, a FoxA1-nek, az ER α -nak és az AP-1-nek a motívuma jellemezte, emellett az inzulátor CTCF fehérje motívuma is megfigyelhető volt **(29. ábra, E-F)**. Az eredmények azt mutatták, hogy az AR esetében az anya enhanszerek létrehozásához a FoxA1 jelenléte is szükséges **(29. ábra, G)**. A leány enhanszerek – talán már nem meglepő módon – sokkal gyengébb androgén válaszadó elemet (ARE; androgen response element) és Fox motívumot eredményeztek, ugyanakkor azonosítottuk az említett két fehérje válaszadó elemeinek kompozit elemét (Fox::ARE) is, amely az AR és FoxA1 fehérjék szoros együttműködésére utal az LNCaP sejtvonalban, és utal egyúttal arra is, hogy együttműködésüket feltehetően más fehérjék jelenléte nem, vagy csak kis mértékben befolyásolja **(29. ábra, H)**.

A JUNB esetében az anya enhanszerek annak minden fontos, specifikus motívumát reprezentálták: feldúsult az AP-1, az AP-1::IRF (interferon regulatory factor) kompozit eleme, az ún. AICE (AP-1::IRF composite element), a cAMP válaszadó elem (CRE; cAMP response element) és a C/EBP (CCAAT/enhancer-binding protein) fehérje válaszadó eleme is **(29. ábra, I)**. Ugyanezeknek a motívumoknak a gyengébb feldúsulása volt megfigyelhető a leány enhanszerek régióiban is, ugyanakkor további fehérjék, mint a PU.1 (purine-rich box-1), az NFκB (nuclear factor kappa-light-chain-enhancer of activated B cells) és a STAT (signal transducer and activator of transcription) motívumai is megjelentek **(29. ábra, J)**.

Az RAR és VDR transzkripciós faktorok anya enhanszerei is erős, a magreceptorokra specifikus, egymás utáni félhely párokat (direct repeat, DR motívumokat) eredményeztek, eltérő hosszúságú (0-3 bp) összekötő szekvenciával (DR0, DR1, DR2, DR3). A VDR esetében feltehetően a GATA motívum szerepe az anya és leány enhanszereknél egyaránt számottevő, a leány enhanszerek pedig további fehérjék, mint például az AP-1 és a KLF (Krüppel-like factor) motívumait és maguknak a fehérjéknek a jelenlétét igénylik **(29. ábra, K-N)**.

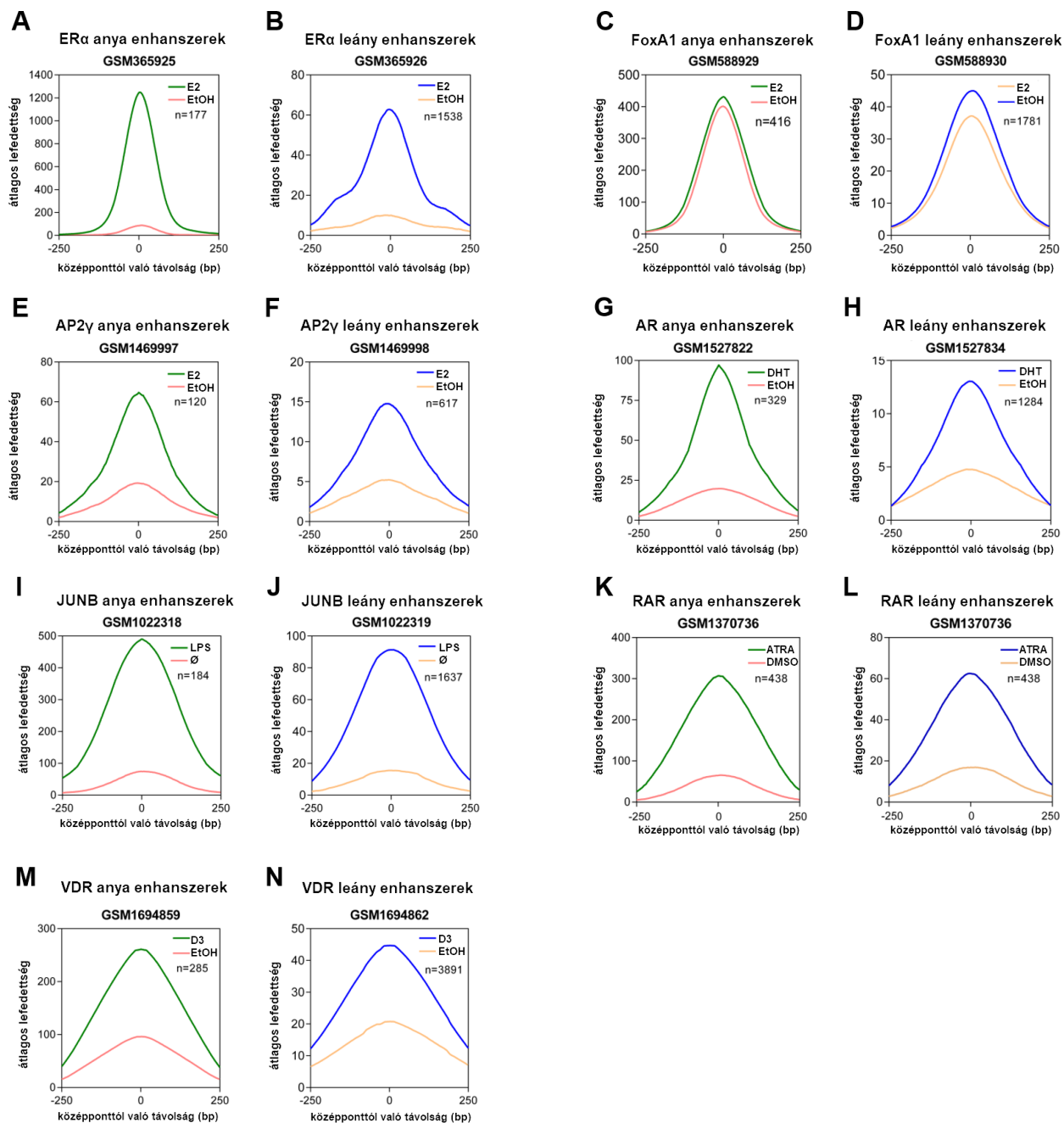


27. ábra. A leány enhanszerek megjelenlése a specifikus ligand hatására. Az ábramagyarázat a következő oldalon található.



27. ábra. A leány enhanszerek megjelenése ligand aktiválást követően. A hő térképek a különböző SE-ek anya (M, mint „mother”) és a top 6 (1-6) leány enhanszereinek lefedettségértékeit (RPKM) reprezentálják az oldószerrel és a specifikus liganddal kezelt ER α (A, B), FoxA1 (C, D), AP2 γ (E,F), AR (G, H), JUNB (I, J), RAR (K, L) és VDR (M, N) ChIP-seq mintákból számolva. Minden sor egy SE-t jelenít meg; az első oszlop minden esetben az anya enhanszer, ugyanazon sorban a további oszlopok pedig a top 6 leány enhanszert jelölik. Az összetartozó hő térképeken a sorok és az oszlopok sorrendjei megegyeznek. A SE-ek számai (n) a mintákban a hő térképek alatt vannak feltüntetve. Ligandok

rövidítései: E2: ösztradiol; DHT: dihidrotesztoszteron; LPS: lipopoliszaharid; ATRA: csupa-transz retinsav; D3: kalcitriol.



28. ábra. Az anya és leány enhanszerek átlagos lefedettsége specifikus ligandkezelés előtt és után. A hisztogramok az ER α , FoxA1, AP2 γ , AR, JUNB, RAR és VDR ChIP-seq mintákban azonosított anya és leány enhanszerek átlagos lefedettségét ábrázolja kezeletlen ($\emptyset$) vagy oldószeres (EtOH vagy DMSO) és specifikus ligand kezelést (ösztadiol (E2), dihidrotesztoszteron (DHT), lipopoliszaharid (LPS), csupa-transz retinsav (ATRA) vagy kalcitriol (D3)) követően. A lefedettség értékek a kötőhelyek középpontjától mért ± 250 bázispárnyi (bp) távolságon, 1 bp felbontással lettek számolva és ábrázolva.

A ER α anya enhanszerek (GSM365925)

Target szekvencia: 147

	P-value	Target %	Bg %	Motivum
	1e-141	58.5 %	0.74 %	ERE
	1e-24	32.65 %	5.21 %	ER0
	1e-18	20.41 %	2.45 %	NR half
	1e-13	12.24 %	1.05 %	NR half

B ER α leány enhanszerek (GSM365925)

Target szekvencia: 1,082

	P-value	Target %	Bg %	Motivum
	1e-57	14.70 %	3.11 %	ERE
	1e-109	20.71 %	4.95 %	DR (-1)

C FoxA1 anya enhanszerek

Target szekvencia: 340

	P-value	Target %	Bg %	Motivum
	1e-204	73.53 %	7.22 %	FoxA1
	1e-17	2.35 %	0.01 %	ERE
	1e-13	11.47 %	2.58 %	Fox(J1)

D FoxA1 leány enhanszerek

Target szekvencia: 1,305

	P-value	Target %	Bg %	Motivum
	1e-258	35.86 %	5.03 %	FoxA1
	1e-40	9.73 %	2.29 %	ERE
	1e-33	3.45 %	0.27 %	FoxA2
	1e-21	8.43 %	2.88 %	AP-1
	1e-21	3.07 %	0.41 %	AP2 γ

E AP2 γ anya enhanszerek

Target szekvencia: 109

	P-value	Target %	Bg %	Motivum
	1e-21	69.72 %	25.09 %	AP2 γ
	1e-13	6.42 %	0.04 %	unknown

F AP2 γ leány enhanszerek

Target szekvencia: 542

	P-value	Target %	Bg %	Motivum
	1e-41	34.32 %	11.97 %	AP2 γ
	1e-23	10.15 %	1.79 %	FoxA1
	1e-19	4.24 %	0.27 %	ERE
	1e-14	13.84 %	4.91 %	CTCF
	1e-13	5.90 %	1.07 %	AP-1

G AR anya enhanszerek

Target szekvencia: 250

	P-value	Target %	Bg %	Motivum
	1e-148	48.8 %	1.52 %	ARE
	1e-63	52 %	9.63 %	FoxA1

H AR leány enhanszerek

Target szekvencia: 821

	P-value	Target %	Bg %	Motivum
	1e-122	22.29 %	2.17 %	ARE
	1e-79	28.26 %	6.69 %	FoxA1
	1e-19	17.9 %	7.92 %	Fox:ARE

29. ábra. A további vizsgált faktorok anya és leány enhanszereinek motívumfeldúsulásai. Az ábramagyarázat a következő oldalon található.

I JUNB anya enhanszerek					J JUNB leány enhanszerek				
Target szekvencia: 187	P-value	Target %	Bg %	Motivum	Target szekvencia: 1,637	P-value	Target %	Bg %	Motivum
	1e-77	48.13 %	3.52 %	AP-1		1e-256	29.14 %	4.07 %	AP-1
	1e-21	12.83 %	0.78 %	AICE		1e-109	20.71 %	4.95 %	PU.1
	1e-15	20.32 %	3.96 %	CRE		1e-59	19.30 %	6.95 %	C/EBP
	1e-14	16.58 %	2.9 %	C/EBP		1e-39	11.42 %	3.74 %	CRE
						1e-38	10.14 %	3.10 %	NFkB
						1e-23	4.09 %	0.78 %	AICE
						1e-23	6.78 %	2.24 %	STAT

K RAR anya enhanszerek					L RAR leány enhanszerek				
Target szekvencia: 93	P-value	Target %	Bg %	Motivum	Target szekvencia: 438	P-value	Target %	Bg %	Motivum
	1e-50	49.46 %	2.12 %	DR0		1e-145	31.28 %	1.26 %	DR0
	1e-24	31.18 %	2.28 %	DR2		1e-23	8.45 %	0.88 %	DR1
	1e-23	12.9 %	0.07 %	DR2		1e-13	18.04 %	7.00 %	unknown
	1e-15	20.32 %	3.96 %	DR3		1e-13	10.5 %	2.78 %	unknown

M VDR anya enhanszerek					N VDR leány enhanszerek				
Target szekvencia: 285	P-value	Target %	Bg %	Motivum	Target szekvencia: 3,891	P-value	Target %	Bg %	Motivum
	1e-101	45.96 %	4.05 %	VDRE		1e-152	22.23 %	8.40 %	DR1
	1e-21	14.74 %	2.16 %	NR half		1e-52	8.4 %	3.21 %	VDRE
	1e-14	5.96 %	0.38 %	GATA		1e-47	5.27 %	1.58 %	AP-1
	1e-14	16.84 %	4.47 %	NR half		1e-46	9.69 %	4.28 %	GATA
						1e-35	16.94 %	10.30 %	KLF

29. ábra. A további vizsgált faktorok anya és leány enhanszereinek motívumfeldúsulásai. Az anya-és leány enhanszerek kötőhelyei alatt a háttérhez képest statisztikailag feldúsult motívumok az ER α , FoxA1, AP2 γ , AR, JUNB, RAR és VDR mintákban. (Target %: megadja, hogy a vizsgált anya és leány régiók hány százalékában található meg az adott motívum; Bg%: megadja, hogy a véletlenszerűen generált, háttérként szolgáló szekvenciákban a feldúsult motívum hány százalékában van jelen; unknown: nem ismert motívumot jelöli).

Eredményeink alapján elmondható, hogy minden vizsgált indukálható transzkripció faktor esetében ugyanazt a jelenséget figyeltük meg, mint eredendően az ER α esetében: az elsődlegesen aktív szabályozó régiók tulajdonában egy kanonikus vagy közel tökéletes válaszadó elem állt, amely specifikus a vizsgált, az adott sejtben/sejtvonalban domináns transzkripció faktorra, míg a másodlagosan elfoglalt kötőhelyek sokkal gyengébb motívumokon keresztül, és sokszor társfaktor(ok)ra hagyatkozva stabilizálták az egyébként is kisebb mennyiségű fehérjét.

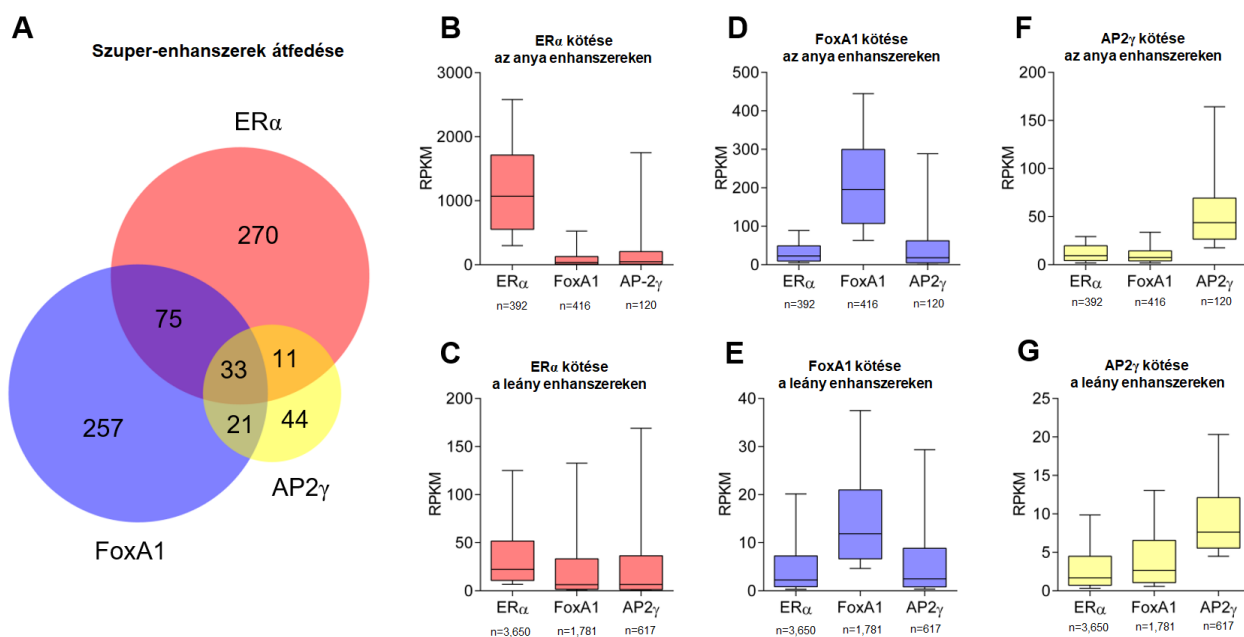
Az eredmények alapján azt feltételezzük, hogy a kanonikus elemeknek a fehérjék általi elfoglalása feltehetően megkönnyíti vagy elősegíti a szomszédos régiók elfoglalását, és ez nemcsak az MCF-7 sejtvonalra és nemcsak az ER α -ra specifikusan jellemző jelenség, hanem egyaránt megfigyelhető más sejtekben és más ligand által indukálható transzkripció faktorok esetében is.

4.4. Az ER α , FoxA1 és AP2 γ jelenlétének vizsgálata egymás szuper-enhanszer régióiban

A következő lépésben megvizsgáltuk az ER α legfontosabb társfaktoraiként számontartott fehérjék, a FoxA1 és az AP2 γ -vezérelt szuper-enhanszerek átfedését az ER α szuper-enhanszereivel, továbbá azt, hogy az egyes faktorok lefedettsége milyen mértékű a másik két faktor anya és leány enhanszerein (**30. ábra**). Bár a motívumfeldúsulások azt mutatták, hogy a másik két faktor motívumán keresztül nagyon kevés százalékban jön létre kötés – igaz ez mindhárom faktorra –, ez nem zárta ki azt, hogy másodlagosan, fehérje-fehérje interakción keresztül nem játszanak szerepet egymás DNS-hez való kihorgonyzásában.

Összesen 416 FoxA1 és csupán 120 AP2 γ szuper-enhanszert sikerült azonosítanunk, melyek a 392 ER α szuper-enhanszerrel – és egymással is – csak részben mutattak átfedést (**30. ábra, A**). Fontos megjegyezni, hogy bár a szuper-enhanszer régiók meghatározásakor gyakran

technikai problémákba ütközhetünk, hiszen két azonos transzkripciós faktor, vagy akár két párhuzamos ChIP-seq mintái közötti minőségbeli eltérés nagyban befolyásolhatja a szuper-enhanszerként azonosított régiókat, jelen összehasonlítás célját ez az esetleges technikai probléma nem torzítja. Ennek oka az, hogy ha csupán a 416 FoxA1, 120 AP2 γ és 392 ER α szuper-enhanszerre támaszkodunk, ezen régiók kötőhelyei elegendőek ahhoz, hogy egy reprezentatív képet adjanak arról, hogy egy transzkripciós faktor jelenléte milyen mértékben jellemzi a másik két faktor szuper-enhanszereit.

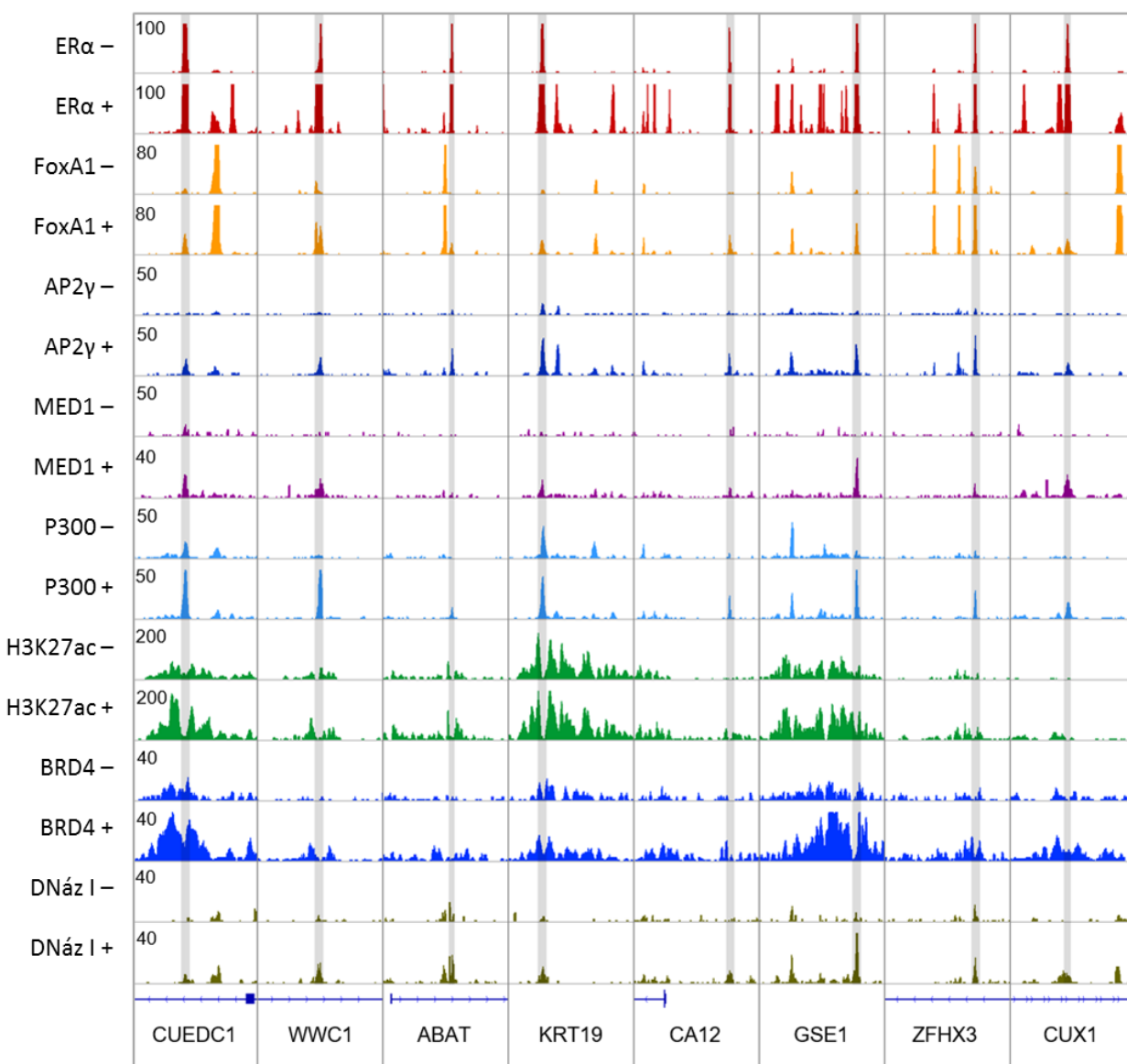


30. ábra. Az ER α , FoxA1 és AP2 γ eltérő szuper-enhanszereken működnek. (A) A Venn diagram az ER α , FoxA1 és az AP2 γ szuper-enhanszereinek átfedését reprezentálja. **(B-G)** A **B**, **D** és **F** panelek az ER α , a FoxA1 és az AP2 γ anya enhanszereken számolt ER α , FoxA1 és AP2 γ lefedettségeket, míg a **C**, **E** és **G** panelek az ER α , a FoxA1 és az AP2 γ leány enhanszereken számolt ER α , FoxA1 és AP2 γ lefedettségeket (RPKM) reprezentálják, E2 kezelést követően.

A **30. ábra B-G** paneljeiből, melyek az ER α , FoxA1 és az AP2 γ lefedettségét reprezentálják egymás anyá és leány enhanszerein, jól kivehető, hogy míg az ER α anyá enhanszerein szinte csak az ER α jelenléte figyelhető meg, a FoxA1 és az AP2 γ anyá enhanszerei sem mutatják a másik két faktor jelentős feldúsulását (**30. ábra, B, D, F**). Hasonló mintázat figyelhető meg a leány enhanszereken is. A domináns transzkripciós faktor jelenléte itt bár sokkal csekélyebb, de így is markánsabb a vizsgált társfaktorokéhoz képest (**30. ábra, B, D, F**). Mindez arra utal, hogy a szuper-enhanszereket jellemzően egy domináns fehérje jelenléte jellemzi és a többi fehérje feltehetően csak elvétve, a szuper-enhanszerek csupán néhány elemén keresztül járul hozzá a végső aktivitás eléréséhez.

4.5. ER α -vezérelt szuper-enhanszerek: áttekintés

Az ER α -vezérelt szuper-enhanszereknek a 4.1 – 4.4 fejezetekben ismertetett tulajdonságait egy összesítő ábra reprezentálja, néhány szuper-enhanszer régió által (**31. ábra**). Az ábrán megfigyelhető, hogy legalább egy FoxA1 kötés minden szuper-enhanszeren belül megtalálható már az E2 kezelés előtt is, de ezek tipikusan a leány enhanszereket jellemzik. Habár a FoxA1 mennyisége E2 kezelést követően globálisan nem emelkedik (**28. ábra, C és D panel**), úgy tűnik, kötődése az ER α szuper-enhanszerek egy kis csoportján, az anyá enhanszereken mégis megfigyelhető. Ezzel szemben az AP2 γ nem detektálható, csupán kezelést követően, mégpedig szintén az anyá enhanszereken, és sok esetben azokon a leány enhanszereken, ahol a FoxA1 is. Mennyiségük azonban az ER α -hoz képest alacsony (**30. ábra, B-G**). Feltehetően a FoxA1 ezeken a régiókon nem klasszikus pioneer faktorként funkcionál, hanem az ER α kötés stabilizálásáért felelős komplex részét képezi, az AP2 γ -val együtt.



31. ábra. Az ER α -vezérelt szuper-enhanszer régiók reprezentatív ábrája. Az ábra az ER α , FoxA1, AP2 γ , MED1, P300, H3K27ac és BRD4 ChIP-seq és DNáz I minták lefedettségét mutatja MCF-7 sejtben, a CUEDC1, WWC1, ABAT, KRT19, CA12, GSE1, ZFH3 és a CUX1 géneket szabályozó ER α -vezérelt szuper-enhanszer régiókon, E2 kezelést megelőzően (–) és azt követően (+).

A MED1 jelenléte, ahogy azt a 23. ábra is érzékeltette, többnyire csak az anya enhanszerekre jellemző E2 hozzáadását követően, és ugyanez jellemző a kromatin nyitottságát mérő DNáz I hiperszenzitivitásra is. A P300 mennyisége az anya enhanszereken fokozódik, de

azokon a leány enhanszereken, melyeket a FoxA1 is köt, eltűnik. Mindez jelentheti azt, hogy a FoxA1 által elfoglalt helyek represszorként működnek, vagy a P300 cisztrómja részben átrendeződik. A H3K27ac többnyire az anya enhanszereken, de gyakran a teljes szuper-enhanszer régióon azonosítható, melynek jellegét a BRD4 is teljes mértékben követi és azzal azonos mértékben is fokozódik (**31. ábra**).

4.6. A nem szuper-enhanszerek részét képező ER α kötőhelyek jellemzése

4.6.1. A kötőhelyek csoportosítása az indukálhatóság és az ERE megléte alapján

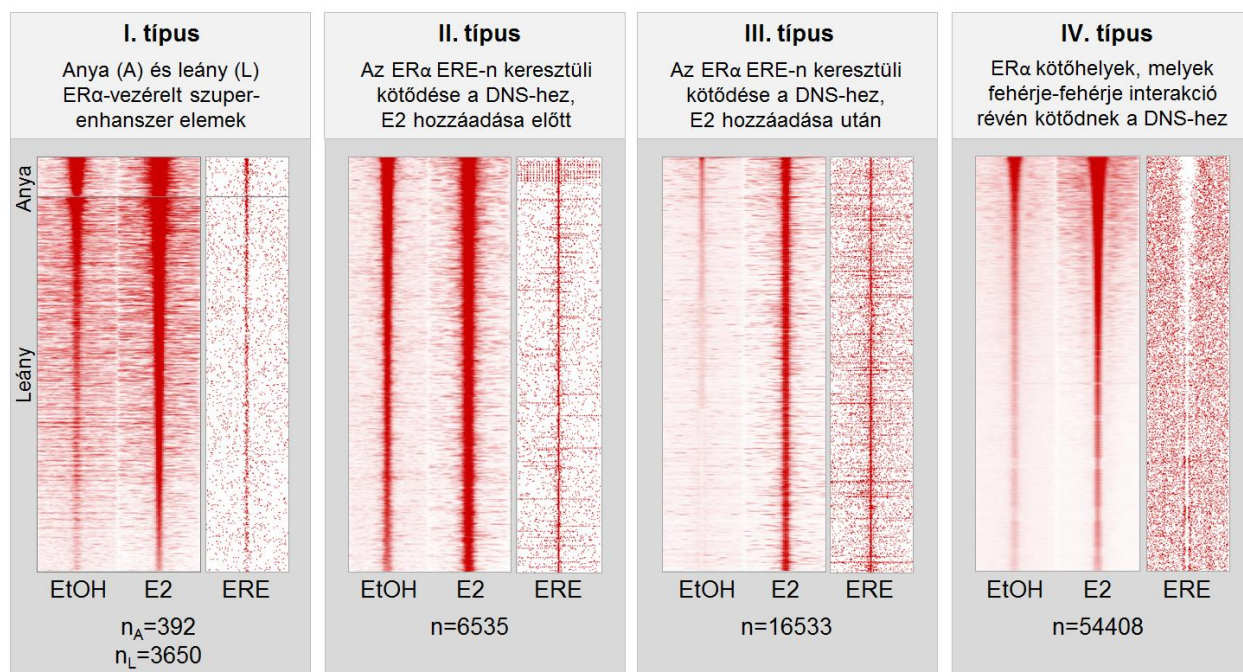
Abból kiindulva, hogy az oldószerrel kezelt ER α ChIP-seq mintában több, mint 20000 kötőhelyet tudtunk azonosítani, azt feltételeztük, hogy lehetnek az anya enhanszerekhez hasonló, olyan ER α kötőhelyek is a genomban, melyek bár nem képezik szuper-enhanszerek részét, de ligand hiányában is kötöttek, és rendelkeznek az anya enhanszerekhez hasonlóan erős, kanonikus válaszadó elemmel. Amennyiben állításunk igaz, a kérdésünk az, van-e különbség a génexpresszióra gyakorolt hatásukban vagy a ko-faktoraik között.

A fentiek megválaszolására az E2 kezelt mintából származó, közel 80000 ER α kötőhelyet osztályoztuk azon tulajdonságaik alapján, hogy az ER α kötődött-e a DNS-hez már ösztadiol hiányában is, vagy csak annak hozzáadását követően, továbbá, hogy a kötőhelyek rendelkeztek-e ERE-vel. Az osztályozáshoz azt az ERE motívum mátrixot használtuk, amely a 392 anya enhanszer motívumainak összetételét reprezentálta a 26. ábrán (**32. ábra**).

GG CA		TG CC	
A	C	G	T
0.520	0.090	0.295	0.095
0.085	0.001	0.824	0.090
0.045	0.015	0.930	0.010
0.105	0.100	0.310	0.485
0.001	0.869	0.050	0.080
0.790	0.015	0.125	0.070
0.160	0.395	0.365	0.080
0.190	0.295	0.315	0.200
0.115	0.305	0.460	0.120
0.090	0.135	0.015	0.760
0.090	0.090	0.819	0.001
0.545	0.270	0.100	0.085
0.025	0.910	0.010	0.055
0.080	0.845	0.005	0.070
0.050	0.330	0.095	0.525

32. ábra. Az ERE motívumhoz tartozó mátrix. A mátrix minden sora a motívum egy bázisát, az oszlopok pedig egy adott nukleotidot reprezentálnak. Egy-egy sor összege 1-et tesz ki, és azt írja le, hogy a vizsgált motívum adott pozíciójában az adott nukleotid milyen százalékos valószínűséggel fordul elő. A mátrix által reprezentált szekvencia a következő: AGGTCASNGTGACCT, melyben az S lehet G vagy C, az N pedig bármely bázis lehet. A mátrixhoz minden esetben tartozik még egy *score*, amely függ a motívum hosszától és a mátrixban nagy százalékos valószínűséggel előforduló bázisok számától. Jelen motívum esetében ez az érték 7,437420 volt, melyet az ERE motívumok térképezése során is megtartottunk.

Az ER α kötőhelyeknek összesen négy csoportját különítettük el a fenti besorolás alapján, melyek közül egy csoportot a szuper-enhanszerek képeztek (I. csoport; n=4042). A második csoportba azok az ER α kötőhelyek kerültek, melyek az anya enhanszerekhez hasonlóan E2 hiányában is ERE-n keresztül voltak elfoglalva (II. csoport; n=6535). A harmadik csoport viszonylag nagyobb elemszámú ER α kötőhelyet foglalt magába, ahol bár azonosítani tudtuk az ERE-t, a receptor kötődése csak kezelést követően volt megfigyelhető (III. csoport; n=16533). Azon típusú kötőhelyek pedig, melyek nem rendelkeztek ERE-vel, a negyedik csoportba lettek sorolva (IV. csoport; n=54408). A négy csoport kötőhelyeit, az ERE motívum meglétét vagy hiányát a következő ábra ismerteti (**33. ábra**).

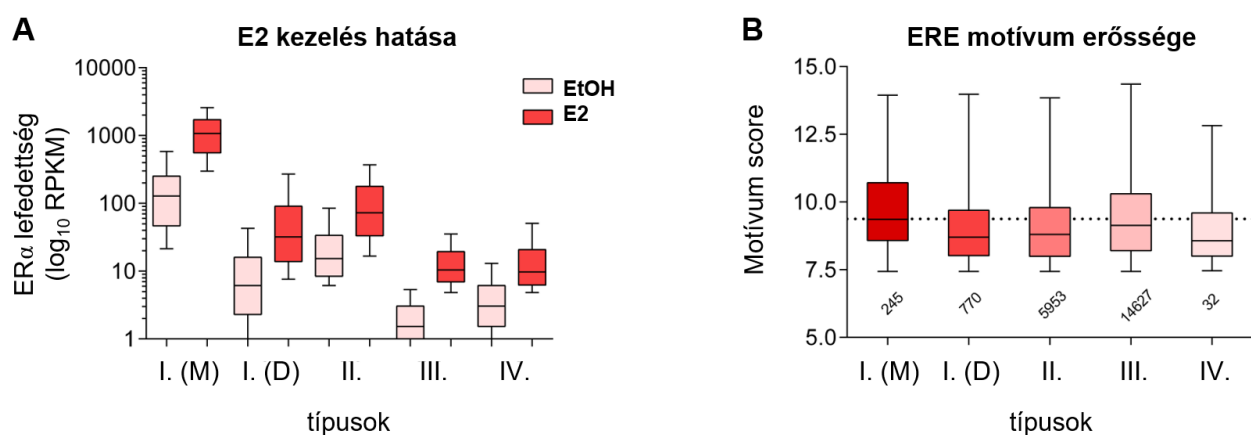


33. ábra. Az ER α kötőhelyek különböző típusai. A hő térképek az ER α kötőhelyek négy típusának (I-IV. típus) kötőhelyeit és az ERE jelenlétét mutatják oldószeres (EtOH) és E2 kezelést követően. A minták lefedettségei és az ERE jelenléte az E2-kezelt ER α minta kötőhelyeinek középpontjához viszonyított ± 1 kbp-nyi régióknak lettek ábrázolva.

4.6.2. Az ösztradiol hatása az ER α kötési gyakoriságára a különböző csoportokban

A négy csoport ER α kötőhelyeinek lefedettség értékeit ábrázolva azt tapasztaltuk, hogy bár a II. típusú kötőhelyek az anya enhanszerekhez hasonló tulajdonságokkal rendelkeznek, és a dinamikája is azonos (értem ez alatt az ERE meglétét és a ligand hiányában történő DNS-kötést), kötési gyakorisága E2 hatására alig haladja meg a leány enhanszerekét (**34. ábra, A**). Ezt a markáns különbséget okozhatja az, hogy az anya enhanszereknek erősebb válaszadó elemeik vannak, illetve a szuper-enhanszerek tagjaiként a leány enhanszerek tovább növelhetik a DNS/fehérje kölcsönhatások számát (**34. ábra, B**).

Az ösztradiol hatása, bár mindegyik csoportban számottevő volt, a III. típusú, újonnan, ERE-n keresztül elfoglalt *de novo* enhanszerek, és az indirekt módon kötöttek, végeredményében sokkal kisebb fehérjemennyiséget mutattak (**34. ábra, A**). Ezeken a régiókon a kromatin feltehetően zárt állapotából válik nyitottá, a fehérje pedig csak ezt követően képes megtalálni a helyét, a IV. típusú kötőhelyek esetében pedig azt a fehérjét vagy komplexet, amelyhez kapcsolódni képes.



34. ábra. Az ösztradiol hatása az ER α kötésére. (A) A box plotok az I-IV. típusú ER α kötőhelyek erősségét mutatják oldószeres (EtOH) és ösztadiolos (E2) kezelést követően. **(B)** A box plotok az ERE motívum erősségét reprezentálják. A számok az adott csoportban azonosított ERE-k számát jelölik. A lefedettségértékek (RPKM) meghatározása és az ERE motívumok (vissza)térképezése az E2-kezelt kötőhelyek középpontjához viszonyított ± 100 bp-nyi régióon történt.

4.6.3. A különböző csoportokban feldúsult szekvenciamotívumok

Annak meghatározására, hogy mely fehérjéken keresztül képesek a IV. típusú kötőhelyek létrejönni, illetve mutatnak-e különbséget a II. típusú kötőhelyek az anya enhanszerekhez képest, *de novo* motívumelemzést végeztünk (**35. ábra**). A IV. típus esetében a magreceptor motívum félhelye bár azonosítható volt, a háttérszekvenciákhoz képest csak két és félszeres feldúsulást

mutatott (56,36% vs. 22,81%). Feldúsult továbbá a Fox, TCF, AP-1, AP2 és GATA fehérjék motívuma. Utóbbi bár nagyon kis százalékban, de nagyon szépen kirajzolódó, fordítottan ismétlődő (inverted repeat, IR) formában, ahol a két motívumot 3 bázispárnyi random szekvencia választja el (GATA-IR3) (35. ábra). A GATA3 motívuma a FoxA1-hez és AP2 γ -hoz hasonlóan – ahogy arról a bevezetésben is szó esett –, jelentős feldúsulást mutat az ER α által kötött genomi régiókban (Lin et al., 2007; Theodorou, Stark, Menon, & Carroll, 2013). Azt, hogy a tökéletesen szimmetrikus 'CAGGTTAACCTG' motívumot mely fehérje képes kötni, nehezen megjósolható. Legközelebb vélhetően a TCF1 (T cell factor 1, melyet a TCF7 gén kódol), fehérje motívuma áll, mellyel 7 bázis átfedést mutat ('AGTTAACTAA'). A II. típusú kötőhelyeknél a több mint 90%-os ERE feldúsulás mellett (92,27%) bizonyára az AP2, AP-1 és Fox fehérjék jelenléte lehet az, amely a viszonylag gyengébb ERE motívumokat kompenzálni tudják és megteremtik azt, hogy az anya enhanszerekéhez hasonló dinamikát produkáljanak. A szuper-enhanszerek esetében ugyanezen fehérjék motívumai a leány enhanszerek kialakulását segítik elő (35. ábra).

II. típus					IV. típus				
	P-érték	Target %	Bg %	Motívum		P-érték	Target %	Bg %	Motívum
	1e-6167	92.27 %	2.80 %	ERE		1e-4267	56.36 %	22.81 %	NR félhely
	1e-117	11.25 %	3.51 %	AP2		1e-1178	16.46 %	5.77 %	FoxA1
	1e-96	9.60 %	3.07 %	AP-1		1e-468	17.59 %	9.79 %	TCF
	1e-55	7.83 %	3.10 %	FoxA1		1e-427	4.98 %	1.48 %	AP-1
III. típus						1e-387	2.96	0.63 %	GATA-IR3
	P-érték	Target %	Bg %	Motívum		1e-286	7.61 %	3.63 %	AP2
	1e-10937	89.35 %	5.66 %	ERE					

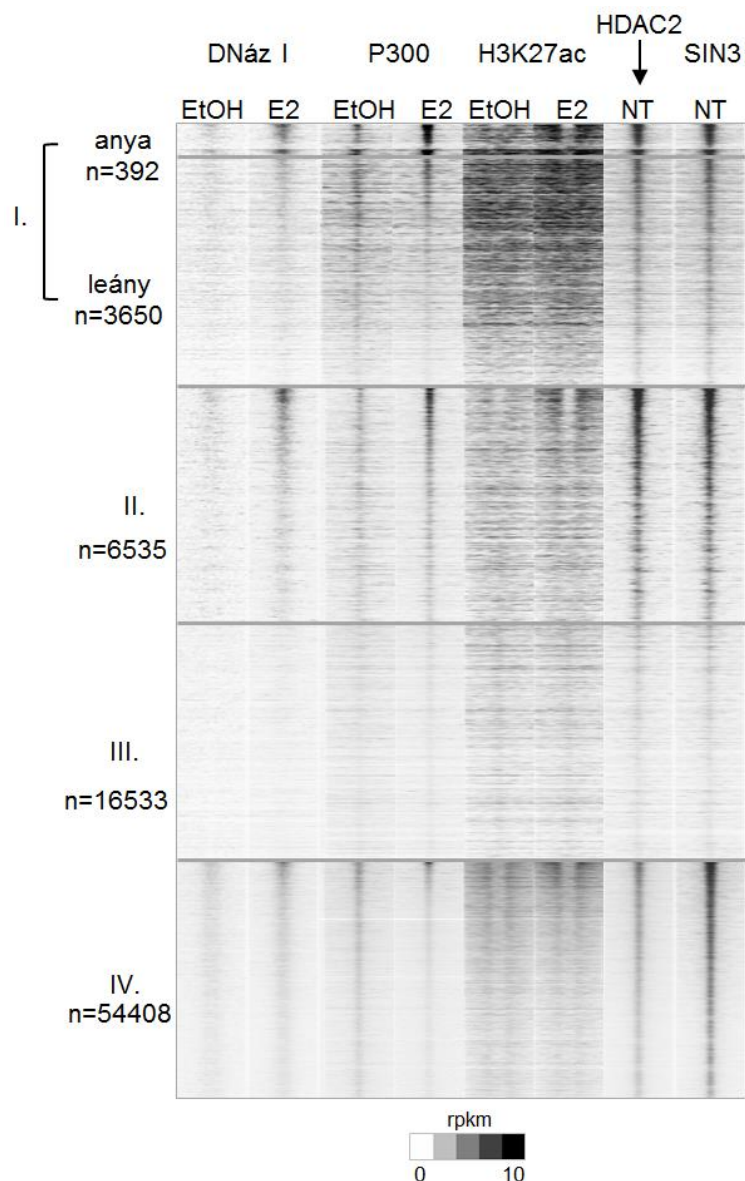
35. ábra. Motívumfeldúsulás. A II-IV. típusú ER α kötőhelyek alatt a háttérhez képest statisztikailag feldúsult motívumok. (Target %: megadja, hogy a vizsgált régiók hány százalékában található meg az adott motívum; Bg%: megadja, hogy a véletlenszerűen generált, háttérként szolgáló szekvenciák alatt az adott motívum hány százalékban van jelen).

4.6.4. A kromatin hozzáférhetőségének vizsgálata a különböző csoportokban

Ahhoz, hogy megértsük, a kromatin nyitott/zárt állapota miként járul hozzá a különböző típusú kötések kialakulásához, megvizsgáltuk a kromatin nyitott állapotát jelző DNáz I, P300, H3K27ac és a represszált, zárt állapotot jelző HDAC2 (histone deacetylase 2) és SIN3A (SIN3 transcription regulator family member A) fehérjék és hiszton módosító enzimek jelenlétét **(36. ábra)**.

A legszembevetőbb talán az, hogy az anya és leány enhanszereket minden további csoporttal összevetve sokkal markánsabb acetiláltság jellemezte, amely csak kis mértékben fokozódott E2 hatására. Míg a P300 mennyisége az anya enhanszereknél és a II. típusú kötőhelyeken megnőtt, a leány enhanszereknél mérséklődött.

Érdekes, hogy a III. típusú *de novo* aktiválódott kötőhelyek nem mutatnak változást a kromatin nyitottságát tekintve; sem aktív, sem represszált állapotot jelző fehérje jelenléte nem tapasztalható. Habár nincs bizonyíték a korábbi tanulmányok alapján arra, hogy az ER α mint *pioneer* faktor funkcionálna, úgy tűnik, több mint 16000 genomi régió minden további segítség nélkül képes ezt a feladatot ellátni. Míg a HDAC2 és a SIN3A feldúsulása eddig következetes volt, a IV. típusú kötőhelyeken a HDAC2 mennyisége a SIN3A-éhoz képest sokkal csekélyebb, mindemellett a P300 és a H3K27ac is csökkent intenzitást mutat **(36. ábra)**.



36. ábra. A kromatin hozzáférhetősége a különböző típusú ERα kötőhelyeknél. A hőterkép az oldószerrel (etanol; EtOH), ösztradiollal (E2) kezelt, vagy a kezeletlen (non-treated; NT) DNáz I enzim által elemésztett, a P300, HDAC2 és SIN3A fehérjék által kötött, továbbá a H3K27ac-val jelölt nukleosóma régiókat reprezentálja az ERα kötőhelyek középpontjához viszonyított ± 1 kbp-nyi régió az anya (n=392) és leány (n=3650) enhanszereken (I. típus), továbbá a II. III. és IV. típusú ERα kötőhelyeken.

4.6.5. A FoxA1, AP2γ és GATA3 fehérjék megléte a különböző csoportokban

A 4 különböző típusba sorolt ~80000 ERα transzkripció faktor kötőhely motívumpreferenciái részben utaltak arra, hogy mely fehérje/fehérjék közvetlen DNS-hez történő kapcsolódásán keresztül jön létre az ERα kötése, ugyanakkor nem zárták ki az ERα transzkripció szabályozásában résztvevő további fehérjék másodlagos jelenlétét. Ebből kiindulva megvizsgáltuk a FoxA1, AP2γ és a GATA3 fehérjék, továbbá magának az ERα-nak a jelenlétét a

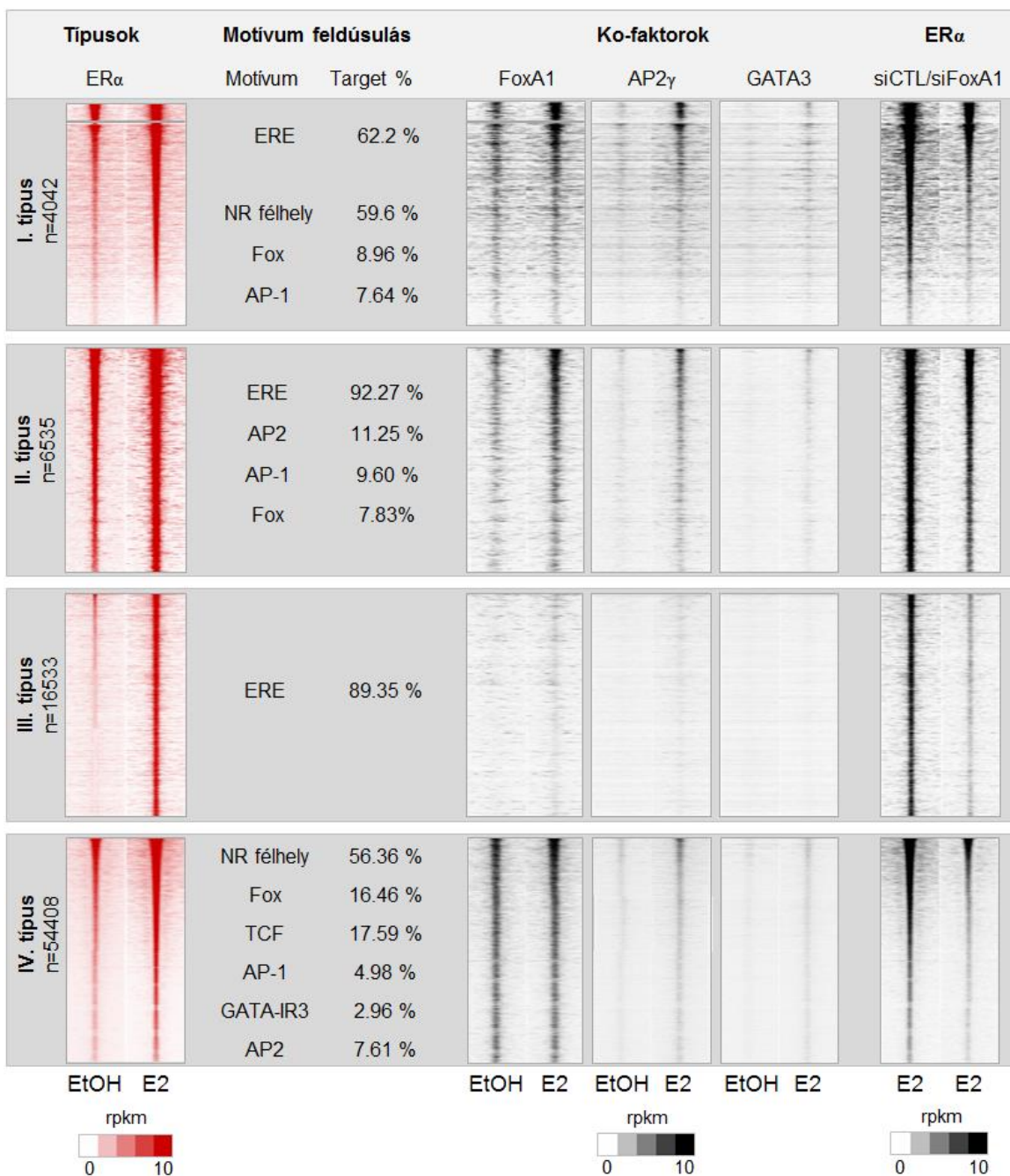
FoxA1 fehérje siRNS-mediált géncsendesítését követően a szuper-enhanszerek egyedi elemein és a II-IV. típusú ER α kötőhelyeken.

Annak ellenére, hogy az E2 teljes genom szinten vizsgálva nem okoz változást a FoxA1 kötésében, az anya enhanszereken a FoxA1 jelentős mértékű feldúsulása volt megfigyelhető (Hurtado et al., 2011) **(37. ábra)**. A FoxA1 meghatározó szerepét az is érzékelteti, hogy hiánya (siFoxA) teljes körű csökkenést eredményez az ER α kötések erősségében, a leány enhanszereknek egy nagy hányada pedig el is tűnik. Mindez arra utal, hogy a FoxA1-nek bár csupán másodlagos, de nagyon fontos szerepe van az ER α -vezérelt szuper-enhanszerek kialakulásában. A FoxA1 hatását tovább fokozza az AP2 γ jelenléte az anya enhanszereken, ugyanakkor a GATA3 csak elenyésző mértékű kötődést mutat **(37. ábra)**.

A II. típusú kötőhelyeket mind a FoxA1, mind az AP2 γ globális jelenléte jellemezte E2 kezelést követően, ugyanakkor a FoxA1 hiánya nem eredményezett olyan drámai hatást az ER α kötésében; alighanem az AP2 γ megléte képes ellensúlyozni a FoxA1 hiányát **(37. ábra)**.

A III. típusú, *de novo* elfoglalt kötőhelyek az ERE-n kívül, bár minden egyéb fehérje motívumától és azok közvetlen jelenlététől függetlenek, a FoxA1 hiányában az ER α kötések teljes egészében eltűntek. A FoxA1 valószínűleg a kromatin hurkolódásának köszönhetően, távoli szabályozás révén kerül fizikai közelségbe ezekkel az ER α kötőhelyekkel és hatását ily módon képes kifejteni **(37. ábra)**.

Azon ER α kötőhelyek jelentős hányada, amelyeket nem jellemez az ERE (IV. típus), szintén eltűnik a FoxA1 hiányában. Bár a Fox motívum csak 16,46%-os feldúsulást mutatott, úgy tűnik, akár a TCF, akár az AP-1, GATA vagy AP2 fehérjéken keresztül jön létre az ER α kötődése, voltaképpen a FoxA1 a fehérjekomplex részét képezi már a kezelést megelőzően is **(37. ábra)**.

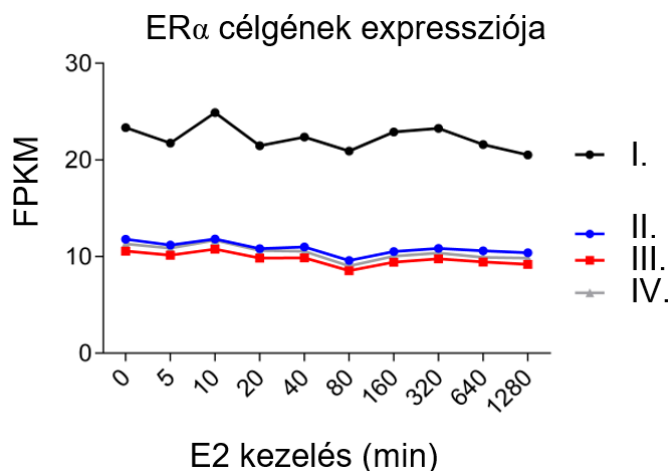


37. ábra. Ko-faktorok jelenléte a különböző típusú ER α kötőhelyeken. A hőterképek az I-IV. típusú ER α kötőhelyeket, azok motívumfeldúsulásait és a FoxA1, AP2 γ és GATA3 fehérjék jelenlétét demonstrálják etanollal (EtOH) és ösztradiállal (E2) történő kezelést követően, továbbá az ER α kötést kontoll (siCTL) és a FoxA1 siRNS-mediált géncsökkentését követő (siFoxA1) kondícióban, E2 kezelést követően, az ER α kötőhelyek középpontjához viszonyított ± 1 kbp-nyi régióban.

4.7. Az ER α célgének expressziója

A szuper-enhanszerek nem csupán a magukhoz toborzott ko-aktivátorok mennyisége és az ebből fakadó aktivitásuk miatt kivételesek; a génekre gyakorolt hatásuk szintén meghatározó. Többnyire az adott sejt jellegéért felelős gének szabályozását koordinálják, mégpedig szokatlanul magas expressziót produkálva (Hnisz et al., 2013; Whyte et al., 2013).

Utolsó lépésben arra voltunk kíváncsiak, hogy a mi vizsgálati rendszerünkben is képesek-e a szuper-enhanszerek valóban markánsabb transzkripciók kimenetet produkálni a tipikus enhanszerek csoportjaival szemben. Ehhez meghatároztuk az ER α kötőhelyekhez/szuper-enhanszerekhez legközelebbi fehérjekódoló géneket, majd ábrázoltuk az E2 kezelést követő különböző időpontokban mért expressziójuk átlagos értékét. Azok a gének, melyek minden vizsgált időpontban 0 expressziós értéket mutattak, nem lettek ábrázolva (**38. ábra**).



38. ábra. A különböző típusú ER α célgének átlagos expressziója. A hisztogram az I-IV. típusú ER α kötőhelyekhez annotált és az általuk potenciálisan szabályozott fehérjekódoló gének expresszióját (FPKM) reprezentálja az E2 kezelést követő 0., 5., 10., 20., 40., 80., 160., 320., 640. és 1280. percben, átlagolva.

Míg a II., III. és IV. típusba sorolt ER α kötőhelyek nagyon hasonló, átlagosan 10 FPKM körüli génexpressziót mutattak, az I-es típusba sorolt szuper-enhanszerek ennek – minden időpontban – legalább a kétszeresét produkálták. Az expresszió időbeli változásának mintázata mindegyik típus esetében hasonló volt, melyek közül az E2 kezelést követő 10. és a 320. percben

két jelentősebb kiugrás volt megfigyelhető. Mindez feltehetően azért rajzolódott ki, mert az ábrázolt gének jelentős hányadát vagy korai, vagy kései kifejeződés jellemzi az MCF-7 sejten **(38. ábra)**.

4.7.1. A szuper-enhanszerek által szabályozott gének

Az ER α -vezérelt szuper-enhanszerek által szabályozott gének között olyan, az emlőrákos folyamatokban szerepet játszó géneket azonosítottunk, mint például a BCAS3, KRT8, KRT19, AZIN1, SLC9A3R1, CXXC5, SLC7A5, HES1 vagy a CCND1.

A BCAS3 (Breast Carcinoma Amplified Sequence-3) míg az emlőrákos sejtekben az ER α ko-aktivátoraként ismert, a KRT8 és KRT19 (Keratin 8 és 19), az AZIN1 (Antizyme inhibitor 1), SLC9A3R1 (SLC9A3 Regulator 1), SLC7A5 (Solute Carrier Family 7 Member 5), CXXC5 (CXXC Finger Protein 5) és a HES1 (Hes Family BHLH Transcription Factor 1) gének megváltozott expressziójuk révén a rossz prognózissal, a kemoterápiás rezisztenciával vagy metasztatikus állapottal mutatnak korrelációt az emlőrákos betegekben (El Ansari et al., 2018; Gururaj, Peng, Vadlamudi, & Kumar, 2007; Saha, Kim, Yang, Choi, & Cho, 2018; Tomaskovic-Crook, Thompson, & Thiery, 2009; Walker et al., 2007). A CCND1 (Cyclin D1), mint a sejtciklus szabályozásában központi szerepet játszó gén, számos daganattípusban – többek közt a primer emlődaganatok több mint 90%-ában is – magas expresszióval jellemezhető, amplifikációja pedig az ER⁺ emlődaganatok 10-20%-ában figyelhető meg (Al-Kuraya et al., 2004; Reis-Filho et al., 2006; Xu & Lin, 2018).

Az általunk azonosított ER α -vezérelt szuper-enhanszerekről is elmondható, hogy többnyire olyan gének közelében helyezkednek el, amelyek az emlő-eredetű MCF-7 sejtvonalban valóban kulcsfontosságú szereppel bírnak, az őket felépítő egyedi enhanszerek együttműködése

révén az általuk produkált transzkripció kimenet pedig sokkal markánsabb a tipikus enhanszerekéhez képest **(38. ábra)**.

5. Diszkusszió

A emberi genom megközelítőleg 99%-a fehérjét nem kódoló régiókból áll, melyet az 1960-as években – az akkor nem létezőnek vélt funkciójuk tekintetében – szemét DNS-nek (ún. *junk DNA*-nek) nevezték el (Ehret & De Haller, 1963). A nem kódoló régiók zöme (~55%-a) bár repetitív elem, ugyanakkor olyan fontos, a transzkripció szabályozásában részt vevő régiókat is tartalmaznak, mint például a promóterek, enhanszerek, silencer-ek vagy az inzulátorok (Lander et al., 2001; Yifeng Li, Chen, Kaye, & Wasserman, 2015). Azonosításuk csak az első lépés ahhoz, hogy a transzkripciós szabályozás további mechanizmusait is megfejthessük, hiszen a kromatin hozzáférhetősége, a transzkripciós faktorok kötődése, a DNS metilációja, a hisztonmódosítások és a különböző interakciók az egyes elemek között mind-mind ennek az összetett szabályozásnak egy újabb szintjét képezik.

A szabályozó régiók genom-szinten való megismerését a szekvenálási technikák folyamatos fejlődése tette lehetővé. A szekvenálásból származó adathalmaz feldolgozásának igénye magával hozta a bioinformatika exponenciális fejlődését, és ma már az újabbnál újabb programcsomagok megjelenésével adatainkat egyre hatékonyabb módon tudjuk elemezni. A transzkripciós faktor kötőhelyek feltérképezése forradalmasította a betegségek vizsgálatára irányuló kutatásokat is. Az ER α , mely az ösztrogén-szenzitív emlőrákos betegekben az elsődleges gyógyszer-célpontot jelenti, szerepe a transzkripciós szabályozásban bár közel három évtizede ismert, annak pontos működése csak az utóbbi egy évtizedben, a ChIP-seq technika révén körvonalazódott (Flach & Zwart, 2016).

Jason S. Carroll és munkatársai 2005-ben a microarray-vel kombinált ChIP technikát (ChIP-on-chip) alkalmazva összesen 3665, nem sokkal később, 2009-ben Willem-Jan Welboren és munkatársai az első ER α ChIP-seq kísérletből pedig már 10205 ER α transzkripciós faktor

kötőhelyet azonosítottak az emberi genomban (Carroll et al., 2006; Welboren et al., 2009). Ezek az eredmények óriási paradigmaváltást eredményeztek. Az addigi tudásunkkal ellentétben a genom-szintű adatok arra utaltak, hogy az ER α transzkripciós faktor kötőhelyeknek csupán 7%-a található közvetlenül a gének promóterén, 93%-uk pedig *cisz* szabályozó elemeken keresztül fejti ki hatását (Welboren et al., 2009). Az enhanszerek és a vizsgált sejtben kifejeződő gének (~2000) száma közötti különbségből feltételezték, hogy egy-egy gént jellemzően több, 2-3 enhanszer is szabályozhat egyidőben, mégpedig a promóterhez való hurkolódásuk révén (Fullwood et al., 2009).

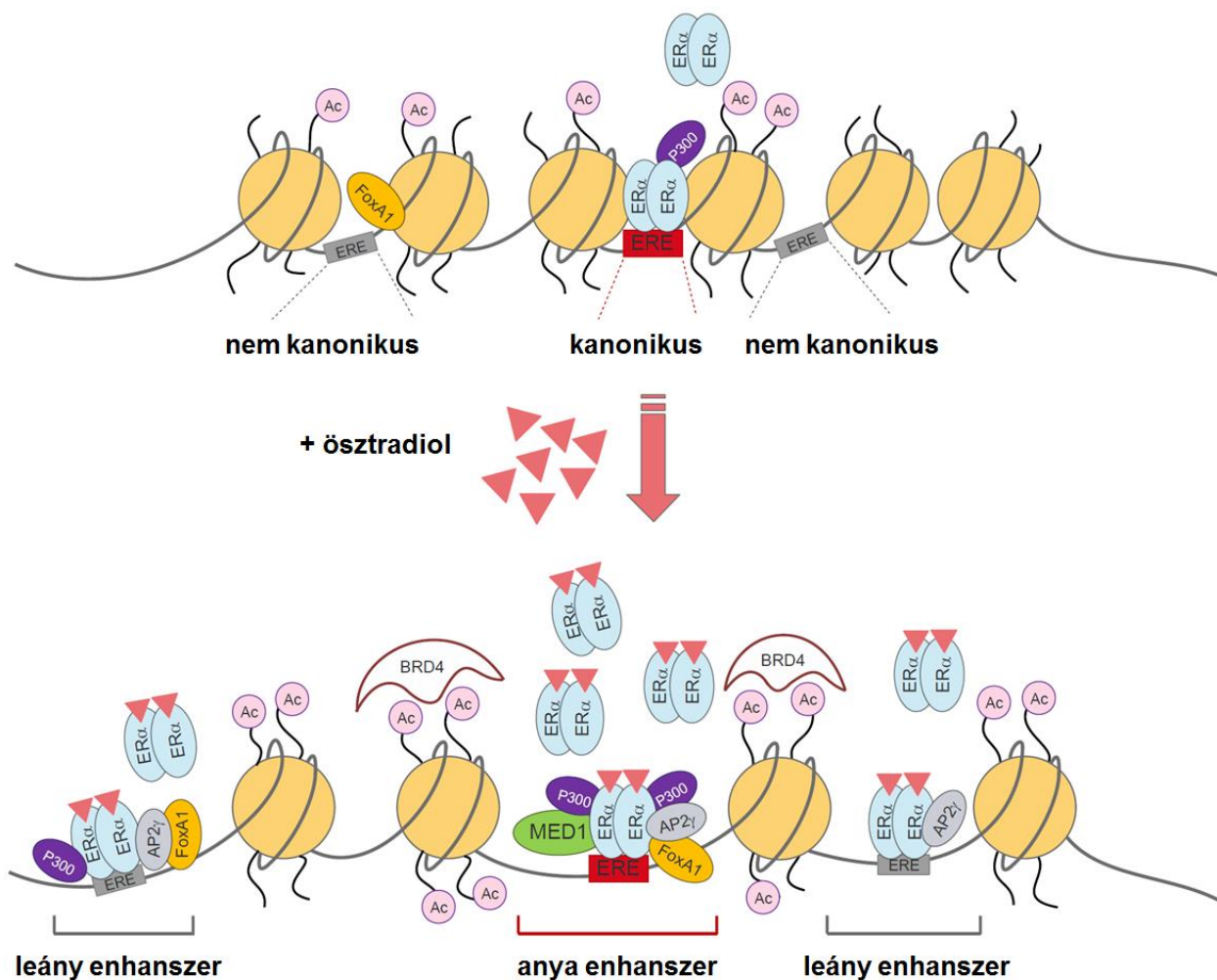
2013-ban *Richard A. Young* és munkacsoportja a különböző sejttypusok specifikációjára alkalmas, ún. sejttypus-specifikus gének közelében olyan szabályozó elemek csoportját azonosította, melyekről bebizonyosodott, hogy hatásuk összeadódik és együtt a szabályozott gének extrém magas expresszióját képesek kiváltani. Ezeket a szabályozó régiókat szuper-enhanszereknek nevezték el (Lovén et al., 2013; Whyte et al., 2013).

Jelen munka alapját egy olyan 2010-ben publikált ER α ChIP-seq minta képezte, amely már 81518 ER α kötőhely azonosítását tette lehetővé az emlő-eredetű MCF-7 sejtvonalon, ösztadiol kezelést követően (Schmidt et al., 2010). A sejtvonalon ER α -vezérelt szuper-enhanszereinek azonosítása során arra lettünk figyelmesek, hogy minden szuper-enhanszerre jellemző egy, legfeljebb kettő ER α kötőhely jelenléte már az ösztadiol kezelést megelőzően is **(16. ábra)**. Feltételeztük, hogy ezek az elsődleges – később anya enhanszereknek elnevezett – régiók alapját képezik a szuper-enhanszerek kialakulásának, ezért célunk volt az anya-, és a másodlagosan megjelenő, ún. leány enhanszerek jellemzése és szerepük megértése az ER α általi transzkripciós szabályozásban.

Az anya enhanszerek nagy affinitású, kanonikus válaszadó elemekkel rendelkeznek; ez jelentheti a kiindulópontját a szuper-enhanszerek kialakulásának – már a ligand kezelést megelőzően **(25. ábra)**. Az ER α kötése szinte kizárólag a saját válaszadó elemén, az ERE-n keresztül jön létre, és a kötés fenntartásához alapvetően nincs szükség ko-faktorok közvetlen jelenlétére sem. Jellemző rájuk a MED1 és különböző ko-regulátorok, mint a P300 és a BRD4 jelenléte, továbbá az aktív hiszton (H3K27ac) jelek feldúsulása **(24. ábra)**. Ligandkezelést követően a ko-regulátorok mennyisége fokozódik, a FoxA1 és AP2 γ fehérjék is megjelennek, de kizárólag másodlagos, stabilizáló szerepet látnak el **(31. ábra)**.

A másodlagosan, ösztadiol stimulációt követően megjelenő leány enhanszerek dinamikájukat tekintve sokkal inkább jellemezhetőek tipikus enhanszerekként. Többnyire kevésbé specifikus, nem kanonikus ösztrogén válaszadó elemén, vagy a magreceptorokra jellemző félhelyen és az irodalomból is jól ismert FoxA1 és AP-1 ko-faktorok motívumán keresztül foglalják el helyüket a DNS-en, kötésük gyakorisága pedig az anya enhanszerekének töredéke, nagyjából tizede **(18. ábra)**. A különbséget a tipikus és leány enhanszerek között mégis az jelentheti, hogy a leány enhanszerek feladatukat más enhanszerekkel együtt, egy funkcionális egységként látják el. Kialakulásukat illetően azt feltételezzük, hogy az anya enhanszerekhez nagy affinitással kötődő, valamint a környező ER α fehérjék sokkal nagyobb valószínűséggel foglalják el az anya enhanszerekhez közeli leány enhanszereket, mint a távolabbi, egyéb kötőhelyeket. A jelenséget 5 különböző egér és humán sejtvonalon, 7 transzkripciós faktor vonatkozásában validáltuk.

A ligand-indukálható ER α -vezérelt szuper-enhanszerek működését a **39. ábra** reprezentálja.



39. ábra. Az ER α -vezérelt szuper-enhanszerek kialakulásának sematikus ábrája. A felső panel a kromatin egy zárt, fehérjék számára kevésbé hozzáférhető formáját reprezentálja, ahol az ER α a DNS-hez csak a specifikus, kanonikus ERE-n keresztül képes kötődni. Az alsó panel az ösztadiol kezelést követő kromatin átrendeződéseket illusztrálja. A kromatin nyitottá válik, az ösztadiol pedig elősegíti számos ER α dimernek a nem kanonikus ERE-hez való kötődését. A hormon-receptor komplex kialakulásán és DNS-hez való kötődésén felül a kialakult szuper-enhanszer régiók acetiláltabbá válnak, a ko-faktorok (főként FoxA1, AP2 γ) és ko-aktivátor fehérjék (P300, MED1, BRD4) pedig megerősítik, stabilizálják a kötődésüket.

Bár jelentek meg tanulmányok, melyek megkérdőjelezték a szuper-enhanszer mint fogalom újdonságát, valamint olyanok, amelyek azt igazolták, hogy az egyes elemek transzkripcióra gyakorolt hatása csak összeadódik, de egymás hatását nem fokozzák, és kooperáció sem figyelhető meg közöttük, ezen állítások ellenkezőjére is van példa (Dukler, Gulko, Huang, & Siepel, 2017; Hay et al., 2016; Moorthy et al., 2017; Rickels & Shilatifard, 2018). *Saravanan* és munkatársai a mi tanulmányunkból kiindulva külön-külön vizsgálták a ligand aktiválható ER α szuper-enhanszerek elemeit. Elsőként az anya enhanszerek működését gátolták a CRISPR-Cas9 módszer segítségével a TFF1 és a GREB1 gének közelében, és azt találták, hogy a leány enhanszereket (melyeket ők „tranzienst kötőhelyeknek” neveztek) az ER α nem fogja elfoglalni az anya enhanszerek (ún. „perzisztens kötőhelyek”) hiányában (Saravanan et al., 2019).

Az a tény, hogy a szuper-enhanszerek elemei egymástól akár több, mint 10 kb-nyi távolságra is elhelyezkedhetnek, csak a hurokképződéssel magyarázható. Alátámasztja mindezt a MED1 fehérjének a jelenléte az anya enhanszereken, amely hidat képez a szabályozott gén promótere és maga az enhanszer között, biztosítva ezáltal az interakciójukat. *Saravanan* és munkatársai ChIA-PET és 5C módszerekkel kimutatták azt is, hogy az anya és leány enhanszerek, valamint a célgének között valóban erős fizikai interakció van, és azt, hogy a leány enhanszerek hiányában az anya enhanszerek önmagukban nem képesek a célgének expresszióját fenntartani (Saravanan et al., 2019). Mindez megerősíti azt a felvetésünket, hogy az anya enhanszerek a szuper-enhanszerek alappilléreiként funkcionálnak, de összességében a leány enhanszerekkel egy egységként működnek.

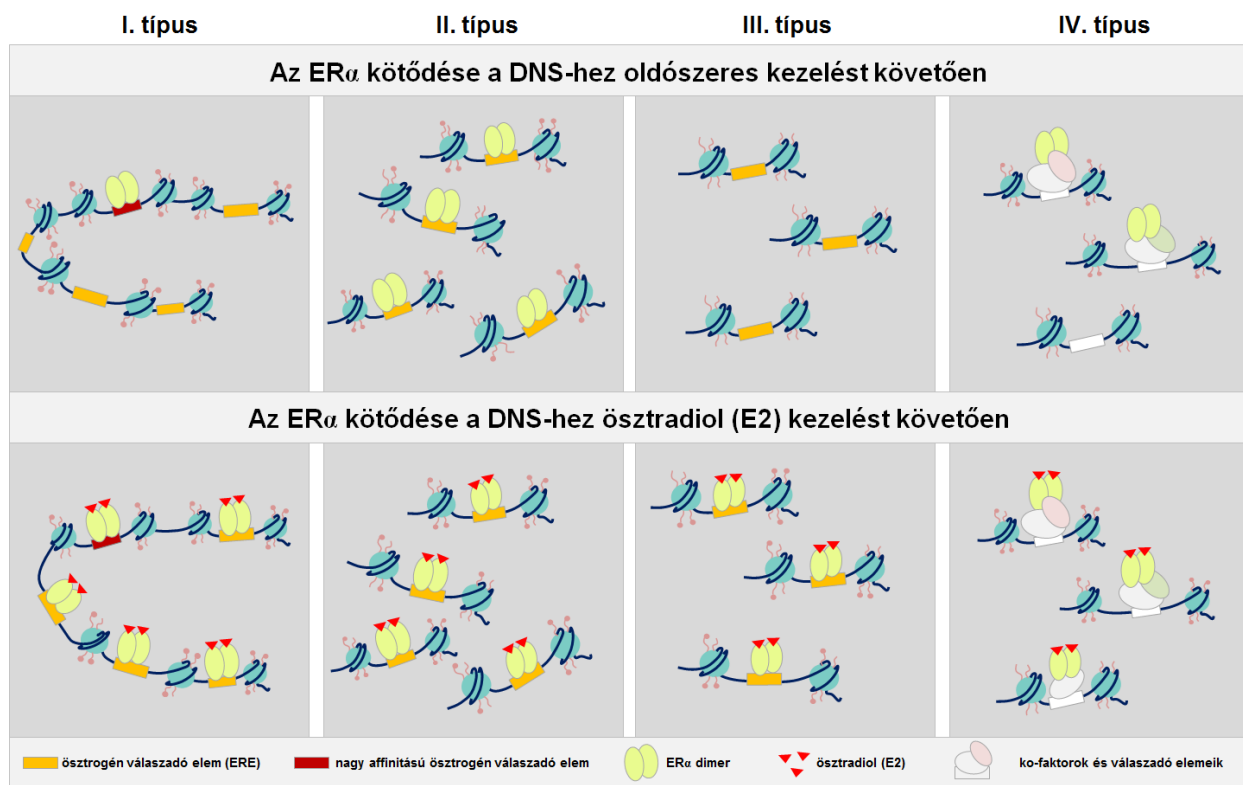
Huang és munkatársai a szuper-enhanszerek egyedi elemei közötti kromatin kölcsönhatásokon alapuló szerkezeti hierarchiát vizsgálva a szuper-enhanszereken belül olyan csomópontokat, ún. „hub” enhanszereket azonosítottak, melyeket nagy mennyiségű CTCF

fehérje jelenléte jellemzett. Azt találták, hogy ezek a „hub” enhanszerek szokatlanul sok távoli interakció kialakításáért felelősek és ezen tulajdonságuk révén képesek a szuper-enhanszerek szerkezetét és működését fenntartani (Huang et al., 2018). Mindez összhangban van az általunk azonosított anya enhanszerek domináns szerepével. Abból kiindulva, hogy ezek a „hub” enhanszerek sokkal nagyobb hatással vannak a génaktivációra, mint a „nem hub” társaik ugyanazon szuper-enhanszeren belül, a *Huang*-ék által felvázolt modell szerint a szuper-enhanszerek szerkezeti hierarchiája előrevetítheti a funkcionális hierarchiájukat is.

Egy másik, az általunk leírt koncepcióval összefüggésbe hozható fogalom és egyben szabályozó egység az ún. „hot spot”. A „hot-spot”-ok a genom olyan egyedi lókuszait jelentik, ahol fehérje-fehérje interakciók révén számos különböző transzkripció faktor jelenléte figyelhető meg (Siersbæk, Baek, et al., 2014). A két szabályozó egység, a szuper-enhanszerek és a „hot-spot”-ok integrálása révén *Siersbæk* és munkatársai leírták, hogy a szuper-enhanszereken belül legalább egy „hot-spot” azonosítható (Siersbæk, Rabiee, et al., 2014). Az ER α -vezérelt szuper-enhanszerek vonatkozásában mi is ehhez hasonló jelenséget figyeltünk meg, attól eltekintve, hogy az anya enhanszerek nem feltétlen mutatták szükségét különféle, akár több tíz eltérő transzkripció faktor jelenlétének.

In vitro kísérletek révén több tanulmányban bizonyították azt is, hogy egy adott sejtre specifikus szuper-enhanszer teljes mértékű eltávolítása vagy akár csak a részleges megváltoztatása képes a sejttípus-specifikus gén(ek) csökkent expresszióját eredményezni, ugyanakkor egy-egy szakasza – mesterséges rendszerbe történő klónozást követően – is magasabb génexpressziót eredményez, mint az adott sejt tipikus enhanszerei (Adam et al., 2015; Huang et al., 2016; Yan Li et al., 2014; Lovén et al., 2013).

A rendelkezésünkre álló 81518 ER α transzkripciós faktor kötőhelyet aszerint rendszerezve, hogy kötődnek-e a DNS-hez már ösztadiol hiányában vagy csak annak hozzáadását követően, továbbá rendelkeznek-e ERE-vel – a szuper-enhanszereken felül – három további típusra különítettünk el (33. ábra). A különböző típusok kötési mechanizmusát a 40. ábra reprezentálja.



40. ábra. Az ER α kötőhelyek négy típusának kötési mechanizmusa oldószerrel és ösztadiollal történő kezelést követően. Az I. típus az ER α -vezérelt szuper-enhanszereket foglalja magába; ösztadiol kezelést követően az ER α a leány enhanszereket is elfoglalja. A II. típusú kötőhelyek az anya enhanszerhez hasonlóan működnek, de nem képezik szuper-enhanszerek részét. A III. típusba sorolt kötőhelyek az II. típusúakhoz hasonlóan rendelkeznek ERE-vel, de a kötésükhöz minden esetben szükséges a ligand általi aktiválás. A IV. típusú kötőhelyek nem rendelkeznek ERE-vel; az ER α kötése más ko-faktorokon keresztül jön létre és tipikusan ösztadiol kezelést követően.

Azon túlmenően, hogy jellemeztük a kötésük dinamikáját és megvizsgáltuk a különböző ko-faktorok jelenlétét, arra voltunk kíváncsiak, van-e különbség a különböző csoportok célgénjeinek expressziója között. Azt találtuk, hogy sem a II. típusba sorolt ER α kötőhelyek, melyek az anya enhanszerekhez hasonlóan már ligand hiányában az ERE-n keresztül képesek voltak kötődni a DNS-hez, sem a III. típus kötőhelyei, melyek szintén kizárólag az ERE-n keresztül, de csak a ligand hatására foglalták el helyüket, sem a IV. típus kötőhelyei, melyek kötését társfaktorok segítették elő, nem voltak képesek olyan erős génexpressziót produkálni, mint a szuper-enhanszerek (38. ábra).

Eredményeink tovább erősítik azt a sokak által vitatott nézetet, hogy a szuper-enhanszerek egy szabályozó egységként működnek, továbbá a tipikus enhanszerekhez képest sokkal markánsabb transzkripciós kimenetet képesek eredményezni, még akkor is, ha azok az anya vagy leány enhanszerekhez hasonló tulajdonságokkal és dinamikával rendelkeznek. Elmondhatjuk, hogy elsőként jellemeztük a ligand-indukálható szuper-enhanszerek működését és eredményeink más kutatócsoportokat is arra ösztönöztek, hogy tovább vizsgálják az ER α -vezérelt szuper-enhanszerek egyedi elemeinek működését, fontosságát és a közöttük megfigyelhető – általunk addig csak feltételezett – interakció meglétét.

6. Összefoglalás

A szuper-enhanszerek a genom olyan szabályozó egységei, ahol az enhanszerek klaszterekbe rendeződnek és hatásukat a célgénekre – a nevükhöz hűen – „szuper fokozókként”, együtt fejtik ki. A fogalmat először 2013-ban említették, azóta több mint 500 olyan publikáció jelent meg, mely a szuper-enhanszerek vizsgálatára irányul. Közöttük számos olyan található, amelyben a különböző betegségekre hajlamosító gének szuper-enhanszerek általi szabályozását azonosították.

Kutatásunk során az ER α -vezérelt szuper-enhanszerek működését vizsgáltuk az emlő eredetű MCF-7 sejtvonalban. Azt találtuk, hogy az ER α -vezérelt szuper-enhanszerek helyét tipikusan már a ligand általi aktivációt megelőzően kijelöli a kanonikus válaszadó eleméhez kötődő ER α . Az ilyen válaszadó elemek olyan specifikusak és olyan nagy mennyiségű fehérjét képesek a genom adott szakaszához toborozni E2 hatására, hogy az egyúttal elősegíti a szomszédos, gyengébb válaszadó elemek ER α általi elfoglalását is – kialakítva így a szuper-enhanszereket.

Az elsődleges, „kijelölő” szereppel rendelkező kötőhelyeket anya enhanszereknek, míg a ligand általi aktivációt követően elfoglalt kötőhelyeket leány enhanszereknek neveztük el. A szuper-enhanszerek kialakulásának ezen mechanizmusát 5 további humán és egér sejtvonalon, 7 ligand által indukálható transzkripció faktor vonatkozásában validáltuk.

A genomban azonosítható ~80000 ER α transzkripció faktor kötőhely csoportosítása és vizsgálata során azt találtuk, hogy az anya enhanszerekkel azonos tulajdonságokkal és dinamikával rendelkező tipikus enhanszerek önmagukban nem képesek ugyanazt a transzkripció kimenetet eredményezni, mint az egységként működő szuper-enhanszerek. Eredményeink nemcsak az anya enhanszerek kivételes szerepét igazolják a szuper-enhanszereken belül, hanem a

fokozó szereppel bíró leány enhanszerek fontosságát is, és indokoltta teszik a szuper-enhanszerek különböző sejtekben történő azonosítását és az általuk szabályozott gének megismerését.

7. Summary

Super-enhancers are regulatory units of the genome where enhancers are clustered and act together and have an extraordinary effect on the target genes. The concept was first mentioned in 2013. Since then, more than 500 studies have been published focusing on super-enhancers. Many of these studies identified the role of super-enhancers in regulating genes predisposing to various diseases.

In this present work, we investigated the function of the ER α -driven super-enhancers in breast cancer-derived MCF-7 cell line and found that super-enhancers are typically marked prior to ligand activation by the presence of a pre-occupied ER α binding and these primary bindings can occur through a canonical response element in all cases. These response elements are so specific and capable of attracting a large amount of ER α molecules to a particular region of the genome. They also facilitate the occupation of the neighboring weaker response elements by ER α molecules rather than a distant genomic region thus forming super-enhancers. We referred to ER α -dense primary, “nucleating” binding sites as ‘mother enhancers’, while the additional, secondary regions appearing upon ligand activation were designated as ‘daughter enhancers’.

The mechanism of super-enhancer formation was validated on 5 additional human and mouse cell lines for 7 further ligand-inducible transcription factors as well. By grouping and examining approximately 80,000 ER α transcription factor binding sites identifiable in the genome, it has been found that typical enhancers with the same properties and dynamics as the mother-enhancers are not capable of inducing the same transcription output; presumably because mother enhancers function as units of super-enhancers.

Our results confirm not only the exceptional role of mother enhancers but also the importance of daughter enhancers. Altogether, these justify the identification of super-enhancers in different cells and understanding of the genes that they are regulated.

8. Kulcsszavak

ösztrogén receptor alfa, transzkripció szabályozás, transzkripció faktor, enhanszer, szuper-enhanszer, MCF-7 sejtvonal, ChIP-seq, genomika, bioinformatika

8. Keywords

estrogen receptor alpha, transcriptional regulation, transcription factor, enhancer, super-enhancer, MCF-7 cell line, ChIP-seq, genomics, bioinformatics

9. Rövidítések jegyzéke

ABAT	4-Aminobutyrate Aminotransferase
AF-1, AF-2	activation function 1/2 (AF-1 és AF-2 transzaktivációs domének)
AICE	AP-1/IRF composite element
AP-1	Activator protein 1
AP2γ	Activator Protein 2 gamma
AR	Androgen Receptor
ARE	Androgen Response Element (androgén válaszadó elem)
ATRA	csupa-transz retinsav
BAM	Binary Alignment Map
bp	bázispár
BRD4	Bromodomain-containing protein 4
CA12	Carbonic Anhydrase 12
CCND1	Cyclin D1
C/EBP	CCAAT/enhancer-binding protein
ChIA-PET	Chromatin Interaction Analysis by Paired-End Tag Sequencing
ChIP-seq	Chromatin Immunoprecipitation followed by sequencing (újgenerációs szekvenálással kapcsolt kromatin immunprecipitáció)
CRE	cis-regulatory element (<i>cisz</i> szabályozó elem)
CRE	cAMP Response Element
CRM	cis-regulatory module (<i>cisz</i> szabályozó modul)
CTCF	CCTC-binding factor
CTD	C-terminal domain (C-terminális domén)
CUEDC1	CUE Domain Containing 1
CUX1	Cut Like Homeobox 1
D3	kalcitriol
DBD	DNA-binding domain (DNS-kötő domén)
DHT	dihidrotesztoszteron
DMSO	dimetil-szulfoxid
DNáz I	Dezoxiribonukleáz I (enzim)

DNS	deoxiribonukleinsav
DR	direct repeat (direkt módon ismétlődő)
E1	ösztrol
E2	ösztradiol
E3	ösztriol
ER	ösztrogén receptor
ERE	Estrogen Response Element (ösztrogén válaszadó elem)
eRNA	enhanszer RNS
ERα	estrogen receptor alpha (ösztrogén receptor alfa)
ERβ	estrogen receptor beta (ösztrogén receptor béta)
ESR1	Estrogen Receptor 1
EtOH	etanol
FoxA1	Forkhead box A1
FPKM	Fragments Per Kilobase of transcripts per Million mapped reads (fragmentek száma 1000 bázisonként, egy millió térképezett leolvasásra vetítve)
GATA3	GATA-binding protein 3
GEO	Gene Expression Omnibus (adatbázis)
GR	glükokortikoid receptor
H3K27ac	Histone H3 lisine 27 acetylation
H3K27me3	Histone H3 lisine 27 trimethylation
H3K4me1/2	Histone H3 lisine 2 mono/dimethylation
H3K9me3	Histone H3 lisine 9 trimethylation
HDAC2	histone deacetylase 2
HGP	Humán Genom Projekt
H-híd; H-kötés	hidrogénhíd-kötés
HOMER	Hypergeometric Optimization of Motif EnRichment
hsp90	heat shock protein 90 (hő sokk fehérje)
IGV	Integrative Genomics Viewer
IRF	Interferon Regulatory Factor
JUNB	JunB Proto-Oncogene
kbp	kilobázispár

kDa	kilodalton
KLF	Krüppel-like factor
KRT19	Keratin 19
LBD	Ligand-binding domain (ligand-kötő domén)
LCR	locus control region (lókuszt kontroll régió)
LPS	lipopoliszaharid
MACS2	Model-based Analysis of ChIP-Seq 2
MARE	MAF recognition element
Mbp	megabázispár
MCF-7	Michigan Cancer Foundation-7 (sejtvonal)
MDa	megadalton
MED1	Mediator complex subunit 1 (ko-faktor)
mRNS	messenger RNS (hírvivő RNS)
MYOF	Myoferlin
NES	nuclear export signal (magi export szignál)
NFκB	Nuclear Factor Kappa-light-chain-enhancer of activated B cells
NLS	nuclear localization signal (magi lokalizációs szignál)
NOS1AP	Nitric Oxide Synthase 1 Adaptor Protein
NR	nuclear receptor (magreceptor)
NTR, NTD	N-terminal region/domain (N-terminális régió/domén)
P300	E1A protein 300 kDa
PTM	post-translational modification (poszt-transzlációs módosítás)
PU.1	Purine-rich box-1
RARα	Retinoic acid receptor alpha
RARγ	Retinoic acid receptor gamma
RNA-seq	RNA-sequencing (RNS szekvenálás)
RNS	ribonukleinsav
RPKM	Reads Per Kilober per Million mapped reads (leolvasások száma 1000 bázisonként, egy millió térképezett leolvasásra vetítve)
SE	super-enhancer (szuper-enhanszer)
SIN3A	SIN3 transcription regulator family member A

siRNA	small interfering RNA (kis interferáló RNS)
Sp-1	Specificity protein 1
SRA	Sequence Read Archive (adatbázis)
STAT1	Signal transducer and activator of transcription 1
TAD	Topologically Associated Domain (topológiaiilag asszociált domén)
TCF1	T cell factor 1
TF	transzkripció faktor
TMEM120B	Transmembrane Protein 120B
TSS	transcription start site (transzkripció start hely)
VDR	Vitamin D Receptor
WWC1	WW And C2 Domain Containing 1
ZFH3	Zinc Finger Homeobox 3

10. Függelék

10.1. A magreceptor szupercsalád tagjai és ligandjaik

Osztály	Név	Rövidítés	Általános név	Ligand(ok)
I. Szteroid receptorok	NR3A1	ER α	Estrogen receptor alpha	ösztrogének
	NR3A2	ER β	Estrogen receptor beta	
	NR3C1	GR	Glucocorticoid receptor	glükokortikoidok
	NR3C2	MR	Mineralocorticoid receptor	mineralokortikoidok glükokortikoidok
	NR3C3	PR	Progesterone receptor	progeszteron
	NR3C4	AR	Androgen receptor	androgének

II. RXR heterodimerek	NR1A1	TR α	Thyroid hormone receptor alpha	pajzsmirigyhormonok
	NR1A2	TR β	Thyroid hormone receptor beta	
	NR1B1	RAR α	Retinoic acid receptor alpha	csupa-transz retinsav
	NR1B2	RAR β	Retinoic acid receptor beta	9-cisz retinsav
	NR1B3	RAR γ	Retinoic acid receptor gamma	csupa-transz retinsav
	NR1C1	PPAR α	Peroxisome proliferator-activated receptor alpha	oxidált koleszterin származékok (oxiszterolok)
	NR1C2	PPAR β/δ	Peroxisome proliferator-activated receptor beta/delta	
	NR1C3	PPAR γ	Peroxisome proliferator-activated receptor gamma	
	NR1H2	LXR β	Liver X receptor beta	oxiszterolok
	NR1H3	LXR α	Liver X receptor alpha	
	NR1H4	FXR α	Farnesoid X receptor alpha	epesavak
	NR1H5	FXR β	Farnesoid X receptor beta	lanoszterol
	NR1I1	VDR	Vitamin D receptor	12,5-(OH) $_2$ -VD
				litokolsav
	NR1I2	PXR	Pregnane X receptor	endo- és xenobiotikumok
	NR1H3	CAR	Constitutive androstane receptor	androsztán

III. Dimerizáló árva receptorok	NR1D1	REV-ERB α	Reverse-ERB alpha	hem
	NR1D2	REV-ERB β	Reverse-ERB beta	
	NR2A1	HNF4 α	Hepatocyte nuclear factor 4 alpha	zsírsavak
	NR2A2	HNF4 γ	Hepatocyte nuclear factor 4 gamma	
	NR2B1	RXR α	Retinoid X receptor alpha	9-cisz retinsav
	NR2B2	RXR β	Retinoid X receptor beta	dokozahexaénsav
	NR2B3	RXR γ	Retinoid X receptor gamma	<i>nem ismert</i>
	NR2C1	T2	Testicular orphan receptor 2	<i>nem ismert</i>
	NR2C2	T4	Testicular orphan receptor 4	
	NR2F1	COUP-TF α	Chicken ovalbumin upstream promoter-transcription factor alpha	<i>nem ismert</i>
	NR2F2	COUP-TF β	Chicken ovalbumin upstream promoter-transcription factor beta	
	NR2F6	COUP-TF γ	Chicken ovalbumin upstream promoter-transcription factor gamma	
	NR6A1	GCNF	Germ cell nuclear factor	<i>nem ismert</i>

IV. Monomer árva receptorok	NR0B1	DAX1	Dosage-sensitive sex reversal-adrenal hypoplasia congenital critical region on the X chromosome, gene 1	<i>nem ismert</i>
	NR0B2	SHP	Short heterodimeric partner	<i>nem ismert</i>
	NR1F1	ROR α	RAR-related orphan receptor alpha	szterolok
	NR1F2	ROR β	RAR-related orphan receptor beta	
	NR1F3	ROR γ	RAR-related orphan receptor gamma	
	NR2E3	TLX	Talless homolog orphan receptor	<i>nem ismert</i>
	NR3B1	PNR	Photoreceptor-cell-specific nuclear receptor	<i>nem ismert</i>
	NR3B2	ERR α	Estrogen related receptor alpha	<i>nem ismert</i>
	NR3B3	ERR β	Estrogen related receptor beta	
	NR4A1	ERR γ	Estrogen related receptor gamma	
	NR4A2	NGFI-B	Nerve-growth-factor-induced gene	<i>nem ismert</i>
	NR4A3	NOR1	Neuron-derived orphan receptor 1	<i>nem ismert</i>
	NR5A1	SF-1	Steroidogenic factor 1	foszfolipidek
	NR5A2	LRH1	Liver receptor homolog-1	foszfolipidek

10.2. Az értekezés alapjául szolgáló közlemények



**DEBRECENI
EGYETEM**

**DEBRECENI EGYETEM
EGYETEMI ÉS NEMZETI KÖNYVTÁR**

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

Nyilvántartási szám: DEENK/296/2019.PL
Tárgy: PhD Publikációs Lista

Jelölt: Bojcsuk Dóra
Neptun kód: BD1BZT
Doktori Iskola: Molekuláris Sejt- és Immunbiológia Doktori Iskola

A PhD értekezés alapjául szolgáló közlemények

1. **Bojcsuk, D.**, Bálint, B. L.: Classification of different types of estrogen receptor alpha binding sites in MCF-7 cells.
J. Biotechnol. 299, 13-20, 2019.
DOI: <http://dx.doi.org/10.1016/j.jbiotec.2019.04.016>
IF: 3.163 (2018)
2. **Bojcsuk, D.**, Nagy, G., Bálint, B. L.: Inducible super-enhancers are organized based on canonical signal-specific transcription factor binding elements.
Nucleic Acids Res. 45 (7), 3693-3706, 2017.
DOI: <http://dx.doi.org/10.1093/nar/gkw1283>
IF: 11.561

További közlemények

3. Divoux, A., Sándor, K., **Bojcsuk, D.**, Talukder, A., Li, X., Bálint, B. L., Osborne, T. F., Smith, S. R.: Differential open chromatin profile and transcriptomic signature define depot-specific human subcutaneous preadipocytes: primary outcomes.
Clin Epigenet. 10 (1), 1-15, 2018.
DOI: <http://dx.doi.org/10.1186/s13148-018-0582-0>
IF: 5.174
4. Márton, J., Péter, M., Balogh, G., Bódi, B., Vida, A., Szántó, M., **Bojcsuk, D.**, Jankó, L., Bhattoa, H. P., Gombos, I., Uray, K., Horváth, I., Török, Z., Bálint, B. L., Papp, Z., Vigh, L., Bai, P.: Poly(ADP-ribose) polymerase-2 is a lipid-modulated modulator of muscular lipid homeostasis.
Biochim. Biophys. Acta. Mol. Cell Biol. Lipids. 1863 (11), 1399-1412, 2018.
DOI: <http://dx.doi.org/10.1016/j.bbalip.2018.07.013>
IF: 4.402





5. Vető, B., **Bojcsuk, D.**, Bacquet, C., Kiss, J., Sipeki, S., Martin, L., Buday, L., Bálint, B. L., Arányi, T.: The transcriptional activity of hepatocyte nuclear factor 4 alpha is inhibited via phosphorylation by ERK1/2.
PLoS One. 12 (2), 1-19, 2017.
DOI: <http://dx.doi.org/10.1371/journal.pone.0172020>
IF: 2.766
6. Ozgyin, L., Erdős, E., **Bojcsuk, D.**, Bálint, B. L.: Nuclear receptors in transgenerational epigenetic inheritance.
Prog. Biophys. Mol. Biol. 118 (1-2), 34-43, 2015.
DOI: <http://dx.doi.org/10.1016/j.pbiomolbio.2015.02.012>
IF: 2.581
7. **Bojcsuk, D.**, Erdős, E., Bálint, B. L.: Az ösztrogénreceptor működése a legújabb genomikai kutatások tükrében.
LAM KID. 4 (2), 79-84, 2014.
8. **Bojcsuk, D.**, Sipos, L., Bálint, B. L.: A mikro-RNS-ek mint a hormonok egy új családja.
LAM KID. 3 (4), 29-33, 2013.

A közlő folyóiratok összesített impakt faktora: 29,647

**A közlő folyóiratok összesített impakt faktora (az értekezés alapján szolgáló közleményekre):
14,724**

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

Debrecen, 2019.08.22.



11. Köszönetnyilvánítás

Mindenekelőtt szeretném hálámat és köszönetemet kifejezni témavezetőmnek, *Dr. Bálint Bálint Lászlónak*, aki MSc tanulmányaim során arra ösztönzött, hogy a munkacsoportja tagjaként a számomra addig teljesen ismeretlen bioinformatikai elemző módszereket megismerhessem. Köszönöm a PhD tanulmányaim során nyújtott folyamatos segítségét, támogatását és értékes szakmai tanácsait.

Köszönettel tartozom *Dr. Horváth Attilának* és *Czipa Eriknek*, hogy az első hónapokban türelemmel és erő felett segítették munkámat; nélkülük talán nem sajátíthattam volna el a bioinformatika alapjait. Köszönettel tartozom *Erdős Edinának*, *Ozgyin Lillának* és *Noura Farajnak*, akik a munkacsoport tagjaiként nem csupán szakmai támogatást nyújtottak számomra, hanem barátként a mindennapokban is mellettem álltak.

Köszönöm a Molekuláris Sejt- és Immunbiológiai Doktori Iskola jelenlegi és korábbi vezetőinek, *Prof. Dr. Tózsér Józsefnek* és *Prof. Dr. Fésüs Lászlónak*, hogy PhD tanulmányaimat a doktori iskola keretein belül, egy nagyon jó légkörben és nagy szakmai tudással rendelkező kutatók között végezhettem. Köszönettel tartozom a kollaborátoroknak, *Dr. Dánielné Sándor Katalinnak*, *Dr. Adeline Divoux-nak*, *Dr. Timothy F. Osborne-nak* és *Dr. Steven R. Smith-nek*, hogy a közös munka során egyúttal tanulási lehetőséget is biztosítottak számomra. Külön köszönettel tartozom *Szántó-Kiss Líviának* és *Elek Nórának* mérhetetlen türelmükért és segítségükért – nélkülük számos szakmai rendezvény adta lehetőségről kellett volna lemondjak.

Legnagyobb hálával azonban a családomnak, *Szüleimnek* és *Gergőnek* tartozom. Köszönöm a szüleim pótolhatatlan szeretetét, Gergőnek pedig köszönöm, hogy kritikus hozzáállásával mindig a jobbra ösztönzött, emellett barátként és társként is mindvégig mellettem állt.

A disszertációhoz kapcsolódó kutatások az Emberi Erőforrások Minisztériuma ÚNKP-17-3-DE-140 és az ÚNKP-18-3-III-DE-253 Új Nemzeti Kiválóság Programjának, továbbá az EFOP-3.6.3-VEKOP-16-2017-00009 ösztöndíjnak a támogatásával készültek.

12. Hivatkozások

- Aagaard, M. M., Siersbæk, R., & Mandrup, S. (2011). Molecular basis for gene-specific transactivation by nuclear receptors. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1812(8), 824–835. <http://doi.org/10.1016/J.BBADIS.2010.12.018>
- Abderrahman, B., & Jordan, V. C. (2018). Steroid Receptors in Breast Cancer. In *The Breast* (p. 272–281.e2). Elsevier. <http://doi.org/10.1016/B978-0-323-35955-9.00021-0>
- Adam, R. C., Yang, H., Rockowitz, S., Larsen, S. B., Nikolova, M., Oristian, D. S., ... Fuchs, E. (2015). Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. *Nature*, 521(7552), 366–370. <http://doi.org/10.1038/nature14289>
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). *Molecular biology of the cell*. Garland Science.
- Al-Kuraya, K., Schraml, P., Thorhorst, J., Tapia, C., Zaharieva, B., Novotny, H., ... Sauter, G. (2004). Prognostic Relevance of Gene Amplifications and Coamplifications in Breast Cancer. *Cancer Research*, 64(23), 8534–8540. <http://doi.org/10.1158/0008-5472.CAN-04-1945>
- Arzate-Mejía, R. G., Recillas-Targa, F., & Corces, V. G. (2018). Developing in 3D: the role of CTCF in cell differentiation. *Development*, 145(6), dev137729. <http://doi.org/10.1242/DEV.137729>
- Badve, S., Turbin, D., Thorat, M. A., Morimiya, A., Nielsen, T. O., Perou, C. M., ... Nakshatri, H. (2007). FOXA1 Expression in Breast Cancer Correlation with Luminal Subtype A and Survival. *Clinical Cancer Research*, 13(15), 4415–4421. <http://doi.org/10.1158/1078-0432.CCR-07-0122>
- Banerji, J., Rusconi, S., & Schaffner, W. (1981). Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell*, 27(2 Pt 1), 299–308. [http://doi.org/10.1016/0092-8674\(81\)90413-x](http://doi.org/10.1016/0092-8674(81)90413-x)
- Beatson, G. T. (1896). On the Treatment of Inoperable Cases of Carcinoma of the Mamma: Suggestions for a New Method of Treatment, with Illustrative Cases. *Transactions. Medico-Chirurgical Society of Edinburgh*, 15, 153–179. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/29584099>
- Bell, A. C., West, A. G., & Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, 98(3), 387–96. [http://doi.org/10.1016/s0092-8674\(00\)81967-4](http://doi.org/10.1016/s0092-8674(00)81967-4)
- Bian, Q., & Belmont, A. S. (2012). Revisiting higher-order and large-scale chromatin organization. *Current Opinion in Cell Biology*, 24(3), 359–366. <http://doi.org/10.1016/j.ceb.2012.03.003>
- Bojcsuk, D., Nagy, G., & Balint, B. L. (2017). Inducible super-enhancers are organized based on canonical signal-specific transcription factor binding elements. *Nucleic Acids Research*, 45(7), 3693–3706. <http://doi.org/10.1093/nar/gkw1283>
- Brown, R., & Feder, M. E. (2005). Reverse transcriptional profiling: non-correspondence of transcript level variation and proximal promoter polymorphism. *BMC Genomics*, 6(1), 110. <http://doi.org/10.1186/1471-2164-6-110>
- Brunelle, M., Nordell Markovits, A., Rodrigue, S., Lupien, M., Jacques, P.-É., & Gévry, N. (2015). The histone variant H2A.Z is an important regulator of enhancer activity. *Nucleic Acids Research*,

43(20), gkv825. <http://doi.org/10.1093/nar/gkv825>

- Candelaria, N. R., Liu, K., & Lin, C.-Y. (2013). Estrogen receptor alpha: Molecular mechanisms and emerging insights. *Journal of Cellular Biochemistry*, 114(10), 2203–2208. <http://doi.org/10.1002/jcb.24584>
- Carroll, J. S., Liu, X. S., Brodsky, A. S., Li, W., Meyer, C. A., Szary, A. J., ... Brown, M. (2005). Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell*, 122(1), 33–43. <http://doi.org/10.1016/j.cell.2005.05.008>
- Carroll, J. S., Meyer, C. A., Song, J., Li, W., Geistlinger, T. R., Eeckhoute, J., ... Brown, M. (2006). Genome-wide analysis of estrogen receptor binding sites. *Nature Genetics*, 38(11), 1289–1297. <http://doi.org/10.1038/ng1901>
- Chatagnon, A., Veber, P., Morin, V., Bedo, J., Triqueneaux, G., Sémon, M., ... Benoit, G. (2015). RAR/RXR binding dynamics distinguish pluripotency from differentiation associated cis-regulatory elements. *Nucleic Acids Research*, 43(10), 4833–54. <http://doi.org/10.1093/nar/gkv370>
- Chipumuro, E., Marco, E., Christensen, C. L., Kwiatkowski, N., Zhang, T., Hatheway, C. M., ... George, R. E. (2014). CDK7 Inhibition Suppresses Super-Enhancer-Linked Oncogenic Transcription in MYCN-Driven Cancer. *Cell*, 159(5), 1126–1139. <http://doi.org/10.1016/j.cell.2014.10.024>
- Cirillo, L. A., Lin, F. R., Cuesta, I., Friedman, D., Jarnik, M., & Zaret, K. S. (2002). Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4. *Molecular Cell*, 9(2), 279–289. [http://doi.org/10.1016/S1097-2765\(02\)00459-8](http://doi.org/10.1016/S1097-2765(02)00459-8)
- Cirillo, L. A., McPherson, C. E., Bossard, P., Stevens, K., Cherian, S., Shim, E. Y., ... Zaret, K. S. (1998). Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *The EMBO Journal*, 17(1), 244–54. <http://doi.org/10.1093/emboj/17.1.244>
- Clark, K. L., Halay, E. D., Lai, E., & Burley, S. K. (1993). Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*, 364(6436), 412–420. <http://doi.org/10.1038/364412a0>
- Cui, J., Shen, Y., & Li, R. (2013). Estrogen synthesis and signaling pathways during aging: from periphery to brain. *Trends in Molecular Medicine*, 19(3), 197–209. <http://doi.org/10.1016/j.molmed.2012.12.007>
- Dauvois, S., White, R., & Parker, M. G. (1993). The antiestrogen ICI 182780 disrupts estrogen receptor nucleocytoplasmic shuttling. *Journal of Cell Science*, 1377–88. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8126115>
- Davies, E., & Hiscox, S. (2011). New therapeutic approaches in breast cancer. *Maturitas*, 68(2), 121–8. <http://doi.org/10.1016/j.maturitas.2010.10.012>
- Di Micco, R., Fontanals-Cirera, B., Low, V., Ntziachristos, P., Yuen, S. K., Lovell, C. D., ... Hernando, E. (2014). Control of embryonic stem cell identity by BRD4-dependent transcriptional elongation of super-enhancer-associated pluripotency genes. *Cell Reports*, 9(1), 234–47. <http://doi.org/10.1016/j.celrep.2014.08.055>
- Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., ... Ren, B. (2012). Topological domains in

- mammalian genomes identified by analysis of chromatin interactions. *Nature*, 485(7398), 376–380. <http://doi.org/10.1038/nature11082>
- Doisy, E. A., Veler, C. D., & Thayer, S. (1929). Folliculin from urine of pregnant women. *Am. J. Physiol.*, 90(2), 329–330. Retrieved from <https://www.popline.org/node/509158>
- Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R. R., & Richmond, T. J. (2004). Nucleosome Arrays Reveal the Two-Start Organization of the Chromatin Fiber. *Science*, 306(5701), 1571–1573. <http://doi.org/10.1126/science.1103124>
- Dukler, N., Gulko, B., Huang, Y.-F., & Siepel, A. (2017). Is a super-enhancer greater than the sum of its parts? *Nature Genetics*, 49(1), 2–3. <http://doi.org/10.1038/ng.3759>
- Dunham, I., Kundaje, A., Aldred, S. F., Collins, P. J., Davis, C. A., Doyle, F., ... Birney, E. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), 57–74. <http://doi.org/10.1038/nature11247>
- Dupont, C., Armant, D., & Brenner, C. (2009). Epigenetics: Definition, Mechanisms and Clinical Perspective. *Seminars in Reproductive Medicine*, 27(5), 351–357. <http://doi.org/10.1055/s-0029-1237423>
- Ehret, C. F., & De Haller, G. (1963). Origin, development, and maturation of organelles and organelle systems of the cell surface in *Paramecium*. *Journal of Ultrastructure Research*, 9, 1–42. [http://doi.org/10.1016/S0022-5320\(63\)80088-X](http://doi.org/10.1016/S0022-5320(63)80088-X)
- El Ansari, R., Craze, M. L., Miligy, I., Diez-Rodriguez, M., Nolan, C. C., Ellis, I. O., ... Green, A. R. (2018). The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in luminal B tumours. *Breast Cancer Research*, 20(1), 21. <http://doi.org/10.1186/s13058-018-0946-6>
- ENCODE Project Consortium. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature*, 489(7414), 57–74. <http://doi.org/10.1038/nature11247>
- Enmark, E., Peltö-Huikko, M., Grandien, K., Lagercrantz, S., Lagercrantz, J., Fried, G., ... Gustafsson, J.-Å. (1997). Human Estrogen Receptor β -Gene Structure, Chromosomal Localization, and Expression Pattern¹. *The Journal of Clinical Endocrinology & Metabolism*, 82(12), 4258–4265. <http://doi.org/10.1210/jcem.82.12.4470>
- Evans, R. (1988). The steroid and thyroid hormone receptor superfamily. *Science*, 240(4854), 889–895. <http://doi.org/10.1126/science.3283939>
- Felsenfeld, G., & Groudine, M. (2003). Controlling the double helix. *Nature*, 421(6921), 448–453. <http://doi.org/10.1038/nature01411>
- Feuerborn, A., & Cook, P. R. (2015). Why the activity of a gene depends on its neighbors. *Trends in Genetics*, 31(9), 483–490. <http://doi.org/10.1016/J.TIG.2015.07.001>
- Flach, K. D., & Zwart, W. (2016). The first decade of estrogen receptor cistromics in breast cancer. *Journal of Endocrinology*, 229(2), R43–R56. <http://doi.org/10.1530/JOE-16-0003>
- Fortin, J.-P., & Hansen, K. D. (2015). Reconstructing A/B compartments as revealed by Hi-C using long-

- range correlations in epigenetic data. *Genome Biology*, 16(1), 180. <http://doi.org/10.1186/s13059-015-0741-y>
- Franco, H. L., Nagari, A., & Kraus, W. L. (2015). TNF α signaling exposes latent estrogen receptor binding sites to alter the breast cancer cell transcriptome. *Molecular Cell*, 58(1), 21–34. <http://doi.org/10.1016/j.molcel.2015.02.001>
- Fullwood, M. J., Liu, M. H., Pan, Y. F., Liu, J., Xu, H., Mohamed, Y. Bin, ... Ruan, Y. (2009). An oestrogen-receptor- α -bound human chromatin interactome. *Nature*, 462(7269), 58–64. <http://doi.org/10.1038/nature08497>
- Fulton, D. L., Sundararajan, S., Badis, G., Hughes, T. R., Wasserman, W. W., Roach, J. C., & Sladek, R. (2009). TFCat: the curated catalog of mouse and human transcription factors. *Genome Biology*, 10(3), R29. <http://doi.org/10.1186/gb-2009-10-3-r29>
- Gaszner, M., & Felsenfeld, G. (2006). Insulators: exploiting transcriptional and epigenetic mechanisms. *Nature Reviews Genetics*, 7(9), 703–713. <http://doi.org/10.1038/nrg1925>
- Gerhart-Hines, Z., Feng, D., Emmett, M. J., Everett, L. J., Loro, E., Briggs, E. R., ... Lazar, M. A. (2013). The nuclear receptor Rev-erb α controls circadian thermogenic plasticity. *Nature*, 503(7476), 410–413. <http://doi.org/10.1038/nature12642>
- Gertz, J., Savic, D., Varley, K. E., Partridge, E. C., Safi, A., Jain, P., ... Myers, R. M. (2013). Distinct Properties of Cell-Type-Specific and Shared Transcription Factor Binding Sites. *Molecular Cell*, 52(1), 25–36. <http://doi.org/10.1016/j.molcel.2013.08.037>
- Gibcus, J. H., Samejima, K., Goloborodko, A., Samejima, I., Naumova, N., Nuebler, J., ... Dekker, J. (2018). A pathway for mitotic chromosome formation. *Science*, 359(6376), eaao6135. <http://doi.org/10.1126/science.aao6135>
- Giguère, V., Hollenberg, S. M., Rosenfeld, M. G., & Evans, R. M. (1986). Functional domains of the human glucocorticoid receptor. *Cell*, 46(5), 645–52. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3742595>
- Glass, C. K., & Ogawa, S. (2006). Combinatorial roles of nuclear receptors in inflammation and immunity. *Nature Reviews Immunology*, 6(1), 44–55. <http://doi.org/10.1038/nri1748>
- Gong, Y., Lazaris, C., Sakellaropoulos, T., Lozano, A., Kambadur, P., Ntziachristos, P., ... Tsirigos, A. (2018). Stratification of TAD boundaries reveals preferential insulation of super-enhancers by strong boundaries. *Nature Communications*, 9(1), 542. <http://doi.org/10.1038/s41467-018-03017-1>
- Gorski, J., Toft, D., Shyamala, G., Smith, D., & Notides, A. (1968). Hormone receptors: studies on the interaction of estrogen with the uterus. *Recent Progress in Hormone Research*, 24, 45–80. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/4885833>
- Green, S., Walter, P., Kumar, V., Krust, A., Bornert, J.-M., Argos, P., & Chambon, P. (1986). Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A. *Nature*, 320(6058), 134–139. <http://doi.org/10.1038/320134a0>
- Guertin, M. J., Zhang, X., Coonrod, S. A., & Hager, G. L. (2014). Transient estrogen receptor binding and p300 redistribution support a squelching mechanism for estradiol-repressed genes. *Molecular*

Endocrinology (Baltimore, Md.), 28(9), 1522–33. <http://doi.org/10.1210/me.2014-1130>

- Gururaj, A. E., Peng, S., Vadlamudi, R. K., & Kumar, R. (2007). Estrogen Induces Expression of BCAS3, a Novel Estrogen Receptor- α Coactivator, through Proline-, Glutamic Acid-, and Leucine-Rich Protein-1 (PELP1). *Molecular Endocrinology*, 21(8), 1847–1860. <http://doi.org/10.1210/me.2006-0514>
- Gustafsson, J.-Å. (2003). What pharmacologists can learn from recent advances in estrogen signalling. *Trends in Pharmacological Sciences*, 24(9), 479–485. [http://doi.org/10.1016/S0165-6147\(03\)00229-3](http://doi.org/10.1016/S0165-6147(03)00229-3)
- Harris, H. A. (2007). Estrogen Receptor- β : Recent Lessons from in Vivo Studies. *Molecular Endocrinology*, 21(1), 1–13. <http://doi.org/10.1210/me.2005-0459>
- Hay, D., Hughes, J. R., Babbs, C., Davies, J. O. J., Graham, B. J., Hanssen, L. L. P., ... Higgs, D. R. (2016). Genetic dissection of the α -globin super-enhancer in vivo. *Nature Genetics*, 48(8), 895–903. <http://doi.org/10.1038/ng.3605>
- He, H. H., Meyer, C. A., Chen, M. W., Jordan, V. C., Brown, M., & Liu, X. S. (2012). Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. *Genome Research*, 22(6), 1015–1025. <http://doi.org/10.1101/gr.133280.111>
- Heintzman, N. D., Hon, G. C., Hawkins, R. D., Kheradpour, P., Stark, A., Harp, L. F., ... Ren, B. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, 459(7243), 108–112. <http://doi.org/10.1038/nature07829>
- Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., ... Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell*, 38(4), 576–89. <http://doi.org/10.1016/j.molcel.2010.05.004>
- Heldring, N., Pike, A., Andersson, S., Matthews, J., Cheng, G., Hartman, J., ... Gustafsson, J.-Å. (2007). Estrogen Receptors: How Do They Signal and What Are Their Targets. *Physiological Reviews*, 87(3), 905–931. <http://doi.org/10.1152/physrev.00026.2006>
- Hnisz, D., Abraham, B. J., Lee, T. I., Lau, A., Saint-André, V., Sigova, A. A., ... Young, R. A. (2013). Super-Enhancers in the Control of Cell Identity and Disease. *Cell*, 155(4), 934–947. <http://doi.org/10.1016/J.CELL.2013.09.053>
- Honkela, A., Peltonen, J., Topa, H., Charapitsa, I., Matarese, F., Grote, K., ... Rattray, M. (2015). Genome-wide modeling of transcription kinetics reveals patterns of RNA production delays. *Proceedings of the National Academy of Sciences*, 112(42), 13115–13120. <http://doi.org/10.1073/pnas.1420404112>
- Horwitz, K. B., Costlow, M. E., & McGuire, W. L. (1975). MCF-7; a human breast cancer cell line with estrogen, androgen, progesterone, and glucocorticoid receptors. *Steroids*, 26(6), 785–95. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/175527>
- Huang, J., Li, K., Cai, W., Liu, X., Zhang, Y., Orkin, S. H., ... Yuan, G.-C. (2018). Dissecting super-enhancer hierarchy based on chromatin interactions. *Nature Communications*, 9(1), 943. <http://doi.org/10.1038/s41467-018-03279-9>

- Huang, J., Liu, X., Li, D., Shao, Z., Cao, H., Zhang, Y., ... Xu, J. (2016). Dynamic Control of Enhancer Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. *Developmental Cell*, 36(1), 9–23. <http://doi.org/10.1016/j.devcel.2015.12.014>
- Hug, C. B., & Vaquerizas, J. M. (2018). The Birth of the 3D Genome during Early Embryonic Development. *Trends in Genetics*, 34(12), 903–914. <http://doi.org/10.1016/J.TIG.2018.09.002>
- Hulsen, T., de Vlieg, J., & Alkema, W. (2008). BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics*, 9(1), 488. <http://doi.org/10.1186/1471-2164-9-488>
- Hurtado, A., Holmes, K. A., Ross-Innes, C. S., Schmidt, D., & Carroll, J. S. (2011). FOXA1 is a key determinant of estrogen receptor function and endocrine response. *Nature Genetics*, 43(1), 27–33. <http://doi.org/10.1038/ng.730>
- Hübner, M. R., Eckersley-Maslin, M. A., & Spector, D. L. (2013). Chromatin organization and transcriptional regulation. *Current Opinion in Genetics & Development*, 23(2), 89–95. <http://doi.org/10.1016/j.gde.2012.11.006>
- International HapMap Consortium. (2003). The International HapMap Project. *Nature*, 426(6968), 789–796. <http://doi.org/10.1038/nature02168>
- International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011), 931–945. <http://doi.org/10.1038/nature03001>
- Istrail, S., & Davidson, E. H. (2005). *Logic functions of the genomic cis-regulatory code*. Retrieved from www.pnas.org/cgi/doi/10.1073/pnas.0409624102
- Jacob, F., & Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology*, 3(3), 318–356. [http://doi.org/10.1016/S0022-2836\(61\)80072-7](http://doi.org/10.1016/S0022-2836(61)80072-7)
- Jensen, E. V., Suzuki, T., Kawashima, T., Stumpf, W. E., Jungblut, P. W., & DeSombre, E. R. (1968). A two-step mechanism for the interaction of estradiol with rat uterus. *Proceedings of the National Academy of Sciences*, 59(2), 632–638. <http://doi.org/10.1073/pnas.59.2.632>
- Jensen, E. V., & Jordan, V. C. (2003). The estrogen receptor: a model for molecular medicine. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, 9(6), 1980–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12796359>
- Jeziorska, D. M., Jordan, K. W., & Vance, K. W. (2009). A systems biology approach to understanding cis-regulatory module function. *Seminars in Cell & Developmental Biology*, 20(7), 856–862. <http://doi.org/10.1016/J.SEMCDB.2009.07.007>
- Jiang, Y., Qian, F., Bai, X., Liu, Y., Wang, Q., Ai, B., ... Li, C. (2019). SEdb: a comprehensive human super-enhancer database. *Nucleic Acids Research*, 47(D1), D235–D243. <http://doi.org/10.1093/nar/gky1025>
- Johnson, D. S., Mortazavi, A., Myers, R. M., & Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science (New York, N.Y.)*, 316(5830), 1497–502. <http://doi.org/10.1126/science.1141319>

- Johnston, S. D., Liu, X., Zuo, F., Eisenbraun, T. L., Wiley, S. R., Kraus, R. J., & Mertz, J. E. (1997). Estrogen-Related Receptor $\alpha 1$ Functionally Binds as a Monomer to Extended Half-Site Sequences Including Ones Contained within Estrogen-Response Elements. *Molecular Endocrinology*, 11(3), 342–352. <http://doi.org/10.1210/mend.11.3.9897>
- Joseph, R., Orlov, Y. L., Huss, M., Sun, W., Kong, S. L., Ukil, L., ... Liu, E. T. (2010). Integrative model of genomic factors for determining binding site selection by estrogen receptor- α . *Molecular Systems Biology*, 6, 456. <http://doi.org/10.1038/msb.2010.109>
- Khan, A., Mathelier, A., & Zhang, X. (2018). Super-enhancers are transcriptionally more active and cell type-specific than stretch enhancers. *Epigenetics*, 13(9), 910–922. <http://doi.org/10.1080/15592294.2018.1514231>
- Khan, A., & Zhang, X. (2016). dbSUPER: a database of super-enhancers in mouse and human genome. *Nucleic Acids Research*, 44(D1), D164–D171. <http://doi.org/10.1093/nar/gkv1002>
- Klein-Hitpass, L., Schorpp, M., Wagner, U., & Ryffel, G. U. (1986). An estrogen-responsive element derived from the 5' flanking region of the *Xenopus vitellogenin A2* gene functions in transfected human cells. *Cell*, 46(7), 1053–1061. [http://doi.org/10.1016/0092-8674\(86\)90705-1](http://doi.org/10.1016/0092-8674(86)90705-1)
- Kleinjan, D. A., & van Heyningen, V. (2005). Long-range control of gene expression: emerging mechanisms and disruption in disease. *American Journal of Human Genetics*, 76(1), 8–32. <http://doi.org/10.1086/426833>
- Klinge, C. M. (2001). Estrogen receptor interaction with estrogen response elements. *Nucleic Acids Research*, 29(14), 2905–19. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11452016>
- Kong, S. L., Li, G., Loh, S. L., Sung, W.-K., & Liu, E. T. (2011). Cellular reprogramming by the conjoint action of ER α , FOXA1, and GATA3 to a ligand-inducible growth state. *Molecular Systems Biology*, 7, 526. <http://doi.org/10.1038/msb.2011.59>
- Kouzarides, T. (2007). Chromatin Modifications and Their Function. *Cell*, 128(4), 693–705. <http://doi.org/10.1016/j.cell.2007.02.005>
- Kuiper, G. G. J. M., Carlsson, B., Grandien, K., Enmark, E., Häggblad, J., Nilsson, S., & Gustafsson, J.-Å. (1997). Comparison of the Ligand Binding Specificity and Transcript Tissue Distribution of Estrogen Receptors α and β . *Endocrinology*, 138(3), 863–870. <http://doi.org/10.1210/endo.138.3.4979>
- Kushiro, T., Nambara, E., & McCourt, P. (2003). Hormone evolution: The key to signalling. *Nature*, 422(6928), 122–122. <http://doi.org/10.1038/422122a>
- Kwiatkowski, N., Zhang, T., Rahl, P. B., Abraham, B. J., Reddy, J., Ficarro, S. B., ... Gray, N. S. (2014). Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature*, 511(7511), 616–620. <http://doi.org/10.1038/nature13393>
- Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., ... Weirauch, M. T. (2018). The Human Transcription Factors. *Cell*, 172(4), 650–665. <http://doi.org/10.1016/j.cell.2018.01.029>
- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., ... International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome.

- Nature*, 409(6822), 860–921. <http://doi.org/10.1038/35057062>
- Lawrence, M., Daujat, S., & Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. <http://doi.org/10.1016/j.tig.2015.10.007>
- Layman, W. S., & Zuo, J. (2015). Epigenetic regulation in the inner ear and its potential roles in development, protection, and regeneration. *Frontiers in Cellular Neuroscience*, 8, 446. <http://doi.org/10.3389/fncel.2014.00446>
- Ledford, H. (2018). Cancer researchers seek to harness mysterious DNA “super-enhancers.” *Nature* 2018 564:7735.
- Lee, S. M., Riley, E. M., Meyer, M. B., Benkusky, N. A., Plum, L. A., DeLuca, H. F., & Pike, J. W. (2015). 1,25-Dihydroxyvitamin D3 Controls a Cohort of Vitamin D Receptor Target Genes in the Proximal Intestine That Is Enriched for Calcium-regulating Components. *The Journal of Biological Chemistry*, 290(29), 18199–215. <http://doi.org/10.1074/jbc.M115.665794>
- Lelli, K. M., Slattery, M., & Mann, R. S. (2012). Disentangling the Many Layers of Eukaryotic Transcriptional Regulation. *Annual Review of Genetics*, 46(1), 43–68. <http://doi.org/10.1146/annurev-genet-110711-155437>
- Levenson, A. S., & Jordan, V. C. (1997). MCF-7: the first hormone-responsive breast cancer cell line. *Cancer Research*, 57(15), 3071–8. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9242427>
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754–1760. <http://doi.org/10.1093/bioinformatics/btp324>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <http://doi.org/10.1093/bioinformatics/btp352>
- Li, Q., Peterson, K. R., Fang, X., & Stamatoyannopoulos, G. (2002). Locus control regions. *Blood*, 100(9), 3077–3086. <http://doi.org/10.1182/BLOOD-2002-04-1104>
- Li, Y., Chen, C., Kaye, A. M., & Wasserman, W. W. (2015). The identification of cis-regulatory elements: A review from a machine learning perspective. *Biosystems*, 138, 6–17. <http://doi.org/10.1016/J.BIOSYSTEMS.2015.10.002>
- Li, Y., Rivera, C. M., Ishii, H., Jin, F., Selvaraj, S., Lee, A. Y., ... Ren, B. (2014). CRISPR Reveals a Distal Super-Enhancer Required for Sox2 Expression in Mouse Embryonic Stem Cells. *PLoS ONE*, 9(12), e114485. <http://doi.org/10.1371/journal.pone.0114485>
- Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., ... Dekker, J. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science*, 326(5950), 289–293. <http://doi.org/10.1126/science.1181369>
- Lin, C.-Y., Vega, V. B., Thomsen, J. S., Zhang, T., Kong, S. L., Xie, M., ... Liu, E. T. (2007). Whole-Genome Cartography of Estrogen Receptor α Binding Sites. *PLoS Genetics*, 3(6), e87. <http://doi.org/10.1371/journal.pgen.0030087>
- Liu, M., Maurano, M. T., Wang, H., Qi, H., Song, C.-Z., Navas, P. A., ... Stamatoyannopoulos, G. (2015).

- Genomic discovery of potent chromatin insulators for human gene therapy. *Nature Biotechnology*, 33(2), 198–203. <http://doi.org/10.1038/nbt.3062>
- Liu, Z., Merkurjev, D., Yang, F., Li, W., Oh, S., Friedman, M. J., ... Rosenfeld, M. G. (2014). Enhancer Activation Requires trans-Recruitment of a Mega Transcription Factor Complex. *Cell*, 159(2), 358–373. <http://doi.org/10.1016/j.cell.2014.08.027>
- Lovén, J., Hoke, H. A., Lin, C. Y., Lau, A., Orlando, D. A., Vakoc, C. R., ... Young, R. A. (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*, 153(2), 320–34. <http://doi.org/10.1016/j.cell.2013.03.036>
- Maeshima, K., Hihara, S., & Eltsov, M. (2010). Chromatin structure: does the 30-nm fibre exist in vivo? *Current Opinion in Cell Biology*, 22(3), 291–297. <http://doi.org/10.1016/j.ceb.2010.03.001>
- Mahony, S., & Pugh, B. F. (2015). Protein–DNA binding in high-resolution. *Critical Reviews in Biochemistry and Molecular Biology*, 50(4), 269–283. <http://doi.org/10.3109/10409238.2015.1051505>
- Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schütz, G., Umesono, K., ... Evans, R. M. (1995). The nuclear receptor superfamily: the second decade. *Cell*, 83(6), 835–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8521507>
- Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., ... Look, A. T. (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science (New York, N.Y.)*, 346(6215), 1373–7. <http://doi.org/10.1126/science.1259037>
- Mardis, E. R. (2007). ChIP-seq: welcome to the new frontier. *Nature Methods*, 4(8), 613–614. <http://doi.org/10.1038/nmeth0807-613>
- Moorthy, S. D., Davidson, S., Shchuka, V. M., Singh, G., Malek-Gilani, N., Langroudi, L., ... Mitchell, J. A. (2017). Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes. *Genome Research*, 27(2), 246–258. <http://doi.org/10.1101/gr.210930.116>
- Murdoch, F. E., Meier, D. A., Furlow, J. D., Grunwald, K. A. A., & Gorski, J. (1990). Estrogen receptor binding to a DNA response element in vitro is not dependent upon estradiol. *Biochemistry*, 29(36), 8377–8385. <http://doi.org/10.1021/bi00488a026>
- Niehrs, C., & Pollet, N. (1999). Synexpression groups in eukaryotes. *Nature*, 402(6761), 483–487. <http://doi.org/10.1038/990025>
- Niu, M., Tabari, E., Ni, P., & Su, Z. (2018). Towards a map of cis-regulatory sequences in the human genome. *Nucleic Acids Research*, 46(11), 5395–5409. <http://doi.org/10.1093/nar/gky338>
- Nora, E. P., Lajoie, B. R., Schulz, E. G., Giorgetti, L., Okamoto, I., Servant, N., ... Heard, E. (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature*, 485(7398), 381–385. <http://doi.org/10.1038/nature11049>
- Nuclear Receptors Nomenclature Committee. (1999). A unified nomenclature system for the nuclear receptor superfamily. *Cell*, 97(2), 161–3. Retrieved from

<http://www.ncbi.nlm.nih.gov/pubmed/10219237>

- Ogbourne, S., & Antal, T. M. (1998). Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes. *The Biochemical Journal*, 331 (Pt 1)(Pt 1), 1–14. <http://doi.org/10.1042/bj3310001>
- Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H., & Nakatani, Y. (1996). The Transcriptional Coactivators p300 and CBP Are Histone Acetyltransferases. *Cell*, 87(5), 953–959. [http://doi.org/10.1016/S0092-8674\(00\)82001-2](http://doi.org/10.1016/S0092-8674(00)82001-2)
- Olefsky, J. M. (2001). Nuclear receptor minireview series. *The Journal of Biological Chemistry*, 276(40), 36863–4. <http://doi.org/10.1074/jbc.R100047200>
- Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., ... Fraser, P. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature Genetics*, 36(10), 1065–1071. <http://doi.org/10.1038/ng1423>
- Ostuni, R., Piccolo, V., Barozzi, I., Polletti, S., Termanini, A., Bonifacio, S., ... Natoli, G. (2013). Latent Enhancers Activated by Stimulation in Differentiated Cells. *Cell*, 152(1), 157–171. <http://doi.org/10.1016/j.cell.2012.12.018>
- Overington, J. P., Al-Lazikani, B., & Hopkins, A. L. (2006). How many drug targets are there? *Nature Reviews Drug Discovery*, 5(12), 993–996. <http://doi.org/10.1038/nrd2199>
- Oyekan, A. (2011). PPARs and their Effects on the Cardiovascular System. *Clinical and Experimental Hypertension*, 33(5), 287–293. <http://doi.org/10.3109/10641963.2010.531845>
- Parker, S. C. J., Stitzel, M. L., Taylor, D. L., Orozco, J. M., Erdos, M. R., Akiyama, J. A., ... NISC Comparative Sequencing Program Authors. (2013). Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proceedings of the National Academy of Sciences of the United States of America*, 110(44), 17921–6. <http://doi.org/10.1073/pnas.1317023110>
- Pott, S., & Lieb, J. D. (2015). What are super-enhancers? *Nature Genetics*, 47(1), 8–12. <http://doi.org/10.1038/ng.3167>
- Pribnow, D. (1975). Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proceedings of the National Academy of Sciences*, 72(3), 784–788. <http://doi.org/10.1073/pnas.72.3.784>
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)*, 26(6), 841–2. <http://doi.org/10.1093/bioinformatics/btq033>
- Ranhotra, H. S. (2013). The orphan nuclear receptors in cancer and diabetes. *Journal of Receptors and Signal Transduction*, 33(4), 207–212. <http://doi.org/10.3109/10799893.2013.781624>
- Reis-Filho, J. S., Savage, K., Lambros, M. B. K., James, M., Steele, D., Jones, R. L., & Dowsett, M. (2006). Cyclin D1 protein overexpression and CCND1 amplification in breast carcinomas: an immunohistochemical and chromogenic in situ hybridisation analysis. *Modern Pathology*, 19(7), 999–1009. <http://doi.org/10.1038/modpathol.3800621>

- Reiter, F., Wienerroither, S., & Stark, A. (2017). Combinatorial function of transcription factors and cofactors. *Current Opinion in Genetics & Development*, 43, 73–81. <http://doi.org/10.1016/j.gde.2016.12.007>
- Rickels, R., & Shilatifard, A. (2018). Enhancer Logic and Mechanics in Development and Disease. *Trends in Cell Biology*, 28(8), 608–630. <http://doi.org/10.1016/J.TCB.2018.04.003>
- Robinson, P. J. J., Fairall, L., Huynh, V. A. T., & Rhodes, D. (2006). EM measurements define the dimensions of the ‘‘30-nm’’ chromatin fiber: Evidence for a compact, interdigitated structure. *Proceedings of the National Academy of Sciences*, 103(17), 6506–6511. <http://doi.org/10.1073/pnas.0601212103>
- Ross-Innes, C. S., Stark, R., Teschendorff, A. E., Holmes, K. A., Ali, H. R., Dunning, M. J., ... Carroll, J. S. (2012). Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*, 481(7381), 389–93. <http://doi.org/10.1038/nature10730>
- Sabari, B. R., Dall’Agnese, A., Boija, A., Klein, I. A., Coffey, E. L., Shrinivas, K., ... Young, R. A. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science*, 361(6400), eaar3958. <http://doi.org/10.1126/science.aar3958>
- Safe, S., Kim, K., & Kim, K. (2008). Non-classical genomic estrogen receptor (ER)/specificity protein and ER/activating protein-1 signaling pathways. *Journal of Molecular Endocrinology*, 41(5), 263–75. <http://doi.org/10.1677/JME-08-0103>
- Saha, S. K., Kim, K., Yang, G.-M., Choi, H. Y., & Cho, S.-G. (2018). Cytokeratin 19 (KRT19) has a Role in the Reprogramming of Cancer Stem Cell-Like Cells to Less Aggressive and More Drug-Sensitive Cells. *International Journal of Molecular Sciences*, 19(5). <http://doi.org/10.3390/ijms19051423>
- Saldanha, A. J. (2004). Java Treeview--extensible visualization of microarray data. *Bioinformatics*, 20(17), 3246–3248. <http://doi.org/10.1093/bioinformatics/bth349>
- Salmon-Divon, M., Dvinge, H., Tammoja, K., & Bertone, P. (2010). PeakAnalyzer: Genome-wide annotation of chromatin binding and modification loci. *BMC Bioinformatics*, 11(1), 415. <http://doi.org/10.1186/1471-2105-11-415>
- Saravanan, B., Soota, D., Islam, Z., Jayani, R., Mann, R., Farooq, U., ... Notani, D. (2019). Ligand dependent gene regulation by transient ERα clustered enhancers. *BioRxiv*. <http://doi.org/10.1101/541540>
- Schalch, T., Duda, S., Sargent, D. F., & Richmond, T. J. (2005). X-ray structure of a tetranucleosome and its implications for the chromatin fibre. *Nature*, 436(7047), 138–141. <http://doi.org/10.1038/nature03686>
- Schmidt, D., Schwalie, P. C., Ross-Innes, C. S., Hurtado, A., Brown, G. D., Carroll, J. S., ... Odom, D. T. (2010). A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Research*, 20(5), 578–88. <http://doi.org/10.1101/gr.100479.109>
- Schwabe, J. W., Chapman, L., Finch, J. T., & Rhodes, D. (1993). The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements. *Cell*, 75(3), 567–78. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8221895>

- Sengupta, D., Kannan, A., Kern, M., Moreno, M. A., Vural, E., Stack, B., ... Gao, L. (2015). Disruption of BRD4 at H3K27Ac-enriched enhancer region correlates with decreased c-Myc expression in Merkel cell carcinoma. *Epigenetics*, 10(6), 460–6. <http://doi.org/10.1080/15592294.2015.1034416>
- Shin, H. Y., Willi, M., Yoo, K. H., Zeng, X., Wang, C., Metser, G., & Hennighausen, L. (2016). Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nature Genetics*, 48(8), 904–911. <http://doi.org/10.1038/ng.3606>
- Siersbæk, R., Baek, S., Rabiee, A., Nielsen, R., Traynor, S., Clark, N., ... Mandrup, S. (2014). Molecular Architecture of Transcription Factor Hotspots in Early Adipogenesis. *Cell Reports*, 7(5), 1434–1442. <http://doi.org/10.1016/j.celrep.2014.04.043>
- Siersbæk, R., Rabiee, A., Nielsen, R., Sidoli, S., Traynor, S., Loft, A., ... Mandrup, S. (2014). Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. *Cell Reports*, 7(5), 1443–55. <http://doi.org/10.1016/j.celrep.2014.04.042>
- Song, Y., Ahn, J., Suh, Y., Davis, M. E., & Lee, K. (2013). Identification of Novel Tissue-Specific Genes by Analysis of Microarray Databases: A Human and Mouse Model. *PLoS ONE*, 8(5), e64483. <http://doi.org/10.1371/journal.pone.0064483>
- Soule, H. D., Vazquez, J., Long, A., Albert, S., & Brennan, M. (1973). A human cell line from a pleural effusion derived from a breast carcinoma. *Journal of the National Cancer Institute*, 51(5), 1409–16. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/4357757>
- Spitz, F., & Furlong, E. E. M. (2012). Transcription factors: from enhancer binding to developmental control. *Nature Reviews Genetics*, 13(9), 613–626. <http://doi.org/10.1038/nrg3207>
- Sun, F. L., & Elgin, S. C. (1999). Putting boundaries on silence. *Cell*, 99(5), 459–62. [http://doi.org/10.1016/s0092-8674\(00\)81534-2](http://doi.org/10.1016/s0092-8674(00)81534-2)
- Swinstead, E. E., Miranda, T. B., Paakinaho, V., Baek, S., Goldstein, I., Hawkins, M., ... Hager, G. L. (2016). Steroid Receptors Reprogram FoxA1 Occupancy through Dynamic Chromatin Transitions. *Cell*, 165(3), 593–605. <http://doi.org/10.1016/j.cell.2016.02.067>
- Szabó, T. A. (2010). Valók gráditsonkénti lépegetése (1818) és a Természet genetikai törvényei (1819). *Kaleidoscope History*, 1(1), 237–271. <http://doi.org/10.17107/KH.2010.1.237-271>
- Takayama, K., Suzuki, T., Tsutsumi, S., Fujimura, T., Urano, T., Takahashi, S., ... Inoue, S. (2015). RUNX1, an androgen- and EZH2-regulated gene, has differential roles in AR-dependent and -independent prostate cancer. *Oncotarget*, 6(4), 2263–76. <http://doi.org/10.18632/oncotarget.2949>
- Tan, S. K., Lin, Z. H., Chang, C. W., Varang, V., Chng, K. R., Pan, Y. F., ... Cheung, E. (2011). AP-2 γ regulates oestrogen receptor-mediated long-range chromatin interaction and gene transcription. *The EMBO Journal*, 30(13), 2569–81. <http://doi.org/10.1038/emboj.2011.151>
- Theodorou, V., Stark, R., Menon, S., & Carroll, J. S. (2013). GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. *Genome Research*, 23(1), 12–22. <http://doi.org/10.1101/gr.139469.112>
- Thorvaldsdottir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, 14(2), 178–192. <http://doi.org/10.1093/bib/bbs017>

- Thurman, R. E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M. T., Haugen, E., ... Stamatoiyannopoulos, J. A. (2012). The accessible chromatin landscape of the human genome. *Nature*, 489(7414), 75–82. <http://doi.org/10.1038/nature11232>
- Tolani, B., Gopalakrishnan, R., Punj, V., Matta, H., & Chaudhary, P. M. (2014). Targeting Myc in KSHV-associated primary effusion lymphoma with BET bromodomain inhibitors. *Oncogene*, 33(22), 2928–37. <http://doi.org/10.1038/onc.2013.242>
- Tomaskovic-Crook, E., Thompson, E. W., & Thiery, J. P. (2009). Epithelial to mesenchymal transition and breast cancer. *Breast Cancer Research*, 11(6), 213. <http://doi.org/10.1186/bcr2416>
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., ... Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, 28(5), 511–515. <http://doi.org/10.1038/nbt.1621>
- Truong, H., Geynet, C., Millet, C., Soullignac, O., Bucourt, R., Vignau, M., ... Baulieu, E.-E. (1973). *PURIFICATION OF ESTRADIOL RECEPTOR BY AFFINITY CHROMATOGRAPHY. REPRESENTATIVE EXPERIMENTS*. Retrieved from <https://febs.onlinelibrary.wiley.com/doi/pdf/10.1016/0014-5793%2873%2980306-0>
- Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., & Luscombe, N. M. (2009). A census of human transcription factors: function, expression and evolution. *Nature Reviews Genetics*, 10(4), 252–263. <http://doi.org/10.1038/nrg2538>
- Vega, V. B., Lin, C.-Y., Lai, K. S., Kong, S. L., Xie, M., Su, X., ... Liu, E. T. (2006). Multiplatform genome-wide identification and modeling of functional human estrogen receptor binding sites. *Genome Biology*, 7(9), R82. <http://doi.org/10.1186/gb-2006-7-9-r82>
- Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., ... Pennacchio, L. A. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature*, 457(7231), 854–858. <http://doi.org/10.1038/nature07730>
- Waddington, C. H. (1968). Towards a theoretical biology. *Nature*, 218(5141), 525–527. <http://doi.org/10.1038/218525a0>
- Waddington, C. H. (2012). The Epigenotype. *International Journal of Epidemiology*, 41(1), 10–13. <http://doi.org/10.1093/ije/dyr184>
- Walker, L. C., Harris, G. C., Holloway, A. J., McKenzie, G. W., Wells, J. E., Robinson, B. A., & Morris, C. M. (2007). Cytokeratin KRT8/18 expression differentiates distinct subtypes of grade 3 invasive ductal carcinoma of the breast. *Cancer Genetics and Cytogenetics*, 178(2), 94–103. <http://doi.org/10.1016/j.cancergencyto.2007.06.002>
- Wang, Z., Benoit, G., Liu, J., Prasad, S., Aarnisalo, P., Liu, X., ... Perlmann, T. (2003). Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors. *Nature*, 423(6939), 555–560. <http://doi.org/10.1038/nature01645>
- Wei, Y., Zhang, S., Shang, S., Zhang, B., Li, S., Wang, X., ... Zhang, Y. (2016). SEA: a super-enhancer archive. *Nucleic Acids Research*, 44(D1), D172–D179. <http://doi.org/10.1093/nar/gkv1243>

- Welboren, W.-J., van Driel, M. A., Janssen-Megens, E. M., van Heeringen, S. J., Sweep, F. C., Span, P. N., & Stunnenberg, H. G. (2009). ChIP-Seq of ERalpha and RNA polymerase II defines genes differentially responding to ligands. *The EMBO Journal*, 28(10), 1418–28. <http://doi.org/10.1038/emboj.2009.88>
- Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., ... Young, R. A. (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, 153(2), 307–19. <http://doi.org/10.1016/j.cell.2013.03.035>
- Wilson, T. E., Paulsen, R. E., Padgett, K. A., & Milbrandt, J. (1992). Participation of non-zinc finger residues in DNA binding by two nuclear orphan receptors. *Science (New York, N.Y.)*, 256(5053), 107–10. <http://doi.org/10.1126/science.1314418>
- Wittkopp, P. J., & Kalay, G. (2012). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics*, 13(1), 59–69. <http://doi.org/10.1038/nrg3095>
- Wray, G. A. (2007). The evolutionary significance of cis-regulatory mutations. *Nature Reviews Genetics*, 8(3), 206–216. <http://doi.org/10.1038/nrg2063>
- Writing Group for the Women’s Health Initiative Investigators, J. E., Anderson, G. L., Prentice, R. L., LaCroix, A. Z., Kooperberg, C., Stefanick, M. L., ... Writing Group for the Women’s Health Initiative Investigators. (2002). Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results From the Women’s Health Initiative Randomized Controlled Trial. *JAMA: The Journal of the American Medical Association*, 288(3), 321–333. <http://doi.org/10.1001/jama.288.3.321>
- Wu, C. -t., & Morris, J. R. (2001). Genes, Genetics, and Epigenetics: A Correspondence. *Science*, 293(5532), 1103–1105. <http://doi.org/10.1126/science.293.5532.1103>
- Xie, W., & Evans, R. M. (2001). Orphan nuclear receptors: the exotics of xenobiotics. *The Journal of Biological Chemistry*, 276(41), 37739–42. <http://doi.org/10.1074/jbc.R100033200>
- Xu, J., & Lin, D. I. (2018). Oncogenic c-terminal cyclin D1 (CCND1) mutations are enriched in endometrioid endometrial adenocarcinomas. *PLOS ONE*, 13(7), e0199688. <http://doi.org/10.1371/journal.pone.0199688>
- Yang, X., Lamia, K. A., & Evans, R. M. (2007). Nuclear Receptors, Metabolism, and the Circadian Clock. *Cold Spring Harbor Symposia on Quantitative Biology*, 72(1), 387–394. <http://doi.org/10.1101/sqb.2007.72.058>
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., ... Liu, X. S. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biology*, 9(9), R137. <http://doi.org/10.1186/gb-2008-9-9-r137>
- Zhong, J., Luo, K., Winter, P. S., Crawford, G. E., Iversen, E. S., & Hartemink, A. J. (2016). Mapping nucleosome positions using DNase-seq. *Genome Research*, 26(3), 351–64. <http://doi.org/10.1101/gr.195602.115>