

DOKTORI (PhD) ÉRTEKEZÉS

Madarász Kristóf

**Onkopatológiai génvariánsok- és fúziók klinikai
jelentőségének tanulmányozása *in silico* fehérje modellezés
segítségével**

DEBRECENI EGYETEM

KLINIKAI ORVOSTUDOMÁNYOK DOKTORI ISKOLA

Debrecen, 2025

DOKTORI (PhD) ÉRTEKEZÉS

**Onkopatológiai génvariánsok- és fúziók klinikai
jelentőségének tanulmányozása *in silico* fehérje modellezés
segítségével**

Madarász Kristóf

Témavezető: Dr. Mokánszki Attila



DEBRECENI EGYETEM

KLINIKAI ORVOSTUDOMÁNYOK DOKTORI ISKOLA

Debrecen, 2025

Tartalomjegyzék

1. Rövidítések jegyzéke.....	4
2. Bevezetés.....	7
3. A precíziós onkológia és a bioinformatika szerepe az ismeretlen jelentőségű génvariánsok patogenitásának meghatározásában.....	9
3.1 A precíziós onkológia és bioinformatika	9
3.2 A <i>TP53</i> gén és a mieloid neopláziák	10
3.2.1 A <i>TP53</i> gén.....	10
3.2.2 A <i>TP53</i> gén szerepe a mieloid neopláziában	12
3.2.3 Mieloid neopláziával összefüggő akut mieloid leukémia	13
3.3 A <i>BCOR</i> és a <i>BCOR</i> -átrendeződéssel járó szarkómák.....	14
3.3.1 A génátrendeződések szerepe tumoros megbetegedésekben	14
3.3.2 A szarkómás megbetegedések általános jellemzői.....	14
3.3.3 A <i>BCOR</i> és szerepe a daganatokban	15
3.3.4 A <i>BCOR</i> -átrendeződéssel járó szarkómák	16
4. Célkitűzések	19
5. Anyagok és módszerek.....	20
5.1 <i>TP53</i> mutációk molekuláris és <i>in silico</i> fehérjeelemzése mielodiszpláziás neopláziákban és akut mieloid leukémiában	20
5.1.1 A minták kiválasztása és a vizsgálat felépítése	20
5.1.2 Immunhisztokémiai vizsgálatok.....	20
5.1.3 DNS izolálás FFPE mintákból	20
5.1.4 Újgenerációs szekvenálás.....	20
5.1.5 <i>In silico</i> fehérjeelemzés	21
5.2 A <i>BCOR</i> -génátrendeződéses szarkómák <i>in silico</i> elemzése	22
5.2.1 Fehérjeinformációk	22
5.2.2 Fizikai-kémiai tulajdonságok számítása.....	22
5.2.3 Szipnálpeptidek és rendezetlen fehérje szakaszok predikciója	23
5.2.4 Fehérjeszerkezetek elemzése.....	23

5.2.5	Molekuláris dinamikai szimulációk	24
5.3	Statisztikai elemzés	25
6.	Eredmények.....	27
6.1	A TP53 molekuláris és <i>in silico</i> fehérjeelemzése mielodiszplázias neopláziákban és akut mieloid leukémiában	27
6.1.1	Klinikopatológiai jellemzők	27
6.1.2	Újgenerációs szekvenálás.....	29
6.1.3	A citogenetikai, az IHC eredmények, a túlélés és a TP53 mutációs státusz közötti kapcsolat vizsgálata.....	33
6.1.4	A mutációk patogenitásának <i>in silico</i> predikciója.....	35
6.1.5	Az IARC TP53 adatbázis <i>in vitro</i> kísérleti eredményeinek analízise	36
6.1.6	Mutáns TP53 fehérjék stabilitás analízise	36
6.1.7	A mutáns TP53 fehérje-fehérje kölcsönhatások vizsgálata.....	38
6.1.8	A TP53 fehérje-DNS kölcsönhatások vizsgálata a mutációk hatására	40
6.1.9	A mutáns TP53 fehérje patogenitási és szerkezeti stabilitásának statisztikai elemzése AML/MDS betegekben	40
6.2	BCOR-gént érintő génfúziók <i>in silico</i> funkcionális elemzése szarkómákban.....	44
6.2.1	A fúziós fehérjék szekvencia- és doménjellemzői	44
6.2.2	A fúziós fehérjék fizikokémiai tulajdonságai.....	47
6.2.3	Génontológia	49
6.2.4	Előre jelzett sejten belüli lokalizáció és jelzőpeptidek.....	51
6.2.5	A fehérjék rendezetlen régióinak előrejelzése	52
6.2.6	A BCOR-PCGF1 dimer kötési affinitásának statisztikai elemzése	54
6.2.7	A BRS fúziós fehérjék hatása a nem kanonikus Polycomb represszív komplex 1 kötődési affinitásaira	60
6.2.8	A dimer komplexek molekuláris dinamikai elemzése	62
7.	Megbeszélés	67
7.1	TP53 mutációk potenciális patogén hatásának vizsgálata és klinikai relevanciája AML és MDS esetekben	67

7.2	A <i>BCOR</i> fúziók patogenitási vizsgálata <i>in silico</i> eszközökkel	71
8.	Összefoglalás.....	79
9.	Summary	81
10.	Irodalomjegyzék.....	81
11.	Publikációs lista (A Kenézy Élettudományi Könyvtár jóváhagyása által)	101
12.	Tárgyszavak	105
13.	Köszönetnyilvánítás	106

1. Rövidítések jegyzéke

- **3D:** Háromdimenziós
- **AA:** Aminosav
- **ADD:** Acetiltranszferáz domén
- **AF:** AlphaFold
- **AHR:** Aromásszénhidrogén-receptor
- **AI:** Alifás index
- **AML:** Akut mieloid leukémia
- **AML-MR:** Akut mieloid leukémia mielodispláziával összefüggő eltérésekkel
- **ANK:** Ankyrin ismétlődések
- **APR-246:** Eprenetapopt
- **BCOR:** BCL6 korepresszor
- **BCOR-2:** BCOR 2-es izoforma
- **BRS:** *BCOR*-átrendeződéssel járó szarkóma
- **CCNB3:** G2/Mitotikus-specifikus Ciklin-B3
- **CDD:** Konzervált domén adatbázis
- **CDK:** Ciklin-dependens kináz
- **CIC:** Capicua transzkripció represszor
- **CIITA:** MHC osztály II transzaktivátor
- **Cry:** Kristályszerkezet
- **D-box:** Destrukciós Box
- **DBD:** DNS-kötő domén
- **ERAD:** Endoplazmatikus retikulum-asszociált degradáció
- **EWSR1:** RNS-kötő fehérje EWS
- **EZH2:** Enhancer of Zeste Homológ 2
- **FFPE:** Formaldehiddel fixált, paraffinba ágyazott szövetminta
- **FISH:** Fluoreszcens in situ hibridizáció
- **FYRC:** FY-gazdag domén C-terminális régió
- **FYRN:** FY-gazdag domén N-terminális régió
- **GO:** Génontológia

- **GOF:** Funkciónyerés (gain of function)
- **GPI:** Glikozilfoszfátidilinozitol
- **GRAVY:** Hidropátia átlagérték
- **hESC:** Humán embrionális őssejt
- **HMG-box:** High mobility group box domén
- **HOX:** Homeobox
- **IHC:** Immunhisztokémia
- **IDRs:** Intrinszikusan rendezetlen régiók
- **IPSS-M:** Molecular International Prognostic Scoring System
- **IPSS-R:** Revised International Prognostic Scoring System
- **ipTM:** Interfész prediktált templát modellezés (interface predicted template modeling)
- **ITD:** Internal tandem duplikációk
- **KDM2B:** Lizin-demetiláz 2B
- **KMT2D:** Hiszton-Lizin N-Metiltransferáz 2D
- **LD motif:** Leucin-Aszpartát ismétlődő motívum
- **LOF:** Funkcióvesztés (loss of function)
- **MAML1:** Mastermind-szerű fehérje 1
- **MAML1/3 TAD1/2:** Mastermind-like transzaktivációs domének 1 és 2
- **MAML3:** Mastermind-szerű fehérje 3
- **MD:** Molekuláris dinamikai
- **MDS:** Mielodiszpláziás szindróma
- **MDS-IB:** Mielodiszpláziás szindróma emelkedett blasztarány mellett
- **MDS-LB:** Mielodiszpláziás szindróma alacsony blasztarány mellett
- **mESC:** Egér embrionális őssejt
- **MM-PBSA:** Molekulamechanikai energiák kombinálva a Poisson-Boltzmann és a felületi kontinuum szolvatációval
- **NES/NLS:** Nukleáris export/lokalizációs szignálok
- **NGS:** Újgenerációs szekvenálás
- **NLS1:** Nukleáris lokalizációs szignál 1
- **OFCD-szindróma:** Oculofaciocardiodental szindróma

- **PANTHER GO:** Protein ANalysis THrough Evolutionary Relationships Génontológiai
- **PARP:** Poli ADP-Ribóz Polimeráz
- **PCGF1:** Polycomb csoport RING finger homolog
- **PRC1.1/ncPRC1.1 komplexek:** Nem-kanonikus Polycomb represszív komplexek
- **PRC1:** Polycomb represszív komplex 1
- **pTM:** Predicted template modeling
- **PUFD:** PCGF Ubikvitin-szerű fold diszkriminátor domén
- **RAWUL:** RING-ujj és WD40-asszociált ubikvitinszerű domén
- **RGAG1:** Retrotranszpozon Gag domént tartalmazó fehérje 1
- **Rg:** Girációs sugár
- **RMSD:** Négyzetes középeltérés
- **RMSF:** Négyzetes középfluktuáció
- **RNS-pII T+/-:** RNS-polimeráz II által szabályozott transzkripció pozitív/negatív szabályozása
- **RTL9:** Retrotranszpozon Gag-szerű fehérje 9
- **SD:** Standard deviáció (szórás)
- **Sec/SPI:** Szekretoros szignálpeptidek
- **SET domén:** Specifikus metiltranszferáz aktivitással rendelkező domén
- **SKP1:** S-fázis kináz-asszociált fehérje 1
- **SNPs:** Egy pontos nukleotid polimorfizmusok
- **TCR-T:** TCR-módosított T-sejtek
- **TpI:** Elméleti izoelektromos pont
- **UPR:** Kibomlott fehérje válasz (Unfolded Protein Response)
- **WHO:** Egészségügyi Világszervezet
- **ZC3H7B:** Cinkujj CCCH domént tartalmazó fehérje 7B

2. Bevezetés

A mielodiszpláziás szindrómák (MDS) és az akut mieloid leukémia (AML) a vérképző rendszer klonális rendellenességeiből eredő hematológiai malignitások, amelyeket a mieloid őssejtek genomikai instabilitása és differenciációs zavara jellemez [1]. Az MDS egy heterogén betegségcsoport, amelyet a vörsejtek elégtelen termelése és a perifériás vér citopéniája jellemez, míg az AML gyorsan progrediáló, agresszív megbetegedés, amelyet a leukémiás blasztok kontrollálatlan proliferációja dominál [2]. Az MDS és AML közötti kapcsolatot alátámasztja, hogy az MDS esetek 20-30%-a alakul át AML-é, amelyet a csontvelőben $\geq 20\%$ -os blasztarány és klonális genetikai elváltozások jelenléte igazol [3,4].

Globálisan az MDS incidenciája 4-5 eset/100 000 fő/év, míg az AML előfordulása 3-4/100 000 fő/év [5,6]. Az időskorú populációban (70+ év) ezek az értékek 5-10-szeresére nőnek, különösen komplex kariotípusú vagy *TP53* mutációval rendelkező esetekben [7]. Genetikai szempontból mindkét betegségben kiemelkedő szerepet játszanak a DNS metilációt (*TET2*, *DNMT3A*), a kromatin struktúra átalakulását (*ASXL1*, *EZH2*), valamint az RNS splicing-ot (*SF3B1*, *SRSF2*) szabályozó gének mutációi, amelyek elősegítik a klonális evolúciót és a terápiás rezisztenciát [8,2,9]. A komplex kariotípusokkal (pl. 5q-, 7q- deléción) és *TP53* mutációkkal (a betegek ~10-30%-ában) járó alcsoportok prognózisa kedvezőtlen, a medián túlélési idő <6 hónap intenzív kezelés mellett is, különösen biallelikus mutációk vagy magas *TP53* variáns allél frekvencia (VAF) (>40%) esetén [10–13].

A Molecular International Prognostic Scoring System (IPSS-M) új prognosztikai modellje integrálja a klinikai paramétereket, citogenetikai rendellenességeket és 31 gén szomatikus mutációit, pontosítva a betegek kockázat besorolását [2]. Ez a modell 46%-ban módosítja a Revised International Prognostic Scoring System (IPSS-R) kategóriákat, különösen *TP53* mutációkkal vagy *FLT3*-ITD-vel rendelkező betegeknél, ami hangsúlyozza a genetikai profil szerepét a kezelési döntésekben [2].

Az ismeretlen klinikai jelentőségű variánsok (variants of uncertain significance, VUS) értelmezése továbbra is az onkológiai kutatások egyik legnagyobb kihívása. Az újgenerációs szekvenálási (NGS) technológiák elterjedésével egyre több genetikai eltérést azonosítanak daganatokban, melyek közül számos variáns funkcionális és klinikai hatása nem tisztázott [14,15]. Ezek az ismeretlen variánsok különösen fontosak a daganatos betegségek esetében, mivel a genetikai eltérések pontos értelmezése alapvető feltétele a hatékony célzott terápiák alkalmazásának és a precíziós onkológia sikerességének. A VUS megfelelő értékelése nemcsak a diagnosztikai pontosságot javítja, hanem lehetővé teszi a daganatok genetikai profiljának pontosabb megértését is [16]. Ez hozzájárulhat a személyre szabott terápiás stratégiák kidolgozásához, amelyek figyelembe veszik a beteg egyedi genetikai adottságait. A klinikai

döntéshozatalban azonban ezeknek a variánsoknak az értelmezése komplex feladatot jelent. Ez integrált megközelítést igényel, amely magában foglalja a klinikai adatok elemzését, családi kórtörténetet, szövettani vizsgálatokat és bioinformatikai predikciós eszközök alkalmazását [17].

A génfúziók, mint például az *ALK*, *ROS1* és *NTRK* géneket érintők, kiemelt szerepet játszanak számos daganat patogenezisében. Ezek az eltérések onkogén fehérjéket hoznak létre, amelyek célzott terápiák hatékony célpontjai lehetnek [18,19]. Az *NTRK* génfúziók különösen ígéretesek, mivel tumor-agnosztikus terápiás célpontot jelentenek, azaz különböző szövettani típusú daganatokban is jelen lehetnek [18,20]. Ugyanakkor a célzott terápiákra adott válasz után gyakran kialakul terápiás rezisztencia, ami további kutatásokat igényel [21].

Az NGS technológiák és a bioinformatikai eszközök fejlődése lehetővé tette ezen fúziók pontosabb detektálását, azonban a klinikai jelentőségük értelmezése továbbra is kihívást jelent. A szolid tumorokban előforduló génfúziók gyakorisága 0,3%-tól 5%-ig terjed, és számos esetben célzott terápiák hatékony célpontjai lehetnek, különösen *NTRK*, *ALK* vagy *ROS1* fúzióval rendelkező esetekben [18,19].

Az *ALK*, *ROS1* és *NTRK* génfúziók esetében alkalmazott kinázgátlók (pl. entrectinib, larotrectinib) jelentős klinikai válaszokat eredményeznek, azonban a terápiás rezisztencia gyakori [19,20]. A génfúziók detektálására a fluoreszcens *in situ* hibridizáció (FISH), az immunhisztokémia (IHC) és az NGS alapú módszerek alkalmazása a leggyakoribb [22,18]. A hagyományos módszerek korlátai miatt a bioinformatikai eszközök (pl. FusionCatcher, STAR-Fusion) váltak kulcsfontosságúvá az összetett fúziós szerkezetek (pl. dupla vagy hármas fúziók) érzékeny és specifikus azonosításában [23,24]. A harmadik generációs, egy molekula szekvenálásán alapuló technológiák (pl. Oxford Nanopore) használata lehetővé teszi a fúziók teljes szerkezetének feltárását, különösen olyan esetekben, ahol a hagyományos módszerek nem hatékonyak, például komplex kromoszóma-rendellenességek vagy repetitív régiók esetén [25].

A bioinformatikai eszközök integrációja elengedhetetlen a fúziók funkcionális következményeinek értelmezéséhez. Az olyan platformok, mint a Fusion-sq, kombinálják az RNS-szekvenálást és a teljes genom szekvenálást (WGS), csökkentve ezzel a fals pozitív eredmények kockázatát. Ezen túlmenően a több omikai adat (pl. transzkriptomika, proteomika) integrált elemzése lehetővé teszi a fúziók által aktivált jelátviteli útvonalak feltérképezését, lehetővé téve új terápiás célpontok azonosítását [26].

3. A precíziós onkológia és a bioinformatika szerepe az ismeretlen jelentőségű génvariánsok patogenitásának meghatározásában

3.1 A precíziós onkológia és bioinformatika

Az NGS technológiák és a bioinformatikai eszközök integrációja alapvetően átalakította a precíziós onkológia területét, különösen a VUS patogenitásának értékelésében. Az NGS alapú nagy áteresztőképességű molekuláris profilozás lehetővé teszi a genomikus mutációk, génfüziók és más molekuláris jellemzők gyors és pontos detektálását, amelyek kulcsfontosságúak a célzott terápiák kiválasztásában. Az ilyen innovációk révén a betegspecifikus tumorprofil alapján személyre szabott kezelési stratégiák alakíthatók ki [27–29].

A variánsok patogenitásának előrejelzésére számos klasszikus és modern bioinformatikai eszköz áll rendelkezésre. A klasszikus algoritmusok, mint például a SIFT (Sorting Intolerant from Tolerant) és a PolyPhen-2 (Polymorphism Phenotyping v2), továbbra is széles körben alkalmazott módszerek. A SIFT az aminosav szekvenciák evolúciós konzervációján alapulva jósolja meg, hogy egy adott aminosav csere milyen mértékben zavarhatja meg a fehérje funkcióját [30], míg a PolyPhen-2 strukturális és evolúciós információkat integrálva osztályozza a variánsokat "benignus", "lehetségesen patogén" vagy "valószínűleg patogén" kategóriákba [31].

A modern predikciós eszközök, mint például a REVEL (Rare Exome Variant Ensemble Learner) és a FATHMM-XF (Functional Analysis through Hidden Markov Models eXtended Features) továbbfejlesztett algoritmusokat alkalmaznak. A REVEL több különböző prediktív modell eredményét egyesíti egy konszenzusos pontszám érdekében, amely pontosabban jelzi a variánsok patogenitását [32]. A FATHMM-XF pedig kiterjesztett genomikus jellemzőket használ, beleértve a kontextuális információkat, hogy azonosítsa a potenciálisan patogén mutációkat [33].

A fehérjestabilitás változásainak előrejelzésére szolgáló módszerek közül kiemelkedik az I-Mutant2.0, amely az aminosav szekvencia vagy fehérjestruktúra alapján becsüli meg a mutációk által okozott stabilitás változásokat [34]. Az újabb eszközök, mint például a DynaMut2 és a DDGun, képesek a fehérjestruktúrák részletes elemzésére, továbbá az AlphaFold által generált modellek is felhasználhatók bemenetként az előrejelzésekhez, így pontosabb becsléseket nyújtanak a stabilitási változásokról és azok funkcionális következményeiről [35–37].

A *TP53* gén mutációinak vizsgálatában alkalmazott specifikus módszerek közé tartozik az Align-GVG, amely elemzi a mutációk elhelyezkedését és azok funkcionális doménekhez való viszonyát. Emellett az olyan prediktív modellek, mint a PhD-SNP és a BayesDel, különösen hasznosak az SNP-k patogenitásának bináris osztályozásában [38,39]. Ezek az eszközök különböző szempontokat vizsgálnak: míg egyesek (pl. SIFT, PolyPhen-2) elsősorban evolúciós

konzervációs adatokat használnak, mások (pl. DynaMut2, DDGun3D) strukturális változásokat és fehérje-fehérje vagy fehérje-DNS kölcsönhatásokat is figyelembe vesznek [27,29].

Az ismeretlen jelentőségű variánsok klinikai értelmezése azonban továbbra is kihívást jelent. Az olyan bioinformatikai eszközök, mint például a Cancer Genome Interpreter (CGI) vagy a PODS (Precision Oncology Decision Support) szabályalapú és gépi tanulási megközelítéseket alkalmaznak ezen variánsok klinikai relevanciájának és terápiás célozhatóságának értékelésére. A CGI például olyan jellemzőket vizsgál, mint a mutáció helye a gén funkcionális doménjeiben és annak prevalenciája az emberi populációban, míg a PODS rendszer az ismert onkogén variánsok terápiás lehetőségeit is vizsgálja [27,29].

A mesterséges intelligencia (AI) és mélytanulási modellek szintén ígéretes eszközök ezen variánsok értelmezésében. Az olyan modellek, mint az AlphaMissense és az ESM1b, bizonyították hatékonyságukat örökletes emlőrák génekben található missense variánsok patogenitásának előrejelzésében [29,40,41]. Ezek az algoritmusok képesek komplex mintázatok felismerésére nagy adathalmazokban, így támogatva új biomarkerek felfedezését és driver mutációk azonosítását [42].

A multiomikai adatok integrációja szintén kulcsfontosságú szerepet játszik a precíziós onkológiában. A genomika, a transzkriptomika és a proteomika együttes elemzése lehetővé teszi a tumorbiológia mélyebb megértését és új terápiás célpontok azonosítását. Az ilyen integrált megközelítések nemcsak a molekuláris eltérések közötti összefüggések feltárását teszik lehetővé, hanem hozzájárulnak új biomarkerek azonosításához is, amelyek prognosztikai vagy terápiás jelentőséggel bírhatnak [43–45].

Összességében ezek az innovatív módszerek nemcsak pontosabbá teszik a VUS-ok patogenitásának meghatározását, hanem hozzájárulnak ahhoz is, hogy ezek alapján személyre szabott terápiák kerüljenek kifejlesztésre.

3.2 A *TP53* gén és a mieloid neopláziák

3.2.1 A *TP53* gén

A *TP53* gén, amely a 17-es kromoszóma rövid karján (17p13) helyezkedik el, az emberi daganatokban leggyakrabban érintett gén. Mutációi az összes malignus megbetegedés több mint 50%-ban előfordulnak, ami kiemeli a gén központi szerepét a daganatképződésben és a progresszióban [46,47]. A *TP53* által kódolt TP53 fehérje egy 393 aminosavból álló transzkripció faktor, amely alapvető szerepet játszik a sejtciklus szabályozásában, a DNS javításban, az apoptózis kiváltásában és a genom stabilitásának fenntartásában. E funkciók miatt a TP53-at gyakran nevezik a „genom őrzőjének” (guardian of the genome), mivel kulcsszerepet tölt be a sejtek normális működésének fenntartásában és a daganatképződés megakadályozásában [47].

A TP53 fehérje tumorszuppresszív funkciói közé tartozik a káros sejtsztódás megakadályozása azáltal, hogy lehetővé teszi a DNS károsodások kijavítását még a sejtsztódás előtt. Amennyiben a károsodás irreverzibilis, a TP53 aktiválja az apoptotikus útvonalakat, emellett transzkripciós aktivitása révén számos DNS javító gént aktivál, amelyek hozzájárulnak a genom integritásának helyreállításához [47]. A TP53 aktivációját különböző stresszhatások váltják ki, például hipoxia, hősokk vagy DNS-károsodás. Az aktiváció mechanizmusa főként poszttranszlációs módosításokkal történik, amelyek közül az N-terminális transzaktivátor doménben (TAD) található szerin- és treonin foszforilációk kulcsszerepet játszanak [48].

A TP53 mutációi két fő típusba sorolhatók: szomatikus és csíravonalbeli mutációk. A szomatikus mutációk egy adott daganatban fordulnak elő és lokális genetikai károsodások következményei, míg a csíravonalbeli mutációk örökölhetők, és az egész szervezetben jelen vannak. A csíravonalbeli TP53 mutációk felelősek olyan örökletes malignus szindrómákért, mint például a Li-Fraumeni szindróma (LFS), amelyet korai életkorban kialakuló emlőrák, szarkómák, agydaganatok és adrenokortikális karcinómák jellemeznek [49,50].

A TP53 mutációk funkcionálisan három fő következménnyel járhatnak: a fehérje tumorszuppressziós funkciójának elvesztése (loss of function, LOF), domináns negatív hatás kifejtése (dominant negative effect), vagy új onkogén aktivitások szerzése (gain of function, GOF) [51,52]. A leggyakoribb mutációk misszensz típusúak, amelyek a TP53 DNS-kötő doménjében (DBD) fordulnak elő (5-8. exon), és jelentős mértékben hozzájárulnak a daganatok agresszivitásához és rossz prognózisához [50,53].

A TP53 mutációk előfordulási gyakorisága és prognosztikai jelentősége jelentősen változik a daganatok típusától függően. Emlődaganatokban körülbelül az esetek 30%-ában található meg TP53 mutáció, amely különösen rossz prognózissal jár az ösztrogénreceptor-pozitív altípusokban [54], és rosszabb túléléssel társul mind ösztrogénreceptor-pozitív (OS), mind betegségspecifikus túlélés tekintetében [54]. Vastagbél-daganatok esetében ez az arány elérheti az 50%-ot is, különösen előrehaladott stádiumokban [50], ahol a transzaktivációs képességet inaktíváló mutációk gyakoribbak és rosszabb prognózissal társulnak, valamint befolyásolják a kezelésre adott választ, különösen az 5-fluorouracil (5-FU) alapú kemoterápiák esetében [55]. Nem kissejtes tüdődaganatban (NSCLC) a TP53 mutációk gyakorisága körülbelül 40%, míg hepatocelluláris karcinómákban ez az arány 30% [56,57]. A méhnyakot érintő malignus elváltozásokban azonban nem maga a TP53 gén mutációja dominál, hanem annak indirekt inaktíválása a magas kockázatú humán papillómavírus (HPV) E6 onkoproteinje által, amely elősegíti a TP53 ubiquitin-függő lebomlását [58].

Összességében a TP53 gén központi szerepet játszik mind a daganatok kialakulásában, mind azok progressziójában. Mutációi nemcsak diagnosztikai és prognosztikai értékkel bírnak számos

daganattípus esetében, hanem potenciális célpontokat is kínálnak új terápiás stratégiák fejlesztéséhez. Az ilyen célzott terápiák kidolgozása azonban további kutatásokat igényel annak érdekében, hogy jobban megértsük a *TP53* mutáns tumorok biológiai sajátosságait és azok terápiára adott válaszait.

3.2.2 A *TP53* gén szerepe a mieloid neopláziában

A *TP53* gén mutációi kiemelkedő jelentőséggel bírnak a MDS és az AML patogenezisében, továbbá meghatározó prognosztikai értékkel rendelkeznek. Az MDS és AML esetében a *TP53* mutációk jelenléte gyakran társul komplex kariotípussal és terápia rezisztenciával, amelyek jelentősen rontják a betegek túlélését. Ezen mutációk egyedi biológiai entitást hoznak létre, amelyet rossz prognózis és terápiás válasz jellemez [59,60]. A legújabb eredmények szerint a *TP53*-mutációk - többnyire misszensz típusúak - az újonnan diagnosztizált (*de novo*) MDS/AML esetek 5-10%-ában, míg a terápiával összefüggő MDS/AML esetek 30-40%-ában fordulnak elő, és a mutációk által kialakuló domináns negatív hatás illetve funkcióvesztés a rendkívül rossz prognózis fő oka [61].

A *TP53* mutációk MDS-ben és AML-ben szenvedő betegek teljes túlélésének egyik legerősebb független prediktorai. Bejar és munkatársai (2011) [62] kimutatták, hogy a *TP53* mutáns betegek medián túlélése mindössze 5-12 hónap volt, szemben a vad típusú *TP53*-val rendelkező betegek 66 hónapos medián túlélésével. A prognosztikai hatás szoros összefüggést mutat a VAF-val: magasabb VAF (>40%) esetén rövidebb túlélést figyeltek meg, míg alacsonyabb VAF-értékekkel rendelkező betegeknél kedvezőbb kimenetel tapasztalható [63]. A biállélikus inaktivációhoz vezető "multihit" *TP53* mutációk különösen rossz prognózissal társulnak, gyakran gyors betegségprogresszióval és terápia rezisztenciával járnak [3].

Az intratumorális genetikai heterogenitás jelentős szerepet játszik az MDS és AML progressziójában. A *TP53* függő és független klónok egyidejű jelenléte hozzájárulhat a betegség biológiai sokszínűségéhez és a terápiás rezisztencia kialakulásához [62,63]. A *TP53* mutációk általában korai, "trunkális" eseményekként jelennek meg, és domináns klónokká fejlődnek, különösen kemoterápia vagy sugárkezelés által kiváltott szelekciós nyomás hatására [64]. Számos *TP53* függő és független klón együtt létezhet, ami alátámasztja a tumorok széles genetikai intratumorális heterogenitását [65]. Az érzékeny szekvenálási technikák alkalmazása kulcsfontosságú az MDS diagnosztikájában, mivel már a kezdeti diagnózis során is kimutathatók olyan kis szubklónok, amelyek későbbi terápiás rezisztenciához vezethetnek [62].

Az új terápiás stratégiák területén számos ígéretes megközelítés jelent meg a *TP53* mutációval rendelkező MDS és AML kezelésében. Az immunterápiák, például a PD-1/PD-L1 gátlók, bispecifikus antitestek és kimérikus antigén receptor (CAR) T-sejt terápiák, új lehetőségeket

kínálnak az immunrendszer daganatellenes aktivitásának fokozására [66,67]. A TP53 szerkezetet moduláló szerek, mint az eprenetapopt (APR-246), célzottan próbálják helyreállítani a TP53 funkcióit, ígéretes eredményeket mutatva korai klinikai vizsgálatokban [68]. Egy másik ígéretes megközelítés a magrolimab nevű anti-CD47 monoklonális antitest, amely elősegíti a daganatos sejtek fagocitózist a CD47-SIRP α jelátviteli útvonal blokkolásával, amely hatékonyságát jelenleg klinikai vizsgálatokban értékeli [69,70]. Ezenkívül olyan kombinációs terápiák, mint az arzén-trioxid és a decitabin alkalmazása, valamint új gyógyszerek, például az idasanutlin (MDM2 inhibitor), szintén biztató eredményeket mutattak MDS-ben és AML-ben [3]. E stratégiák jelentősége különösen azoknál a betegeknél nő meg, akiknél - a gyakori terápiarezisztencia és relapszus okán - az allogén hematopoetikus sejttranszplantáció (allo-HCT) nem hoz tartós betegségmentes túlélést, így az új terápiák célja nemcsak a teljes remisszió arányának javítása, hanem a páciensek életminőségének fokozása is [10,71].

3.2.3 Mieloid neopláziával összefüggő akut mieloid leukémia

Az akut mieloid leukémia mielodiszpláziával összefüggő változásokkal (AML-MR) egy különálló, magas kockázatú betegségformának tekinthető, amely az AML-esetek 24-35%-át teszi ki [72,73]. Az AML-MR diagnózisa a legújabb WHO és ICC (International Consensus Classification of Myeloid Neoplasms and Acute Leukemias) osztályozások szerint morfológiai, genetikai vagy klinikai jellemzők (például megelőző az MDS-hez köthető citogenetikai eltérések, illetve a korábbi MDS/MPN transzformáció) alapján történik. Ezeknél a betegeknél rendszerint komplex kariotípus figyelhető meg, és a terápiarezisztencia is gyakoribb, ami összességében alacsonyabb remissziós arányt és rövidebb túlélést (gyakran kevesebb mint 1 év) eredményez. A standard intenzív indukciós protokollok hatékonysága általában erősen korlátozott, részben a magas életkor és társbetegségek, részben pedig a leukémiás klónok genetikai sokfélesége miatt. A liposzómás daunorubicin-citarabin (CPX-351) alkalmazása az AML-MR esetében - különösen a 60 év feletti betegek körében - hosszabb teljes túlélést eredményezett a hagyományos 7+3 protokollhoz képest, ugyanakkor a teljes remisszió arányának javítása továbbra is kihívást jelent. Az allo-HCT egyes betegeknél potenciálisan gyógyuláshoz vezethet, de az AML-MR-es páciensekben e beavatkozást követően is magas a relapszus kockázata és korlátozott a hosszú távú túlélés. A célzott terápiák, köztük az immunterápiás megközelítések és a TP53 funkcionalitását helyreállító gyógyszerek ígéretesnek bizonyulhatnak a magas kockázatú betegcsoport kezelésében, ám további, jól megtervezett klinikai vizsgálatok szükségesek annak érdekében, hogy az AML-MR-ben szenvedő páciensek számára hosszú távú terápiás előnyt és érdemi életminőség javulást lehessen elérni [10,61,64].

3.3 A *BCOR* és a *BCOR*-átrendeződéssel járó szarkómák

3.3.1 A génátrendeződések szerepe tumoros megbetegedésekben

A kromoszomális átrendeződések - beleértve a génfüziók kialakulását is - a daganatok egyik kritikus genetikai eseményének számítanak, és számos különböző tumor (köztük hematológiai és szolid daganatok) patogenezisét meghatározhatják [74,75]. A génátrendeződések többféle módon járulhatnak hozzá a karcinogenezishez. Egyes esetekben új, konstitutív onkogén fehérjék jönnek létre, amelyek állandó aktivációt vagy jelátviteli zavarokat idéznek elő (*ALK*-, *RET*-, vagy *BCR::ABL*-füziók), míg máskor a génfüziók meghatározott transzkripciós kontrollt, epigenetikai szabályozást vagy alternatív genetikai útvonalakat aktiválhatnak [76]. Számos diagnosztikai és terápiás protokoll ma már kifejezetten a génfüziók jelenlétének kimutatására támaszkodik: a célzott gyógyszerek - például tirozin-kináz gátlók - alkalmazása akkor lehet a leghatékonyabb, ha a tumor onkogén füziós fehérjéit gátolják [77]. A kromoszomális átrendeződések következményei nem korlátozódnak az onkoproteinek termelődésére: a füziós gének olyan nem kódoló RNS-eket (például hosszú nem kódoló RNS-eket vagy cirkuláris RNS-eket) is létrehozhatnak, amelyek közvetve módosíthatják a sejtek proliferációs, apoptotikus vagy differenciálódási folyamatait [75].

A génszerkezeti változások sokszínű kvantitatív és kvalitatív eltéréseket hozhatnak létre: a génátrendeződés gyakran felelőssé tehető a primer klonális eseményért (mint például a klasszikus Philadelphia-kromoszóma esetében), de a tumor evolúciója során újabb füziós variánsok is megjelenhetnek, amelyek terápiareszisztenciát és betegségkiújulást okozhatnak [74]. Mindezek alapján napjainkban már rutinszerű a molekuláris diagnosztika keretében végzett célzott szekvenálási, FISH, vagy real-time qPCR vizsgálat, amely képes fényt deríteni az esetlegesen jelen lévő, daganatképződésért felelős kromoszomális átrendeződésekre. A modern, „precíziós onkológiai” megközelítések egyik célja, hogy az adott betegnél kimutatott füziós gének és azok fehérjetermékei specifikus támadáspontként szolgálhassanak, így maximalizálva a terápiás válasz és minimalizálva a mellékhatások esélyét [76,78].

3.3.2 A szarkómás megbetegedések általános jellemzői

A szarkómák a mesenchimális eredetű rosszindulatú daganatok közé tartoznak, amelyek a test különböző kötő- és támasztószöveteiben alakulnak ki, beleértve a csontot, porcot, izmot, zsírszövetet és ereket [75,79,80]. Rendkívül heterogén csoportot alkotnak, több mint 100 különböző altípussal, amelyek klinikai és molekuláris jellemzőik alapján osztályozhatók. Ezek a daganatok két fő kategóriába sorolhatók: csontszarkómák és lágyrészsarkómák. A csontszarkómák közé tartozik az osteoszarkóma, Ewing-szarkóma és kondroszarkóma, míg a lágyrészsarkómák közé sorolhatók például a liposzarkóma, leiomyoszarkóma és

gastrointestinális stróma tumorok (GIST). A szarkómák egyes altípusai - például a differenciálatlan kisméretű kereksejtes szarkómák - különösen fontosak a molekuláris diagnosztika és célzott terápiák szempontjából, mivel ezek gyakran genetikai átrendeződésekkel, például *BCOR* génfüziókkal társulnak [81–85].

A szarkómák genetikai hátterének egyik legfontosabb jellemzője a kromoszomális átrendeződések jelenléte, amelyek génfüziókat eredményeznek. Ezek a fúziók transzkripciós faktorokat vagy epigenetikai szabályozókat érintenek, és alapvetően megváltoztatják a sejtek biológiai viselkedését. Például az *EWSR1::FLII* fúzió az Ewing-szarkómában egy kimerikus transzkripciós faktort hoz létre, amely számos downstream célgén expresszióját szabályozza, elősegítve a tumorsejtek túlélését és invázióját [86,87]. Hasonlóképpen, a *BCOR::CCNB3* fúzió a *BCOR*-átrendeződéssel járó szarkóma (BRS) patogeneziséhez vezet [82].

3.3.3 A *BCOR* és szerepe a daganatokban

3.3.3.1 A *BCOR* gén

A *BCOR* (BCL-6 korepresszor) gén az X kromoszóma rövid karján, az Xp11.4 régióban helyezkedik el, és egy 1755 aminosavból álló fehérjét kódol, amely transzkripciós represszorként működik [88,89]. A gén számos alternatív exonból áll, amelyek különböző izoformák létrejöttét teszik lehetővé, de ezek közül csak néhány képes ismert fehérje interakciókra, attól függően, hogy a splicing során mely domének maradnak meg [90].

3.3.3.2 A *BCOR* fehérje, funkcionális doménjei és interakciós partnerei

A *BCOR* fehérje széles körben expresszálódik az emberi szövetekben, és alapvető szerepet játszik az embrionális fejlődésben, a sejt differenciáció szabályozásában, valamint az epigenetikai mechanizmusokon keresztül történő génszabályozásban [88,89,91–95]. A *BCOR* fehérje több kulcsfontosságú funkcionális doménnel és interakciós hellyel rendelkezik. A BCL-6 kötő domén (Bbs) (498-514 aminosav) a BCL-6 transzkripciós represszonnal lép kölcsönhatásba a BCL-6 POZ doménjén keresztül, elősegítve annak géncsendesítő aktivitását [88,96]. Ez a kapcsolat kritikus a B-sejt germinális központok kialakításában, és a *BCOR* itt specifikus I. és II. osztályú hiszton-deacetilázokkal (HDAC) is kölcsönhat, tovább fokozva a repressziót [88,97,98].

Meghatározó régiója a PUF₁D (PCGF Ubiquitin-szerű fold diszkriminátor) domén, amely a Polycomb csoport RING finger homolog 1 (PCGF1) RAWUL doménjéhez kötődik. Ez az interakció alapvető a nem-kanonikus Polycomb represszív komplex (ncPRC1.1) összeszerelésében [93,94,99]. Az ncPRC1.1 komplex hiszton modifikációk (pl. H2AK119ub) révén csendesíti a célgének expresszióját, így létfontosságú a sejtek differenciációjának és sejsorsának meghatározásában [92,94,100]. A *BCOR* az ncPRC1.1 komplex részeként részt vesz

a homeobox (HOX) gének repressziójában is [93,99], ami kritikus az embrionális fejlődés során [101,94,102]. Az epigenetikai szabályozáson keresztül a BCOR részt vesz olyan gének csendesítésében is, amelyek kontrollálják a sejtciklust vagy apoptózis útvonalakat [92,94]. A PRC1.1 komplexen belüli szerepe mellett a BCOR a H2AK119ub révén elősegítheti a PRC2 kapcsolódását is bizonyos kromatin régiókhoz [94,103].

További fontos interakciós hely az AF9-cel (MLLT3) való kötési régió (Mbs) (1123-1245. aminosavak). Ez az interakció az AF9 AHD régiójával történik, és szerepet játszik az MLL komplexekhez kapcsolódó epigenetikai szabályozásban [104].

A BCOR szorosan kölcsönhat a KDM2B hiszton-demetilázssal is. Ez az interakció alapvető fontosságú a ncPRC1.1 összeszerelése során, mivel lehetővé teszi annak pontos lokalizációját specifikus CpG-szigetekhez [105,106]. A KDM2B-hez való kötődést elsősorban két kulcsfontosságú motívum mediálja. A PUF1 domén hierarchikus módon kölcsönhatásba lép a PCGF1 RAWUL doménjével, amely lehetővé teszi a KDM2B komplexben való részvételét. Ez az interakció stabilizálja a PRC1.1 komplexet és biztosítja annak epigenetikai aktivitását [93,99,107]. A BCOR fehérje specifikus szekvenciái - különösen azok, amelyek prolinban és savas-oldallánccal rendelkező aminosavakban gazdagok - közvetlenül kapcsolódnak a KDM2B katalitikus régiójához. Ezek a motívumok lehetővé teszik a PRC1.1 célzott lokalizációját olyan génszabályozási helyekre, ahol CpG-szigetek találhatóak [103,108].

3.3.3.3 A *BCOR* gén szerepe a tumorgenezisben

A *BCOR* gén mutációit számos betegségben leírták [90]. Germline mutációk esetén heterozigóta nőknél oculofaciocardiodental (OFCD) szindrómát okozhatnak, amely ritka genetikai rendellenességként kraniofaciális, szív- és szemfejlődési rendellenességekkel jár együtt [109]. Szomatikus mutációk gyakran fordulnak elő hematológiai malignitásokban, például MDS és AML-ben [60,110]. Az MDS-ben szenvedő betegek körében a *BCOR* mutációk rossz prognosztikai tényezőt jelentenek, mivel alacsonyabb túléléssel és nagyobb AML transzformációs kockázattal társulnak [111,112]. Ezen kívül a *BCOR* internal tandem duplikációi (ITD) és fúziós eseményei is jelentős szerepet játszanak különböző daganatok kialakulásában. Az ITD-k leggyakrabban a PUF1 domént érintik, ami megváltoztatja annak funkcionális aktivitását. Ezeket az eltéréseket gyakran figyelték meg vesedaganatokban, például a vese világos sejt szarkómájában, központi idegrendszeri daganatokban, magas rizikójú méhszarkómákban és primer csontszarkómákban [113–118].

3.3.4 A *BCOR*-átrendeződéssel járó szarkómák

A BRS egy ritka mesenchimális daganattípus, amely elsősorban fiatal betegeket érint. A WHO 5. kiadású klasszifikációja szerint a lágyszövet- és csontdaganatok kategóriáján belül a "csont és

lágyrész differenciálatlan kisméretű kereksejtes szarkómáinak" harmadik különálló alcsoportjába tartozik [119]. A BRS hisztológiai jellemzői, immunprofilja és génexpressziós mintázatai eltérnek azoktól a kereksejtes szarkómáktól, amelyekben az *EWSR1* gén fúziója nem-ETS (non-E26 transformation-specific) géncsaládba tartozókkal fordul elő, valamint a Capicua transzkripciós represszorral (*CIC*) átrendezett szarkómáktól [120]. A *BCOR*-hoz kapcsolódó leggyakoribb genetikai eltérések közé tartoznak az ITD-k és a transzlokációk, amelyek génfúziókat eredményeznek. A G2/mitotikus-specifikus ciklin-B3 (*CCNB3*) a BRS-ben azonosított leggyakoribb fúziós partner [82,121]. A molekuláris diagnosztikai módszerek, különösen az NGS elterjedésével újabb fúziókat azonosítottak: 2016-ban Spacht és munkatársai két új fúziót írtak le: *BCOR::MAML3* (mastermind-like protein 3) és *ZC3H7B::BCOR* (zinc finger CCCH domain-containing protein 7B) [122], ezt követően Kao és munkatársai 2018-ban fedezték fel a *KMT2D::BCOR* (histone-lysine N-methyltransferase 2D) [120], míg Yoshida és munkatársai 2020-ban a *CIITA::BCOR* (MHC class II transactivator) génátrendeződést [123]. Legutóbb, 2022-ben Vassella és munkatársai azonosították az *RTL9::BCOR* génfúziót (retrotransposon gag-like protein 9, amelyet retrotransposon gag domén-tartalmú fehérje 1-ként is említene). Ez a fúziós fehérje egyedi módon tartalmaz egy nem-kanonikus BCOR-t, amelyből hiányzik a 1168-1201 régió, így megfelel a *BCOR-2* fehérje izoformának [124]. Kutatócsoportunk 2023-ban egy további génfúziót írt le BRS-ben: *BCOR::CLGN* (calmegin) [125]. Ezen kívül Cocchi és munkatársai 2024-ben két új BCOR-2 izoform fúziós fehérjét azonosítottak, a *BCOR::MAML1*-t (mastermind-like protein 1) és az *AHR::BCOR*-t (aryl hydrocarbon receptor) [126].

3.3.4.1 Diagnosztika és kezelés

A BRS leggyakrabban gyermekeknél, serdülőknél és fiatal felnőtteknél fordul elő, döntően csontokban (pl. végtagokban, gerincben) vagy lágyrészekben [120,127]. Klinikai képükre jellemző a gyors növekedés és a metasztázis képző hajlam, noha a végleges diagnózist leginkább összetett molekuláris vizsgálatok - például RT-qPCR, FISH, illetve NGS - segítségével lehet felállítani [122,128,129]. A diagnózis megállapításában a morfológiai és immunhisztokémiai jellemzők játszanak kulcsszerepet. A tumorok sejtes állományát kisméretű, kerek vagy esetenként orsó alakú sejtek alkotják, amelyek BCOR IHC esetén erős nukleáris pozitivitást mutathatnak; ezzel szemben a CD99 általában gyengén, foltszerűen pozitív vagy negatív [120,127]. Radiológiai vizsgálatok, például az MRI, extraosseális lágyrészbe terjedő folyamatot tárnak fel, ami bizonyos esetekben hasonlíthat az Ewing-szarkómák képére [128,130].

A terápiás megközelítések tekintetében jelenleg nincs egységesen elfogadott protokoll, jórészt a daganat ritka előfordulása miatt [127]. Gyakran alkalmazzák az Ewing-szarkómára kifejlesztett kemoterápiás séma (vincristin, doxorubicin, cyclophosphamid) különböző variációit, melyet

sebészi reszekció és szükség szerint sugárkezelés egészít ki [120,131]. Az irodalmi adatok szerint azonban a nem-Ewing-alapú terápiás megközelítések is közel azonos eredményeket mutathatnak, ezért elképzelhető, hogy a BRS számára saját, specifikusabb kezelési algoritmus kidolgozása válik szükségessé [128]. A lokalizált esetekben az öt éves túlélési arány általában 70-80% körül mozog, míg a metasztatikus megbetegedések prognózisa jelentősen rosszabb [120,127,131]. A molekuláris diagnosztikai módszerek fejlődése kiemelt fontosságú a BRS korai felismerésében, a helyes terápiás döntéshozatalban és a betegek hosszú távú életminőségének javításában [122,131].

4. Célkitűzések

Munkánk célul tűzi ki a malignus elváltozások háttérében álló két alapvető molekuláris genetikai eltérés, a kisebb méretű, nukleotid bázisokat érintő aberrációk (SNV, indel) és a nagyobb kiterjedésű, strukturális változásokat eredményező génfüziók (transzlokációk) tanulmányozását. Ennek megvalósítása érdekében az egyik leggyakoribb onkohematológiai betegségekhez vezető eltérés, a *TP53* gén variánsok, és az egyik legkritikább lágyrész szarkómákban előforduló *BCOR* génátrendeződések vizsgálatát tervezzük elvégezni. Az újgenerációs szekvenálás során kimutatott ismeretlen klinikai jelentőségű nukleotidokat érintő eltérések és génfüziók fehérje szinten történő elemzése *in silico* bioinformatikai programok segítségével kívánjuk megvalósítani.

I. A *TP53* gén mutációinak klinikai és molekuláris jelentőségének vizsgálata mielodiszpláziás szindrómában (MDS) és akut mieloid leukémiában (AML) szenvedő betegekben.

- A *TP53* gén mutációinak klonális heterogenitásának vizsgálata MDS és AML csontvelőmintákban.
- A *TP53* mutációs státusz és a hematopoetikus rendellenességek súlyossága közötti összefüggések vizsgálata.
- A *TP53* gén mutációi által okozott TP53 fehérje szerkezeti és funkcionális változásainak *in silico* elemzése, különös tekintettel a fehérje-fehérje és fehérje-DNS kölcsönhatásokra.
- Különböző patogenitási pontozási rendszerek összehasonlítása a *TP53* génmutációk értelmezésének megkönnyítésére.
- A TP53 fehérje stabilitásának változásait vizsgáló *in silico* szekvenacia- és szerkezetalapú elemzések elvégzése.

II. A *BCOR*-átrendeződéssel járó szarkómák (BRS) onkogén mechanizmusának feltárása *in silico* megközelítésekkel.

- A vizsgált betegcsoport (BRS) jelenleg elérhető irodalmi adatainak összegyűjtése, beleértve a fúziós gének cDNS szekvenciáit, amelyek reverz transzkripcióval való idejű PCR (RT-qPCR) és szekvenálási adatokon alapulnak, valamint a fúziós fehérjék aminosavszekvenciáinak összeállítása.
- *In silico* megközelítések alkalmazása a fúziós fehérjék szekvenciáinak és szerkezeteinek elemzésére, különös tekintettel a funkcionális domének változásaira és a fehérje-fehérje kölcsönhatásokra.
- A fúziós fehérjék 3D szerkezeteinek felépítése, a *BCOR* fúzióval rendelkező szarkómák onkogén mechanizmusának tanulmányozása, ideértve a RAWUL-PUFD domének kötési affinitásának változásait és azok hatását a PRC1 komplexen belüli kölcsönhatásokra.

5. Anyagok és módszerek

5.1 TP53 mutációk molekuláris és *in silico* fehérjeelemzése mielodiszplázias neopláziákban és akut mieloid leukémiában

5.1.1 A minták kiválasztása és a vizsgálat felépítése

A betegeket a Debreceni Egyetem, Klinikai Központ, Hematológiai Tanszékén kezelték. Formaldehiddel fixált, paraffinba ágyazott csontvelő-biopszia szövetmintákat (formaldehyde-fixed paraffin-embedded, FFPE) retrospektív módon elemeztünk összesen 77 olyan betegről, akiket a WHO 2022-es irányelvei szerint AML-MR, megnövekedett blasztarányú MDS (MDS-IB; 12 eset), illetve alacsony blasztarányú MDS (MDS-LB; 39 eset) kategóriákba soroltak át a Debreceni Egyetem, Klinikai Központ, Patológiai Intézetében. A hematoxin-eozinnal (H&E) festett metszeteket patológus kategorizálta. A citogenetikai elemzést a rutindiagnosztikai eljárás részeként végezték el. A vizsgálatokhoz etikai engedélyekkel rendelkezünk (60355-2/2016/EKU és IV/8465-3/2021/EKU). A tanulmányok a Helsinkai Nyilatkozat irányelveinek megfelelően zajlottak.

5.1.2 Immunhisztokémiai vizsgálatok

A H&E metszetek vizsgálatát követően TP53 IHC elemzést végeztünk Do-07 klónnal (Dako, Agilent Technologies Company, Santa Clara, CA, USA). Az IHC pozitivitást akkor állapítottuk meg, ha a TP53 festési intenzitása magas (3+) volt, és legalább a sejtek 10%-a pozitívnak bizonyult.

5.1.3 DNS izolálás FFPE mintákból

Az FFPE szövetek genomikus DNS-ének (gDNS) kinyeréséhez a QIAamp DNA FFPE Tissue Kit-et (Qiagen, Hilden, Németország) használtuk. Az izolálást a gyártó utasításai szerint végeztük, és a gDNS-t 50 µL eluáló pufferben oldottuk fel. A DNS koncentrációkat a Qubit dsDNS HS Assay Kit segítségével mértük Qubit 4.0 Fluorométerrel (Thermo Fisher Scientific, Waltham, MA, USA).

5.1.4 Újgenerációs szekvenálás

A gDNS fragmentálását követően a könyvtárakat az Accel-Amplicon Comprehensive TP53 panel (Swift Biosciences, Ann Arbor, MI, USA) segítségével hoztuk létre. A szekvenálást MiSeq rendszerrel végeztük (MiSeq Reagent Kit v3 600 cycles, Illumina, San Diego, CA, USA). A könyvtárakat (végső koncentráció: 4 nM, egyenlő molaritás szerint összemérve) 0,2 nM NaOH

hozzáadásával denaturáltuk, majd Illumina hibridizációs pufferrel (San Diego, CA, USA) 40 pM-ra hígítottuk. A végső betöltési koncentráció 10 pM könyvtár és 5% PhiX volt. A szekvenálást a MiSeq használati utasításának megfelelően végeztük. Az elkészített könyvtárakat multiplex formában szekvenáltuk páros végű futtatással, hogy 2×150 bp hosszúságú readeket nyerjünk legalább 250X lefedettséggel. Az adapterektől trimmert fastq fájlokat a MiSeq Reporter (Illumina, San Diego, CA, USA) segítségével generáltuk.

A nyers szekvencia adatokat a NextGENe szoftverrel (v.2.4.3.; SoftGenetics, State College, PA, USA) elemeztük SNV és indel jelenlétének kimutatására. A szekvenciák illesztéséhez a GRCh37 (UCSC hg19) referencia genomot használtuk. Az egyes variánsok esetében a cut-off 5% VAF volt.

Az NGS vizsgálatokat a Debreceni Egyetem Klinikai Központ, Patológiai Intézet, Molekuláris Daganatpatológiai Laboratóriumában végeztem, a laboratórium munkatársainak közreműködésével.

5.1.5 *In silico* fehérjeelemzés

A TP53 fehérjére (P04637, P53_HUMAN) vonatkozó információkat az UniProt adatbázisból és az RCSB Protein Data Bankból gyűjtöttük össze. A trunkált fehérjék háromdimenziós (3D) szerkezeteit a Robetta szoftver segítségével építettük fel [132,133]. A mutációk által érintett TP53 fehérje poszttranszlációs módosítási (PTM) helyeit a PhosphoSite Plus weboldal segítségével azonosítottuk [134]. A szekvenciák strukturális rendezetlenségi meghatározását az IUPred3 webszerverrel végeztük [135]. A másodlagos szerkezet becslésére a GORIV alkalmazást használtuk [136]. A variánsok szekvenciaalapú (Seq) stabilitásváltozásait az I-Mutant2.0 [34] és a DDGun [137,138] szerverekkel, alapértelmezett paraméterekkel állapítottuk meg. A szerkezetalapú (struk) stabilitási tesztekhez az I-Mutant2.0 [34], a DynaMut2 [36] és a DDGun3D [137,138] alkalmazásokat használtuk. A predikciókat számos TP53 kristály-szerkezet (Cry) (5MCT, 5MG7, 5MF7 [139], 1AIE [140], 1C26 [141], 2FOO [142], 1YC5 [143]) és az AlphaFold (AF) [35,144] nagy pontosságú mélytanulási algoritmus által épített 3D modellek felhasználásával végeztük. A mutációk fehérjestabilitásra gyakorolt prediktált hatását ($\Delta\Delta G_{\text{stabilitás}}$ érték) a kísérletileg meghatározott eltérésekkel a p.G245S [145] és p.R248Q [146] mutációk esetében hasonlítottuk össze, mivel kísérleti adatok csak ezekben az esetekben álltak rendelkezésre. A DynaMut2 és DDGun3D stabilitásvizsgálatokhoz olyan emberi TP53 kristályszerkezeteket (5MCT, 5MG7, 5MF7) használtunk, amelyek csak a 94-293-as DNS-kötő domént tartalmazzák, így a rendezetlen szakaszok közvetlen hatása a $\Delta\Delta G_{\text{stabilitás}}$ becslésekben nem jelentkezett.

A fehérje-fehérje kölcsönhatások (PPI) vizsgálatához az mCSM-PPI2 [147] programot alkalmaztuk, hogy a p.G334R mutáció esetében kiszámítsuk a TP53 monomerek egymásra

gyakorolt kölcsönhatás változásait homotetramer szerkezetben. Valamint meghatározzuk a p.S362N mutációt érintő kölcsönhatás változásokat az ubiquitin-karboxil-terminális hidroláz 7 (vagy herpeszvírus-asszociált ubiquitin-specifikus proteáz; USP7/HAUSP) fehérjével történő interakció során. A PPI elemzéshez a p.G334R mutáció esetében a 2FOO [142] kristályszerkezetet használtuk, mivel ez tartalmazta az USP7/HAUSP N-terminális doménjét TP53 peptiddel komplexben. A p.S362N mutáció esetében olyan kristályszerkezeteket használtunk, amelyek tartalmazzák a TP53 tetramerikus oligomerizációs doménjét: 1OLG [148] és 1SAL [149]. A mutáns TP53 fehérjék DNS-hez való affinitásának változásait az mCSM-NA [150] segítségével számítottuk ki.

A variáns fehérjék patogénitásának meghatározására a National Cancer Institute International Agency for Research on Cancer (IARC) TP53 adatbázisát (R20, 2019. július) [151], és annak komponenseit (Align-GVGD [152,153], BayesDel [154,155], REVEL [32], Sift class [30], PolyPhen2 [31], Transactivation, TransactivationClass [156], DNE_LOFclass [157], DNE class [158,159], Structure/Function class [156]), valamint a ClinVar [160–164], Varsity [39], Phd-SNP^s [38] és FATHMM-XF [33] patogénitást meghatározó pontrendszereket használtuk.

5.2 A *BCOR*-génátrendeződéses szarkómák *in silico* elemzése

5.2.1 Fehérjeinformációk

A szakirodalomban eddig ismert *BCOR* gént érintő kilenc fúzió nukleotid szekvenciáit vizsgáltuk [82,120,122–125]. Az Ensemble adatbázist (GRCh37.p13; GCA_000001405.14) [165] használtuk a fúziók töréspontjainak lokalizálására és a szekvenciák manuális összeállítására. A kódoló aminosav szekvenciákat *in silico* módon határoztuk meg az ExPASy: Translate segítségével [166]. A *BCOR/BCOR-2*, *CCNB3*, *MAML3*, *ZC3H7B*, *KMT2D*, *CIITA*, *RTL9*, *CLGN*, *MAML1* és *AHR* fehérjék, valamint az összeállított fúziós fehérjék információit az UniProt adatbázisból szereztük be [167] (a gének és fehérjék fúzióját a legújabb nomenklátúra szerint dupla kettősponttal (::) jelöltük [168]). A doméninformációkat az UniProt [167], a Conserved Domain Database (CDD) [169] és az InterPro (verzió 5.67-99.0) adatbázisokból nyertük [170]. A Gén Ontológiai kifejezéseket (GO) az összeállított aminosavszekvenciák alapján határoztuk meg az InterPro segítségével, specifikusan a PANTHER (Protein ANalysis THrough Evolutionary Relationships) GO kifejezéseket alkalmaztuk [171–174].

5.2.2 Fizikai-kémiai tulajdonságok számítása

A fizikai-kémiai paramétereket az ExPASy ProtParam eszközzel számoltuk [175], beleértve az elméleti izoelektromos pontot (TpI), a molekulatömeget, a pozitív és negatív töltésű

aminosavak összesített számát, az extinkciós együtthatót (EC) [176–178], az instabilitási indexet (II) [179], az alifás indexet (AI) [180], valamint a hidropátia átlagértékét (GRAVY) [181]. A vad típusú fehérjék és a fúziós fehérjék közötti fizikai-kémiai értékek változásának kiszámításához a következő képletet alkalmaztuk:

$$\text{változás (\%)} = \frac{\text{fúziós fehérje értéke} - \text{vad típusú fehérje értéke}}{|\text{vad típusú fehérje értéke}|}$$

5.2.3 Szignálpeptidek és rendezetlen fehérje szakaszok predikciója

A SignalP 6.0 programot alkalmaztuk a fehérjékben található szignálpeptidek azonosítására. Csak az eukariótákban előforduló szekretoros szignálpeptideket (Sec/SPI) kerestük, mivel a bemeneti organizmusként az „Eukarióták” opciót választottuk [182]. A DeepLoc 2.0 többcímkes prediktort használtuk a fehérjék intracelluláris lokalizációjának becslésére; ez az eszköz transzformátor-alapú fehérje nyelvi modellt használ, pontos előrejelzéseket és interpretálhatóságot biztosítva. Csak azokat a valószínűségi értékeket vettük figyelembe, amelyek meghaladták a DeepLoc 2.0 alapértelmezett küszöbértékeit: citoplazma (0,4761), sejtmag (0,5014) és endoplazmatikus retikulum (0,6090) [183]. Az előrejelzések érvényességének ellenőrzéséhez az eredményeket összevetettük az UniProt adatbázis és a The Human Protein Atlas adataival [167,184]. Emellett a NetGPI-1.1 GPI-anchored prediktort alkalmaztuk a fehérjék glikozilfoszfatidilinozitol (GPI) horgonyozottság vizsgálatára. A három előrejelzőt a <https://services.healthtech.dtu.dk/> weboldalon értük el.

A fehérjékben található intrinszikusan rendezetlen régiók (IDRs) előrejelzéséhez az IUPred3-at (<https://iupred3.elte.hu/>) használtuk, amely biofizikai modellen alapuló online eszközzel azonosítja a natív körülmények között stabil szerkezet nélküli területeket. Az IUPred3 webes szerver segítségével összehasonlító elemzést végeztünk a vad típusú BCOR (vBCOR) és annak fúziós fehérjeinek 1448-1633 (region^{ANK+linker}) és 1634-1748 aminosavig terjedő (region^{PUFD}) szakaszain [135].

5.2.4 Fehérjeszerkezetek elemzése

A BCOR-PCGF1 és PRC1.1 komplexekben részt vevő fehérjék háromdimenziós (3D) modelljeit az AlphaFold3 (AF3) segítségével generáltuk [37]. A BCOR és a PCGF1 dimerjeit, valamint a BCOR, a PCGF1, a KDM2B és az S-fázis kináz-asszociált fehérje 1 (SKP1) tetramerjeit építettük fel. Minden komplexhez 50 modellt készítettünk. Az általunk generált modelleket összevetettük a kísérletileg meghatározott BCOR-PCGF1 dimer szerkezettel (PDB ID: 4HPL) [93] és a vad típusú fehérjék modelljeivel. Az AF Server (béta) szekvencia hossza vonatkozó limitációi miatt az 5480 aminosavból álló KMT2D::BCOR szerkezetét nem tudtuk elkészíteni.

A PRODIGY (PROtein binDIng enerGY prediction) web szerveret [185] alkalmaztuk a kötési affinitás becslésére, amelyet a Gibbs-féle szabadenergia-változás (ΔG , kcal mol⁻¹) ír le. A PRODIGY webszerver strukturális és energetikai paraméterek kombinálásával becsli meg a kötés során bekövetkező ΔG változást a fehérje-fehérje komplexek 3D szerkezete alapján. Az algoritmus intramolekuláris kontaktokat és nem kölcsönható felületi tulajdonságokat integrál a kötési affinitás pontos becsléséhez. A BCOR-PCGF1 dimerek ΔG -változásait a teljes fehérjeszekvencia (teljes hossz), különösen a BCOR PUF_D régió (1634-1748) és a PCGF1 RAWUL régió (167-255; RAWUL-PUF_D doménhossz) mentén vizsgáltuk. A RAWUL domén 167-177 és 185-254, valamint a PUF_D domén 1636-1748 (4HPL hossz) tartományait elemeztük [93]. Ezenkívül különböző kölcsönhatási felületeket és kulcsfontosságú aminosav reziduumokat vizsgáltunk. Ez lehetővé tette a modellek összehasonlítását a BCOR-PCGF1 dimer kísérleti szerkezetével, valamint a PRODIGY alkalmazását a releváns szekvenciákra. A fehérjeszerkezetek vizualizációjához a PyMOL szoftvert használtuk [186]. A RAWUL domén 178-184-es szakaszait kihagytuk az elemzésből, mivel ez a hurok struktúra hiányzik a kristályszerkezetből (PDB ID: 4HPL). Bár ez a régió nem lép kölcsönhatásba a PUF_D doménnel, a konzisztencia érdekében eltávolítottuk a modellezett komplexekből PyMOL segítségével [186], ahol a 4HPL kontrollként szerepelt. További kontrollként BCOR-PUF_D vad típusú dimereket generáltunk a 4HPL kristályszerkezet alapján AF3 felhasználásával. A BCOR, PCGF1, KDM2B és SKP1 heterotetramereket teljes hosszúságú szekvenciákkal (BCOR 1-1755, PCGF1 1-259, KDM2B 1-1336, SKP1 1-163) modelleztük és elemeztük. A BCOR-2 izoforma jelenléte esetén a kanonikus fehérjének megfelelő izoformszekvenciát használtuk. A kontakttérképet a MAPIYA webszerverrel készítettük 5.5 Å távolságra vonatkozó küszöbérték alkalmazásával. A MAPIYA az aminosav párok közötti kölcsönhatásokat a távolság függvényében határozza meg a fehérjeszerkezet alapján. A megadott küszöbértéken belüli reziduumok lehetővé teszik a fehérje szerkezeti szerveződésének és kölcsönhatási hálózatának vizualizációját [187]. A PDB 4HPL szerkezetet az RCSB Protein Data Bank-ból töltöttük le [188].

5.2.5 Molekuláris dinamikai szimulációk

A dimer szerkezetek vizsgálatára 10 ns hosszúságú molekuláris dinamikai (MD) szimulációkat végeztünk a GROMACS szoftvercsomag segítségével [189–196]. A kilenc BCOR-PCGF1 dimer (egy vad típusú és nyolc génfüzió) az AF által generált 50 replikátumából egyetlen szerkezetet választottunk ki, a legmagasabb ipTM (interface predicted template modelling) és pTM (predicted template modelling) pontszámok alapján, biztosítva, hogy a legstabilabb és legmegbízhatóbb konformációkat elemezzük. A BCOR fehérjék és a PCGF1 közötti kötődési affinitásokat a gmx_MMPBSA program segítségével számítottuk ki, amely a molekulamechanika

energiákat kombinálja Poisson-Boltzmann és felületi kontinuum szolvatációs (MM-PBSA) módszert alkalmazva, hogy nagy pontossággal becsülje meg az interakciós energiákat [197,198].

5.3 Statisztikai elemzés

A statisztikai elemzéseket a GraphPad Prism 8.0.1 és 9.5.1 szoftverrel végeztük (GraphPad Software, San Diego, CA, USA) [199]. A statisztikai próbák kiválasztása minden esetben az adatok típusának (folytonos, kategoriális), az összehasonlítandó csoportok számának (kettő vagy több), a csoportok függetlenségének vagy összetartozásának, valamint az adatok eloszlásának gondos mérlegelésén alapult. Minden összehasonlító statisztikai elemzés előtt megvizsgáltuk az adatok eloszlását és a varianciák homogenitását. A GraphPad Prism beépített diagnosztikai eszközeit használtuk, beleértve a normalitásvizsgálatokat (pl. Shapiro-Wilk, Anderson-Darling, D'Agostino & Pearson, Kolmogorov-Smirnov tesztek), valamint a reziduálisok diagnosztikáját. A szoftver lehetőséget biztosít annak megadására, hogy feltételezzük-e a Gauss-eloszlást és az egyenlő szórást. Amennyiben a normalitás és/vagy a szórások egyenlőségének feltétele nem teljesült, és ezt a szoftver diagnosztikája is alátámasztotta, a megfelelő nem-parametrikus tesztek vagy az ANOVA robusztusabb változatait (pl. Brown-Forsythe ANOVA) választottuk.

Annak vizsgálatára, hogy van-e különbség a TP53 fehérje patogénitására vonatkozó patogénitási pontszámok (különböző adatbázisokból és predikciós eszközökből származó pontszámok, stabilitási értékek) között a három független klinikai csoportban (AML-MR, MDS-IB, MDS-LB), egyszempontos varianciaanalízist (ordinary one-way ANOVA) alkalmaztunk, amennyiben az adatok csoportonként normális eloszlást és homogén varianciát mutattak. Szignifikáns ANOVA eredmény esetén Tukey többszörös összehasonlítási tesztet használtunk annak azonosítására, hogy mely specifikus csoportátlagok különböznek egymástól szignifikánsan. Amennyiben a varianciák homogenitásának feltétele sérült, a Brown-Forsythe ANOVA tesztet alkalmaztuk, amelyet Dunnett T3 többszörös összehasonlítási teszt követett, mivel ez utóbbi nem feltételezi az egyenlő varianciákat. Ha az adatok eloszlása jelentősen eltért a normalitástól, a nem-parametrikus Kruskal-Wallis tesztet választottuk, amit szignifikáns eredmény esetén Dunn többszörös összehasonlítási teszt követett a csoportok páronkénti összehasonlítására, korrigálva a többszörös tesztelésből adódó hibát.

Korrelációs mátrixot készítettünk Pearson-féle korrelációs együttható (r) felhasználásával annak meghatározására, hogy milyen lineáris kapcsolat van a stabilitást előrejelző módszerek által generált folytonos változók között (I-Mutant2.0, DynaMut2, DDGun és DDGun3D). A Pearson-korreláció alkalmazása előtt meggyőződünk a változók közötti legalább közelítőleg lineáris kapcsolatról és a normalitásról.

Az intrinzikusan rendezetlen régiók (IDRs) IUPred3 pontszámainak összehasonlítására, ahol ugyanazon fehérjéken több, összetartozó (nem független) mérést hasonlítottunk össze, a nem-parametrikus Friedman-próbát (az ANOVA nem-parametrikus megfelelője összetartozó minták esetén) alkalmaztuk. Szignifikáns eredmény esetén a páronkénti különbségek azonosítására Dunn többszörös összehasonlítási tesztet használtunk.

A BCOR-PCGF1 és PRC1.1 komplexek kötési affinitásértékeinek összehasonlítására a különböző fúziós fehérjék és a vad típusú kontrollcsoportok között a fentebb már részletezett elvek alapján, az adatok előzetes vizsgálatát (normalitás, szórás egyenlőség) követően ordinary one-way ANOVA-t (Tukey post-hoc teszttel), Brown-Forsythe ANOVA-t (Dunnett T3 post-hoc teszttel), vagy Kruskal-Wallis tesztet (Dunn post-hoc teszttel) használtunk [200–202]. A választás minden esetben a GraphPad Prism által végzett diagnosztikai tesztek eredményein alapult.

A $P < 0,05$ értéket tekintettük statisztikailag szignifikánsnak minden elvégzett teszt esetében. A grafikonok és ábrák elkészítéséhez a GraphPad Prism 9.5.1 verzióját [202] és az IBS (Illustrator of Biological Sequences) szoftvert használtuk [203].

6. Eredmények

6.1 A TP53 molekuláris és *in silico* fehérjeelemzése mielodiszpláziás neopláziákban és akut mieloid leukémiában

6.1.1 Klinikopatológiai jellemzők

A betegek átlagéletkora 64,1 év (tartomány: 25-90) volt. Az alcsoportok átlagéletkora a következő volt: AML-MR esetén 64,7 év (tartomány: 33-88), MDS-IB esetén 63,1 év (tartomány: 25-90), valamint MDS-LB esetén 63,9 év (tartomány: 25-89). A női/férfi arány: 42/35 (AML-MR: 13/13, MDS-IB: 4/8, MDS-LB: 25/14). A női betegek átlagéletkora 68,9 év (AML-MR: 47-88), 62,5 év (MDS-IB: 25-90) és 68,9 év (MDS-LB: 43-89) volt, míg a férfiak átlagéletkora 60,6 év (AML-MR: 33-78), 63,5 év (MDS-IB: 48-90) és 55 év (MDS-LB: 25-77) volt. A TP53 IHC, az NGS, a citogenetikai eredmények és az ösztülélés (OS) adatait az **1. táblázat** foglalja össze.

1. táblázat. A betegek klinikopatológiai eredményei és az ösztülélés (OS) adatai. AML-MR: akut mieloid leukémia mielodiszpláziával összefüggő eltérésekkel, MDS-IB: mielodiszpláziás neoplázia emelkedett blasztarány mellett, MDS-LB: mielodiszpláziás neoplázia alacsony blasztarány mellett, CK: komplex kariotípus. Túlélési státusz: 1=elhunyt, 0=követés megszakadt.

Esetek	Nem	Életkor (Év)	Diagnózis	OS (napok)	Túlélési státusz	TP53 IHC*	TP53 NGS**	Kariotípus
1	N	65	AML-MR	0	0	+	+	CK
2	F	73		119	1	+	+	n.a.
3	F	35		169	1	-	-	CK
4	N	47		2506	0	-	+	normál
5	F	76		1244	1	-	+	CK
6	F	64		308	1	+	+	n.a.
7	F	69		171	0	-	-	CK
8	N	76		3	0	-	-	n.a.
9	F	33		131	1	+	-	normál
10	F	78		153	1	-	+	CK
11	N	53		291	1	+	+	CK
12	F	64		822	1	-	-	normál
13	N	73		70	1	-	+	tetraszómia
14	N	88		95	1	-	-	normál
15	N	83		1	0	+	+	CK
16	F	66		174	1	-	+	CK
17	F	61		224	0	-	+	normál
18	N	64		186	1	-	-	47,XX,+8
19	N	66		341	1	-	-	CK
20	F	63		19	0	+	+	CK
21	N	86		224	1	+	+	n.a.
22	F	41		58	0	-	-	normál

23	N	56		613	1	-	-	CK
24	F	65		10	1	+	+	CK
25	N	69		16	1	+	+	CK
26	N	69		1398	1	-	-	normál
27	F	48		89	1	-	-	normál
28	F	56		32	0	-	+	n.a.
29	F	59		85	0	-	+	CK
30	N	69		989	0	-	-	normál
31	N	25		260	1	-	-	46,XX,t(8;21)/46,XX
32	F	80	MDS-IB	20	1	+	+	CK
33	F	61		199	0	-	-	normál
34	F	79		122	1	+	+	n.a.
35	F	59		49	0	-	-	CK
36	N	90		97	0	-	-	normál
37	F	65		309	0	-	-	normál
38	N	66		522	1	-	-	normál
39	N	52		626	0	-	+	normál
40	N	80		312	1	-	-	normál
41	N	71		494	1	-	+	46,XX,t(2;12)/46,XX,t(17;17)
42	F	64		2296	1	-	+	normál
43	N	64		754	1	-	-	n.a.
44	F	70		450	1	-	-	normál
45	F	64		918	0	-	-	normál
46	F	41		294	1	-	+	n.a.
47	N	74		710	1	-	-	46,XX,del(20q)
48	N	61		75	0	-	+	normál
49	N	58		1086	1	-	-	normál
50	N	60		993	1	-	+	normál
51	F	36		274	1	-	-	normál
52	F	60		175	1	-	-	normál
53	F	72		513	1	-	-	normál
54	N	71	MDS-LB	1105	0	-	-	normál
55	N	78		1559	0	-	-	46,XX,del(5q)/46,XX
56	N	64		20	1	-	-	n.a.
57	N	89		541	1	-	-	n.a.
58	F	68		1101	1	-	-	CK
59	F	36		203	0	-	-	n.a.
60	N	64		2137	1	-	+	normál
61	F	77		1144	0	-	-	CK
62	N	43		726	1	-	-	normál
63	N	74		947	0	-	-	n.a.
64	F	70		680	0	-	-	normál
65	F	30		135	1	-	-	normál
66	F	25		635	0	-	-	normál
67	N	84		801	0	-	-	normál
68	N	76		1245	1	-	-	normál
69	N	75		1183	0	-	-	normál

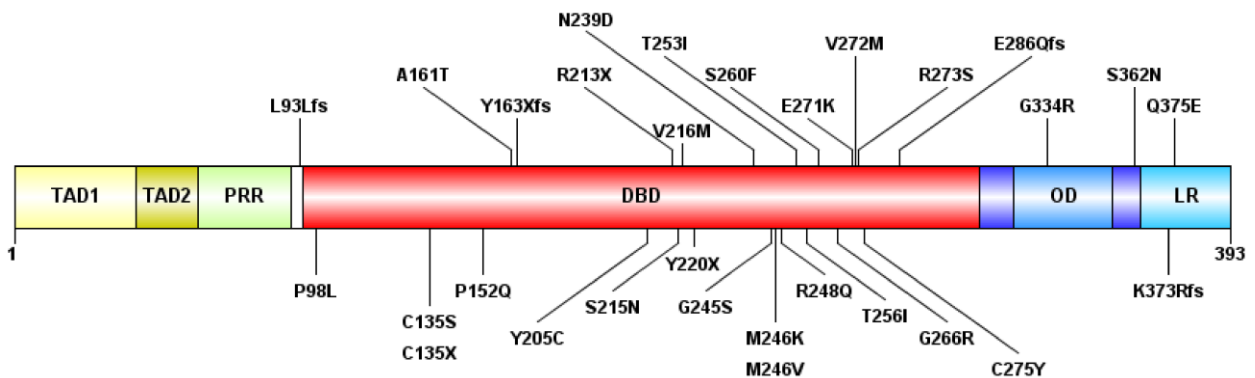
70	N	69	988	1	-	-	normál
71	N	73	957	1	-	-	normál
72	F	58	397	0	-	-	normál
73	N	85	108	0	-	-	46,XX,del(7q)
74	N	46	358	0	-	-	46,XX,del(5q)/46,XX
75	N	71	206	1	+	-	normál
76	N	67	317	0	-	-	normál
77	N	73	117	1	-	-	normál

* A TP53 IHC pozitivitást akkor definiáltuk, ha a TP53 festődés intenzitása magas (3+) és a pozitív sejtek aránya legalább 10% volt.

** A TP53 NGS pozitivitás olyan eset, ahol variánst legalább 5%-os variáns allél frekvencia és 250X lefedettség jellemezte.

6.1.2 Újgenerációs szekvenálás

A 77 vizsgált esetből 26-ban legalább egy TP53 mutációt azonosítottunk, összesen 41 mutációt detektálva (**1. ábra**), amelyek 30 különböző genotípust foglaltak magukban. Csoportonként az alábbi megoszlást figyeltük meg: az AML-MR minták közül 15 a 26-ból (57,69%), az MDS-IB minták közül 4 a 12-ből (33,33%), és az MDS-LB minták közül 7 a 39-ből (17,95%) mutatott TP53 génelteréseket. Hét esetben egyetlen mintában két vagy több mutációt is találtunk: ötöt az AML-MR csoportban, egyet az MDS-IB csoportban és egyet az MDS-LB csoportban. Az átlagos VAF összességében 22,75% volt; csoportonkénti bontásban az AML-MR esetében 34,12%, az MDS-IB esetében 35,59%, míg az MDS-LB esetében 22,83% volt. Az NGS eredményeit a **2. táblázat** foglalja össze.



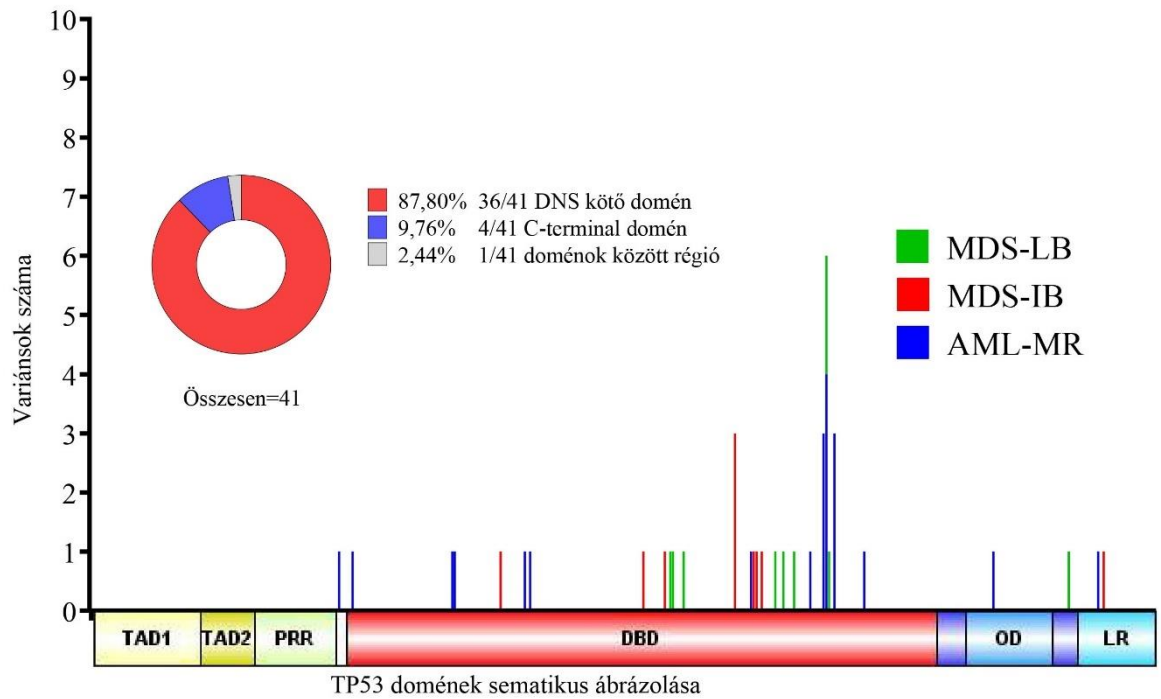
1. ábra. A TP53 fehérje szerkezetének sematikus ábrázolása és a kimutatott mutációk. Domének: transzaktivációs domén 1 (TAD1) (1-40 aminosav, AA), TAD2 (40-60 AA), prolinban gazdag régió (PRR) (60-90 AA), DNS-kötő domén (DBD) (94-312 AA), oligomerizációs domén (OD) (323-355 AA), valamint lizinben gazdag C-terminális szakasz (LR) (364-393 AA). Az „fs” a frameshift típusú aminosav-változást, az „X” pedig a stop kodont jelöli.

A 30 egyedi mutációból 25 (83,3%) a DBD-ben volt megtalálható (**1. ábra** és **2. ábra**). A 30 különböző TP53 mutáció közül 23 (77%) misszensz variáns, négy (13%) frameshift variáns, valamint három (10%) stop kodont eredményezett. A hét nem-misszensz mutációból hat esetben a fehérjetermék hossza csökkent, ami trunkált fehérje képződését eredményezte (**3. ábra**).

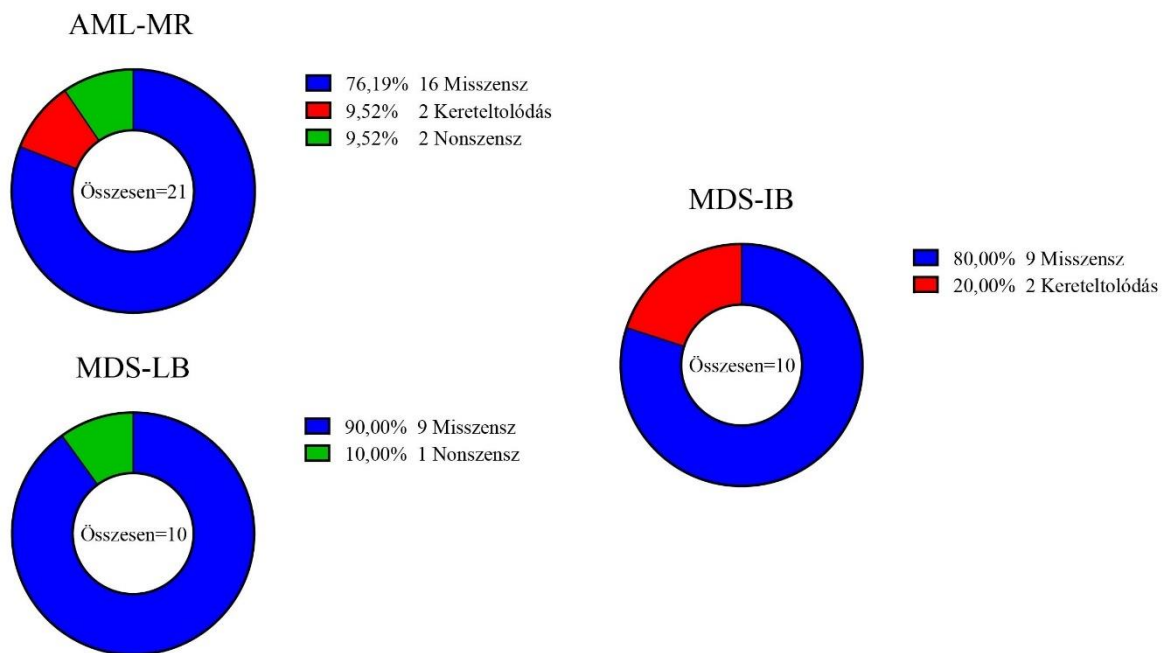
2. táblázat. A kimutatott *TP53* mutációk és variáns allélfrekvenciák (VAF).

Esetek	Diagnózis	<i>TP53</i> nukleotid variánsok	<i>TP53</i> AA variánsok	VAF (%)
1	AML-MR	c.811G>A	p.E271K	27,8
		c.715A>G	p.N239D	29,14
2		c.824G>A	p.C275Y	88,04
4		c.779C>T	p.S260F	16,2
5		c.279delG	p.L93Lfs	36,66
6		c.814G>A	p.V272M	30,7
		c.660T>G	p.Y220X	8,38
10		c.811G>A	p.E271K	6,08
		c.715A>G	p.N239D	6,67
11		c.736A>G	p.M246V	28,39
13		c.811G>A	p.E271K	8,28
		c.737T>A	p.M246K	8,35
15		c.405C>A	p.C135X	5,65
		c.817C>A	p.R273S	41,02
16		c.856_869del14	p.E286Qfs	16,96
17		c.824G>A	p.C275Y	12
20		c.796G>A	p.G266R	41,73
21		c.481G>A	p.A161T	73,42
24		c.403T>A	p.C135S	45,06
25		c.646G>A	p.V216M	39,44
	c.733G>A	p.G245S	38,22	
28	MDS-IB	c.644G>A	p.S215N	30,65
		c.1118delA	p.K373Rfs	11,59
		c.767C>T	p.T256I	6,82
		c.293C>T	p.P98L	6,25
		c.1123C>G	p.Q375E	5,6
		c.758C>T	p.T253I	5,45
29	c.489delC	p.Y163Xfs	70,27	
32	MDS-LB	c.743G>A	p.R248Q	39,93
		c.455C>A	p.P152Q	38,21
34	c.614A>G	p.Y205C	13,49	
39	c.824G>A	p.C275Y	12,2	
41	c.814G>A	p.V272M	13,06	
42	MDS-LB	c.1085G>A	p.S362N	5,19
		c.1000G>A	p.G334R	5,04
		c.814G>A	p.V272M	6,27
		c.715A>G	p.N239D	10,81
46	c.814G>A	p.V272M	10,42	
48	c.814G>A	p.V272M	7,52	
50	c.814G>A	p.V272M	19,77	
60	c.637C>T	p.R213X	5,94	

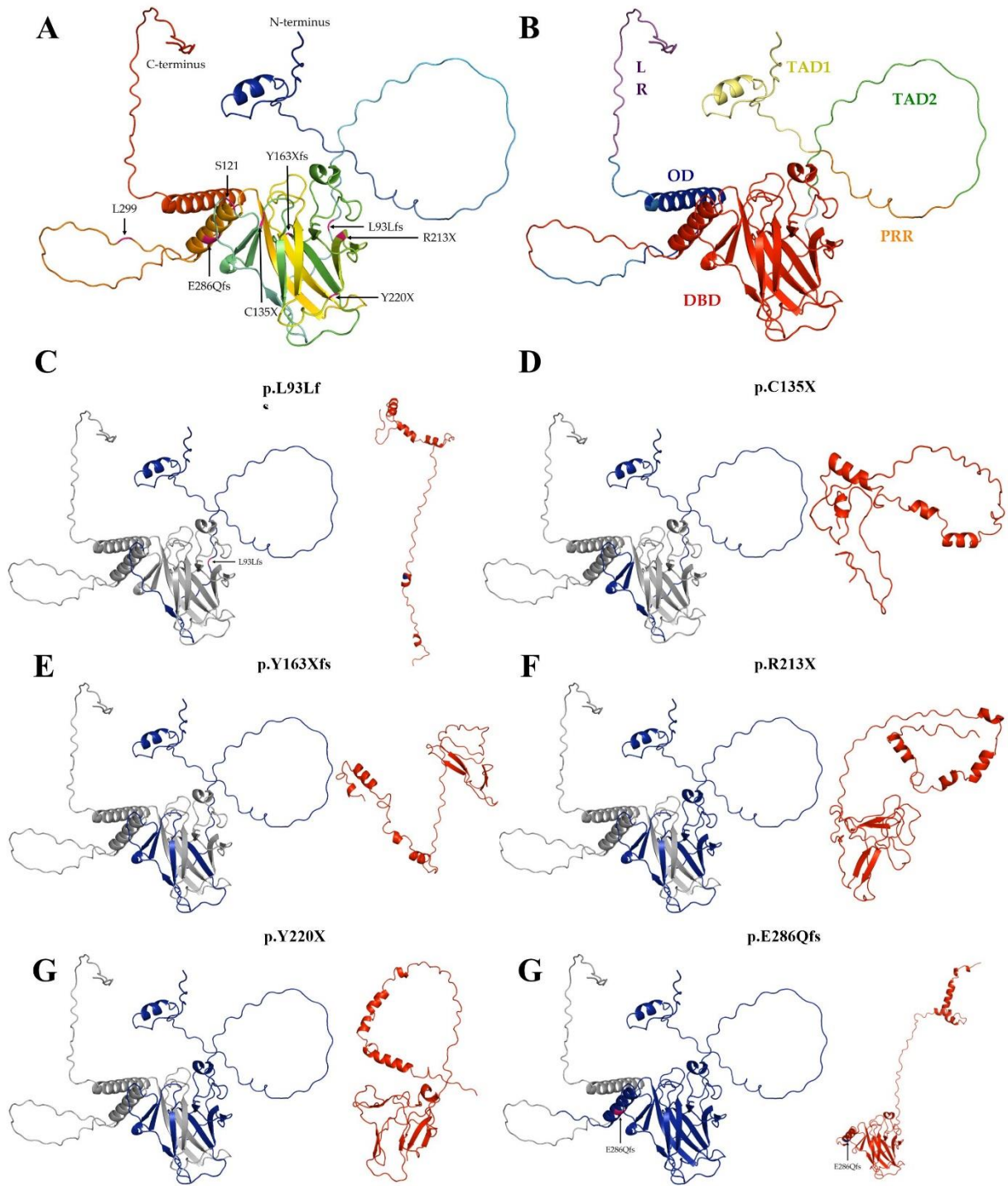
A



B



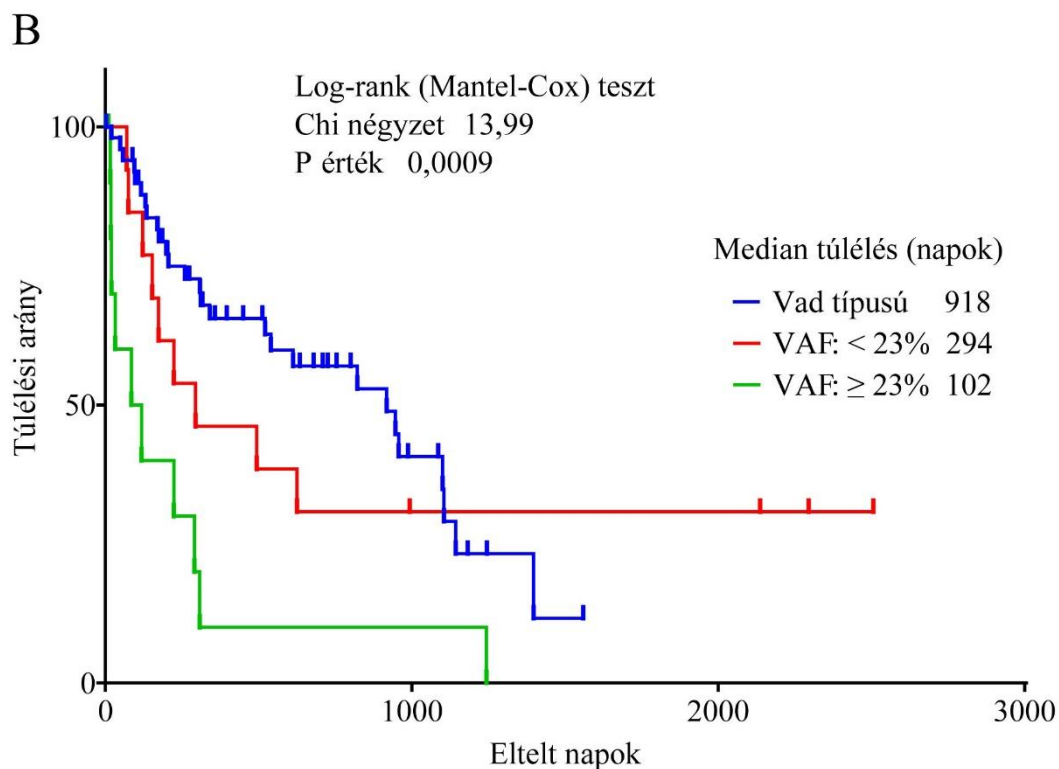
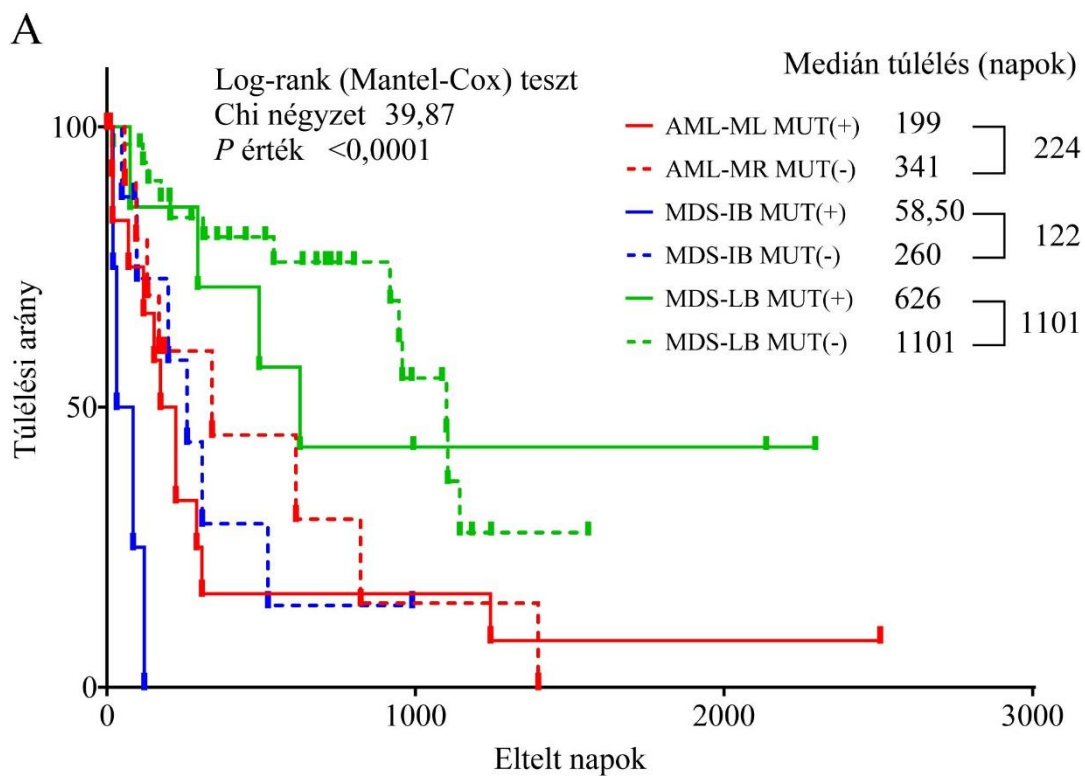
2. ábra. A *TP53* génelterések azonosítása újgenerációs szekvenálással (NGS). **A:** A mutációk gyakorisága a három betegcsoportban. A *TP53* fehérje szerkezetén jelöltük a kimutatott eltérések helyét. **B:** A mutációk megoszlása az alcsoportok között.



3. ábra. A trunkált TP53 fehérjék szerkezetének Robetta-alapú predikciója. **A:** A vad típusú TP53 fehérje harmadlagos szerkezete (AlphaFold). **B:** A vad típusú fehérje doménszerkezetének ábrázolása. **C-H:** A vad típusú fehérje hossz kék színnel, a mutáns (piros) pedig a rövidült részek hosszáig van színezve: **C:** p.L93Lfs (121 AA), **D:** p.C135X (134 AA), **E:** p.Y163Xfs (162 AA), **F:** p.R213X (212 AA), **G:** Y220X (219 AA), **H:** p.E286Qfs (299 AA).

6.1.3 A citogenetikai, az IHC eredmények, a túlélés és a *TP53* mutációs státusz közötti kapcsolat vizsgálata

A *TP53* mutáns esetek közül 11/26 (42,3%) esetben komplex kariotípus (CK) volt kimutatható. A 15 *TP53* mutáns AML-MR beteg közül 9-nél (60%) igazolódott CK, míg az MDS-IB csoportban 4/2 (50%) esetben, az MDS-LB mutáns pozitív csoportban pedig egyetlen esetben sem észleltek CK-t (a citogenetikai vizsgálat 12 esetben nem történt meg). Az AML-MR csoportban 15/26 (57,7%) citogenetikai eltérést detektáltunk, míg az MDS csoportokban a kromoszómaeltérések alacsonyabb gyakorisággal fordultak elő (MDS-IB: 4/12 - 33,3%; MDS-LB: 7/38 - 18,4%). Összesen 13 esetben (16,9%) értékeltük pozitívnak az IHC eredményt: 10 (38,5%) az AML-MR csoportban, 2 (16,7%) az MDS-IB csoportban, és 1 (2,6%) az MDS-LB csoportban. Az összes eset 14,3%-ában IHC és NGS pozitivitás együtt mutatkozott, míg az összes NGS mutáns eset 42%-a volt IHC pozitív. Az AML-MR csoportban 15 NGS mutáns közül 9 (60%), az MDS-IB csoportban 50% volt IHC pozitív, az MDS-LB betegeknél pedig egyetlen IHC és NGS pozitív eset sem került kimutatásra. A hat trunkált fehérjét eredményező minta közül öt esetben negatív IHC eredményt kaptunk. A p.R213X, p.Y163Xfs, p.E286Qfs és p.L93Lfs mutációk (utóbbi egy mintában két másik mutációval együtt) IHC negatívnak bizonyultak, míg a p.Y220X mutáció (egy másik mutációval párhuzamosan) IHC pozitív volt. A medián össztúlélés (OS) szignifikáns különbséget mutatott a három csoport között ($P \leq 0,0001$) a *TP53* mutáns/vad típusú állapot függvényében (**4. ábra**). A vad típusú esetek medián OS-értéke 918 nap volt (AML-MR: 341, MDS-IB: 260, MDS-LB: 1101 nap), míg a mutáns betegeknél 224 nap (AML-MR: 199, MDS-IB: 58,5, MDS-LB: 626 nap).

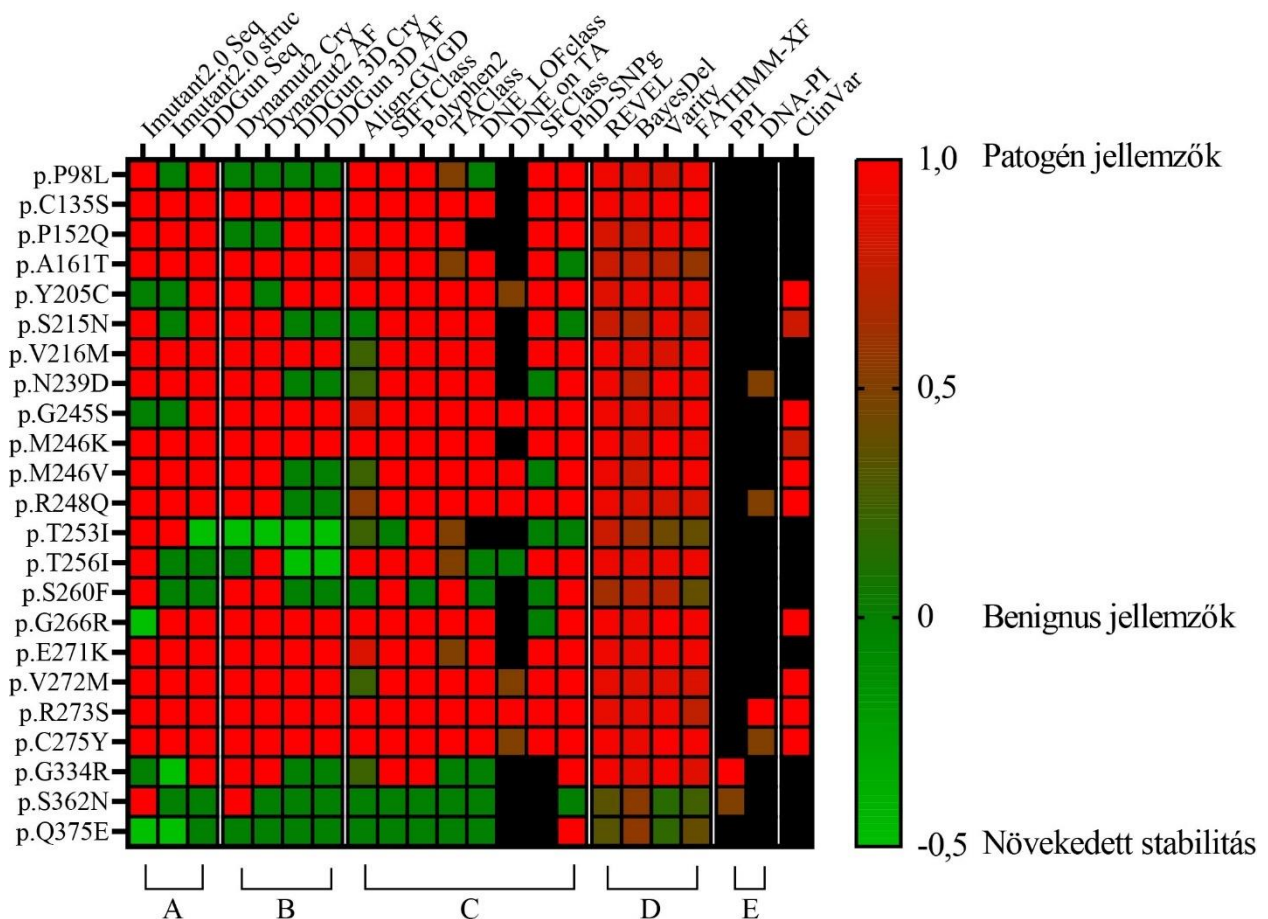


4. ábra. A *TP53* mutáns és vad típusú AML/MDS betegek túlélése. A vízszintes tengely a túlélési időt napokban, a függőleges tengely a túlélési arányt (%) ábrázolja. **A:** A betegek hat csoportokba soroltuk a mutációs állapot és diagnózis (AML/MDS) alapján. **B:** A betegek csoportosítása a variáns allélfrekvencia (VAF) szerint. VAF >23% esetén rosszabb a túlélés a vad típusú betegekhez képest, míg a VAF ≤23% mutáns esetek túlélése hasonló volt a vad típusú kontrollcsoportéhoz [204].

6.1.4 A mutációk patogenitásának *in silico* predikciója

A hat trunkált fehérjét (3. ábra) LOF-nak tekintettük a C-terminális domén hiánya miatt. A stabilitási vizsgálatok teljes körű elvégzése és a patogenitás pontos meghatározása nem volt lehetséges. Bár a p.K373Rfs hossza azonos a vad típusú fehérjével, a szekvencia jelentősen eltér a frameshift mutáció miatt, így a stabilitást és patogenitást értékelő rendszerünkben nem volt összehasonlítható a többi misszensz SNV mutánszal.

A 23 különböző mutáció klinikai relevanciáját először a ClinVar adatbázisban vizsgáltuk, amely szerint hét mutációt patogénnek, kettőt valószínűleg patogénnek, egyet patogénnek/valószínűleg patogénnek klasszifikált, míg 11 esetben az értékelés bizonytalan maradt (hat esetben „bizonytalan jelentőségű”, öt esetben „ellentmondásos értelmezések a patogenitásról”). Két mutáció (p.S260F és p.Q375E) nem szerepelt az adatbázisban. Az adatok bizonytalansága miatt a 23 mutációt további elemzésnek vetettük alá különböző adatbázisok és pontozási rendszerek segítségével (5. ábra).



5. ábra. A kimutatott TP53 mutációk klinikai relevanciájának összehasonlítása. Az ábra a TP53 misszensz mutációk klinikai jelentőségét mutatja be különböző elemző módszerek alapján. Minden sor egy mutációnak felel meg, míg az oszlopok a fehérje-stabilitás előrejelzéseit (A, B), pontozási rendszereket (C, D) és interakciós affinitás-vizsgálatokat (E) ábrázolják. A színskála zöldtől (nem patogén) pirosig (patogén) változik, 0-1 közötti számértékekkel. A stabilitás-előrejelzéshez szekvencia-alapú (Seq), kristályszerkezeti (Cry) és AlphaFold által

modellezett (AF) módszereket alkalmaztunk. A stabilitásváltozást a mutációk ΔG értéke alapján kategorizáltuk: $<-0,5$ kcal/mol esetén csökkenő (piros), $>0,5$ kcal/mol esetén növekvő (világoszöld), egyébként semleges (zöld). A pontozási rendszerek közé tartozik az Align-GVG (C0-C65 skála), a SIFTClass (0 = benignus, 1 = patogén), a Polyphen2 (három kategória), valamint a TAClass (funkcionalitás alapján). A domináns-negatív hatás (DNE, LOFclass) és a transzaktivációs befolyás (DNE on TA) szintén 0-1 skálán van értékelve. A REVEL, BayesDel, Varity és FATHMM-XF pontok 0-1 közötti értékekkel jelzik a patogenitási potenciált. Az interakciós affinitás-vizsgálatoknál a fehérje-fehérje (PPI) és DNS-fehérje (DNA-PI) kölcsönhatások változásait mértük. A ClinVar adatbázis szerint csak a patogén vagy valószínűleg patogén mutációk kaptak számértéket (1 vagy 0,8), míg a bizonytalan vagy hiányzó adatok fekete négyzetekkel vannak jelölve. Az eredmények a TP53 adatbázisból és kiegészítő algoritmusokból (pl. Varity) származnak, hangsúlyozva a módszerek közötti egyezéseket és eltéréseket.

6.1.5 Az IARC TP53 adatbázis *in vitro* kísérleti eredményeinek analízise

Az adatbázis nyolc típusú promoter specifikus transzkripció aktivitást (TA) tartalmaz (WAF1, MDM2, BAX, h1433s, AIP1, GADD45, NOXA, P53R2), amelyeket élesztőben végzett funkcionális tesztekkel mérték és a vad típusú aktivitás százalékos értékében fejezték ki. A 30 típusú mutáció közül a p.L93fs, p.Y163fs, p.E286Qfs és p.K373fs esetében nem álltak rendelkezésre adatok. A 26 elérhető mutációt két csoportba oszthatjuk a doménszerkezet alapján (23 a DBD-ben és három a fehérje C-terminális részében). Átlagosan a DBD-ben található mutációk promoter specifikus TA-ja csupán 13,84% a vad típusú fehérjéhez képest (tartomány: 0,09% a p.C275Y esetében és 40,24% a p.P98L esetében). A C-terminális doménben detektált p.G334R, p.S362N és p.Q375E variánsok aktivitása 87,72%, 78,06% és 80,33% volt.

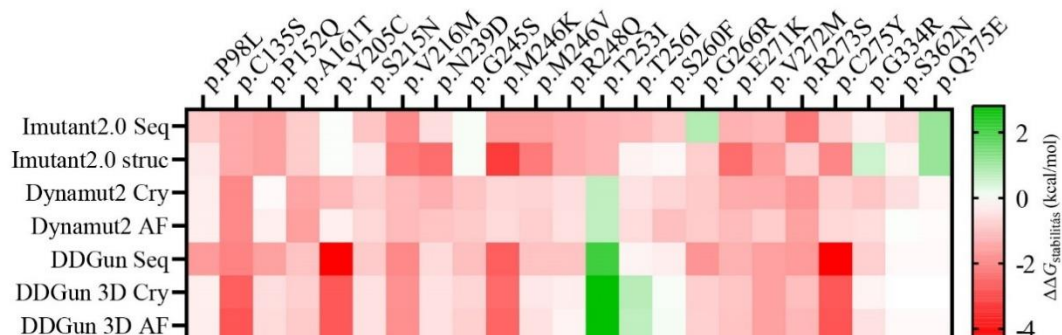
Az *in vitro* kísérletek alapján a p.C135X, a p.C135S, a p.A161T, a p.Y163fs, a p.Y205C, a p.R213X, a p.S215N, a p.V216M, a p.Y220X, a p.N239D, a p.G245S, a p.M246K, a p.M246V és a p.R248Q funkcióvesztő mutáció a növekedésgátlás szempontjából. Egy mutáció, a p.C135S, LOF az apoptózis indukciójában és funkciónyerő (GOF) a RAS vagy más transzformáló onkogénnel való kölcsönhatásban, hasonlóan a p.Y205C, a p.G245S, a p.M246V, a p.R248Q és a p.C275Y variánsokhoz. Domináns-negatív hatás hét mutáció (p.G245S, p.M246V, p.R248Q, p.G266R, p.V272M, p.R273S, p.C275Y) esetében igazolódott. Az adatbázis szerint a p.G245S, a p.R248Q, a p.G266R és a p.R273S mutációk funkciónyerőként viselkednek, és gátolják a p73 aktivitását, ha mindkét fehérje jelen van a sejtben.

6.1.6 Mutáns TP53 fehérjék stabilitás analízise

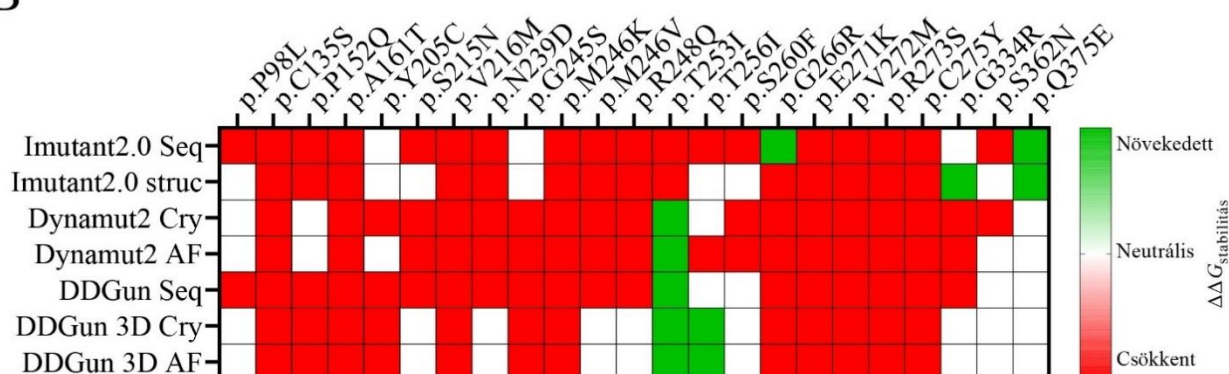
A TP53 fehérje stabilitására gyakorolt hatást *in silico* elemzések segítségével vizsgáltuk az I-Mutant2.0, DynaMut2 és DDGun eszközökkel (6. ábra). Az elemzést fehérje szekvencia (I-Mutant2.0 seq, DDGun seq) vagy szerkezet alapján végeztük, felhasználva a vad típusú fehérje kristályszerkezetét (I-Mutant2.0 Cry, DynaMut2 Cry, DDGun 3D Cry) és egy AF által generált modellszerkezetet (DynaMut2 AF, DDGun 3D AF).

A prediktált és a rendelkezésre álló kísérletes eredmények összehasonlítása (6. ábra) szerint a p.G245S esetén pontos egyezést kaptunk (DDGun 3D AF), míg a p.R248Q mutációnál a stabilitásváltozás csökkenése hasonló volt a kísérleti eredményhez.

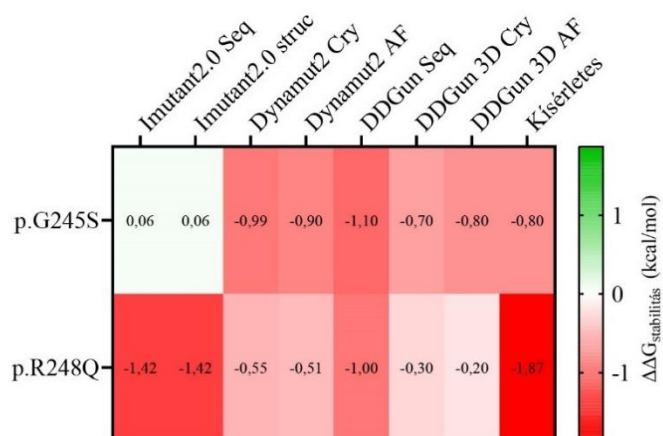
A



B

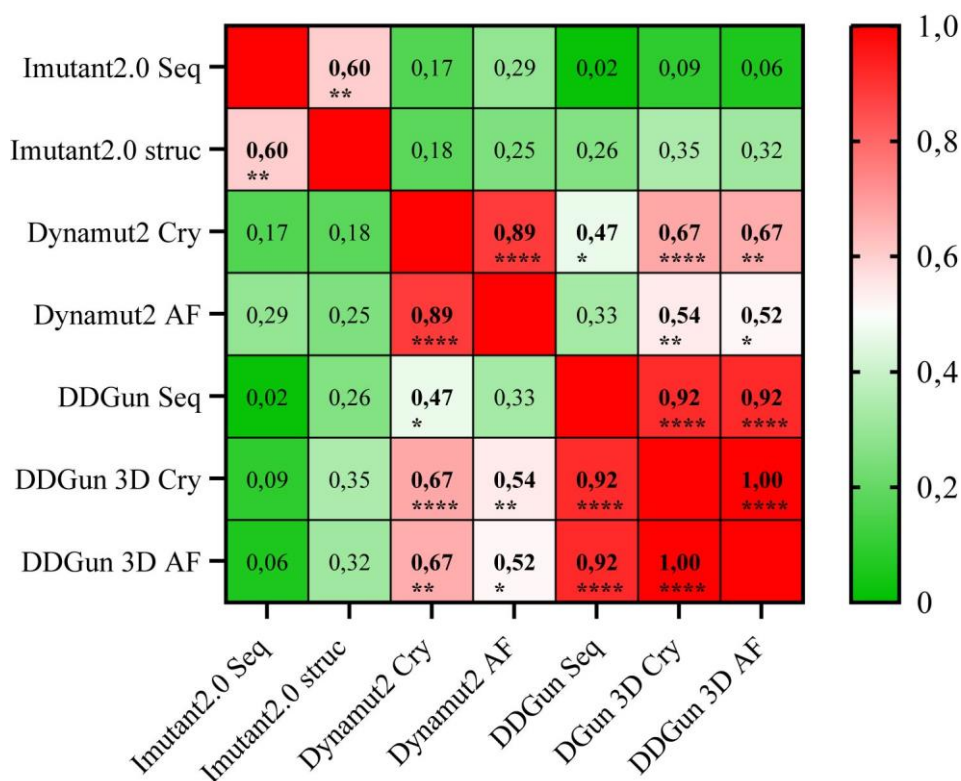


C



6. ábra. A mutáns TP53 fehérjék stabilitásváltozása ($\Delta\Delta G_{\text{stabilitás}}$). **A:** A számított eredmények piros (stabilitás csökkenése) és zöld (stabilitás növekedése) színnel jelölve. Az ábra mutatja, mely mutációkhoz tartozik csökkenő vagy növekvő $\Delta\Delta G_{\text{stabilitás}}$ érték kcal/mol mértékegységben. **B:** A stabilitás csökkent, ha $<-0,5$ kcal/mol (piros), növekedett, ha $>0,5$ kcal/mol (zöld), egyébként semlegesnek (fehér). **C:** Két kísérletes úton már meghatározott stabilitású mutáció (p.G245S [145] és p.R248Q [146]) és a predikciók összehasonlítása.

Hasonló eredményeket kaptunk a kristály (Cry) és modellszerkezetek (AF) elemzésével ugyanezen módszerekkel (DynaMut2, DDGun 3D), de szignifikáns eltérések voltak megfigyelhetők az egyes módszerek között (7. ábra). Ezért kiszűrtük azokat a módszereket, amelyek Pearson-féle korrelációs együtthatója $<0,5$ volt a mátrix többi tagjához képest (I-Mutant 2.0 seq és struc, DDGun seq). A Cry alapú predikciók és a teljes AF modell alapján végzett stabilitási becslések nagyon magas korrelációt mutattak (DynaMut2 Cry és AF között: $r = 0,89$; DDGun3D Cry és AF között: $r = 1,0$).



7. ábra. A hét predikciós módszer Pearson-féle korrelációs mátrixa. Az I-Mutant 2.0 elemzések egyike sem mutatott $>0,5$ Pearson-r értéket a többi módszerhez viszonyítva. A táblázat cellái a Pearson-r értékek szerint színezettek (színskála). * $P < 0,05$, ** $P < 0,01$, *** $P < 0,0001$, **** $P < 0,0001$

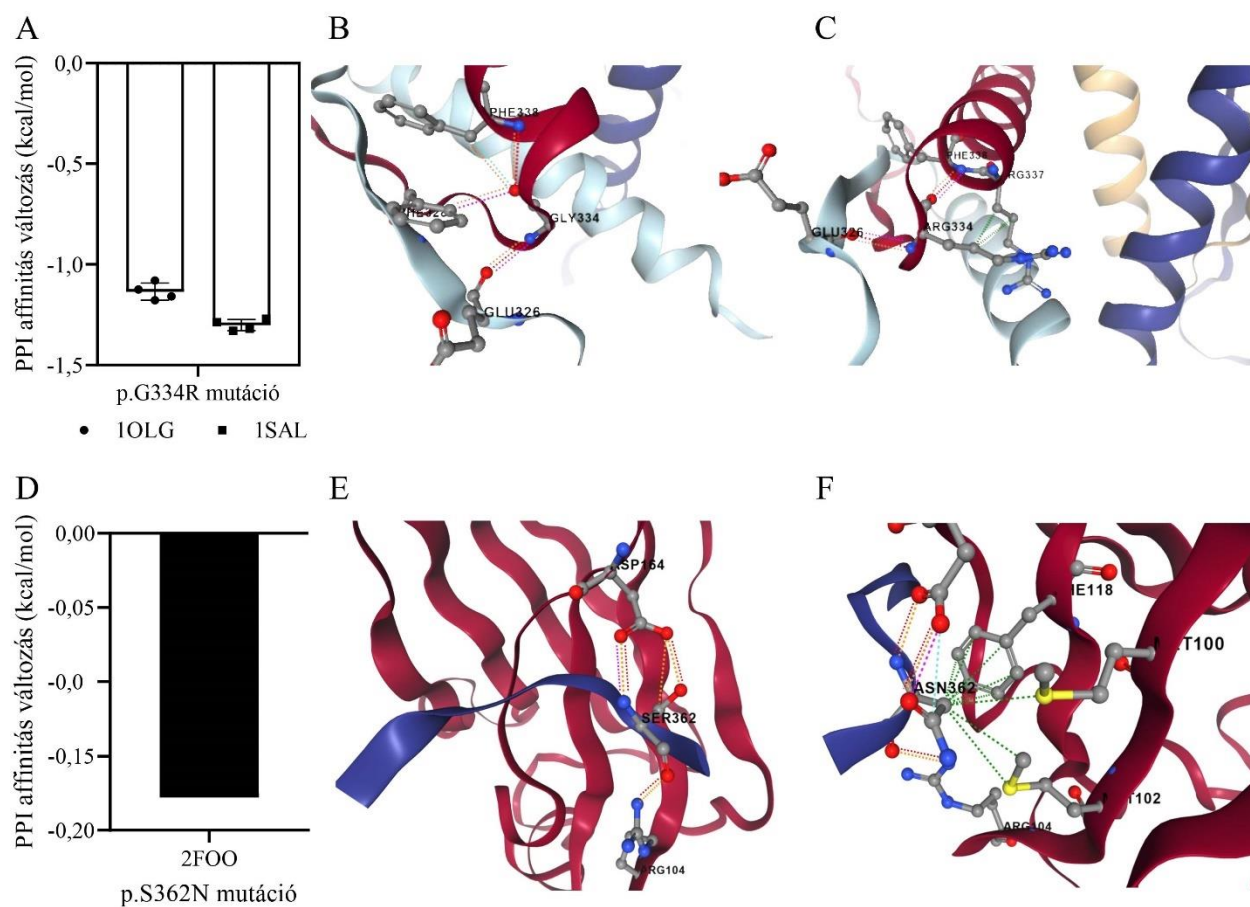
A maradék négy előrejelzés (DynaMut2, DDGun 3D eszközök) alapján szignifikáns különbségeket észleltünk a TP53 fehérjék stabilitásváltozásaiban ($\Delta\Delta G_{\text{stabilitás}}$) a három klinikai csoport (AML-MR, MDS-IB, MDS-LB) között one-way ANOVA statisztikai próba segítségével.

6.1.7 A mutáns TP53 fehérje-fehérje kölcsönhatások vizsgálata

A p.G334R mutáció (az MDS-LB csoportban) az oligomerizációs doménben helyezkedik el, és így befolyásolhatja a TP53 monomerek tetramerizációját. A mutáció hatását a tetramer szerkezet kialakításában részt vevő monomer kölcsönhatásokra az mCSM-PPI2 programmal vizsgáltuk. Ehhez két, tetramer konformációjú TP53 fehérje kristály szerkezetet használtunk

(PDB: 1OLG és 1SAL). Négy-négy lánc elemzése alapján csökkent affinitást (átlagosan -1,218 kcal/mol) prediktáltunk. Ez a mutáció potenciálisan gyengítheti a monomer kölcsönhatásokat az oligomerizáció során (**8. ábra**).

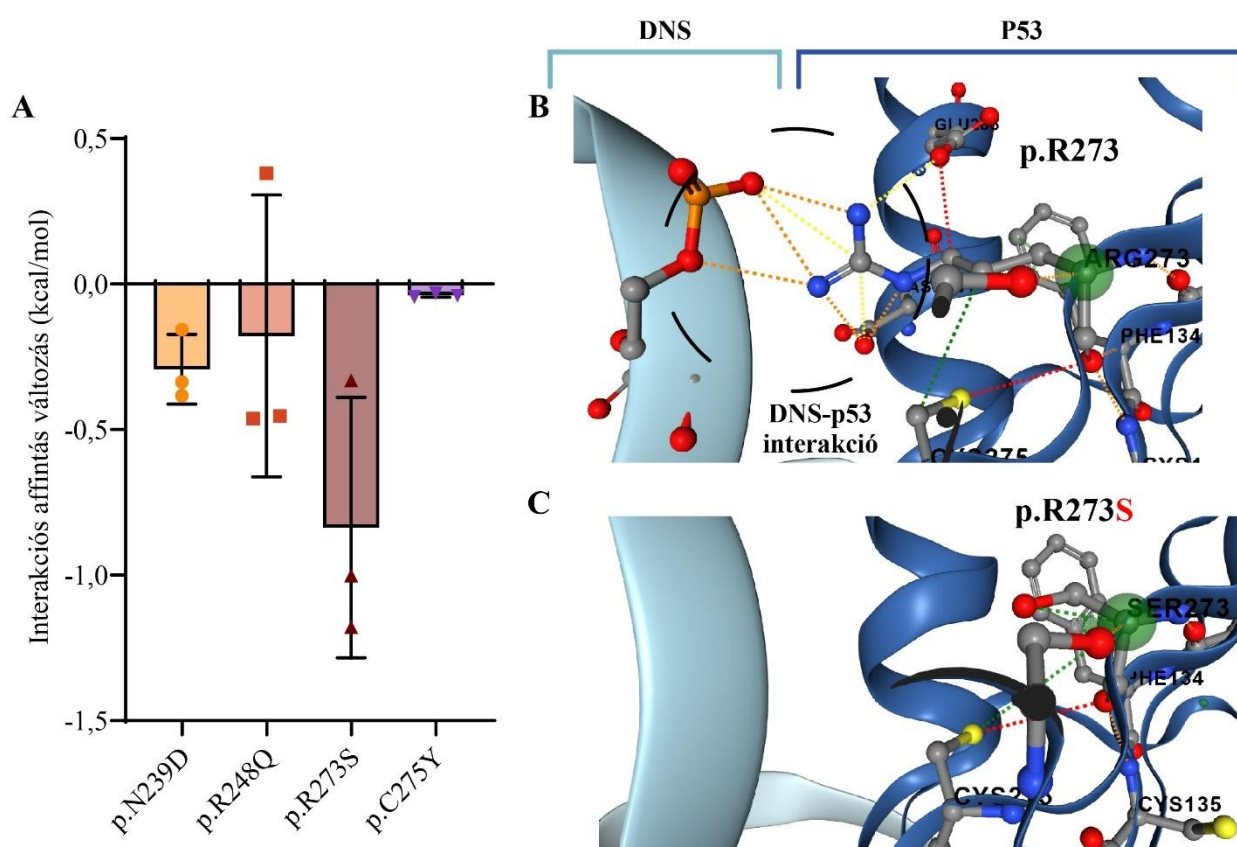
A p.S362N mutáció (MDS-LB csoport) a TP53 és a USP7/HAUSP fehérjék kölcsönhatási felületén (359-362) helyezkedik el. Annak vizsgálatára, hogy a mutáció befolyásolja-e ezt a kölcsönhatást, az mCSM-PPI2 programmal szimuláltuk az interakciós energia változását (-0,178 kcal/mol; **8. ábra**). A mutáció módosítja a két fehérje közti kölcsönhatás erősségét a kölcsönhatási felület változása miatt.



8. ábra. A fehérje-fehérje kölcsönhatások (PPI) változása a p.G334R (**A-C**) és a p.S362N (**D-F**) mutációk hatására. **A:** A kötési affinitás csökkenése a tetramer szerkezetben a p.G334R mutáció hatására. **D:** A kötési affinitás csökkenése a dimer szerkezetben a p.S362N mutáció következtében. Az 1OLG, az 1SAL és a 2FOO a vizsgált kristály szerkezetek PDB azonosítói. A p.G334R mutáció a tetramer TP53 monomer kölcsönhatásait befolyásolja (**A** és **C**), míg a p.S362N a TP53 és a USP7/HAUSP fehérjék kölcsönhatását módosítja (**D** és **F**). **B:** A vad típusú láncok közötti, **C:** a vad típusú és mutáns lánc közötti, **E:** a vad típusú fehérje és az USP7 közötti, **F:** a mutáns fehérje és az USP7 közötti PPI. A szaggatott vonalak az interakció típusát jelölik: rózsaszín = ütközés, világoskék = van der Waals, piros = hidrogénkötés, sárga = ionos kötés, zöld = hidrofób, narancs = poláris.

6.1.8 A TP53 fehérje-DNS kölcsönhatások vizsgálata a mutációk hatására

Négy olyan mutációt azonosítottunk (p.R273S az AML-MR csoportban; p.C275Y és p.N239D az AML-MR és MDS-LB csoportokban; p.R248Q az MDS-IB betegekben), amelyek közvetlenül befolyásolják a TP53 és a DNS közötti kölcsönhatást. Az mCSM-NA program segítségével három különböző TP53-DNS komplex kristály szerkezetet elemeztünk az intermonomer kölcsönhatás változások előrejelzéséhez (9. ábra). Szinte minden esetben a mutációk gyengítették a TP53 és a DNS közötti kölcsönhatást. A legnagyobb csökkenést a p.R273S mutációnál (-0,8370 kcal/mol), míg a leggyengébb változást a p.C275Y mutációnál (-0,03867 kcal/mol) mértük.



9. ábra. A TP53 és a DNS közötti kölcsönhatások változása a mutációk hatására. **A:** Kötési affinitás változás a talált mutációk esetében. **B:** DNS-fehérje kölcsönhatás (fekete szaggatott kör) a vad típusú p.R273 (zöld jelölés) és a DNS között. **C:** A DNS-fehérje kölcsönhatás hiánya a p.R273S mutáns fehérje (zöld jelölés) és a DNS között.

6.1.9 A mutáns TP53 fehérje patogénitási és szerkezeti stabilitásának statisztikai elemzése AML/MDS betegekben

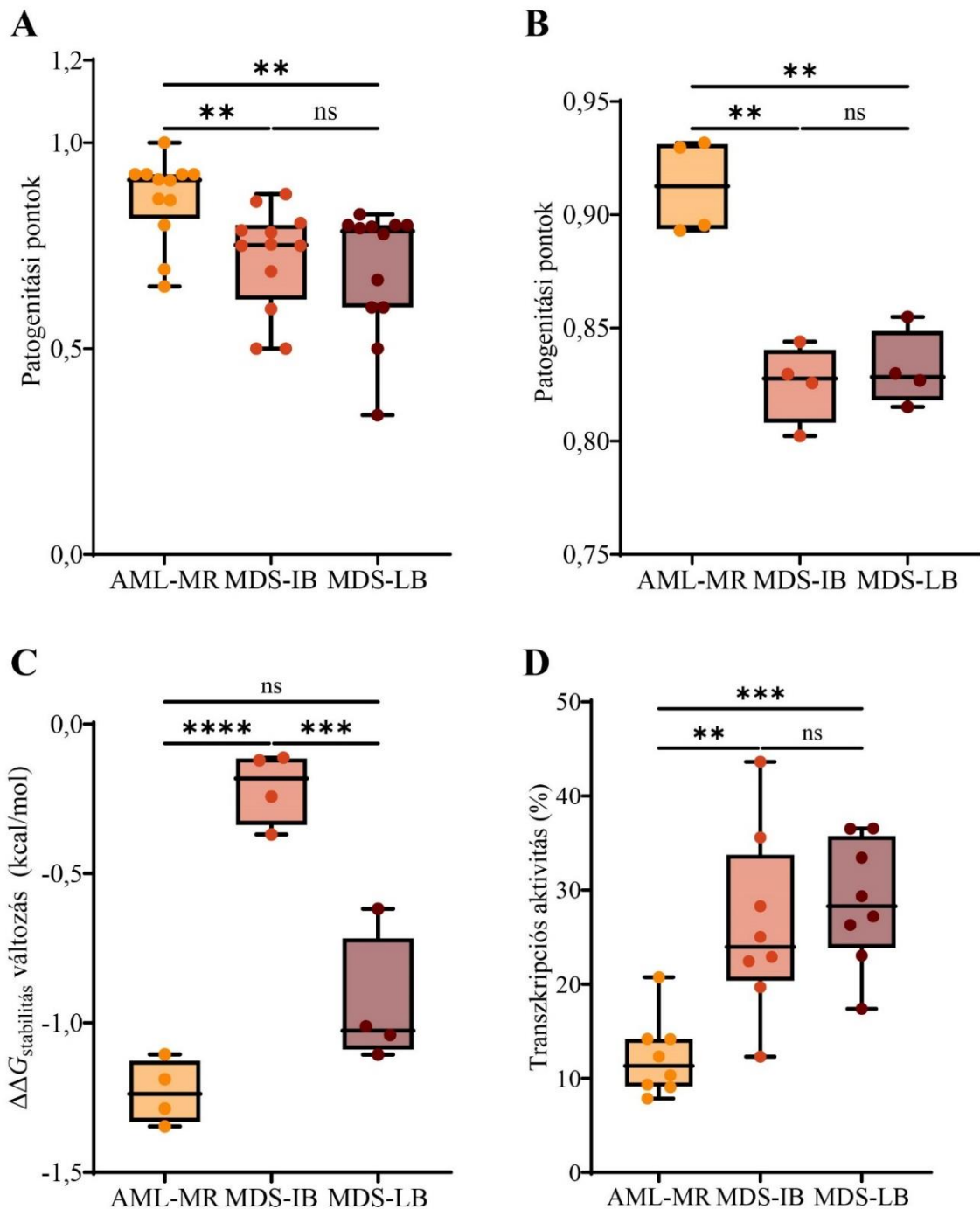
A három csoport (AML-MR, MDS-IB, MDS-LB) közötti különbségek értékelésére Kruskal-Wallis tesztet végeztünk az adatbázisokon alapuló pontozási rendszerek alapján (5. ábra), amely szignifikáns különbséget mutatott legalább két csoport között ($P = 0,0021$). A Dunn-féle többszörös összehasonlítási teszt szignifikáns eltérést adott az AML-MR és az MDS-LB ($P =$

0,0081), illetve az AML-MR és az MDS-IB ($P = 0,0061$) csoportok között. Nem volt szignifikáns különbség az MDS-IB és az MDS-LB csoportok között ($P > 0,9999$) (**10. ábra**).

Az elemzést négy olyan módszerrel végeztük, ahol minden mutációhoz elérhetőek voltak az adatok, és a patogenitási osztályozás eredetileg nem volt egyértelműen meghatározott (**5. ábra**). A REVEL, BayesDel, Varity és FATHMM-XF pontozások alapján Brown-Forsythe and Welch ANOVA tesztet alkalmaztunk a patogenitási státusz összehasonlítására, mely szignifikáns különbséget igazolt legalább két csoport között ($P = 0,0002$). A Dunnett-féle T3 többszörös összehasonlítási teszt szerint szignifikánsan eltért az AML-MR és az MDS-IB ($P = 0,0019$), illetve az AML-MR és az MDS-LB ($P = 0,0004$) csoport. Nem volt szignifikáns különbség az MDS-IB és az MDS-LB csoportok között ($P = 0,9330$) (**10. ábra**).

Négy predikciós módszert alkalmaztunk a szerkezeti stabilitási különbségek elemzésére. A DynaMut2 Cry (kristályos szerkezet alapú), az AF, a DDGun 3D Cry és az AF módszereket használtuk. Az I-Mutant2.0 Seq, I-Mutant2.0 struc és DDGun Seq módszereket kizártuk az elemzésből az alacsony Pearson-r érték miatt (**7. ábra**). A one-way ANOVA teszt szignifikáns különbséget adott legalább két csoport között ($P < 0,0001$). A Tukey-féle többszörös összehasonlítási teszt szerint az AML-MR és az MDS-IB ($P < 0,0001$), valamint az MDS-IB és az MDS-LB ($P = 0,0003$) csoportok szignifikánsan eltértek egymástól. Az AML-MR és az MDS-LB csoportok között nem volt szignifikáns különbség ($P = 0,0698$) (**10. ábra**).

Összehasonlítottuk még a 26 különböző nem kereteltolódással járó mutáció promóter specifikus transzkripció aktivitási értékeit. A one-way ANOVA teszt legalább két csoport között szignifikáns különbséget mutatott ($P = 0,0003$). A Tukey-féle többszörös összehasonlítási teszt alapján szignifikáns eltérések voltak az AML-MR és az MDS-IB ($P < 0,0024$), valamint az MDS-IB és az MDS-LB ($P = 0,0005$) csoportok között. Az AML-MR és az MDS-LB csoportok között azonban nem találtunk szignifikáns különbséget ($P = 0,7716$) (**10D. ábra**).



10. ábra. A mutáns TP53 fehérje előrejelzett patogén hatásának és stabilitásának összehasonlítása AML/MDS beteg csoportok között. **A:** A TP53 mutációk súlyosságának összehasonlítása a patogenitási pontozási rendszerek alapján (**5C** és **5D. ábra** alapján) az AML-MR, az MDS-IB és az MDS-LB esetekben. A Kruskal-Wallis teszt alapján szignifikáns különbségek vannak a csoport között. **B:** A mutációk patogenitási képességének összehasonlítása REVEL, BayesDel, Varity és FATHMM-XF pontozásokkal (**5D. ábra**). A Brown-Forsythe and Welch ANOVA teszt szignifikáns különbségeket mutatott ki a csoportok között. **C:** A TP53 fehérje stabilitásváltozásainak statisztikai különbségei a csoportok között az **5B. ábra** módszereivel. A one-way ANOVA teszt szignifikáns különbséget mutatott. **D:** A TP53 mutációk promóter

specifikus transzkripciós aktivitásának mértéke a vad típusú TP53 (100%) fehérjéhez képest (%).
* $P < 0,05$, ** $P < 0,01$, *** $P < 0,0001$, **** $P < 0,0001$

6.2 *BCOR*-gént érintő génfúziók *in silico* funkcionális elemzése szarkómákban

6.2.1 A fúziós fehérjék szekvencia- és doménjellemezői

A BRS diagnosztikája és terápiája jelentős kihívás a ritkasága és más szarkómákkal való átfedő jellemzői miatt. A *BCOR*-fúziók azonban diagnosztikus markerekként szolgálnak, segítve a BRS elkülönítését hasonló malignitásoktól [128].

Ennek a ritka malignitásnak a vizsgálatához összegyűjtöttük a BRS-hez köthető szakirodalomban fellelhető *BCOR* génfúziókat (**3. táblázat**). A nukleotid szekvenciákon *in silico* átírást végeztünk, majd elemeztük a fúziós fehérjék szekvenciáit, és összevetettük őket a vad típusúakéval (Blastp eszköz segítségével). A kilenc vizsgált fúziós fehérjét két csoportba soroltuk a vBCOR pozíciója alapján. Kilenc esetből négyben a vBCOR fehérje a fúziós fehérje N-terminális részén, míg az öt további esetben a C-terminális részén helyezkedett el ($BCOR^{NT}$ és $BCOR^{CT}$).

A CDD, az UniProt és az InterPro adatbázisokat használtuk a domének és kölcsönhatási helyek feltérképezésére a vad típusú és fúziós fehérjék esetén. Kiemelendő, hogy a BCOR ankyrin (ANK) ismétlődései és a PUFD doménje minden fúziós fehérjében megmaradt. Az Mbs-t minden fúziós fehérjében azonosítottuk. Azonban a BCOR-2 izoformákban csak 89 AA található, szemben a kanonikus BCOR 123 aminosavával. A Bbs minden $BCOR^{NT}$ fúzióban jelen van, de a $BCOR^{CT}$ csoportból csak az RTL9::*BCOR*-2 fúziós fehérjében található meg. A többi partnerrel kialakított fúziós fehérjék funkcionális doménjeiben különböző mértékű változások figyelhetők meg (**11. ábra**).

A *BCOR*::CCNB3 (UniProt ID H9A532 [82]) 3038 aminosavból áll és tartalmazza a teljes hosszúságú vBCOR-t (1755 AA). A fúzió a CCNB3 fehérje rövid N-terminális régiójának (a teljes fehérjeszekvencia 10%-a) elvesztését okozza. A fúzió következtében a CCNB3 rövid N-terminális régiója, amely a destruktív boxot (D-box) tartalmazza, elveszett, míg a C-terminális ciklin boxok megmaradtak.

A *BCOR*::MAML3 fúziós fehérje szinte teljes hosszúságú vBCOR-t tartalmazza, az 1-1751. aminosavak megmaradnak, és csak az 1751-1755. C-terminális részek hiányoznak. A fúziós fehérje 1752-2734. régiója a MAML3 fehérje trunkált N-terminális részének felel meg, ahol az 1-156. aminosavak hiányoznak a 1138 hosszúságú fehérjéből. Az 1752. pozícióban lévő aminosavat a *BCOR* adenin és guanin, illetve a *MAML3* guanin bázisa kódolja. A *BCOR*::MAML3 fúziós fehérje elvesztette a neurogén mastermind-szerű N-terminális (MamL-1) domént és a MamL1/3 transzaktivációs domén 1 (MamL1/3 TAD1) egy részét, de a MamL1/3 TAD2 doménje megmaradt.

A *BCOR*::CLGN (2070 AA) tartalmazza a teljes hosszúságú vBCOR-t (1755 AA). Ebben az esetben jelentősebb trunkáció (~50%) figyelhető meg a CLGN esetében, ahol az 1-295.

aminosavak hiányoznak a 610-ből [125]. A BCOR::CLGN fúzió a kalretikulin domén 368 aminosavából 232-t veszít el, amely kulcsfontosságú a kalciumion kötéshez.

A BCOR-2::MAML1 (2699 AA) fúziós fehérje megtartja a BCOR-2 (UniProt ID Q6W2J9-2) integritását, még a MAML1 elveszíti kezdeti 39 szekvenciáját, ami a MamL-1 domén 26 aminosavját érinti.

A ZC3H7B::BCOR fúziós fehérje hossza 1108 AA, amelynek 1-433. régiója a trunkált ZC3H7B fehérjének, míg a 434-1108. régiója a vBCOR 1081-1755. aminosavainak felel meg. A ZC3H7B::BCOR fúziós fehérje megtartja az LD motívumot és a három tetratrikopeptid ismétlést (TPR), de elveszíti az összes cinkujját.

A KMT2D::BCOR fúzió esetében egy 5480 AA hosszúságú fúziós fehérje keletkezik, amely a KMT2D 1-4667. és a vBCOR 943-1755. aminosav szekvenciáját tartalmazza. Ennek megfelelően a KMT2D 16%-a és a BCOR 54%-a hiányzik a fúziós fehérjéből. A KMT2D-ben csak a kiterjesztett növényi homeodomén (ePHD) cinkujj, az öt PHD és a high mobility group (HMG) box domének maradtak meg, míg a második ePHD, a F/Y-gazdag N-terminális (FYRN), az F/Y-gazdag C-terminális (FYRC) és a SET (beleértve a post-SET-et) domének elvesznek.

A CIITA::BCOR fúziós fehérje a vizsgált fúziós fehérjék közül a legrövidebb (821 AA). A fúziós fehérje a CIITA 13%-át (1-145. N-terminális régió) és a vBCOR 38%-át (1081-1755. C-terminális régió) tartalmazza. A 146. pozícióban lévő argininnak megfelelő tripletet a *CIITA* citozinja, illetve a *BCOR* gén guaninja és citozinja kódolja. A CIITA::BCOR fúziós fehérje csak a CIITA acetiltranszferáz doménjét (ADD) tartalmazza, miközben a prolin-szerin-treonin gazdag (P/S/T) és a NACHT domének, valamint a leucin-gazdag ismétlődések (LRR) kiesnek.

Az RTL9::BCOR-2 fúziós fehérje - 2891 aminosavból áll - tartalmazza a C-terminális részén trunkált RTL9 fehérjét (az 1200-1388. régió hiányzik), valamint a szinte teljes hosszúságú BCOR-2-t, amelyből csak az 1-29. N-terminális régió hiányzik a fúziós fehérjéből. Az RTL9::BCOR-2 fúzió az RTL9 1199. pozíciójánál töréspontot okoz, ami a C-terminális régió közelében lévő teljes kapszid-szerű domén (1169-1316. régió) kiesését eredményezi.

Az AHR::BCOR-2 fúziós fehérje 1676 AA hosszúságú. Az AHR 801. aminosavánál törést szenvedett és a fúziós esemény miatt a 10-es intronból származó 802-806. aminosavakkal kapcsolódott. A fúziós fehérje C-terminálisa a BCOR-2 fehérje nagymértékben trunkálódott 868-1721. részével ér véget. A BCOR-2::AHR fúzióban az AHR C-terminálisa 39 aminosavval rövidebb. Az AHR::BCOR-2 fúzió esetében az AHR minden funkcionális doménjét megtartja, de elveszít 47 C-terminális aminosavát, míg a BCOR-2 elveszíti a Bbs doménjét (**11. ábra**).

Összefoglalva, a szekvenciaelemzés jelentős változatosságot mutatott mind a BCOR, mind annak fúziós partnerei esetében, ezzel hangsúlyozva a fúziós események által okozott szerkezeti sokszínűséget. A BCOR-nál megfigyelt szekvencia megtartás mértéke 38% és 100% között

változott, átlagos értéke 79% volt. Fontos megjegyezni, hogy a legnagyobb mértékű trunkáció a CIITA::BCOR fúzióban volt megfigyelhető, ahol a BCOR fehérjeszekvenciájának csak 38%-át őrizte meg, míg a BCOR teljesen érintetlen maradt olyan fúziókban, mint a BCOR::CCNB3, BCOR::MAML3 és BCOR::CLGN. A BCOR-hoz képest a fúziós partnerek nagyobb mértékű trunkációt mutattak, eredeti szekvenciájuk 13-98%-át őrizték meg (átlagosan 65%).



11. ábra. A *BCOR*-átrendeződéses szarkómák ismert fúziós fehérjének doménjei és funkcionális helyei. Minden fúziós partner és fúziós fehérje sematikus ábrázolása a fúziós fehérjékben elfoglalt N- vagy C-terminális pozícióknak megfelelő sorrendben látható. A piros szaggatott vonalak a töréspontok helyét jelölik. **Bbs:** BCL6 kötési hely, **Mbs:** MLLT kötési hely, **ANK:** Ankyrin ismétlődések, **PUFD:** PCGF Ubiquitin-szerű fold diszkriminátor; **D-box:** Destrukciós box, **MamL-1:** Neurogén mastermind-szerű N-terminális domén, **Mam1_3 TAD 1-2:** Mastermind-szerű 1/3: transzaktivációs domén, **TPR:** Tetratrikopeptid ismétlődések, **LD motívum:** Leucin-asparaginsav ismétlődési motívum, **C3H1-type:** CCCH típus, **C2H2-type:** CCHH típus, **ePHD:** Kiterjesztett növényi homeodomén, **PHD:** Növényi homeodomén, **HMG-box:** High mobility group -bocsz, **FYRN:** FY-gazdag domén N-terminális régió, **FYRC:** FY-gazdag domén C-terminális régió, **SET:** Pre-SET és post-SET régió, **ADD:** Acetiltranszferáz domén, **P/S/T:** Prolin-szerin-treonin gazdag domén, **LRR:** Leucin-gazdag ismétlődésű ribonukleáz inhibitor, **NLS1:** Nukleáris lokalizációs szignál 1, **bHLH:** Alap hélix-hurok-hélix, **PAS:** Per-ARNT-Sim domén, **PAC:** PAS C-terminális alegység.

3. táblázat. A fúziós események és a töréspontok (BP) *BCOR*-átrendeződéses szarkómákban. Az *RTL9*, a *MAML1* és az *AHR* génp partnerek kivételével a *BCOR* a kanonikus formának felel meg (1755 AA). A *BCOR-2* izoforma 1721 aminosav hosszú. Partner 1 vagy 2: N és C terminális fehérje partner. Partner 1 vagy 2 / VT: a vad típusú fehérjéhez (VT) viszonyított AA hossz.

Fúziók	Gén 5' BP	Gén 3' BP	Exonikus BP 5' - 3'	Partner 1 / VT (AA)	Partner 2 / VT (AA)	Fúziós fehérje (AA)
<i>BCOR::CCNB3</i>	39 911 366	50 051 504	15. - 5.	1755 / 1755	1283 / 1395	3038
<i>BCOR::MAML3</i>	39 911 374	141 074 014	15. - 1.	1751 / 1755	982 / 1138	2734
<i>BCOR::CLGN</i>	39 911 366	141 317 359	15. - 9.	1755 / 1755	316 / 610	2070
<i>BCOR::MAML1</i>	40 051 795	179 733 227	10. - 4.	1721/1721	978 / 1016	2699
<i>ZC3H7B::BCOR</i>	41 738 632	39 923 852	12. - 7.	433 / 977	675 / 1755	1108
<i>KMT2D::BCOR</i>	49 424 063	39 931 774	42. - 4.	4667 / 5537	813 / 1755	5480
<i>CIITA::BCOR</i>	10 992 859	39 923 852	5. - 7.	145 / 1130	675 / 1755	821
<i>RTL9::BCOR</i>	109 697 442	39 937 097	2. - 3.	1199 / 1388	1692 / 1721	2891
<i>AHR::BCOR</i>	17 340 228	40 072 790	10. (+intron 10.) 4.	801 / 848 + 6 (intron 10)	853/1721	1676

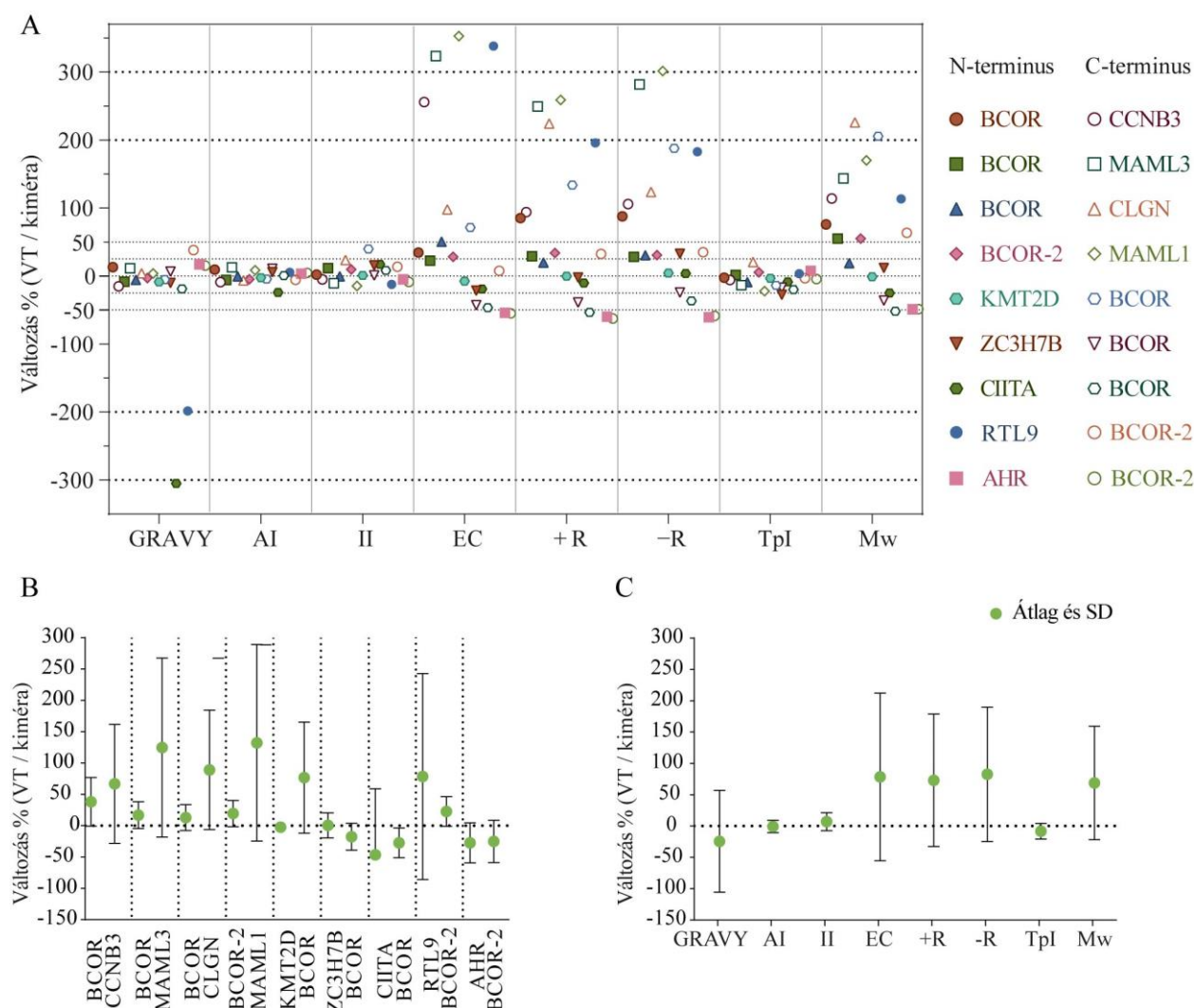
6.2.2 A fúziós fehérjék fizikokémiai tulajdonságai

A fúziós fehérjék fizikokémiai tulajdonságait a ProtParam web szerver segítségével számítottuk ki. Minden esetben megfigyelhető volt a fúziós fehérjék tulajdonságainak változása a vad típusú fehérjékhez képest (**12. ábra**)

Az összes esetre vonatkozó átlagos eltérés 29,54% volt. A +R (%) és -R (%) értékek - amelyek a töltéssel rendelkező aminosavak arányát jelzik a teljes szekvenciához viszonyítva, 9,79% és voltak. A *BCOR*^{NT} csoport átlagos eltérése 62,47%, míg a *BCOR*^{CT} csoportnál ezek az értékek 3,20% volt.

A *MAML1* fehérje mutatta a legnagyobb átlagos eltérést a vizsgált tulajdonságok esetében (132,20%), ami a fúziós esemény következtében bekövetkezett jelentős trunkáció magyarázható.

Ezzel szemben a ZC3H7B és KMT2D esetében a legkisebb átlagos eltérés (0,47%) volt megfigyelhető. A GRAVY értékek átlagos eltérése 26,13% volt, a legkisebb és legnagyobb eltérések a BCOR-2::MAML1 (BCOR-2: 3,18%) és a CIITA::BCOR (305,10%) esetében fordultak elő. Az AI értékek átlagos változása elhanyagolható volt (0,19%), bár a CIITA 24,29%-ot csökkent, ami a termikus stabilitás romlására utal.



12. ábra. A fúziós fehérjék fizikai-kémiai tulajdonságainak változása a vad típusú partnerfehérjékhez képest. **A:** A nyolc vizsgált jellemző (%) eltérései a vad típusú és fúziós fehérjék között. GRAVY: Átlagos hidrofóbicitás, AI: Alifás index, II: Instabilitási index, EC: Abszorpciós együttható 280 nm-en, +R: Pozitív töltésű aminosavak száma, -R: Negatív töltésű aminosavak száma, Tpl: Elméleti izoelektromos pont, Mw: Molekulatömeg, VT: Vad típusú. **B:** A fehérjékben a vizsgált fizikokémiai tulajdonságok százalékos változásának átlaga és szórása. **C:** A nyolc paraméter százalékos változásának átlaga és szórása a fúziók esetén a vad típusú fehérjéhez képest. SD: Standard deviáció, szórás.

Az instabilitási index (II) átlagos változása kifejezettebb volt (4,69%), a legnagyobb változás a KMT2D::BCOR fúzióban lévő BCOR fehérjénél (39,63%) jelentkezett, ami a fehérje termikus stabilitásának csökkenésére utal. A +R és -R százalékos változásának átlaga 62,61% és 69,71% között mozgott, a legmagasabb értékeket (258,67% és 301,41%) a MAML1 mutatta, ami a fúzió

következtében bekövetkezett masszív trunkáció eredménye. A fúziós fehérjék EC értékének átlagos változása 74,18% volt a vad típusúakhoz képest. A BCOR-2::MAML1 mutatta a legnagyobb EC növekedést (352,6%), míg a CIITA::BCOR a legnagyobb csökkenést (19,2%). A fúziós fehérjék TpI értéke átlagosan 6,31%-kal csökkent, a ZC3H7B::BCOR mutatta a legnagyobb TpI csökkenést (27,67%), míg a BCOR::CLGN esetén a TpI 20,6%-kal magasabb volt. Azonban egyik fúziós fehérje sem lépte túl a 4,5-10 pH-tartományt, ami nukleáris lokalizációra utal [205].

A CIITA fehérje AI és GRAVY értékei 305,10%-kal és 24,29%-kal csökkentek, jelezve a termikus stabilitás romlását és a hidrofilitás növekedését. Kiemelendő, hogy a KMT2D::BCOR fúzióban lévő BCOR fehérje II értéke nőtt, annak ellenére, hogy a BCOR már eredetileg magas II értékkel és hosszabb szekvenciával rendelkezett. Ez jelentős stabilitáscsökkenést okozhat a BCOR-nál, míg a KMT2D stabilitása csak mérsékelten változott.

Az AI, az II és a TpI értékeknél jelentősen alacsonyabb szórás (SD) volt megfigyelhető, mint a többi paraméternél. A hasonló AI értékek hasonló termikus stabilitást jeleznek a fúziós fehérjéknél. Az instabilitási indexek is hasonlóak voltak, a legmagasabb értéket a KMT2D::BCOR mutatta, ami a legkisebb relatív stabilitásra utal.

A partnerfehérjék és fúziós fehérjék átlagértékeinek összehasonlítása alapján a GRAVY, az AI, -R(%) és +R(%) értékek csökkentek, míg az II nőtt. Ez hidrofóbicitás növekedését, termikus stabilitás csökkenését, töltött aminosavak arányának romlását és a fehérjekárosodás valószínűségének növekedését jelzi.

6.2.3 Génontológia

A PANTHER GO tulajdonságokat az InterPro keresővel értük el, hogy meghatározzuk a vad típusú és fúziós fehérjék funkcionális jellemzőit és sejten belüli lokalizációjukat (**4. táblázat**). Az eredmények szerint a BCOR fehérjéhez kapcsolódó GO tulajdonságok - például a transzkripció negatív szabályozása RNS-polimeráz II által, a transzkripció korepresszor aktivitás és a sejtmagi lokalizáció - a fúziós fehérjékben is jelen voltak. A vad típusú fehérjék GO értékei csaknem minden fúziós partner esetében megváltoztak, kivéve a BCOR-t. Érdekes módon a KMT2D::BCOR fehérje csak a vad típusú KMT2D GO tulajdonságát őrizte meg, de a BCOR-ét nem. Az RTL9 és az RTL9::BCOR-2 fehérjékre nem találtunk GO tulajdonságokat, mivel ezekre a fehérjékre nem állt rendelkezésre domén információ az adatbázisban.

4. táblázat. A vad típusú fehérjék és fúziós alakjainak PANTHER Gén Ontológiás tulajdonságai. **RNS-pII:** RNS polimeráz II, **TKRA:** Transzkripció korepresszor aktiválás, **TKAA:** Transzkripció koaktivátor aktiválás, **T+:** Transzkripció pozitív szabályozása, **T-:** Transzkripció negatív szabályozása, **CDK:** Ciklin-dependens kináz, **ERAD:** Endoplazmatikus retikulum-asszociált degradáció, **H3K4 MA:** Hiszton H3K4 metiltransferáz aktivitás, **MHC:** Fő hisztokompatibilitási komplex, **MLL3/4:** Mixed Lineage Leukemia 3/4.

Fehérjék	Biológiai folyamat	Molekuláris funkció	Sejten belüli lokalizáció
BCOR és BCOR-2	RNS-pII általi T-	TKRA	sejtmag
CCNB3	CDK aktivitás szabályozása Mitotikus ciklus fázisváltás	CDK szabályozó akt.	CDK komplex citoplazma
BCOR::CCNB3	RNS-pII általi T-	TKRA	sejtmag
MAML3	Notch célpont T+	TKAA	nukleoplazma
BCOR::MAML3	RNS-pII általi T-	TKRA	sejtmag
CLGN	Fehérje folding Ubiquitin-függő ERAD	Ca ²⁺ -kötés	ER membrán
BCOR::CLGN	RNS-pII általi T-	TKRA	sejtmag
MAML1	Notch célpont T+	TKAA	nukleoplazma
BCOR-2::MAML1	RNS-pII általi T-	TKRA	sejtmag
KMT2D	RNS-pII általi T+ Hisztonkötés H3K4 MA	TKAA	MLL3/4 komplex
KMT2D::BCOR	RNS-pII általi T+ TKAA Hisztonkötés	H3K4 MA	MLL3/4 komplex
ZC3H7B	miRNS feldolgozás	miRNS kötés	-
ZC3H7B::BCOR	RNS-pII általi T-	TKRA	sejtmag
CIITA	MHC I bioszintézis + MHC II bioszintézis + RNS-pII általi T+	-	-
CIITA::BCOR	RNS-pII általi T-	TKRA	sejtmag
AHR	RNS-pII általi T+ Nukleáris receptor akt.	Cis-regulációs régió kötése Aريل-szénhidrogén receptor	sejtmag -
AHR::BCOR-2	RNS-pII általi T-	TKRA	sejtmag

6.2.4 Előre jelzett sejten belüli lokalizáció és jelzőpeptidek

A SignalP 6.0 eszközt a Sec/SPI-ek azonosítására, a DeepLoc 2.0-t pedig a sejten belüli lokalizáció predikciójára használtuk. Emellett a NetGPI-1.1 prediktort alkalmaztuk GPI-horgonyzottsági jelek azonosítására (**5. táblázat**).

A SignalP 6.0 elemzés szerint csak a CLGN fehérje tartalmaz szignálpeptidet, pontosabban 99,92%-os valószínűséggel Sec/SPI-et. Azonban a CLGN N-terminális régiója, amely ezt a szekvenciamotívumot tartalmazza, hiányzik a BCOR::CLGN fúziós fehérjéből.

A DeepLoc 2.0 számításai szerint a tíz vad típusú fehérjéből hét (BCOR, CCNB3, MAML3, MAML1, KMT2D, CIITA, AHR) valószínűleg a sejtmagban lokalizálódik, míg a CLGN az ER-ben, ami összhangban van az UniProt adatbázis adataival. A Human Protein Atlas adatbázis szerint a KMT2D a plazmamembránban és citoszolban is jelen lehet. A ZC3H7B szubcelluláris lokalizációja bizonytalan: az UniProt sejtmagi jelenlétet ír le, de a Human Protein Atlas adataival összhangban citoplazmában lokalizálódik. Az RTL9 fehérjére vonatkozó adatok hiányoztak az UniProt-ból. Legújabb eredmények szerint az egér RTL9 fehérje intracelluláris, és erősen expresszálódik az újszülöttkori agy mikrogliális lizoszómáiban [206]. Minden fúziós fehérjének sejtmagi elhelyezkedést prediktáltunk, kivéve a BCOR::CLGN fehérjét, amely az ER-ben lokalizálódhat.

A DeepLoc 2.0 előrejelzések alapján minden vad típusú és fúziós fehérje (a CLGN, BCOR::CLGN, a CIITA::BCOR és az RTL9 kivételével) tartalmaz nukleáris export szignált (NES). A CLGN fehérjében szignálpeptidet és transzmembrán domént jósolt, még a BCOR::CLGN esetében csak szignálpeptidet. A ZC3H7B::BCOR, a KMT2D::BCOR és az RTL9::BCOR-2 esetében a NES mellett nukleáris lokalizációs szignál (NLS) jelenlétét valószínűsítettük, még a CIITA::BCOR és az RTL9 fehérje a DeepLoc 2.0 predikció szerint tartalmaz NLS-t. Fontos megjegyezni, hogy az NES előrejelzés pontossága a legalacsonyabb (<50%), ezért mutathattunk ki NES-t a legtöbb vizsgált fehérjében.

Összességében a DeepLoc 2.0 által prediktált lokalizáció a legtöbb esetben összhangban volt az UniProt és a Human Protein Atlas adatbázisok adataival. A NetGPI-1.1 programmal nem azonosítottunk GPI-horgonyzottsági szignálokat, ami azt jelzi, hogy sem a vad típusú, sem a fúziós fehérjék nem tartalmaznak lipidhorgonyzót és nem sejt felszíni fehérjék.

5. táblázat. A vad típusú és a fúziós fehérjék intracelluláris lokalizációja és a DeepLoc 2.0 segítségével előre jelzett szignálok. NES: Nukleáris export szignál, SP: szekretoros szignálpeptidek, TD: Transzmembrán domén, NLS: Nukleáris lokalizációs domén, ER: Endoplazmatikus retikulum, PML: Promielocita leukémia magtestek, egyéb: nem találtunk szignálpeptidet. A DeepLoc predikciók esetében csak a küszöbértéket elérő lokalizációk vannak feltüntetve.

Fehérjék	DeepLoc 2.0				Prediktált szignálok	UniProt adat
	Prediktált lokalizáció		Valószínűség			
BCOR	Nukleusz		0,8702		NES	Nukleusz
CCNB3	Citoplazma	Nukleusz	0,5879	0,7553	NES	Nukleusz
BCOR::CCNB3	Citoplazma	Nukleusz	0,5760	0,7130	NES	NA
MAML3	Nukleusz		0,8818		NES	Nukleusz magfolt
BCOR::MAML3	Citoplazma	Nukleusz	0,4820	0,7875	NES	NA
CLGN	Endoplazmatikus retikulum		0,8276		SP, TD	Endoplazmatikus retikulum
BCOR::CLGN	Endoplazmatikus retikulum		0,6739		SP, TD	NA
MAML1	Nukleusz		0,8781		NLS, NES	Nukleusz
BCOR-2::MAML1	Citoplazma	Nukleusz	0,4892	0,7800	NES	NA
ZC3H7B	Citoplazma		0,6175		NES	Nukleusz
ZC3H7B::BCOR	Citoplazma	Nukleusz	0,5808	0,7169	NLS, NES	NA
KMT2D	Nukleusz		0,7295		NES	Nukleusz
KMT2D::BCOR	Nukleusz		0,8675		NLS, NES	NA
CIITA	Citoplazma	Nukleusz	0,5189	0,6742	NLS, NES	Nukleusz, PML body
CIITA::BCOR	Nukleusz		0,9452		NLS	NA
RTL9	Citoplazma	Nukleusz	0,6454	0,5456	NLS	NA
RTL9::BCOR-2	Nukleusz		0,8576		NLS, NES	NA
AHR	Citoplazma	Nukleusz	0,5320	0,802	NES	Citoplazma, Nukleusz
AHR::BCOR-2	Citoplazma	Nukleusz	0,4890	0,8070	NES	NA

6.2.5 A fehérjék rendezetlen régióinak előrejelzése

Az IUPred3 online eszközt használtuk a fehérjék IDRs-jainak azonosítására. Ezt követően összehasonlítottuk a region^{ANK+linker} és a region^{PUFD} szakaszokat a fúziós fehérjékben a BCOR fehérjével, és az IUPred3 pontszámokat statisztikai elemzéseinkbe is integráltuk, külön-külön vizsgálva minden régió értékeit Friedman-tesztek segítségével (**13. ábra**).

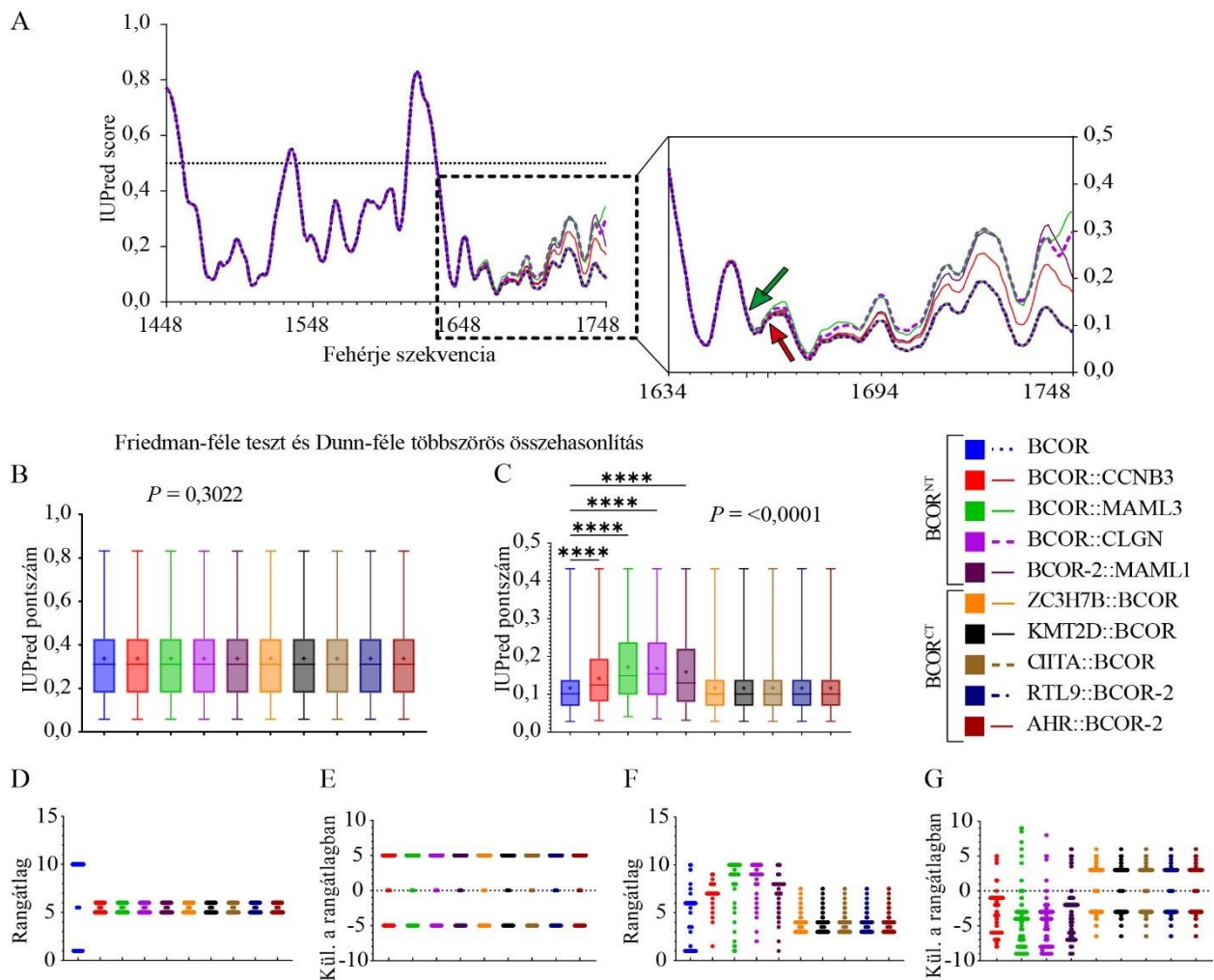
A vBCOR region^{ANK+linker} régiója (1448-1633 AA) ANK ismétlésekből és egy linker régióból áll, míg a region^{PUFD} a PUFDD domént (1634-1748) tartalmazza.

Nem találtunk szignifikáns különbséget a fehérjék region^{ANK+linker} IUPred3 IDRs pontszámai között ($P = 0,3022$). Ugyanakkor szignifikáns különbségeket figyeltünk meg a region^{PUFD} esetében legalább két csoport között ($P < 0,0001$). A Dunn-féle többszörös összehasonlítási teszt

szignifikáns különbségeket adott az IUPred3 IDR pontszámok rangátlaiban a vBCOR és BCOR^{NT} fehérjék között. Szignifikáns különbségeket találtunk a rangátlagokban az alábbi párok között: BCOR vs. BCOR::CCNB3 ($P < 0,0001$), BCOR vs. BCOR::MAML3 ($P < 0,0001$), BCOR vs. BCOR::CLGN ($P < 0,0001$) és BCOR vs. BCOR-2::MAML1 ($P < 0,0001$). Nem volt statisztikailag szignifikáns különbség a rangátlagokban a vBCOR és az alábbi fúziós fehérjék között: ZC3H7B::BCOR, KMT2D::BCOR, CIITA::BCOR, RTL9::BCOR-2 és AHR::BCOR-2.

A Dunn-féle többszörös összehasonlítás Z-pontszáma statisztikai mutatóként szolgál a két BCOR csoport közötti különbség nagyságának jelzésére. A legnagyobb különbséget a BCOR^{NT} csoporton belül a vBCOR-hoz képest a BCOR::CLGN fúzióban figyeltük meg ($Z = 12,59$), amelyet a BCOR::MAML3 fúzió ($Z = 11,63$), a BCOR-2::MAML1 ($Z = 9,256$) és végül a BCOR::CCNB3 fúziós fehérje ($Z = 8,13$) követett. E jelentős különbségek ellenére fontos megjegyezni, hogy egyik IDR pontszám sem érte el az rendezetlenségre vonatkozó 0,5-ös küszöbértéket, ami azt jelzi, hogy ez a régió megőrzi strukturált természetét.

Az IPred3 előrejelzése szerint a BCOR^{NT} fúziós fehérjék megnövekedett rendezetlenségi hajlamot mutatnak a region^{PUFD}-ben, ami potenciálisan befolyásolhatja a fehérje kölcsönhatását a PCGF1 fehérje RAWUL doménjével és a KDM2B fehérje LRR régiójával. Ilyen változásokat nem prediktáltunk a BCOR^{CT} csoport tagjai esetében.

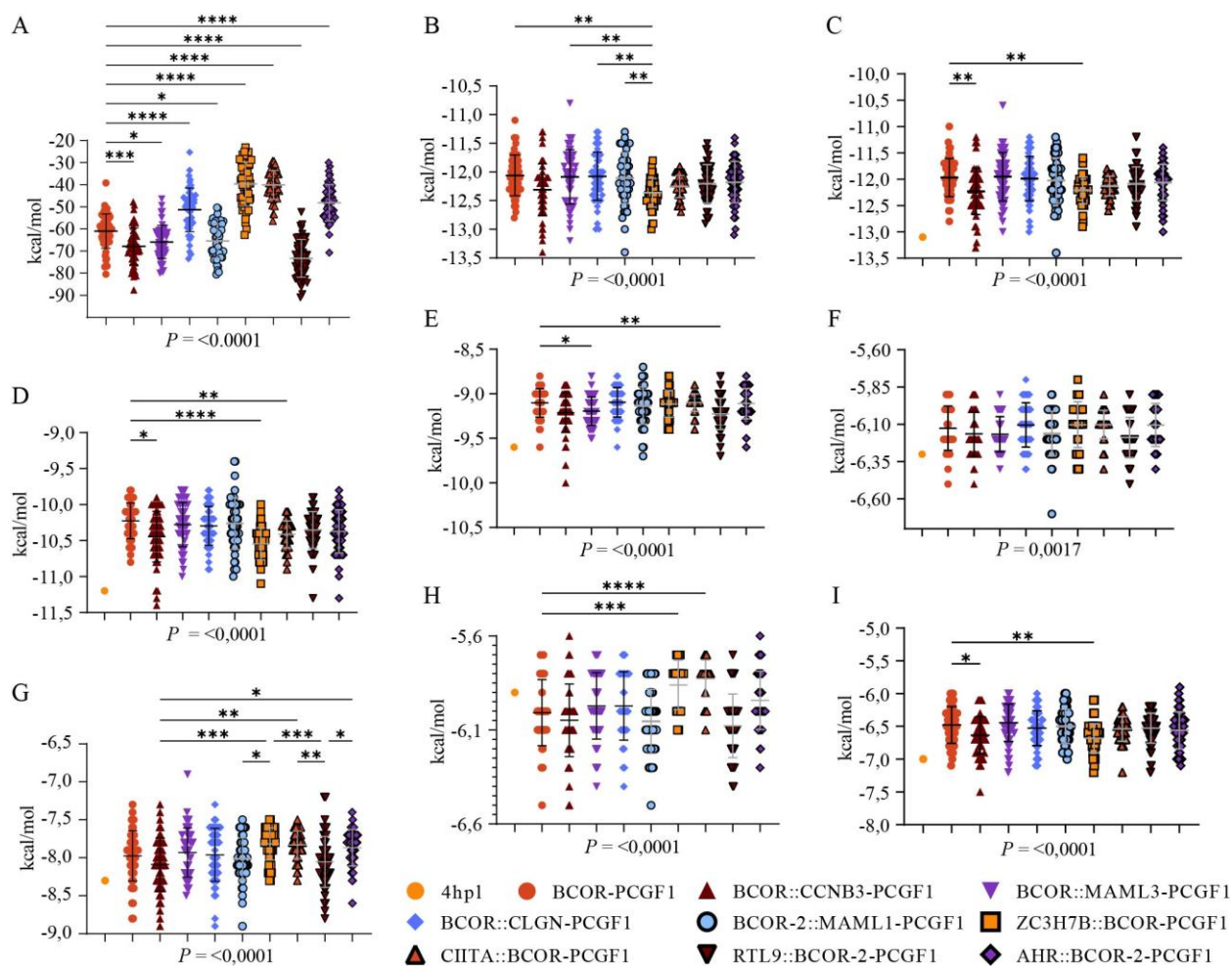


13. ábra. A vad típusú BCOR és a fúziós fehérjék rendezetlenségi hajlamának összehasonlítása. A rendezetlenségi hajlamokat az IUPred3 online eszköz segítségével számítottuk ki. A BCOR region^{ANK+linker} (1448-1633 AA) és region^{PUFD}-t (1634-1748 AA) vizsgáltuk mind a vBCOR-ban, mind pedig a fúziós fehérjékben. **A:** A fehérjék intruztikusan rendezetlen régióinak (IDRs) IUPred3 pontszámai koordináta-rendszerben ábrázolva. A zöld nyíl az első különbséget mutató pontot jelöli. A piros nyíl a csoportok szétválási pontjára mutat. **B** és **C:** Friedman-teszt és Dunn-féle többszörös összehasonlítás vBCOR-ra és fúziós fehérjékre vonatkozóan ($n = 8$). A vízszintes vonal mediánt jelöl, míg a pluszjel (+) az átlagértéket mutatja. **D** és **F:** A Friedman-teszt rangátlag diagramja. **E** és **G:** A Friedman-teszt rangátlag különbségei a csoportpárok között. A region^{ANK+linker} (**A, B, D, E**) és a region^{PUFD} (**A, C, F, G**) szakaszok mutatják a csoportok közötti különbséget. BCOR^{NT}: A BCOR a fúziós fehérje N-terminális végénél található. BCOR^{CT}: A BCOR a fúziós fehérje C-terminális végénél lokalizálódik. Kül.: Különbségek. **** $P < 0,0001$

6.2.6 A BCOR-PCGF1 dimer kötési affinitásának statisztikai elemzése

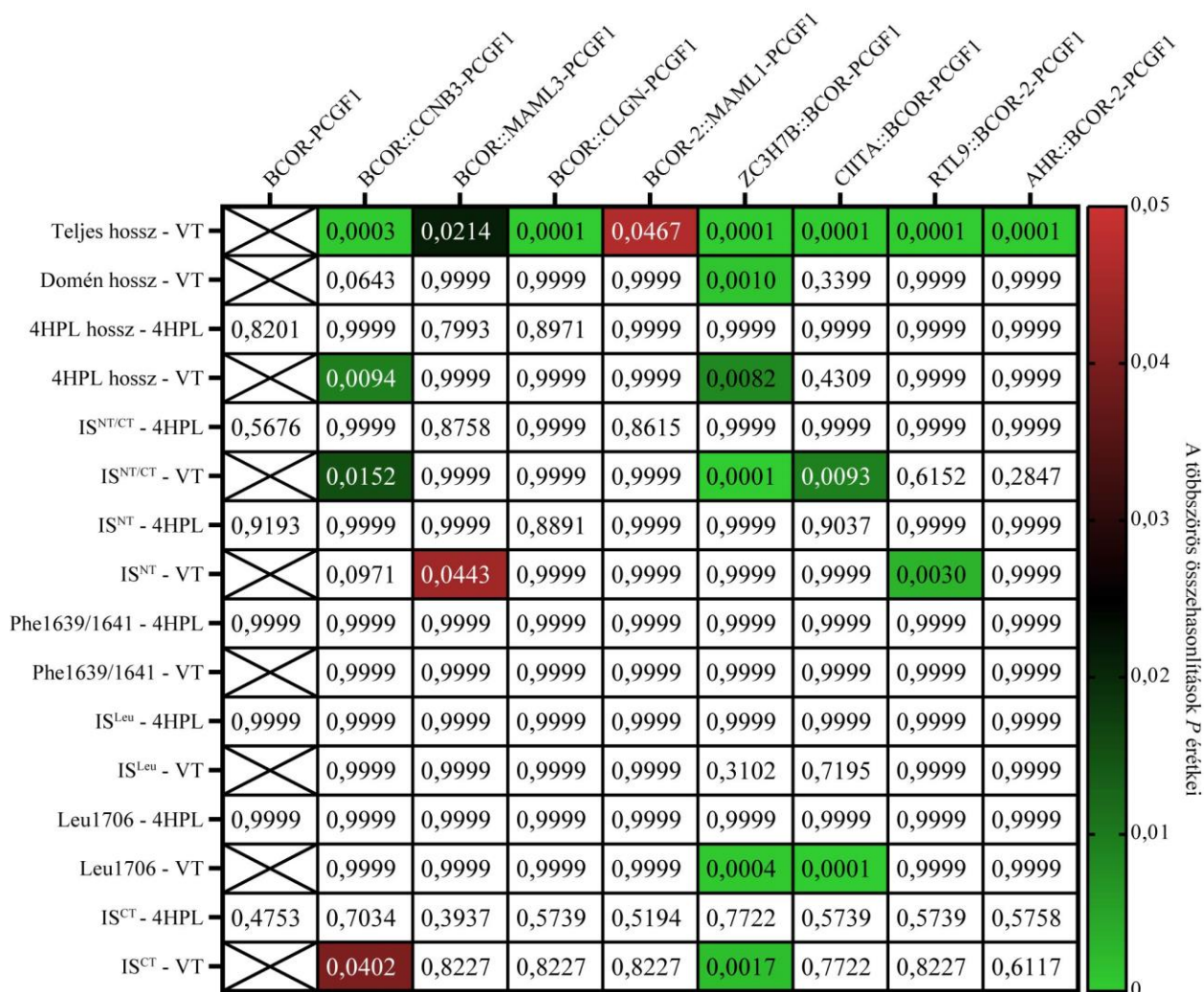
Amint azt fentebb leírtuk (**11. ábra**), a PUFD domén megőrzi szekvenciájának integritását, de ennek ellenére várható, hogy a fúziók befolyásolják a BCOR és más fehérjék közötti kölcsönhatásokat. Ezt a lehetőséget a fúziós fehérjék, vBCOR és a PCGF1 komplex prediktált háromdimenziós szerkezetei alapján vizsgáltuk statisztikai próbákkal (**14. ábra**). Ezt követően post-hoc tesztekkel végeztünk a vBCOR és fúziós dimerek közti szignifikáns különbséget mutató párosítások meghatározására (**14. ábra**). A szükséges statisztikai elemzések összetettsége miatt

fehérjénként 50 predikciós szerkezetet vizsgáltunk, ami várhatóan elegendő megbízhatóságot nyújtott az elemzésekhez. A dimer ΔG értékeit a PRODIGY segítségével számítottuk ki. Átfogó vizsgálatot végeztünk a teljes hosszúságú fehérjéken, a RAWUL-PUFD domén hossz, valamint a 4HPL hossz tartományokban [93]. Emellett hat különböző kölcsönhatási felszínt és az interakció szempontjából kiemelkedő aminosavat elemeztünk. A RAWUL-PUFD kötődési helyeket Junco *et al.* írta le [93]. Összesen kilenc régióon vizsgáltuk a kötési affinitási értékeket, hogy meghatározzuk a fúziós esemény hogyan befolyásolja a BCOR és a PCGF1 interakcióját.



14. ábra. A BCOR-PCGF1 dimerek kötődési affinitásának (ΔG) összehasonlítása különböző BRS fúziós fehérjékkel. Ötven predikciót végeztünk mind a kilenc fehérjedimerre (egy vad típusú és nyolc fúziós fehérje), amelyek BCOR-t vagy BRS fúziós fehérjét és PCGF1-et tartalmaztak. Egy kísérletileg meghatározott RAWUL-PUFD dimer (PDB ID: 4HPL) további kontrollként szolgált (C-I). A statisztikai elemzések közé tartozott a one-way ANOVA (A, I) és a Kruskal-Wallis tesztek (B-H). **A:** A teljes hosszúságú BCOR-PCGF1 dimer fehérjék összehasonlítása szignifikáns különbségeket mutatott ΔG értékekben a BCOR-PCGF1 komplex és a nyolc fúziós dimer között. **B:** A RAWUL-PUFD domén hosszában végzett összehasonlítás (BCOR: 16341748, PCGF1: 167-255) statisztikailag szignifikáns különbségeket mutatott az ΔG értékekben. **C:** A prediktált szerkezetek kötődési affinitások összehasonlítása a 4HPL tartományon belül és a generált vad típusú szerkezettel (BCOR-PCGF1) két fúziós dimer esetében szignifikánsan alacsonyabb ΔG értékeket mutatott, mint a BCOR-PCGF1 esetében. **D-I:** A különböző tartományokban végzett kötési affinitások (**D:** IS^{NT/CT}, **E:** IS^{NT}, **F:** Phe1639/Phe1641, **G:** IS^{Leu}, **H:**

Leu1706, I: IS^{CT}) szignifikáns eltéréseket mutattak a fúziós fehérjét tartalmazó dimerek és a BCOR-PCGF1 kontroll között. A kölcsönhatási felszínek közé tartoznak: a Val1636-Asn1651, a Ser1704-Asp1712 és a Gly1738-Asp1748, specifikus régiókkal, az Interaction-Surface^{NT/CT} (IS^{NT/CT}) és az Interaction-Surface^{Leu} (IS^{Leu}), valamint egyedi aminosavakkal, mint a Leu1706 és a Phe1639/Phe1641. * $P < 0,05$, ** $P < 0,01$, *** $P < 0,001$, **** $P < 0,0001$



15. ábra. A kötődési affinitások poszt-hoc statisztikai összehasonlításainak P értékei hőtérképen. A one-way ANOVA és a Kruskal-Wallis tesztek statisztikailag szignifikáns különbségeket mutattak a csoport között minden vizsgált tartományban. Többszörös összehasonlító teszteket alkalmaztunk a kontrollcsoportok (BCOR-PCGF1 (VT) és 4HPL) és a fúziós fehérjék közötti különbségek feltárására. Az oszlopok a tartományokat és a kontrollcsoportokat (domain hossz - VT), míg a sorok a PCGF1-el dimerizált fúziós fehérjéket jelölik.

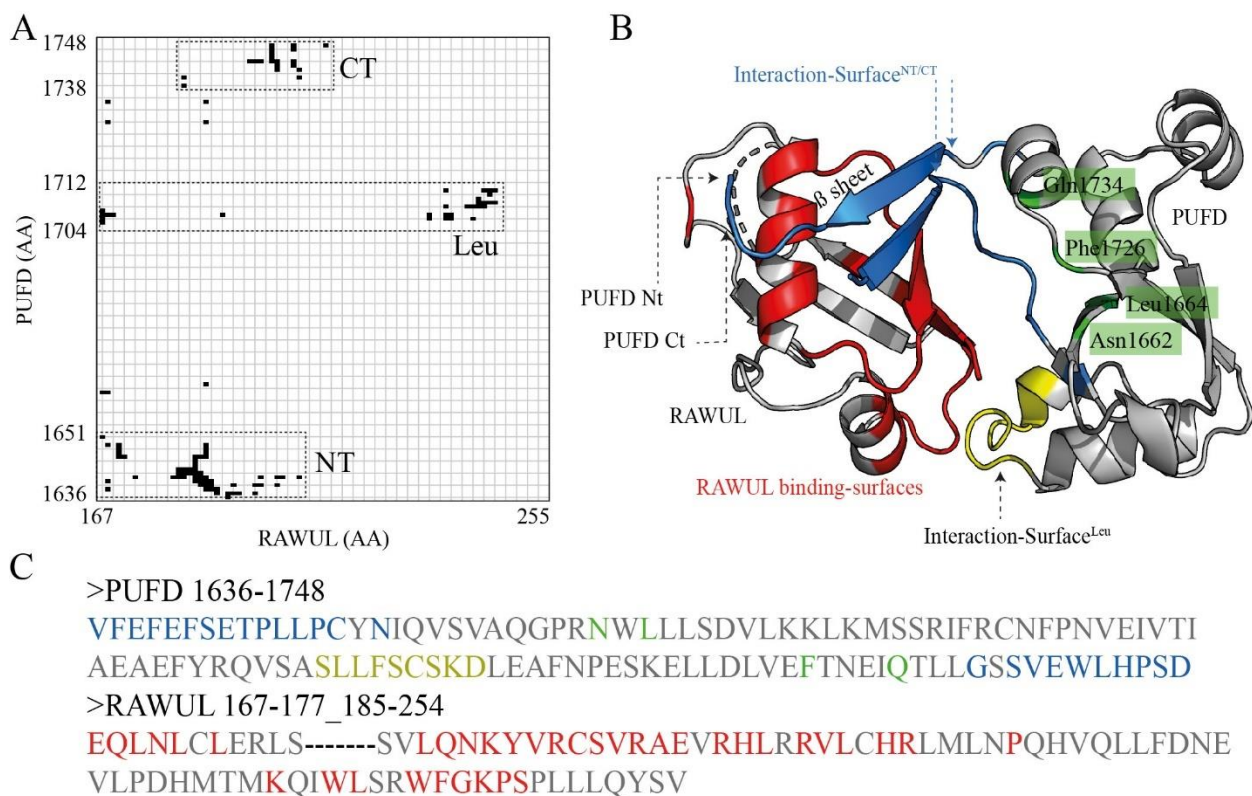
Először one-way ANOVA tesztet alkalmaztunk a teljes hosszúságú fehérjék dimerjeinek ΔG értékeinek elemzésére. Statisztikailag szignifikáns különbséget találtunk az átlagos ΔG értékek között legalább két csoport esetében ($P < 0,0001$). A Dunnett-féle többszörös összehasonlító teszt alapján az átlagos ΔG értékek szignifikánsan eltértek a BCOR-PCGF1 és a fúziós típusú csoportok között (**14. ábra és 15. ábra**). Az átlagos ΔG értékek szerint a BCOR::CCNB3-, BCOR::MAML3-, BCOR-2::MAML1- és RTL9::BCOR-2-PCGF1 dimerek szignifikánsan alacsonyabb, míg ezzel

ellentétben a BCOR::CLGN-, ZC3H7B::BCOR-2-, CIITA::BCOR- és AHR-BCOR-2-PCGF1 dimerek magasabb ΔG értékeket mutattak.

A második összehasonlítás során a RAWUL-PUFD domén hosszában lévő kilenc csoport ΔG értékeit elemeztük. A vizsgált régiók magukban foglalták a BCOR 1634-1748 és a PCGF1 167-255 aminosav szakaszait. A Kruskal-Wallis teszt szignifikáns különbséget mutatott ki a RAWUL-PUFD domén hosszában a kilenc csoport között ($P < 0,0001$). A domén tartományon belüli ΔG értékek rangátlagai szignifikánsan eltértek a csoportok között (**14. ábra**). Dunn-féle többszörös összehasonlító tesztet is végeztünk, amely szignifikáns különbségeket adott a BCOR-PCGF1 és a CIITA::BCOR-PCGF1 között, ahol a fúziós fehérje alacsonyabb ΔG értéket mutatott.

A harmadik elemzés során kontrollként használtuk a 4HPL kristályszerkezetet, mivel az tartalmazza az ismert RAWUL-PUFD dimert. A PyMOL-t használtuk az AF3 szerkezetek hosszának uniformizálására, hogy azok megegyezzenek a kristályszerkezetben (4HPL) szereplő fehérje szekvenciájával. Ezután összehasonlítást végeztünk ezen szerkezetek és a BCOR-PCGF1 prediktált szerkezete között fúziós típusú dimerekkel. A Kruskal-Wallis teszt szignifikáns különbséget mutatott a 4HPL hosszában (PCGF1 RAWUL: 167-177 és 185-254; BCOR PUFD: 185-254) a tíz csoport között ($P < 0,0001$) (**14. ábra**). A Dunn-féle többszörös összehasonlító teszt kimutatta, hogy a BCOR::CCNB3- és a ZC3H7B::BCOR-PCGF1 alacsonyabb ΔG értékekkel rendelkeztek. Ennek ellenére nem volt statisztikailag szignifikáns különbség a 4HPL és a fúziós dimerek között.

A PRODIGY elemzés feltárta a két fehérje (BCOR-PCGF1) közötti intermolekuláris kontaktusokat (IC-k), amelyek összhangban vannak Junco *et al.* által korábban leírt kötődési felszínekkel [93]. Megállapítottuk, hogy a RAWUL és a PUFD domének közötti kölcsönhatások két fő területre koncentráálódtak (**16. ábra**). Az első, ami a RAWUL kötési helyeivel interakcióba lép a PUFD Val1636 aminosava, az N-terminális β -lemez (Phe1637-Ser1642), az azt követő Glu1643-Asn1651 szakasz, valamint a hurok (Gly1738, Ser1739), amely átmenetet képez a C-terminális β -lemezbe (Ser1740-Leu1744) illetve az ezutáni His1745-Asp1748 aminosavak (Interaction-Surface^{NT/CT}). A RAWUL-PUFD kontaktus másik megfigyelt IC felszíne a Leu1706 körül koncentrálódtak, amely a Ser1704-Asp1712 régióban található (Interaction-Surface^{Leu}). Az Interaction-Surface^{NT/CT} magába foglalja a Phe1639/Phe1641 aminosavakat, amelyek hidrofób kontaktot létesítenek RAWUL Val206 aminosavával, míg az Interaction-Surface^{Leu} Leu1706 aminosava RAWUL Leu Cage-vel lép interakcióba [93].



16. ábra. A PUFID domén kölcsönhatási felszínei a RAWUL-PUFID dimerben. A PRODIGY elemzés kimutatta, hogy a PUFID domén három kölcsönhatási felszínnel rendelkezik: Interaction-Surface^{NT} (Val1636-Asn1651), Interaction-Surface^{CT} (Gly1738-Asp1748) és Interaction-Surface^{Leu} (Ser1704-Asp1712). Ezekben a szakaszokban koncentráltak az intermolekuláris kölcsönhatások. **A:** A RAWUL-PUFID dimer kölcsönhatási térképe. **B:** A RAWUL-PUFID dimer reakciós felszínei és egyéb kontaktok (PDB ID: 4HPL). **C:** A fehérje dimer aminosavszekvenciái. Kék színnel jelölve az Interaction-Surface^{NT/CT}, zölddel az egyéb interakciós aminosavak, sárgával az Interaction-Surface^{Leu}, szürkével a nem interaktáló aminosavak és pirossal a RAWUL domén kölcsönhatásban résztvevő aminosavai. **Nt:** N-terminális, **Ct:** C-terminális.

A PUFID kötődési affinitás predikciói során különös figyelmet fordítottunk a Phe1639/Phe1641 és a Leu1706 aminosavakra [93], valamint azok RAWUL doménnel való kapcsolatára.

A Kruskal-Wallis teszt szignifikáns különbséget mutatott az IS^{NT/CT} ΔG értékei között a tíz csoportban ($P < 0,0001$) (**14. ábra**). A Dunn-féle többszörös összehasonlítási teszt alapján szignifikánsan alacsonyabb ΔG értékeket kaptunk a BCOR::CCNB3-, a ZC3H7B::BCOR- és a CIITA::BCOR-PCGF1 csoportok esetében a BCOR-PCGF1-hez képest.

A Kruskal-Wallis teszt szignifikáns különbséget mutatott az IS^{NT} tartomány ΔG értékei között a tíz csoportban ($P < 0,0001$) (**14. ábra**). A Dunn-féle többszörös összehasonlítási teszt szerint a BCOR::MAML3- és az RTL9::BCOR-2-PCGF1 csoportok szignifikánsan alacsonyabb ΔG értékeket mutattak a BCOR-PCGF1-hez képest, de a 4HPL-hez viszonyítva nem. A Kruskal-Wallis teszt szintén szignifikáns különbséget mutatott a Phe1639/Phe1641 aminosavak ΔG értékei

között a tíz csoportban ($P = 0,0017$) (**14F. ábra**), de a Dunn-féle többszörös összehasonlítási teszt nem talált szignifikáns különbségeket a csoportok között, még a fúziós fehérjék esetében sem.

A Kruskal-Wallis teszt szintén szignifikáns különbséget mutatott az IS^{Leu} ΔG értékei között a tíz csoportban ($P < 0,0001$) (**14. ábra**). A Dunn-féle többszörös összehasonlítási teszt nem mutatott szignifikáns különbségeket a kontrollcsoportok és a fúziós fehérje csoport között, csak a fúziós fehérje dimerek értékei között.

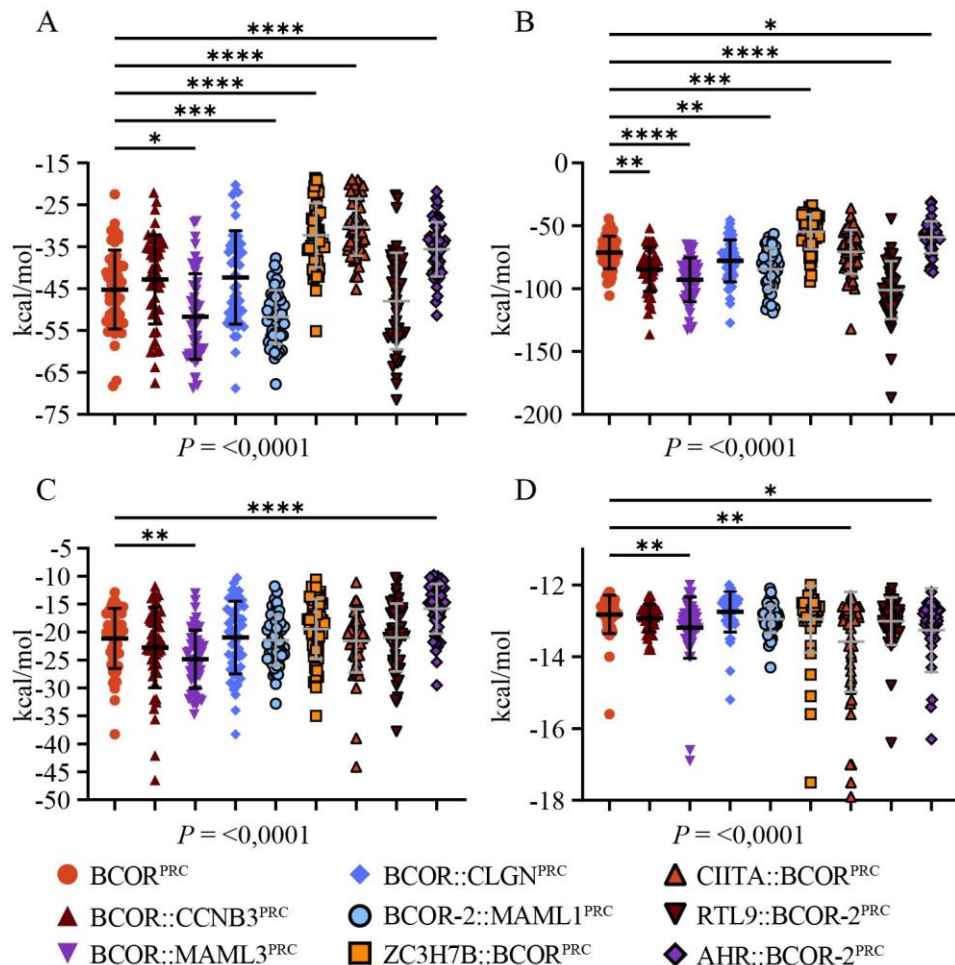
A Kruskal-Wallis teszt szignifikáns különbséget mutatott a Leu1706 aminosav ΔG értékei között a tíz csoportban ($P < 0,0001$) (**14. ábra**). Szignifikánsan gyengébb kötődési affinitást figyeltünk meg a ZC3H7B::BCOR- és a CIITA::BCOR-PCGF1 csoportokban a BCOR-PCGF1-hez képest.

Végül statisztikailag szignifikáns különbséget találtunk az IS^{CT} tartomány átlagos ΔG értékei között legalább két csoport esetében ($P < 0,0001$) (**14I. ábra**). A Holm-Šídák többszörös összehasonlítási teszt szignifikánsan alacsonyabb ΔG értékeket talált a BCOR::CCNB3- és a ZC3H7B::BCOR-PCGF1 csoportokban a BCOR-PCGF1-hez képest.

Átfogó elemzésünk szerint a BCOR-PCGF1 dimerek jelentős különbségeket mutattak kötődési affinitásukban a vad típushoz képest, különösen teljes hosszúságú szerkezetükben (**14. ábra**). Kiemelkedően négy BCOR fúziós fehérje - BCOR::CCNB3, BCOR::MAML3, BCOR-2::MAML1 és RTL9::BCOR-2 - növekedett, míg másik négy - BCOR::CLGN, ZC3H7B::BCOR, CIITA::BCOR, és AHR::BCOR-2 - csökkent kötési affinitást mutatott a vad típushoz képest. A kilenc összehasonlítás közül mindegyikben jelentős eltéréseket találtunk a csoportok között, ezek közül hét esetben a poszt-hoc tesztek statisztikailag szignifikáns különbségeket mutattak a BCOR-PCGF1 és legalább egy fúziós dimer között. A vad típushoz viszonyítva a legjelentősebb változásokat a ZC3H7B::BCOR fúziós fehérje esetében találtuk (hat esetben szignifikáns eltérés), amely két összehasonlítás során csökkent, míg három alkalommal növekedett kötődési affinitást mutatott. Négy szignifikáns eltérés volt megfigyelhető a BCOR::CCNB3 esetében, ami mindig növekedett kötődési affinitást mutatott. Az eredmények szerint az összesen végzett 72 BCOR-PCGF1 és fúziós típusú többszörös összehasonlításból 20 szignifikáns különbséget találtunk a kötődési affinitásban (28%) (**15. ábra**). Ezek az eredmények azt jelzik, hogy a fúziós fehérjék jelentősen befolyásolják a BCOR-PCGF1 dimerizációs tulajdonságait, ami potenciális funkcionális következményekkel járhat. Nem figyeltünk meg jelentős eltéréseket a kontrollcsoportként használt 4HPL-hez képest, valószínűleg annak limitált száma miatt.

6.2.7 A BRS fúziós fehérjék hatása a nem kanonikus Polycomb represszív komplex 1 kötődési affinitásaira

A BRS fehérjefúziók PRC1.1-ben betöltött szerepének további megértése érdekében modelleztük a vad típusú és fúziós fehérjéket a PCGF1, a KDM2B és az SKP1 alegységekkel. A PRODIGY segítségével kiszámított ΔG értékekkel statisztikai elemzéseket végeztünk, hogy meghatározzuk a kölcsönhatások közötti különbségeket, összehasonlítva a nyolc fúziós fehérje (Fúziós fehérje^{PRC}) kötődési affinitását egy vad típusú komplexszel (BCOR^{PRC}) (17. ábra).



17. ábra. A PRC1.1 partnerfehérjék kötődési affinitásának összehasonlítása vad típusú és fúziós BCOR fehérjék között. Az elemzés során vizsgáltuk a BRS fúziós fehérjék alegységeinek és három további PRC1.1 komponensnek (PCGF1, KDM2B, SKP1) a kötődési affinitását. **A:** A BCOR és a PCGF1 közötti kötődési affinitást Brown-Forsythe ANOVA teszttel elemeztük. **B, C és D:** A BCOR és a KDM2B, illetve a PCGF1 és a KDM2B, valamint a KDM2B és a SKP1 közötti kötődési affinitásokat Kruskal-Wallis tesztekkel vizsgáltuk. Az interakciós erősség kcal/mol-ban van kifejezve, ahol a nagyobb negatív érték erősebb interakciót jelent. * $P < 0,05$, ** $P < 0,01$, *** $P < 0,001$, **** $P < 0,0001$

A Brown-Forsythe és a Welch-féle ANOVA tesztekkel használtuk a PRC1.1 komplexekben található BCOR-PCGF1 párok ΔG értékeinek összehasonlítására (17. ábra). Az elemzés statisztikailag szignifikáns különbségeket talált a csoportok között ($P < 0,0001$). A Dunnett-féle

T3 többszörös összehasonlítási teszt szerint szignifikánsan eltérő értékeket találtunk a BCOR^{PRC} és az alábbi csoportok között: BCOR::MAML3^{PRC} ($P = 0,0114$), BCOR-2::MAML1^{PRC} ($P = 0,0008$), ZC3H7B::BCOR^{PRC} ($P < 0,0001$), CIITA::BCOR^{PRC} ($P < 0,0001$) és AHR::BCOR-2^{PRC} ($P < 0,0001$). Nem találtunk statisztikailag szignifikáns különbséget a BCOR^{PRC} és a következő csoportok között: BCOR::CCNB3^{PRC} ($P = 0,8719$), BCOR::CLGN^{PRC} ($P = 0,7544$) és RTL9::BCOR-2^{PRC} ($P = 0,8188$) (17A. ábra).

A BCOR-KDM2B fehérjék ΔG értékeinek kilenc csoport közötti összehasonlításához Kruskal-Wallis tesztet alkalmaztunk (17. ábra). Statisztikailag szignifikáns különbség volt a csoportok között ($P < 0,0001$). A Dunnett-féle többszörös összehasonlítási teszt szignifikáns különbségeket talált a BCOR^{PRC} és a következő fúziók között: BCOR::CCNB3^{PRC} ($P = 0,0080$), BCOR::MAML3^{PRC} ($P < 0,0001$), BCOR-2::MAML1^{PRC} ($P = 0,0076$), ZC3H7B::BCOR^{PRC} ($P = 0,0004$) és RTL9-BCOR-2^{PRC} ($P < 0,0001$). Nem találtunk szignifikáns különbséget a BCOR^{PRC} és a BCOR::CLGN^{PRC} ($P = 0,7025$), valamint a BCOR^{PRC} és a CIITA-BCOR^{PRC} között ($P > 0,9999$).

Elemzésre került továbbá a PCGF1-KDM2B kölcsönhatások kötődési affinitása is (17. ábra). A Kruskal-Wallis teszt statisztikailag szignifikáns különbségeket mutatott ki a csoportok között ($P < 0,0001$). A Dunnett-féle többszörös összehasonlítási tesztek szignifikáns különbségeket találtak a BCOR^{PRC} és a BCOR::MAML3^{PRC} ($P = 0,0082$), illetve a RTL9::BCOR-2^{PRC} ($P < 0,0001$) között. Nem volt statisztikailag szignifikáns különbség a BCOR^{PRC} és a másik hat fúziós fehérjét tartalmazó komplexek között ($P > 0,9999$).

Végül megvizsgáltuk és összehasonlítottuk a KDM2B-SPK1 párokat a PRC1.1 komplexben (17. ábra). A Kruskal-Wallis teszt statisztikailag szignifikáns különbségeket mutatott ki a csoportok között ($P < 0,0001$). A Dunnett-féle többszörös összehasonlítási teszt statisztikailag szignifikáns különbséget mutatott ki a BCOR^{PRC} és következő fúziós tetramerek között BCOR::MAML3^{PRC} ($P = 0,0029$), CIITA-BCOR^{PRC} ($P = 0,0047$) és AHR::BCOR-2^{PRC} ($P = 0,0283$). Nem figyeltünk meg statisztikailag szignifikáns különbséget a BCOR^{PRC} és az öt fennmaradó fúziós fehérjét tartalmazó tetramer között.

A ΔG értékek jelentősen eltértek a PRC1.1 komplexen belül lévő BCOR-PCGF1-, BCOR-KDM2B-, PCGF1-KDM2B- és KDM2B-SKP1 fehérjekölcsönhatások esetében. Az ANOVA és a Kruskal-Wallis tesztekkel végzett statisztikai elemzések során összesen 32 többszörös összehasonlítást végeztünk, ezek közül tizenhat esetben mutattunk ki változásokat a fúziós fehérjék kötődési affinitásában a vad típushoz képest (50%) (17. ábra). Nyolc fúziós fehérje közül öt (BCOR::CCNB3, BCOR::CLGN, ZC3H7B::BCOR, CIITA::BCOR és AHR::BCOR-2) csökkent kötődési affinitást mutatott, míg három esetben (BCOR::MAML3, BCOR-2::MAML1 és RTL9::BCOR-2) növekedett az affinitás. Csak két fúziós fehérje - a BCOR::MAML3 és az

AHR::BCOR-2 - mutatott statisztikailag szignifikáns eltéréseket mind a négy összehasonlításban. Eredményeink azt mutatják, hogy bár a PUFD domén szekvenciája nem trunkálódik a fúziós esemény következtében, az interakciók mégis jelentős változásokon mentek át, ami módosítja a PRC1.1 komplex interakciós dinamikáját és funkcióját.

6.2.8 A dimer komplexek molekuláris dinamikai elemzése

Az MD szimulációkat a szerkezeti és kötődési affinitás elemzések kiegészítéseként végeztük el, hogy dinamikus betekintést nyerjünk a BCOR-PCGF1 dimer komplexek - beleértve a génfúziókat is - stabilitásába, interakciós erősségébe és termodinamikai tulajdonságaiba.

A BCOR-PCGF1 dimereken - beleértve a különböző BRS fúziós fehérjéket is - végzett MD szimulációk jelentős különbségeket tártak fel a szerkezeti stabilitásban és a kölcsönhatási dinamikában. A négyzetes középeltérés (Root Mean Square Deviation, RMSD) analízisek (**18A.**, **B. ábra**) szignifikáns eltérést mutattak a dimerek szerkezeti stabilitásában, a Kruskal-Wallis teszt pedig statisztikailag szignifikáns különbségeket igazolt ($P < 0,0001$). A Dunnett-féle többszörös összehasonlító tesztek szignifikáns különbségeket találtak a BCOR-PCGF1 és az összes fúziós fehérje dimer között (**18B. ábra**).

Az RMSD analízis kimutatta, hogy a vBCOR-PCGF1, valamint a BCOR::CCNB3-PCGF1, BCOR-2::MAML1-PCGF1 és RTL9::BCOR-2-PCGF1 fúziós komplexek gyorsabban érték el a stabilitást (3000 ps-on belül), és alacsonyabb átlagos RMSD értékeket ($<0,4$ nm) tartottak fenn a szimuláció során, ami stabilabb konformációra utal. A BCOR::CCNB3-PCGF1, a BCOR-2::MAML1-PCGF1 és az RTL9::BCOR-2-PCGF1 fúziós variánsok szintén viszonylag gyorsan stabilizálódtak (3000 ps-on belül) mérsékelt RMSD értékekkel ($<0,4$ nm). Ezzel szemben más fúziós variánsok, különösen a BCOR::MAML3, a BCOR::CLGN, és főként a ZC3H7B::BCOR, CIITA::BCOR és az AHR::BCOR-2, emelkedett RMSD értékeket mutattak, és még 5000 ps után is folytatódó fluktuációkat jeleztek, ami ezeknek a fúziós fehérjéknek a tartós szerkezeti instabilitására és konformációs flexibilitására utal.

A négyzetes középfluktuáció (Root Mean Square Fluctuation, RMSF) analízis során a fúziós típusú szekvenciákat a vBCOR-hoz illesztettük. A fúziós fehérjék megváltozott fluktuációkat mutattak a vBCOR-hoz képest, különösen a funkcionálisan kritikus PUFD doménben (1634-1748). A BCOR::MAML3 és CIITA::BCOR variánsok magasabb aminosav szintű fluktuációkat mutattak ebben a doménben, ahol az RMSF csúcserkékek elérték a 0,3-0,4 nm-t, szemben a stabilabb vad típusú komplexszel (RMSF értékek $<0,3$ nm). Figyelemre méltó, hogy a legalacsonyabb fluktuációt a BCOR::CCNB3 esetében figyeltük meg (**18G. ábra**), ami alátámasztja az RMSD analízisben megfigyelt szerkezeti stabilitását.

Hasonlóképpen, a gmx_MMPBSA kötési affinitás analízisek (**18C., D. ábra**) eltérő dimer termodinamikai profilokat jeleztek. A Kruskal-Wallis teszt statisztikailag szignifikáns különbségeket igazolt ($P < 0,0001$). A Dunnett-féle többszörös összehasonlító tesztek szignifikáns különbségeket találtak a BCOR-PCGF1 és a következő fúziós fehérje dimerek között: BCOR::CLGN ($P < 0,0001$), ZC3H7B::BCOR ($P < 0,0001$), CIITA::BCOR ($P < 0,0001$), RTL9::BCOR-2 ($P = 0,0007$), AHR::BCOR-2 ($P < 0,0001$). Az összes fúziós fehérje csökkent kötési affinitást mutatott a vBCOR-PCGF1 komplexhez képest (átlagos $\Delta G = -142$ kcal/mol). A leggyengébb átlagos ΔG értéket ($-90,24$ kcal/mol) a CIITA::BCOR-PCGF1 esetben találtuk, míg a legerősebbet a BCOR::CCNB3-PCGF1 komplexre számítottuk ($-141,4$ kcal/mol) (**18D. ábra**).

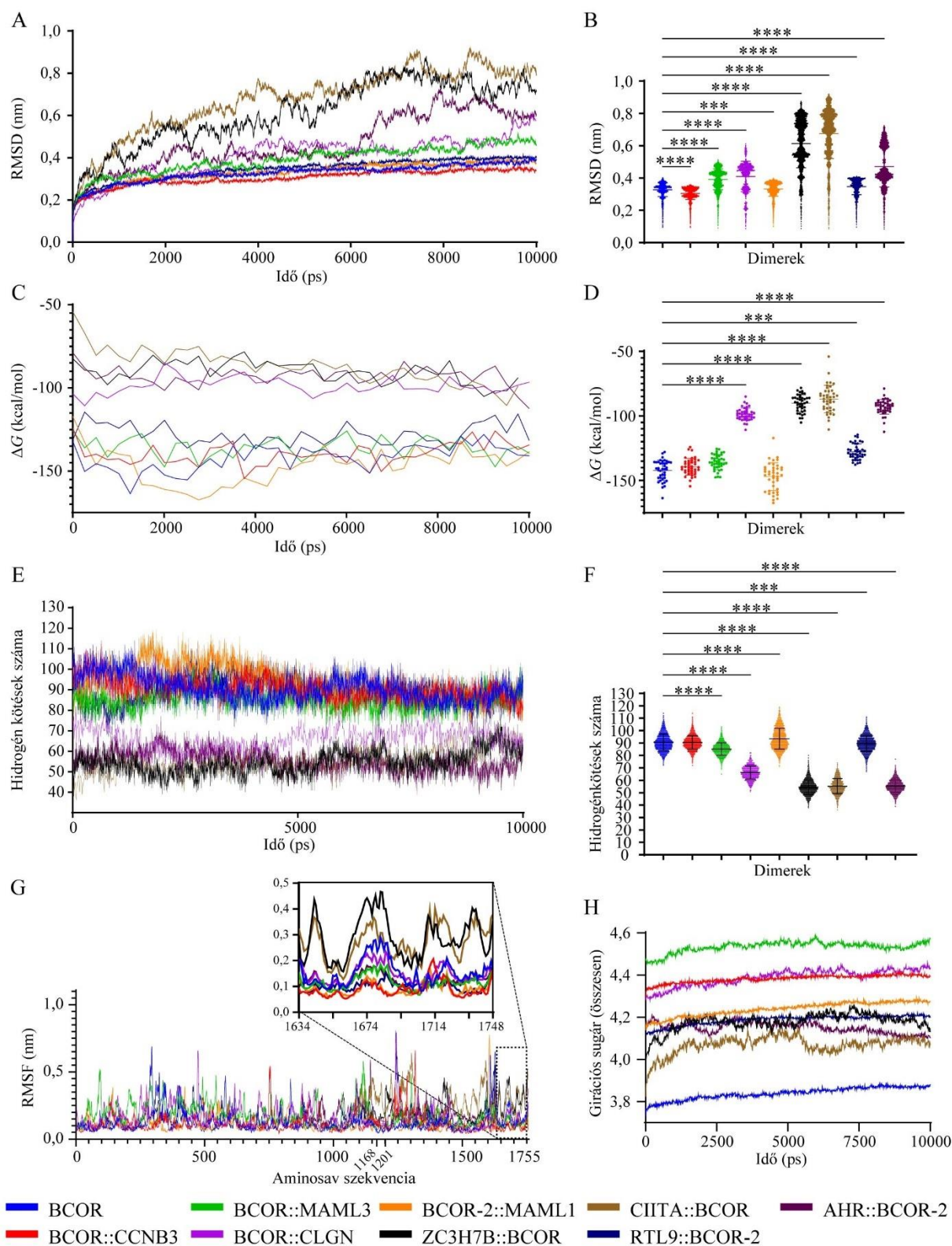
A hidrogénkötések számának meghatározása - kizárólag a két fehérjelánc között, mivel a víz által közvetített hidrogénkötéseket nem vontuk be az elsődleges elemzésbe - (**18E., F. ábra**) tovább hangsúlyozta a kölcsönhatások erősségében mutatkozó különbségeket. A Kruskal-Wallis teszt szerint szignifikáns eltérés volt a kötések számában a dimerek között ($P < 0,0001$). A post-hoc Dunnett-féle többszörös összehasonlítások statisztikailag szignifikáns különbségeket mutattak ki a BCOR-PCGF1 és az összes fúziós dimer között (a P értékek $0,0009$ és $< 0,0001$ között mozogtak), kivéve a BCOR::CCNB3 fúziót tartalmazó dimert ($P > 0,9999$). A legalacsonyabb hidrogénkötés számot a ZC3H7B::BCOR-PCGF1 dimer esetében találtuk ($54,4$ átlagérték), míg a legmagasabbat a vad típusú ($90,41$ átlagérték) és a BCOR::CCNB3-PCGF1 dimer ($90,27$ átlagérték) esetében.

A gmx_MMPBSA kötési affinitások korrelációt mutattak a PRODIGY eredményekkel (Pearson-teszt: $P = 0,00014$, $R^2 = 0,8885$). Korreláció volt továbbá a hidrogénkötések száma és az RMSD (Spearman-teszt: $P < 0,0001$, $R^2 = -0,530$), valamint a ΔG értékek között ($P < 0,007$, $R^2 = -0,417$), de nem volt korreláció az RMSD és a ΔG értékek között ($P = 0,862$). Ezek a korrelációk azt mutatják, hogy a magas RMSD értékekkel rendelkező régiók kevesebb stabil hidrogénkötésnek felelnek meg, mechanisztikus kapcsolatot teremtve a szekvenciák flexibilitása és a csökkent kölcsönhatási erősség között. A vBCOR fehérjét és fúziós partnereket, köztük a CCNB3, a MAML3, a MAML1 és az RTL9 fehérjét tartalmazó BCOR-PCGF1 dimerek egy külön csoportot alkottak magasabb hidrogénkötés számmal és erősebb kötési affinitással, összehasonlítva azokkal, amelyek a CLGN, a ZC3H7B, a CIITA és az AHR fúziós partnereket tartalmazták (**18C-D. ábra**). A kevesebb hidrogénkötéssel rendelkező dimerek gyengébb kötési affinitást mutattak, ami kevesebb stabil kölcsönhatást jelez. A ZC3H7B::BCOR, a CIITA::BCOR és az AHR::BCOR-2 fúziós fehérjék csökkent kötési affinitása annak tudható be, hogy ezekben a fúziókban a BCOR fehérje trunkálódik, ami megzavarja a kulcsfontosságú kölcsönhatási felületeket. A BCOR::CLGN esetében azonban a teljes BCOR fehérje megmarad a fúziós fehérjében. Ennek ellenére a BCOR::CLGN dimer csökkent kötési affinitást és redukált hidrogénkötés számot mutatott. Ez a

megfigyelés arra utal, hogy a CLGN partner szerkezeti változásai akadályozhatják a PCGF1-gyel való hatékony interakciót, ami potenciálisan a dimer általános stabilitásának megváltozását eredményezheti.

A girációs sugár (Radius of Gyration, Rg) analízise további betekintést nyújtott a szerkezeti kompaktságba a vBCOR-PCGF1 dimerek és fúziós variánsaik esetében (**18H. ábra**). A vad típusú komplex mutatta a legkompaktabb konformációt, a legalacsonyabb Rg értékeket (3,8-3,9 nm) produkálva, ami szoros szerkezeti illeszkedést jelez a szimuláció során. Ezzel szemben az összes fúziós fehérje magasabb Rg értékeket (4,0-4,6 nm) mutatott, ami kiterjedtebb konformációkra utal. Figyelemre méltó, hogy bár a fehérje hosszúságának változásai befolyásolták az Rg görbék magasságát, sok fúziós fehérje a szekvenciahosszával aránytalan Rg értékeket mutatott, ami arra utal, hogy ezeket a konformációs változásokat inkább szerkezeti átrendeződések, mintsem pusztán méretbeli különbségek okozzák. A BCOR::MAML3 fúzió mutatta a legmagasabb Rg értékeket (4,4-4,6 nm), jelentős szerkezeti kiterjedést sugallva, míg a CIITA::BCOR mutatta a legalacsonyabb Rg értéket a fúziós variánsok között (4,0-4,1 nm). Megfigyeltük, hogy minden rendszer 2500 ps-on belül elérte az egyensúlyt, ezt követően az Rg értékek kisebb fluktuációkkal stabilizálódtak, kivéve a CIITA::BCOR és a ZC3H7B::BCOR fúziókat. Az eredmények azt jelzik, hogy a fúziós események jelentősen megváltoztatják a BCOR komplex térbeli elrendeződését és kompaktságát, potenciálisan befolyásolva annak kölcsönhatását a PCGF1 fehérjével és a későbbi funkcionális tulajdonságokat a PRC1.1 komplex összeállításában.

Az RMSD, RMSF, hidrogénkötés-dinamika és Rg értékek együttes elemzése átfogó képet ad arról, hogyan változtatják meg a fúziós események a BCOR-PCGF1 dimer stabilitását. Két fúzió, a ZC3H7B::BCOR és a CIITA::BCOR mutatta a legkifejezettebb szerkezeti instabilitást, amelyet a legalacsonyabb kötési affinitás, a legkevesebb hidrogénkötés, a legmagasabb RMSD értékek, a legnagyobb PUF_D domén flexibilitás és a leginstabilabb Rg trajektóriák jellemezték. Ez a szerkezetdinamika funkció összefüggés magyarázza, hogy a megőrzött domén architektúrával rendelkező fúziós fehérjék miért mutatnak mégis csökkent kötési tulajdonságokat, mivel a lokális flexibilitási változások globális szerkezeti destabilizációhoz vezetnek. Ezek az eredmények összhangban vannak a **14A. ábrán** ezen dimerek esetében megfigyelt kötési affinitási tendenciákkal.



18. ábra. A BCOR-PCGF1 dimerek molekuláris dinamikai szimulációi különböző BRS fúziós fehérjékkel 10 ns időtartamon keresztül. **A:** Négyzetes középeltérés (RMSD) elemzés grafikonon. **B:** Az RMSD értékek statisztikai összehasonlítása. Vizuálisan és statisztikailag is kimutatható eltéréseket detektáltunk az RMSD értékek között **C:** A kötődési affinitás (ΔG) változása grafikonon ábrázolva. **D:** A mért ΔG adatok statisztikai összehasonlítása. A legerősebb ΔG -val rendelkező dimer a vad típusú volt. **E:** A hidrogénkötések száma grafikonon ábrázolva. **F:** A hidrogén kötések számának statisztikai összehasonlítása. A hidrogén kötések aszáma a BCOR és a BCOR::CCNB3 dimererk esetében volt a legmagasabb. **G:** A BCOR szekvenciák négyzetes középfluktuáció (RMSF) analízise, amely kiemeli a PUFD domén megváltozott flexibilitását a

fúziós variánsok esetében a vad típushoz képest. Az 1168-1201 számok a BCOR-2 fehérjeszekvenciákban lévő specifikus hiányt jelzik. **H**: A girációs sugár analízise a fúziós variánsok esetében a vad típusú dimerekhez képest a szerkezeti kiterjedés jelentős növekedését mutatja. A statisztikai szignifikanciát Kruskal-Wallis teszttel határoztuk meg, amely szignifikáns különbségeket mutatott ki ($P < 0,0001$) minden tesztben a kilenc dimer között az RMSD (**B**), ΔG (**D**) és a hidrogénkötések száma (**F**) tekintetében. *** $P < 0,001$, **** $P < 0,0001$

7. Megbeszélés

7.1 *TP53* mutációk potenciális patogén hatásának vizsgálata és klinikai relevanciája AML és MDS esetekben

A precíziós onkológia területén a nagy átírási képességű molekuláris elemzés alapvetővé vált a biológiailag eltérő betegség alcsoportok azonosításához és a leghatékonyabb kezelési lehetőségek személyre szabásához. Az MDS a klonális hematopoetikus őssejt rendellenességek csoportja, amelyek gyakran AML-be progresszióznak, így preleukémiás állapotnak tekinthetők. A szomatikus *TP53* génmutációk kulcsfontosságú tényezők az MDS/AML betegek progressziójában és betegség túlélésében [207]. Az MDS-ben a *TP53* mutációval rendelkező betegek egy egységesen rossz prognózisú molekuláris kohortot képviselnek, azonban a *TP53* mutáció státusza továbbra is az egyik legfontosabb kockázati tényező [60,62–64,208]. A mechanizmusok, amelyek révén a *TP53* mutációk ezeket a kedvezőtlen kimeneteket okozzák, még nem tisztázottak.

Munkánk során NGS analízist végeztünk a *TP53* génre célzottan 77 AML-MR, MDS-IB és MDS-LB beteg mintáin. Összesen 26 *TP53* mutációval rendelkező beteget azonosítottunk és 30 különböző variációt azonosítottunk a 41 mutációból. A *TP53* mutációk legmagasabb arányát AML esetekben detektáltuk (57,69%), majd 33,33%-ot az MDS-IB mintákban és 17,95%-ot az MDS-LB mintákban. A *TP53* mutációs státusz és citogenetikai kép összehasonlításakor az irodalmi adatokkal összehangban az AML és az MDS-IB csoportokban detektáltuk a legtöbb citogenetikai aberrációt és *TP53* mutációt [209,204]. *TP53* pozitívitas 60%-ban társult CK-val az AML-MR és 50%-ban az MDS-IB betegek esetében, míg az MDS-LB betegeknél nem detektáltunk CK-t.

A VAF és a variánsok számának együttes figyelembevételével következtetni lehet, hogy a betegmintákban egy vagy több klón van-e jelen. Hét esetben két vagy több mutációt találtunk ugyanazon mintában. A hasonló VAF (1., 10., 13., 32., 42. eset) azt sugallja, hogy párhuzamos mutációk lehetnek mindkét *TP53* allélon (összetett heterozigóta) egyetlen klónban vagy több mutáció ugyanazon az allélon multi-hit-ként értelmezhető [204]. Két vagy több, eltérő gyakorisággal észlelt *TP53* mutáció arra utal, hogy ezek különböző klónokból származnak (6., 28. eset) (2. táblázat).

A *TP53* mutációk prognosztikai jelentősége részben a variáns allélfrekvenciától függ és a kevésbé gyakori klónok kevésbé kedvezőtlen hatással bírnak [210,211]. Az irodalmi adatok szerint az alacsonyabb VAF-val rendelkező betegeknek jobb a túlélésük. Míg a *TP53* mutációk VAF >23% nagyobb halálozási kockázatot hordoznak a vad típusú betegekhez képest, addig a VAF ≤23% esetek túlélése hasonló volt a vad típusú betegekhez képest [212] (4. ábra). Eredményeink magasabb átlagos VAF-t mutattak ki az AML-MR és az MDS-IB csoportokban az MDS-LB esetekhez képest (34,12%, 35,59%, 22,83%).

A TP53 immun pozitivitását akkor állapítottuk meg, ha a TP53 festési intenzitása magas volt (3+) legalább 10%-os pozitív sejtszám mellett. Ez összesen 13/77 esetben (16,9%) fordult elő. A TP53 mutáció nélküli betegeknél nem volt TP53 IHC pozitivitás. A stop kodonnal járó esetekben (5., 16., 29., és 60.) az IHC nem detektálta a módosult TP53 fehérjét, amit a felgyorsult felezési időre utal.

A variáns fehérjék *in silico* szerkezeti elemzése feltárja a TP53 és az AML/MDS progressziójának fehérjeszintű összefüggéseit. A pontmutációk által okozott szerkezeti változások potenciálisan befolyásolják a fehérje funkcióját, és a TP53 fehérje prediktált szerkezeti változásai klonális módon összefügghetnek a klinikai viselkedéssel. A különböző mutáció csoportok eltérő hatásokat okoznak, amelyeket szekvencia- és szerkezet alapú számítógépes megközelítésekkel lehet prediktálni. Például a DBD-ben nem minden mutáció vezet feltétlenül funkcióvesztéshez. Ezeket a kategóriákat a variáns helye (N-terminális, DNS-kötő vagy oligomerizációs domén) és a mutációnak a TP53 funkciójára gyakorolt hatásai alapján határozzák meg, mint a teljes vagy részleges funkcióvesztés, a domináns-negatív hatás és/vagy a funkcionyerési tulajdonságok [213]. Összességében a TP53 mutációk 87,8%-át (36) a fehérje DBD-ben mutattuk ki.

In silico bioinformatikai módszereket alkalmaztunk a leggyakoribb TP53 hotspot mutációk validálására az alkalmazott adatbázisok segítségével, azonban a ritka mutációk AML/MDS-re gyakorolt hatása továbbra is nagyrészt ismeretlen. Ezért megvizsgáltuk a TP53 mutációk patogenitását, hogy új információkat szerezzünk a rendelkezésre álló adatok alapján, valamint stabilitási és interakciós elemzéseket végeztünk biostatistikai algoritmusok segítségével. A detektált összesen 30 mutációból 20 nincs kategorizálva a ClinVar adatbázisban, így klinikai jelentőségük továbbra is ismeretlen. Ezért *in silico* elemzéseket és adatgyűjtést végeztünk a variánsok patogenitásának meghatározására. Szignifikáns különbségeket találtunk az AML-MR és az MDS-IB, valamint az AML-MR és az MDS-LB csoportok között 12 pontozási kategória alapján (**10. ábra**). Ugyanezeket a szignifikáns különbségeket figyeltük meg a REVEL BayesDel, Varity és FATHMM-XF-re szűkített elemzési csoportban a patogenitási pontszámok esetében. A stabilitási vizsgálatok (DynaMut2, DDGun) szintén kimutatták a $\Delta\Delta G_{\text{stabilitás}}$ (kcal/mol) változásait a mutációk következtében, bár szignifikáns különbségeket (a három betegcsoport között) csak az AML-MR és az MDS-IB, valamint az MDS-IB és az MDS-LB csoportok között figyeltünk meg. Ahogy a Cry és AF alapú $\Delta\Delta G_{\text{stabilitás}}$ vizsgálati eredmények korrelációja kimutatta a rendezetlen régiók bevonása nem módosította érdemben a stabilitási eredményeket, vagyis a DNS-kötő domén szerkezeti állapota önmagában jól reprezentálja a TP53 stabilitásváltozását.

A TP53 fehérje stabilitásának *in silico* predikciója során a DynaMut2 és DDGun3D programokat olyan TP53 kristályszerkezetekre (5MCT, 5MG7, 5MF7) alapoztuk, amelyek csak a 94-293-as DNS-kötő domént tartalmazzák - ebben a régióban fordul elő a mutációk túlnyomó

többsége. Ezeken a modelleken a rendezetlen szakaszok nem szerepelnek, így azok közvetlen hatása a $\Delta\Delta G_{\text{stabilitás}}$ -ra nem jelentkezett. Ugyanakkor a teljes, rendezetlen régiókat is tartalmazó AF modell alapján végzett stabilitási predikciók nagyon magas korrelációt mutattak a Cry alapú eredményekkel (DynaMut2: $r = 0,89$; DDGun3D: $r = 1,0$). Ez azt mutatja, hogy a rendezetlen régiók jelenléte nem módosította érdemben a stabilitási eredményeket, vagyis a DBD szerkezeti állapota önmagában jól reprezentálja a TP53 stabilitásváltozását.

A különböző mutációk promóter-specifikus TA értékeit is megvizsgáltuk és szintén szignifikáns különbséget találtunk az AML-MR és az MDS-IB, illetve az AML-MR és az MDS-LB csoportok között. Az rosszabb prognózissal járó AML-MR és az enyhébb MDS-LB csoport közötti szignifikáns különbség hiánya, valamint az MDS-LB csoportban tapasztalt alacsonyabb átlagos stabilitásértékek oka valószínűleg az MDS-IB csoport kis mintaszámában keresendő. Ugyanakkor mind a patogenitási pontozási módszerek, a stabilitásváltozást előre jelző és mind pedig a promóter-specifikus TA értékeit összehasonlító módszerek is képesek voltak megkülönböztetni a három csoportban lévő mutációk előrejelzett patogenitását (**10. ábra**).

Megállapítottuk, hogy a 30 mutációtípus közül 23 variáns (p.E271K, p.S260F, p.T256I, p.P98L, p.Q375E, p.T253I, p.R248Q, p.P152Q, p.S362N, p.N239D, p.C275Y, p.V272M, p.M246V, p.M246K, p.R273S, p.G266R, p.A161T, p.C135S, p.V216M, p.G245S, p.S215N, p.Y205C és p.G334R) csökkent stabilitási értékeket mutatott. Hét variáns (p.S260F, p.T256I, p.P98L, p.Q375E, p.T253I, p.R248Q és p.P152Q) semleges hatásának bizonyult, míg egy (p.T253I) az aberráns TP53 fehérje stabilitásának növekedését mutatta a vad típushoz képest (**6. ábra**). Néhány csökkent stabilitású mutáns TP53 fehérjét klinikai vizsgálatokban is tanulmányoztak [214,215], ahol rosszabb kimenetelt igazoltak komplex kromoszóma aberrációval összefüggésben. A fennmaradó hét nem misszensz mutáció közül hat (négy kereteltolódásos és két stopkodon mutáció) olyan trunkált fehérjéket eredményezett, amelyek elvesztették a teljes C-terminális domént és sérült DBD-vel rendelkeztek. Ezek a fehérjék valószínűleg nem funkcionálisak és közvetlenül a transláció után lebomlanak. Csoportokra bontva megfigyeltük, hogy az AML-MR esetekben a 13 misszensz mutáció közül 12-nek csökkent a fehérje stabilitása (**6. ábra**). Ezen kívül két kereteltolódásos és két stopkodon mutációt is detektáltunk. Ebben a csoportban mind a hét nem misszensz mutáció és három variáns a fehérje-DNS interakciós felületein helyezkedtek el, ami gyengébb DNS-kötődést eredményez (**9. ábra**). Az MDS-IB csoportban a nyolc misszensz variáns közül hatnak semleges hatása volt a fehérjére, kettő csökkentette és egy pedig növelte annak stabilitását. Ebben a csoportban két kereteltolódásos mutációt és egy fehérje-DNS interakciós partnert is azonosítottunk (**9. ábra**). Az MDS-LB csoportban - amely a legjobb prognózissal rendelkezik - az öt misszensz variáns közül négy esetben csökkent stabilitást prediktáltunk, egy esetben pedig semlegesnek osztályoztuk, ezen kívül

egy stopkodon mutáció is jelen volt (**6. ábra**). Érdekes módon ebben a csoportban figyeltük meg mindkét fehérje-fehérje interakcióban közvetlenül részt vevő mutációt (p.G334R és p.S362N) (**8. ábra**).

In silico elemzéseket végeztünk a fehérje-fehérje és fehérje-DNS interakciók lehetséges változásainak kiszámítására (**8. ábra** és **9. ábra**). Ezek az elemzések kimutatták, hogy az oligomerizációs domén p.G334R mutációja csökkentheti a TP53 monomerek közötti intermolekuláris interakciót (-1,218 kcal/mol), ami hatással lehet a homotetramer szerkezet kialakulására (**8. ábra**). Továbbá, a korábbi tanulmányok által már vizsgált p.S362N mutáció megakadályozza az USP7/HAUSP és a TP53 fehérje közötti interakciót (**8. ábra**). A Ser362-nek nemcsak az USP7/HAUSP-hoz való kötődésben van szerepe hanem PTM hely is (foszforiláció). A Ser-362 és Ser-366 alaninra történő cseréje a TP53 IKK2 általi foszforilációjának csökkenéséhez, a TrCP1-gyel való társulás csökkenéséhez, valamint a TP53 stabilitásának és célgénjeinek upregulációjához vezet, ami megváltoztatja a sejtciklus G1 fázisát. Egy másik általunk azonosított mutáció (p.S215N) is PTM helyként funkcionál. A p.Ser215 a PAK4 kináz foszforilációs helye, ennek a helynek a módosítása csökkenti a TP53 aktivitását hepatocelluláris karcinóma sejtekben [216]. Az MDS-IB esetekben e módosítás elvesztésének hatása azonban még nem ismert.

A promóter-specifikus TA [156,217,218] adatok összhangban vannak az *in silico* eredményeinkkel (**10. ábra**), amelyek azt mutatják, hogy a p.S362N és a p.Q375E variánsok rendelkeztek a legalacsonyabb patogenitási tulajdonságokkal. A p.G334R variáns mutatta a leginkább vad típusú promóter-specifikus transzkripciós aktivitást, és az *in vitro* adatok szerint képes tetramer szerkezetet kialakítani annak ellenére, hogy az *in silico* adataink csökkent PPI affinitást prediktáltak.

A p.T253I mutáció esetében az általunk alkalmazott módszerek megnövekedett TP53 fehérjestabilitást jósoltak, ellentétben a patogenitási osztályozási rendszerek egyes eredményeivel (**6. ábra**), ami részlegesen funkcionális mutáns fehérjére utal (TransactivationClass [156]). A p.S362N variánst szinte minden módszer benignusnak értékelte (**5. ábra**), azonban - ahogy fentebb említettük - a p.S362 egy PTM hely, ezért a p.S362N mutáció megakadályozhatja ennek az aminosavnak a foszforilációját, és potenciálisan befolyásolhatja a TP53 fehérje oligomerizációját. Következésképpen a p.S362N mutációt nem lehet teljes bizonyossággal benignusnak minősíteni. Ezzel szemben a p.Q375E nagyobb valószínűséggel tekinthető benignus mutációnak az általunk gyűjtött adatok és elvégzett számítások alapján. A többi mutáció patogén jellemzőket mutatott még akkor is, ha predikciós módszereink szerint nem csökkentették a fehérjestabilitást.

7.2 A BCOR fúziók patogénitási vizsgálata *in silico* eszközökkel

A BCOR-t először, mint a BCL-6 új interakciós korepresszorát azonosították, amely fokozza a BCL-6 által közvetített transzkripciós repressziót [88]. Később felfedezték, hogy a BCOR 498-514-es helyén található a specifikus BCL6 kötődési motívum. Emellett kiderült, hogy a BCOR közvetlenül kölcsönhatásba lép az AF9 (MLLT3) [219] transzkripciós szabályozóval, és kötődik az MLL-AF9 kimérikus fehérjékhez (MLL: mixed lineage leukemia), bár csak néhány BCOR izoform képes kötődni az AF9-hez. A BCOR ismert része a PRC1.1 komplexnek. A komplex összeállítása során interakció alakul ki a BCOR PUFD doménje és a PCGF1 fehérje RAWUL doménje között. Ezenkívül a PUFD ANK ismétlődéseinek és kapcsoló régiójának is kölcsönhatásba kell lépnie a KDM2B C-terminálisával [107,99].

Egy új génfúziós alcsoportot írtak le csontszarkómában 2013-ban Pierron és munkatársai [82]. Azóta nyolc további fúziós gént azonosítottak [82,122,120,123–126], amelyeket a „csont és lágyrész differenciálatlan kis kerek sejtes szarkómáinak” harmadik különálló alcsoportjába soroltak [82,119]. Bár a patológiai jellemzőiket alaposan megvizsgálták, a fúziós fehérjéket eddig nem elemezték szerkezeti szinten. Munkánk során *in silico* megközelítéseket alkalmaztunk ezen ritka fúziós fehérjék tulajdonságainak feltárására. Ennek érdekében elemeztük a szarkómákban ismert BCOR-átrendeződések szekvenciáit, összehasonlítottuk a fúziós fehérjék doménarchitektúráit (**11. ábra**) és fizikokémiai jellemzőit (**12. ábra**), valamint elemeztük a PANTHER GO tulajdonságaikat (**5. táblázat**), a szignálpeptideket, az intramolekuláris lokalizációkat (**5. táblázat**) és az IDR-eket (**13. ábra**). A BCOR-PCGF1 komplexek 3D szerkezetét is vizsgáltuk (**14. ábra**,

15. ábra, **16. ábra** és **17. ábra**), különös tekintettel a RAWUL-PUFD domének közötti kötődési affinitásra (**14. ábra**) a BCOR és partnereik között. Emellett új RAWUL-PUFD dimer interakciós felszíneket is feltártunk (**16. ábra**). Továbbá megvizsgáltuk a fehérje-fúziók hatását a kötődési affinitásokra tetramer PRC1.1 szinten (**17. ábra**).

A kilenc fúziós fehérjét két csoportra osztottuk a BCOR lokalizációja alapján: BCOR^{NT} és BCOR^{CT}. A BCOR::CCNB3 fúziós fehérje elvesztette destrukciós boxát. Ez a motívum az ubiquitin-proteaszóma útvonal felismerési jeleként működik, és elengedhetetlen a ciklin B3 mitotikus lebontásához, ami meghosszabbított sejtciklus-progressziót és kontrollálatlan sejtosztódást eredményez [220,221]. Emellett a fúziós fehérje ektopikus expressziója NIH3T3 sejtekben rendellenes CCNB3 kifejeződést eredményezett, valamint növelte az S-fázisba belépő sejtek számát [82,222,120]. A MAML3 és MAML1 megrövidülése érintette azok 1-157-es és 1-39-es aminosav régióját (N-terminális MamL-1 domén), amely felelős az olyan Notch-fehérjék ankyrin ismétlődési régióival való kölcsönhatásért, mint például NOTCH1, NOTCH2, NOTCH3 és NOTCH4 [223–226]. A MamL-1 domén megrövidülése a BCOR::MAML3 fúzióban

megzavarja a Notch jelátvitelt, ami hozzájárul az onkogén potenciálhoz [227,228]. A BCOR::CLGN közel felét elvesztette szekvenciájának, beleértve kalretikulin szekvenciájának nagy részét is. Ez valószínűleg hatással van a kalmegin funkciójára, rontva a fehérje foldingját és megzavarva az ER chaperon funkciókat. Ez potenciálisan ER stresszt okozhat, elősegítve ezzel a tumor progresszióját a diszfunkcionális fehérjeminőség ellenőrzés révén [229], ami befolyásolhatja a férfi spermiumok funkcionális képességet is [230].

A ZC3H7B::BCOR fúzió, amely a BCOR^{CT} csoportba tartozik, jelentős szerkezeti és funkcionális változásokat mutat. A BCOR fehérje elvesztette a Bbs helyét, ezáltal megszűnt a BCL6-tal való korepresszor funkciója. Bár a ZC3H7B megőrizte TPR és LD motívumait, elvesztette C2H2- és C3H1-típusú cinkujjait. A ZC3H7B felismeri a miR-7-1, miR-16-2 és a miR-29a hajtúszerkezeit [231]. Korábban kimutatták, hogy a *ZC3H7B* hiánya csökkentette az érett miR-7 szintjét HEK293 sejtekben. Ezért a fehérje trunkáció és a C2H2, valamint a C3H1 cinkujjainak elvesztése, amelyek kölcsönhatásba léphetnek az RNS-sel [232–234], az RNS kötési képességek romlásához vezet és hatással lehet a miR-7-tel való interakcióra. Ez megzavarhatja a poszttranszkripció génszabályozást és hozzájárulhat olyan kulcsfontosságú útvonalak diszregulációjához, mint például a miRNS biogenezis [235]. A *KMT2D* és a *BCOR* gének fúziója következtében kialakuló kiméra fehérje nem tartalmazza a BCOR Bbs helyét, így elvesztette a BCL6 korepresszor funkcióját. Az elveszett ePHD domén feltételezhetően DNS kötési funkcióval rendelkezett [236], míg a FYRN/C és a SET domének kulcsfontosságúak a COMPASS komplexen belüli fehérje-fehérje interakciókhoz. Ez a komplex elengedhetetlen a hiszton H3 lizin 4 (H3K4) metilációjához, amely alapvető szerepet játszik a génszabályozásban és fejlődésben [237]. A FYRN/C, az ePHD és a SET domének elvesztése destabilizálja a kromatin átrendeződést, ami rendellenes génextpresszióhoz és genom instabilitáshoz vezet [238]. A CIITA::BCOR fúzió esetében a BCOR elvesztette Bbs helyét, míg a CIITA fehérje fő része (146-1130) hiányzik, kivéve az N-terminális régiót. Ez a változás valószínűleg hatással lehet a fehérje önasszociációjára [239], valamint az MHC II osztály fő hiszton kompatibilitási komplex transzkripció aktivációjára. Ez potenciálisan károsíthatja az MHC II osztály génextpresszióját, ami elősegítheti a tumor sejtek immunelkerülését [240]. Az RTL9 más néven sushi-ichi retrotranszpozon homológ 10 (SIRH10) vagy retrotranszpozon Gag domént tartalmazó fehérje 1 (RGAG1) az úgynevezett gag-szerű fehérjék családjába tartozik, amelyek doménjei homológok a retrovírusok és retrotranszpozonok poliproteinjeinek doménjeivel, beleértve a kapszid fehérjét is [241,206]. Az RTL9 1169-1316 régiója, amely közel van annak C-terminálisához, magában foglalja mind az N-, mind pedig a C-terminális aegységeket a kapszid-szerű doménen belül. Ez a domén nagy szerkezeti hasonlóságot mutat az emberi immundeficiencia vírus kapszid doménjével és más emlős gag-szerű fehérjék kapszid-szerű doménjeivel, mint például az aktivitás-szabályozott citoszkeleton-asszociált

fehérjék (Arc/Arg3.1). Ezekről a fehérjékről és más gag-szerű fehérjékről - beleértve az RTL család tagjait, például az RTL1-et (más néven PEG11) és az RTL2-t (más néven PEG10 vagy RGAG3) - ismert, hogy képesek önállóan összeállni vírus kapszid-szerű részecskékké kapszid-szerű doménjeik oligomerizációján keresztül [242]. Az ép RTL9 fehérje potenciálisan szintén képes lehet önálló összeállásra, bár ezt kísérletileg eddig nem igazolták. Az RTL9::BCOR-2 fúziós fehérje csak az RTL9 1-1199 aminosavait tartalmazza, ennek következtében az RTL9 kapszid-szerű doménjének szinte teljes egésze elveszett. Így az RTL9 nem képes elősegíteni a fúziós fehérje oligomerizációját kapszid-szerű doménjén keresztül. Az AHR::BCOR-2 kiméra fehérjében az AHR funkcionális doménjei nem szenvedtek deléció, azonban a BCOR szekvenciák hozzáadása valószínűleg megváltoztatja interakciós dinamikáját, potenciálisan megzavarva olyan útvonalakat, mint például xenobiotikum-metabolizmus és immunválaszok, amelyek kritikusak lehetnek tumor szupresszor folyamatok szempontjából [243] (**11. ábra**).

A fúziós fehérjék fizikokémiai tulajdonságait vizsgáltuk és összehasonlítottuk a vad típusú fehérjékével (**12. ábra**). A vizsgálat kimutatta, hogy a fúziós fehérjék GRAVY értékei 35,95%-kal csökkentek a vad típusú fehérjékhez képest (**12. ábra**). Ez a csökkenés a hidrofóbicitás csökkenésére utal, amely jelentős hatással lehet a fehérje foldingjára és interakcióira. Ez helytelen összeszerelést és funkcióvesztést eredményezhet a sejtkörnyezetben. A hidrofób kölcsönhatások kulcsszerepet játszanak a fehérje térszerkezetének stabilizálásában és a többfehérjés komplexek összeszerelésében [244]. Ezek a hatások aggregációhoz és funkcióvesztéshez vezetnek [245]. Továbbá, a hidrofóbicitás változásai befolyásolhatják a fehérje-fehérje kölcsönhatásokat és a fehérjeszerkezetek általános stabilitását, különösen stresszhelyzetekben [246]. Azonban az értékek nem értek el pozitív tartományt, ami arra utal, hogy a fehérjeszerkezetek továbbra is globuláris konfigurációt vesznek fel [181]. Az AI értékek nagyrészt változatlanok maradtak. Az II növekedése a fehérje stabilitásának és az *in vivo* felezési idő csökkenésére utal [179] (**12. ábra**). Ez az instabilitás megzavarhatja a proteosztázist, amely kritikus fontosságú a sejtfunkciók és jelátviteli útvonalak fenntartásához. A proteosztázis zavara túlterheli a sejt minőségellenőrzési mechanizmusait, beleértve a molekuláris chaperonokat és proteaszómákat, ami sejtműködési zavarhoz és betegségekhez vezet [74]. Már leírták a BCOR fehérje magas expresszióját BRS esetekben [247], ami a proteosztázis zavarával függhet össze. Kimutatták, hogy a BCOR túlzott expressziója nukleáris aggregátumok kialakulását eredményezi, amelyek eltérnek a klasszikus aggregátumoktól [248]. Tumorsejtekben ez a dereguláció elősegíti a túlélést és a proliferációt proteotoxikus stressz alatt, támogatva ezzel a tumor progresszióját [249]. Az instabilitás kiváltja az unfolded protein response (UPR) folyamatát, amelynek célja a proteosztázis helyreállítása úgy, hogy fokozza a fehérje folding kapacitást, lebontja a hibás folding-al rendelkező fehérjéket és

csökkenti a fehérjeszintézist. Azonban kimutatták, hogy az UPR krónikus aktivációja támogatja a tumorképződést azáltal, hogy lehetővé teszi a tumoros sejtek túlélését stressz hatásra [250].

A fúziós fehérjék EC értékeinek változásai (**12. ábra**) jelentős hatással lehetnek funkcionális tulajdonságaikra. Az emelkedett EC érték az aromás aminosavak - különösen triptofán, tirozin és cisztein - megnövekedett tartalmára utal, amelyek ultraibolya fényt nyelnek el 280 nm-en [177]. Az abszorpció növekedése pontosabb kvantifikálást tesz lehetővé spektrofotometriás vizsgálatokkal, ami kritikus tényező a biokémiai elemzések és fehérjejellemzés szempontjából [251]. Az aromás aminosav összetétel változásai azonban befolyásolhatják a fehérje folding-ját és stabilitását. Az aromás oldalláncok fontos szerepet játszanak azokban a hidrofób kölcsönhatásokban, amelyek stabilizálják a fehérjék harmadlagos szerkezetét [252]. Az ilyen változások hatással lehetnek folding útvonalakra, potenciálisan konformációs változásokat idézve elő, amelyek befolyásolják a fehérje biológiai aktivitását és más makromolekulákkal való kölcsönhatását. Továbbá az EC változásai jelezhetik a fehérje felszíni tulajdonságainak módosulását is, amelyek hatással lehetnek sejten belüli lokalizációjára, fehérje-fehérje kölcsönhatásokra és poszttranszlációs módosításokra való hajlamára [176].

A fúziós fehérjék pI értékei partnerfehérjéikhez viszonyított változásai (**12. ábra**) jelentős biokémiai változásokra utal. A legtöbb fúziós fehérje pI értéke csökkent a natív megfelelőikhez képest (**12. ábra**). A pI csökkenése az oldhatóság csökkenését eredményezheti fiziológias pH-n, ami befolyásolhatja a fehérje stabilitását és növelheti az aggregáció valószínűségét [253]. Az alacsonyabb pI értékű fehérjék semleges pH-n nagyobb negatív töltést hordoznak, ami hatással lehet sejten belüli lokalizációjukra, más makromolekulákkal való kölcsönhatásokra és fehérjekomplexekben való részvételükre [254]. Ezzel szemben a megfigyelt pI emelkedés pozitívabb nettó töltés felé tolódást jelez fiziológias körülmények között, ami szintén megváltoztathatja a fehérje kölcsönhatási dinamikáját és funkcionális aktivitását [255].

A SignalP 6.0 és DeepLoc 2.0 eredményeink nagyrészt összhangban voltak az irodalmi adatokkal (**13. ábra**). Azonban néhány korábban leírt szignálpeptidet ezek az eszközök nem ismertek fel. Csak a kalmegin esetében találtunk szekréción szignálpeptideket, annak ellenére, hogy az AHR két NLS-t is tartalmaz a 13-16 és a 37-42 aminosavaknál [256], amelyek nem veszték el a fúziós esemény során. A DeepLoc 2.0 eredményei szerint szinte minden fúziós fehérje a sejtmagban található, kivéve a BCOR::CLGN-t, amely az ER-ban helyezkedhet el. Ezek az eredmények összhangban vannak a PANTHER GO eredményeivel, amely szintén nukleáris lokalizációt jelez a legtöbb fúziós fehérje esetében. Csak a BCOR::CLGN esetében figyeltünk meg ER lokalizációt a DeepLoc 2.0 használatával, míg nukleáris lokalizációt mutatott a PANTHER GO. Ismert, hogy a CLGN egy N-terminális hidrofób szignálpeptidet tartalmaz [257] a korábban általunk prediktált 1-19 pozícióban [125], amely az ER lokalizációért felelős, valamint egy

hidrofób C-terminális régiót transzmembrán doménnel és ER szignállal [230]. A fúziós esemény következtében az N-terminális hidrofób szignál elveszett a fehérje első 295 aminosavával együtt, de a C-terminális helyek érintetlenek maradtak. Az N-terminális hidrofób szignálpeptid alapvető a kalmegin ER-be irányításához. Hiánya megakadályozhatja a fehérje megfelelő lokalizációját, ami funkcióvesztést eredményez. A CLGN C-terminálisa rögzíti a fehérjét az ER membránjához, biztosítva annak ottani chaperon funkcióit. Ez elengedhetetlen a calmegin ER-en belüli szerepéhez, amely magában foglalja a fehérje folding-ját, minőségellenőrzését, valamint a spermatogenezisben és spermium-petesejt kölcsönhatásokban betöltött funkcióit [230,258]. A DeepLoc 2.0 azonosította a CLGN C-terminálisát és azt az ER-ben helyezte el annak ellenére, hogy az N-terminális hiányzik (**5. táblázat**). A KMT2D és KMT2D::BCOR sejtes komponens terminusa az MLL3/4 komplex, amely egy hiszton-metiltranszferáz komplex, és a sejtmag nukleoplazmájában található (**5. táblázat**). Az RTL9::BCOR-2 esetében nem álltak rendelkezésre irodalmi adatok.

A BCOR-átrendeződéses szarkómát kis méretű daganatok jellemzik, amelyek fibrovaszkuláris stromával rendelkeznek, néha mixoid változással. A daganat sejtmagjai lehetnek kerek vagy ovális alakúak, de közös jellemzőjük a kevés citoplazma, hasonlóan az Ewing-szarkómához. Bár a fúziós fehérjék elméletileg kimutathatók lennének a citoplazmában vagy az ER-ben, minimális mennyiségük megnehezíti az immunhisztokémiai festések értelmezését.

Az IUPred3 segítségével végzett rendezetlenségi elemzés jelentős növekedést mutatott a BCOR^{NT} fúziós fehérjék PUF^D régiójában a vBCOR-hoz képest (**13. ábra**). Ez a növekedés befolyásolhatja a BCOR interakcióit a PCGF1 RAWUL doménjével és a KDM2B C-terminálisával, potenciálisan zavarva a PRC1.1 komplex kialakulását. Feltételeztük, hogy a fúziós események befolyásolják a BCOR és kötődési partnerei közötti interakciós dinamikát. Ennek vizsgálatára felépítettük a BCOR-PCGF1 és PRC1.1 komplexeket AF3-mal és értékeltük a kötési affinitásaikat (**14. ábra**, **15. ábra** és **17. ábra**). A komplexek szerkezeti elemzése jelentős eltéréseket tárt fel a kötési affinitásban mind a dimer (**14. ábra** és **15. ábra**), mind pedig a tetramer formák esetében, valamint azok interakciós partnereinél (**17. ábra**). A 92 elvégzett összehasonlításból (

15. ábra és **17. ábra**) 32 (35%) mutatott megváltozott ΔG értékeket. Továbbá két interakciós felszín azonosítottunk a BCOR-ban, IS^{NT/CT} és IS^{Leu} néven (**16. ábra**), amelyek mindegyikét érintették a fúziós események (**14. ábra**). Ezenkívül a BRS fehérjéket érintő interakciók közvetetten befolyásolták a tetramer kötődési affinitásait, különösen a PCGF1-KDM2B és a KDM2B-SKP1 között (**17. ábra**).

A fúziós fehérjék közül például a BCOR::CCNB3, a BCOR::MAML3, a BCOR-2::MAML1 és az RTL9::BCOR-2 megnövekedett kötési affinitást, míg a BCOR::CLGN, a ZC3H7B::BCOR,

a CIITA::BCOR és az AHR::BCOR-2 csökkent affinitást mutattak (**14. ábra**). A csökkent affinitás gyengült kötődést jelez, a megnövekedett affinitás fokozott, merevebb komplexre utal. Azonban a PRC1.1 által közvetített epigenetikai szabályozáshoz egyensúly szükséges a stabilitás és rugalmasság között annak érdekében, hogy fenntartsák az összeállítás és szétszerelés dinamikáját, amely elengedhetetlen a precíz génszabályozáshoz [259]. A túlzott stabilizáció akadályozhatja ezeket a dinamikus folyamatokat, rontva a sejtek jelátviteli válaszképességét és megzavarva a génexpressziót [260].

A molekuladinamikai (MD) szimulációk (10 ns) szignifikáns változásokat mutattak az RMSD-ben, a hidrogénkötés mintázatban és a kötési affinitásokban a kilenc legpontosabb dimer szerkezet között. A gmx_MMPBSA kötési affinitás predikciók eredményei erős korrelációt mutattak a PRODIGY pontszámokkal, ezáltal alátámasztva az eredmények megbízhatóságát. A kapott eredmények kiemelik a BCOR fúziós fehérjék szerkezeti és energetikai hatását a dimer stabilitására és a PCGF1 fehérjével való kölcsönhatási dinamikájára.

A fúziós fehérjék MD szimulációkban megfigyelt dinamikus viselkedése mechanisztikus betekintést nyújt abba, hogyan zavarják meg ezek az átrendeződések a PRC1.1 komplex összeállítását, ami túlmutat azon, amit csupán a statikus szerkezetekből lehetne következtetni. A megnövekedett konformációs flexibilitás konzisztens mintázata arra utal, hogy még akkor is, ha az esszenciális kötő domének megmaradnak a fúziós fehérjékben, azok dinamikus tulajdonságai alapvetően megváltoznak. A domén destabilizációja, a ZC3H7B::BCOR és a CIITA::BCOR fúziók esetében, úgy tűnik, hogy végig terjed a fehérje szerkezetén, amit a megnövekedett girációs sugaruk és csökkent kompaktságuk is bizonyít.

Ezek a dinamikai változások valószínűleg magyarázatot adnak arra a paradox megfigyelésre, hogy a teljes BCOR szekvenciát megtartó fúziós fehérjék mégis funkcionális hiányosságokat mutatnak az epigenetikai szabályozásban. A hidrogénkötés instabilitás és a csökkent kötési affinitás közötti korreláció bemutatja, hogyan befolyásolhatják drámaian az aminosav szekvencia flexibilitásában bekövetkező finom változások a kritikus kölcsönhatási felületeken a kölcsönhatás erősségét. Ez a paradigma új keretet biztosít azon mechanizmusok megértéséhez, amelyek révén a tumoros megbetegedésekkel összefüggő fúziós események megzavarják a biológiai folyamatokat, túlmutatva az egyszerű domén trunkálódás hatókörén. A tanulmány egyik korlátja, hogy hosszabb, 100 ns-os szimulációk helyett 10 ns-os MD szimulációkat alkalmaztunk. Számítási erőforrás korlátok és a kilenc BCOR-PCGF1 dimer elemzésének szükségessége miatt 10 ns-os szimulációkat alkalmaztunk, de ez az időtartam elegendő a kulcsfontosságú szerkezeti és kölcsönhatási trendek megfigyeléséhez.

A túlzott kötési affinitás olyan konformációs változásokat idézhet elő, amelyek gátolják a PRC1.1 működéséhez szükséges kofaktorok vagy enzimatis tevékenységek, például a RING1B

E3 ubikvitin ligáz aktivitását, amely kritikus a hiszton H2A ubikvitinációjához [261]. A fúziós fehérjék domináns-negatív hatást is kifejthetnek azáltal, hogy kiszorítják a vBCOR-t a PRC1.1 komplexekből, de a szerkezeti változások miatt nem képesek fenntartani a normál funkciót [262]. Az új fehérjeszregmensek beépítése a fúziók révén megzavarhatja a BCOR szabályozó szerepét, még akkor is, ha az esszenciális domének érintetlenek maradnak [263,264]. Ez a jelenség más tumoros megbetegedésekben is megfigyelhető, ahol a kiméra fehérjék megváltoztatják lokalizációjukat és funkciójukat annak ellenére, hogy szabályozó elemeket tartanak meg [265].

Az ilyen zavarok veszélyeztethetik a BCOR képességét a PRC1.1 komplex összeállítására, ami onkogének, különösen a HOX génklaszter génjeinek hatástalan represszióját eredményezi [266]. A PRC1.1 komplex károsodása csökkenti a H2A-K119 monoubiquitinációját (H2A-K119-ub1), amely egy alapvető epigenetikai jel a géncsendesítés fenntartásához [267]. Ennek eredményeként a géncsendesítési mechanizmusok zavara miatt bekövetkező HOX gének rendellenes felregulációja elősegítheti az onkogenezist azáltal, hogy kontrollálatlan sejtosztódást váltanak ki és megzavarják a normál differenciációs útvonalakat [266]. Ezenkívül ezek a *BCOR*-átrendeződések megváltoztathatják az epigenetikai mintázatot, hozzájárulva a génexpresszió diszregulációjához és a tumorgenezishez [113,114]. Így a *BCOR*-átrendeződések által okozott molekuláris változások jelentősen károsítják annak szerepét a géncsendesítésben, összességében hozzájárulva az onkológiai betegségek kialakulásához és progressziójához.

A fizikokémiai változások, például a csökkent hidrofóbicitás és a megnövekedett instabilitás súlyosbíthatják a diszfunkciót azáltal, hogy befolyásolják a fehérje folding-ját vagy elősegítik az aggregációt [245,246]. A fehérje aggregáció funkcionális komponenseket vonhat el vagy diszfunkcionális komplexeket képezhet, veszélyeztetve ezzel a PRC1.1 kromatin módosító kapacitását és génszabályozó funkcióját. Ezek az eredmények kiemelik, hogy az esszenciális domének megtartása nem elegendő, ha a fúzióból származó szekvenciák károsítják a fehérje működését vagy kölcsönhatásait.

Összességében eredményeink azt mutatják, hogy a BCOR fúziós fehérjék megzavarják a PRC1.1-en belüli molekuláris mechanizmusokat és az onkogenezist indukálnak szabályozatlan génexpresszió keresztül. Annak ellenére, hogy elméletileg képesek komplexekbe összeállni, ezek a fúziós fehérjék nem működnek hatékony epigenetikai szabályozóként, ami hangsúlyozza szerepüket a BRS patogenezisében.

A vizsgált genetikai aberrációk célzott kezelése sokoldalú megközelítést igényel, mivel jelenleg nincsenek kifejezetten a BRS-re tervezett hatékony terápiás lehetőségek [128]. Ugyanakkor elengedhetetlen annak vizsgálata, hogy a meglévő vagy potenciális terápiás szerek hogyan alkalmazhatóak. A CDK4/6 inhibitorok, például a palbociklib, ígéretesnek bizonyultak különböző szarkóma altípusok kezelésében a sejtciklus deregulációjának célzott gátlásával.

Preklinikai modellek és II. fázisú klinikai vizsgálatok is igazolták a betegség kontrollját [268,269]. A klinikai tesztek ígéretes progressziómentes és teljes túlélést mutattak előrehaladott szarkómákban ahol CDK4 túlzott expressziót figyeltek meg [269]. Ezzel párhuzamosan a hagyományos kemoterápiás protokollok, mint például a vincristin/doxorubicin/cyclophosphamid alkalmazása ifosfamid/etoposiddal, kemoterápiás érzékenységet és terápiás választ mutattak csecsemőkori BRS esetén [270]. Bár a PARP-inhibitorokat különböző daganatokban vizsgálták, azok specifikus szerepe BRS-ben - különösen HMGA1-kromatin módosítókkal vagy ATR/CHK1 inhibitorokkal kombinálva homológ rekombinációs hiány indukálására - további kutatásokat igényel [271,272]. Az epigenetikai modulációt célzó EZH2-inhibitorok és RING1B ubiquitin ligáz blokkolók célja a PRC1.1 funkció helyreállítása és a PROTAC technológiák pedig képesek lehetnek szelektíven lebontani a diszfunkcionális BCOR fúziós fehérjét, bár közvetlen bizonyítékok BRS esetében még nincsenek [273,274]. Preklinikai adatok arra utalnak, hogy a BCOR génfúzió által indukált PRC1.1 destabilizáció mieloid differenciációs gének deregulációjához vezet, azonban további kutatások szükségesek [275]. Az immunterápiás stratégiák, mint például a TCR-módosított T-sejtek (TCR-T), amelyek a tumorspecifikus antigéneket célozzák, szintén ígéretesek, különösen az olyan intracelluláris antigéneket kifejező tumorok esetében, mint a *BCOR* génfúziók [276–279]. Ezenkívül kombinált terápiák is jelentős potenciállal rendelkeznek. Egy metaanalízis szerint, a nem-Ewing kezelés (pl. doxorubicin-alapú protokollok) biztonságos opciónak tekinthető. A túlélési arányok összehasonlíthatók az Ewing-protokollokkal BRS esetén (20% és 21,8%-os mortalitás) [128]. Klinikai szempontból molekuláris diagnosztikai módszerek (pl. *BCOR* FISH/RNS-szekvenálás) nélkülözhetetlenek a fúzióspecifikus célpontok azonosításához (pl. *BCOR::CCNB3*), mivel a terápiás eredmények jelentősen függenek a genetikai variánsoktól [130,247]. E daganatok ritkasága és a konszenzusos irányelvek hiánya miatt elengedhetetlen a folyamatos kutatás (pl. NCT04889924) és a klinikai vizsgálatok, amelyek PARP- vagy checkpoint-inhibitorokat alkalmaznak. Ezek kulcsfontosságúak az optimális terápiás stratégiák kidolgozásához és a személyre szabott kezelések alkalmazásához [280–286].

8. Összefoglalás

Az értekezés célja az ismeretlen génvariánsok és fúziók vizsgálata *in silico* eszközökkel, amely során a *TP53* gén mutációit MDS-ben és AML-ben, illetve a *BCOR* fúziót vizsgáltuk BRS-ben.

Az NGS eredmények szerint a *TP53* mutációk többsége a DBD-ben helyezkedett el, amelyek jelentős változásokat okoztak a fehérje stabilitásában és funkciójában. Az *in silico* modellezés kimutatta, hogy ezek a mutációk csökkentették a fehérje stabilitását és gyengítették a DNS-kötődést, ami domináns-negatív hatást vagy funkcióvesztést eredményezett. Klinikai szempontból a *TP53* mutációk jelenléte rosszabb túlélési aránnyal társult, különösen magas allélfrekvencia esetén. Az AML-MR csoportban kimagaslóan magas volt a komplex kariotípusok előfordulása, amely tovább rontotta a prognózist. A fehérje-fehérje kölcsönhatások elemzése során kimutattuk, hogy bizonyos mutációk, például a p.G334R, gyengítették a homotetramer szerkezet kialakulását, míg mások, mint például a p.S362N, megváltoztatták az USP7/HAUSP fehérjével való kölcsönhatást. Az MDS/AML csoportokban detektált mutációkat összehasonlítva szignifikánsan kitűnt az AML-MR csoport, ahol a legtöbb és legsúlyosabb következményekkel járó rendellenességet találtuk beleértve a komplex karitípust, az ösztülélést, a patogenitási pontokat, a fehérje stabilitást és a promoter specifikus TA értékeket. Molekuláris és fehérje *in silico* adataink alapján az MDS-IB egy köztes csoport az AML-MR és az MDS-LB betegek között, amely gyakran progresszálódik AML-be, és így pre-leukémiás állapotnak tekinthető. Az egyes, nem egyértelmű klinikai jelentőségű variánsok *in silico* modellezéssel tovább értékelhetők, lehetővé téve patogenitásuk előrejelzését.

Munkánk második részében a BRS eddig felfedezett kilenc fúziós fehérjéjét vizsgáltunk *in silico* módszerekkel. A fúziós események jelentős szerkezeti változásokat, funkcionális domének elvesztését vagy módosulását okozták. A fizikokémiai tulajdonságok elemzése csökkent hidrofóbicitást és növekvő instabilitási indexet mutatott. A további szekvencia alapú analízisek is jelentős elváltozásokat eredményeztek az intracelluláris lokalizáció, a rendezetlenség növekedés, mind pedig a gének ontológiás tulajdonságainak változásában. Nyolc fúziós fehérje 3D struktúráját is felépítettük, dimer és tetramer komplexben szimulálva a PRC1.1 komplexen belüli interakcióit. Szignifikáns kötési affinitás változásokat találtunk, egyes fúziós fehérjék növekedett, míg mások gyengített kölcsönhatásokkal bírtak, ami az epigenetikai szabályozás zavarához vezet. Dimer eredményeinket az MD szimuláció is megerősítette.

Összességében az értekezés eredményei rávilágítanak arra, hogy mind a *TP53* mutációk, mind pedig a *BCOR* génfúziók jelentős szerepet játszanak a daganatok patogenezisében. Az *in silico* módszerek pedig hatékony eszközök ismeretlen variánsok és fúziók szerkezeti és funkcionális

következményeinek feltárására. Az eredményeink hozzájárulhatnak új terápiás célpontok azonosításához és személyre szabott kezelési stratégiák kidolgozásához.

9. Summary

The dissertation aims to investigate unknown gene variants and fusions using *in silico* tools, focusing on *TP53* mutations in MDS and AML and *BCOR* fusions in *BCOR*-rearranged sarcomas.

According to NGS results, most *TP53* mutations were located in the DNA-binding domain, resulting in significant changes to protein stability and function. *In silico* modeling demonstrated, these mutations reduced protein stability and weakened DNA binding, leading to dominant-negative effects or loss of function. Clinically, the presence of *TP53* mutations was associated with poorer survival rates, particularly in cases with high allele frequency. In the AML-MR group, a remarkably high occurrence of complex karyotypes was observed, further worsening the prognosis. Protein-protein interaction analysis revealed that certain mutations, such as p.G334R, weakened the formation of the homotetramer structure, while others, like p.S362N, altered interactions with the USP7/HAUSP protein. When comparing mutations detected in MDS/AML groups, the AML-MR group stood out significantly due to the most severe abnormalities found, including complex karyotypes, overall survival rates, pathogenicity scores, protein stability, and promoter-specific TA values. Based on molecular and protein *in silico* data, MDS-IB appears to be an intermediate group between AML-MR and MDS-LB patients, often progressing to AML and thus considered a pre-leukemic state. Certain variants of uncertain clinical significance can be further evaluated using *in silico* modeling to predict their pathogenicity.

In the second part of our work, we investigated nine fusion proteins identified to date in *BCOR*-rearranged sarcomas using *in silico* methods. The fusion events caused significant structural changes and loss or modification of functional domains. Analysis of physicochemical properties revealed reduced hydrophobicity and increased instability indexes. Further sequence-based analyses also indicated major alterations in intracellular localization, increased disorder, and changes in gene ontology terms. We constructed 3D structures for eight fusion proteins and simulated their interactions within dimeric and tetrameric complexes of PRC1.1. Significant changes in binding affinities were identified; some fusion proteins exhibited increased, while others showed weakened interactions, leading to disruptions in the epigenetic regulation. MD simulations confirmed our dimer results.

In summary, the dissertation's results highlight that both *TP53* mutations and *BCOR* gene fusions play significant roles in tumor pathogenesis. Moreover, *in silico* methods are effective tools for uncovering the structural and functional consequences of unknown variants and fusions. Our findings may contribute to identifying new therapeutic targets and developing personalized treatment strategies.

10. Irodalomjegyzék

1. Alaggio, R.; Amador, C.; Anagnostopoulos, I.; Attygalle, A.D.; Araujo, I.B. de O.; Berti, E.; Bhagat, G.; Borges, A.M.; Boyer, D.; Calaminici, M.; et al. The 5th Edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia* **2022**, *36*, 1720–1748, doi:10.1038/s41375-022-01620-2.
2. Bernard, E.; Tuechler, H.; Greenberg, P.L.; Hasserjian, R.P.; Arango Ossa, J.E.; Nannya, Y.; Devlin, S.M.; Creignou, M.; Pinel, P.; Monnier, L.; et al. Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. *NEJM Evidence* **2022**, *1*, EVIDoa2200008, doi:10.1056/EVIDoa2200008.
3. Döhner, H.; Wei, A.H.; Appelbaum, F.R.; Craddock, C.; DiNardo, C.D.; Dombret, H.; Ebert, B.L.; Fenaux, P.; Godley, L.A.; Hasserjian, R.P.; et al. Diagnosis and Management of AML in Adults: 2022 Recommendations from an International Expert Panel on Behalf of the ELN. *Blood* **2022**, *140*, 1345–1377, doi:10.1182/blood.2022016867.
4. Ambinder, A.J.; DeZern, A.E. Navigating the Contested Borders between Myelodysplastic Syndrome and Acute Myeloid Leukemia. *Front. Oncol.* **2022**, *12*, doi:10.3389/fonc.2022.1033534.
5. Acute Myeloid Leukemia - Cancer Stat Facts Available online: <https://seer.cancer.gov/statfacts/html/amyl.html> (accessed on 10 February 2025).
6. Ma, X. Epidemiology of Myelodysplastic Syndromes. *Am J Med* **2012**, *125*, S2–S5, doi:10.1016/j.amjmed.2012.04.014.
7. Chen, P.; Liu, X.; Zhao, Y.; Hu, Y.; Guo, J.; Wang, H. Global, National, and Regional Burden of Acute Myeloid Leukemia among 60–89 Years-Old Individuals: Insights from a Study Covering the Period 1990 to 2019. *Front Public Health* **2024**, *11*, 1329529, doi:10.3389/fpubh.2023.1329529.
8. Papaemmanuil, E.; Gerstung, M.; Bullinger, L.; Gaidzik, V.I.; Paschka, P.; Roberts, N.D.; Potter, N.E.; Heuser, M.; Thol, F.; Bolli, N.; et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med* **2016**, *374*, 2209–2221, doi:10.1056/NEJMoa1516192.
9. Pellagatti, A.; Boulwood, J. Splicing Factor Mutations in the Myelodysplastic Syndromes: Role of Key Aberrantly Spliced Genes in Disease Pathophysiology and Treatment. *Advances in Biological Regulation* **2023**, *87*, 100920, doi:10.1016/j.jbior.2022.100920.
10. Daver, N.G.; Maiti, A.; Kadia, T.M.; Vyas, P.; Majeti, R.; Wei, A.H.; Garcia-Manero, G.; Craddock, C.; Sallman, D.A.; Kantarjian, H.M. TP53-Mutated Myelodysplastic Syndrome and Acute Myeloid Leukemia: Biology, Current Therapy, and Future Directions. *Cancer Discov* **2022**, *12*, 2516–2529, doi:10.1158/2159-8290.CD-22-0332.
11. Weinberg, O.K.; Siddon, A.; Madanat, Y.F.; Gagan, J.; Arber, D.A.; Dal Cin, P.; Narayanan, D.; Ouseph, M.M.; Kurzer, J.H.; Hasserjian, R.P. TP53 Mutation Defines a Unique Subgroup within Complex Karyotype de Novo and Therapy-Related MDS/AML. *Blood Adv* **2022**, *6*, 2847–2853, doi:10.1182/bloodadvances.2021006239.
12. Adès, L. Impact of TP53 in MDS with Isolated Del(5q). *Blood* **2024**, *144*, 1655–1656, doi:10.1182/blood.2024026010.
13. Shahzad, M.; Amin, M.K.; Daver, N.G.; Shah, M.V.; Hiwase, D.; Arber, D.A.; Kharfan-Dabaja, M.A.; Badar, T. What Have We Learned about TP53-Mutated Acute Myeloid Leukemia? *Blood Cancer J.* **2024**, *14*, 1–9, doi:10.1038/s41408-024-01186-5.
14. Lucci-Cordisco, E.; Amenta, S.; Panfili, A.; Del Valle, J.; Capellá, G.; Pineda, M.; Genuardi, M. Variants of Uncertain Significance (VUS) in Cancer Predisposing Genes: What Are We Learning from Multigene Panels? *Eur J Med Genet* **2022**, *65*, 104400, doi:10.1016/j.ejmg.2021.104400.
15. Adam, F.; Fluri, M.; Scherz, A.; Rabaglio, M. Occurrence of Variants of Unknown Clinical Significance in Genetic Testing for Hereditary Breast and Ovarian Cancer Syndrome and

- Lynch Syndrome: A Literature Review and Analytical Observational Retrospective Cohort Study. *BMC Med Genomics* **2023**, *16*, 7, doi:10.1186/s12920-023-01437-7.
16. Makhnoon, S.; Levin, B.; Ensinger, M.; Mattie, K.; Volk, R.J.; Zhao, Z.; Mendoza, T.; Shete, S.; Samiian, L.; Grana, G.; et al. A Multicenter Study of Clinical Impact of Variant of Uncertain Significance Reclassification in Breast, Ovarian and Colorectal Cancer Susceptibility Genes. *Cancer Med* **2023**, *12*, 2875–2884, doi:10.1002/cam4.5202.
 17. Henkel, J.; Laner, A.; Locher, M.; Wohlfrom, T.; Neitzel, B.; Becker, K.; Neuhan, T.; Abicht, A.; Steinke-Lange, V.; Holinski-Feder, E. Diagnostic Yield and Clinical Relevance of Expanded Germline Genetic Testing for Nearly 7000 Suspected HBOC Patients. *Eur J Hum Genet* **2023**, *31*, 925–930, doi:10.1038/s41431-023-01380-2.
 18. Schram, A.M.; Chang, M.T.; Jonsson, P.; Drilon, A. Fusions in Solid Tumours: Diagnostic Strategies, Targeted Therapy, and Acquired Resistance. *Nat Rev Clin Oncol* **2017**, *14*, 735–748, doi:10.1038/nrclinonc.2017.127.
 19. Drilon, A.; Laetsch, T.W.; Kummar, S.; DuBois, S.G.; Lassen, U.N.; Demetri, G.D.; Nathenson, M.; Doebele, R.C.; Farago, A.F.; Pappo, A.S.; et al. Efficacy of Larotrectinib in TRK Fusion-Positive Cancers in Adults and Children. *New England Journal of Medicine* **2018**, *378*, 731–739, doi:10.1056/NEJMoal714448.
 20. Doebele, R.C.; Drilon, A.; Paz-Ares, L.; Siena, S.; Shaw, A.T.; Farago, A.F.; Blakely, C.M.; Seto, T.; Cho, B.C.; Tosi, D.; et al. Entrectinib in Patients with Advanced or Metastatic NTRK Fusion-Positive Solid Tumours: Integrated Analysis of Three Phase 1–2 Trials. *Lancet Oncol* **2020**, *21*, 271–282, doi:10.1016/S1470-2045(19)30691-6.
 21. Engelman, J.A.; Jänne, P.A. Mechanisms of Acquired Resistance to Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Non-Small Cell Lung Cancer. *Clinical Cancer Research* **2008**, *14*, 2895–2899, doi:10.1158/1078-0432.CCR-07-2248.
 22. Shaw, A.T.; Kim, D.-W.; Nakagawa, K.; Seto, T.; Crinó, L.; Ahn, M.-J.; Pas, T.D.; Besse, B.; Solomon, B.J.; Blackhall, F.; et al. Crizotinib versus Chemotherapy in Advanced ALK-Positive Lung Cancer. *New England Journal of Medicine* **2013**, *368*, 2385–2394, doi:10.1056/NEJMoal214886.
 23. Nicorici, D.; Şatalan, M.; Edgren, H.; Kangaspeska, S.; Murumägi, A.; Kallioniemi, O.; Virtanen, S.; Kilkku, O. FusionCatcher – a Tool for Finding Somatic Fusion Genes in Paired-End RNA-Sequencing Data 2014, 011650.
 24. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast Universal RNA-Seq Aligner. *Bioinformatics* **2013**, *29*, 15–21, doi:10.1093/bioinformatics/bts635.
 25. Chen, Y.; Wang, Y.; Chen, W.; Tan, Z.; Song, Y.; Chen, H.; Chong, Z. Gene Fusion Detection and Characterization in Long-Read Cancer Transcriptome Sequencing Data with FusionSeeker. *Cancer Res* **2023**, *83*, 28–33, doi:10.1158/0008-5472.CAN-22-1628.
 26. Bailey, M.H.; Tokheim, C.; Porta-Pardo, E.; Sengupta, S.; Bertrand, D.; Weerasinghe, A.; Colaprico, A.; Wendl, M.C.; Kim, J.; Reardon, B.; et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell* **2018**, *173*, 371–385.e18, doi:10.1016/j.cell.2018.02.060.
 27. Tamborero, D.; Rubio-Perez, C.; Deu-Pons, J.; Schroeder, M.P.; Vivancos, A.; Rovira, A.; Tusquets, I.; Albanell, J.; Rodon, J.; Tabernero, J.; et al. Cancer Genome Interpreter Annotates the Biological and Clinical Relevance of Tumor Alterations. *Genome Med* **2018**, *10*, 25, doi:10.1186/s13073-018-0531-8.
 28. Clark, A.J.; Lillard, J.W. A Comprehensive Review of Bioinformatics Tools for Genomic Biomarker Discovery Driving Precision Oncology. *Genes (Basel)* **2024**, *15*, 1036, doi:10.3390/genes15081036.
 29. Nicora, G.; Zucca, S.; Limongelli, I.; Bellazzi, R.; Magni, P. A Machine Learning Approach Based on ACMG/AMP Guidelines for Genomic Variant Classification and Prioritization. *Sci Rep* **2022**, *12*, 2517, doi:10.1038/s41598-022-06547-3.

30. Ng, P.C.; Henikoff, S. Predicting Deleterious Amino Acid Substitutions. *Genome Res* **2001**, *11*, 863–874, doi:10.1101/gr.176601.
31. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A Method and Server for Predicting Damaging Missense Mutations. *Nat Methods* **2010**, *7*, 248–249, doi:10.1038/nmeth0410-248.
32. Ioannidis, N.M.; Rothstein, J.H.; Pejaver, V.; Middha, S.; McDonnell, S.K.; Baheti, S.; Musolf, A.; Li, Q.; Holzinger, E.; Karyadi, D.; et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am J Hum Genet* **2016**, *99*, 877–885, doi:10.1016/j.ajhg.2016.08.016.
33. Rogers, M.F.; Shihab, H.A.; Mort, M.; Cooper, D.N.; Gaunt, T.R.; Campbell, C. FATHMM-XF: Accurate Prediction of Pathogenic Point Mutations via Extended Features. *Bioinformatics* **2018**, *34*, 511–513, doi:10.1093/bioinformatics/btx536.
34. Capriotti, E.; Fariselli, P.; Casadio, R. I-Mutant2.0: Predicting Stability Changes upon Mutation from the Protein Sequence or Structure. *Nucleic Acids Res* **2005**, *33*, W306–W310, doi:10.1093/nar/gki375.
35. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Židek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589, doi:10.1038/s41586-021-03819-2.
36. Rodrigues, C.H.M.; Pires, D.E.V.; Ascher, D.B. DynaMut2: Assessing Changes in Stability and Flexibility upon Single and Multiple Point Missense Mutations. *Protein Sci* **2021**, *30*, 60–69, doi:10.1002/pro.3942.
37. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A.J.; Bambrick, J.; et al. Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3. *Nature* **2024**, *630*, 493–500, doi:10.1038/s41586-024-07487-w.
38. Capriotti, E.; Fariselli, P. PhD-SNPg: A Webserver and Lightweight Tool for Scoring Single Nucleotide Variants. *Nucleic Acids Res* **2017**, *45*, W247–W252, doi:10.1093/nar/gkx369.
39. Wu, Y.; Li, R.; Sun, S.; Weile, J.; Roth, F.P. Improved Pathogenicity Prediction for Rare Human Missense Variants. *Am J Hum Genet* **2021**, *108*, 1891–1906, doi:10.1016/j.ajhg.2021.08.012.
40. Khandakji, M.N.; Mifsud, B. Gene-Specific Machine Learning Model to Predict the Pathogenicity of BRCA2 Variants. *Front Genet* **2022**, *13*, 982930, doi:10.3389/fgene.2022.982930.
41. Cheng, J.; Novati, G.; Pan, J.; Bycroft, C.; Žemgulytė, A.; Applebaum, T.; Pritzel, A.; Wong, L.H.; Zielinski, M.; Sargeant, T.; et al. Accurate Proteome-Wide Missense Variant Effect Prediction with AlphaMissense. *Science* **2023**, *381*, eadg7492, doi:10.1126/science.adg7492.
42. Acharya, D.; Mukhopadhyay, A. A Comprehensive Review of Machine Learning Techniques for Multi-Omics Data Integration: Challenges and Applications in Precision Oncology. *Briefings in Functional Genomics* **2024**, *23*, 549–560, doi:10.1093/bfgp/ela013.
43. Wang, C.; Lue, W.; Kaalia, R.; Kumar, P.; Rajapakse, J.C. Network-Based Integration of Multi-Omics Data for Clinical Outcome Prediction in Neuroblastoma. *Sci Rep* **2022**, *12*, 15425, doi:10.1038/s41598-022-19019-5.
44. Ivanisevic, T.; Sewduth, R.N. Multi-Omics Integration for the Design of Novel Therapies and the Identification of Novel Biomarkers. *Proteomes* **2023**, *11*, 34, doi:10.3390/proteomes11040034.
45. Hernández-Lemus, E.; Ochoa, S. Methods for Multi-Omic Data Integration in Cancer Research. *Front. Genet.* **2024**, *15*, doi:10.3389/fgene.2024.1425456.
46. Hollstein, M.; Rice, K.; Greenblatt, M.S.; Soussi, T.; Fuchs, R.; Sørliie, T.; Hovig, E.; Smith-Sørensen, B.; Montesano, R.; Harris, C.C. Database of P53 Gene Somatic Mutations in Human Tumors and Cell Lines. *Nucleic Acids Res* **1994**, *22*, 3551–3555.

47. Levine, A.J. P53, the Cellular Gatekeeper for Growth and Division. *Cell* **1997**, *88*, 323–331, doi:10.1016/S0092-8674(00)81871-1.
48. Jenkins, L.M.M.; Durell, S.R.; Mazur, S.J.; Appella, E. P53 N-Terminal Phosphorylation: A Defining Layer of Complex Regulation. *Carcinogenesis* **2012**, *33*, 1441–1449, doi:10.1093/carcin/bgs145.
49. Li, F.P.; Fraumeni, J.F.; Mulvihill, J.J.; Blattner, W.A.; Dreyfus, M.G.; Tucker, M.A.; Miller, R.W. A Cancer Family Syndrome in Twenty-Four Kindreds. *Cancer Res* **1988**, *48*, 5358–5362.
50. Olivier, M.; Hollstein, M.; Hainaut, P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harb Perspect Biol* **2010**, *2*, a001008, doi:10.1101/cshperspect.a001008.
51. Brosh, R.; Rotter, V. When Mutants Gain New Powers: News from the Mutant P53 Field. *Nat Rev Cancer* **2009**, *9*, 701–713, doi:10.1038/nrc2693.
52. Muller, P.A.J.; Vousden, K.H. Mutant P53 in Cancer: New Functions and Therapeutic Opportunities. *Cancer Cell* **2014**, *25*, 304–317, doi:10.1016/j.ccr.2014.01.021.
53. Kandoth, C.; McLellan, M.D.; Vandin, F.; Ye, K.; Niu, B.; Lu, C.; Xie, M.; Zhang, Q.; McMichael, J.F.; Wyczalkowski, M.A.; et al. Mutational Landscape and Significance across 12 Major Cancer Types. *Nature* **2013**, *502*, 333–339, doi:10.1038/nature12634.
54. Silwal-Pandit, L.; Vollan, H.K.M.; Chin, S.-F.; Rueda, O.M.; McKinney, S.; Osako, T.; Quigley, D.A.; Kristensen, V.N.; Aparicio, S.; Børresen-Dale, A.-L.; et al. TP53 Mutation Spectrum in Breast Cancer Is Subtype Specific and Has Distinct Prognostic Relevance. *Clinical Cancer Research* **2014**, *20*, 3569–3580, doi:10.1158/1078-0432.CCR-13-2943.
55. Iacopetta, B.; Russo, A.; Bazan, V.; Dardanoni, G.; Gebbia, N.; Soussi, T.; Kerr, D.; Elsaleh, H.; Soong, R.; Kandioler, D.; et al. Functional Categories of TP53 Mutation in Colorectal Cancer: Results of an International Collaborative Study. *Annals of Oncology* **2006**, *17*, 842–847, doi:10.1093/annonc/mdl035.
56. Zhou, Y.; Cui, G.; Xu, H.; Chun, J.; Yang, D.; Zhang, Z.; Yang, L.; Wang, J.; Wan, M.; Calvisi, D.F.; et al. Loss of TP53 Cooperates with C-MET Overexpression to Drive Hepatocarcinogenesis. *Cell Death Dis* **2023**, *14*, 476, doi:10.1038/s41419-023-05958-y.
57. Chen, Y.; Zhu, Z.; Wu, X.; Li, H.; Guan, W.; Ren, H. CHEK2 Knockout Is a Therapeutic Target for TP53-Mutated Hepatocellular Carcinoma. *Cell Death Discov* **2024**, *10*, 1–12, doi:10.1038/s41420-023-01777-4.
58. Tommasino, M.; Accardi, R.; Caldeira, S.; Dong, W.; Malanchi, I.; Smet, A.; Zehbe, I. The Role of TP53 in Cervical Carcinogenesis. *Human Mutation* **2003**, *21*, 307–312, doi:10.1002/humu.10178.
59. Rücker, F.G.; Schlenk, R.F.; Bullinger, L.; Kayser, S.; Teleanu, V.; Kett, H.; Habdank, M.; Kugler, C.-M.; Holzmann, K.; Gaidzik, V.I.; et al. TP53 Alterations in Acute Myeloid Leukemia with Complex Karyotype Correlate with Specific Copy Number Alterations, Monosomal Karyotype, and Dismal Outcome. *Blood* **2012**, *119*, 2114–2121, doi:10.1182/blood-2011-08-375758.
60. Haferlach, T.; Nagata, Y.; Grossmann, V.; Okuno, Y.; Bacher, U.; Nagae, G.; Schnittger, S.; Sanada, M.; Kon, A.; Alpermann, T.; et al. Landscape of Genetic Lesions in 944 Patients with Myelodysplastic Syndromes. *Leukemia* **2014**, *28*, 241–247, doi:10.1038/leu.2013.336.
61. Santini, V.; Stahl, M.; Sallman, D.A. TP53 Mutations in Acute Leukemias and Myelodysplastic Syndromes: Insights and Treatment Updates. *Am Soc Clin Oncol Educ Book* **2024**, *44*, e432650, doi:10.1200/EDBK_432650.
62. Bejar, R.; Stevenson, K.; Abdel-Wahab, O.; Galili, N.; Nilsson, B.; Garcia-Manero, G.; Kantarjian, H.; Raza, A.; Levine, R.L.; Neuberg, D.; et al. Clinical Effect of Point Mutations in Myelodysplastic Syndromes. *New England Journal of Medicine* **2011**, *364*, 2496–2506, doi:10.1056/NEJMoa1013343.
63. Papaemmanuil, E.; Gerstung, M.; Malcovati, L.; Tauro, S.; Gundem, G.; Van Loo, P.; Yoon, C.J.; Ellis, P.; Wedge, D.C.; Pellagatti, A.; et al. Clinical and Biological Implications of

- Driver Mutations in Myelodysplastic Syndromes. *Blood* **2013**, *122*, 3616–3627, doi:10.1182/blood-2013-08-518886.
64. Haferlach, C.; Dicker, F.; Herholz, H.; Schnittger, S.; Kern, W.; Haferlach, T. Mutations of the TP53 Gene in Acute Myeloid Leukemia Are Strongly Associated with a Complex Aberrant Karyotype. *Leukemia* **2008**, *22*, 1539–1541, doi:10.1038/leu.2008.143.
 65. Kulasekararaj, A.G.; Smith, A.E.; Mian, S.A.; Mohamedali, A.M.; Krishnamurthy, P.; Lea, N.C.; Gäken, J.; Pennaneach, C.; Ireland, R.; Czepulkowski, B.; et al. TP53 Mutations in Myelodysplastic Syndrome Are Strongly Correlated with Aberrations of Chromosome 5, and Correlate with Adverse Prognosis. *British Journal of Haematology* **2013**, *160*, 660–672, doi:10.1111/bjh.12203.
 66. Chong, E.A.; Melenhorst, J.J.; Lacey, S.F.; Ambrose, D.E.; Gonzalez, V.; Levine, B.L.; June, C.H.; Schuster, S.J. PD-1 Blockade Modulates Chimeric Antigen Receptor (CAR)–Modified T Cells: Refueling the CAR. *Blood* **2017**, *129*, 1039–1041, doi:10.1182/blood-2016-09-738245.
 67. Daver, N.; Garcia-Manero, G.; Basu, S.; Boddu, P.C.; Alfayez, M.; Cortes, J.E.; Konopleva, M.; Ravandi-Kashani, F.; Jabbour, E.; Kadia, T.; et al. Efficacy, Safety, and Biomarkers of Response to Azacitidine and Nivolumab in Relapsed/Refractory Acute Myeloid Leukemia: A Non-Randomized, Open-Label, Phase 2 Study. *Cancer Discov* **2019**, *9*, 370–383, doi:10.1158/2159-8290.CD-18-0774.
 68. Sallman, D.A.; DeZern, A.E.; Garcia-Manero, G.; Steensma, D.P.; Roboz, G.J.; Sekeres, M.A.; Cluzeau, T.; Sweet, K.L.; McLemore, A.; McGraw, K.L.; et al. Eprenetapopt (APR-246) and Azacitidine in TP53-Mutant Myelodysplastic Syndromes. *J Clin Oncol* **2021**, *39*, 1584–1594, doi:10.1200/JCO.20.02341.
 69. Liu, J.; Wang, L.; Zhao, F.; Tseng, S.; Narayanan, C.; Shura, L.; Willingham, S.; Howard, M.; Prohaska, S.; Volkmer, J.; et al. Pre-Clinical Development of a Humanized Anti-CD47 Antibody with Anti-Cancer Therapeutic Potential. *PLOS ONE* **2015**, *10*, e0137345, doi:10.1371/journal.pone.0137345.
 70. Sikic, B.I.; Lakhani, N.; Patnaik, A.; Shah, S.A.; Chandana, S.R.; Rasco, D.; Colevas, A.D.; O'Rourke, T.; Narayanan, S.; Papadopoulos, K.; et al. First-in-Human, First-in-Class Phase I Trial of the Anti-CD47 Antibody Hu5F9-G4 in Patients With Advanced Cancers. *J Clin Oncol* **2019**, *37*, 946–953, doi:10.1200/JCO.18.02018.
 71. Zhao, D.; Zarif, M.; Zhou, Q.; Capo-Chichi, J.-M.; Schuh, A.; Minden, M.D.; Atenafu, E.G.; Kumar, R.; Chang, H. TP53 Mutations in AML Patients Are Associated with Dismal Clinical Outcome Irrespective of Frontline Induction Regimen and Allogeneic Hematopoietic Cell Transplantation. *Cancers* **2023**, *15*, 3210, doi:10.3390/cancers15123210.
 72. Vardiman, J.W.; Thiele, J.; Arber, D.A.; Brunning, R.D.; Borowitz, M.J.; Porwit, A.; Harris, N.L.; Le Beau, M.M.; Hellström-Lindberg, E.; Tefferi, A.; et al. The 2008 Revision of the World Health Organization (WHO) Classification of Myeloid Neoplasms and Acute Leukemia: Rationale and Important Changes. *Blood* **2009**, *114*, 937–951, doi:10.1182/blood-2009-03-209262.
 73. SEER Hematopoietic and Lymphoid Neoplasm Database Available online: <https://seer.cancer.gov/seertools/hemelymph/51f6cf58e3e27c3994bd53ae/> (accessed on 14 January 2025).
 74. Powers, M.P. The Ever-Changing World of Gene Fusions in Cancer: A Secondary Gene Fusion and Progression. *Oncogene* **2019**, *38*, 7197–7199, doi:10.1038/s41388-019-1057-2.
 75. Taniue, K.; Akimitsu, N. Fusion Genes and RNAs in Cancer Development. *Noncoding RNA* **2021**, *7*, 10, doi:10.3390/ncrna7010010.
 76. Gillani, R.; Seong, B.K.A.; Crowdis, J.; Conway, J.R.; Dharia, N.V.; Alimohamed, S.; Haas, B.J.; Han, K.; Park, J.; Dietlein, F.; et al. Gene Fusions Create Partner and Collateral Dependencies Essential to Cancer Cell Survival. *Cancer Res* **2021**, *81*, 3971–3984, doi:10.1158/0008-5472.CAN-21-0791.

77. Khuntia, S.P.; Mukherjee, N.; Javle, V.; Tanwar, N.A.; Sreekanthreddy, P.; Varghese, L.; Gowda, P.; Kottlahouse, A.; Chandrani, P.; Choughule, A.; et al. Deep Diving Into the Fusion Across Cancer Types in the Indian Population From Formalin-Fixed Paraffin-Embedded RNA-Exome Data: A Road to Discovering Novel Rearrangements With Clinical Relevance. *JCO Glob Oncol* **2024**, e2400289, doi:10.1200/GO-24-00289.
78. Sportoletti, P.; Sorcini, D.; Falini, B. BCOR Gene Alterations in Hematologic Diseases. *Blood* **2021**, *138*, 2455–2468, doi:10.1182/blood.2021010958.
79. Borden, E.C.; Baker, L.H.; Bell, R.S.; Bramwell, V.; Demetri, G.D.; Eisenberg, B.L.; Fletcher, C.D.M.; Fletcher, J.A.; Ladanyi, M.; Meltzer, P.; et al. Soft Tissue Sarcomas of Adults: State of the Translational Science. *Clin Cancer Res* **2003**, *9*, 1941–1956.
80. Parker, B.C.; Zhang, W. Fusion Genes in Solid Tumors: An Emerging Target for Cancer Diagnosis and Treatment. *Chin J Cancer* **2013**, *32*, 594–603, doi:10.5732/cjc.013.10178.
81. Burningham, Z.; Hashibe, M.; Spector, L.; Schiffman, J.D. The Epidemiology of Sarcoma. *Clin Sarcoma Res* **2012**, *2*, 14, doi:10.1186/2045-3329-2-14.
82. Pierron, G.; Tirode, F.; Lucchesi, C.; Reynaud, S.; Ballet, S.; Cohen-Gogo, S.; Perrin, V.; Coindre, J.-M.; Delattre, O. A New Subtype of Bone Sarcoma Defined by BCOR-CCNB3 Gene Fusion. *Nat Genet* **2012**, *44*, 461–466, doi:10.1038/ng.1107.
83. Fiedorowicz, M.; Bartnik, E.; Sobczuk, P.; Teterycz, P.; Czarnecka, A.M. Molecular biology of sarcoma. *Oncology in Clinical Practice* **2018**, *14*, 307–330, doi:10.5603/OCP.2018.0045.
84. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer Statistics, 2018. *CA: A Cancer Journal for Clinicians* **2018**, *68*, 7–30, doi:10.3322/caac.21442.
85. Nacev, B.A.; Sanchez-Vega, F.; Smith, S.A.; Antonescu, C.R.; Rosenbaum, E.; Shi, H.; Tang, C.; Socci, N.D.; Rana, S.; Gularte-Mérida, R.; et al. Clinical Sequencing of Soft Tissue and Bone Sarcomas Delineates Diverse Genomic Landscapes and Potential Therapeutic Targets. *Nat Commun* **2022**, *13*, 3405, doi:10.1038/s41467-022-30453-x.
86. Mossé, Y.P.; Voss, S.D.; Lim, M.S.; Rolland, D.; Minard, C.G.; Fox, E.; Adamson, P.; Wilner, K.; Blaney, S.M.; Weigel, B.J. Targeting ALK With Crizotinib in Pediatric Anaplastic Large Cell Lymphoma and Inflammatory Myofibroblastic Tumor: A Children’s Oncology Group Study. *JCO* **2017**, *35*, 3215–3221, doi:10.1200/JCO.2017.73.4830.
87. Yamazaki, F.; Nakatani, F.; Asano, N.; Wakai, S.; Sekimizu, M.; Mitani, S.; Kubo, T.; Kawai, A.; Ichikawa, H.; Yoshida, A. Novel NTRK3 Fusions in Fibrosarcomas of Adults. *The American Journal of Surgical Pathology* **2019**, *43*, 523, doi:10.1097/PAS.0000000000001194.
88. Huynh, K.D.; Fischle, W.; Verdin, E.; Bardwell, V.J. BCoR, a Novel Corepressor Involved in BCL-6 Repression. *Genes Dev* **2000**, *14*, 1810–1823, doi:10.1101/gad.14.14.1810.
89. Pagan, J.K.; Arnold, J.; Hanchard, K.J.; Kumar, R.; Bruno, T.; Jones, M.J.K.; Richard, D.J.; Forrest, A.; Spurdle, A.; Verdin, E.; et al. A Novel Corepressor, BCoR-L1, Represses Transcription through an Interaction with CtBP. *Journal of Biological Chemistry* **2007**, *282*, 15248–15257, doi:10.1074/jbc.M700246200.
90. Astolfi, A.; Fiore, M.; Melchionda, F.; Indio, V.; Bertuccio, S.N.; Pession, A. BCOR Involvement in Cancer. *Epigenomics* **2019**, *11*, 835–855, doi:10.2217/epi-2018-0195.
91. Albagli, O.; Lantoine, D.; Quief, S.; Quignon, F.; Englert, C.; Kerckaert, J.-P.; Montarras, D.; Pinset, C.; Lindon, C. Overexpressed BCL6 (LAZ3) Oncoprotein Triggers Apoptosis, Delays S Phase Progression and Associates with Replication Foci. *Oncogene* **1999**, *18*, 5063–5075, doi:10.1038/sj.onc.1202892.
92. Wamstad, J.A.; Corcoran, C.M.; Keating, A.M.; Bardwell, V.J. Role of the Transcriptional Corepressor Bcor in Embryonic Stem Cell Differentiation and Early Embryonic Development. *PLoS ONE* **2008**, *3*, e2814, doi:10.1371/journal.pone.0002814.
93. Junco, S.E.; Wang, R.; Gaipa, J.C.; Taylor, A.B.; Schirf, V.; Gearhart, M.D.; Bardwell, V.J.; Demeler, B.; Hart, P.J.; Kim, C.A. Structure of the Polycomb Group Protein PCGF1 in

- Complex with BCOR Reveals Basis for Binding Selectivity of PCGF Homologs. *Structure* **2013**, *21*, 665–671, doi:10.1016/j.str.2013.02.013.
94. Blackledge, N.P.; Rose, N.R.; Klose, R.J. Targeting Polycomb Systems to Regulate Gene Expression: Modifications to a Complex Story. *Nat Rev Mol Cell Biol* **2015**, *16*, 643–649, doi:10.1038/nrm4067.
 95. Wang, Z.; Gearhart, M.D.; Lee, Y.-W.; Kumar, I.; Ramazanov, B.; Zhang, Y.; Hernandez, C.; Lu, A.Y.; Neuenkirchen, N.; Deng, J.; et al. A Non-Canonical BCOR-PRC1.1 Complex Represses Differentiation Programs in Human ESCs. *Cell Stem Cell* **2018**, *22*, 235–251.e9, doi:10.1016/j.stem.2017.12.002.
 96. Ghetu, A.F.; Corcoran, C.M.; Cerchietti, L.; Bardwell, V.J.; Melnick, A.; Privé, G.G. Structure of a BCOR Corepressor Peptide in Complex with the BCL6 BTB Domain Dimer. *Mol Cell* **2008**, *29*, 384–391, doi:10.1016/j.molcel.2007.12.026.
 97. Wong, C.-W.; Privalsky, M.L. Components of the SMRT Corepressor Complex Exhibit Distinctive Interactions with the POZ Domain Oncoproteins PLZF, PLZF-RAR α , and BCL-6*. *Journal of Biological Chemistry* **1998**, *273*, 27695–27702, doi:10.1074/jbc.273.42.27695.
 98. Béguelin, W.; Rivas, M.A.; Calvo Fernández, M.T.; Teater, M.; Purwada, A.; Redmond, D.; Shen, H.; Challman, M.F.; Elemento, O.; Singh, A.; et al. EZH2 Enables Germinal Centre Formation through Epigenetic Silencing of CDKN1A and an Rb-E2F1 Feedback Loop. *Nat Commun* **2017**, *8*, 877, doi:10.1038/s41467-017-01029-x.
 99. Wong, S.J.; Senkovich, O.; Artigas, J.A.; Gearhart, M.D.; Ilangovan, U.; Graham, D.W.; Abel, K.N.; Yu, T.; Hinck, A.P.; Bardwell, V.J.; et al. Structure and Role of BCOR PUFD in Noncanonical PRC1 Assembly and Disease. *Biochemistry* **2020**, *59*, 2718–2728, doi:10.1021/acs.biochem.0c00285.
 100. Chittock, E.C.; Latwiel, S.; Miller, T.C.R.; Müller, C.W. Molecular Architecture of Polycomb Repressive Complexes. *Biochemical Society Transactions* **2017**, *45*, 193–205, doi:10.1042/BST20160173.
 101. Sparmann, A.; van Lohuizen, M. Polycomb Silencers Control Cell Fate, Development and Cancer. *Nat Rev Cancer* **2006**, *6*, 846–856, doi:10.1038/nrc1991.
 102. Blackledge, N.P.; Klose, R.J. The Molecular Principles of Gene Regulation by Polycomb Repressive Complexes. *Nat Rev Mol Cell Biol* **2021**, *22*, 815–833, doi:10.1038/s41580-021-00398-y.
 103. Chen, R.; Shen, F.; Zhang, Y.; Sun, M.; Dong, Y.; Yin, Y.; Su, C.; Peng, C.; Liu, J.; Xu, J. Calcium Modulates the Tethering of BCOR-PRC1.1 Enzymatic Core to KDM2B via Liquid-Liquid Phase Separation. *Commun Biol* **2024**, *7*, 1–16, doi:10.1038/s42003-024-06820-3.
 104. Schmidt, C.R.; Achille, N.J.; Kuntimaddi, A.; Boulton, A.M.; Leach, B.I.; Zhang, S.; Zeleznik-Le, N.J.; Bushweller, J.H. BCOR Binding to MLL-AF9 Is Essential for Leukemia via Altered EYA1, SIX, and MYC Activity. *Blood Cancer Discov* **2020**, *1*, 162–177, doi:10.1158/2643-3230.BCD-20-0036.
 105. Farcas, A.M.; Blackledge, N.P.; Sudbery, I.; Long, H.K.; McGouran, J.F.; Rose, N.R.; Lee, S.; Sims, D.; Cerase, A.; Sheahan, T.W.; et al. KDM2B Links the Polycomb Repressive Complex 1 (PRC1) to Recognition of CpG Islands. *eLife* **2012**, *1*, e00205, doi:10.7554/eLife.00205.
 106. Wu, X.; Johansen, J.V.; Helin, K. Fbx110/Kdm2b Recruits Polycomb Repressive Complex 1 to CpG Islands and Regulates H2A Ubiquitylation. *Mol Cell* **2013**, *49*, 1134–1146, doi:10.1016/j.molcel.2013.01.016.
 107. Blackledge, N.P.; Farcas, A.M.; Kondo, T.; King, H.W.; McGouran, J.F.; Hanssen, L.L.P.; Ito, S.; Cooper, S.; Kondo, K.; Koseki, Y.; et al. Variant PRC1 Complex-Dependent H2A Ubiquitylation Drives PRC2 Recruitment and Polycomb Domain Formation. *Cell* **2014**, *157*, 1445–1459, doi:10.1016/j.cell.2014.05.004.

108. Wong, S.J.; Gearhart, M.D.; Taylor, A.B.; Nanyes, D.R.; Ha, D.J.; Robinson, A.K.; Artigas, J.A.; Lee, O.J.; Demeler, B.; Hart, P.J.; et al. KDM2B Recruitment of the Polycomb Group Complex, PRC1.1, Requires Cooperation between PCGF1 and BCORL1. *Structure* **2016**, *24*, 1795–1801, doi:10.1016/j.str.2016.07.011.
109. Ng, D.; Thakker, N.; Corcoran, C.M.; Donnai, D.; Perveen, R.; Schneider, A.; Hadley, D.W.; Tifft, C.; Zhang, L.; Wilkie, A.O.M.; et al. Oculofaciocardiodental and Lenz Microphthalmia Syndromes Result from Distinct Classes of Mutations in BCOR. *Nat Genet* **2004**, *36*, 411–416, doi:10.1038/ng1321.
110. Abáigar, M.; Robledo, C.; Benito, R.; Ramos, F.; Díez-Campelo, M.; Herмосín, L.; Sánchez-del-Real, J.; Alonso, J.M.; Cuello, R.; Megido, M.; et al. Chromothripsis Is a Recurrent Genomic Abnormality in High-Risk Myelodysplastic Syndromes. *PLoS One* **2016**, *11*, e0164370, doi:10.1371/journal.pone.0164370.
111. Damm, F.; Chesnais, V.; Nagata, Y.; Yoshida, K.; Scourzic, L.; Okuno, Y.; Itzykson, R.; Sanada, M.; Shiraiishi, Y.; Gelsi-Boyer, V.; et al. *BCOR* and *BCORL1* Mutations in Myelodysplastic Syndromes and Related Disorders. *Blood* **2013**, *122*, 3169–3177, doi:10.1182/blood-2012-11-469619.
112. Terada, K.; Yamaguchi, H.; Ueki, T.; Usuki, K.; Kobayashi, Y.; Tajika, K.; Gomi, S.; Kurosawa, S.; Saito, R.; Furuta, Y.; et al. Usefulness of BCOR Gene Mutation as a Prognostic Factor in Acute Myeloid Leukemia with Intermediate Cytogenetic Prognosis. *Genes, Chromosomes and Cancer* **2018**, *57*, 401–408, doi:10.1002/gcc.22542.
113. Roy, A.; Kumar, V.; Zorman, B.; Fang, E.; Haines, K.M.; Doddapaneni, H.; Hampton, O.A.; White, S.; Bavle, A.A.; Patel, N.R.; et al. Recurrent Internal Tandem Duplications of BCOR in Clear Cell Sarcoma of the Kidney. *Nat Commun* **2015**, *6*, 8891, doi:10.1038/ncomms9891.
114. Ueno-Yokohata, H.; Okita, H.; Nakasato, K.; Akimoto, S.; Hata, J.; Koshinaga, T.; Fukuzawa, M.; Kiyokawa, N. Consistent In-Frame Internal Tandem Duplications of BCOR Characterize Clear Cell Sarcoma of the Kidney. *Nat Genet* **2015**, *47*, 861–863, doi:10.1038/ng.3338.
115. Mariño-Enriquez, A.; Lauria, A.; Przybyl, J.; Ng, T.L.; Kowalewska, M.; Debiec-Rychter, M.; Ganesan, R.; Sumathi, V.; George, S.; McCluggage, W.G.; et al. BCOR Internal Tandem Duplication in High-Grade Uterine Sarcomas. *The American Journal of Surgical Pathology* **2018**, *42*, 335, doi:10.1097/PAS.0000000000000993.
116. Malik, F.; Zreik, R.T.; Hedges, D.J.; Nakitandwe, J.; Lee, S.; Ward, R.A.; McCarville, M.B.; Pappo, A.; Bahrami, A. Primary Bone Sarcoma with BCOR Internal Tandem Duplication. *Virchows Arch* **2020**, *476*, 915–920, doi:10.1007/s00428-019-02729-z.
117. Al-Ibraheemi, A.; Putra, J.; Tsai, H.K.; Cano, S.; Lip, V.; Pinches, R.S.; Restrepo, T.; Alexandrescu, S.; Janeway, K.A.; Duraisamy, S.; et al. Assessment of BCOR Internal Tandem Duplications in Pediatric Cancers by Targeted RNA Sequencing. *The Journal of Molecular Diagnostics* **2021**, *23*, 1269–1278, doi:10.1016/j.jmoldx.2021.07.006.
118. Wang, W.; Zhang, A.; Li, Y.; Wang, D.; Chen, L.; Li, Q.; Chen, J.; Li, H.; Sun, S.; Pan, M.; et al. Clinical, Pathological, and Molecular Features of Central Nervous System Tumors with BCOR Internal Tandem Duplication. *Pathology - Research and Practice* **2024**, *259*, 155367, doi:10.1016/j.prp.2024.155367.
119. Sbaraglia, M.; Bellan, E.; Dei Tos, A.P. The 2020 WHO Classification of Soft Tissue Tumours: News and Perspectives. *Pathologica* **2020**, *113*, 70–84, doi:10.32074/1591-951X-213.
120. Kao, Y.-C.; Owosho, A.A.; Sung, Y.-S.; Zhang, L.; Fujisawa, Y.; Lee, J.-C.; Wexler, L.; Argani, P.; Swanson, D.; Dickson, B.C.; et al. BCOR-CCNB3 Fusion Positive Sarcomas: A Clinicopathologic and Molecular Analysis of 36 Cases With Comparison to Morphologic Spectrum and Clinical Behavior of Other Round Cell Sarcomas. *Am J Surg Pathol* **2018**, *42*, 604–615, doi:10.1097/PAS.0000000000000965.

121. Kao, Y.-C.; Sung, Y.-S.; Zhang, L.; Huang, S.-C.; Argani, P.; Chung, C.T.; Graf, N.S.; Wright, D.C.; Kellie, S.J.; Agaram, N.P.; et al. Recurrent BCOR Internal Tandem Duplication and YWHAE-NUTM2B Fusions in Soft Tissue Undifferentiated Round Cell Sarcoma of Infancy: Overlapping Genetic Features With Clear Cell Sarcoma of Kidney. *Am J Surg Pathol* **2016**, *40*, 1009–1020, doi:10.1097/PAS.0000000000000629.
122. Specht, K.; Zhang, L.; Sung, Y.-S.; Nucci, M.; Dry, S.; Vaiyapuri, S.; Richter, G.H.; Fletcher, C.D.; Antonescu, C.R. Novel BCOR-MAML3 and ZC3H7B-BCOR Gene Fusions in Undifferentiated Small Blue Round Cell Sarcomas. *Am J Surg Pathol* **2016**, *40*, 433–442, doi:10.1097/PAS.0000000000000591.
123. Yoshida, A.; Arai, Y.; Hama, N.; Chikuta, H.; Bando, Y.; Nakano, S.; Kobayashi, E.; Shibahara, J.; Fukuhara, H.; Komiyama, M.; et al. Expanding the Clinicopathologic and Molecular Spectrum of BCOR-Associated Sarcomas in Adults. *Histopathology* **2020**, *76*, 509–520, doi:10.1111/his.14023.
124. Vasella, M.; Wagner, U.; Fritz, C.; Seidl, K.; Giudici, L.; Exner, G.U.; Moch, H.; Wild, P.J.; Bode-Lesniewska, B. Novel RGAG1-BCOR Gene Fusion Revealed in a Somatic Soft Tissue Sarcoma with a Long Follow-Up. *Virchows Arch* **2022**, *480*, 1107–1114, doi:10.1007/s00428-021-03160-z.
125. Chang Chien, Y.-C.; Madarász, K.; Csoma, S.L.; Mótyán, J.A.; Huang, H.-Y.; Méhes, G.; Mokánszki, A. Molecular Identification and In Silico Protein Analysis of a Novel BCOR-CLGN Gene Fusion in Intrathoracic BCOR-Rearranged Sarcoma. *Cancers* **2023**, *15*, 898, doi:10.3390/cancers15030898.
126. Cocchi, S.; Gambarotti, M.; Gamberi, G.; Magagnoli, G.; Maioli, M.; Stevanin, M.; Samperi, F.; Righi, A.; Benini, S. The Utility of FISH Analysis in the Diagnosis of BCOR-Rearranged Sarcomas. *Pathology - Research and Practice* **2024**, *255*, 155209, doi:10.1016/j.prp.2024.155209.
127. Sparber-Sauer, M.; Corradini, N.; Affinita, M.C.; Milano, G.M.; Pierron, G.; Carton, M.; Tirode, F.; Pissaloux, D.; Alaggio, R.; Vokuhl, C.; et al. Clinical Characteristics and Outcomes for Children, Adolescents and Young Adults with “CIC-Fused” or “BCOR-Rearranged” Soft Tissue Sarcomas: A Multi-Institutional European Retrospective Analysis. *Cancer Medicine* **2023**, *12*, 14346–14359, doi:10.1002/cam4.6113.
128. Kyriazoglou, A.; Bagos, P. Meta-Analysis of BCOR Rearranged Sarcomas: Challenging the Therapeutic Approach. *Acta Oncologica* **2021**, *60*, 721–726, doi:10.1080/0284186X.2021.1890818.
129. Suzuki, K.; Yasuda, T.; Haruhara, Y.; Watanabe, K.; Nomura, K.; Kanamori, M.; Kawaguchi, Y. BCOR-CCNB3 Sarcoma Arising in the Proximal Tibia: A Case Report. *Mol Clin Oncol* **2022**, *16*, 77, doi:10.3892/mco.2022.2510.
130. Gupta, R.K.; Rekhi, B.; Sharma, M.C.; Ganga, R.; Ravina, M.; Kumar, A.; Sahu, D. The BCOR-Rearranged Sarcoma Involving the Lung: Diagnosis with Clinical Outcome and Literature Review. *Indian Journal of Pathology and Microbiology* **2024**, *67*, 865, doi:10.4103/ijpm.ijpm_569_24.
131. Ji, T.; Li, X.; Li, Y.; Zhang, X.; Liu, Q.; Zhang, J.; Sun, N.; Liu, Z.; Liu, Y.; Wang, S.; et al. Clinical Characteristics of Children with Nasal BCOR-Rearranged Ewing-like Sarcoma: A Case Report and Literature Review. *Ear Nose Throat J* **2024**, *103*, 548–553, doi:10.1177/01455613221132392.
132. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate Prediction of Protein Structures and Interactions Using a Three-Track Neural Network. *Science* **2021**, *373*, 871–876, doi:10.1126/science.abj8754.
133. Yang, J.; Anishchenko, I.; Park, H.; Peng, Z.; Ovchinnikov, S.; Baker, D. Improved Protein Structure Prediction Using Predicted Interresidue Orientations. *Proceedings of the National Academy of Sciences* **2020**, *117*, 1496–1503, doi:10.1073/pnas.1914677117.

134. Hornbeck, P.V.; Zhang, B.; Murray, B.; Kornhauser, J.M.; Latham, V.; Skrzypek, E. PhosphoSitePlus, 2014: Mutations, PTMs and Recalibrations. *Nucleic Acids Res* **2015**, *43*, D512-520, doi:10.1093/nar/gku1267.
135. Erdős, G.; Pajkos, M.; Dosztányi, Z. IUPred3: Prediction of Protein Disorder Enhanced with Unambiguous Experimental Annotation and Visualization of Evolutionary Conservation. *Nucleic Acids Research* **2021**, *49*, W297–W303, doi:10.1093/nar/gkab408.
136. Garnier, J.; Gibrat, J.-F.; Robson, B. [32] GOR Method for Predicting Protein Secondary Structure from Amino Acid Sequence. In *Methods in Enzymology*; Computer Methods for Macromolecular Sequence Analysis; Academic Press, 1996; Vol. 266, pp. 540–553.
137. Montanucci, L.; Capriotti, E.; Frank, Y.; Ben-Tal, N.; Fariselli, P. DDGun: An Untrained Method for the Prediction of Protein Stability Changes upon Single and Multiple Point Variations. *BMC Bioinformatics* **2019**, *20*, 335, doi:10.1186/s12859-019-2923-1.
138. Montanucci, L.; Capriotti, E.; Birolo, G.; Benevenuta, S.; Pancotti, C.; Lal, D.; Fariselli, P. DDGun: An Untrained Predictor of Protein Stability Changes upon Amino Acid Variants. *Nucleic Acids Research* **2022**, *50*, W222–W227, doi:10.1093/nar/gkac325.
139. Golovenko, D.; Bräuning, B.; Vyas, P.; Haran, T.E.; Rozenberg, H.; Shakked, Z. New Insights into the Role of DNA Shape on Its Recognition by P53 Proteins. *Structure* **2018**, *26*, 1237-1250.e6, doi:10.1016/j.str.2018.06.006.
140. Mittl, P.R.; Chène, P.; Grütter, M.G. Crystallization and Structure Solution of P53 (Residues 326-356) by Molecular Replacement Using an NMR Model as Template. *Acta Crystallogr D Biol Crystallogr* **1998**, *54*, 86–89, doi:10.1107/s0907444997006550.
141. Jeffrey, P.D.; Gorina, S.; Pavletich, N.P. Crystal Structure of the Tetramerization Domain of the P53 Tumor Suppressor at 1.7 Angstroms. *Science* **1995**, *267*, 1498–1502, doi:10.1126/science.7878469.
142. Sheng, Y.; Saridakis, V.; Sarkari, F.; Duan, S.; Wu, T.; Arrowsmith, C.H.; Frappier, L. Molecular Recognition of P53 and MDM2 by USP7/HAUSP. *Nat Struct Mol Biol* **2006**, *13*, 285–291, doi:10.1038/nsmb1067.
143. Avalos, J.L.; Bever, K.M.; Wolberger, C. Mechanism of Sirtuin Inhibition by Nicotinamide: Altering the NAD(+) Cosubstrate Specificity of a Sir2 Enzyme. *Mol Cell* **2005**, *17*, 855–868, doi:10.1016/j.molcel.2005.02.022.
144. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Yuan, D.; Stroe, O.; Wood, G.; Laydon, A.; et al. AlphaFold Protein Structure Database: Massively Expanding the Structural Coverage of Protein-Sequence Space with High-Accuracy Models. *Nucleic Acids Research* **2022**, *50*, D439–D444, doi:10.1093/nar/gkab1061.
145. Joerger, A.C.; Ang, H.C.; Fersht, A.R. Structural Basis for Understanding Oncogenic P53 Mutations and Designing Rescue Drugs. *Proc Natl Acad Sci U S A* **2006**, *103*, 15056–15061, doi:10.1073/pnas.0607286103.
146. Bullock, A.N.; Henckel, J.; Fersht, A.R. Quantitative Analysis of Residual Folding and DNA Binding in Mutant P53 Core Domain: Definition of Mutant States for Rescue in Cancer Therapy. *Oncogene* **2000**, *19*, 1245–1256, doi:10.1038/sj.onc.1203434.
147. Rodrigues, C.H.M.; Myung, Y.; Pires, D.E.V.; Ascher, D.B. mCSM-PPI2: Predicting the Effects of Mutations on Protein–Protein Interactions. *Nucleic Acids Research* **2019**, *47*, W338–W344, doi:10.1093/nar/gkz383.
148. Clore, G.M.; Omichinski, J.G.; Sakaguchi, K.; Zambrano, N.; Sakamoto, H.; Appella, E.; Gronenborn, A.M. High-Resolution Structure of the Oligomerization Domain of P53 by Multidimensional NMR. *Science* **1994**, *265*, 386–391, doi:10.1126/science.8023159.
149. Clore, G.M.; Ernst, J.; Clubb, R.; Omichinski, J.G.; Kennedy, W.M.; Sakaguchi, K.; Appella, E.; Gronenborn, A.M. Refined Solution Structure of the Oligomerization Domain of the Tumour Suppressor P53. *Nat Struct Biol* **1995**, *2*, 321–333, doi:10.1038/nsb0495-321.

150. Pires, D.E.V.; Ascher, D.B. mCSM–NA: Predicting the Effects of Mutations on Protein–Nucleic Acids Interactions. *Nucleic Acids Research* **2017**, *45*, W241–W246, doi:10.1093/nar/gkx236.
151. de Andrade, K.C.; Lee, E.E.; Tookmanian, E.M.; Kesserwan, C.A.; Manfredi, J.J.; Hatton, J.N.; Loukissas, J.K.; Zavadil, J.; Zhou, L.; Olivier, M.; et al. The TP53 Database: Transition from the International Agency for Research on Cancer to the US National Cancer Institute. *Cell Death Differ* **2022**, *29*, 1071–1073, doi:10.1038/s41418-022-00976-3.
152. Mathe, E.; Olivier, M.; Kato, S.; Ishioka, C.; Hainaut, P.; Tavtigian, S.V. Computational Approaches for Predicting the Biological Effect of P53 Missense Mutations: A Comparison of Three Sequence Analysis Based Methods. *Nucleic Acids Res* **2006**, *34*, 1317–1325, doi:10.1093/nar/gkj518.
153. Tavtigian, S.V.; Deffenbaugh, A.M.; Yin, L.; Judkins, T.; Scholl, T.; Samollow, P.B.; de Silva, D.; Zharkikh, A.; Thomas, A. Comprehensive Statistical Study of 452 BRCA1 Missense Substitutions with Classification of Eight Recurrent Substitutions as Neutral. *J Med Genet* **2006**, *43*, 295–305, doi:10.1136/jmg.2005.033878.
154. Feng, B.-J. PERCH: A Unified Framework for Disease Gene Prioritization. *Hum Mutat* **2017**, *38*, 243–251, doi:10.1002/humu.23158.
155. Fortunato, C.; James, P.A.; Young, E.L.; Feng, B.; Olivier, M.; Pesaran, T.; Tavtigian, S.V.; Spurdle, A.B. Improved, ACMG-Compliant, in Silico Prediction of Pathogenicity for Missense Substitutions Encoded by TP53 Variants. *Hum Mutat* **2018**, *39*, 1061–1069, doi:10.1002/humu.23553.
156. Kato, S.; Han, S.-Y.; Liu, W.; Otsuka, K.; Shibata, H.; Kanamaru, R.; Ishioka, C. Understanding the Function–Structure and Function–Mutation Relationships of P53 Tumor Suppressor Protein by High-Resolution Missense Mutation Analysis. *Proc Natl Acad Sci U S A* **2003**, *100*, 8424–8429, doi:10.1073/pnas.1431692100.
157. Giacomelli, A.O.; Yang, X.; Lintner, R.E.; McFarland, J.M.; Duby, M.; Kim, J.; Howard, T.P.; Takeda, D.Y.; Ly, S.H.; Kim, E.; et al. Mutational Processes Shape the Landscape of TP53 Mutations in Human Cancer. *Nat Genet* **2018**, *50*, 1381–1387, doi:10.1038/s41588-018-0204-y.
158. Dearth, L.R.; Qian, H.; Wang, T.; Baroni, T.E.; Zeng, J.; Chen, S.W.; Yi, S.Y.; Brachmann, R.K. Inactive Full-Length P53 Mutants Lacking Dominant Wild-Type P53 Inhibition Highlight Loss of Heterozygosity as an Important Aspect of P53 Status in Human Cancers. *Carcinogenesis* **2007**, *28*, 289–298, doi:10.1093/carcin/bgl132.
159. Monti, P.; Perfumo, C.; Bisio, A.; Ciribilli, Y.; Menichini, P.; Russo, D.; Umbach, D.M.; Resnick, M.A.; Inga, A.; Fronza, G. Dominant-Negative Features of Mutant TP53 in Germline Carriers Have Limited Impact on Cancer Outcomes. *Mol Cancer Res* **2011**, *9*, 271–279, doi:10.1158/1541-7786.MCR-10-0496.
160. Landrum, M.J.; Chitipiralla, S.; Brown, G.R.; Chen, C.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; Kaur, K.; Liu, C.; et al. ClinVar: Improvements to Accessing Data. *Nucleic Acids Res* **2020**, *48*, D835–D844, doi:10.1093/nar/gkz972.
161. Landrum, M.J.; Kattman, B.L. ClinVar at Five Years: Delivering on the Promise. *Hum Mutat* **2018**, *39*, 1623–1630, doi:10.1002/humu.23641.
162. Landrum, M.J.; Lee, J.M.; Benson, M.; Brown, G.R.; Chao, C.; Chitipiralla, S.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; et al. ClinVar: Improving Access to Variant Interpretations and Supporting Evidence. *Nucleic Acids Res* **2018**, *46*, D1062–D1067, doi:10.1093/nar/gkx1153.
163. Landrum, M.J.; Lee, J.M.; Benson, M.; Brown, G.; Chao, C.; Chitipiralla, S.; Gu, B.; Hart, J.; Hoffman, D.; Hoover, J.; et al. ClinVar: Public Archive of Interpretations of Clinically Relevant Variants. *Nucleic Acids Res* **2016**, *44*, D862–868, doi:10.1093/nar/gkv1222.
164. Landrum, M.J.; Lee, J.M.; Riley, G.R.; Jang, W.; Rubinstein, W.S.; Church, D.M.; Maglott, D.R. ClinVar: Public Archive of Relationships among Sequence Variation and Human Phenotype. *Nucleic Acids Res* **2014**, *42*, D980–985, doi:10.1093/nar/gkt1113.

165. Martin, F.J.; Amode, M.R.; Aneja, A.; Austine-Orimoloye, O.; Azov, A.G.; Barnes, I.; Becker, A.; Bennett, R.; Berry, A.; Bhai, J.; et al. Ensembl 2023. *Nucleic Acids Research* **2023**, *51*, D933–D941, doi:10.1093/nar/gkac958.
166. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A. ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Res* **2003**, *31*, 3784–3788.
167. The UniProt Consortium UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* **2023**, *51*, D523–D531, doi:10.1093/nar/gkac1052.
168. Bruford, E.A.; Antonescu, C.R.; Carroll, A.J.; Chinnaiyan, A.; Cree, I.A.; Cross, N.C.P.; Dagleish, R.; Gale, R.P.; Harrison, C.J.; Hastings, R.J.; et al. HUGO Gene Nomenclature Committee (HGNC) Recommendations for the Designation of Gene Fusions. *Leukemia* **2021**, *35*, 3040–3043, doi:10.1038/s41375-021-01436-6.
169. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The Conserved Domain Database in 2020. *Nucleic Acids Res* **2020**, *48*, D265–D268, doi:10.1093/nar/gkz991.
170. Paysan-Lafosse, T.; Blum, M.; Chuguransky, S.; Grego, T.; Pinto, B.L.; Salazar, G.A.; Bileschi, M.L.; Bork, P.; Bridge, A.; Colwell, L.; et al. InterPro in 2022. *Nucleic Acids Research* **2023**, *51*, D418–D427, doi:10.1093/nar/gkac993.
171. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: Tool for the Unification of Biology. *Nat Genet* **2000**, *25*, 25–29, doi:10.1038/75556.
172. Mi, H.; Muruganujan, A.; Huang, X.; Ebert, D.; Mills, C.; Guo, X.; Thomas, P.D. Protocol Update for Large-Scale Genome and Gene Function Analysis with the PANTHER Classification System (v.14.0). *Nat Protoc* **2019**, *14*, 703–721, doi:10.1038/s41596-019-0128-8.
173. The Gene Ontology Consortium; Aleksander, S.A.; Balhoff, J.; Carbon, S.; Cherry, J.M.; Drabkin, H.J.; Ebert, D.; Feuermann, M.; Gaudet, P.; Harris, N.L.; et al. The Gene Ontology Knowledgebase in 2023. *Genetics* **2023**, *224*, iyad031, doi:10.1093/genetics/iyad031.
174. Mi, H.; Muruganujan, A.; Ebert, D.; Huang, X.; Thomas, P.D. PANTHER Version 14: More Genomes, a New PANTHER GO-Slim and Improvements in Enrichment Analysis Tools. *Nucleic Acids Research* **2019**, *47*, D419–D426, doi:10.1093/nar/gky1038.
175. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*; Walker, J.M., Ed.; Humana Press: Totowa, NJ, 2005; pp. 571–607 ISBN 978-1-58829-343-5.
176. Edelhoch, H. Spectroscopic Determination of Tryptophan and Tyrosine in Proteins. *Biochemistry* **1967**, *6*, 1948–1954, doi:10.1021/bi00859a010.
177. Gill, S.C.; von Hippel, P.H. Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data. *Analytical Biochemistry* **1989**, *182*, 319–326, doi:10.1016/0003-2697(89)90602-7.
178. Pace, C.N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to Measure and Predict the Molar Absorption Coefficient of a Protein. *Protein Sci* **1995**, *4*, 2411–2423.
179. Guruprasad, K.; Reddy, B.V.; Pandit, M.W. Correlation between Stability of a Protein and Its Dipeptide Composition: A Novel Approach for Predicting in Vivo Stability of a Protein from Its Primary Sequence. *Protein Eng* **1990**, *4*, 155–161, doi:10.1093/protein/4.2.155.
180. Ikai, A. Thermostability and Aliphatic Index of Globular Proteins. *J Biochem* **1980**, *88*, 1895–1898.
181. Kyte, J.; Doolittle, R.F. A Simple Method for Displaying the Hydropathic Character of a Protein. *J Mol Biol* **1982**, *157*, 105–132, doi:10.1016/0022-2836(82)90515-0.
182. Teufel, F.; Almagro Armenteros, J.J.; Johansen, A.R.; Gíslason, M.H.; Pihl, S.I.; Tsirigos, K.D.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 Predicts All Five

- Types of Signal Peptides Using Protein Language Models. *Nat Biotechnol* **2022**, *40*, 1023–1025, doi:10.1038/s41587-021-01156-3.
183. Thumulari, V.; Almagro Armenteros, J.J.; Johansen, A.R.; Nielsen, H.; Winther, O. DeepLoc 2.0: Multi-Label Subcellular Localization Prediction Using Protein Language Models. *Nucleic Acids Research* **2022**, *50*, W228–W234, doi:10.1093/nar/gkac278.
 184. Thul, P.J.; Åkesson, L.; Wiking, M.; Mahdessian, D.; Geladaki, A.; Ait Blal, H.; Alm, T.; Asplund, A.; Björk, L.; Breckels, L.M.; et al. A Subcellular Map of the Human Proteome. *Science* **2017**, *356*, eaal3321, doi:10.1126/science.aal3321.
 185. Xue, L.C.; Rodrigues, J.P.; Kastriitis, P.L.; Bonvin, A.M.; Vangone, A. PRODIGY: A Web Server for Predicting the Binding Affinity of Protein–Protein Complexes. *Bioinformatics* **2016**, *32*, 3676–3678, doi:10.1093/bioinformatics/btw514.
 186. The PyMOL Molecular Graphics System, Version 2.5, Schrödinger, LLC.
 187. Badaczewska-Dawid, A.E.; Nithin, C.; Wroblewski, K.; Kurcinski, M.; Kmiecik, S. MAPIYA Contact Map Server for Identification and Visualization of Molecular Interactions in Proteins and Biological Complexes. *Nucleic Acids Research* **2022**, *50*, W474–W482, doi:10.1093/nar/gkac307.
 188. Burley, S.K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chao, H.; Chen, L.; Craig, P.A.; Crichlow, G.V.; Dalenberg, K.; Duarte, J.M.; et al. RCSB Protein Data Bank (RCSB.Org): Delivery of Experimentally-Determined PDB Structures alongside One Million Computed Structure Models of Proteins from Artificial Intelligence/Machine Learning. *Nucleic Acids Research* **2023**, *51*, D488–D508, doi:10.1093/nar/gkac1077.
 189. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25, doi:10.1016/j.softx.2015.06.001.
 190. Páll, S.; Abraham, M.J.; Kutzner, C.; Hess, B.; Lindahl, E. Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS. In *Proceedings of the Solving Software Challenges for Exascale*; Markidis, S., Laure, E., Eds.; Springer International Publishing: Cham, 2015; pp. 3–27.
 191. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M.R.; Smith, J.C.; Kasson, P.M.; van der Spoel, D.; et al. GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit. *Bioinformatics* **2013**, *29*, 845–854, doi:10.1093/bioinformatics/btt055.
 192. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447, doi:10.1021/ct700301q.
 193. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A.E.; Berendsen, H.J.C. GROMACS: Fast, Flexible, and Free. *J Comput Chem* **2005**, *26*, 1701–1718, doi:10.1002/jcc.20291.
 194. Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A Package for Molecular Simulation and Trajectory Analysis. *J Mol Model* **2001**, *7*, 306–317, doi:10.1007/s008940100045.
 195. Berendsen, H.J.C.; van der Spoel, D.; van Drunen, R. GROMACS: A Message-Passing Parallel Molecular Dynamics Implementation. *Computer Physics Communications* **1995**, *91*, 43–56, doi:10.1016/0010-4655(95)00042-E.
 196. Abraham, M.; Alekseenko, A.; Bergh, C.; Blau, C.; Briand, E.; Doijade, M.; Fleischmann, S.; Gapsys, V.; Garg, G.; Gorelov, S.; et al. GROMACS 2023.3 Source Code 2023.
 197. Valdés-Tresanco, M.S.; Valdés-Tresanco, M.E.; Valiente, P.A.; Moreno, E. gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *J Chem Theory Comput* **2021**, *17*, 6281–6291, doi:10.1021/acs.jctc.1c00645.

198. Miller, B.R.I.; McGee, T.D.Jr.; Swails, J.M.; Homeyer, N.; Gohlke, H.; Roitberg, A.E. MMPBSA.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314–3321, doi:10.1021/ct300418h.
199. GraphPad Prism Version 8.0.1. for Windows, GraphPad Software, Boston, Massachusetts USA, Www.Graphpad.Com.
200. Friedman, M. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *Journal of the American Statistical Association* **1937**, *32*, 675–701, doi:10.1080/01621459.1937.10503522.
201. Dunn, O.J. Multiple Comparisons Using Rank Sums. *Technometrics* **1964**, *6*, 241–252, doi:10.2307/1266041.
202. GraphPad Prism Version 10.4.1. for Windows, GraphPad Software, Boston, Massachusetts USA, Www.Graphpad.Com.
203. Liu, W.; Xie, Y.; Ma, J.; Luo, X.; Nie, P.; Zuo, Z.; Lahrmann, U.; Zhao, Q.; Zheng, Y.; Zhao, Y.; et al. IBS: An Illustrator for the Presentation and Visualization of Biological Sequences. *Bioinformatics* **2015**, *31*, 3359–3361, doi:10.1093/bioinformatics/btv362.
204. Bernard, E.; Nannya, Y.; Hasserjian, R.P.; Devlin, S.M.; Tuechler, H.; Medina-Martinez, J.S.; Yoshizato, T.; Shiozawa, Y.; Saiki, R.; Malcovati, L.; et al. Implications of TP53 Allelic State for Genome Stability, Clinical Presentation and Outcomes in Myelodysplastic Syndromes. *Nat Med* **2020**, *26*, 1549–1556, doi:10.1038/s41591-020-1008-z.
205. Schwartz, R.; Ting, C.S.; King, J. Whole Proteome pI Values Correlate with Subcellular Localizations of Proteins for Organisms within the Three Domains of Life. *Genome Res* **2001**, *11*, 703–709, doi:10.1101/gr-1587r.
206. Ishino, F.; Itoh, J.; Irie, M.; Matsuzawa, A.; Naruse, M.; Suzuki, T.; Hiraoka, Y.; Kaneko-Ishino, T. Retrovirus-Derived RTL9 Plays an Important Role in Innate Antifungal Immunity in the Eutherian Brain. *Int J Mol Sci* **2023**, *24*, 14884, doi:10.3390/ijms241914884.
207. Thol, F.; Platzbecker, U. Do Next-Generation Sequencing Results Drive Diagnostic and Therapeutic Decisions in MDS? *Blood Advances* **2019**, *3*, 3449–3453, doi:10.1182/bloodadvances.2018022434.
208. Bally, C.; Adès, L.; Renneville, A.; Sebert, M.; Eclache, V.; Preudhomme, C.; Mozziconacci, M.-J.; de The, H.; Lehmann-Che, J.; Fenaux, P. Prognostic Value of TP53 Gene Mutations in Myelodysplastic Syndromes and Acute Myeloid Leukemia Treated with Azacitidine. *Leukemia Research* **2014**, *38*, 751–755, doi:10.1016/j.leukres.2014.03.012.
209. Haase, D.; Stevenson, K.E.; Neubergh, D.; Maciejewski, J.P.; Nazha, A.; Sekeres, M.A.; Ebert, B.L.; Garcia-Manero, G.; Haferlach, C.; Haferlach, T.; et al. TP53 Mutation Status Divides Myelodysplastic Syndromes with Complex Karyotypes into Distinct Prognostic Subgroups. *Leukemia* **2019**, *33*, 1747–1758, doi:10.1038/s41375-018-0351-2.
210. Sallman, D.A.; Komrokji, R.; Vaupel, C.; Cluzeau, T.; Geyer, S.M.; McGraw, K.L.; Al Ali, N.H.; Lancet, J.; McGinniss, M.J.; Nahas, S.; et al. Impact of TP53 Mutation Variant Allele Frequency on Phenotype and Outcomes in Myelodysplastic Syndromes. *Leukemia* **2016**, *30*, 666–673, doi:10.1038/leu.2015.304.
211. Goel, S.; Hall, J.; Pradhan, K.; Hirsch, C.; Przychodzen, B.; Shastri, A.; Mantzaris, I.; Janakiram, M.; Battini, R.; Kornblum, N.; et al. High Prevalence and Allele Burden-Independent Prognostic Importance of P53 Mutations in an Inner-City MDS/AML Cohort. *Leukemia* **2016**, *30*, 1793–1795, doi:10.1038/leu.2016.74.
212. Cumbo, C.; Tota, G.; Anelli, L.; Zagaria, A.; Specchia, G.; Albano, F. TP53 in Myelodysplastic Syndromes: Recent Biological and Clinical Findings. *Int J Mol Sci* **2020**, *21*, 3432, doi:10.3390/ijms21103432.
213. Sabapathy, K.; Lane, D.P. Therapeutic Targeting of P53: All Mutants Are Equal, but Some Mutants Are More Equal than Others. *Nat Rev Clin Oncol* **2018**, *15*, 13–30, doi:10.1038/nrclinonc.2017.151.
214. Nakano, Y.; Kiyoi, H.; Miyawaki, S.; Asou, N.; Ohno, R.; Saito, H.; Naoe, T. Molecular Evolution of Acute Myeloid Leukaemia in Relapse: Unstable N-Ras and FLT3 Genes

- Compared with P53 Gene. *Br J Haematol* **1999**, *104*, 659–664, doi:10.1046/j.1365-2141.1999.01256.x.
215. Watanabe, A.; Inokuchi, K.; Yamaguchi, H.; Mizuki, T.; Tanosaki, S.; Shimada, T.; Dan, K. Near-Triploidy and near-Tetraploidy in Hematological Malignancies and Mutation of the P53 Gene. *Clinical & Laboratory Haematology* **2004**, *26*, 25–30, doi:10.1111/j.0141-9854.2003.00574.x.
 216. Xu, H.-T.; Lai, W.-L.; Liu, H.-F.; Wong, L.L.-Y.; Ng, I.O.-L.; Ching, Y.P. PAK4 Phosphorylates P53 at Serine 215 to Promote Liver Cancer Metastasis. *Cancer Research* **2016**, *76*, 5732–5742, doi:10.1158/0008-5472.CAN-15-3373.
 217. Kakudo, Y.; Shibata, H.; Otsuka, K.; Kato, S.; Ishioka, C. Lack of Correlation between P53-Dependent Transcriptional Activity and the Ability to Induce Apoptosis among 179 Mutant P53s. *Cancer Res* **2005**, *65*, 2108–2114, doi:10.1158/0008-5472.CAN-04-2935.
 218. Kawaguchi, T.; Kato, S.; Otsuka, K.; Watanabe, G.; Kumabe, T.; Tominaga, T.; Yoshimoto, T.; Ishioka, C. The Relationship among P53 Oligomer Formation, Structure and Transcriptional Activity Using a Comprehensive Missense Mutation Library. *Oncogene* **2005**, *24*, 6976–6981, doi:10.1038/sj.onc.1208839.
 219. Bushweller, J.H.; Schmidt, C.; Achille, N.; Kuntimaddi, A.; Boulton, A.; Leach, B.; Zhang, S.; Zeleznik-Le, N.J. Direct Binding of BCOR, but Not CBX8, to MLL-AF9 Is Essential for MLL-AF9 Leukemia Via Regulation of the EYA1/SIX1 Gene Network. *Blood* **2018**, *132*, 1316, doi:10.1182/blood-2018-09-111388.
 220. Yamano, H.; Tsurumi, C.; Gannon, J.; Hunt, T. The Role of the Destruction Box and Its Neighbouring Lysine Residues in Cyclin B for Anaphase Ubiquitin-Dependent Proteolysis in Fission Yeast: Defining the D-Box Receptor. *EMBO J* **1998**, *17*, 5670–5678, doi:10.1093/emboj/17.19.5670.
 221. Nguyen, T.B.; Manova, K.; Capodici, P.; Lindon, C.; Bottega, S.; Wang, X.-Y.; Refik-Rogers, J.; Pines, J.; Wolgemuth, D.J.; Koff, A. Characterization and Expression of Mammalian Cyclin B3, a Prepachytene Meiotic Cyclin *. *Journal of Biological Chemistry* **2002**, *277*, 41960–41969, doi:10.1074/jbc.M203951200.
 222. Peters, T.L.; Kumar, V.; Polikepahad, S.; Lin, F.Y.; Sarabia, S.F.; Liang, Y.; Wang, W.-L.; Lazar, A.J.; Doddapaneni, H.; Chao, H.; et al. BCOR-CCNB3 Fusions Are Frequent in Undifferentiated Sarcomas of Male Children. *Mod Pathol* **2015**, *28*, 575–586, doi:10.1038/modpathol.2014.139.
 223. Chiang, M.Y.; Xu, M.L.; Histén, G.; Shestova, O.; Roy, M.; Nam, Y.; Blacklow, S.C.; Sacks, D.B.; Pear, W.S.; Aster, J.C. Identification of a Conserved Negative Regulatory Sequence That Influences the Leukemogenic Activity of NOTCH1. *Mol Cell Biol* **2006**, *26*, 6261–6271, doi:10.1128/mcb.02478-05.
 224. Nam, Y.; Sliz, P.; Song, L.; Aster, J.C.; Blacklow, S.C. Structural Basis for Cooperativity in Recruitment of MAML Coactivators to Notch Transcription Complexes. *Cell* **2006**, *124*, 973–983, doi:10.1016/j.cell.2005.12.037.
 225. Wu, L.; Maillard, I.; Nakamura, M.; Pear, W.S.; Griffin, J.D. The Transcriptional Coactivator Maml1 Is Required for Notch2-Mediated Marginal Zone B-Cell Development. *Blood* **2007**, *110*, 3618–3623, doi:10.1182/blood-2007-06-097030.
 226. Liu, H.; Kennard, S.; Lilly, B. NOTCH3 Expression Is Induced in Mural Cells through an Autoregulatory Loop That Requires Endothelial-Expressed JAGGED1. *Circ Res* **2009**, *104*, 466–475, doi:10.1161/circresaha.108.184846.
 227. Oyama, T.; Harigaya, K.; Sasaki, N.; Okamura, Y.; Kokubo, H.; Saga, Y.; Hozumi, K.; Suganami, A.; Tamura, Y.; Nagase, T.; et al. Mastermind-like 1 (MamL1) and Mastermind-like 3 (MamL3) Are Essential for Notch Signaling in Vivo. *Development* **2011**, *138*, 5235–5246, doi:10.1242/dev.062802.
 228. Kitagawa, M. Notch Signalling in the Nucleus: Roles of Mastermind-like (MAML) Transcriptional Coactivators. *The Journal of Biochemistry* **2016**, *159*, 287–294, doi:10.1093/jb/mvv123.

229. Venkatesan, A.; Satin, L.S.; Raghavan, M. Roles of Calreticulin in Protein Folding, Immunity, Calcium Signaling and Cell Transformation. In *Cellular Biology of the Endoplasmic Reticulum*; Agellon, L.B., Michalak, M., Eds.; Springer International Publishing: Cham, 2021; pp. 145–162 ISBN 978-3-030-67696-4.
230. Ikawa, M.; Wada, I.; Kominami, K.; Watanabe, D.; Toshimori, K.; Nishimune, Y.; Okabe, M. The Putative Chaperone Calmegin Is Required for Sperm Fertility. *Nature* **1997**, *387*, 607–611, doi:10.1038/42484.
231. Treiber, T.; Treiber, N.; Plessmann, U.; Harlander, S.; Daiß, J.-L.; Eichner, N.; Lehmann, G.; Schall, K.; Urlaub, H.; Meister, G. A Compendium of RNA-Binding Proteins That Regulate MicroRNA Biogenesis. *Molecular Cell* **2017**, *66*, 270-284.e13, doi:10.1016/j.molcel.2017.03.014.
232. Carballo, E.; Lai, W.S.; Blackshear, P.J. Feedback Inhibition of Macrophage Tumor Necrosis Factor-Alpha Production by Tristetraprolin. *Science* **1998**, *281*, 1001–1005, doi:10.1126/science.281.5379.1001.
233. Lai, W.S.; Carballo, E.; Strum, J.R.; Kennington, E.A.; Phillips, R.S.; Blackshear, P.J. Evidence That Tristetraprolin Binds to AU-Rich Elements and Promotes the Deadenylation and Destabilization of Tumor Necrosis Factor Alpha mRNA. *Mol Cell Biol* **1999**, *19*, 4311–4323, doi:10.1128/mcb.19.6.4311.
234. Brown, R.S. Zinc Finger Proteins: Getting a Grip on RNA. *Curr Opin Struct Biol* **2005**, *15*, 94–98, doi:10.1016/j.sbi.2005.01.006.
235. Li, X.; Han, M.; Zhang, H.; Liu, F.; Pan, Y.; Zhu, J.; Liao, Z.; Chen, X.; Zhang, B. Structures and Biological Functions of Zinc Finger Proteins and Their Roles in Hepatocellular Carcinoma. *Biomarker Research* **2022**, *10*, 2, doi:10.1186/s40364-021-00345-1.
236. Lee, H.S.; Bang, I.; You, J.; Jeong, T.-K.; Kim, C.R.; Hwang, M.; Kim, J.-S.; Baek, S.H.; Song, J.-J.; Choi, H.-J. Molecular Basis for PHF7-Mediated Ubiquitination of Histone H3. *Genes Dev* **2023**, *37*, 984–997, doi:10.1101/gad.350989.123.
237. Cho, Y.-W.; Hong, T.; Hong, S.; Guo, H.; Yu, H.; Kim, D.; Guszczynski, T.; Dressler, G.R.; Copeland, T.D.; Kalkum, M.; et al. PTIP Associates with MLL3- and MLL4-Containing Histone H3 Lysine 4 Methyltransferase Complex*. *Journal of Biological Chemistry* **2007**, *282*, 20395–20406, doi:10.1074/jbc.M701574200.
238. Zhang, Z.-L.; Yu, P.-F.; Ling, Z.-Q. The Role of KMT2 Gene in Human Tumors. *Histol Histopathol* **2022**, *37*, 323–334, doi:10.14670/HH-18-447.
239. Linhoff, M.W.; Harton, J.A.; Cressman, D.E.; Martin, B.K.; Ting, J.P. Two Distinct Domains within CIITA Mediate Self-Association: Involvement of the GTP-Binding and Leucine-Rich Repeat Domains. *Mol Cell Biol* **2001**, *21*, 3001–3011, doi:10.1128/MCB.21.9.3001-3011.2001.
240. León Machado, J.A.; Steimle, V. The MHC Class II Transactivator CIITA: Not (Quite) the Odd-One-Out Anymore among NLR Proteins. *International Journal of Molecular Sciences* **2021**, *22*, 1074, doi:10.3390/ijms22031074.
241. Campillos, M.; Doerks, T.; Shah, P.K.; Bork, P. Computational Characterization of Multiple Gag-like Human Proteins. *Trends in Genetics* **2006**, *22*, 585–589, doi:10.1016/j.tig.2006.09.006.
242. Segel, M.; Lash, B.; Song, J.; Ladha, A.; Liu, C.C.; Jin, X.; Mekhedov, S.L.; Macrae, R.K.; Koonin, E.V.; Zhang, F. Mammalian Retrovirus-like Protein PEG10 Packages Its Own mRNA and Can Be Pseudotyped for mRNA Delivery. *Science* **2021**, *373*, 882–889, doi:10.1126/science.abg6155.
243. Bahman, F.; Choudhry, K.; Al-Rashed, F.; Al-Mulla, F.; Sindhu, S.; Ahmad, R. Aryl Hydrocarbon Receptor: Current Perspectives on Key Signaling Partners and Immunoregulatory Role in Inflammatory Diseases. *Front Immunol* **2024**, *15*, 1421346, doi:10.3389/fimmu.2024.1421346.
244. Matthews, B.W. Hydrophobic Interactions in Proteins. In *Encyclopedia of Life Sciences*; Wiley, 2001 ISBN 978-0-470-01617-6.

245. Pace, C.N.; Scholtz, J.M.; Grimsley, G.R. Forces Stabilizing Proteins. *FEBS Lett* **2014**, *588*, 2177–2184, doi:10.1016/j.febslet.2014.05.006.
246. Ferenczy, G.G.; Kellermayer, M. Contribution of Hydrophobic Interactions to Protein Mechanical Stability. *Computational and Structural Biotechnology Journal* **2022**, *20*, 1946–1956, doi:10.1016/j.csbj.2022.04.025.
247. Li, L.; Zhang, M.; Chen, S.; Sun, X.; Xu, H.; Li, L.; Zhang, T.; Huang, X.; Ye, H.; Ding, Y. Detection of BCOR Gene Rearrangement in Ewing-like Sarcoma: An Important Diagnostic Tool. *Diagnostic Pathology* **2021**, *16*, 50, doi:10.1186/s13000-021-01114-2.
248. Buchberger, E.; Harchi, M.E.; Payrhuber, D.; Zommer, A.; Schauer, D.; Simonitsch-Klupp, I.; Bilban, M.; Brostjan, C. Overexpression of the Transcriptional Repressor Complex BCL-6/BCoR Leads to Nuclear Aggregates Distinct from Classical Aggresomes. *PLOS ONE* **2013**, *8*, e76845, doi:10.1371/journal.pone.0076845.
249. Hanahan, D.; Weinberg, R.A. Hallmarks of Cancer: The Next Generation. *Cell* **2011**, *144*, 646–674, doi:10.1016/j.cell.2011.02.013.
250. Wang, S.; Kaufman, R.J. The Impact of the Unfolded Protein Response on Human Disease. *J Cell Biol* **2012**, *197*, 857–867, doi:10.1083/jcb.201110131.
251. Scopes, R.K. *Protein Purification: Principles and Practice*; Springer Science & Business Media, 1993; ISBN 978-0-387-94072-4.
252. Burley, S.K.; Petsko, G.A. Aromatic-Aromatic Interaction: A Mechanism of Protein Structure Stabilization. *Science* **1985**, *229*, 23–28, doi:10.1126/science.3892686.
253. Bjellqvist, B.; Basse, B.; Olsen, E.; Celis, J.E. Reference Points for Comparisons of Two-Dimensional Maps of Proteins from Different Human Cell Types Defined in a pH Scale Where Isoelectric Points Correlate with Polypeptide Compositions. *ELECTROPHORESIS* **1994**, *15*, 529–539, doi:10.1002/elps.1150150171.
254. Righetti, P.G.; Caravaggio, T. Isoelectric Points and Molecular Weights of Proteins: A Table. *Journal of Chromatography A* **1976**, *127*, 1–28, doi:10.1016/S0021-9673(00)98537-6.
255. Dill, K.A. Dominant Forces in Protein Folding. *Biochemistry* **1990**, *29*, 7133–7155, doi:10.1021/bi00483a001.
256. Haidar, R.; Henkler, F.; Kugler, J.; Rosin, A.; Genkinger, D.; Laux, P.; Luch, A. The Role of DNA-Binding and ARNT Dimerization on the Nucleo-Cytoplasmic Translocation of the Aryl Hydrocarbon Receptor. *Sci Rep* **2021**, *11*, 18194, doi:10.1038/s41598-021-97507-w.
257. Watanabe, D.; Yamada, K.; Nishina, Y.; Tajima, Y.; Koshimizu, U.; Nagata, A.; Nishimune, Y. Molecular Cloning of a Novel Ca(2+)-Binding Protein (Calmegin) Specifically Expressed during Male Meiotic Germ Cell Development. *Journal of Biological Chemistry* **1994**, *269*, 7744–7749, doi:10.1016/S0021-9258(17)37349-0.
258. Ikawa, M.; Nakanishi, T.; Yamada, S.; Wada, I.; Kominami, K.; Tanaka, H.; Nozaki, M.; Nishimune, Y.; Okabe, M. Calmegin Is Required for Fertilin α/β Heterodimerization and Sperm Fertility. *Developmental Biology* **2001**, *240*, 254–261, doi:10.1006/dbio.2001.0462.
259. Francis, N.J.; Kingston, R.E.; Woodcock, C.L. Chromatin Compaction by a Polycomb Group Protein Complex. *Science* **2004**, *306*, 1574–1577, doi:10.1126/science.1100576.
260. Cheutin, T.; McNairn, A.J.; Jenuwein, T.; Gilbert, D.M.; Singh, P.B.; Misteli, T. Maintenance of Stable Heterochromatin Domains by Dynamic HP1 Binding. *Science* **2003**, *299*, 721–725, doi:10.1126/science.1078572.
261. Gao, Z.; Zhang, J.; Bonasio, R.; Strino, F.; Sawai, A.; Parisi, F.; Kluger, Y.; Reinberg, D. PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. *Mol Cell* **2012**, *45*, 344–356, doi:10.1016/j.molcel.2012.01.002.
262. Kadoch, C.; Crabtree, G.R. Reversible Disruption of mSWI/SNF (BAF) Complexes by the SS18-SSX Oncogenic Fusion in Synovial Sarcoma. *Cell* **2013**, *153*, 71–85, doi:10.1016/j.cell.2013.02.036.
263. Mitelman, F.; Johansson, B.; Mertens, F. The Impact of Translocations and Gene Fusions on Cancer Causation. *Nat Rev Cancer* **2007**, *7*, 233–245, doi:10.1038/nrc2091.

264. Mertens, F.; Johansson, B.; Fioretos, T.; Mitelman, F. The Emerging Complexity of Gene Fusions in Cancer. *Nat Rev Cancer* **2015**, *15*, 371–381, doi:10.1038/nrc3947.
265. Stransky, N.; Cerami, E.; Schalm, S.; Kim, J.L.; Lengauer, C. The Landscape of Kinase Fusions in Cancer. *Nat Commun* **2014**, *5*, 4846, doi:10.1038/ncomms5846.
266. Cao, Q.; Gearhart, M.D.; Gery, S.; Shojaee, S.; Yang, H.; Sun, H.; Lin, D.-C.; Bai, J.-W.; Mead, M.; Zhao, Z.; et al. BCOR Regulates Myeloid Cell Proliferation and Differentiation. *Leukemia* **2016**, *30*, 1155–1165, doi:10.1038/leu.2016.2.
267. Gearhart, M.D.; Corcoran, C.M.; Wamstad, J.A.; Bardwell, V.J. Polycomb Group and SCF Ubiquitin Ligases Are Found in a Novel BCOR Complex That Is Recruited to BCL6 Targets. *Mol Cell Biol* **2006**, *26*, 6880–6889, doi:10.1128/MCB.00630-06.
268. Perez, M.; Muñoz-Galván, S.; Jiménez-García, M.P.; Marín, J.J.; Carnero, A. Efficacy of CDK4 Inhibition against Sarcomas Depends on Their Levels of CDK4 and P16ink4 mRNA. *Oncotarget* **2015**, *6*, 40557–40574.
269. Martin-Broto, J.; Martinez-Garcia, J.; Moura, D.S.; Redondo, A.; Gutierrez, A.; Lopez-Pousa, A.; Martinez-Trufero, J.; Sevilla, I.; Diaz-Beveridge, R.; Solis-Hernandez, M.P.; et al. Phase II Trial of CDK4/6 Inhibitor Palbociclib in Advanced Sarcoma Based on mRNA Expression of CDK4/CDKN2A. *Sig Transduct Target Ther* **2023**, *8*, 1–10, doi:10.1038/s41392-023-01661-8.
270. Merjaneh, N.; Kim, H.; Escoto, H.; Metts, J.; Ray, A.; Bukowinski, A.; LeBlanc, Z.; Fair, D.; Watanbe, M.; Alva, E.; et al. Strategies for the Treatment of Infantile Soft Tissue Sarcomas With BCOR Alterations. *J Pediatr Hematol Oncol* **2023**, *45*, 315–321, doi:10.1097/MPH.0000000000002620.
271. Kim, H.; George, E.; Ragland, R.; Rafial, S.; Zhang, R.; Krepler, C.; Morgan, M.; Herlyn, M.; Brown, E.; Simpkins, F. Targeting the ATR/CHK1 Axis with PARP Inhibition Results in Tumor Regression in BRCA Mutant Models. *Clin Cancer Res* **2017**, *23*, 3097–3108, doi:10.1158/1078-0432.CCR-16-2273.
272. Planas-Paz, L.; Pliego-Mendieta, A.; Hagedorn, C.; Aguilera-Garcia, D.; Haberecker, M.; Arnold, F.; Herzog, M.; Bankel, L.; Guggenberger, R.; Steiner, S.; et al. Unravelling Homologous Recombination Repair Deficiency and Therapeutic Opportunities in Soft Tissue and Bone Sarcoma. *EMBO Mol Med* **2023**, *15*, e16863, doi:10.15252/emmm.202216863.
273. Fursova, N.A.; Blackledge, N.P.; Nakayama, M.; Ito, S.; Koseki, Y.; Farcas, A.M.; King, H.W.; Koseki, H.; Klose, R.J. Synergy between Variant PRC1 Complexes Defines Polycomb-Mediated Gene Repression. *Mol Cell* **2019**, *74*, 1020-1036.e8, doi:10.1016/j.molcel.2019.03.024.
274. Békés, M.; Langley, D.R.; Crews, C.M. PROTAC Targeted Protein Degraders: The Past Is Prologue. *Nat Rev Drug Discov* **2022**, *21*, 181–200, doi:10.1038/s41573-021-00371-6.
275. Lloyd, R.L.; Wijnhoven, P.W.G.; Ramos-Montoya, A.; Wilson, Z.; Illuzzi, G.; Falenta, K.; Jones, G.N.; James, N.; Chabbert, C.D.; Stott, J.; et al. Combined PARP and ATR Inhibition Potentiates Genome Instability and Cell Death in ATM-Deficient Cancer Cells. *Oncogene* **2020**, *39*, 4869–4883, doi:10.1038/s41388-020-1328-y.
276. Wei, F.; Cheng, X.-X.; Xue, J.Z.; Xue, S.-A. Emerging Strategies in TCR-Engineered T Cells. *Front. Immunol.* **2022**, *13*, doi:10.3389/fimmu.2022.850358.
277. Shafer, P.; Kelly, L.M.; Hoyos, V. Cancer Therapy With TCR-Engineered T Cells: Current Strategies, Challenges, and Prospects. *Front Immunol* **2022**, *13*, 835762, doi:10.3389/fimmu.2022.835762.
278. Zhang, Y.; Liu, Z.; Wei, W.; Li, Y. TCR Engineered T Cells for Solid Tumor Immunotherapy. *Exp Hematol Oncol* **2022**, *11*, 38, doi:10.1186/s40164-022-00291-0.
279. Baulu, E.; Gardet, C.; Chuvin, N.; Depil, S. TCR-Engineered T Cell Therapy in Solid Tumors: State of the Art and Perspectives. *Sci Adv* **9**, eadf3700, doi:10.1126/sciadv.adf3700.

280. Mahadevia, H.; Ponvilawan, B.; Al-Obaidi, A.; Buckley, J.; Subramanian, J.; Bansal, D. Exceptional Synergistic Response of PARP Inhibitor and Immune Checkpoint Inhibitor in Esophageal Adenocarcinoma with a Germline BRCA2 Mutation: A Case Report. *Ther Adv Med Oncol* **2024**, *16*, 17588359241242406, doi:10.1177/17588359241242406.
281. Whittle, S.B.; Fetzko, S.; Roy, A.; Venkatramani, R. Soft Tissue and Visceral Organ Sarcomas With BCOR Alterations. *J Pediatr Hematol Oncol* **2022**, *44*, 195–200, doi:10.1097/MPH.0000000000002480.
282. Yap, T.A.; Krebs, M.G.; Postel-Vinay, S.; El-Khouiery, A.; Soria, J.-C.; Lopez, J.; Berges, A.; Cheung, S.Y.A.; Irurzun-Arana, I.; Goldwin, A.; et al. Ceralasertib (AZD6738), an Oral ATR Kinase Inhibitor, in Combination with Carboplatin in Patients with Advanced Solid Tumors: A Phase I Study. *Clin Cancer Res* **2021**, *27*, 5213–5224, doi:10.1158/1078-0432.CCR-21-1032.
283. Mateo, J.; Lord, C.J.; Serra, V.; Tutt, A.; Balmaña, J.; Castroviejo-Bermejo, M.; Cruz, C.; Oaknin, A.; Kaye, S.B.; de Bono, J.S. A Decade of Clinical Development of PARP Inhibitors in Perspective. *Ann Oncol* **2019**, *30*, 1437–1447, doi:10.1093/annonc/mdz192.
284. D'Angelo, S.P.; Mahoney, M.R.; Van Tine, B.A.; Atkins, J.; Milhem, M.M.; Jahagirdar, B.N.; Antonescu, C.R.; Horvath, E.; Tap, W.D.; Schwartz, G.K.; et al. A Non-Comparative Multi-Center Randomized Phase II Study of Nivolumab +/- Ipilimumab for Patients with Metastatic Sarcoma (Alliance A091401). *Lancet Oncol* **2018**, *19*, 416–426, doi:10.1016/S1470-2045(18)30006-8.
285. Choy, E.; Butrynski, J.E.; Harmon, D.C.; Morgan, J.A.; George, S.; Wagner, A.J.; D'Adamo, D.; Cote, G.M.; Flamand, Y.; Benes, C.H.; et al. Phase II Study of Olaparib in Patients with Refractory Ewing Sarcoma Following Failure of Standard Chemotherapy. *BMC Cancer* **2014**, *14*, 813, doi:10.1186/1471-2407-14-813.
286. Sleijfer, S.; Ray-Coquard, I.; Papai, Z.; Le Cesne, A.; Scurr, M.; Schöffski, P.; Collin, F.; Pandite, L.; Marreaud, S.; De Brauwier, A.; et al. Pazopanib, a Multikinase Angiogenesis Inhibitor, in Patients with Relapsed or Refractory Advanced Soft Tissue Sarcoma: A Phase II Study from the European Organisation for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group (EORTC Study 62043). *J Clin Oncol* **2009**, *27*, 3126–3132, doi:10.1200/JCO.2008.21.3223.

11. Publikációs lista (A Kenézy Élettudományi Könyvtár jóváhagyása által)



**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/164/2025.PL
Tárgy: PhD Publikációs Lista

Jelölt: Madarász Kristóf
Doktori Iskola: Klinikai Orvostudományok Doktori Iskola
MTMT azonosító: 10088102

A PhD értekezés alapjául szolgáló közlemények

1. **Madarász, K.**, Mótyán, J. A., Chang Chien, Y. C., Bedekovics, J., Csoma, S. L., Méhes, G., Mokánszki, A.: BCOR-rearranged sarcomas: In silico insights into altered domains and BCOR interactions.
Comput. Biol. Med. 191, 1-22, 2025.
DOI: <http://dx.doi.org/10.1016/j.compbiomed.2025.110144>
IF: 7 (2023)
2. **Madarász, K.**, Mótyán, J. A., Bedekovics, J., Miltényi, Z., Ujfalusi, A., Méhes, G., Mokánszki, A.: Deep Molecular and In Silico Protein Analysis of p53 Alteration in Myelodysplastic Neoplasia and Acute Myeloid Leukemia.
Cells. 11 (21), 1-23, 2022.
DOI: <http://dx.doi.org/10.3390/cells11213475>
IF: 6

További közlemények

3. Bedekovics, J., **Madarász, K.**, Mokánszki, A., Deliné Molnár, S., Mester, Á., Miltényi, Z., Méhes, G.: Exploring p53 protein expression and its link to TP53 mutation in myelodysplasia-related malignancies - Interpretive challenges and potential field of applications.
Histopathology. 85 (1), 143-154, 2024.
DOI: <http://dx.doi.org/10.1111/his.15185>
IF: 3.9 (2023)
4. Csoma, S. L., **Madarász, K.**, Chang Chien, Y. C., Emri, G., Bedekovics, J., Méhes, G., Mokánszki, A.: Correlation Analyses between Histological Staging and Molecular Alterations in Tumor-Derived and Cell-Free DNA of Early-Stage Primary Cutaneous Melanoma.
Cancers (Basel). 15 (21), 1-13, 2023.
DOI: <http://dx.doi.org/10.3390/cancers15215141>
IF: 4.5





5. Chang Chien, Y. C., **Madarász, K.**, Csoma, S. L., Mótyán, J. A., Huang, H. Y., Méhes, G.,
Mokánszki, A.: Molecular Identification and In Silico Protein Analysis of a Novel BCOR-CLGN
Gene Fusion in Intrathoracic BCOR-Rearranged Sarcoma.
Cancers (Basel). 15 (3), 1-17, 2023.
DOI: <http://dx.doi.org/10.3390/cancers15030898>
IF: 4.5

A közlő folyóiratok összesített impakt faktora: 25,9

**A közlő folyóiratok összesített impakt faktora (az értekezés alapjául szolgáló közleményekre):
13**

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

Debrecen, 2025.04.25.





Registry number: DEENK/164/2025.PL
Subject: PhD Publication List

Candidate: Kristóf Madarász
Doctoral School: Doctoral School of Clinical Medicine
MTMT ID: 10088102

List of publications related to the dissertation

1. **Madarász, K.**, Mótyán, J. A., Chang Chien, Y. C., Bedekovics, J., Csoma, S. L., Méhes, G., Mokánszki, A.: BCOR-rearranged sarcomas: In silico insights into altered domains and BCOR interactions.
Comput. Biol. Med. 191, 1-22, 2025.
DOI: <http://dx.doi.org/10.1016/j.combiomed.2025.110144>
IF: 7 (2023)
2. **Madarász, K.**, Mótyán, J. A., Bedekovics, J., Miltényi, Z., Ujfalusi, A., Méhes, G., Mokánszki, A.: Deep Molecular and In Silico Protein Analysis of p53 Alteration in Myelodysplastic Neoplasia and Acute Myeloid Leukemia.
Cells. 11 (21), 1-23, 2022.
DOI: <http://dx.doi.org/10.3390/cells11213475>
IF: 6

List of other publications

3. Bedekovics, J., **Madarász, K.**, Mokánszki, A., Deliné Molnár, S., Mester, Á., Miltényi, Z., Méhes, G.: Exploring p53 protein expression and its link to TP53 mutation in myelodysplasia-related malignancies - Interpretive challenges and potential field of applications.
Histopathology. 85 (1), 143-154, 2024.
DOI: <http://dx.doi.org/10.1111/his.15185>
IF: 3.9 (2023)
4. Csoma, S. L., **Madarász, K.**, Chang Chien, Y. C., Emri, G., Bedekovics, J., Méhes, G., Mokánszki, A.: Correlation Analyses between Histological Staging and Molecular Alterations in Tumor-Derived and Cell-Free DNA of Early-Stage Primary Cutaneous Melanoma.
Cancers (Basel). 15 (21), 1-13, 2023.
DOI: <http://dx.doi.org/10.3390/cancers15215141>
IF: 4.5





5. Chang Chien, Y. C., **Madarász, K.**, Csoma, S. L., Mótyán, J. A., Huang, H. Y., Méhes, G., Mokánszki, A.: Molecular Identification and In Silico Protein Analysis of a Novel BCOR-CLGN Gene Fusion in Intrathoracic BCOR-Rearranged Sarcoma. *Cancers (Basel)*. 15 (3), 1-17, 2023.
DOI: <http://dx.doi.org/10.3390/cancers15030898>
IF: 4.5

Total IF of journals (all publications): 25,9

Total IF of journals (publications related to the dissertation): 13

The Candidate's publication data submitted to the Tudóstér have been validated by DEENK on the basis of the Journal Citation Report (Impact Factor) database.

25 April, 2025



12.Tárgyszavak

Onkopatológia, génvariáns, génfúzió, precíziós onkológia, *TP53* mutáció, *BCOR* génfúzió, mieloid neoplázia, bioinformatika, fehérje modellezés

13. Köszönetnyilvánítás

Ezúton szeretnék köszönetet mondani témavezetőmnek, Dr. Mokánszki Attilának a Debreceni Egyetem, Általános Orvostudományi Kar, Patológia Intézet egyetemi adjunktusának a támogatásért és türelméért, amelyet a kutatásom során tanúsított. Az ő szakmai tanácsai és tudományos precizitása nélkül ez a munka nem jöhetett volna létre.

Hálás köszönettel tartozom a Patológiai Intézet igazgatójának Prof. Dr. Méhes Gábornak, aki lehetőséget biztosított kutatásaim elvégzésére és támogatott az elmúlt években. Hálás köszönet illeti a Patológiai Intézet Molekuláris Laboratóriumának munkatársait, akik nemcsak bevezettek a rutin diagnosztika világába, hanem barátságukkal és támogatásukkal jelentősen hozzájárultak kutatási munkám gazdagításához. Külön szeretném megköszönni Mónus Anikónak és Csoma Szilviának a hosszú évek során nyújtott önzetlen segítségüket és szakmai támogatásukat.

Köszönettel tartozom Dr. Mótyán János Andrásnak, a Biokémiai és Molekuláris Biológiai Intézet egyetemi docensének a munkám során nyújtott értékes segítségért, valamint azért, hogy bevezetett az *in silico* módszerek alkalmazásának világába, ezzel is jelentősen hozzájárulva szakmai fejlődésemhez.

Kutatásunkat az alábbi pályázatok támogatták: ÚNKP-23-3-II Felsőoktatási Doktori Hallgatói Kutatói Ösztöndíj, GTIDEA Kiválósági PhD Ösztöndíj, Debreceni Egyetem Doktorandusz Kiegészítő Ösztöndíj, EKÖP-24-3-II Doktori Hallgatói Ösztöndíj II.