

Doktori (PhD) értekezés tézisei

**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

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

Értekezés a doktori (PhD) fokozat megszerzése érdekében a
klinikai orvostudományok tudományágában

Írta: Madarász Kristóf, okleveles biológus

Készült a Debreceni Egyetem Klinikai Orvostudományok
Doktori Iskolája (Konzervatív orvostudományok és klinikai
vizsgálatok programja) keretében

Témavezető: Dr. Mokánszki Attila, PhD

Az értekezés bírálói:

Prof. Dr. Csósz Éva, az MTA doktora

Prof. Dr. Vásárhelyi Barna, az MTA doktora

A bírálóbizottság:

elnök: Prof. Dr. Németh Norbert, az MTA doktora

tagok: Prof. Dr. Csósz Éva, az MTA doktora

Prof. Dr. Vásárhelyi Barna, az MTA doktora

Prof. Dr. Balogh István, az MTA doktora

Dr. Czajlik András, PhD

Az értekezés védésének helyszíne és időpontja:

Debreceni Egyetem ÁOK, 2025. október 07. 13:00 óra,

DE ÁOK Belgyógyászati Intézet A épület tanterme.

1. 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. Az MDS egy heterogén betegségsorozat, amelyet a vörsejtek elégtelen termelése és a perifériás vér citopéniája határoz meg, míg az AML egy gyorsan progrediáló, agresszív megbetegedés, amelyet a leukémiás blasztok kontrollálatlan burjánzása dominál. A két kórkép közötti szoros kapcsolatot igazolja, hogy az MDS esetek jelentős része AML-be progrediál.

Globálisan mindkét betegség incidenciája növekszik az életkorral, különösen a 70 év feletti populációban. Genetikai szempontból a patogenezisben kiemelkedő szerepet játszanak a DNS metilációját, a kromatin struktúra átalakulását, valamint az RNS-splicingot szabályozó gének mutációi, amelyek elősegítik a klonális evolúciót és a terápiás rezisztenciát. A komplex kariotípussal és TP53 mutációkkal jellemezhető alcsoportok prognózisa különösen kedvezőtlen, a medián túlélési idő intenzív kezelés mellett is rövid, főként biállélikus mutációk vagy magas variáns allél frekvencia esetén.

A modern prognosztikai modellek, mint a Molekuláris Nemzetközi Prognosztikai Pontozási Rendszer (IPSS-M), integrálják a klinikai paramétereket, a citogenetikai rendellenességeket és a szomatikus mutációkat, így pontosítva a betegek kockázati besorolását. Ez a megközelítés hangsúlyozza a genetikai profil meghatározó szerepét a kezelési döntésekben.

Az onkológiai kutatások egyik legnagyobb kihívása az ismeretlen klinikai jelentőségű variánsok (VUS) értelmezése. Az újgenerációs

szekvenálási (NGS) technológiák elterjedésével egyre több genetikai eltérést azonosítanak, amelyek funkcionális és klinikai hatása gyakran tisztázatlan. Ezen variánsok pontos értékelése alapvető feltétele a hatékony célzott terápiák alkalmazásának és a precíziós onkológia sikerességének, amely integrált megközelítést, többek között bioinformatikai predikciós eszközök alkalmazását igényli.

A génfüziók szintén kiemelt szerepet játszanak számos daganat, köztük a szarkómák patogenezisében. Ezek az eltérések onkogén fehérjéket hoznak létre, amelyek célzott terápiák hatékony célpontjai lehetnek. Az NGS és a bioinformatikai eszközök fejlődése lehetővé tette ezen füziók pontosabb detektálását, azonban klinikai jelentőségük és funkcionális következményeik értelmezése továbbra is kihívást jelent. A multiomikai adatok integrált elemzése elengedhetetlen ezen füziók által aktivált jelátviteli útvonalak feltérképezéséhez és új terápiás célpontok azonosításához.

2. 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 génfúziók (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 génfúziók 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* szekvencia- 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 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.

3. Anyagok és módszerek

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

3.1.1 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 mielodiszplázias szindróma (MDS-IB; 12 eset), illetve alacsony blasztarányú mielodiszplázias szindróma (MDS-LB; 39 eset) kategóriákba soroltak át a Debreceni Egyetem, Klinikai Központ, Patológiai Intézetében. A hematoxinil-eozinnal (H&E) festett metszeteket patológus kategorizálta. A citogenetikai elemzést a rutindiagnosztikai eljárás részeként végezték el. A vizsgálatokhoz etikai engedélyekkel rendelkeztünk (60355-2/2016/EKU és IV/8465-3/2021/EKU). A tanulmányok a Helsinki Nyilatkozat irányelveinek megfelelően zajlottak.

3.1.2 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.

3.1.3 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 μ L 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).

3.1.4 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 összeszerveve) 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.

3.1.5 *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 3D szerkezeteit a Robetta szoftver segítségével építettük fel. A mutációk által érintett TP53 fehérje poszttranszlációs módosítási (PTM) helyeit a PhosphoSite Plus weboldal segítségével azonosítottuk. A szekvenciák strukturális rendezetlenségi meghatározását az IUPred3 webszerverrel végeztük. A másodlagos szerkezet becslésére a GORIV alkalmazást használtuk. A variánsok szekvenciaalapú (Seq) stabilitásváltozásait az I-Mutant2.0 és a DDGun szerverekkel, alapértelmezett paraméterekkel állapítottuk meg. A szerkezetalapú (struk) stabilitási tesztekhez az I-Mutant2.0, a DynaMut2 és a DDGun3D alkalmazásokat használtuk. A predikciókat számos TP53 kristály-szerkezet (Cry) (5MCT, 5MG7, 5MF7, 1AIE, 1C26, 2FOO, 1YC5) és az AlphaFold (AF) 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 és p.R248Q mutációk esetében hasonlítottuk össze, mivel kísérleti adatok csak ezekben az esetekben álltak rendelkezésre.

A fehérje-fehérje kölcsönhatások (PPI) vizsgálatához az mCSM-PPI2 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 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 és 1SAL. A mutáns TP53 fehérjék DNS-hez való affinitásának változásait az mCSM-NA segítségével számítottuk ki.

A variáns fehérjék patogénitásának meghatározására a National Cancer Institute International Agency for Research on Cancer (IARC) TP53 adatbázisát (R20, 2019. július) , és annak komponenseit (Align-GVGD, BayesDel, REVEL, Sift class, PolyPhen2, Transactivation, TransactivationClass, DNE_LOFclass , DNE class , Structure/Function class), valamint a ClinVar, Varsity, Phd-SNP⁸ és FATHMM-XF patogénitást meghatározó pontrendszereket használtuk.

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

3.2.1 Fehérjeinformációk

A szakirodalomban eddig ismert *BCOR* gént érintő kilenc fúzió nukleotid szekvenciáit vizsgáltuk. Az Ensemble adatbázist (GRCh37.p13; GCA_000001405.14) használtuk a fúziók töréspontjainak lokalizálására és a szekvenciák manuális összeállítására. A kódoló aminosav szekvenciákat *in silico* módon határoztuk meg az ExPASy: Translate segítségével. 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 (a gének és fehérjék fúzióját a legújabb nomenklatúra szerint dupla kettősponttal (::) jelöltük). A doméninformációkat az UniProt , a Conserved Domain Database (CDD) és az InterPro (verzió 5.67-99.0) adatbázisokból nyertük. A Gén Ontológiai kifejezéseket (GO) az

összeállított aminosavszekvenciák alapján határoztuk meg az InterPro segítségével, specifikusan a PANTHER GO kifejezéseket alkalmaztuk.

3.2.2 Fizikai-kémiai tulajdonságok számítása

A fizikai-kémiai paramétereket az ExPASy ProtParam eszközzel számoltuk, beleértve az elméleti izoelektromos pontot (TpI), a molekulatömeget, a pozitív és negatív töltésű aminosavak összesített számát, az extinkciós együtthatót (EC), az instabilitási indexet (II), az alifás indexet (AI), valamint a hidropátia átlagértékét (GRAVY). A vad típusú fehérjék és a fúziós fehérjék közötti fizikai-kémiai értékek változásának kiszámításához a következő képletet alkalmaztuk:

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

3.2.3 Szignálpeptidek és rendezetlen fehérje szakaszok predikciója

A SignalP 6.0 programot alkalmaztuk a fehérjékben található szignálpeptidek azonosítására. Csak az eukariótákban előforduló Sec/jelzőpeptideket (Sec/SPI) kerestük, mivel a bemeneti organizmusként az „Eukarióták” opciót választottuk. 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). Az előrejelzések érvényességének ellenőrzéséhez az eredményeket összevetettük az UniProt adatbázis és a The Human Protein Atlas adataival. Emellett a NetGPI-1.1 GPI-anchored prediktort

alkalmaztuk a fehérjék glikozilfoszfatidilinozitol (GPI) horgonyzottság vizsgálatára.

A fehérjékben található belsőleg rendezetlen régiók (IDR) előrejelzéséhez az IUPred3-at 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 fehérjéinek 1448-1633 (region^{ANK+linker}) és 1634-1748 aminosavig terjedő (region^{PUPD}) szakaszain.

3.2.4 Fehérjeszerkezetek elemzése

A BCOR-PCGF1 és PRC1.1 komplexekben részt vevő fehérjék háromdimenziós (3D) modelljeit az AlphaFold3 (AF3) segítségével generáltuk. 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 és a vad típusú fehérjék modelljeivel. Az AlphaFold Server (béta) szekvencia hosszra vonatkozó limitációi miatt az 5480 aminosavból álló KMT2D::BCOR szerkezetét nem tudtuk elkészíteni.

A PRODIGY (PROtein binDing enerGY prediction) web szervert 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 PUFD régió (1634-1748) és a PCGF1 RAWUL régió (167-255; RAWUL-PUFD doménhossz) mentén vizsgáltuk. A RAWUL domén 167-177 és 185-254, valamint a PUFD domén 1636-1748 (4HPL hossz) tartományait elemeztük. 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. 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 PUFD doménnel, a konzisztencia érdekében eltávolítottuk a modellezett komplexekből PyMOL segítségével, ahol a 4hpl kontrollként szerepelt. További kontrollként BCOR-PUFD 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. A PDB 4hpl szerkezetet az RCSB Protein Data Bank-ból töltöttük le.

3.2.5 Molekuláris dinamikai szimulációk

A dimer szerkezetek vizsgálatára 10 ns hosszúságú molekuláris dinamikai (MD) szimulációkat végeztünk a GROMACS szoftvercsomag segítségével. A kilenc BCOR-PCGF1 dimer (egy vad típusú és nyolc génfüzió) az AlphaFold által generált 50 replikátumból 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.

3.3 Statisztikai elemzés

A statisztikai elemzéseket a GraphPad Prism 8.0.1 és 9.5.1 szoftverrel végeztük (GraphPad Software, San Diego, CA, USA). 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 tesztet vagy az ANOVA robusztusabb változatait (pl. Brown-Forsythe ANOVA) választottuk.

Annak vizsgálatára, hogy van-e különbség a TP53 fehérje patogenitására vonatkozó patogenitási 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. A választás minden esetben a GraphPad Prism által végzett diagnosztikai tesztek eredményeire 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 és az IBS (Illustrator of Biological Sequences) szoftvert használtuk.

4. Eredmények

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

4.1.1 Klinikopatológiai jellemzők

A vizsgálatba bevont 77 beteg átlagéletkora 64,1 év volt, a férfi-nő arány 42:35 volt. A betegek a diagnózis alapján három csoportba sorolódtak: 26 beteg akut mieloid leukémia mielodiszpláziával összefüggő elváltozásokkal (AML-MR), 12 beteg emelkedett blasztarányú mielodiszpláziás szindróma (MDS-IB), és 39 beteg alacsony blasztarányú mielodiszpláziás szindróma (MDS-LB) kategóriába

4.1.2 Újgenerációs szekvenálás

A 77 vizsgált esetből 26-ban (33,8%) azonosítottunk legalább egy TP53 mutációt, összesen 41 különálló mutációt detektálva, amelyek 30 különböző genotípust alkottak. A mutációk előfordulása az AML-MR csoportban volt a legmagasabb (57,7%), ezt követte az MDS-IB (33,3%) és az MDS-LB (17,9%) csoport. Hét esetben egy mintán belül több mutáció is jelen volt. A variáns allélfrekvencia (VAF) átlagértéke a legmagasabb az MDS-IB és AML-MR csoportokban volt. A 30 egyedi mutáció túlnyomó többsége (83,3%) a fehérje DNS-kötő doménjében (DBD) helyezkedett el. A mutációk többsége (77%) misszensz variáns volt, de azonosítottunk frameshift és stop kodont eredményező elváltozásokat is, amelyek közül több a fehérje jelentős rövidülését, úgynevezett trunkációját okozta

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

A TP53 mutációt hordozó betegek jelentős részénél (42,3%) komplex kariotípus volt kimutatható, különösen az AML-MR csoportban (60%). Az immunhisztokémiai (IHC) pozitivitás és az NGS-sel kimutatott mutációk között részleges volt az átfedés; az NGS-sel igazolt mutáns esetek 42%-a mutatott IHC pozitivitást. A hat, fehérje-trunkációt okozó mutációból öt esetben az IHC-vizsgálat negatív eredménnyel zárult. A túlélési analízis szignifikáns különbséget mutatott a TP53 mutáns és vad típusú betegek között. A vad típusú esetek medián túlélése 918 nap volt, míg a mutációt hordozóké csupán 224 nap. A prognózist tovább rontotta a magas, 23%-ot meghaladó variáns allélfrekvencia

4.1.4 A mutációk funkcionális hatásainak *in silico* predikciója

Az azonosított misszensz mutációk jelentős részét a ClinVar adatbázis bizonytalan klinikai jelentőségűnek vagy ellentmondásosnak minősítette, ami indokolta a további, részletes bioinformatikai elemzést. Számos, különböző elven működő patogenitást és fehérjestabilitást előrejelző program segítségével értékeltük a mutációk potenciális kártékony hatását. Az eredmények alapján a legtöbb vizsgált mutáció valószínűsíthetően patogén funkcionális következményekkel járt. Az IARC TP53 adatbázisában elérhető *in vitro* kísérletes adatok megerősítették, hogy a DNS-kötő doménben található mutációk átlagosan a vad típusú fehérje transzkripció aktivitásának csupán 13,8%-át mutatták. Ezen mutációk jelentős része funkcióvesztő (loss-of-function), míg kisebb részük domináns-negatív vagy akár funkciónyerő (gain-of-function) hatással bírt.

4.1.5 A mutáns TP53 fehérjék stabilitásának és kölcsönhatásainak elemzése

Az *in silico* stabilitásvizsgálatok kimutatták, hogy a legtöbb elemzett mutáció szignifikánsan csökkentette a TP53 fehérje szerkezeti stabilitását. A különböző predikciós módszerek eredményei között erős korreláció volt megfigyelhető, ami növelte az előrejelzések megbízhatóságát. A fehérje-fehérje kölcsönhatások modellezése során kimutattuk, hogy a p.G334R mutáció gyengítette a TP53 monomerek közötti kötődést, ami a funkcióhoz nélkülözhetetlen tetramer szerkezet kialakulását gátolhatja. A p.S362N variáns pedig megváltoztatta a TP53 és az azt szabályozó USP7/HAUSP fehérje közötti kölcsönhatás erősségét.

Az AML-MR, MDS-IB és MDS-LB csoportok összehasonlító elemzése kimutatta, hogy az AML-MR csoportban azonosított mutációk jártak a legsúlyosabb funkcionális következményekkel, mind a patogenitási pontszámok, mind a fehérjestabilitás-csökkenés, mind pedig a transzkripciós aktivitás elvesztése tekintetében. Az MDS-IB csoport ezen mutatók alapján egy köztes állapotot képviselt az AML-MR és az alacsonyabb kockázatú MDS-LB között.

4.2 A *BCOR*-génátrendeződéssel járó szarkómák *in silico* elemzése

4.2.1 A fúziós fehérjék szekvencia- és doménjellemzői

A ritka *BCOR*-átrendeződéssel járó szarkómák (BRS) patomechanizmusának vizsgálatához a szakirodalomban leírt kilenc ismert génfúziót elemeztük átfogó bioinformatikai módszerekkel. Az analízis kimutatta, hogy bár a fúziós események jelentős szerkezeti változásokat okoztak, a *BCOR* fehérje kulcsfontosságú, PCGF1-kötő

PUDF doménje minden vizsgált fúziós fehérjében megmaradt. Ugyanakkor a fúziók során gyakran elvesztek más funkcionális domének, mint például a BCOR BCL6-kötő doménje (Bbs), ami a fehérje transzkripciós korepresszor funkciójának károsodásához vezethet. A fúziós partnerek is jelentős trunkálódást szenvedtek, elveszítve kritikus funkcionális régióikat, mint például a CCNB3 destrukciós boxát, a KMT2D a hiszton-metiltransferáz aktivitásért felelős SET domént.

4.2.2 A fúziós fehérjék fizikokémiai és lokalizációs tulajdonságai

A fúziós fehérjék *in silico* elemzése a fizikokémiai tulajdonságok markáns megváltozását mutatta ki a vad típusú fehérjékhez képest. Általánosságban a fúziós fehérjék instabilitási indexe megnőtt, míg hidrofobicitásuk csökkent, ami a fehérjék szerkezeti stabilitásának romlására és potenciálisan megváltozott oldékonyságára utal. A sejten belüli lokalizációt előrejelző modellek a legtöbb fúziós fehérje számára - a vad típusú BCOR-hoz hasonlóan - sejtmagi elhelyezkedést jósoltak. A fehérjék belső rendezetlenségének vizsgálata a fúziós fehérjék PUDF doménjének C-terminális régiójában szignifikánsan megnövekedett strukturális rendezetlenséget tárt fel a fúziós fehérjék egy részében, ami alapvetően befolyásolhatja a fehérje kötődési képességeit és a PRC1.1 komplex megfelelő összeszerelődését.

4.2.3 A komplexek háromdimenziós modellezése és a kötési affinitás statikai elemzése

A fúziós események funkcionális következményeinek mélyebb megértése érdekében a BRS esetében leírt fúziós fehérjék, a vad típusú BCOR vBCOR és a PCGF1 partnerük által alkotott komplexek háromdimenziós szerkezetét modelleztük. A modellezéshez az AF3 mélytanulási algoritmust alkalmaztuk, amely nagy pontossággal képes

előre jelezni a fehérjék és komplexeik térszerkezetét. A modellezés kiterjedt a BCOR-PCGF1 dimerekre, valamint a tágabb, a nem-kanonikus Polycomb represszív komplex 1 (PRC1.1) alegységeit (KDM2B, SKP1) is tartalmazó tetramer komplexekre.

A komplexek stabilitásának és a fehérjék közötti kölcsönhatások erősségének számszerűsítésére a PRODIGY webszervert használtuk, amely a 3D szerkezet alapján megbecsüli a kötési affinitást, a Gibbs-féle szabadenergia-változás (ΔG) értékével kifejezve. Az elemzések szignifikáns változásokat mutattak ki a kötési affinitásban a vad típusú és a fúziós komplexek között. Az eredmények alapján a fúziós fehérjék két csoportra voltak oszthatók: egyes fúziók, mint a BCOR::CCNB3, megnövekedett kötődési affinitást (alacsonyabb ΔG értéket) mutattak, ami egy túlzottan stabil, merev komplex kialakulására utal; míg más fúziók, például a ZC3H7B::BCOR és a CIITA::BCOR, jelentősen csökkent affinitással (magasabb ΔG értékkel) rendelkeztek, ami a komplex instabilitását és gyengébb kötődését jelzi. A PRC1.1 komplex megfelelő epigenetikai működéséhez a stabilitás és a dinamikus rugalmasság finom egyensúlya szükséges a génszabályozáshoz elengedhetetlen összeszerelési és szétszerelési folyamatokhoz. Eredményeink szerint mind a túlzott stabilizáció, mind a gyengült kötődés megzavarja ezt a kényes dinamikát, ami a komplex funkcióvesztéséhez vezethet.

4.2.4 A dimer komplexek molekuláris dinamikai elemzése

A statikus 3D modellek elemzését követően 10 nanoszekundumos MD szimulációkat végeztünk a GROMACS szoftvercsomaggal, hogy dinamikus tekintést nyerjünk a BCOR-PCGF1 dimer komplexek

viselkedésébe. Az MD szimulációk megerősítették és tovább finomították a statikus modellekből levont következtetéseket.

A négyzetes középeltérés elemzése kimutatta, hogy a fúziós komplexek általában nagyobb szerkezeti instabilitást és konformációs rugalmasságot mutattak a vad típusú komplexhez képest, amely a szimuláció során stabilabb maradt. Különösen a ZC3H7B::BCOR és a CIITA::BCOR fúziók mutattak magas és ingadozó RMSD értékeket, ami tartós szerkezeti instabilitásra utal.

A két fehérje közötti hidrogénkötések számának elemzése feltárta, hogy a gyengébb kötési affinitású fúziós komplexek (pl. ZC3H7B::BCOR) szignifikánsan kevesebb stabil hidrogénkötést alakítottak ki a PCGF1 fehérjével, ami közvetlenül magyarázza a csökkent kötőerőt és a komplex instabilitását.

A girációs sugár vizsgálata megmutatta, hogy a fúziós komplexek kevésbé voltak kompakta és kiterjedtebb konformációt vettek fel, mint a szorosán illeszkedő vad típusú dimer. Ez a szerkezeti "fellazulás" arra utal, hogy a fúziós események megváltoztatják a komplex általános térbeli elrendezését, ami befolyásolhatja annak funkcióját.

Az MD szimulációkból számított kötési energiák erős korrelációt mutattak a PRODIGY által becsült statikus ΔG értékekkel, ami alátámasztja az eredmények megbízhatóságát, és konzisztens képet ad a fúziók által okozott termodinamikai változásokról.

Összességében ezek az elemzések mechanisztikus magyarázatot adnak arra a paradoxonra, hogy a BRS-ben a kulcsfontosságú PUF3 domén megmaradása ellenére miért károsodik a PRC1.1 komplex működése. A fúziók nem csupán a domének elvesztésével, hanem a fehérjék finom szerkezeti és dinamikai tulajdonságainak megváltoztatásával - a rugalmasság növelésével, a kölcsönhatások

gyengítésével és a globális szerkezet destabilizálásával - járulnak hozzá a génszabályozás hibáihoz és a daganatok kialakulásához

5. Főbb eredmények és következtetés

Röviden összefoglalva, jelen Ph.D munka főbb megállapításai a következők:

- Sikeresen elvégeztük 77 mielodiszpláziás szindrómában (MDS) és akut mieloid leukémiában (AML) szenvedő beteg *TP53* mutációinak molekuláris és *in silico* elemzését, mely során 26 betegben (33,8%) összesen 41, köztük 30 egyedi típusú *TP53* mutációt azonosítottunk.
- A *TP53* mutációk gyakorisága szignifikánsan magasabb volt az AML-MR (mielodiszpláziával összefüggő AML) alcsoportban (57,7%), mint az MDS-IB (emelkedett blasztarányú MDS, 33,3%) vagy MDS-LB (alacsony blasztarányú MDS, 17,9%) alcsoportokban, és rosszabb túléléssel, valamint komplex kariotípussal társultak.
- Az *in silico* patogenitási, stabilitási és transzkripciók aktivitást előrejelző elemzések szignifikáns különbségeket tártak fel az AML-MR, MDS-IB és MDS-LB csoportokban kimutatott *TP53* mutációk között, ami alátámasztja a különböző klinikai lefolyást és prognózist.
- A fehérje-fehérje (pl. homotetramerizáció, USP7 kötődés) és fehérje-DNS kölcsönhatások *in silico* modellezése rávilágított egyes *TP53* mutációk funkcionális következményeire, mint például a tetramerizáció zavara vagy a DNS kötődés gyengülése, hozzájárulva patogén mechanizmusuk megértéséhez.
- Az alkalmazott *in silico* eszközök (pl. Align-GVGD, REVEL, BayesDel, I-Mutant2.0, DynaMut2, DdGun, mCSM-PPI2, mCSM-NA) hasznosnak bizonyultak a *TP53* gén ismeretlen klinikai

jelentőségű variánsainak (VUS) értékelésében és patogénitásuk előrejelzésében MDS és AML kórképekben.

- Kilenc, az irodalomban korábban leírt, BCOR-átrendeződéssel járó szarkómában (BRS) előforduló BCOR génfüzió átfogó *in silico* funkcionális elemzését végeztük el, vizsgálva a szekvenciális, doménszerkezeti, fiziko-kémiai, lokalizációs és rendezetlenségi tulajdonságokat.
- Megállapítottuk, hogy bár a BCOR fehérje PRC1.1 komplexhez kötődés szempontjából kritikus PUF1 doménje minden vizsgált fúziós fehérjében megmaradt, a fúziók jelentős szerkezeti és fiziko-kémiai változásokat idéztek elő a vad típusú fehérjékhez képest.
- Az AlphaFold3 és PRODIGY segítségével végzett 3D szerkezetmodellezés és kötési affinitás számítások kimutatták, hogy a BCOR fúziók szignifikánsan befolyásolják a BCOR és a PCGF1 fehérje közötti dimerizációt, egyes fúziók (pl. BCOR::CCNB3, BCOR::MAML3) növelték, míg mások (pl. ZC3H7B::BCOR, CIITA::BCOR) csökkentették a kötési affinitást a vad típusú komplexhez képest.
- A nem-kanonikus PRC1.1 komplex (BCOR/Fúzió, PCGF1, KDM2B, SKP1) modellezése és kötési affinitásának elemzése arra utal, hogy a BCOR fúziós fehérjék megzavarhatják a komplex normál összeszerelődését és/vagy epigenetikai funkcióját.
- Az *in silico* modellezés értékes eszköztárnak bizonyult a BRS háttérben álló BCOR fúziók patomechanizmusának vizsgálatára, lehetővé téve a fúziós fehérjék által okozott szerkezeti és kölcsönhatás béli változások részletes elemzését.

6. Publikációs lista



**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.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

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., Mótán, 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, 2025.05.28.



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.