

DIPLOMADOLGOZAT

Blaha Bence

2024



Magyar Agrár- és Élettudományi Egyetem
Szent István Campus
Genetikai és Biotechnológiai Intézet
mesterképzési szak

A Harmonio sgRNS tervező szoftver hatástanulmánya

Belső konzulens:	Dr. Hoffmann Orsolya Ivett Tudományos főmunkatárs
Belső konzulens intézete/tanszéke:	Genetika és Biotechnológia Intézet, Állatbiotechnológia tanszék
Külső konzulens:	Nincs Nincs
Készítette:	Blaha Bence

Gödöllő
2024

Tartalom

Bevezetés és célkitűzések.....	2
1. Szakirodalmi áttekintés	3
2. Alkalmazott módszerek.....	20
2.1. A Harmonio szoftver rendszerkövetelményei:	20
2.2. In silico elemzés	22
2.3. Statisztikai elemzés	22
3. Eredmények és értékelésük	26
3.1. In silico elemzés	26
3.2. Statisztikai elemzés	28
4. Következtetések és javaslatok	31
5. Összefoglalás	32
6. Irodalomjegyzék	33
7. Táblázatok és ábrák jegyzéke.....	37
8. Hallgatói nyilatkozat	46
9. Konzulensi nyilatkozat.....	47

Bevezetés és célkitűzések

Modern, felgyorsult világunk sok kihívás elé néz, ilyen az ember által okozott klímakatasztrófa is (Lynas et al., 2021). Ennek hatására az egyes területeken eddig jó termelőképességű állatfajták nem képesek megfelelő gyorsasággal adaptálódni a gyorsan változó éghajlati viszonyokhoz, mely a termelés visszaesésével jár (Aydinalp et al., 2008). Másik kiemelkedő globális problémánk a túl népesedés, számszerűen előrejelzések szerint a világ népessége 2050-re 7,2-ről 9,6 milliárdra fog nőni. (Tännsjö et al., 2004). A sokasodó népesség következtében egyre több embert kell ellátni élelmiszerrel, ám erre a hagyományos fajták egyre kevésbé alkalmasak. Ezen tényezők teszik az állatok termelőképességének fokozását az állattenyésztés egyik fő célkitűzésévé (Rojas-Downing et al., 2017).

Sajnos a hagyományos nemesítési módszerek alapvetően a spontán mutáción alapulnak, mely következtében az élőlény genomjában véletlenszerű változások alakulnak ki, melyek lehetnek optimálisak vagy hátrányosak az egyed felhasználása szempontjából. Másik megemlíthető szempont, hogy egy-egy megfelelő tulajdonság rögzülése a populációban állatfajtól függően akár évtizedeket is igénybe vehet, és épp ezért rengeteg anyagi ráfordítást igényel (Georges, 2001).

A modern genomszerkesztés ezekre a kihívásokra szolgál megoldásként. A 1986-ban kifejlesztett cink-ujj nukleázok segítségével egy másik eszközt is kapott a tudomány a már ekkor is létező ráadásul jóval olcsóbb kiméra technika mellett. Ahhoz, hogy az adott élőlény genomjában változásokat hajtsunk végre, például géneket üssünk ki, így információt nyerjünk azok funkciójáról, bár a ZFN mind ezt jóval egyszerűbbé tette. Anyagi szempontból viszont nem volt siker történet, ugyanis használatához külön fehérjemérnökségre is szükség volt, mert bonyolulttá, és költségessé tette az eljárást (Burno et al., 2019).

Az ezt követő új megoldás a TALEN (Transcription activator-like effector nucleases) már olcsóbb volt, hisz nem kellett a fehérjék térszerkezetét módosítani minden lókuszhoz. Ezen felül további előnye hogy általában nagyobb célspecifitással rendelkeztek így csökkentve az off-target-ek valószínűségét. A TALEN-eket a ZFN-ekhez képest viszonylag egyszerűbb megtervezni és kifejleszteni. A TALEN-ek moduláris jellege egyszerűbb testreszabást tesz lehetővé, hogy specifikus DNS-szekvenciákat célozzanak meg. A TALEN DNS-kötő domén minden egyes ismétlődése egyetlen DNS-bázisnak felel meg, ami lehetővé teszi a célpont felismerésének pontos tervezését, ezen felül a TALEN rendszerek alacsonyabb szintű sejttotoxicitást mutatnak (Becker-Boch, 2021). Ám ennek a módszernek elterjedése nem történt

meg, ugyan is 2012 évben megjelent a CRISPR Cas9 (clustered regularly interspaced short palindromic repeats), mely elődeinél sokkal könnyebben tervezhető, gyorsabb, egyszerűbb, olcsóbb genomszerkesztés tett lehetővé. A természetben bakteriális immunrendszerként működő rendszert biotechnológiai eljárásokhoz optimalizálták.

Rengeteg kutatás foglalkozott azzal, hogy javítsák az eljárás hatékonyságát, így létrehoztak többféle CRISPR rendszert, más baktériumokból származó Cas fehérjékkel együtt, de még teljesen mesterséges endonukleázokat is létrehoztak már. Ezen vizsgálatok egy másik iránya a gRNS-ek fejlesztése, hogy növeljék az on-target aktivitást (Javaid-Choi, 2021).

A CRISPR rendszerrel való munka egyik első, és fontos lépése a sgRNS megtervezése, melyet manapság külön erre fejlesztett programokkal lehet *in silico* végezni. Jelenleg több különböző szoftver áll a kutatók rendelkezésére (Zhang et al., 2020).

A szakdolgozat témája egy olyan szoftver elkészítését tűzte ki célul, mely a kutatóknak *in silico* segít a megfelelő nukleinsav kiválasztását Mesterséges Inteligencia (MI) alkalmazásával. A jelenleg elérhető legtöbb hasonló szoftverrel ellentétben a jelenlegi verzió házi nyúlra (*Oryctolagus cuniculus var. domestica*) optimalizált.

A szoftver része egy knock-in-t segítő modul, mely segít a homológ karok keresésében, itt meg lehet jelölni a karok kívánt hosszát.

Ezen felül a többi hasonló programmal ellentétben itt specifikusabb guide RNS-eket lehet keresni, melyek off-target hatása jelentősen kisebb lesz a jelenleg elérhető szoftverek által számítottaknál, a kutató dönti el hány bázispárból álló RNS-el kíván dolgozni a 20 bp hossz helyett, mely más szoftvereknél az alapbeállítás.

Célkitűzések

Fő célom az volt, hogy *in silico* összehasonlítsam a Harmonio szoftver által nyúl genomra tervezett sgRNS-ek hatékonyságát a szakirodalomban publikált sgRNS-ekkel, valamint egy gRNS hatékonyságának tesztelése egér embriókon.

1. Szakirodalmi áttekintés

Napjainkban a legelterjedtebb genomszerkesztési eljárás a CRISPR Cas9 (Vasileva et al., 2015). Ezzel a knock out élőlények akár 100%-os hatékonysággal is létrehozhatóak, ám knock in esetében a hatékonyság jóval alacsonyabb csupán 2-38%, bár ez kutatása válogatja (Javaid-

Choi, 2021), (Peng et al., 2016). Bár mind ez nem akadályozza a GMO-k, és vele a CRISPR technológia gazdaságban történő elterjedését. A valódi korlátot az jelenti, hogy a tudomány már régóta nem tudja mindenki számára megnyugtató eredményekkel alátámasztani biztonságos használatukat. A növényi és állati GMO-k közötti különbségtétel azonban jelentős. A növényi fajták, mint például a BT-toxint termelő kukoricahibridek, más kockázatot jelentenek, mint az olyan tulajdonságok miatt módosított állatok, mint például a myostatin gén kiütése miatti fokozott izomnövekedés. Ráadásul egyes nemzetek szabályozási keretei messze elmaradnak a tudományos fejlődés mögött (Lee et al., 2020).

Maga a rendszer három fő alkotórészből áll, a Cas vágófehérjéből, és a vágás helyét meghatározó CRISPR RNS-ből (crRNS), illetve a trRNS-ből, bár a modernebb eljárások ezen két nukleinsav egyesítéséből létrejött single guide RNS-sel (sgRNS) dolgozik.

A CRISPR technológia egyik elterjedt alkalmazása a génkiütés, amikor a Cas endonukleáz egy komplementer gRNS segítségével a genom egy meghatározott célszekvenciájához kapcsolódik. A célhelyet elérve és a Protospacer Adjacent Motif (PAM) felismerését követően a vágóenzim kettős szálú törést (DSB) idéz elő. Ezt követően a DNS-javító mechanizmusok vagy pontosan, vagy hibásan javítják a vágást, ami potenciálisan működésképtelenné teszi a gént. (Javaid-Choi, 2021).

Másik fontos eljárás a knock in, mely esetben nem a gén elrontása a cél, hanem egy új genetikai elem beütése a kívánt eredmény. Ez a KO-tól ott tér el, hogy a sejtbe nem csak az RNS-t, és a Cas fehérjét juttatják be, hanem egy inszert DNS-t is, mely a beütni kívánt genetikai információt tartalmazza. Ennek két végére úgynevezett homológ karokat helyeznek, melyek a DSB-kor keletkezett két DNS véggel megegyező bázissorrendű nukleinsavak. A DNS javító mechanizmus így a templát alapú javítás során a bejuttatni kívánt szekvenciát beépíti a törés helyére (Jackson et al., 2017).

A modern genomszerkesztés a biológia sok gyakorlati területén próbálják alkalmazni a kutatók, ilyen a mezőgazdaság mellett a gyógyítás is (Luthra et al., 2021). Történetileg az egyik legnagyobb áttörést a gén targetálás elérése jelentette Homológ Rekombináció (HR) segítségével. Így lehetőség nyílt az egyes gének pontos manipulálására génkiütött állatok létrehozásával, ám ennek a kiméra technikának hatékonysága kezdetben kifejezetten alacsony volt, így nagyon sok időt, és pénzt emésztett fel egy-egy állatmodell létrehozása (Hyun-Los Angeles, 2019). Épp ezen javítandó, a kutatók programozható nukleázokkal kezdtek

dolgozni, melynek eredménye a fehérje mérnökök munkáját is igénylő cink-ujjak (ZNFs) és később a jóval olcsóbb transcription activator-like effector nucleases (TALEN) (Peng et al., 2016).

Azonban a jelenlegi technológiák közül a leghatékonyabb a CRISPR, melyet baktériumokban fedeztek fel, elsőként *E. coli*-ban írták le. A természetben a CRISPR/Cas rendszer a prokarióták adaptív, egyfajta immunrendszereként funkcionál, mely bármilyen idegen DNS-t felismer (Luthra et al., 2021). Az I. és III. típusú CRISPR/Cas rendszerek immunmechanizmusai bonyolultak, és a genomszerkesztési alkalmazásokban nem általánosan használtak. Ezzel szemben a II. típusú CRISPR/Cas rendszer egyszerűbb és egyetlen többfunkciós Cas9 fehérjére támaszkodik a behatoló genetikai elemekkel szembeni védelemként.

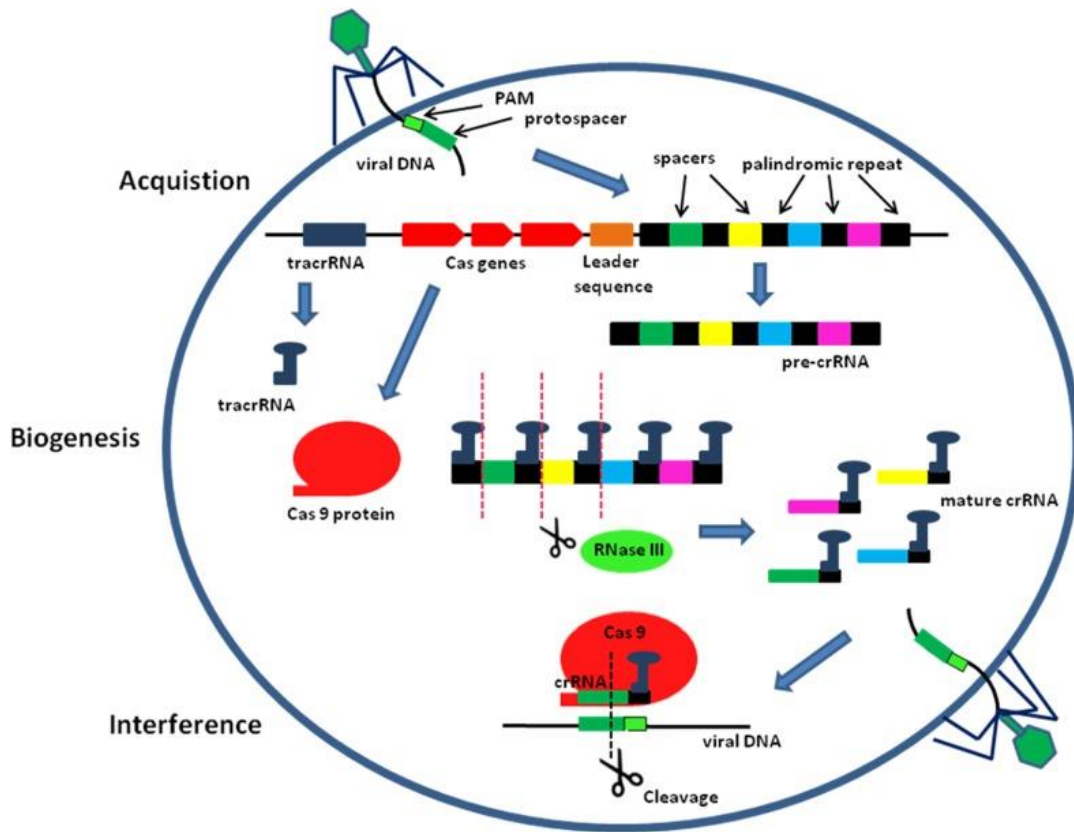
A natív II. típusú CRISPR/Cas9 rendszerben három komponens elengedhetetlen a célpontok hasításához: a Cas9 fehérje, a CRISPR RNS (crRNS) és a transzaktiváló crRNS (tracrRNS). A tracrRNS segíti a crRNS érését és megkönnyíti a Cas9-komplex összeszerelését. A II. típusú CRISPR/Cas rendszer három fő szakaszból áll (1. Ábra):

Az első fázisban a Cas nukleáz a behatoló fág DNS-ét kis DNS-töredékekké, úgynevezett protospacer-szekvenciákká aprítja fel. Ezek a szekvenciák ezután új spacer-ként beépülnek a baktérium genom CRISPR-lokuszába. Minden egyes CRISPR array így szerzett spacereket tartalmaz, amelyeket ismétlődő szekvenciák választanak el egymástól. A protospacerek kiválasztását a vírusgenomban található protospacer szomszédos motívumok (PAM) befolyásolják. A CRISPR-lokuszs protospacer-szekvenciái azonban nem tartalmazzak PAM-helyeket. A specifikus PAM-szekvencia attól függően változik, hogy a Cas9 fehérje milyen fajból származik, például 5'NGG-3' a *Streptococcus pyogenes*-ben és 5'-NGGNG-3' a *Streptococcus thermophilus*-ban.

crRNS biogenezis: Ebben a lépésben a CRISPR-lokuszs átíródik egy prekursor CRISPR RNS-é (pre-crRNS). A tracrRNS kötődik a pre-crRNS repeat szekvenciáihoz, és az endogén RNáz III hasítja ezt a komplexet, hogy érett crRNS-ek keletkezzenek. Minden crRNS tartalmaz egy spacert és egy részleges repeat szekvenciát.

Interferencia a behatoló DNS-sel: Az érett crRNS a Cas9 fehérjét a komplementer idegen nukleinsavakhoz vezeti, ami a betolakodó fág DNS lebontásához vezet. A Cas9 fehérje HNH és RuvC nukleáz doménekkel rendelkezik, amelyek a komplementer, illetve a nem-komplementer DNS-szálakat hasítják. Ez az endonukleáz kettősszál-töréseket (DSB-eket) hoz

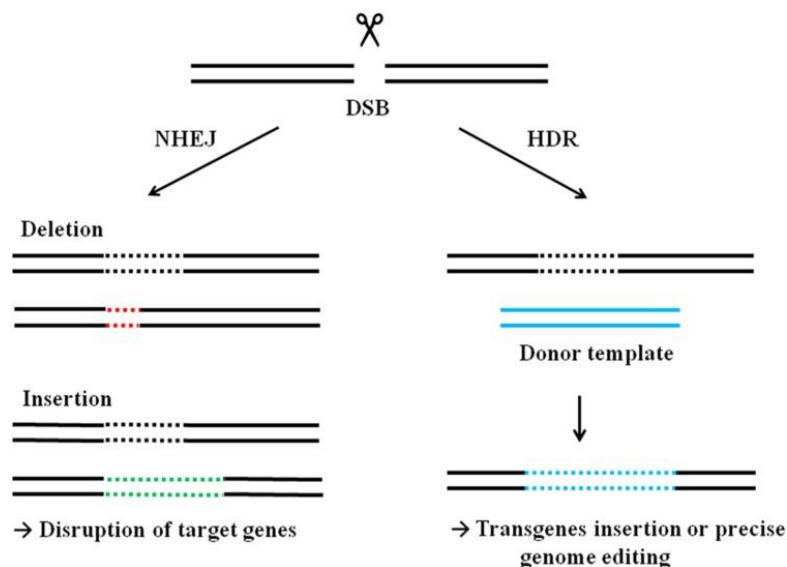
létre a vírusgenomban 3 bázispárral a megfelelő PAM előtt. A PAM-szekvencia létfontosságú ehhez a beavatkozási szakaszhoz, mivel különbséget tesz az idegen behatoló DNS és a gazdagenom CRISPR-lókuszai között, amelyekből hiányzik a PAM (Hryhorowicz et al., 2017).



1. ábra: A bakteriális CRISPR/Cas adaptív immunrendszer három fázisa: a spacer megszerzése, a crRNS érése és a vírus DNS-sel való interferencia. (Forrás: Hryhorowicz et al., 2017 nyomán)

Ezt a rendszert optimalizálták biotechnológiai alkalmazásokra, ennek részeként ma már a kutatások többsége nem két RNS-el dolgozik, hanem egy kimérikus sgRNS-el, melynek 20 bázispárnyi guide RNS része a cél lókusznak megfelelően variálható. Ha egy élő sejtet akarunk genom editálni, miután a Cas-RNS komplex létrehozta a DSBs-t, két fajta DNS javító mechanizmus indulhat el. Az egyik neve nem homológ végek összekapcsolása (NHEJ), a másik a Homológján Alapuló Javítás (HDR). Ha nem áll rendelkezésre templát, az NHEJ indul el, és összeilleszti a törés két végét ezzel inszerciákat vagy deléciókat (indel) okozva (2. Ábra). Ez, ha az elcsúszást okoz a leolvasási keretben, akkor korai stop-kodont hozhat létre. Ha az inszerció vagy deléció hárommal osztható számú nukleotidot érint, akkor ugyan nem változik a leolvasási keret, de módosulhat a fehérje funkciója vagy korai stop kodont

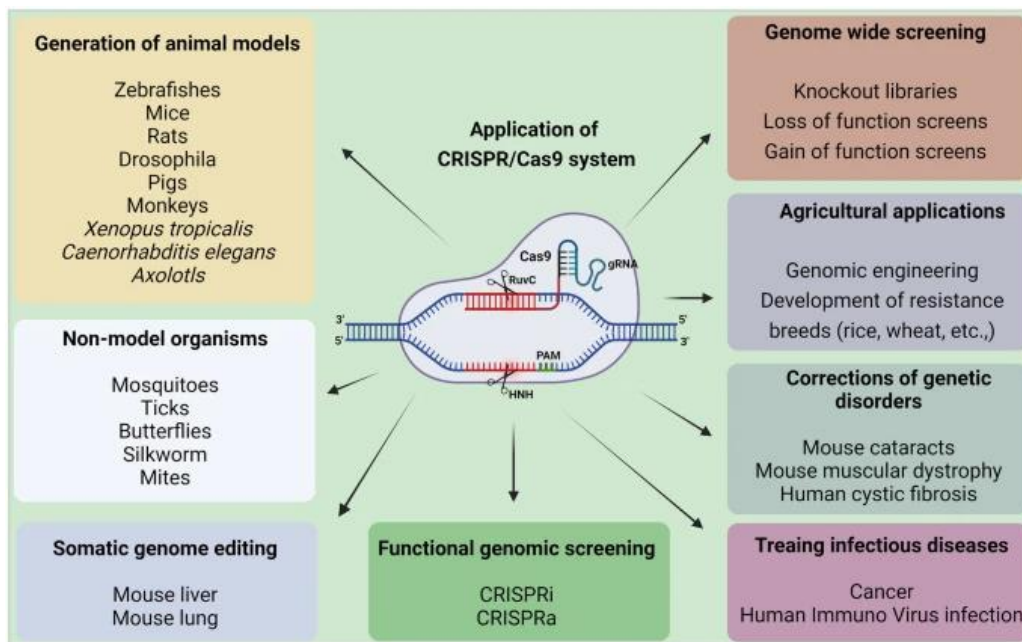
hozhat létre. Donor templát jelenlétében a HDR segítségével specifikus mutációkat, deléciókat, inszerciákat lehet végrehajtani (Peng et al., 2016).



2. ábra: Génszerkesztéssel mesterségesen módosított nukleázok segítségével kettősszálú törést (DSB) indukálhatunk a célszekvencián, amelyet a nem-homológ végcsatlakozási (NHEJ) vagy homológia irányított (HDR) DNS repair mechanizmusok javítanak. (Forrás: Hryhorowicz et al., 2017 nyomán)

A CRISPR/Cas rendszer széleskörű alkalmazásra talált mind az alaputatásban, mind a klinikai környezetben, különböző betegségek esetében, valamint kórokozó diagnosztikában, környezeti károk mérséklésénél, élelmiszer biztonságban, növény nemesítésben és állattenyésztésben. Érdeemes megjegyezni, hogy a CRISPR/Cas rendszer felhasználása komplex munkafolyamatot foglal magában, amely sejtes munkánál a konkrét CRISPR/Cas rendszer kiválasztásából, a gRNS-tervezésből, a transzfekcióból, az egysejtes klónok létrehozásából, a klónok szűréséből és a szisztematikus mutációelemzésből áll (Zhang et al., 2020). E lépések mindegyike jelentős időbeli, pénzügyi és emberi erőforrásokat igényel. Szerencsére az informatika fejlődése lehetőséget kínál e kihívások kezelésére és a teljes folyamat hatékonyságának növelésére. A különböző algoritmusokat és keretrendszereket használó *in silico* módszerek különböző előnyöket kínálnak, és számos alkalmazáshoz optimálisak (3. Ábra) (Hillary-Caesar, 2023).

Egyszerű felépítésük és könnyű használhatóságuk miatt a II. típusú CRISPR/Cas rendszerek szélesebb körben elterjedtek. Ennek eredményeképpen a jelenlegi *in silico* gRNS-tervezési eszközök többsége a II. típusú rendszerekhez igazodik.



3. ábra: A CRISPR/Cas9 rendszer precizitása, gyorsasága és költséghatékonysága révén átalakította a genomszerkesztést, és alkalmazások széles körét tette lehetővé. Bevezetése óta ezt a rendszert betegségek gyógyításában, új terápiák kifejlesztésében és a precíz orvoslás fejlesztésében alkalmazzák. Az emberi egészségügyön túl a CRISPR/Cas9 rendszert többek között a modell és nem modell állatok életműködéseinek tanulmányozására, szomatikus genomszerkesztésre, bioüzemanyagok előállítására és olyan növényfajok módosítására is használják, mint a rizs és a búza. (Forrás: Hillary-Caesar, 2022 nyomán)

Különböző alapelvek alapján a gRNS-tervezési eszközöket három fő csoportba soroljuk (Zhang et al., 2020):

1. A mintafelismerés kategória a bázispárosítás elvén működik. Ebben a típusban az eszközök olyan szekvenciafragmentumokat keresnek, amelyek egy rövid PAM és körülbelül 20 bázispárnyi guideRNS (gRNS) szekvenciát tartalmaznak, amelyek kiegészítik a célszekvenciát egy adott genomon belül. A jelölt gRNS-nek az on-target hatás valószínűsége annál nagyobb, minél kevesebb a nukleotid különbség. Ezenkívül a specifikus PAM-ot előre meg kell határozni, mivel a különböző CRISPR/Cas-rendszerek között ez változik. A gRNS-mintázat másik meghatározó tényezője a transzkripció módszer, az U6 és a T7 promóterek működéséhez a gRNS-en az 5' végén G és GG bázisoknak kell lenniük. Ez főleg sejtes rendszereknél kell figyelembe venni, amikor a CRISPR részei egy plazmidon helyezkednek el. Az U6 promóterből expresszált gRNS-ek potenciális target-einek speciális megkötéseknek kell megfelelniük. Ezek közé tartozik, hogy a gRNS 5' végén egy G-nek kell lennie, és a target mellett egy NGG-szekvenciának. Míg a T7 promóter esetében az első két nukleotidnál GG-re, a target mellett pedig NGG-szekvenciára van szükség.

2. A "jellemző szabály" (Feature rule) kategória a különböző target-ek változó szerkesztési aktivitásának megfigyelései nyomán alakult ki, ami a hasítással szembeni érzékenységben rejlő különbségekből ered. Ez ösztönözte a célzás hatékonyságát befolyásoló kulcsfontosságú jellemzők vizsgálatát. Ilyen jellemzők például a gRNS-ek G/C-tartalma (a magas vagy alacsony G/C-tartalom alacsonyabb aktivitást jelez), a frameshift mutációk gyakorisága (negatív korreláció a CRISPR hatékonyságával), a poli-T szekvenciák jelenléte (a gRNS-átírás terminátoraként működnek), a Cas kötődési preferenciáját befolyásoló bázis-összetétel (a PAM-ot megelőző G előnyben részesítése és a pirimidinek hiánya a gRNS-spacerek utolsó 4 nukleotidjában), az exonok helyzete (alacsonyabb hatékonyságot észleltek, ha a terminális kódoló exont célozzák meg, mint a korábbi exonokat), a PAM-hoz közeli, motívumokkal és jellemzőkkel dúsított körülbelül 10-12 nukleotidos, magterületnek nevezett spacer-szekvenciák jellemzői (befolyásolják a párosítási folyamatot), stb.

Az ebbe a kategóriába tartozó eszközök különböző mérhető jellemzőket integrálnak alapvető mintafelismerési megközelítésekkel, hogy átfogó információt nyújtsanak a gRNS-jelöltekről és a célpontokról. A felhasználók saját szűrési szabályokat állíthatnak fel a jellemzőindexek és a megfelelő küszöbértékek alapján, hogy kizárják a kevésbé megbízható vagy érdektelen gRNS-eket. A Cas-Designer (<http://www.rgenome.net/cas-designer/>) (Park et al. 2015) például a feltételezett gRNS-eket a G/C arányokkal és a kereten belüli mutációk gyakoriságát jelző out-of-frame pontszámokkal együtt mutatja be.

Az ebbe a kategóriába tartozó eszközök azonban gyakran különálló értékeléseket vagy több jellemző önkényes kombinációját nyújtják, ahelyett, hogy integrált elemzéseket végeznének felhasználói interakciók segítségével. Ez a megközelítés összezavarhatja a felhasználókat azzal kapcsolatban, hogy miként egyeztethetők össze a több jellemzőből származó, potenciálisan ellentmondásos eredmények. E kihívás kezelésére a gépi tanulási algoritmusok kínálnak megoldást.

3. Gépi tanulás kategória: A több jellemző súlyozását övező bizonytalanság fényében a kutatók olyan matematikai algoritmusokhoz fordulnak, amelyek szisztematikusan integrálják a jellemzőket az optimális gRNS-kiválasztás finomítása érdekében. Ezek a modellek a képzési adataikban felhasznált algoritmusok és információk tekintetében különböznek. Doench és munkatársai (Rule set 1) például a sejtfelszíni markereket

célzó gRNS-ek kimerülési arányát figyelték meg egér és emberi sejtekben, és a célszekvenciák belső nukleotid-összetételének tulajdonították (Doench et al. 2016). Ezt az adatot aztán felhasználták egy logisztikus regressziós osztályozó megalkotásához a gRNS-aktivitás előrejelzésére.

CRISPRscan, egy lineáris regressziós modell, a nukleotid-összetétel CRISPR/Cas9 hatékonyságra gyakorolt hatását vizsgálta a célszekvenciák gRNS-indukált mutációs rátáit zebrahal-embriókban aktivitási jelként használva. Hasonlóképpen, az sgRNA Scorers v2.0, amely támogató vektor gépeken alapul, szekvenálási adatokat használt fel (humán HEK293T sejtekben lévő célszekvenciák mutációs rátái). A TUSCAN újra elemezte a publikált adatokat, és a flanking célterületekre vonatkozó jellemzők beépítésével és az algoritmus véletlen erdő döntési fa módszerrel való felváltásával javította a predikciós teljesítményt.

A mélytanulás-alapú eszközök a jellemző (feature) kivonás automatizálásával minimalizálják a torzításokat. A DeepCRISPR egyesíti az on-target és az off-target előrejelzéseket egy keretrendszerben, és lehetővé teszi az epigenetikai jellemzők alkalmazását a fenotípus-alapú adatok mellett. Míg a fenotípusvezérelt modelleket befolyásolják a célpontok pozíciói, a szekvencia-alapú modellek a genetikai mutációk közvetlen mérését kínálják, ami jobb általánosíthatóságot eredményez. A fenotípusvezérelt modellek akkor jeleskednek, ha a felhasználók a gRNS-indukált mutációk funkcionális eredményeit helyezik előtérbe, míg a szekvencia-alapú modellek szélesebb körben alkalmazhatók, ha kizárólag a genotípus-változásokra összpontosítanak.

Bár az *in silico* gRNS-tervezők erőteljesen fejlődtek, a gépi tanuláson alapuló eszközök teljesítményének általánosítása, több fajra, és eljárásra való kivetítése továbbra is kihívást jelent a fajok és Cas enzimek eltérő jellemzői miatt, ami egyesével történő feltöltési folyamatokat tesz szükségessé. A feature rule-on kívül ezek a gRNS-tervezők más különálló specifikációkkal is rendelkeznek, mint például a párosított gRNS-ek (pgRNS) egylépéses testre szabása a nagy fragmentumok törléséhez, a CRISPR-aktiválás vagy -interferencia (CRISPRa/i), az alkalmazási platformok, az off-target előrejelzés. Ezek a specialitások egyedi képességekkel ruházzák fel az eszközöket az egyes területeken, így a felhasználóknak több, a saját céljaikra szabott választási lehetőséget biztosítanak. Emellett egyes kereskedelmi eszközök vizuális felületeket, online konzultációt és egyablakos rendelési szolgáltatásokat kínálnak, mint például a Synthego és az IDT, bár a legtöbb kereskedelmi eszköz elsősorban a

CRISPR/Cas9 rendszerhez készült, és korlátozott támogatást nyújt más CRISPR-rendszerekhez.

Mivel egyetlen eszköz sem képes minden igényt kielégíteni, a gRNS tervező kiválasztásakor elengedhetetlen az előfeltételek és a konkrét célok gondos mérlegelése.

A CRISPR/Cas felhasználása jelentősen kiszélesedett a génfunkció feltárásában, ami a validált adatok gyűjtésére szolgáló nyílt adatbázisok iránti fokozott igényt eredményezett. Ezek az adattárak megkönnyítik a kutatók hozzáférését a nyers vagy feldolgozott adatokhoz. E sürgető igény kielégítésére számos adattárat hoztak létre. A GenomeCRISPR a screening eredmények kutatásközi összehasonlítása nélküli, kizárólag az eredményeket dokumentáló adatbázisokkal (mint például a CRISPRz, CrisprGE, CRISPRInc és BioGRID ORCS) összehasonlítva, a GenomeCRISPR 84 nagy áteresztőképességű szűrésen alapuló, intuitív összehasonlítást kínál a gRNS hatékonyságáról és a perturbációs fenotípusokról meghatározott körülmények között. A gRNS-információk pusztán aggregálása helyett a PICKLES újra elemzi a nyers szűrés adatokat, és összehasonlítja az egyes gének lényegességét több kísérletben, szövetben vagy sejttípusban. Két független, humán rákkutatásra szabott adatbázis, a Sanger DepMap és a Broad DepMap a CRISPR/Cas9 screening adatok elemzésével a génfüggőségek rögzítésére összpontosít a rákos sejtvonalakban (Zhang et al., 2020).

Egy sgRNS tervező szoftver tesztelésének első fázisa *in silico* zajlik, mely során összevetik a generált oligonukleotidokat egy adatbázissal, melyben megtalálhatóak a már ismert RNS-ek, és azok hatékonysága, erre példa a crisprSQL (Störtz-Minary, 2020). A jelenleg elérhető adatbázisok viszont főleg emberi sejtvonalakon, valamint egereken végzett kísérletek eredményeit tartalmazzák.

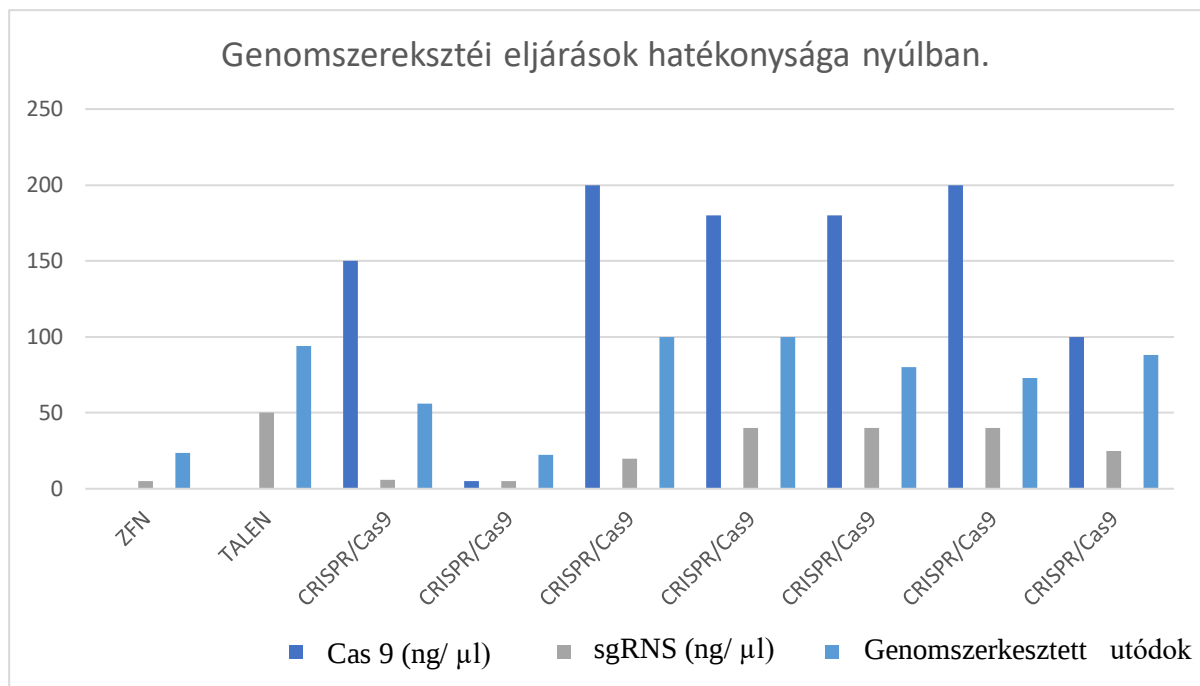
A házi nyúl (*Oryctolagus cuniculus* var. *domestica*) jelentős szerepet játszott a biológia és az orvostudomány kulcsfontosságú vizsgálataiban, a mai napig klasszikus állatmodellnek tekintik (Honda-Ogura, 2017). A nyulak kísérletekhez való felhasználásának előnyei közé tartozik, hogy áthidalják a kis- és nagyállatmodellek közötti szakadékot. Költséghatékonyabbak, mint a nagyobb állatok, és a laboratóriumi környezetben megfizethető elhelyezéssel és fenntartással rendelkeznek. Emellett könnyen tenyészthetők, közelebbi rokonságban állnak a főemlősökkel, mint a rágcsálók, és változatos genetikai hátterük révén közelebb állnak az emberhez (Bösze et al. 2016). A nyulak használatának

további előnyei közé tartozik, hogy szelíd és nem agresszív természetük miatt könnyen kezelhetőek és megfigyelhetőek. Rövid életciklusuk van a vemhesség, a szoptatás és a pubertás eléréséhez. Mivel a nyulak kisállatoknak minősülnek, a nyulakkal végzett kutatások a helyi etikai bizottságok hatáskörébe tartoznak. Ezzel szemben a nagyobb állatokkal végzett vizsgálatokhoz a központi etikai bizottságok engedélye szükséges, ami egy időigényesebb, szigorú előírásokkal járó folyamat. A nagyobb állatok etikai engedélyezéséhez és a nagyobb állatokon végzett sebészeti beavatkozásokhoz szükséges létesítmények bonyolultak és költségesek, és jellemzően csak a központi állatkutató létesítményekben valósíthatók meg (Mapara et al., 2012). Az európai nyúl az egerhez képest alkalmasabb modell például a veleszületett immunrendszer gégeinek tanulmányozására. A nyulak és az emberek közötti genetikai sokféleség lényegesen kisebb, mint az egerek és az emberek közötti. Az egerek és az emberek, illetve a nyulak és az emberek közötti genetikai távolságban mutatkozó különbség valószínűleg a rágcsálók vonalában a mutációs ráta növekedésének köszönhető (Soares et al., 2022).

A nyulak a nyúlfélék családjába és a nyúlfélék (*Lagomorpha*) rendjébe tartoznak. A taxonómusok 47 különböző nyúltörzset ismernek, de az Amerikai Nyúltenyésztők Szövetsége (ARBA) nemrégiben 49 törzset sorolt be. Míg az új-zélandi fehér (NZW) nyulakat az orvosbiológiai kutatásokban előnyben részesítik, más fontos fajták, például a holland öves, a kaliforniai fehér, és az új-zélandi vörös nyulak is gyakran modell állatként jelennek meg. Az Európai Unió 2010-es adatai szerint a nyulak az egerek és a patkányok után a harmadik leggyakrabban használt laboratóriumi emlősök között szerepelnek világszerte (Sengputa-Dutta, 2020). A nyulak felhasználása 1950 és 2010 között jelentősen csökkent, és ez a tendencia az Egyesült Államok Mezőgazdasági Minisztériumának (USDA)/Az Állat- és Növényegészségügyi Ellenőrző Szolgálat (APHIS) 1973 és 2015 közötti adataiból is kitűnik. Ez a csökkenés valószínűleg az állatkísérletek elleni globális tiltakozásoknak tulajdonítható, amelyek az orvosbiológiai kutatást az *in vitro* vizsgálatok irányába terelték. Az USDA/APHIS 2014 és 2016 között az állatkísérleti kutatások felhasználásának összességében 8%-os csökkenéséről számolt be, ami hozzájárult a laboratóriumi nyulak felhasználásának csökkenéséhez a 2000. évi 258754-ről 2016-ban 139391-re. A nyulak az USDA/APHIS adatai szerint 2016-ban a második legtöbbet használt állatok voltak az Egyesült Államokban a kutatásban és az oktatásban (Sengputa-Dutta, 2020).

Az EU-ban a BfR adatai szerint 2020-ban 70 840 nyulat használtak fel kísérleti célokra, mely a 3,7%-a az összes kutatásra használt állatnak (BfR, 2020).

Napjainkban mégis a legelterjedtebben használt modell állat a házi vagy laboratóriumi egér (*Mus musculus*), melynek laboratórium használata az elmúlt három évtizedben lett népszerű, elsősorban az embrionális őssejtek hozzáférhetőségének köszönhetően, amely elsőként lehetővé tette a génkiütéses (KO) állatok létrehozását. Az elmúlt néhány évben több faj csatlakozott a KO állatok sorához. Ezek közül a nyúl kiemelkedik a genomszerkesztés lehetőségeiben rejlő előnyeivel. Ezek közé tartozik a szuperovuláció könnyű indukációja, a mára kifinomult embriómanipulációs technikák és a rövid vemhességi idő. A CRISPR/Cas9 rendszer különböző állatfajok esetében bizonyítottan egyszerű és figyelemre méltóan hatékony eszköz, amely a nyulak genomszerkesztés technikájaként is megbízhatóan működik. 2014-ben úttörő munka mutatta be a CRISPR/Cas9 rendszer első alkalmazását a génkiütés (KO) indukálására nyulakban, ami kulcsfontosságú mérföldkövet jelentett. Ez a kezdeti kísérlet négy gén kiütésére összpontosított, konkrétan hiperlipidémia modelljeinek kifejlesztése volt a cél, és 55,9%-os átlagos hatékonyságot ért el, az off-target mutációk észlelhető jelei nélkül (Honda-Ogura, 2017). Figyelemre méltó, hogy több gén (három és öt gén) egyidejű kiütése ugyanabban az évben sikeresnek bizonyult, különösen az immunhiányos nyúlmodell létrehozásában kulcsfontosságú gének célzásával, ami felülmúlta a TALEN-rendszert alkalmazó korábbi erőfeszítéseket (Honda-Ogura, 2017). A azon módosítások, mint a Cas9 mRNS koncentrációjának 150 ng/μl-ről 200 ng/μl-re történő emelése és a gRNS koncentrációjának 6 ng/μl-ről 20 ng/μl-re történő növelése, valószínűleg hozzájárultak ehhez a sikerhez, bár a KO-utódok aránya 22,6%-ról 11,0%-ra csökkent, és azonosított off-target-ek is előfordultak. Ezek a felismerések hangsúlyozzák a CRISPR/Cas9 technológia alkalmazásakor a célszekvencia körültekintő kiválasztásának kritikus fontosságát. Az ezt követő erőfeszítések során a CRISPR/Cas9 rendszerrel a betegségek modellezése és a nagy szekvenciák (105 kb) delécioja terén elért eredményeket mutattak be, és optimális körülmények között 73% és 100% közötti szerkesztési hatékonyságot értek el (Honda-Ogura, 2017). A cikk adatait a 4. Ábrán összegzem.



4. ábra: Genomszereksztett nyulak előállítási hatékonysága. (Forrás: Saját szerkesztés Honda-Ogura, 2017-es cikkének adatai alapján.)

A Harmonio *in silico* teszteléséhez sajnos nem találtam olyan jelenleg az interneten elérhető adatbázist, amely validált sgRNS-eket tartalmazna, melyekkel KO/KI nyulakat állítottak volna elő, így már megjelent publikációkhoz fordultam.

Petheő L. Gábor és munkatársai a CRISPR/Cas9 technikát alkalmazták a NOX5 gén kiütésére új-zélandi fehér nyulakban. A Cas9 mRNS-t és a NOX5 CRISPR-vezetőt mikroinjektálták a zigótákba, amelyeket ezt követően álvemhes nyulakba ültettek. Az újszülöttekből genomi DNS-t izoláltak, és a kettős szálú DNS-töréseket T7 nukleáz teszt segítségével azonosították. A NOX5 gén pontos módosításának sikerességét DNS-szekvenálással tisztázták. Két vonalat hoztak létre, és a kísérletekhez kiválasztott vonal 10 bázispár deléciót tartalmazott a 3. exonban. Az ebből a vonalból származó knockout (KO) állatok nem termeltek funkcionális fehérjét. Az off-target hatások elemzése és a NOX1, NOX2 és NOX4 cDNS-ek szekvenálása megerősítette a CRISPR által kiváltott mutáció NOX5 génre való specititását (Petheő et al., 2021).

Qingyan Lv és munkatársai miosztatin KO állatokat hoztak létre. Az MSTN knockout nyulakat Cas9 mRNS és sgRNS együttes a zigótákba való befecskendezésével hozták létre a . Ezek az MSTN KO nyulak a kettős izom jellegzetes fenotípusát mutatták az izomrostok hiperpláziájával vagy hipertrófiájával. Ezenkívül hasonló fenotípus volt megfigyelhető az F1 generációban is, ami az MSTN mutáció stabil öröklődését jelzi az MSTN KO nyulakban.

Ezekből következik, hogy, az MSTN KO nyulak sikeres generálása a CRISPR/Cas9 rendszer segítségével nagy hatékonyságot mutat, és megbízható és hatékony állatmodellt biztosít az izomfejlődés és a kapcsolódó betegségek vizsgálatához (Lv et al., 2016).

Tingting Sui és munkatársai (Sui et al., 2018) a Duchenne-izomdisztrófiát (DMD) vizsgálták, ami egy X-kromoszómás izomsorvadási rendellenesség, amelyet a dystrophin gén mutációi okoznak. Ebben a tanulmányban a Duchenne-izomdisztrófia (DMD) nyúlmodelljét sikeresen hozták létre Cas9 mRNS és az 51-es exont célzó sgRNS együttes befecskendezésével nyúlzigótákba. A DMD knockout (KO) nyulak a DMD jellegzetes tulajdonságait mutatják, mint például súlyosan csökkent fizikai aktivitás, emelkedett kreatin-kináz szérumszint, valamint progresszív izomnekrózis és fibrózis. Figyelemre méltó, hogy 5 hónapos korban a rekeszizomban és a szívben is nyilvánvaló patológiát figyeltek meg, ami hasonlít a DMD-s betegeknek tapasztaltakhoz. Az echokardiográfiás felvételek kamratágulatot mutattak ki, csökkent ejekciós frakcióval és frakciórövidüléssel együtt a DMD KO nyulakban. Összefoglalva, a DMD ezen újonnan létrehozott, CRISPR/Cas9 rendszerrel létrehozott nyúlmodellje hűen reprodukálja a DMD-s betegeknek megfigyelt szövettani és funkcionális rendellenességeket, így jelentős lehetőségeket hordoz a preklinikai vizsgálatok számára (Sui et al., 2018).

Tingting Sui egy másik tanulmányában (Sui et al., 2018) az LGMD2L és az MMD3-al foglalkozott. A 2L típusú végtagi izomdisztrófia (LGMD2L) és a Miyoshi 3 típusú myopathia (MMD3) autoszomális recesszív izomdisztrófiák, amelyek az anoctamin-5 (ANO5), az anoctamin fehérjecsald egyik tagját kódoló gén mutációjából erednek. Sok ANO5-ös rendellenességben szenvedő betegnél az ANO5 génben pontmutációk vagy kis beillesztések/deleciók (indelek) találhatók. Az emberi ANO5 mutációk jobb reprodukálása érdekében a kutatók KO ANO5 nyulakat hoztak létre Cas9 mRNS és sgRNS együttes befecskendezésével a zigótákba. A mutáns nyulakban a CRISPR-mediált kis indelek a 12. és/vagy 13. exonban a tipikus izomdisztrófia tüneteinek megjelenését eredményezték, beleértve az emelkedett kreatin-kináz (CK) szérumszintet, izomnekrózist, regenerációt, zsírosodást és fibrózist. Ez az innovatív ANO5 mutáns nyúlmodell ígéretes a betegség patogenezisének vizsgálatához és az ANO5-hiányos izomdisztrófia terápiás megközelítéseinek feltárásához (Sui et al., 2018).

Tingting Sui egy másik cikkében (Sui et al., 2016) PHEX génkiütött nyulakat hozott létre. Az X-hez kötött hipofoszfátémia (XLH) az örökítő angolkór vezető oka, előfordulási

gyakorisága az emberekben 1/20 000. A foszfát-szabályozó endopeptidázt kódoló PHEX gén inaktíválódása vagy mutációja hipofoszfatémiát és károsodott csontmineralizációt okoz az XLH betegeknél. Sajnos jelenleg nincsenek megfelelő állatmodellek a fizioterápiák biztonságosságának értékelésére és az XLH angolkórságra vonatkozó gyógyszerek szűrésére. Ebben a vizsgálatban egy XLH modellt hoztunk létre a PHEX gén kiütésével (KO) a CRISPR (clustered regularly interspaced short palindromic repeats) 9-es (Cas9)/sgRNS mRNS-sel együttesen bejuttatott nyúl zigótákba. A PHEX KO nyulakban a növekedési elmaradás, a hipofoszfatémia, az emelkedett szérum FGF23 szérum szint és a csontmineralizáció károsodása jellegzetes fenotípusait figyelték meg, ami megkülönbözteti őket a normál kontrolltól. Lényegében ebben a tanulmányban először sikerült PHEX KO nyulakat generálni, amelyek hatékonyan rekonstruálják a humán XLH-t a CRISPR/Cas9 rendszer segítségével. Ez az úttörő XLH nyúlmodell jelentős potenciállal rendelkezik az XLH megelőzését és preklinikai terápiáját célzó gyógyszer-szűrések platformjaként (Sui et al., 2016).

Tingting Yuan és kollégái (Yuan et al. 2019) célja, hogy új nyúlmodellek létrehozása volt az LDL-receptor (LDLR) és/vagy apolipoprotein (apoE) gének nagyfragmentumú deléciójával, a hiperlipidémia és az ateroszklerózis vizsgálata céljából. A CRISPR/Cas9 rendszert többszörös sgRNS-rendszerrel irányítva nyúlembriókban alkalmazták az LDLR és apoE gének módosítására. Az alapító nyulak LDLR- és apoE-génjeinek szekvenálását végezték el, valamint a plazma lipid- és lipoproteinprofiljaik elemzését normál táp mellett. Emellett western blottingot alkalmaztak az apolipoprotein expressziójának értékelésére. Az aorta Szudán IV és HE festését végezték az ateroszklerózis kialakulásának megerősítésére. Hat knockout (KO) nyulat hoztak létre LDLR és apoE sgRNS-ek befecskendezésével. Közülük négy nyúl volt LDLR KO, kettő pedig LDLR/apoE kettős KO. E KO nyulak szekvenciaelemzése több mutációt fedezett fel, beleértve indeleket, deléciókat és szubsztitúciókat. Ezenkívül két nyúlvonalban biallelikus nagy fragmentum-deléciót mutattak ki az LDLR régióban. A plazma lipid- és lipoprotein-profiljaik vizsgálata normális táplálkozás mellett valamennyi KO nyúlnál jelentős hyperlipidaemiát mutatott ki, az összkoleszterinszint akár 10-szeresére is megnövekedett a vad típusú nyulakhoz képest. Két alapító nyúl patológiai vizsgálata a KO nyulaknál jelentős aorta- és koszorúér-atheroszklerózist mutatott ki (Yuan et al. 2019).

Rihong Guo és munkatársai (Guo et al., 2016) szintén miosztatin kiütött állatokkal kísérleteztek. A miosztatin (Mstn) az emlősökben a vázizomzat tömegének konzervált negatív

szabályozója. Mindazonáltal továbbra is bizonytalan, hogy az Mstn pontos megzavarása az állatállományban a hústermelékenység biztonságos növelése érdekében megvalósítható-e. Ennek megoldására CRISPR/Cas9 rendszert alkalmaztak Mstn knock-out (KO) nyulak és kecskék létrehozására, elemezve a fenotípusukban bekövetkező változásokat. Sikeresen létrehoztak 24 Mstn KO nyulat 32 újszülöttből, miután két, a nyúl Mstn-t célzó sgRNS-t befecskendeztek az embrióba. Ezek az Mstn KO nyulak megnövekedett születési súlyt mutattak, valamint a quadricepsz és bicepsz izmok súlyarányának az egész testhez viszonyított jelentős növekedését. Az Mstn KO azonban a megnagyobbodott nyelv jelenségének és a súlyos egészségügyi szövődményeknek, köztük a halvaszületésnek és a korai stádiumú elhalálózásnak a megnövekedett valószínűségét is eredményezte. Hasonlóképpen, ugyanezzel a módszerrel négy kecske közül egynél sikerült elérni, az Mstn lókuszon a sikeres szerkesztést. Ennek a kecskének a korai stádiumban elért növekedési rátája meghaladta a kontroll kecskéét. Összefoglalva, a CRISPR/Cas9 technológiával hatékonyan állítottak elő Mstn KO nyulakat és kecskéket. Mindazonáltal az Mstn KO súlyos egészségügyi problémákat eredményezett, és potenciálisan hasonló hatásokat idézhet elő más fajokban is. Ez a biztonsági aggály további vizsgálatokat tesz szükségessé az állati szaporítási folyamatokban való alkalmazás előtt (Guo et al., 2016).

Lin Yuan és kollégái (Yuan et al., 2016) GJA8 knockout nyulakat hoztak létre. A szürkehályog a globális látásromlás elsődleges oka, bár a sebészeti beavatkozások hatékonyan helyreállíthatják az érintettek látását. Eddig a megfelelő állatmodellek hiánya akadályozta a szürkehályog kezelésére szolgáló mesterséges lencsék biztonságosságának és a gyógyszerkölsönhatásoknak az in vivo vizsgálatát. A genetikai kutatások rávilágítottak a GJA8 szerepére a lencse átlátszatlanságának fenntartásában és a megfelelő lencsefejlődés biztosításában. Ebben a tanulmányban Cas9/sgRNS mRNS nyúl-zigótákba történő együttes befecskendezésével GJA8 génkiütéssel rendelkező szürkehályogmodellt hoztunk létre. Eredményeink 98,7%-os génmutációs hatékonyságot mutattak ki az embriókban és 100%-os hatékonyságot a kölykökben, ami aláhúzza a Cas9/sgRNS rendszer figyelemre méltó hatékonyságát a génszerkesztésben nyulakban. A korábbi kutatásokkal összhangban genetikai és szövettani elemzéseink kimutatták, hogy a GJA8 funkciójának megzavarása mikrophthalmiához, csökkent lencsemérethez és szürkehályoghoz vezetett. Összefoglalva, a szürkehályog innovatív nyúlmodellje jelentős ígéretet hordoz magában, mint a szürkehályog megelőzését és kezelését célzó gyógyszerészítés értékes eszköze (Yuan et al., 2016).

Jie Xu és munkatársai (Xu et al. 2021) a CFTR gént vizsgálták. A cisztás fibrózis (CF) jelenlegi állatmodelljei jelentősen hozzájárultak a CF patogenezisének megértéséhez, de rövid élettartamuk, a kulcsfontosságú fenotípusok hiánya és a magas fenntartási költségek miatt korlátozottan alkalmazhatóak. Ebben a tanulmányban bemutatták a CF nyulak létrehozását a CRISPR/Cas9 technológia segítségével, amely viszonylag hosszabb élettartamú és megfizethetőbb fenntartási költségekkel rendelkező modellt kínál. A CF nyulak, amikor kizárólag orális ozmotikus hashajtókat kaptak, körülbelül 40 napos medián túlélést mutattak, és gyomor-bélrendszeri betegségben pusztultak el. A gasztrointesztinális tranzit helyreállítását célzó terápiás beavatkozások azonban a medián túlélést körülbelül 80 napra hosszabbították meg. A CF-nyulak egészségi állapotuk romlásával az exokrin hasnyálmirigy rendellenességekre utaló surrogátum markereket mutattak. A CFTR expressziós mintázata a WT nyulak légútjaiban nagyon hasonlított az emberére, széles körű eloszlást mutatott az orr légző- és szaglóhámában, valamint a proximális és distális alsó légutakban. Ezenkívül a CF nyulak az emberi CF-hez hasonló bioelektromos tulajdonságok rendellenességeit mutatták az orr- és légcsőhámában. Míg a fiatal CF nyulakban nem észleltek spontán légzőszervi betegséget, a túlélő 1 éves CF nyulakban a WT alomtestvérekhez képest abnormális fenotípusok jelentek meg, különösen az orr légzőszervi és szaglóhámjában (Xu et al. 2021).

Eun Jung Lee és munkatársai (Lee et al., 2021) CTGF csendesítést alkalmaztak nyulaknál. A glaukómaszűrő műtétek (GFS) során a túlzott sebgyógyulási reakciók kezelése kulcsfontosságú a műtéti siker szempontjából. A GFS-ben használt jelenlegi kiegészítő anyagok a sejtek nem szelektív gátlása miatt mellékhatásokat okozhatnak, ami ismételt műtétet tehet szükségessé. A CRISPR-Cas9 rendszer ígéretes megoldást jelent a glaukóma sebészetben, mivel nagy szelektivitást és tartós hatást biztosít. A kötőszöveti növekedési faktor (CTGF) a szem szöveti fibrózisának erős stimulátora. Ezért megvizsgálták a CTGF CRISPR-Cas9 rendszerrel történő elnyomásának hatását a fibrózisra GFS-ben. Adeno-asszociált vírus (AAV)-CRISPR-Cas9 rendszert alkalmazva sikeresen elnyomták a CTGF expresszióját fibroblasztokban in vitro, amit western blot elemzés és mélyszekvenálás is megerősített. Egy in vivo szemközi összehasonlító nyúl GFS modellben a CRISPR-CTGF-fel kezelt szemek szignifikánsan jobb túlélést mutattak a műtét helyén, csökkent szubkonjunktivális fibrózist, korlátozott kollagén lerakódást és csökkent sejtszintűséget a kezeletlen szemekhez képest. Ezek az eredmények a CRISPR-Cas9 által közvetített CTGF-

szuppresszió új megközelítését javasolják, amely potenciálisan javíthatja a humán GFS-eljárások eredményeit (Lee et al., 2021).

Dongshan Yang és kollégái (Yang et al., 2016) Knock in nyulakat állítottak elő. Annak érdekében, hogy a nyúlmodellek alkalmasságát növeljék a célzott gén-koppintás és a stabil gén-túlkifejezés céljára, a kutatók genomszekvencia-homológia elemzéssel azonosították az egér Rosa26 lokusz nyúl megfelelőjét. Valós idejű PCR és 5' RACE és 3' RACE kísérletek segítségével azonosították egy hosszú, nem kódoló RNS (lncRNS) két transzkriptváltozatát, amelyeket ez a lokusz kódol, és amelyeket rbRosaV1-nek és rbRosaV2-nek neveztek el. Ezek a változatok ubiquitikusan és következetesen expresszálódnak a különböző szövetekben. Ezt követően a nyúl Rosa26 (rbRosa26) lokuszát CRISPR/Cas9 segítségével célba vették, aminek eredményeképpen két knock-in nyúl vonalat hoztak létre (rbRosa26-EGFP és rbRosa26-Cre-reporter). Mindkét vonal valamennyi alapítója és utódai normális egészségi állapotot és szaporodási képességet mutattak. Az F1 generációban az rbRosa26-EGFP nyulak EGFP-t, az rbRosa26-Cre-reporter nyulak pedig tdTomatot fejeztek ki egységesen minden vizsgált szövetben. Fontos, hogy az rbRosa26 lokusz megszakítása nem volt káros hatással az állatok egészségére vagy szaporodására. Ez a kutatás az rbRosa26-ot nukleáz-közvetített géncélzásra alkalmas biztonságos kikötőnek tekinti, ami bővíti a nyulakon végzett transzgenikus kutatások eszköztárát. Az rbRosa26 hozzáadása várhatóan lehetővé teszi a különböző genetikai manipulációkat, beleértve a gain-of-function vizsgálatokat, a feltételes knockoutokat és a nyulakban végzett vonalkövetési vizsgálatokat (Yang et al., 2016).

Yuning Song és munkatársai (Song et al., 2016) a T373K SNP-vel dolgoztak. Az okulokután albinizmus (OCA) autoszomális recesszív rendellenességek egy csoportját foglalja magában, amelyeket csökkent melanintermelés jellemez, és a melanin szintézisében kulcsfontosságú enzimet, a tirozinázt (TYR) kódoló gén mutációi okoznak. Bár számos tanulmány és metaanalízis utal a TYR T373K egynukleotid-polimorfizmus (SNP) és az OCA1 közötti összefüggésre, az ezt alátámasztó biokémiai és genetikai bizonyítékok továbbra is korlátozottak. Ennek az összefüggésnek a vizsgálatára a TYR-WT és TYR-T373K mutánsokat overexpresszáltak HK293T sejtekben, és vizsgálták a melanintermelés és a tirozináz aktivitás változásait. Ezt követően TYR-K373T knock-in (KI) nyulakat hoztak létre egyszálú DNS mikroinjekciójával, és CRISPR/Cas9-homológia irányított javítási (HDR) technológia segítségével a C1118A lókuszt célzó szintetizált RNS-eket állítottak elő, ami lehetővé tette a melanin képződés megfigyelését. Eredményik azt mutatták, hogy a TYR T373K mutációja csökkentette a tirozináz aktivitását, ami sejtszinten csökkent

melaninszintézist eredményezett. Ezen túlmenően a génkezelt TYR-K373T nyulak helyreállt melanintermelést mutattak a szőrtüszőkben és az íriszben, amit a pigmentáció észrevehető csökkenése jelzett a TYR-T373K nyulakhoz képest. Az albinizmussal kapcsolatos T373K SNP-nek ez a funkcionális validálása állati szinten aláhúzza jelentőségét az OCA patogenezésében (Song et al., 2016).

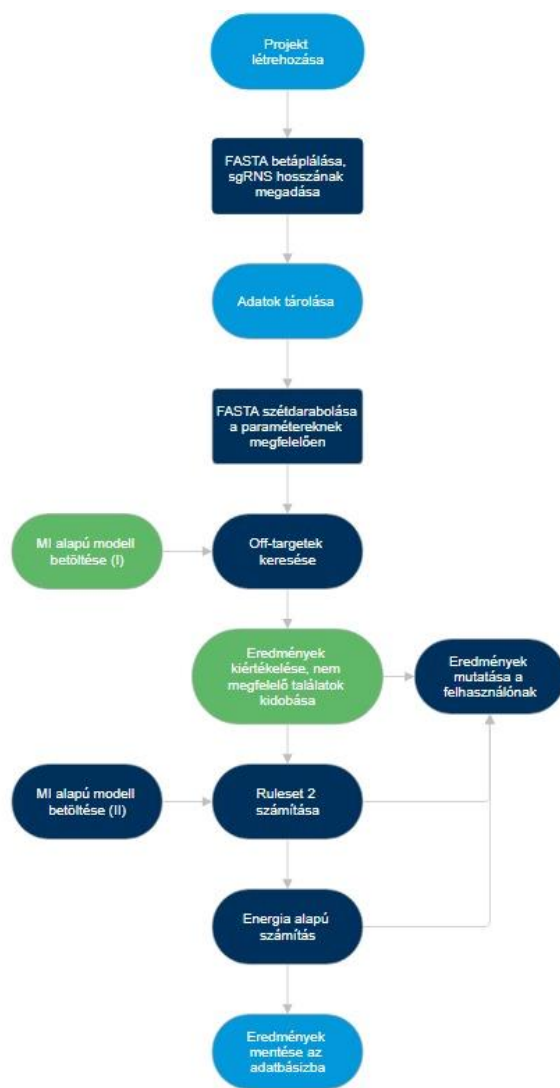
Weihua Jiang és kollégái (Jiang et al., 2018) a Wilson-kór kutatásához használható modell nyulat hoztak létre. Ebben a vizsgálatban a kutatók a CRISPR/Cas9 rendszerrel és egyszálú DNS-oligonukleotidokkal (ssODN-ok) kombinált CRISPR/Cas9-rendszerrel ATP7B-helyi irányított pontmutációjú nyulakat hoztak létre, hogy utánozzák a Wilson-kórral (WD) összefüggésbe hozható, ázsiaiaknál elterjedt mutációt (p. Arg778Leu). A precíziós pontmutáció hatékonysága elérte az 52,94%-ot, amikor a zigótákat 14 órával a HCG-kezelés után injektálták, jelentősen meghaladva a 19 órával a HCG-kezelés után injektált zigótáknál megfigyelt 14,29%-os hatékonyságot. A mutáns ATP7B allélt hordozó nyulak körülbelül három hónapos korukban elpusztultak. Ráadásul ezekben a nyulakban a máj rézsztintje a WD kezdetén kilencszeresére emelkedett, tükrözve a WD-ben szenvedő embereknél megfigyelt ötszörös emelkedést. Ezért a precíziós pontmutációk hatékonysága fokozódik, ha az RNS-eket a fejlődés korábbi szakaszaiban adják be a zigótákba. Az ATP7B mutáns nyulak a humán WD betegség ígéretes modelljét jelentik, amely lehetőséget kínál a patológiai elemzésre, a klinikai beavatkozásra és a génterápiás kutatásra (Jiang et al., 2018).

2. Alkalmazott módszerek

2.1. A Harmonio szoftver rendszerkövetelményei:

Windows 7 vagy újabb, 2GB RAM, 0 GB HDD, böngésző.

A szoftver működését az 5. Ábra szemlélteti. A böngésző követelmények a 11. Ábrán találhatók. A 6. Ábrán a Harmonio felülete látható.



5. ábra: A Harmonio szoftver működését ábrázoló workflow.

Sequencia	gcContent	selfComplementarity	Energo offset	offtargets	efficiency (MENICA)	Műveletek
ATATTACTGTCTTTGTCTCCAAG kód - posztó: 224	30%	6.7%	0.0945	415	66%	[Icons]
GTGATATCGTCATCTTCCAAG kód - posztó: 36	40%	26.09%	0.1057	122	64%	[Icons]
ACCCAAAGCTCTCCACTCCGGG kód - posztó: 243	40%	23.9%	0.2814	236	64%	[Icons]
ACCCAAAGCTCTCCACTCCGGG kód - posztó: 243	40%	23.9%	0.2814	236	64%	[Icons]
GTTCCCGGAGTGAGGAGCTTTG kód - posztó: 166	45%	6.7%	0.1054	99	63%	[Icons]
TATAGAGCACTTTTACCCAAG	30%	15.22%	0.0976	203	63%	[Icons]

6. ábra: A Harmonio project oldala.

2.2. *In silico* elemzés

Az *in silico* elemzés módszertana Ajay Prakash Uniyal és munkatársai tanulmányának vonatkozó részén alapszik. Kutatásukban két gént, az IDM2-t (AT1G54840) és az IDM3-at (AT1G20870) önkényesen választottak ki három kiválasztott számítási eszköz működésének vizsgálatára és összehasonlítására. Ezek a gének az Arabidopsis genomjának egyes kromoszómáján találhatók. Az IDM2 (AT1G54840), más néven a "DNS-metiláció növekedése 2" egy 1657 bázispáron átívelő, fehérjét kódoló gén. Az IDM2 1-3. exonját és promóter régióit használták fel sgRNS-ek tervezéséhez. Egy másik célpont, az IDM3 (AT1G20870 vagy "a DNS-metiláció növekedése 3") szintén egy 1674 bázispár hosszú fehérjekódoló gén. Az IDM3 esetében az 1. és 2. exont, valamint a promóter régiót használták az sgRNS-tervezéshez (Uniyal et al., 2019).

Én a véltelenszerű választás helyett a szakirodalmi feldolgozás részben említett tanulmányok által validált sgRNS-eket kerestem meg úgy, hogy a cikkekben említett géneket futtattam le nyúl genomra, illetve két szoftvert, a Harmonio-t és napjainkban az egyik legszélesebb körben használt ilyen programot, a CHOPCHOP v3-at (<https://chopchop.cbu.uib.no/>) teszteltem mely eredményeit össze tudtam hasonlítani. A vizsgált sgRNS-eket az 1. Táblázatban sorolom fel. Megjegyzendő, hogy az itt említett sgRNS hatékonyság alatt nem az RNS tervező szoftverek által prediktált hatékonyságot értem, hanem a konkrét kísérletben nyújtott hatékonyságot számoltam ki az injektált zigóták, és a mutáns utódok számából, egy esetben a cikkben szövegesen leírt „legjobb” minősítést vettem át mivel nem találtam más adatot. Az egyes sgRNS-ekhez tartozó Off-target adatok a 2. Táblázatban láthatók. A CHOPCHOP v3 eredményei a 6. Táblázatban találhatók.

2.3. Statisztikai elemzés

Összehasonlítottam a Harmonio által tervezett sgRNS-ek és az irodalomból származó sgRNS-ek előre jelzett hatékonyságát előjelpróbát alkalmazva.

Az előjelpróba olyan statisztikai módszer, amelyet a megfigyelési párok közötti konzisztens különbségek vizsgálatára használnak, gyakran akkor, amikor az adatok nem felelnek meg a parametrikus tesztekhez, például a t-próbához szükséges feltételezéseknek. Ez egy nem

parametrikus teszt, ami azt jelenti, hogy nem tesz feltevéseket az adatok mögöttes eloszlásáról.

A null- és alternatív hipotézisek megfogalmazásával kezdjük. A nullhipotézis jellemzően azt állítja, hogy nincs különbség a párosított megfigyelések között, míg az alternatív hipotézis azt állítja, hogy konzisztens különbség van.

Páros adatokat gyűjtünk, ahol minden pár két mérést vagy megfigyelést jelent, amelyeket ugyanattól az egyéntől vagy alanytól különböző körülmények, kezelések vagy időpontok között végeztek.

Minden egyes megfigyelési pár esetében kiszámítjuk a két érték közötti különbség előjelét. Ha az első érték nagyobb, mint a második, akkor pozitív előjelet adunk (+1); ha a második érték nagyobb, akkor negatív előjelet adunk (-1); ha az értékek egyenlőek, akkor a párt elvetjük.

Ezután megszámoljuk a pozitív és negatív előjelek számát az adatainkban.

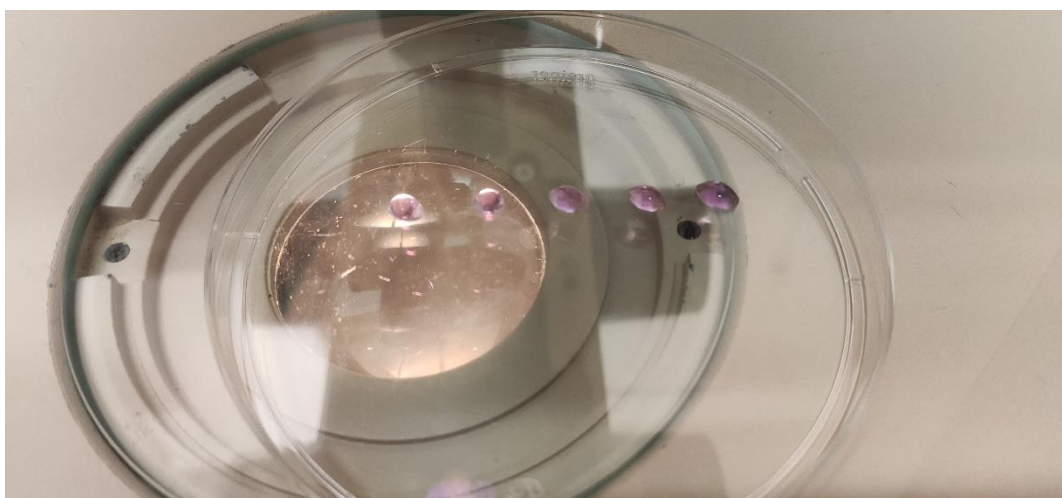
A binomiális eloszlás segítségével meghatározzuk, hogy a megfigyelt pozitív (vagy negatív) előjelek száma szignifikánsan eltér-e attól, ami a nullhipotézis szerint várható lenne. Ehhez ki kell számítani az adott számú előjel vagy több szélsőérték megfigyelésének valószínűségét, figyelembe véve a megfigyelések teljes számát és a nullhipotézis alapján várható arányt.

A teszt eredménye alapján vagy elutasítjuk, vagy megtartjuk a nullhipotézist. Ha a p-érték kisebb, mint az általunk választott szignifikancia szint (gyakran 0,05), akkor elutasítjuk a nullhipotézist, és arra következtetünk, hogy a párosított megfigyelések között konzisztens különbség van (Dixon-Mood, 2012).

Minden egyes sgRNS-hez párosítottam adatokat, ahol minden pár a szoftver által megjósolt hatékonyságból és az irodalomból kapott tényleges hatékonyságból áll. Ezt megismételtem a CHOPCHOP v3 hatékonyság adataival is. Az alfa szint 0,05. E mellett a statisztikai módszer mellett a normalitásvizsgálat (Shapiro-Wilk) és a Q-Q diagram eredményei, illetve az alacsony mintaszám miatt döntöttem. Az előjelpróba és a normalitás teszt az IBM SPSS Statistic 27 szoftverrel készült.

A 6. táblázatban található NOX5 gén specifikus gRNS-t megrendeltük (Sigma-Aldrich) és természetes pároztatással létrehozott egér embriókon előzetesen teszteltük. A hím és nőstény FVB/N törzsből származó egereket bigám párokban pároztattuk. Azon egerek esetén, ahol a termékenyítés megtörtént az embriókat a petevezetőből kimostam, a kumulusz sejteket

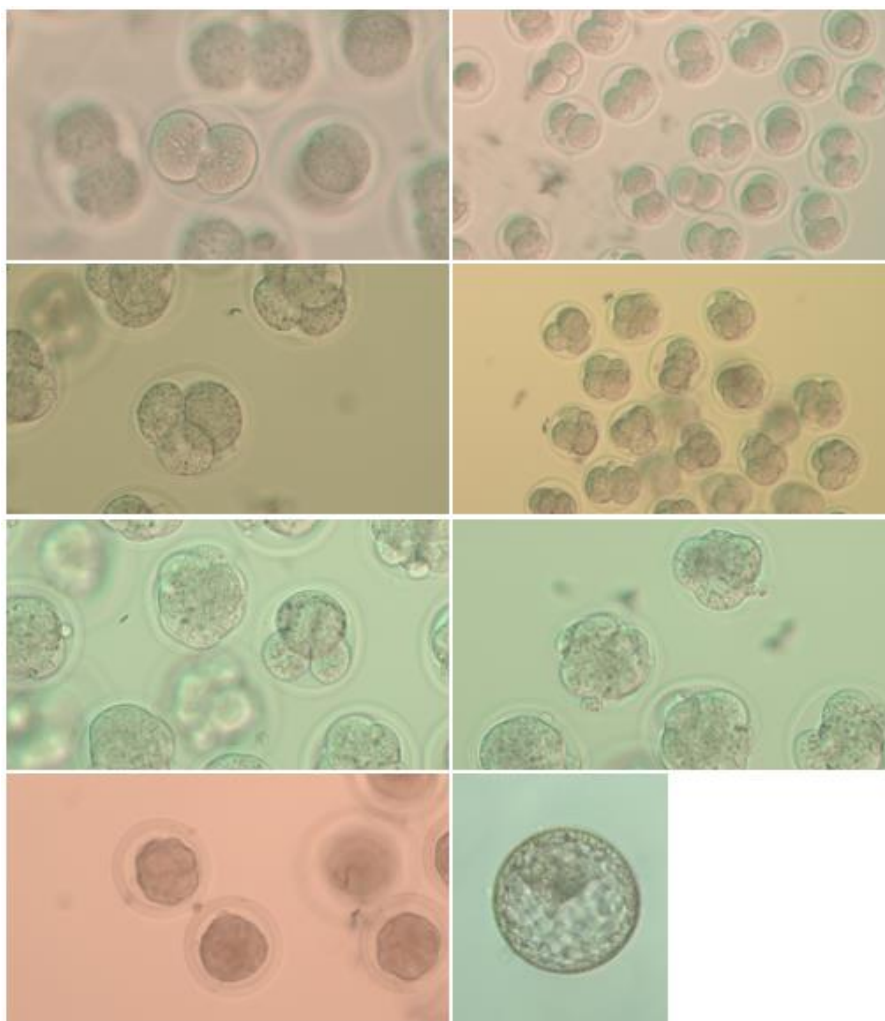
hyaluronidáz enzim segítségével 5 perc alatt 37°C-on leemésztettük. A termékenyült egysejtes embriókat öt darab M2 médium cseppen (7. ábra) átmosva megszabadítottam a sejttörmeléktől. A megtisztított embriókat 20ng/ul-es koncentrációjú NOX5 gén specifikus gRNS-t és 200 ng/ul Cas9 tartalmazó oldattal mikroinjektáltuk. Az embriókat ezután *in vitro* tenyésztettük KSOM médiumban (8. ábra) és fejlődésüket figyelemmel követtük (9. ábra). A blasztociszta stádiumig fejlődött embriókat DNS izoláló pufferbe összegyűjtöttük és mélyhűtve tároljuk (10. ábra). További kísérleteink célja lesz, hogy a blasztocisztákból DNS-t izoláljuk és PCR technika, valamint Sanger szekvenálás segítségével megállapítsuk a genomszerkesztés hatékonyságát.



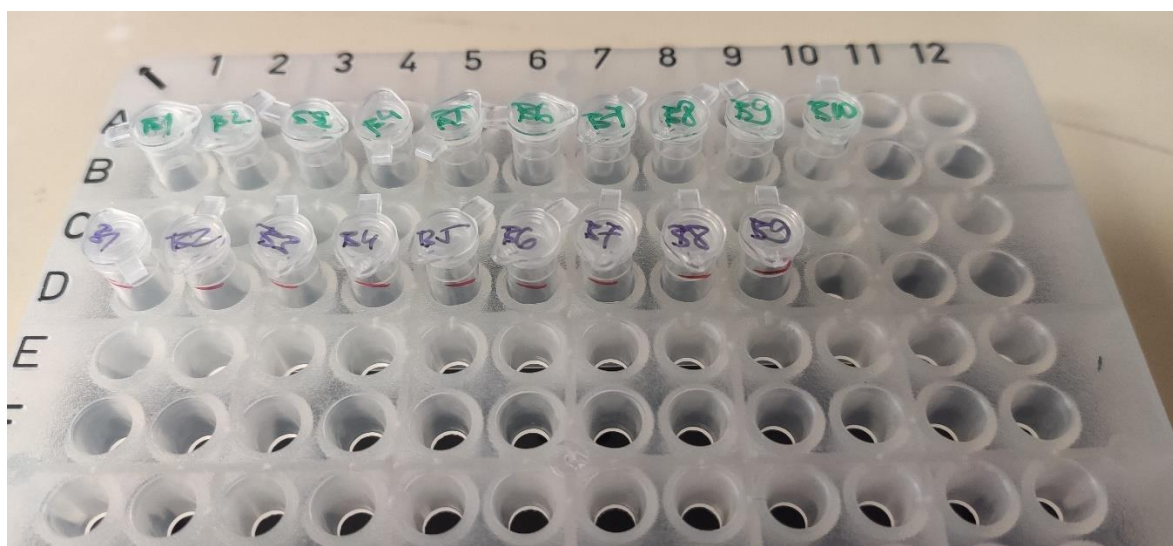
7. ábra: Az M2 médiumcseppek az embrió mosáshoz.



8. ábra: Tenyésztő petricsésze, KSOM médium cseppek olaj alatt.



9. ábra: Kétsejtes, négysejtes, nyolcsejtes, tizenhatsejtes, morula és blasztociszta stádiumú nyúl embriók.



10. ábra: A PCR csövekbe eltett blasztociszta minták.

A vizsgálat ezen részének sajnos a szakdolgozat leadási határ idejéig nem lett eredménye, az később várható.

3. Eredmények és értékelésük

3.1. *In silico* elemzés

A szakirodalmi fejezetben említett cikkekből származó sgRNS-eket és azok hatékonysági adatait az 1. Táblázatban soroltam fel, míg az Off-target elemzések adatait a 2. Táblázatban soroltam fel. Ezek után összehasonlítottam több módon a Harmonio ajánlásaival, melyeket a 3. Táblázat, 4. Táblázat és 5. Táblázat tartalmazza. Valamint a CHOPCHOP v3 által prediktált hatékonysági adatokat a 6. Táblázatban olvasható.

1. táblázat részlete: gRNS értékelés, a PAM pirossal kiemelve

Gén	sgRNS	Injektált zigóták	Élve született utódok	Mutáns utódok	Bi allélikus utódok	sgRNS hatékonysága
Nox5 Exon 3	GCTCATCCATGGGAGCCCTA TGG	n/a	n/a	n/a	n/a	„Legjobb”
MSTN Exon 1	GCTGATCGTGGCTGGCC CAGTGG	158	20	16	6	10,13%
MSTN Exon 1	GGCTGTGTAATGCATGCA CTTGG	158	20	16	6	10,13%
DMD Exon 51	CTTGGACAGAACTTACCGAG TGG	123	33	26	22	21,1%
DMD Exon 51	ATGGCATTCTAGTTCGGAG AGG	123	33	26	22	21,1%

sgRNS hatékonyság előrejelzés

Ahhoz, hogy megtudjam a Harmonio mennyivel működik megbízhatóbban, mint a jelenlegi szoftverek többsége, a szakirodalomban talált gének megfelelő exonjaira vagy ha arra volt a szükség intronjaira lefuttattam egy-egy keresést, melynek eredményeit a 3. Táblázat foglalja össze. Ebből látható, hogy a Harmonio jóval magasabb hatékonysági értékeket prediktál, mint amik a kísérleti adatokból látszanak, bár így is más sgRNS-eket számolt leghatékonyabbnak az egyes génekre, mint amiket a szakirodalomban használtak. A jelenleg kicsi számítási kapacitás miatt, csupán a legjobb 30 találatot tudja kilistázni mert a memória ennyit tesz lehetővé. Sajnos voltak olyan szakirodalmi sgRNS-ek amik hatékonyságát vagy az adott cikk nem írta le számszerű adatokkal, vagy az imént említett okokból nem találta meg a tesztelés alatt lévő szoftver, mert olyan alacsony pontot adott nekik. Ezeket a statisztikai vizsgálatnál nem tudtam felhasználni.

Megjegyzendő, hogy a 4. Táblázat adatai alapján a sgRNS-ek hatékonyságának túlbecslése a CHOCHOP v3-nak is hibája, illetve összehasonlítva itt a túlbecslés mértéke jóval nagyobb, mint a Harmonio esetében.

3. táblázat részlete: A szakirodalmi sgRNS-ek összehasonlítása a Harmonio eredményeivel. Gén: A szakirodalomban felsorolt gének, sgRNS: a szakirodalomban alkalmazott sgRNS, sgRNS hatékonyság: a cikkek vonatkozó részeiből (mikroinjektált zigóták száma, mutáns élve született utódok száma) számított érték, Harmonio hatékonyság: A Harmonio által prediktált hatékonysága a szakirodalmi sgRNS-eknek, ahol „nincs találat” szerepel az azt jelenti, hogy az említett RNS nem tartozik a leghatékonyabb 30 közé. Harmonio sgRNS: A Harmonio által a leghatékonyabbnak ítélt RNS. A piros betűk a PAM-ot jelzik.

Gén	sgRNS	sgRNS hatékonysága	Harmonio hatékonyság	Harmonio sgRNS
Nox5 Exon 3	GCTCATCCATGGGAGCCCTA TGG	„Legjobb”	50%	GGAA CTTGAGTTTGTCCATAGGG
MSTN Exon 1	GCTGATCGTGGCTGGCCAG TGG	10,13%	31%	ATATTACTGTTTTGTCTC CAAG
MSTN Exon 1	GGCTGTGTAATGCATGCACT TGG	10,13%	32%	ATATTACTGTTTTGTCTC CAAG
DMD Exon 51	CTTGGACAGAACTTACCGAG TGG	21,1%	Nincs találat	CCTG CTCTGGCAGATTCAACCG
DMD Exon 51	ATGGCATTCTAGTTCGGAG AGG	21,1%	31%	CCTG CTCTGGCAGATTCAACCG

Off-target előrejelzés

A megbízhatóság másik fontos kritériuma, hogy a Harmonio milyen pontossággal határozza meg egy-egy sgRNS potenciális Off-target-eit. Ezeket az adatokat a 5. és 6. Táblázat tartalmazza.

5. táblázat részlete: Az off-target adatok összevetése a Harmonio adataival. Gén: A szakirodalomban felsorolt gének, sgRNS: a szakirodalomban alkalmazott sgRNS, Potenciális Off-target: A vonatkozó cikkben említett lehetséges Off-target-ek száma, Off-target: A ténylegesen bekövetkezett, kimutatott és a vonatkozó cikkben leközölt off-target-ek száma, Harmonio potenciális Off-target: A vonatkozó cikkben szereplő sgRNS Harmonio által prediktált összes potenciális Off-target-ei. A piros betűk a PAM-ot jelzik.

Gén	sgRNS	Potenciális Off-target	Off-target	Harmonio poteciális Off-target
Nox5 Exon 3	GCTCATCCATGGGAGCCCTA TGG	3	0	104
MSTN Exon 1	GCTGATCGTGGCTGGCCAG TGG	5	0	180

MSTN Exon 1	GGCTGTGTAATGCATGCACTTGG	5	0	160
DMD Exon 51	CTTGGACAGAACTTACCGAGTGG	10	0	120
DMD Exon 51	ATGGCATTCTAGTTCGGAGAGG	10	0	96

A vonatkozó táblázatok alapján kijelenthetjük, hogy a szoftver nagy érzékenységet mutatott a potenciális off-target-ek azonosításában, bár a megjósolt hibák nagy száma finomításra szoruló területeket jelezhetnek.

3.2. Statisztikai elemzés

Harmonio előjelpróba:

Pozitív előjelek száma: 16 olyan eset van, amikor az előre jelzett hatékonyság meghaladja a megfigyelt hatékonyságot. Ez azt jelenti, hogy a szoftver szinte minden sgRNS esetében nagyobb hatékonyságot jósolt, mint amit megfigyeltek.

Negatív előjelűek száma: 0 olyan eset van, amikor a megfigyelt hatékonyság meghaladja az előre jelzett hatékonyságot.

Teljes szám: A nem kötő párok teljes száma 16. Úgy tűnik, hogy egy párt kizártunk az elemzésből, mivel az döntetlen volt, így 16 pár maradt a teszteléshez.

Z-érték: A 4-es z-érték a nullhipotézis alapján várható értéktől való szignifikáns eltérést jelzi. Konkrétan azt sugallja, hogy a megfigyelt eredmény (minden pozitív különbség) nagy valószínűséggel nem véletlenszerűen következett be, az sgRNS-hatékonysági előrejelzések irányában nem lenne valódi különbség.

P-érték: p-értékével (0,00006) az eredmény statisztikailag rendkívül szignifikáns, messze a hagyományos 0,05-ös alfa-szint alatt van. Ez azt jelzi, hogy nagyon kicsi a valószínűsége annak, hogy a pozitív vs. negatív különbségek ilyen ferde eloszlását figyeljük meg, ha a nullhipotézis igaz lenne.

Az előjelpróba eredményei határozottan arra utalnak, hogy a validálandó szoftver szisztematikusan magasabb sgRNS-hatékonyságot jósol, mint amit az irodalomban megfigyeltek. Mivel a 16 nem-kötőjeles párból 16 pozitív előjelet és szignifikáns p-értéket mutat, arra következtethetünk, hogy a szoftver előrejelzései statisztikailag szignifikánsan torzítanak az sgRNS-hatékonyságok túlbecslése felé.

CHOPCHOP v3 előjelpróba:

Pozitív előjelek száma: Mind a 22 összehasonlítás pozitív különbséget mutat, ami azt jelzi, hogy minden egyes sgRNS esetében a validált szoftver által előre jelzett hatékonyság magasabb, mint az irodalomban közölt hatékonyság.

Negatív előjelek száma: Nincs olyan eset, amikor az irodalomban megfigyelt hatékonyság meghaladja a jósolt hatékonyságot, így a negatív előjelek száma nulla.

Teljes szám: Az összes elemzett pár száma 22, ami azt jelenti, hogy nem volt döntetlen (olyan párok, ahol az előre jelzett és a megfigyelt hatékonyság pontosan megegyezik).

Z-érték: A 4,69042-es z-érték szignifikánsan magas, ami azt jelzi, hogy a nullhipotézis alapján (miszerint a pozitív és negatív különbségek egyenlő eséllyel fordulnak elő) várható értéktől erősen eltér. Ez a magas z-érték az előre jelzett és a megfigyelt hatékonyságok közötti különbségek irányának következetességét tükrözi.

P-érték: A 0,00001-nél kisebb p-értékkel a teszteredmény statisztikailag igen szignifikáns. Ez a rendkívül kis p-érték határozottan arra utal, hogy a megfigyelt mintázat (minden pozitív különbség) nem a véletlen műve.

Statisztikai szignifikancia: Az eredmény $<0,05$ -nél szignifikáns.

$p < 0,05$, ami jelentősen meghaladja a statisztikai szignifikancia hagyományos küszöbértékeit. Ez azt jelenti, hogy magabiztosan elvethetjük a nullhipotézist, és arra a következtetésre juthatunk, hogy jelentős torzítás van abban, hogy a megjósolt hatékonyságok magasabbak, mint a szakirodalmiak.

A előjelpróba egyértelműen bizonyítja, hogy a CHOPCHOP v3 mind a 22 elemzett sgRNS-pár esetében következetesen túlbecsüli az sgRNS hatékonyságát a szakirodalomban dokumentáltakhoz képest. Ez a következetes túlbecslés statisztikailag szignifikáns, ami szisztematikus torzítást jelez a szoftver hatékonysági előrejelzéseiben.

Harmonio-CHOPCHOP v3. előjelpróba:

Ahhoz hogy eldönthessem a Harmonio vagy a CHOPCHOP v3 képes pontosabb predikciókra meg kellett vizsgálnom hogy melyiknek jobb a hatékonyság becslése. Ezt azon a 16 sgRNS-

en tudtam elvégezni melyre a Harmonio és a CHOPCHOP v3 is adott eredményt, úgy hogy kiszámítottam a szakirodalmi hatékonyság és az előre jelzett hatékonyság közötti abszolút különbséget mindkét szoftver esetében, így megkaptam az abszolút különbségek két csoportját, amelyek az egyes sgRNS-ek túlbecslésének nagyságát jelentik. Ezek után a kiszámított abszolút különbségeket az új összehasonlítási adatállományként használtam.

Átlagok: Az átlagos abszolút különbség a Harmonio esetében 15,5, ami alacsonyabb átlagos túlbecslést jelez, míg a CHOPCHOP v3 esetében ez az érték jelentősen magasabb, 54,62.

Pozitív előjelek száma: 0 olyan eset van, amikor a CHOPCHOP v3 túlbecslése kisebb volt, mint a Harmonio túlbecslése.

Negatív előjelek száma: 16 olyan eset van (16-ból), amikor a Harmonio túlbecslése kisebb volt, mint a CHOPCHOP v3-é.

Összesen: 16 pár került elemzésre.

Z-pontszám: A számítás szerint az értéke 4, ami a nullhipotézis elvárásától való jelentős eltérést jelzi.

P-érték: A .00006-os p-érték rendkívül szignifikáns, ami határozottan arra utal, hogy az előjelek megfigyelt eloszlása (mind negatív) nem a véletlen műve.

Az előjelpróba eredményei statisztikailag szignifikáns különbséget mutatnak a túlbecslés nagyságrendjében a két szoftver között. A Harmonio következetesen kisebb mértékű túlbecslést mutat a már CHOPCHOP v3-hoz képest az összes összehasonlított sgRNS esetében.

Az abszolút különbségek átlagértékei (15,5 vs. 54,62) aláhúzzák a szoftvereszközök közötti jelentős eltérést a túlbecslésben. A CHOPCHOP v3 sokkal nagyobb mértékben becsüli túl az sgRNS hatékonyságát, mint a Harmonio.

A magas z-érték és a rendkívül alacsony p-érték megerősíti az ezen eredményekbe vetett bizalmat, bizonyítva, hogy a két szoftver közötti túlbecslés nagyságrendje közötti különbség statisztikailag szignifikáns, és nem a véletlen szórás eredménye.

4. Következtetések és javaslatok

A célkitűzéseknek megfelelően megvizsgáltuk a szoftvert, és összevetettük a CHOPCHOP v3.-al. Vizsgálataim kiterjedtek a sgRNS-ek hatékonyságának, és az off target predikciós képesség *in silico* elemzésére, illetve statisztikai elemzést is végeztem.

Az *in silico* vizsgálat eredményei arra engednek következtetni, hogy a Harmonio bár túl nagy hatékonyság értékeket ad az egyes sgRNS-eknek, mégis megállja a helyét olyan széleskörben használt szoftverek mellett, mint a CHOPCHOP v3. Ez a megállapítás fontos az sgRNS-tervezési szoftver felhasználói számára, mivel rávilágít az sgRNS-ek előre jelzett hatékonyságának következetes túlbecslésére. Ez a torzítás befolyásolhatja a kísérleti tervezést és a várható eredményeket, ami aláhúzza a szoftver előrejelző algoritmusának kalibrálásának és kiigazításának szükségességét, hogy az jobban igazodjon a megfigyelt kísérleti eredményekhez.

A jövőben mind az egyes számításokhoz szükséges algoritmusok tökéletesítésére mind az MI algoritmus további fejlesztésére kell törekedni. Ezenfelül hosszútávú cél, hogy a Harmonio ne csak nyúl genomra legyen alkalmazható, hanem a modell organizmusok, vagy egynap az agárirumban vagy gyógyászatban is alkalmazható legyen génmódosított organizmusok vagy génterápiák esetén.

5. Összefoglalás

Az emberiség egyre több globális problémával néz szembe, mint a klímaváltozás, és a túlnépesedés. Előbbi következtében a jelenleg termelésbevont hagyományos fajták termelőképesége csökken, ám a termékekre való igény az emberi népességének növekedésével folyamatosan nő. Megoldást jelenthet a problémára a modern biotechnológiai módszerek alkalmazása, elsősorban a fejlett genomszerkesztés módszerek. A gazdaságban való megjelenésük még is lassan következik be, ennek oka jelentős mértékben ezen technológiák alacsony hatékonysága, és ebből fakadó költségességük. A kutatók folyamatosan dolgoznak az új módszerek tökéletesítésén. Ennek a CRISPR esetében két fő ága a Cas molekula további módosítása, valamint a jobb gRNS-ek tulajdonságainak keresése. Az általunk fejlesztett Harmonio egy olyan sgRNS tervező szoftver, mely a jelenleg elérhető legkorszerűbb értékelő rendszerrel dolgozik, mesterséges intelligenciával megtámogatva. Jelen kutatás ennek hatástanulmánya. Ennek részeként korábbi szakirodalomból származó nyúl genomszerkesztésnél felhasznált sgRNS-ek hatékonyságát vizsgáltam, azok tényleges hatékonysága, illetve a Harmonio, és a CHOPCHOP v3 által előrejelzett hatékonyság tekintetében. A statisztikai elemzést előjel próbával végeztem, mely alkalmas arra, hogy konzistens különbségeket mutasson ki.

Az elemzés egyértelműen bizonyítja, hogy bár mindkét szoftver túlbecsüli az sgRNS hatékonyságát, a Harmonio ezt lényegesen kisebb mértékben teszi, mint a CHOPCHOP v3. Ez arra utal, hogy a Harmonio pontosabb és megbízhatóbb előrejelzést adhat az sgRNS-hatékonyságra vonatkozóan, így potenciálisan hasznosabb lehet olyan alkalmazásokban, ahol a túlbecslés minimalizálása kulcsfontosságú. A túlbecslés nagyságrendjében mutatkozó jelentős különbség, amelyet az átlagok abszolút különbségei kiemelnek, és amelyet az előjelpróba is alátámaszt, a Harmonio egyértelmű előnyét jelzi az sgRNS-hatékonyság előrejelzési pontossága tekintetében.

6. Irodalomjegyzék

- Aktas Ö., Dogan E., Ensari T., "CRISPR/CAS9 Target Prediction with Deep Learning," 2019 *Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*, Istanbul, Turkey, 2019, pp. 1-5, doi: 10.1109/EBBT.2019.8741648.
- Ameen S. Z., Ozsoz M., Mubarak A. S., Turjman F. A., Serte S. (2021): C-SVR Crispr: Prediction of CRISPR/Cas12 guideRNA activity using deep learning models. *Alexandria Engineering Journal*, Volume 60, Issue 4, Pages 3501-3508, ISSN 1110-0168, <https://doi.org/10.1016/j.aej.2021.02.007>.
- Amee, S.Z., Mubarak, A.S., Süleyman, A., Mehmet, O. (2020). Development of CNN Model for Prediction of CRISPR/Cas12 Guide RNA Activity. In: Aliev, R., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M., Sadikoglu, F. (eds) 10th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions - ICSCCW-2019. ICSCCW 2019. Advances in Intelligent Systems and Computing, vol 1095. Springer, Cham. https://doi.org/10.1007/978-3-030-35249-3_90
- Anderson R., Bayer P. E., Edwards D. (2020): Climate change and the need for agricultural adaptation. *Current Opinion in Plant Biology*, Volume 56, Pages 197-202, ISSN 1369-5266, <https://doi.org/10.1016/j.pbi.2019.12.006>.
- Aydinalp C., Cresser M. (2008): The Effects of Global Climate Change on Agriculture. *American-Eurasian J. Agric. Environ. Sci.* 3.
- Bae S, Park J, Kim JS. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*. 2014 May 15;30(10):1473-5. doi: 10.1093/bioinformatics/btu048. Epub 2014 Jan 24. PMID: 24463181; PMCID: PMC4016707.
- Becker S., Boch J. (2021): TALE and TALEN genome editing technologies. *Gene and Genome Editing*, Volume 2, 100007, ISSN 2666-3880, <https://doi.org/10.1016/j.ggedit.2021.100007>.
- BfR (2020): Use of Laboratory Animals in 2020. BfR honlapja. Letöltés dátuma: 2024. 04. 13. forrás: https://www.bfr.de/en/use_of_laboratory_animals_in_2020-291690.html
- Bösze Z, Major P, Baczkó I, Odening KE, Bodrogi L, Hiripi L, Varró A. The potential impact of new generation transgenic methods on creating rabbit models of cardiac diseases. *Prog Biophys Mol Biol*. 2016 Jul;121(2):123-30. doi: 10.1016/j.pbiomolbio.2016.05.007. Epub 2016 May 19. PMID: 27210304.
- Bruno M., Mahgoub M., Macfarlan T. S. (2019): The Arms Race Between KRAB–Zinc Finger Proteins and Endogenous Retroelements and Its Impact on Mammals. *ANNUAL REVIEW OF GENETICS*, Volume 53, <https://doi.org/10.1146/annurev-genet-112618-043717>
- Cancellieri S., Canver M. C., Bombieri N., Giugno R., Pinello L., CRISPRitz: rapid, high-throughput and variant-aware *in silico* off-target site identification for CRISPR genome editing, *Bioinformatics*, Volume 36, Issue 7, April 2020, Pages 2001–2008, <https://doi.org/10.1093/bioinformatics/btz867>
- Cui Y, Xu J, Cheng M, Liao X, Peng S. Review of CRISPR/Cas9 sgRNA Design Tools. *Interdiscip Sci*. 2018 Jun;10(2):455-465. doi: 10.1007/s12539-018-0298-z. Epub 2018 Apr 11. PMID: 29644494.
- Chen L. Wang S.; Zhang Y.; Li J.; Xing Z.; Yang J; Huang T.; Cai Y. (2017): "Identify Key Sequence Features to Improve CRISPR sgRNA Efficacy," in *IEEE Access*, vol. 5, pp. 26582-26590, doi: 10.1109/ACCESS.2017.2775703.
- Chuai G., Wang Q., Liu Q., (2017): *In Silico* Meets In Vivo: Towards Computational CRISPR-Based sgRNA Design, *Trends in Biotechnology*, Volume 35, Issue 1, Pages 12-21, ISSN 0167-7799, <https://doi.org/10.1016/j.tibtech.2016.06.008>.
- Dixon, W. J., & Mood, A. M. (1946). The Statistical Sign Test. *Journal of the American Statistical Association*, 41(236), 557–566. <https://doi.org/10.1080/01621459.1946.10501898>
- Doench, J., Hartenian, E., Graham, D. et al. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* 32, 1262–1267 (2014). <https://doi.org/10.1038/nbt.3026>

- Georges M. (2001): Recent progress in livestock genomics and potential impact on breeding programs. *Theriogenology*, Volume 55, Issue 1, Pages 15-21, ISSN 0093-691X, [https://doi.org/10.1016/S0093-691X\(00\)00442-8](https://doi.org/10.1016/S0093-691X(00)00442-8).
- Gkazi SA. Quantifying CRISPR off-target effects. *Emerg Top Life Sci.* 2019 May 31;3(3):327-334. doi: 10.1042/ETLS20180146. PMID: 33523136.
- Guo, R., Wan, Y., Xu, D. et al. Generation and evaluation of Myostatin knock-out rabbits and goats using CRISPR/Cas9 system. *Sci Rep* 6, 29855 (2016). <https://doi.org/10.1038/srep29855>
- Hillary, V.E., Ceasar, S.A. A Review on the Mechanism and Applications of CRISPR/Cas9/Cas12/Cas13/Cas14 Proteins Utilized for Genome Engineering. *Mol Biotechnol* 65, 311–325 (2023). <https://doi.org/10.1007/s12033-022-00567-0>
- Honda A, Ogura A. Rabbit models for biomedical research revisited via genome editing approaches. *J Reprod Dev.* 2017 Oct 18;63(5):435-438. doi: 10.1262/jrd.2017-053. Epub 2017 Jun 2. PMID: 28579598; PMCID: PMC5649091.
- Hryhorowicz M, Lipiński D, Zeyland J, Słomski R. CRISPR/Cas9 Immune System as a Tool for Genome Engineering. *Arch Immunol Ther Exp (Warsz).* 2017 Jun;65(3):233-240. doi: 10.1007/s00005-016-0427-5. Epub 2016 Oct 3. PMID: 27699445; PMCID: PMC5434172.
- Hyams G., Abadi S., Lahav S., Avni A., Halperin E., Shani E., Mayrose I. CRISPyS: Optimal sgRNA Design for Editing Multiple Members of a Gene Family Using the CRISPR System *J. Mol. Biol.*, 430 (2018), pp. 2184-2195
- Hyun I., Los Angeles A. D. (2019): *Chimerea Research (Methods and Protocols)*. Humana New York, NY, 234 p. DOI: <https://doi.org/10.1007/978-1-4939-9524-0>
- Jackson S. A., MCKENZIE R. E., FAGERLUND R. D., KIEPER S. N., FINERAN P. C., BROUNS S. J. J. (2017): CRISPR-Cas: Adapting to change. *Science* 356, eaal5056. DOI:10.1126/science.aal5056
- Jiang, W., Liu, L., Chang, Q. et al. Production of Wilson Disease Model Rabbits with Homology-Directed Precision Point Mutations in the ATP7B Gene Using the CRISPR/Cas9 System. *Sci Rep* 8, 1332 (2018). <https://doi.org/10.1038/s41598-018-19774-4>
- Kanca O., Zirin J., Garcia-Marques J., Knight S. M., Yang-Zhou D., Amador G., Chung H., Zuo Z., Ma L., He Y., Lin W., Fang Y., Ge M., Yamamoto S., Schulze K. L., Hu Y., Spradling A C., Mohr S E., Perrimon N., Bellen H. J (2019): An efficient CRISPR-based strategy to insert small and large fragments of DNA using short homology arms. *eLife* 8:e51539. <https://doi.org/10.7554/eLife.51539>
- Karlapudi A. P., Venkateswarulu T.C, Tammineedi J., Srirama K., Kanumuri L. Kodali V.P. (2018): *In silico* sgRNA tool design for CRISPR control of quorum sensing in *Acinetobacter* species. *Genes & Diseases*, Volume 5, Issue 2, June 2018, Pages 123-129, <https://doi.org/10.1016/j.gendis.2018.03.004>
- Kleinstiver, B., Prew, M., Tsai, S. et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485 (2015). <https://doi.org/10.1038/nature14592>
- Labun K., Montague T. G., Krause M., Cleuren Y. N. T., Tjeldnes H., Valen E., CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing, *Nucleic Acids Research*, Volume 47, Issue W1, 02 July 2019, Pages W171–W174, <https://doi.org/10.1093/nar/gkz365>
- Lau CH, Tin C, Suh Y. CRISPR-based strategies for targeted transgene knock-in and gene correction. *Fac Rev.* 2020 Dec 4;9:20. doi: 10.12703/r/9-20. PMID: 33659952; PMCID: PMC7886068.
- Lee, E.J., Han, J.C., Park, D.Y. et al. Effect of connective tissue growth factor gene editing using adeno-associated virus-mediated CRISPR–Cas9 on rabbit glaucoma filtering surgery outcomes. *Gene Ther* 28, 277–286 (2021). <https://doi.org/10.1038/s41434-020-0166-4>
- Liu G., Zhang Y., Zhang T. (2020): Computational approaches for effective CRISPR guide RNA design and evaluation, *Computational and Structural Biotechnology Journal*, Volume 18, Pages 35-44, ISSN 2001-0370, <https://doi.org/10.1016/j.csbj.2019.11.006>.
- Louie W, Shen MW, Tahiry Z, Zhang S, Worstell D, Cassa CA, et al. (2021) Machine learning based CRISPR gRNA design for therapeutic exon skipping. *PLoS Comput Biol* 17(1): e1008605. <https://doi.org/10.1371/journal.pcbi.1008605>
- Luthra R., Kaur S., Bhandari K. (2021): Applications of CRISPR as a potential therapeutic. *Life Sciences*, 284, 9 p. DOI: 10.1016/j.lfs.2021.119908

- Lv, Q., Yuan, L., Deng, J. et al. Efficient Generation of Myostatin Gene Mutated Rabbit by CRISPR/Cas9. *Sci Rep* 6, 25029 (2016). <https://doi.org/10.1038/srep25029>
- Lynas M., Houlton B. Z., Perry S. (2021): Greater than 99% consensus on human caused climate change in the peer-reviewed scientific literature *Environ. Res. Lett.* 16 114005
- Mak, J.K., Störtz, F. & Minary, P. Comprehensive computational analysis of epigenetic descriptors affecting CRISPR-Cas9 off-target activity. *BMC Genomics* 23, 805 (2022). <https://doi.org/10.1186/s12864-022-09012-7>
- Mapara M, Thomas BS, Bhat KM. Rabbit as an animal model for experimental research. *Dent Res J (Isfahan)*. 2012 Jan;9(1):111-8. doi: 10.4103/1735-3327.92960. PMID: 22363373; PMCID: PMC3283968.
- Martin, F.; Sánchez-Hernández, S.; Gutiérrez-Guerrero, A.; Pinedo-Gomez, J.; Benabdellah, K. Biased and Unbiased Methods for the Detection of Off-Target Cleavage by CRISPR/Cas9: An Overview. *Int. J. Mol. Sci.* 2016, 17, 1507. <https://doi.org/10.3390/ijms17091507>
- Moreno-Mateos MA, Vejnar CE, Beaudoin JD, Fernandez JP, Mis EK, Khokha MK, Giraldez AJ. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods*. 2015 Oct;12(10):982-8. doi: 10.1038/nmeth.3543. Epub 2015 Aug 31. PMID: 26322839; PMCID: PMC4589495.
- Park J., Bae S., and Kim J.-S. Cas-Designer: A web-based tool for choice of CRISPR-Cas9 target sites. *Bioinformatics* 31, 4014-4016 (2015).
- Peng R., Lin G. and Li J. (2016), Potential pitfalls of CRISPR/Cas9-mediated genome editing. *FEBS J*, 283: 1218-1231. <https://doi.org/10.1111/febs.13586>
- Petheő G. L., Kerekes A., Mihálffy M., Donkó Á., Bodrogi L., Skoda G., Baráth M., Hoffmann O. I., Szeles Z. Balázs B., Sirokmány G., Fábíán J. R., Tóth Z. E., Baksa I., Kacs Kovics I., Hunyady L., Hiripi L., Bösze Z., Geiszt M. (2021): Disruption of the NOX5 Gene Aggravates Atherosclerosis in Rabbits. *Circulation Research*, 128(9), 1320-1322 p. doi: 10.1161/CIRCRESAHA.120.318611
- Rojas-Downing M. M, Nejadhashemi A. P., Harrigan T., Woznicki S. A. (2017): Climate change and livestock: Impacts, adaptation, and mitigation. *Climate Risk Management*, Volume 16, Pages 145-163, ISSN 2212-0963, <https://doi.org/10.1016/j.crm.2017.02.001>.
- Sapoval, N., Aghazadeh, A., Nute, M.G. et al. Current progress and open challenges for applying deep learning across the biosciences. *Nat Commun* 13, 1728 (2022). <https://doi.org/10.1038/s41467-022-29268-7>
- Sengupta P, Dutta S. Mapping the Age of Laboratory Rabbit Strains to Human. *Int J Prev Med*. 2020 Dec 11;11:194. doi: 10.4103/ijpvm.IJPVM_530_18. PMID: 33815718; PMCID: PMC8000163.
- Soares J., Pinheiro A., Esteves P. J. (2022): The rabbit as an animal model to study innate immunity genes: Is it better than mice?. *Frontiers in Immunology*, 13, DOI=10.3389/fimmu.2022.981815, ISSN=1664-3224
- Song, Y., Yuan, L., Wang, Y. et al. Efficient dual sgRNA-directed large gene deletion in rabbit with CRISPR/Cas9 system. *Cell. Mol. Life Sci.* 73, 2959–2968 (2016). <https://doi.org/10.1007/s00018-016-2143-z>
- Song Y., Zhang Y., Chen M., Deng J., Sui T., Lai L. (2018): Functional validation of the albinism-associated tyrosinase T373K SNP by CRISPR/Cas9-mediated homology-directed repair (HDR) in rabbits. *eBioMedicine*, VOLUME 36, P517-525, DOI:<https://doi.org/10.1016/j.ebiom.2018.09.041>
- Störtz F., Minary P., crisprSQL: a novel database platform for CRISPR/Cas off-target cleavage assays, *Nucleic Acids Research*, Volume 49, Issue D1, 8 January 2021, Pages D855–D861, <https://doi.org/10.1093/nar/gkaa885>
- Sui T., Lau Y. S., Liu D., Liu T., Xu L., Gao Y., Lai L., Li Z., Han R.; A novel rabbit model of Duchenne muscular dystrophy generated by CRISPR/Cas9. *Dis Model Mech* 1 June 2018; 11 (6): dmm032201. doi: <https://doi.org/10.1242/dmm.032201>
- Sui T., Yuan L., Liu H., Chen M., Deng J., Wang Y., Li Z., Lai L., CRISPR/Cas9-mediated mutation of *PHEX* in rabbit recapitulates human X-linked hypophosphatemia

(XLH), *Human Molecular Genetics*, Volume 25, Issue 13, 1 July 2016, Pages 2661–2671, <https://doi.org/10.1093/hmg/ddw125>

- Sui, T., Xu, L., Lau, Y.S. et al. Development of muscular dystrophy in a CRISPR-engineered mutant rabbit model with frame-disrupting ANO5 mutations. *Cell Death Dis* 9, 609 (2018). <https://doi.org/10.1038/s41419-018-0674-y>
- Sun W, Liu H, Yin W, Qiao J, Zhao X, Liu Y. Strategies for Enhancing the Homology-Directed Repair Efficiency of CRISPR-Cas Systems. *CRISPR J.* 2022 Feb;5(1):7-18. doi: 10.1089/crispr.2021.0039. Epub 2022 Jan 24. PMID: 35076280.
- Tännsjö T., Ryberg J. (2004): *The Repugnant Conclusion*. Springer Netherlands, Dordrecht Library Of Ethics And Applied Philosophy 15, DOI: 10.1007/978-1-4020-2473-3
- Uniyal AP, Mansotra K, Yadav SK, Kumar V. An overview of designing and selection of sgRNAs for precise genome editing by the CRISPR-Cas9 system in plants. *3 Biotech.* 2019 Jun;9(6):223. doi: 10.1007/s13205-019-1760-2. Epub 2019 May 21. PMID: 31139538; PMCID: PMC6529479.
- Vasileva, E., Shuvalov, O., Garabadiu, A. et al. Genome-editing tools for stem cell biology. *Cell Death Dis* 6, e1831 (2015). <https://doi.org/10.1038/cddis.2015.167>
- Wang, D., Zhang, C., Wang, B. et al. Optimized CRISPR guide RNA design for two high-fidelity Cas9 variants by deep learning. *Nat Commun* 10, 4284 (2019). <https://doi.org/10.1038/s41467-019-12281-8>
- Wang, J., Zhang, X., Cheng, L., & Luo, Y. (2020). An overview and metanalysis of machine and deep learning-based CRISPR gRNA design tools. *RNA Biology*, 17(1), 13–22. <https://doi.org/10.1080/15476286.2019.1669406>
- Xie J., Liu M., Zhou L. (2023): CRISPR-OTE: Prediction of CRISPR On-Target Efficiency Based on Multi-Dimensional Feature Fusion, *IRBM*, Volume 44, Issue 1, 100732, ISSN 1959-0318, <https://doi.org/10.1016/j.irbm.2022.07.003>.
- Xie S, Shen B, Zhang C, Huang X, Zhang Y (2014) sgRNAcas9: A Software Package for Designing CRISPR sgRNA and Evaluating Potential Off-Target Cleavage Sites. *PLoS ONE* 9(6): e100448. <https://doi.org/10.1371/journal.pone.0100448>
- Xu J, Livraghi-Butrico A, Hou X, Rajagopalan C, Zhang J, Song J, Jiang H, Wei HG, Wang H, Bouhamdan M, Ruan J, Yang D, Qiu Y, Xie Y, Barrett R, McClellan S, Mou H, Wu Q, Chen X, Rogers TD, Wilkinson KJ, Gilmore RC, Esther CR Jr, Zaman K, Liang X, Sobolic M, Hazlett L, Zhang K, Frizzell RA, Gentzsch M, O'Neal WK, Grubb BR, Chen YE, Boucher RC, Sun F. Phenotypes of CF rabbits generated by CRISPR/Cas9-mediated disruption of the CFTR gene. *JCI Insight.* 2021 Jan 11;6(1):e139813. doi: 10.1172/jci.insight.139813. PMID: 33232302; PMCID: PMC7821608.
- Yang, D., Song, J., Zhang, J. et al. Identification and characterization of rabbit ROSA26 for gene knock-in and stable reporter gene expression. *Sci Rep* 6, 25161 (2016). <https://doi.org/10.1038/srep25161>
- Yuan, L., Sui, T., Chen, M. et al. CRISPR/Cas9-mediated GJA8 knockout in rabbits recapitulates human congenital cataracts. *Sci Rep* 6, 22024 (2016). <https://doi.org/10.1038/srep22024>
- Yuan, T., Zhong, Y., Wang, Y. et al. Generation of hyperlipidemic rabbit models using multiple sgRNAs targeted CRISPR/Cas9 gene editing system. *Lipids Health Dis* 18, 69 (2019). <https://doi.org/10.1186/s12944-019-1013-8>
- Yuwei Z., Guofang Z., Hadi A. F. Y., Tianfei Y., Shiyun H., Ting C., Qi L. (2020): *In silico* Method in CRISPR/Cas System: An Expedite and Powerful Booster. *Frontiers in Oncology*, 10, DOI:10.3389/fonc.2020.584404, ISSN:2234-943X
- Zhang X., Tee L., Wang X., Huang Q., Yang S. (2015): Off-target Effects in CRISPR/Cas9-mediated Genome Engineering. *Molecular Therapy - Nucleic Acids*, Volume 4, e264, ISSN 2162-2531, <https://doi.org/10.1038/mtna.2015.37>.

7. Táblázatok és ábrák jegyzéke

BROWSER	SUPPORTED VERSIONS
Chrome	2 most recent versions
Firefox	latest and extended support release (ESR)
Edge	2 most recent major versions
Safari	2 most recent major versions
iOS	2 most recent major versions
Android	2 most recent major versions

11. ábra: A Harmonio szoftver böngésző követelményei

1. táblázat: gRNS értékelés, a PAM pirossal kiemelve..

Gén	sgRNS	Injektált zigóták	Élve született utódok	Mutáns utódok	Bi allélikus utódok	sgRNS hatékonysága
Nox5 Exon 3	GCTCATCCATGGGAGCCCTA TGG	n/a	n/a	n/a	n/a	„Legjobb”
MSTN Exon 1	GCTGATCGTGGCTGGCC CAGTGG	158	20	16	6	10,13%
MSTN Exon 1	GGCTGTGTAATGCATGCACT TGG	158	20	16	6	10,13%
DMD Exon 51	CTTGACAGAACTTACCGAG TGG	123	33	26	22	21,1%
DMD Exon 51	ATGGCATTCTAGTTCGGAG AGG	123	33	26	22	21,1%
ANO5 Exon 12	ACTCTGTGTCGTCGCCTGT ATGG	302	26	13	n/a	4,3%
ANO5 Exon 13	CCT CGAACTTACCAGGAGTACG	302	26	13	n/a	4,3%
PHEX Exon 1	CCA CCAGAGGCACTCGAATTG	153	26	19	11	12,41%
PHEX Exon 1	GCTTGGCAGGATCCTCTT TC TGG	153	26	19	11	12,41%
LDLR Exon 2	CCG CAGCAGGGGACAAGTGTGGC	109	3	2	n/a	1,83%
LDLR Exon 7	GGAATGAGTTCCAATGC CGG	109	3	2	n/a	1,83%
LDLR	CCT CAGAATCGGCCACGAGTGTC	96	10	3	n/a	3,12%

Exon 7						
ApoE intron	AAGACTGACCAACCGCAGGC GGG	96	10	3	n/a	3,12%
ApoE Exon 1	GGTGTGGTGGGCTGTGTTGG CGG	109	3	2	n/a	1,83%
LDLR Exon 2	GGAATGAGTTCCAATGC CGG	96	10	3	n/a	3,12%
MSTN Exon 1	GCTGATCGTGGCTGGCCCAG TGG	250	34	14	n/a	5,6%
MSTN Exon 3	GCTTCAAAATCCACAGTTAG AGG	65	34	14	n/a	21,5%
GJA8 Exon 1	CCC TCAATGTACGTGGGCCATGCC	90	11	11	0	12,22%
GJA8 Exon 1	GGAGAAGCGACCGGACCGCG AGG	90	11	11	0	12,22%
CFTR Exon 11	GGA GAGTTGGAGCCTTCAGA	162	11	3	n/a	1,85%
CTGF Exon 1	ACTCT CTGTGACTTTGGCTCGCCGGC	n/a	6	6	n/a	n/a
rbROSA 26 Uncharacterized locus LOC103345511 (Genomic Sequence: NW_003159428)	GGAGCATGCAGCTTTTCTT GGG	100	20	7	n/a	7%
TYR Exon 3	TGTTCAATTTCCCTGGGGCGG TGG	158	31	23	n/a	14,5%
ATP7B Exon 8	GGG ACCCCGCCACCGACCTTGTG	118	31	23	n/a	19,49%

2. táblázat: Off-target előrejelzés, a PAM pirossal kiemelve.

Gén	sgRNS	Potenciális Off-target	Off-target
Nox5 Exon 3	GCTCATCCATGGGAGCCCTA TGG	3	0
MSTN Exon 1	GCTGATCGTGGCTGGCCCAG TGG	5	0
MSTN Exon 1	GGCTGTGTAATGCATGCACT TGG	5	0
DMD Exon 51	CTTGGACAGAACTTACCGAG TGG	10	0
DMD Exon 51	ATGGCATTCTAGTTCGGAG AGG	10	0
ANO5 Exon 12	ACTCTTGTCGTCGCCTGTAT TGG	12	0
ANO5 Exon 13	CCT CGAACTTACCAGGAGTACG	12	0

PHEX Exon 1	CCACCAGAGGCACTCGAATTG	14	0
PHEX Exon 1	GCTTGGCACGATCCTCTTTCTGG	14	0
LDLR Exon 2	CCGCAGCAGGGGACAAGTGTGGC	30	0
LDLR Exon 7	GGAATGAGTTCCAATGCCGG	30	0
LDLR Exon 7	CCTCAGAATCGGCCACGAGTGTC	30	0
ApoE intron	AAGACTGACCAACCGCAGGC	30	0
ApoE Exon 1	GGTGTGGTGGGCTGTGTTGGCGG	30	0
LDLR Exon 2	GGAATGAGTTCCAATGCCGG	30	0
MSTN Exon 1	GCTGATCGTGGCTGGCCAGTGG	27	0
MSTN Exon 3	GCTTCAAAATCCACAGTTAGAGG	27	0
GJA8 Exon 1	CCCTCATGTACGTGGGCCATGCC	7	0
GJA8 Exon 1	GGAGAAGCGACCGGACCGCGAGG	7	0
CFTR Exon 11	GGAGAGTTGGAGCCTTCAGA	51	0
CTGF Exon 1	ACTCTTCTGTGACTTTGGCTCGCCGGC	51	0
rbROSAV 26 Uncharacterized locus LOC103345511 (Genomic Sequence: NW_003159428)	GGAGCATGCAGCTTTTCTTGGG	24	0
TYR Exon 3	TGTTCAATTCCTGGGGCGGTGG	10	0
ATP7B Exon 8	GGGACCCCGCCACCGACCTTGTG	3	1

3. táblázat: A szakirodalmi sgRNS-ek összehasonlítása a Harmonio eredményeivel. Gén: A szakirodalomban felsorolt gének, sgRNS: a szakirodalomban alkalmazott sgRNS, sgRNS hatékonyság: a cikkek vonatkozó részeiből (mikroinjektált zigóták száma, mutáns élve született utódok száma) számított érték, Harmonio hatékonyság: A Harmonio által prediktált hatékonysága a szakirodalmi sgRNS-eknek, ahol „nincs találat” szerepel az azt jelenti, hogy az említett RNS nem tartozik a leghatékonyabb 30 közé. Harmonio sgRNS: A Harmonio által a leghatékonyabbnak ítélt RNS. A piros betűk a PAM-ot jelzik.

Gén	sgRNS	sgRNS hatékonysága	Harmonio hatékonyság	Harmonio sgRNS
Nox5 Exon 3	GCTCATCCATGGGAGCCCTATGG	„Legjobb”	50%	GGAAGTTGAGTTGTCCATAGGG
MSTN Exon 1	GCTGATCGTGGCTGGCCAGTGG	10,13%	31%	ATATTACTGTTTGTCTCCAGG
MSTN Exon 1	GGCTGTGTAATGCATGCACTGG	10,13%	32%	ATATTACTGTTTGTCTCCAGG
DMD Exon 51	CTTGGACAGAACTACCGAGTGG	21,1%	Nincs találat	CCTGCTCTGGCAGATTCAACCG
DMD Exon 51	ATGGCATTTCTAGTTCGGAGAGG	21,1%	31%	CCTGCTCTGGCAGATTCAACCG

ANO5 Exon 12	ACTCTTGTGCTCGCCTGTATGG	4,3%	Nincs találat	AATATTGCAAAGAACACTGTG
ANO5 Exon 13	CCTCGAACTTACCAGGAGTACG	4,3%	Nincs találat	ACTGGTATTCCAATGTGGCTTG
PHEX Exon 1	CCAACAGAGGCACTCGAATTG	12,41%	23%	CCTGTTTCTGCTTCCATCAGAAG
PHEX Exon 1	GCTTGGCAGCATCTCTTTCTGG	12,41%	Nincs találat	CCTGTTTCTGCTTCCATCAGAAG
LDLR Exon 2	CCGCAGCAGGGGACAAGTGTGGC	1,83%	Nincs találat	TACACTTCCCGTTCGGGATTGG
LDLR Exon 7	GGAATGAGTTCCAATGCCGG	3,12%	Nincs találat	TGATATCTTCGACGCGTCGCTGG
LDLR Exon 7	CCTCAGAATCGGCCACGAGTGTC	1,83%	28%	AGGTGTGGGAGCAGCCTCCGTG
ApoE intron	AAGACTGACCAACCGCAGGCGGG	3,12%	10%	CAGCTGTGTGTAACCTGGCGGG
ApoE Exon 1	GGTGTGGTGGGCTGTGTTGGCGG	3,12%	22%	AGGCGACTGACCAACGCAGGCGG
LDLR Exon 2	GGAATGAGTTCCAATGCCGG	1,83%	11%	CTTCATCTTCCCGCTGCGTTGG
MSTN Exon 1	GCTGATCGTGGCTGGCCAGTGG	5,6%	11%	ATATTACTGTTTTGTCTCCAAG
MSTN Exon 3	GCTTCAAAATCCACAGTTAGAGG	21,5%	50%	AGTAATTGGCCTTGATCTCTTG
GJA8 Exon 1	CCCTCATGTACGTGGGCCATGCC	12,22%	26%	CTCCGGGGCATGTCCGCTAGGG
GJA8 Exon 1	GGAGAAGCGACCGGACCGCGAGG	12,22%	33%	CGAAGCAGTCCACCACATTGGG
CFTR Exon 11	GGAGAGTTGGAGCCTTCAGA	1,85%	13%	GTAGCCCTTCGCTTTCTGGATTG
CTGF Exon 1	ACTCTTCTGTGACTTTGGCTCGCCGGC	n/a	24%	AGTGTCTCTTGGTAACTCCGGGG
rbROSA 26 Uncharacterized locus LOC103345511 (Genomic Sequence: NW_003159428)	GGAGCATGCAGCTTTTTCTTGGG	7%	11%	TTCGAGGAAAAGATCAGCACTGG
TYR Exon 3	TGTTCAATTCCTGGGGCGGTGG	14,5%	33%	CAGTGTCTCTTGGTAACTCCGGG
ATP7B Exon 8	GGGACCCCGCCACCGACCTGTG	19,49%	41%	CATAGGAAGGCAGTGTCTCTTGG

4. táblázat: A szakirodalmi sgRNS-ek CHOPCHOP v3 által jósolt hatékonysága. Gén: Gén: A szakirodalomban felsorolt gének, sgRNS: a szakirodalomban alkalmazott sgRNS, sgRNS hatékonyság: a cikkek vonatkozó részeiből (mikroinjektált zigóták száma, mutáns élve született utódok száma) számított érték, CHOPCHOP hatékonyság: A CHOPCHOP v3 által prediktált hatékonysága a szakirodalmi sgRNS-eknek. A piros betűk a PAM-ot jelzik.

Gén	sgRNS	sgRNS hatékonyság	CHOPCHOP hatékonyság
Nox5 Exon 3	GCTCATCCATGGGAGCCCTATGG	„Legjobb”	41%

MSTN Exon 1	GCTGATCGTGGCTGGCCCAGTGG	10,13%	48%
MSTN Exon 1	GGCTGTGTAATGCATGCACTTGG	10,13%	54%
DMD Exon 51	CTTGGACAGAACTTACCGAGTGG	21,1%	59%
DMD Exon 51	ATGGCATTCTAGTTCGGAGAGG	21,1%	56%
ANO5 Exon 12	ACTCTTGTCGTCGCCTGTATGG	4,3%	60%
ANO5 Exon 13	CCTCGAACTTACCAGGAGTACG	4,3%	71%
PHEX Exon 1	CCACCAGAGGCACTCGAATTG	12,41%	51%
PHEX Exon 1	GCTTGGCACGATCCTCTTTCTGG	12,41%	72%
LDLR Exon 2	CCGCAGCAGGGGACAAGTGTGGC	1,83%	65%
LDLR Exon 7	GGAATGAGTTCCAATGCCGG	1,83%	58%
LDLR Exon 7	CCTCAGAATCGGCCACGAGTGTC	3,12%	50%
ApoE intron	AAGACTGACCAACCGCAGGC	3,12%	59%
ApoE Exon 1	GGTGTGGTGGGCTGTGTTGGCGG	1,83%	75%
LDLR Exon 2	GGAATGAGTTCCAATGCCGG	3,12%	68%
MSTN Exon 1	GCTGATCGTGGCTGGCCCAGTGG	5,6%	81%
MSTN Exon 3	GCTTCAAAAATCCACAGTTAGAGG	21,5%	72%

GJA8 Exon 1	CCCTCATGTACGTGGGCCATGCC	12,22%	58%
GJA8 Exon 1	GGAGAAGCGACCGGACCGCGAGG	12,22%	78%
CFTR Exon 11	GGAGAGTTGGAGCCTTCAGA	1,85%	66%
CTGF Exon 1	ACTCTTCTGTGACTTTGGCTCGCCGGC	n/a	54%
rbROSA 26 uncharacterized locus LOC103345511 (Genomic Sequence: NW_003159428)	GGAGCATGCAGCTTTTTCTTGGG	7%	50%
TYR Exon 3	TGTTCAATTTCCCTGGGGCGGTGG	14,5%	72%
ATP7B Exon 8	GGGACCCCGCCACCGACCTTGTG	19,49%	73%

5. táblázat: Az off-target adatok összevetése a Harmonio adataival. Gén: A szakirodalomban felsorolt gének, sgRNS: a szakirodalomban alkalmazott sgRNS, Potenciális Off-target: A vonatkozó cikkben említett lehetséges Off-target-ek száma, Off-target: A ténylegesen bekövetkezett, kimutatott és a vonatkozó cikkben leközölt off-target-ek száma, Harmonio potenciális Off-target: A vonatkozó cikkben szereplő sgRNS Harmonio által prediktált összes potenciális Off-target-ei. A piros betűk a PAM-ot jelzik.

Gén	sgRNS	Potenciális Off-target	Off- target	Harmonio potenciális Off-target
Nox5 Exon 3	GCTCATCCATGGGAGCCCTATGG	3	0	104
MSTN Exon 1	GCTGATCGTGGCTGGCCAGTGG	5	0	180
MSTN Exon 1	GGCTGTGTAATGCATGCACTTGG	5	0	160
DMD Exon 51	CTTGGACAGAACTTACCGAGTGG	10	0	120
DMD Exon 51	ATGGCATTCTAGTTCGGAGAGG	10	0	96
ANO5 Exon 12	ACTCTTGTCGTCGCCTGTATGG	12	0	90
ANO5 Exon 13	CCTCGAACTTACCAGGAGTACG	12	0	105
PHEX Exon 1	CCACCAGAGGCACTCGAATTG	14	0	89
PHEX Exon 1	GCTTGGCACGATCCTCTTCTGG	14	0	220
LDLR Exon 2	CCGCAGCAGGGGACAAGTGTGGC	30	0	156

LDLR Exon 7	GGAATGAGTTCCAATGCC CGG	30	0	164
LDLR Exon 7	CCT CAGAATCGGCCACGAGTGTC	30	0	187
ApoE intron	AAGACTGACCAACCGCAGGC GGG	30	0	82
ApoE Exon 1	GGTGTGGTGGGCTGTGTTGG CGG	30	0	43
LDLR Exon 2	GGAATGAGTTCCAATGCC CGG	30	0	36
MSTN Exon 1	GCTGATCGTGGCTGGCC CAGTGG	27	0	180
MSTN Exon 3	GCTTCAAAATCCACAGTTAG AGG	27	0	236
GJA8 Exon 1	CCCT CATGTACGTGGGCCATGCC	7	0	230
GJA8 Exon 1	GGAGAAGCGACCGGACCGCG AGG	7	0	12
CFTR Exon 11	GGAGAGTTGGAGCCTTCAGA	51	0	95
CTGF Exon 1	ACTCTT CTGTGACTTTGGCTCGCCGGC	51	0	68
rbROSA 26 Uncharacterized locus LOC103345511 (Genomic Sequence: NW_003159428)	GGAGCATGCAGCTTTTCTT GGG	24	0	55
TYR Exon 3	TGTTCAATTCCTGGGGCGGT TGG	10	0	32
ATP7B Exon 8	GGG ACCCCGCCACCGACCTTGTG	3	1	11

6. táblázat: A Harmonio által javasolt sgRNS Off-target és hatékonyság adatai. Gén: A szakirodalomban felsorolt gének, Harmonio sgRNS: A Harmonio által a leghatékonyabbnak ítélt RNS, Harmonio potenciális Off-target: A vonatkozó cikkben szereplő sgRNS Harmonio által prediktált összes potenciális Off-target-ei. Harmonio hatékonyság: Az adott sgRNS Harmonio által prediktált hatékonysága. A piros betűk a PAM-ot jelzik.

Gén	Harmonio sgRNS	Harmonio potenciális target	Off- target	Harmonio hatékonyság
Nox5 Exon 3	GGA ACTTGAGTTGTCCATAGGG	526		70%
MSTN Exon 1	ATATTTACTGTTTTGTCTCC AAG	415		66%
MSTN Exon 1	ATATTTACTGTTTTGTCTCC AAG	98		60%
DMD Exon 51	CCT GCTCTGGCAGATTTCACCG	78		65%

DMD Exon 51	CCTGCTCTGGCAGATTCAACCG	120	56%
ANO5 Exon 12	AATATTGCAAAGAACTGTGG	3972	49%
ANO5 Exon 13	ACTGGTATTCCAATGTGGCTGG	4145	55%
PHEX Exon 1	CCTGTTTCTGCTTCCATCAGAAG	323	51%
PHEX Exon 1	CCTGTTTCTGCTTCCATCAGAAG	256	70%
LDLR Exon 2	TACACTTCCCGTTCCGGCATTGG	155	62%
LDLR Exon 7	TGATATCTTCGCAGCGTCGCTGG	5787	53%
LDLR Exon 7	AGGTGTGGGAGCAGCCTCCGTGG	459	49,5%
ApoE intron	CAGCTGTGTGTGAACCTGGCGGG	253	57%
ApoE Exon 1	AGGCGACTGACCAACGCAGGCGG	42	70%
LDLR Exon 2	CTTCATCTTCCCGCCTGCGTTGG	120	65%
MSTN Exon 1	ATATTTACTGTTTTGTCTCCAAG	153	75%
MSTN Exon 3	AGTAATTGGCCTTGTATCTCTTG	143	71%
GJA8 Exon 1	CTCCGGGGCATGTCCGCTAGGGG	3	57,61%
GJA8 Exon 1	CGAAGCAGTCCACCACATTGGGG	9	72%
CFTR Exon 11	GTAGCCCTTCGCTTTCTGGATGG	10	61%

CTGF Exon 1	AGTGTCTCTTGGTAACTCCG GGG	9	52%
rbROSA 26 uncharacterized locus LOC103345511 (Genomic Sequence: NW_003159428)	TTCGAGGAAAAGATCAGCAC TGG	6	49,3%
TYR Exon 3	CAGTGTCTCTTGGTAACTCC GGG	11	70%
ATP7B Exon 8	CATAGGAAGGCAGTGTCTCT TGG	8	72%

8. Hallgatói nyilatkozat

1. melléklet: Hallgatói nyilatkozat

NYILATKOZAT

Alulírott BLAHA BENCE büntetőjogi felelősségem tudatában kijelentem, hogy az általam benyújtott, ÁHARMOZÓKÖZTERVEZŐ SZOFTVER HATÁSTANULMÁNYA című szakdolgozat (diplomadolgozat) önálló szellemi termékem. Amennyiben mások munkáját felhasználtam, azokra megfelelően hivatkozom, beleértve a nyomtatott és az internetes forrásokat is.

Tudomásul veszem, hogy a szakdolgozat/diplomadolgozat elektronikus példánya a védés után a Magyar Agrár- és Élettudományi Egyetem könyvtárába kerül elhelyezésre, ahol a könyvtár olvasói hozzájuthatnak.

Kelt: VÁC, 2024 év 04 hó 10 nap.

BLAHA BENCE

aláírás

9. Konzulensi nyilatkozat

NYILATKOZAT

Blaha Bence (név) (hallgató Neptun azonosítója: D0WHCQ) konzulenseként nyilatkozom arról, hogy a záródolgozatot/szakdolgozatot/diplomadolgozatot/portfóliót¹ áttekintettem, a hallgatót az irodalmi források korrekt kezelésének követelményeiről, jogi és etikai szabályairól tájékoztattam.

A záródolgozatot/szakdolgozatot/diplomadolgozatot/portfóliót a záróvizsgán történő védésre javaslom / nem javaslom².

A dolgozat állam- vagy szolgálati titkot tartalmaz: igen nem³

Kelt: Gödöllő 2024. év 04. hó 19 nap


belső konzulens

¹ A megfelelő dolgozattípus meghagyása mellett a többi típus törlendő.

² A megfelelő aláhúzendő.

³ A megfelelő aláhúzendő.