

Genomszerkesztés Eukariótákban

Dr. Henn László, tudományos munkatárs

BBTE, Haladó Genetika

Jelen tananyag a Szegedi Tudományegyetemen
készült az Európai Unió támogatásával.

Projektazonosító: EFOP-3.4.3-16-2016-00014

Olvasási idő: 45m perc.

Genomszerkesztés Eukariótákban

Henn László, Szegedi Biológiai Kutatóközpont

A gének funkciójának megértése alapvető kérdése a genetikának. A génfunkciók megértésének két útja van. A forward genetika fenotípus vizsgálatára támaszkodik, tehát egy adott fenotípusért felelős géneket vizsgálja. Ezzel szemben a reverz genetikai megközelítésekben egy adott gén funkciójára vagyunk kíváncsiak. Ez utóbbi esetben egy génkifejeződés változásának (kiütésének, inaktiválásának, expresszió erősségének vagy mintázatának megváltoztatása) fenotípusos hatását vizsgáljuk. Az örökítőanyag mesterséges megváltoztatása régóta lehetséges, például DNS-károsító ágensek (Röntgen-sugárzás, UV-sugárzás, EMS) vagy vírusok és transzpozonok által. Ez utóbbiak előnye, hogy a gazda örökítőanyagába beépülő DNS szerkeszthető, így tetszőleges DNS beépíthető a genomba. Hátrányuk azonban az, hogy az idegen DNS a gazdagenom véletlenszerű pontjára épül be. Az új generációs genommanipulálási technikák azonban lehetővé teszik, hogy akár bázispár pontossággal meghatározzuk a genommanipulálás pozícióját, tehát egy adott gént tetszőlegesen, saját genomi helyén tetszőlegesen módosítsunk.

Az új generációs genommanipulálási technikák elvileg legegyszerűbb módja a **homológ rekombináción alapuló helyspecifikus mutagenézis** adta. Ennek lényege, hogy a gazdasejtbe egy olyan DNS darabot kell bejuttatni (jellemzően plazmidon), amely nagymértékben azonos a módosítani kívánt DNS környezetével. A módosítást hordozó DNS és a genomi DNS között megfelelően nagy az egyező szakasz, akkor kis valószínűséggel ugyan, de átkereszteződés történhet az idegen és saját DNS szálak között, azaz a tetszőlegesen módosított szakasz bázispár pontossággal beépíthető a gazdagenomba. A módszer hátránya, hogy nagyon alacsony hatékonyságú, bonyolult, drága, és ahhoz, hogy öröklődő mutációkat legyenek létrehozhatók, embrionális őssejteken kell kísérletezni. A homológ rekombináción alapuló helyspecifikus mutagenézis csak néhány fajnál végezhető rutinszerűen (pl. egér).

A homológ rekombináción alapuló helyspecifikus mutagenézis hatékonysága azonban nagymértékben növelhető úgy, hogy a módosítani kívánt genomi szekvencia közvetlen közelében kettős szálú DNS törést hozunk létre. Ugyanis ebben az esetben a gazdasejt saját hibajavító mechanizmusai életbe lépnek és a DNS károsodását kijavítják. Ennek két módja:

- Non-homologous end joining (NHEJ): azaz az eltört szálak összeragasztása. Ebben az esetben bizonyos gyakorisággal előfordul, hogy az eltört végek összeragasztása pontatlan (pl: bázisok hiányoznak, vagy ismétlődnek), így adott genomi pozícióban apró deléciók, beépülések állíthatók elő (indel-ek).
- Homológ rekombináció (HR): hibajavítás templát alapján, ami normál esetben a homológ kromoszóma. Ez a hibajavítási mód „átverhető” úgy, hogy a nagyfokú homológiát mutató DNS-t juttatunk a károsodott DNS-ű sejtbe (mint a „homológ rekombináción alapuló helyspecifikus mutagenézis” módszerénél), így a sejt a módosított DNS alapján javítja ki a saját DNS-ét.

A kérdés már csak az, hogy **hogyan lehetséges a genom egyetlen és meghatározott pozíciójában kettős szálú DNS törést létrehozni?**

A **Zinkfinger nukleázok** (ZFN) olyan mesterséges DNS-kötő fehérjék, melyeket transzkripciós faktorokról mintáztak, azzal a különbséggel, hogy a DNS-kötő doménjeiket modulárisan szerkezthetővé tették, valamint hozzákapcsoltak egy dimerként aktív DNáz domént. A ZFN-ok tehát dimerként működnek, és ez a mesterséges fehérje pár 18-23 bp-os szakaszt (ami legtöbb eukarióta genomba már egyedi szekvencia) képes felismerni és elhasítani a gazdasejt genomjában. ZFN-okat kódoló gének tervezése igen bonyolult, de ezek megvásárolhatók. Jellemzően a ZNF géneket kódoló plazmidokat kell bejuttatni kívánt sejtekbe. A módszer hátránya, hogy nem lehet minden szekvenciára ZNF-t előállítani, valamint az, hogy drága.

A **TALEN** technológiánál a *Xanthomonas* baktérium TALE (transcription activator-like effector protein) fehérjéjének használták fel DNS kötő doménjeit. Ezek az apró domének egyetlen nukleotid felismerésére képesek, és minden nukleotid felismerésére létezik domén. E domének felhasználásával összelegőzhető olyan mesterséges DNáz amely tetszőleges DNS szekvenciát képes felismerni és hasítani. A ZFN-ekhez hasonlóan ezt is dimerként használják, általában plazmidon juttatják a sejtekbe. Szintén megfelelő szakértelmet igénylő, drága technológia.

A **CRISPR-Cas9** alapú technológiát a 2010-es években fejlesztették ki, mára mégis a legelterjedtebb módja az eukarióta genomok pontos módosításának. A ZFN és TALEN módszerektől eltérően a genom megfelelő szekvenciárészletének felismerésében nem fehérje, hanem egyes szálú RNS (ssRNS) molekula vesz részt, a hasítást pedig legtöbb esetben egy monomerként működő helikáz és DNáz aktivitású fehérje (Cas9) végzi. A módszer alapját egy prokarióta antivirális rendszer adja, amely képes a baktériumokat megvédeni az őket fertőző fágoktól, úgy hogy azok RNS-ét lebontja. Ez a bakteriális védelmi mechanizmus egyfajta adaptív immunitást biztosít a fágok ellen, hiszen a fágok szekvencia lenyomata a baktérium genomjába beépül és újrafertőzés esetén nagyon hatékonyan vesz részt a fág RNS eliminálásában. A mechanizmus lényege, hogy első fertőzéskor a fág RNS rövid szekvenciárészlete (protospacer) bemásolódik a baktériumgenom adott pontjára (spacer szekvencia), ahonnan újrafertőzés során átíródik (crRNS), kapcsolódik egy másik rövid egyszálú RNS-hez (tracrRNS), és ez a részlegesen kettős szálú RNS a Cas9 fehérjét targetálja azt a fertőző fág RNS-éhez, amit a Cas9 fehérje elhasít. Az eukarióta sejtekben való genommodosításhoz ezt a bakteriális rendszert úgy módosították, hogy Cas9 fehérjét eukarióta expresszióra alkalmassá tették (megfelelő promóter és nukleáris lokalizációs szignál [NLS] használata, kodonoptimalizálás) valamint a tracr- és crRNS-t egy, hajtű alakot felvenni képes RNS-ben egyesítették (chimeraRNS [chiRNS], vagy guideRNS [gRNS]). A CRISPR-Cas9 alapú technológia tehát egy kétkomponensű rendszer, ahol a megfelelő genomi szekvenciát targetáló gRNS-t valamint a gRNS-hez kötődő, és a célszekvenciát elhasító Cas9 fehérjét kell a gazdasejtekbe juttatni, vagy ott megtermeltetni, hogy a kívánt genomi pozícióban kettős szálú DNS-törést lehessen indukálni. A targetszekvencia 20bp hosszúságú, és egy ún. PAM szekvenciával kell határos legyen. A PAM (protospacer adjecent motif) egy rövid, 3 bp-os motívum amelynek az első nukleotidja bármi lehet, és általában a 2. és 3., de legalább a 3. nukleotidnak G-nek kell lennie (NGG vagy NNG). Ez a PAM szekvencia igény a Cas9 fehérje jellegzetessége, az eredeti bakteriális funkcionál is így van. A CRISPR-Cas9 technológia

előnye a ZFN és TALEN módszerekkel szemben, hogy a szekvencia specificitás tervezése nem igényel magasfokú bioinformatikai képzettséget, a hozzá szükséges anyagok (Cas9 és gRNS kódoló gének) hozzáférhetőek, ezért olcsó és viszonylag könnyen használható és hatékony.

A CRISPR Cas9 technológia felhasználásai (a legtöbb módszer természetesen ZFN és TALEN technológiával is működik):

- Indel jellegű mutációk létrehozás a genom tetszőleges pozíciójában: A kívánt genomi pozícióval komplementer gRNS a Cas9-et a targetszekvenciára viszi, ahol az kettős szálú DNS törést indukál. A NHEJ hibajavító mechanizmus azonban az eltört DNS végeket sokszor nem tökéletesen ragasztja össze, ezért ott extra bp beépülések (inszerciók), vagy hiányok (deléciók) lehetnek. Kén fehérjekódoló részében bekövetkező indel-ek nagy valószínűséggel frameshift mutációhoz vezethetnek, aminek következtében a mutáns génről funkcióvesztéses fehérje termelődhet. Hatékony módszere a génkiütésnek. Az indelek előfordulási gyakorisága fokozható a HR hibajavító útvonal blokkolásával (hiszen az tökéletesen javítaná a hibát). A genom két közeli pozíciójába egyszerre targetált Cas9 fehérjékkel a két töréspont között viszonylag pontos deléció hozható létre.
- HR-ón alapuló génszerkesztés: A genom bármely pontjára tetszőleges DNS (max. néhány kb-nyi) bevihető. A módszer alapja, hogy abba a sejtmagba, ahol kettős szálú DNS törést indukáltunk bevigyünk egy olyan DNS-szekvenciát (általában plazmidon) amit a HR hibajavító mechanizmus templátként tud használni. CRISPR-Cas9 módszer esetében fontos, hogy a templát szekvencia ne tartalmazzon PAM motívumot, hiszen akkor a Cas9 fehérje azt is elhasítja. A módszer előnye, hogy gyakorlatilag bázispár pontossággal szerkeszthető a genom. Két közeli pozícióba targetált Cas9 fehérjével elérhető hogy a töréspontok közötti szekvencia helyére tetszőleges DNS-t építsünk, ha ez a tetszőleges DNS megfelelő hosszúságú homológ szakaszokkal (homológia karokkal) határolt. A módszer előnye, hogy egy tetszőleges gént saját pozíciójában módosíthatunk, megőrizve (vagy éppen direkt megváltoztatva) kifejezési mintázatát. Néhány példa:
 - Riporter konstrukt: adott fehérjekódoló szekvencia kicserélése valamilyen riporter fehérjeszekvenciára (pl. LacZ, fluoreszcens fehérje)
 - Taggelés: adott fehérje epitóp- (pl. FLAG, HA, myc) vagy fluoreszcens (GFP, RFP, mCherry) taggel való ellátása
 - Doméncsere/kitörlés/beépítés
 - Génkicserélés: pl. genetikai hiba kijavítása
 - Kondicionális génkiütés: DNS rekombináz (pl Cre vagy Flp rekombináz) célszekvenciák (loxP, FRT) beépítése a génbe. A rekombinázok sejtbe juttatásával, vagy termeltetésével a gén kivágódik a genomból.
- Fehérjék targetálása kívánt genomi pozícióba: A CRISPR-Cas9 módszernek egy olyan típusa, mikor a Cas9 fehérjét úgy módosítják, hogy az ne hasítsa el a DNS-t (deadCas9 vagy dCas9). A cas9 fehérjéhez kapcsolhat sokféle fehérje, amit a gRNS-ek a kívánt

genomi pozícióba irányítanak. Ilyen fehérjék lehetnek például: transzkripciós faktorok, szupresszorok, kromatinmódosító fehérjék, jelmolekulák.

Az utóbbi évtizedben elterjedt génmanipulációs technológiáknak köszönhetően ez eukarióta genomok gyakorlatilag tetszőlegesen módosíthatók. Ez hatalmas technikai fejlődést tesz lehetővé a élettudományokban. Természetesen óriás lehetőséget és egyben kockázatot is jelent az orvosi, állattenyésztési, növénytermesztési és biotechnológiai felhasználások tekintetében.